Impulse Control and Optimal Portfolio Selection
with General Transaction Cost

Jin Ma∗, Qingshuo Song†, Jing Xu‡ and Jianfeng Zhang§

March 20, 2009

Abstract

In this paper we study an optimal portfolio selection problem under general trans-
action cost. We consider a simplified financial market that consists of only a risk free
asset and a risky asset, but the admissible portfolios are only allowed to have piecewise
constant paths, reflecting a more practical perspective. The problem is then reduced to
an impulse control problem, but it is different from the standard models. We prove the
existence of the optimal strategy for a fairly large class of cost functionals, and we show
that the number of trading times is a random variable with finite expectation. Our
result covers the cost functionals such as the commonly used fixed cost case as well as
the more general ones that are Hölder continuous, which essentially covers most of the
existing cost functions in the literature. Unlike the commonly seen results that char-
acterize the value function via variational inequalities, we attack the optimal strategy
directly.

∗Department of Mathematics, University of Southern California, Los Angels, CA 90089; email:
jinma@usc.edu. This author is supported in part by NSF grant #DMS 0505472 and 0806017.
†Department of Mathematics, University of Southern California, Los Angels, CA 90089; email:
qingshu@usc.edu.
‡School of Economics and Business Administration, Chongqing University, Chongqing 400030, China;
email: xujing8023@yahoo.com.cn.
§Department of Mathematics, University of Southern California, Los Angels, CA 90089; email: jian-
fenz@usc.edu. This author is supported in part by NSF grant #DMS 0631360.
1 Introduction

The study of the financial markets with frictions, either due to transaction costs or due to constraints on trade, has long been considered of great interest from both researchers and practitioners, for various practical reasons. These models give much more realistic characterizations of the market, compared to the classical ones that are essentially based on the simplified ideal circumstances. A widely recognized market friction is the transaction costs, and the literature on the subject is rather extensive (see, for example, [7], [14], [10], [13], [5], [6], [11], and the references therein). It should be noted that most existing works assume only two types of transaction costs: the fixed cost and the proportional cost, or the linear combination of them, and the commonly used method is dynamic programming, which leads to the free-boundary problems via a quasilinear variational inequality.

In a recent paper, along with some subsequent works, Cetin, Jarrow, and Protter [2] introduced an interesting model of liquidity risk, which relaxed both frictionless and competitiveness assumptions of the classical arbitrage pricing theory. While the theory of liquidity risk, and the cost of it, is not the main subject of this paper, several observations should be worth noting. First, if the impact of the liquidity cost on the prices is only instantaneous, then the liquidity cost is essentially the same as transaction cost. Second, it was indicated in [3] that in the cases where the dependence of the supply curve on the transaction size has a fixed rate (called the liquidity rate $\alpha$), then the liquidity cost is essentially quadratic when $\alpha$ is small, and finally, the liquidity cost would vanish if the trading strategy has continuous and finite variation paths. The significance of the last point was later amplified by Bank and Baum [1]. They proved that in some cases the liquidity cost can be avoided completely if the market is approximate complete (cf. [2]), since one can always approximate a trading strategy by those that have continuous and finite variation paths. We should note that the liquidity risk and its cost have also been looked at from different perspectives. For example, by considering the Gamma constraint on the admissible portfolios and by using the so-called second order backward SDEs, Cetin, Soner, and Touzi [4] recently showed that the super-hedging price is in general higher than the Black-Scholes price, under the continuous time trading strategies, from which they conclude that the liquidity cost does exist.

In this paper we shall focus on the issue of utility optimization instead of hedging, but at the same time paying strong attention to two main issues. First, we shall consider a more general class of transaction costs, which bridges the usual classes of fixed and proportional ones. Roughly speaking, we shall consider the transaction cost function $c$, where $c(z)$ is of the order $|z|^\alpha$ for small trading size $z$. We should note that, if $\alpha > 1$, then by allowing
multiple instantaneous trading, the transaction cost vanishes in the approximating sense, as we saw in [2]. In the case when $\alpha = 1$, we essentially have the proportional transaction cost, and we know (cf. e.g., [14]) that in general one can find the optimal strategy that has finite variation path, but not necessarily piecewise constant. This brings out the second main feature of our framework: we shall require that our optimal strategy to have only piecewise constants. In other words, we consider only “impulse controls”, the arguably most sensible strategies in practice. We should note that with such strategy the liquidity cost does exist.

At this point we would like to remark that the compatibility between the path regularity of the optimal trading strategy and the growth of the cost functional seems to be an interesting issue in its own right. The fact that the trading strategies with continuous and finite variation paths eliminates the liquidity cost, which is essentially quadratic as we mentioned before, is a clear evidence. On the other hand, it is intuitively clear that one should avoid fluctuation of a trading strategy (hence considering the piecewise constant ones) under fixed cost. It could also be argued that finite variational strategies under proportional cost (e.g, [14]), and quadratic variational strategies under quadratic cost (e.g., [3]). The question is then whether there is a definitive answer? In particular, it might be worth investigating those cases where the optimal strategies have to be piecewise constants. Putting the non-financial rationales for using the piecewise constant strategies (e.g. computational cost, etc.), our main result is that, for $\alpha \in [0, 1)$, the optimal strategy is necessarily piecewise constant and its number of trading times is integrable. To our best knowledge, this result is new. We should note that, when $\alpha = 0$ (the case of fixed transaction cost), some results in this direction have been obtained in [7] and [11].

We should point out that a utility optimization problem with piecewise constant strategies is essentially an impulse control problem, and large amount of literature exist in this field. However, we note that most results in the literature assume that the cost function is strictly positive, and the solution methods often fall into one of the two categories that are typical in the singular/impulse stochastic control literature: one is to show that the value function is the unique viscosity solution to a quasi-variational inequality (QVI) (cf., e.g. [11] and [13]); and the other is to prove a verification theorem, that is, any smooth solution to the QVI must be the value function, provided that one can construct an optimal strategy through the smooth solution of the QVI. However, since the QVI does not have a smooth solution in general, the construction of the optimal strategy, whence in many cases the existence of it, becomes problematic.

In this paper we shall attack the existence of optimal strategy directly. We first consider a sequence of approximating problems for which the strategies are restricted to the fixed
number (say, \( n \)) trades. We show that for each \( n \) the optimal strategy, denoted by \( Z^n \) exists. We then show that as \( n \to \infty \) the value functions will converge to the original one. The main technical part in this analysis turns out to be some uniform estimates of the number of jumps of \( Z^n \). These estimates will enable us to study the regularity of the value function and to construct the optimal strategy. We should note that the regularity of the value function, which we need to construct the optimal strategy, is weaker than those that are commonly seen in the literature.

The rest of the paper is organized as follows. In Section 2 we formulate the problem. In Sections 3 and 4 we study the approximating value function \( V^n \) and its corresponding optimal strategy \( Z^n \). In Section 5 we obtain uniform estimates of \( Z^n \), which leads to the regularity of the value function \( V \). In Section 6 we study the optimal strategy of the original problem. Finally in Section 7 we give the technical proofs needed in previous sections.

## 2 Problem Formulation

Let \((\Omega, \mathcal{F}, P; F)\) be a complete filtered probability space on which is defined a standard Brownian motion \( W \). We assume that the filtration \( F = \{F_t\}_{t \geq 0} \) is actually generated by \( W \), augmented by all the \( P \)-null sets as usual. For simplicity, we assume that the financial market consists of only two assets, a bank account and a stock. We further assume that the interest rate is 0, and that the price of the stock, denoted by \( X = \{X_t : t \geq 0\} \), follows the stochastic differential equation:

\[
\begin{aligned}
&dX_r = b(r, X_r)dr + \sigma(r, X_r)dW_r, \\
&X_t = x, \quad r \geq t.
\end{aligned}
\]

We shall denote the solution to the SDE (2.1) by \( X^{t,x} \) as usual.

Fix a time duration \([t, T]\), the rule of the trading is the following. An investor has an initial wealth \( y \) and holds \( z \) shares of stock at price \( x \) (hence she puts \( y - zx \) in the bank). The investor is allowed to trade finitely many times during \([t, T]\), at the time and amount of her choosing. We denote the trading times by \( t = \tau_0 \leq \tau_1 \leq \tau_2 \leq \cdots \leq T \), and the number of shares held during \([\tau_i, \tau_{i+1}]\) by \( Z_i, \ i = 0, 1, 2, \cdots \), respectively. We should note that \( Z_0 \neq z \) is possible, which means that the investor is allowed to make an immediate trade at time \( t \). To make this point more explicit in what follows we shall make the convention that \( Z_{t-} = z \). Furthermore, we assume that at each trade, the investor is obliged to pay a transaction cost \( c(\Delta Z_i, X_{\tau_i}) \), where \( c \geq 0 \) is the transaction cost function with \( c(0, x) = 0 \), and \( \Delta Z_i = Z_i - Z_{i-1} \) is the number of shares traded at \( \tau_i \) with the obvious convention that
$\Delta Z_0 = Z_0 - z$. Clearly, if $c(\Delta Z_i, X_{\tau_i}) = c(\Delta Z_i X_{\tau_i})$, then the cost function would depend on the trading volume in terms of dollar amount, as we often saw in literature. We note at this point we allow $Z_{i+1} \neq Z_i$ on the set $\{\tau_{i+1} = \tau_i\}$, and we consider such a situation as “multiple trading”. We will see that such strategies can be avoided if the cost function is sub-additive.

Now given initial data $(t, x, y, z)$, let us denote the total wealth of the investor at time $r \in [t, T]$ by $Y_r = Y_{t,x,y,z}^r$, and hence the the bank account amount at time $r$ is $\tilde{Y}_r = Y_r - Z_r X_r$. Then a self-financing strategy $\{(\tau_i, Z_i)\}_{i=0}^\infty$ must satisfy the budget equation:

$$Y_{\tau_i} - Y_{\tau_{i-1}} - Z_{i-1}(X_{\tau_i} - X_{\tau_{i-1}}) + c(\Delta Z_i, X_{\tau_i}) = 0, \quad i = 0, 1, 2, \ldots,$$

Noting that $\tau_0 = t$, $Y_t = y$, and $Z_{-1} \triangleq Z_{t-} = z$, and recalling that the interest rate is 0, we can easily derive the dynamics of the process $Y_r, r \in [t, T]$:

$$Y_r = y + \sum_{i=0}^\infty Z_{i-1}[X_{\tau_i \wedge r}^{t,x} - X_{\tau_{i-1} \wedge r}^{t,x}] - \sum_{i=0}^\infty c(\Delta Z_i, X_{\tau_i}), \quad r \in [t, T]. \tag{2.2}$$

We now give the rigorous description of the admissible strategies. We denote $F = \{\mathcal{F}_t^W\}_{t \geq 0}$ be the filtration generated by $W$.

**Definition 2.1** Given $(t, z)$ with $t < T$ and $|z| \leq M$, the set of admissible trading strategies, $\mathcal{Z}(z)$, is the collection of all impulse controls $\{(\tau_i, Z_i)\}_{i=0}^\infty$, such that:

(i) $t = \tau_0 \leq \tau_1 \leq \cdots \leq T$ are a sequence of $\mathcal{F}$-stopping times.

(ii) $Z_{-1} \triangleq Z_{t-} = Z_{t-} = z$ and $Z_i \in \mathcal{F}_{\tau_i}, i = 0, 1, 2, \cdots$, such that $|Z_i| \leq M, \ a.s.$

(iii) $P\{\tau_i = T, \text{for all but finitely many } i \text{'s}\} = 1$.

(iv) $Z_i = 0, \text{ whenever } \tau_i = T$.

**Remark 2.2** In Definition 2.1 the assumption (ii) states that an initial “jump” at time $t$ is possible, if $Z_0 \neq z$. However, the restriction on the limit of the number of shares to be held is merely technical. This restriction can be removed by requiring that the transaction cost $c$ satisfies certain growth condition, for example,

$$\lim_{|z| \to \infty} \inf_x \frac{|c(z, x)|}{|z|} = \infty. \tag{2.3}$$

In fact, under (2.3) one can show that there exists a bound $M > 0$, depending only on the expected bound of the underlying price $X$, such that any strategy with trading size exceed $M$ at some time will make the transaction cost so big that the total terminal wealth becomes negative. Thus such a strategy can never be optimal. But in this paper we prefer not to pursue these details, and simply assume that optimal strategy are bounded.
Moreover, the assumption means the investor will trade only finitely many times during \([t, T]\), almost surely; and assumption (iv) indicates that the investor will always clear her stock position in the end and will hold only cash in the bank (i.e., \(Y_T = \tilde{Y}_T\)). Such an assumption is not unusual, see for example, [2] and [4].

Clearly, under Definition 2.1 the terminal wealth \(Y_T = \tilde{Y}_T\) can be written as:

\[
Y_T = y + \sum_{i=0}^{\infty} Z_{i-1}[X_{\tau_i}^x - X_{\tau_{i-1}}^x] - \sum_{i=0}^{\infty} c(Z_i - Z_{i-1}, X_{\tau_i}).
\]  

(2.4)

Our optimization problem is then to find, for a given utility function \(U\), an admissible strategy \(\{(\tau_i^*, Z_i^*)\}_{i=0}^{\infty}\) such that the corresponding wealth \(Y^*\) satisfies

\[
V(t, x, y, z) \overset{\triangle}{=} \sup_{\tilde{Z}(\cdot)} E\{U(Y_T)\} = E\{U(Y^*_T)\}.
\]

(2.5)

We now consider our problem in a more generalizable framework. First, note that an impulse control \(\{(\tau_i, Z_i)\}_{i=0}^{\infty}\) can be written as a continuous time, piecewise constant process:

\[
Z_r \overset{\triangle}{=} \sum_{i=0}^{\infty} Z_{i-1}1[\tau_{i-1}, \tau_i) (r), \quad r \in [t, T].
\]

(2.6)

Next, note that for any process \(Z\) in the form (2.6), we can define the stopping times:

\[
\tau_0 \overset{\triangle}{=} t, \quad \text{and}
\]

\[
\tau_{i+1} \overset{\triangle}{=} \inf\{s > \tau_i : Z_s \neq Z_{\tau_i}\} \wedge T, \quad i = 0, 1, \ldots
\]

so that \(Z_r = \sum_{i=0}^{\infty} Z_{\tau_{i-1}}1[\tau_{i-1}, \tau_i) (r), r \in [t, T]\). It is then natural to try to recast our admissible strategies in terms of the simple processes with the form (2.6). To do this, we first need to make the following important observation.

Note that the pairs \((\tau_i, Z_{\tau_i})\), \(i = 1, 2, \ldots\), in the representation (2.6) must satisfy: a) \(\tau_{i+1} > \tau_i\) whenever \(\tau_i < T\), and b) \(Z_{\tau_{i+1}} \neq Z_{\tau_i}\) whenever \(\tau_{i+1} < T\). However, in Definition 2.1 we allow instantaneous multiple trading, that is, we allow \(\tau_i = \tau_{i+1}\) but \(Z_i \neq Z_{i+1}\). To avoid such a discrepancy we now assume that the cost function \(c\) is sub-additive in the variable \(z\). That is, for any \(x\), it holds that

\[
c(z_1 + z_2, x) \leq c(z_1, x) + c(z_2, x), \quad |z_1|, |z_2|, |z_1 + z_2| \leq 2M.
\]  

(2.7)

Here we use the bound \(2M\) because the variable of \(c\) is actually \(\Delta z\). We claim that if in an admissible control \(\{(\tau_i, Z_i)\}_{i=0}^{\infty}\), \(\tau_{i+1} = \tau_i\) but \(Z_{i+1} \neq Z_i\), for some \(i\), then one can always find another admissible control \(\{\tilde{\tau}_i, \tilde{Z}_i\}_{i=0}^{\infty}\) such that \(\tilde{Z}_{i+1} = Z_i\) whenever \(\tau_{i+1} = \tau_i\). Indeed, without loss of generality let us assume that \(\tau_2 = \tau_1\) but \(Z_2 \neq Z_1\). We reset \(\tilde{Z}_2 \overset{\triangle}{=} Z_1\), and
\( \tilde{Z}_i = Z_i, \) for \( i \neq 2. \) Denote the corresponding terminal wealth by \( \tilde{Y}_T. \) From (2.4) we see that the only difference between \( Y_T \) and \( \tilde{Y}_T \) is the two terms in the transaction cost, namely the terms \( c(Z_2 - Z_1, X_{\tau_1}) + c(Z_3 - Z_2, X_{\tau_2}) \) will become

\[
\begin{align*}
&c(\tilde{Z}_2 - Z_1, X_{\tau_1}) + c(Z_3 - \tilde{Z}_2, X_{\tau_2}) = c(Z_3 - Z_1, X_{\tau_1}) \\
&\leq c(Z_3 - Z_2, X_{\tau_2}) + c(Z_2 - Z_1, X_{\tau_1}).
\end{align*}
\]

Thus \( \tilde{Y}_T \geq Y_T \) and hence \( E\{U(\tilde{Y}_T)\} \geq E\{U(Y_T)\}. \) In other words, we can always improve the expected utility by eliminating the instantaneous multiple trading. Therefore, in what follows we shall assume \( Z_{i+1} = Z_i \) whenever \( \tau_{i+1} = \tau_i. \)

We now redefine the set of admissible strategies.

**Definition 2.3** Given \((t, z) \in [0, T) \times [-M, M],\) the set of admissible strategies, denoted as \( Z_t(z), \) is the space of \( F \)-adapted processes \( Z \) over \([t, T)\) such that

(i) There exist \( F \)-stopping times \( \tau_0 \leq \tau_1 \leq \cdots \leq T \) such that \( \tau_{i+1} > \tau_i \) whenever \( \tau_i < T. \)

(ii) \( Z_{\tau_0} = Z_t = z \) and \( Z_s = \sum_{i=0}^{\infty} Z_{\tau_i-1} \mathbf{1}_{[\tau_i-1, \tau_i)}(s) \) for \( s \in [t, T]. \)

(iii) \( Z_{\tau_i+1} \neq Z_{\tau_i} \) whenever \( \tau_{i+1} < T. \)

(iv) \( P\{\tau_i < T, \text{for all but finitely many } i \text{'s}\} = 1. \)

Given \((t, x, y, z)\) and \( Z \in Z_t(z),\) denote \( \Delta Z_{\tau_i} \triangleq Z_{\tau_i} - Z_{\tau_i-1} \) and

\[
Y^{t, x, y, Z}_T \triangleq y + \int_t^T Z_s dX^{t, x}_s - \sum_{i=0}^{\infty} c(\Delta Z_{\tau_i}, X_{\tau_i}).
\]  

(2.8)

Then our optimization problem (2.5) is equivalent to

\[
V(t, x, y, z) \triangleq \sup_{Z \in Z_t(z)} E\{U(Y^{t, x, y, Z}_T)\}.
\]  

(2.9)

**Remark 2.4** (i) We should note that in the case \( Z_{\tau_0} \neq z, \) an initial jump \( \Delta Z_t = \Delta Z_{\tau_0} = Z_{\tau_0} - z \) will appear.

(ii) The optimization problem (2.8) and (2.9) can be extended to the cases where admissible strategies are allowed to be general \( F \)-adapted, càdlàg processes. But in that case we need to redefine the aggregated transaction cost. For example, we can consider the aggregated transaction cost in the following forms:

\[
\begin{align*}
&\sup_\pi \sum_{i=0}^{\infty} c(\Delta Z_{\tau_i}, X_{\tau_i}), \quad \text{or} \quad \lim_{|\pi| \to 0} \sum_{i=0}^{\infty} c(\Delta Z_{\tau_i}, X_{\tau_i}).
\end{align*}
\]  

(2.10)
where the supreme is over all possible random partitions $\pi: t = \tau_0 < \tau_1 < \cdots \leq T$ of $[t, T]$; and $|\pi|$ is the “mesh size” of the partition. Then, under appropriate conditions on the function $c$, one can show that the value function $V$ would be the same if the supreme is taken over only piecewise constant strategies. Namely, it suffices to consider only $Z_t(z)$, and thus the aggregated cost (2.10) is again reduced to that in (2.8).

It is worth noting that our controlled dynamics is general enough to contain many existing ones as special cases. Here are some examples.

**Example 2.5** Assume that $c(z, x) = K + k|z|$. If $K > 0$, then a fixed cost is involved, and in this case we essentially have an impulse control problem (cf. e.g., [7], [10]). However, if $K = 0$, then $\sup \pi \sum_{i=1}^{\infty} c(\Delta Z_{\tau_i}, X_{\tau_i}) = k \int_t^T |dZ_r|$, where the integral is the total variation of the process $Z$ over $[t, T]$. This problem then becomes a singular stochastic control problem (cf. [11]). In this case the optimal controls are of bounded variation, but not necessarily piecewise constant.

**Example 2.6** We now modify our definition of the transaction function slightly. Consider the case where $c(z, x) = zx(e^z - 1)$, and $X_t = S_t$, $t \in [0, T]$, to be the standard Black-Scholes stock prices. Then for any (random) partition $\pi: t = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_n = T$ and any $F$-adapted process $Z$, we have

$$\sum_{i=0}^{n} c(\Delta Z_{\tau_i}, X_{\tau_i}) = \sum_{i=0}^{\infty} \Delta Z_{\tau_i} S_{\tau_i} (e^{\Delta Z_{\tau_i}} - 1) = \sum_{i=0}^{n} \Delta Z_{\tau_i} (S_{\tau_i} e^{\Delta Z_{\tau_i}} - S_{\tau_i})$$

$$= \sum_{i=0}^{n} \Delta Z_{\tau_i} (S(\tau_i, \Delta Z_{\tau_i}) - S(\tau_i, 0)),$$

where $S(t, \Delta Z_t) = S_t e^{\Delta Z_t}$, $t \geq 0$ is the supply curve in an extended Black-Scholes economy, proposed in [2] and [3]. Clearly, as $|\pi| \triangleq \max_i |\tau_{i+1} - \tau_i| \to 0$, as $n \to \infty$, we have

$$\lim_{n \to \infty} \sum_{i=0}^{n} c(\Delta Z_{\tau_i}, X_{\tau_i}) = \sum_{t \leq r \leq T} \delta Z_r [S(r, \Delta Z_r) - S(r, 0)] + \int_t^T \frac{\partial S}{\partial x}(r, 0) d[Z, Z]|^r_c. \quad (2.11)$$

Namely, our transaction cost essentially contains the liquidity cost of [2] as a special case.

Note that if we require that $Z$ be of finite variation paths, then the second term on the right side of (2.11) vanishes. Furthermore, since $e^x - 1 \sim x$ for $|x|$ sufficiently small, if we assume further that $Z$ is piecewise constant and all $|\Delta Z_{\tau_i}|$s are small, then we see from (2.11) that

$$\sum_{i=0}^{n} c(\Delta Z_{\tau_i}, X_{\tau_i}) \sim \sum_{t \leq r \leq T} |\Delta Z_r|^2 S_r.$$
In other words, in this case the transaction cost has a quadratic growth in the trading sizes, when $|\Delta Z_i|$’s are small.

To conclude this section we give the precise description of the transaction cost function $c$. Bearing in mind that an important assumption that we need is the subadditivity (2.7). We can justify this assumption as follows. Note that $Z_{\tau_i} - Z_{\tau_{i-1}}$ may take values in $[-2M, 2M]$, let us define, for any $|z| \leq 2M$,

$$\tilde{c}(z) \triangleq \inf \{ c(z_1) + \cdots + c(z_n) : |z_i| \leq 2M, z_1 + \cdots + z_n = z, \forall n \}.$$  

Then it is easy to see that $\tilde{c} \leq c$ and $\tilde{c}$ satisfies (2.7). Now replacing $c$ by $\tilde{c}$ in (2.4) and (2.5) we have

$$\tilde{Y}_T \triangleq y + \sum_{i=0}^{\infty} Z_{\tau_i} - \sum_{i=0}^{\infty} \tilde{c}(\Delta Z_i) ; \quad \tilde{V}(t, x, y, z) \triangleq \sup E\{U(\tilde{Y}_T)\}.$$  

Now let us assume that the utility function $U$ is continuous. Then since it is increasing, by dividing each trading size into smaller volumes if necessary, it is fairly easy to show that $\tilde{V} = V$. In other words, we can always replace the transaction cost function $c$ to one that satisfies (2.7).

We should point out that if a transaction function $c$ satisfies $c(z) \leq C|z|^\alpha$ for some constants $C > 0$ and $\alpha > 1$ near $z = 0$, then the corresponding $\tilde{c}(z) \equiv 0$. To see this, note that for any $z$, and large $n$ we have

$$\tilde{c}(z) \leq \sum_{i=1}^{n} c\left(\frac{z_i}{n}\right) \leq C \sum_{i=1}^{n} \left|\frac{z_i}{n}\right|^\alpha \leq \frac{CM^\alpha}{n^{\alpha-1}} \to 0, \quad \text{as } n \to \infty.$$  

But on the other hand, since $V = \tilde{V}$ as we argued before, the problem is then reduced to a standard utility optimization without transaction cost. While this might be an intended consequence in some situations (see, e.g., [2], where $\alpha = 2$ and the liquidity cost is eliminated by the approximate completeness of the market), this will more or less trivialize our optimization problem.

Taking all the factors discussed above into account, we shall now consider only those transaction cost functions that satisfy the subadditivity (2.7), and our main focus will be the cases that $|c(z)| \geq C|z|^\alpha$, $0 \leq \alpha < 1$, near $z = 0$. For simplicity we shall assume that all processes in this paper are one dimensional. To be more precise, we shall make use of the following \textit{Standing Assumptions}:

\textbf{(H1)} The coefficients $b$ and $\sigma$ in (2.1) are bounded and uniformly Lipschitz continuous in $x$, with a common Lipschitz constant $K > 0$.  

9
(H2) The utility function $U$ is increasing such that $0 < \lambda \leq U' \leq \Lambda$.

(H3) The transaction cost function $c$ depends only the number of shares traded. Namely, $c = c(z)$, and it satisfies:

(i) $c(z) \geq 0$ for all $z \neq 0$, and $c(0) = 0$;

(ii) $c(z_1 + z_2) \leq c(z_1) + c(z_2)$, for any $z_1, z_2$ such that $|z_1|, |z_2|, |z_1 + z_2| \leq 2M$.

(iii) $c$ is uniformly continuous in $[-2M, 0)$ and in $(0, 2M]$ with the same modulus of continuity $\rho$.

We note that in general $c(0+) \neq c(0-)$. We will specify more assumptions on $c$ later. As standard, we assume $\rho$ is a continuous increasing function on $[0, \infty)$ with $\rho(0) = 0$. The assumption $U' \geq \lambda$ is somewhat nonstandard because it excludes the exponential utility function. Our approach should still work if we replace it with a more standard assumption that $U' > 0$ and $U$ is concave, but the analysis will be much more sophisticated.

Finally, we remark that, although in the definition $Z$ is required only to be $\mathcal{F}$-adapted, the optimal $Z^*$ we will construct is in fact adapted to the smaller filtration generated by $X^{t,x}$.

3 The Approximating Problems

Our plan of attack is the following. We shall approximate the original optimization problem (2.8) and (2.9) by those with only fixed number of transactions, for which the optimal strategies are easier to find. The value function and the optimal strategy of the original problem will depend heavily on the understanding of those of the approximating problems.

To begin with, for any $n \geq 1$ we consider a reduced problem with only $n$ transactions: denote $Z^n_t(z) \triangleq \{Z \in Z_t(z) : \tau_n = T\}$, and define

$$V^n(t, x, y, z) \triangleq \sup_{Z \in Z^n_t(z)} E\{U(Y^n_{T,x,y,z})\}. \quad (3.1)$$

Note that, if $n = 1$, then

$$V^1(t, x, y, z) = E\{U(y + z(X^T_T - x) - c(-z))\}. \quad (3.2)$$

It is then readily seen, assuming (H1)–(H3), that

$$|V^1(t, x, y, z)| \leq C[1 + |y|], \quad (t, x, y, z) \in [0, T] \times \mathbb{R}^2 \times [-M, M]. \quad (3.3)$$
Here and in the sequel $C > 0$ is a generic constant depending only on $T, M, \lambda, \Lambda, K$, and $|U(0)|$ in (H1)–(H3), as well as $\sup_{|z| \leq 2M} c(z)$, and it is allowed to vary from line to line.

We first extend this fact as follows.

**Proposition 3.1** Assume (H1)–(H3). Then $V^n(t, x, y, z) \uparrow V(t, x, y, z)$, as $n \to \infty$; and

$$V^n(t, x, y, z) \leq V(t, x, y, z) \leq C[1 + |y|], \quad (t, x, y, z) \in [0, T] \times \mathbb{R}^2 \times [-M, M]. \quad (3.4)$$

**Proof.** That $V^n$ is increasing and $V^n \leq V$ is clear by definition. We first show that (3.4) holds for $V$ (whence for $V^n$ as well). For any $Z \in Z_t(z)$, let us denote $X = X^{t,x}$ and $Y = Y^{t,x,y,Z}$ for simplicity. By removing the transaction cost we have

$$Y_T \leq y + \int_t^T Z_s dX_s = y + \int_t^T Z_s b(s, X_s) ds + \int_t^T Z_s \sigma(s, X_s) dW_s.$$ 

Then, using the boundedness of the coefficients $b$ and $\sigma$, as well as the process $Z$, and some standard arguments including the Burkholder-Davis-Gundy inequality, we have

$$EU(Y_T) \leq E\left\{U(y + \int_t^T Z_s dX_s)\right\} \leq |U(0)| + \Lambda \left\{|y| + E\left|\int_t^T Z_s dX_s\right|\right\} \leq C[1 + |y|]. \quad (3.5)$$

Hence (3.4) holds.

We now show that $V^n \to V$, as $n \to \infty$. We first note that $V^n$ is non-decreasing, and bounded from above, thanks to (3.4). Thus $V^\infty(t, x, y, z) \triangleq \lim_{n \to \infty} V^n(t, x, y, z)$ exists, and $V^\infty(t, x, y, z) \leq V(t, x, y, z)$, for all $(t, x, y, z)$. We need only show that $V^\infty \geq V$. To this end, for any $Z \in Z_t(z)$ we define $Z^n_s \triangleq Z_s 1_{\{s \leq \tau_n\}}$, $s \in [t, T]$. Clearly, $Z^n \in Z^n_t(z)$. Again, we denote $X \triangleq X^{t,x}$, $Y \triangleq Y^{t,x,y,Z}$, and $Y^n \triangleq Y^{t,x,y,Z^n}$, then by Assumption (H3) (ii),

$$Y_T - Y^n_T = \int_{\tau_n}^T Z_s dX_s - \sum_{i \geq n} c(Z_{\tau_i} - Z_{\tau_{i-1}}) + c(-Z_{\tau_{n-1}}) \leq \int_{\tau_n}^T Z_s dX_s. \quad (3.6)$$

Now, for any $n$, using (H2), (3.4), and (3.6) we have

$$E\{U(Y_T)\} = E\{U(Y^n_T)\} + E\{U(Y_T) - U(Y^n_T)\}$$
$$= E\{U(Y^n_T)\} + E\left\{\left[\int_0^1 U'(Y^n_T + \theta(Y_T - Y^n_T)) d\theta\right] [Y_T - Y^n_T]\right\}$$
$$\leq V^\infty(t, x, y, z) + \Lambda E\left\{\left|\int_{\tau_n}^T Z_s dX_s\right|\right\}. \quad (3.7)$$

Next, Definition 2.3 (iv) implies that

$$\lim_{n \to \infty} \left\{\left|\int_{\tau_n}^T Z_s dX^{t,x}_s\right|\right\} = 0, \quad P - \text{a.s.}$$
This enables us to let $n \to \infty$ in (3.7) and apply the Dominated Convergence Theorem to get $E\{U(Y_T)\} \leq V^\infty(t, x, y, z)$. Since this is true for any $Z \in \mathcal{Z}_t(z)$, we conclude that $V(t, x, y, z) \leq V^\infty(t, x, y, z)$, proving the proposition.

The next two results concern the uniform regularity of $\{V^n : n \geq 1\}$. The first result gives the “equi-continuity” of $V^n$’s in the variables $(t, x, y)$.

**Proposition 3.2** Assume (H1)–(H3). Then, for any $n$, it holds that

$$|V^n(t, x_1, y, z) - V^n(t, x_2, y, z)| \leq C|\delta x|; \quad (3.8)$$

$$\lambda \delta y \leq V^n(t, x, y, z) - V^n(t, x, y_1, z) \leq \Lambda \delta y, \quad \delta y \geq 0; \quad (3.9)$$

$$|V^n(t_1, x, y, z) - V^n(t_2, x, y, z)| \leq C|\delta t|^{1/2}. \quad (3.10)$$

Here and in the sequel, $\delta \xi \triangleq \xi_1 - \xi_2, \xi = t, x, y$, respectively.

**Proof.** First let us denote $X_i = X_{t,x_i}, i = 1, 2$, and $\delta X \triangleq X_1 - X_2$. Then by the standard arguments in SDEs we know that

$$E\{\sup_{s \in [t,T]} |\delta X_s|^2\} \leq C|\delta x|^2. \quad (3.11)$$

Next, for any $Z \in \mathcal{Z}_n^\infty(z)$, denote $Y_i = Y^{t,x_i,z}, i = 1, 2$, and $\delta Y \triangleq Y^1 - Y^2$. Then

$$|\delta Y_T| \leq \int_t^T |Z_s||b(s, X^{t,x_1}_s) - b(s, X^{t,x_2}_s)|ds + \int_t^T Z_s[|\sigma(s, X^{t,x_1}_s)| - |\sigma(s, X^{t,x_2}_s)|]dW_s.$$

Then, using the Lipschitz continuity of $b$ and $\sigma$ and applying the Burkholder-Davis-Gundy inequality, noting the boundedness of $Z$ and (3.11), we have

$$\left|E\left\{U(Y^1_T) - U(Y^2_T)\right\}\right|^2 \leq CE\left\{|\delta Y_T|^2\right\} \leq CE\left\{\int_t^T |Z_s\delta X_s|^2ds\right\} \leq C|\delta x|^2.$$

Since $Z$ is arbitrary, (3.8) follows easily.

To prove (3.9) we denote, for any $Z \in \mathcal{Z}_n^\infty(z)$ and $y_1 > y_2, Y^i = Y^t,x_i,y_i,z, i = 1, 2$, and $\delta Y = Y^1 - Y^2$. Then, note that $\delta Y_T = \delta y$, we have

$$E\left\{U(Y^1_T) - U(Y^2_T)\right\} = E\left\{\left[\int_0^1 U'(Y^1_T + \theta\delta y) + \theta\delta y\right]d\theta\right\}\delta y.$$

Thus (3.9) follows from (H3) immediately.

It remains to prove (3.10). Assume $t_1 < t_2$. It is then standard to show that

$$E\left\{|X^{t_1,x}_t - X^{t_2,x}_t|^2\right\} \leq C|\delta t|, \quad t \geq t_2 > t_1. \quad (3.12)$$
Now for any \( Z \in \mathcal{Z}_{t_2}(z) \), define \( \tilde{Z}_t \triangleq 1_{[t_1, t_2]}(t) + Z_{t_1}(t_2) \). Then \( \tilde{Z} \in \mathcal{Z}_{t_1}(z) \), and if we denote \( X^i = X^i_x \), \( Y^i = Y^i_y \), \( \hat{Y}^i = Y^i_y \), \( i = 1, 2 \), respectively, then

\[
EU(Y_T^2) - V^n(t_1, x, y, z) \leq \mathbb{E}\{U(Y_T^2) - U(Y_T^1)\} \leq \mathbb{E}\{|Y_T^2 - Y_T^1|\}
\]

\[
= CE\left\{\left|\int_{t_2}^T Z_t dX_t^1 - \int_{t_1}^T \tilde{Z}_t dX_t^1\right|\right\}
\]

\[
\leq CE\left\{|z| |X_{t_2}^1 - x| + \left|\int_{t_2}^T Z_t [b(t, X_t^2) - b(t, X_t^1)] dt\right| + \left|\int_{t_2}^T Z_t [\sigma(t, X_t^2) - \sigma(t, X_t^1)] dW_t\right|\right\}
\]

\[
\leq C|\delta t|^{1/2}.
\]

Since \( Z \in \mathcal{Z}_{t_1}(z) \) is arbitrary, we get

\[
V^n(t_2, x, y, z) - V^n(t_1, x, y, z) \leq C|\delta t|^{1/2}, \quad (3.13)
\]

On the other hand, for any \( Z = \sum_{i=0}^{n} Z_{\tau_{i-1}} 1_{[\tau_{i-1}, \tau_i]} \in \mathcal{Z}_{t_1}(z) \), let \( \tilde{\tau}_i \triangleq \tau_i \vee t_2 \) and \( \tilde{Z}_i \triangleq Z_{\tilde{\tau}_i} \), \( i = 0, \ldots, n \). Recall Definition 2.1 and (2.5), we can easily see that \( \tilde{Z} \triangleq \{(\tilde{\tau}_i, \tilde{Z}_i)\} \in \tilde{Z}_{t_2}(z) \), and \( \hat{Y}^2 = Y_{t_2,x,y,Z} = Y_{t_2,x,y,Z} = Y^2 \). Therefore, we have \( V^n(t_2, x, y, z) \geq E\{U(Y_T^2)\} \), and consequently,

\[
E\left\{U(Y_T^1)\right\} - V^n(t_2, x, y, z) \leq E\{U(Y_T^1) - U(Y_T^2)\} \leq CE\left\{\left|\int_{t_1}^T Z_t dX_t^1 - \int_{t_2}^T \tilde{Z}_t dX_t^2\right|\right\}
\]

\[
\leq CE\left\{|z| |X_{t_2}^1 - x| + \left|\int_{t_2}^T Z_t [b(t, X_t^2) - b(t, X_t^1)] dt\right| + \left|\int_{t_2}^T Z_t [\sigma(t, X_t^2) - \sigma(t, X_t^1)] dW_t\right|\right\}
\]

\[
\leq C|\delta t|^{1/2}.
\]

Since \( Z \in \mathcal{Z}_{t_1}(z) \) is arbitrary, we get

\[
V^n(t_1, x, y, z) - V^n(t_2, x, y, z) \leq C|\delta t|^{1/2},
\]

which, together with (3.13), implies (3.10).

**Remark 3.3** Similar to the proof of (3.10), we can easily show that

\[
|V^n(t, x, y, z) - U(y - c(-z))| \leq C\sqrt{T - t}, \quad |V(t, x, y, z) - U(y - c(-z))| \leq C\sqrt{T - t}.
\]

In other words, it would be reasonable to define

\[
V^n(T, x, y, z) \triangleq V(T, x, y, z) \triangleq U(y - c(-z)). \quad (3.14)
\]

We shall use this as the terminal conditions for \( V^n \) and \( V \) in the sequel.
We now turn our attention to the regularity of the $V^n$'s with respect to the variable $z$. The discussion is a little more involved, and we introduce the following notion of “domination” of strategies. Assume $Z^j \in Z^n(z_j)$, $j = 1, 2$, where either $z_1 > z_2 > 0$, or $z_1 < z_2 < 0$. Denote $\delta Z = Z^1 - Z^2$, as usual. We say that $Z^1$ dominates $Z^2$ if $Z^1$ and $Z^2$ has the same jump times $\tau_i$'s, and

$$\delta z = \delta Z_{\tau_0} \geq \delta Z_{\tau_1} \geq \ldots \geq \delta Z_{\tau_n} = 0$$ or $$\delta z = \delta Z_{\tau_0} \leq \delta Z_{\tau_1} \leq \ldots \leq \delta Z_{\tau_n} = 0,$$ (3.15)

and, by denoting $sgn (0) \triangleq 0$ and $\Delta Z^j_{\tau_i} \triangleq Z^j_{\tau_i} - Z^j_{\tau_{i-1}}$,

$$sgn (\Delta Z^1_{\tau_i}) = sgn (\Delta Z^2_{\tau_i}).$$ (3.16)

**Remark 3.4** We remark that the requirements (3.15) and (3.16) guarantee not only that $Z^1$ and $Z^2$ stay close, but that they are on the same side of the origin. This is mainly due to the fact that the transaction function $c$ is allowed to behave differently on the two sides of the origin (i.e., $c(0^+) \neq c(0^-)$).

Note that if $z_1 > z_2 > 0$ and $Z^1$ dominates $Z^2$, then, denoting $Y_i = Y^{t,x,y,z^i}$, $i = 1, 2$, and $X = X^{t,x}$, by induction one can easily check that

$$\|E\{U(Y^1_T)\} - E\{U(Y^2_T)\}\| \leq CE\left\{ \left| \int_t^T \delta Z_s dX_s \right| + \left| \sum_{i=0}^n [c(\Delta Z^1_{\tau_i}) - c(\Delta Z^2_{\tau_i})] \right| \right\}$$

$$\leq C|\delta z| + CE\left\{ \sum_{i=0}^n \rho(|\Delta Z^1_{\tau_i} - \Delta Z^2_{\tau_i}|) \right\}$$ (3.17)

$$= C|\delta z| + CE\left\{ \sum_{i=0}^n \rho(\delta Z_{\tau_{i-1}} - \delta Z_{\tau_i}) \right\} \leq C|\delta z| + C\rho_n(|\delta z|),$$

where $\rho$ is the modulus of continuity of $c$ in (H2), and

$$\rho_n(|\delta z|) \overset{\Delta}{=} \sup \left\{ \sum_{i=0}^n \rho(\theta_i|\Delta z|) : \theta_0, \theta_1, \ldots, \theta_n \geq 0, \sum_{i=0}^n \theta_i = 1 \right\}. $$ (3.18)

We have the following theorem.

**Proposition 3.5** Assume (H1)–(H3). Then for any $z_1, z_2 \neq 0$ with the same sign, it holds that

$$|V^n(t,x,y,z_1) - V^n(t,x,y,z_2)| \leq C|\delta z| + \rho_n(|\delta z|), n = 1, 2, \ldots,$$ (3.19)

where $\rho_n$ is defined by (3.18)
Proof. The estimate is obvious for \( t = T \), thanks to (3.14). So we may assume \( t < T \).

Without loss of generality assume \( z_1 > z_2 > 0 \).

In light of the estimate (3.17), we need only prove the following claim: For any \( Z^1 \in \mathcal{Z}^n(z_1) \), there exists \( Z^2 \in \mathcal{Z}^n(z_2) \) dominated by \( Z^1 \), and for any \( Z^2 \in \mathcal{Z}^n(z_2) \), there exists \( Z^1 \in \mathcal{Z}^n(z_1) \) dominating \( Z^2 \). Indeed, if the claim is true, then for any \( \varepsilon > 0 \), we can find \( Z^{1,\varepsilon} \in \mathcal{Z}(z_1) \) such that \( E\{U(Y^t,x,y,Z^{1,\varepsilon})\} > V^n(t,x,y,z_1) - \varepsilon \), and (3.17) leads to that

\[
V^n(t,x,y,z_1) \leq C[|\delta z| + \rho_n(|\delta z|)] + \rho_n(|\delta z|).
\]

Letting \( \varepsilon \to 0 \) we obtain

\[
V^n(t,x,y,z_1) - V^n(t,x,y,z_2) \leq C[|\delta z| + \rho_n(|\delta z|)].
\]

(3.20)

The similar argument can lead also to

\[
V^n(t,x,y,z_2) - V^n(t,x,y,z_1) \leq C[|\delta z| + \rho_n(|\delta z|)],
\]

(3.21)

and the Theorem will follow. Therefore it remains to validate the claim.

First, let \( Z^1 = \sum_{i=0}^{n-1} Z^1_{\tau^i} 1_{[\tau^i,\tau^{i+1})} \in \mathcal{Z}^n_t(z_1) \) be given. We construct \( Z^2 \in \mathcal{Z}^n_t(z_2) \) as follows. We begin by choosing the same jump times \( \tau^i \)'s, and define \( Z^2_{\tau^i} \triangleq z_2 \). Suppose that we have defined \( Z^2_{\tau^i} \) such that either \( Z^2_{\tau^i} = Z^1_{\tau^i} \) or \( 0 < Z^2_{\tau^i} < Z^1_{\tau^i} \), we then define \( Z^2_{\tau^{i+1}} \) in the following way: if \( \tau^i+1 = T \) or \( Z^{2}_{\tau^i} = Z^{1}_{\tau^i} \), then simply set \( Z^2_{\tau^{i+1}} \triangleq Z^1_{\tau^{i+1}} \). Assume \( \tau^{i+1} < T \) and \( 0 < Z^2_{\tau^{i+1}} < Z^1_{\tau^{i+1}} \). Note that in this case, by Definition 2.3 (iii), \( Z^2_{\tau^{i+1}} \neq Z^1_{\tau^{i+1}} \). If \( Z^1_{\tau^{i+1}} > Z^2_{\tau^{i+1}} \) or \( Z^2_{\tau^{i+1}} < Z^1_{\tau^{i+1}} \), define \( Z^2_{\tau^{i+1}} \triangleq Z^1_{\tau^{i+1}} \). Otherwise, we have \( Z^2_{\tau^{i+1}} > Z^1_{\tau^{i+1}} \), then define \( Z^2_{\tau^{i+1}} \triangleright Z^1_{\tau^{i+1}} \). Note that we still have either \( Z^2_{\tau^{i+1}} = Z^1_{\tau^{i+1}} \) or \( 0 < Z^2_{\tau^{i+1}} < Z^1_{\tau^{i+1}} \), so we may continue to define \( Z^2 \). One can check straightforwardly that \( Z^2 \) constructed in such a way satisfies both (3.15) and (3.16), hence \( Z^1 \) dominates \( Z^2 \).

Conversely, let \( Z^2 = \sum_{i=0}^{n-1} Z^2_{\tau^i} 1_{[\tau^i,\tau^{i+1})} \in \mathcal{Z}^n_t(z_2) \) be arbitrarily chosen. We define \( Z^1 \in \mathcal{Z}^n_t(z_1) \) recursively as follows. First, let \( Z^1_{\tau^0} \triangleq z_1 \). Assume we have defined \( Z^1_{\tau^i} \) such that either \( Z^1_{\tau^i} = Z^2_{\tau^i} \) or \( 0 < Z^1_{\tau^i} < Z^2_{\tau^i} \). If \( \tau^i+1 = T \) or \( Z^1_{\tau^i} = Z^2_{\tau^i} \), define \( Z^1_{\tau^{i+1}} \triangleright Z^2_{\tau^{i+1}} \). Now assume \( \tau^{i+1} < T \) and \( 0 < Z^1_{\tau^{i+1}} < Z^2_{\tau^{i+1}} \). Note that in this case \( Z^2_{\tau^{i+1}} \neq Z^1_{\tau^{i+1}} \). If \( Z^2_{\tau^{i+1}} < Z^1_{\tau^{i+1}} \), define \( Z^1_{\tau^{i+1}} \triangleright Z^2_{\tau^{i+1}} \). Otherwise, we have \( Z^1_{\tau^{i+1}} > Z^2_{\tau^{i+1}} \), then define \( Z^1_{\tau^{i+1}} \triangleright Z^2_{\tau^{i+1}} \). Note that we still have either \( Z^1_{\tau^{i+1}} = Z^2_{\tau^{i+1}} \) or \( 0 < Z^1_{\tau^{i+1}} < Z^2_{\tau^{i+1}} \), so we may continue to define \( Z^1 \). One may check that (3.16) still holds, and for each \( \omega \), there exists \( k \) such that

\[
\delta Z_{\tau_0} = \cdots = \delta Z_{\tau_k} = \delta z \quad \text{and} \quad \delta Z_{\tau_{k+1}} = \cdots = \delta Z_{\tau_n} = 0,
\]

(3.22)
which obviously implies (3.17). The proof is now complete.

We note that, by (3.22) we actually have, for \( z_1 > z_2 > 0 \) and thus similarly for \( z_1 < z_2 < 0 \),

\[
V^n(t, x, y, z_2) - V^n(t, x, y, z_1) \leq C[|\delta z| + \rho(|\delta z|)],
\]

(3.23)

This will be important in Theorem 5.6 below.

We will need the following result in next section.

**Proposition 3.6** Assume (H1)-(H3). Then for any \( n \) and any \( (t, x, y) \),

\[
V^n(t, x, y, 0+) \leq V^n(t, x, y, 0); \quad V^n(t, x, y, 0-) \leq V^n(t, x, y, 0).
\]

**Proof.** First by Proposition 3.5 we know \( V^n(t, x, y, 0+) \) and \( V^n(t, x, y, 0-) \) exist.

For \( z > 0 \) and \( Z^1 = \sum_{i=0}^{n} Z^1_{\tau_i}1_{[\tau_i, \tau_{i+1}]} \in Z^n(z) \), we define \( Z^2 \in Z^n(0) \) as follows. Let

\[
k = \inf \{i : Z^1_{\tau_i} \leq 0\}. \quad \text{We note that} \quad k \leq n \quad \text{since} \quad Z_{\tau_k} = 0.
\]

Define \( Z^2_s = Z^1_s - Z^2 \) for \( s \geq \tau_k \). One can check straightforwardly that

\[
0 \leq \delta Z_{\tau_i} \leq z, \quad i = 0, \cdots, n, \quad \text{and that} \quad \Delta Z^1_{\tau_i} \Delta Z^2_{\tau_i} \geq 0.
\]

Note that

\[
Y^t_{\tau_i, y, Z^1} - Y^t_{\tau_i, y, Z^2} = \int_{\tau_i}^{\tau_k} \delta Z_s dX^t_s + \sum_{i=0}^{k} [c(\Delta Z^2_s) - c(\Delta Z^1_s)].
\]

For \( i \leq k \), by Definition 2.3 we have \( Z^1_{\tau_i} \neq Z^1_{\tau_{i-1}} \). If \( \Delta Z^1_{\tau_i} \Delta Z^2_{\tau_i} > 0 \), then by Assumption (H3) (iii) we get

\[
c(\Delta Z^2_{\tau_i}) - c(\Delta Z^1_{\tau_i}) \leq \rho(|\Delta Z^2_{\tau_i} - \Delta Z^1_{\tau_i}|) = \rho(|\delta Z_{\tau_{i-1}} - \delta Z_{\tau_i}|) \leq \rho(z).
\]

If \( \Delta Z^1_{\tau_i} \Delta Z^2_{\tau_i} = 0 \), then \( \Delta Z^2_{\tau_i} = 0 \) and thus

\[
c(\Delta Z^2_{\tau_i}) - c(\Delta Z^1_{\tau_i}) \leq 0 \leq \rho(z).
\]

Therefore,

\[
E\{U(Y^t_{\tau_i, y, Z^1}) - V^n(t, x, y, 0) \leq E\{U(Y^t_{\tau_i, y, Z^1}) - U(Y^t_{\tau_i, y, Z^2})\}
\]

\[
= E\{U'(\cdot)[Y^t_{\tau_i, y, Z^1} - Y^t_{\tau_i, y, Z^2}]\} \leq E\{U'\left[\int_{\tau_i}^{\tau_k} \delta Z_s dX^t_s + k\rho(z)\right]\}
\]

\[
\leq E\left[\int_{\tau_i}^{\tau_k} \delta Z_s dX^t_s + n\rho(z)\right] \leq C[z + n\rho(z)].
\]

This implies that

\[
V^n(t, x, y, z) - V^n(t, x, y, 0) \leq C[z + n\rho(z)],
\]

and thus \( V^n(t, x, y, 0+) \leq V^n(t, x, y, 0) \). Similarly, we can prove \( V^n(t, x, y, 0-) \leq V^n(t, x, y, 0) \). The proof is now complete.
4 The Approximating Optimal Strategies

In this section we solve the approximating problems. More specifically we shall construct the optimal strategy \( Z^n \in \mathcal{Z}^n(z) \) for each \( n \). We will provide the uniform estimate on \( Z^n \)'s in next section.

We begin by making the following observation. For any \( n > 1 \), \( (t, x, y, z) \in [0, T] \times \mathbb{R}^2 \times [-M, M] \), and \( Z \in \mathcal{Z}^n(z) \), we can write \( Z_n = z1_{[z_n,\tau_n]}(s) + \tilde{Z}_s1_{[\tau_n, T]}(s) \). Then, we have \( \tilde{Z} \in \mathcal{Z}^{n-1}(\tau_n) \), and we can write

\[
Y^t,x,y,z = y + \int_t^T Z_s dX^t_s - \sum_{i=0}^{\infty} c(\Delta Z_{\tau_i})
\]

where \( \bar{Y} \in \mathbb{Z} \)

\[
\bar{Y} = y + z(X^t_t - x) - c(\tilde{Z}_{\tau_1} - z) + \int_{\tau_1}^T \tilde{Z}_s dX^t_s - \sum_{i=2}^{\infty} c(\Delta \tilde{Z}_{\tau_i}).
\]

Let us denote \( X \triangleq X^t_t \) and \( Y_s \triangleq y + z[X_s - x] \) for \( s \in [0, T] \). Then (4.1) states that

\[
E\{Y_s(X^t_s, Y_s, \bar{Y})\} = E\{U(Y^t_t, Y_{\tau_1} - c(\bar{Y}_1 - z), \tilde{Z})\}
\]

where \( \bar{Y} = \sup_{\bar{z} \in [-M, M]} V^{n-1}(t, x, y - c(\bar{z} - z), \bar{z}) \). Thus it follows that

\[
V^n(t, x, y, z) \leq \sup_{\tau \geq t} E\{V^{n-1}(\tau, X_{\tau}, Y_{\tau}, z)\}, \text{ for all } (t, x, y, z).
\]

We shall prove that the equality actually holds. Thus relation (4.2) (with equality) can be thought of as a variation of the Bellman Principle, and it indicates a possible recursive relation among the approximating problems.

In light of the function \( V^{n-1} \), we now define, for any function \( \varphi(t, x, y, z) \),

\[
\bar{\varphi}(t, x, y, z) \triangleq \sup_{\bar{z} \in [-M, M]} \varphi(t, x, y - c(\bar{z} - z), \bar{z}).
\]

The following lemma plays an important role in our discussion.

**Lemma 4.1** Assume (H3). Suppose that a function \( \varphi : [0, T] \times \mathbb{R}^3 \to \mathbb{R} \) enjoys the following properties:

(i) The mapping \( y \mapsto \varphi(t, x, y, z) \) is increasing, for fixed \( (t, x, z) \);

(ii) The mapping \( (t, x, y) \mapsto \varphi(t, x, y, z) \) is uniformly continuous, uniformly in \( z \);

(iii) The mapping \( z \mapsto \varphi(t, x, y, z) \) is uniformly continuous in \([-M, 0]\) and in \((0, M]\), for fixed \( (t, x, y) \); and
(iv) \( \varphi(t, x, y, 0+) \leq \varphi(t, x, y, 0), \varphi(t, x, y, 0-) \leq \varphi(t, x, y, 0) \).

Then \( \tilde{\varphi} \) defined by (4.3) is uniformly continuous in \((t, x, y)\), uniformly in \(z\). Moreover, there exists a Borel measurable function \( \psi(t, x, y, z) \) such that \( |\psi| \leq M \) and

\[
\tilde{\varphi}(t, x, y, z) = \varphi(t, x, y - c(\psi(t, x, y, z) - z)) = \psi(t, x, y, z)).
\] (4.4)

**Proof.** First note that by assumption (ii) and (H3) we see that \( \varphi(t, x, y - c(\tilde{z} - z), \tilde{z}) \)
is uniformly continuous in \((t, x, y)\), uniformly in \((z, \tilde{z})\). Thus \( \tilde{\varphi} \) is uniformly continuous in \((t, x, y)\), uniformly in \(z\).

It remains to construct the function \( \psi \). To this end, for any fixed \((t, x, y, z)\), we denote \( \tilde{\varphi}(\tilde{z}) = \varphi(t, x, y - c(\tilde{z} - z), \tilde{z}) \). Without loss of generality we assume \( z > 0 \). Then \( \tilde{\varphi} \) is uniformly continuous in \([-M, 0)\), in \((0, z)\), and in \((z, M)\). We extend the function \( \tilde{\varphi} \) to \([-M, 0], [0, z], \) and \([z, M]\) in the following way. Define

\[
\begin{align*}
\varphi_1(\tilde{z}) &\triangleq \tilde{\varphi}(\tilde{z})1_{[-M,0]}(\tilde{z}) + \tilde{\varphi}(0\mathbin{-})1_{\{0\}}(\tilde{z}); \\
\varphi_2(\tilde{z}) &\triangleq \tilde{\varphi}(\tilde{z})1_{(0,z]}(\tilde{z}) + \tilde{\varphi}(0\mathbin{+})1_{\{0\}}(\tilde{z}) + \tilde{\varphi}(z\mathbin{-})1_{\{z\}}(\tilde{z}); \\
\varphi_3(\tilde{z}) &\triangleq \tilde{\varphi}(\tilde{z})1_{(z,M]}(\tilde{z}) + \tilde{\varphi}(z\mathbin{+})1_{\{z\}}(\tilde{z}).
\end{align*}
\]

Then one can easily check that \( \varphi_1, \varphi_2, \) and \( \varphi_3 \) are uniformly continuous on \([-M, 0], [0, z], \) and \([z, M]\), respectively. Furthermore, since each \( \varphi_i \) attains its supremum on their respective domain, and applying the measurable selection theorem if necessary, we can assume that the maximizers are all Borel measurable functions of the parameters \((t, x, y, z)\). More precisely, there exist Borel measurable functions \( \psi_i(t, x, y, z), i = 1, 2, 3 \), taking values in \([-M, 0], [0, z], \) and \([z, M]\), respectively, such that

\[
\tilde{\varphi}_1(\psi_1) = \sup_{\tilde{z} \in [-M,0]} \varphi_1(\tilde{z}); \quad \tilde{\varphi}_2(\psi_2) = \sup_{\tilde{z} \in [0,z]} \varphi_2(\tilde{z}); \quad \tilde{\varphi}_3(\psi_3) = \sup_{\tilde{z} \in [z,M]} \varphi_3(\tilde{z}),
\]

and consequently

\[
\tilde{\varphi}(t, x, y, z) = \sup_{\tilde{z} \in [-M,M]} \tilde{\varphi}(\tilde{z}) = \max \left\{ [\varphi_1(\psi_1), \varphi_2(\psi_2), \varphi_3(\psi_3)](t, x, y, z), \tilde{\varphi}(0), \tilde{\varphi}(z) \right\}.
\] (4.5)

Let us now define a maximizer function \( \psi \) as follows (suppressing all variables):

\[
\psi = \begin{cases} 
0 & \text{if } \tilde{\varphi} = \tilde{\varphi}(0) \\
\psi_1 & \text{if } \tilde{\varphi} = \varphi_1(\psi_1) > \max \{ \tilde{\varphi}(0), \tilde{\varphi}(z) \} \\
\psi_2 & \text{if } \tilde{\varphi} = \varphi_2(\psi_2) > \max \{ \tilde{\varphi}(0), \tilde{\varphi}(z), \varphi_1(\psi_1) \} \\
\psi_3 & \text{if } \tilde{\varphi} = \varphi_3(\psi_3) > \max \{ \tilde{\varphi}(0), \tilde{\varphi}(z), \varphi_1(\psi_1), \varphi_2(\psi_2) \}.
\end{cases}
\]
We note, from the definition of \( \psi \)'s, that the only possible places where the function \( \psi \) could be multiply defined are when the maximizers \( \tilde{\varphi} \)'s occur at 0 or \( z \). But since
\[
\varphi_1(0) = \tilde{\varphi}(0) = \varphi(t, x, y - c(-z), 0) \leq \varphi(t, x, y - c(-z), 0) = \tilde{\varphi}(0);
\]
\[
\varphi_2(z) = \tilde{\varphi}(z) = \varphi(t, x, y - c(0), z) \leq \varphi(t, x, y, z) = \tilde{\varphi}(z),
\]
and similarly, \( \varphi_2(0) \leq \tilde{\varphi}(0) \), \( \varphi_3(z) \leq \tilde{\varphi}(z) \), one can then show that \( \psi \) is well-defined for all \((t, x, y, z)\). In fact, if for example \( \varphi_1(\psi_1(t, x, y, z)) > \tilde{\varphi}(0) \), then \( \psi_1(t, x, y, z) \in [-M, 0) \). Similarly, if \( \varphi_2(\psi_2(t, x, y, z)) > \max(\tilde{\varphi}(0), \tilde{\varphi}(z)) \), then \( \psi_2(t, x, y, z) \in (0, z) \); and if \( \varphi_3(\psi_3(t, x, y, z)) > \tilde{\varphi}(z) \), then \( \psi_3(t, x, y, z) \in (z, M] \). Consequently \( \psi \) is the desired Borel measurable implicit function.

We now give the main existence result.

**Theorem 4.2** Assume \((H1)-(H3)\). For each \( n \) and any fixed \((t, x, y, z)\), there exists optimal \( Z^n \in Z^n_t(z) \) such that \( V^n(t, x, y, z) = E\{U(Y^n_{t,x,y,Z^n})\} \).

**Proof.** We prove by induction on \( n \). When \( n = 1 \), \( Z^1_t(z) \) is a singleton and thus \( Z^1_s \triangleq z\mathbf{1}_{(t,T)}(s) \). Assume the result is true for \( n - 1 \).

By virtue of Propositions 3.2, 3.5, and 3.6, we see that all \( V^n \)'s satisfy the condition of Lemma 4.1. Thus we conclude that \( \tilde{V}^{n-1} \) is uniformly continuous in \((t, x, y)\), thanks to Lemma 4.1. Next, recall (3.14) and note that by \((H3)\),
\[
\tilde{V}^{n-1}(T, x, y, z) = \sup_{\hat{z}} V^{n-1}(T, x, y - c(\hat{z} - z), \hat{z}) = \sup_{\hat{z}} U(y - c(\hat{z} - z) - c(-\hat{z})) = U(y - c(-z)) = V^n(T, x, y, z).
\]

We now consider the optimal stopping problem:
\[
\tilde{V}^{n-1}(t, x, y, z) = \sup_{\tau \geq t} E\{\tilde{V}^{n-1}(\tau, X_\tau, Y_\tau, z)\}.
\]

Since \( t \mapsto \tilde{V}^{n-1}(t, X_t, Y_t, z) \), is continuous, uniformly in \( z \), thanks to Lemma (4.1), it is known that (cf., e.g., [9, Appendix D]), we know that an optimal stopping time for (4.7) is
\[
\tau^n_1 \triangleq \inf\{s \geq t : \tilde{V}^{n-1}(s, X_s, Y_s, z) = \tilde{V}^{n-1}(s, X_s, Y_s, z)\}.
\]

Moreover, if we define \( Z^n_{\tau^n_1} \triangleq \psi(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1}, z) \) and \( Y^n_{\tau^n_1} \triangleq Y_{\tau^n_1} - c(Z^n_{\tau^n_1} - z) \), with \( \psi \) being the maximizer function in Lemma 4.1, we see that
\[
\tilde{V}^{n-1}(t, x, y, z) = E\{\tilde{V}^{n-1}(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1}, z)\} = E\{V^{n-1}(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1}, z, Z^n_{\tau^n_1})\} \leq V^n(t, x, y, z).
\]
This, together with (4.2), shows that
\[ \hat{V}^{n-1}(t,x,y,z) = V^{n}(t,x,y,z) = E\left\{ \hat{V}^{n-1}(\tau_i^n, X_{\tau_i^n}, Y_{\tau_i^n}, z) \right\}. \]
Namely the equality holds in (4.2). We can now argue by induction that there exist \( \tau_i^n, i = 1, \ldots, n \), \( Z^n \), so that the corresponding \( Y^n \) satisfies \( V^n(t,x,y,z) = E\{U(Y^n_T)\} \).

To conclude the proof we note that the \( \tau^n \) and \( Z^n \) constructed above may violate Definition 2.3 (i) and (iii). However, by the standard procedure introduced in the paragraph before Definition 2.3, by otherwise redefining \( \tau^n \) and \( Z^n \) we may assume without loss of generality that \( Z^n \in Z^n_t(z) \). The proof is now complete.

5 Regularity of the Value Function

In this section we give some uniform estimates of the value function \( V \). We should note that the regularities of \( V \) with respect to the variables \( (t,x,y) \) are clear, since the estimates (3.8), (3.9), and (3.10) in Proposition 3.2 are already uniform with respect to \( n \). The estimate (3.19), however, depends heavily on \( n \). In fact, in the case \( |z| = |z|^\alpha, 0 < \alpha < 1 \), one can check that \( \rho_n(|z|) = n^{1-\alpha}|z|^{\alpha} \rightarrow \infty \). Therefore the regularity of \( V \) with respect to \( z \) is by no means clear.

We first take a closer look at the approximating optimal strategies \( \{Z^n\}_{n=1}^\infty \). Since our purpose is to construct the optimal piecewise constant control, it is thus conceivable that uniform bound on the number of jumps for the approximating sequence \( \{Z^n\}_{n=1}^\infty \) would be extremely helpful.

We begin by considering the case where a fixed cost is present. Let us denote, for any \( Z \in Z_t(z) \), \( N(Z) \triangleq \sum_{i=1}^{\mathbb{1}\{Z_0 \neq Z_{i-1}\}} \) to be the number of jumps of \( Z \). For each \( (t,x,y,z) \), we denote \( Z^n \) to be the optimal portfolio for \( V^n(t,x,y,z) \), when the context is clear.

**Proposition 5.1** Assume (H1)–(H3), and assume further that \( c(z) \geq c_0 > 0 \) for any \( z \neq 0 \). Then there exists a constant \( C > 0 \) such that
\[ E\{N(Z^n)\} \leq \frac{C}{\lambda c_0}, \quad \text{for all } n \text{ and all } (t,x,y,z). \] (5.1)

**Proof.** The proof is straightforward. Note that \( V^n \)'s are non-decreasing in \( n \), and when \( n = 0 \), \( Z^1(z) = \{z\} \), a singleton. Thus, denoting \( Y_T^0 = Y_T^{t,x,y,z} = y + z[X_T^{t,x} - x] - c(z) \) and \( Y_T^n = Y_T^{t,x,y,Z^n} = y + \int_t^T Z_s^n dX_s^{t,x} - \sum_{i=0}^n c(\Delta Z^n_i) \), we have
\[
0 \geq V^0(t,x,y,z) - V^n(t,x,y,z) = E\{U(Y_T^0)\} - E\{U(Y_T^n)\}
\]
The result follows immediately. 

In the general case where \( c \) does not have a positive lower bound, the situation is more complicated. We need the following extra technical assumptions.

**H4** Assume that there exist constants \( \varepsilon_0 > 0 \) and \( \alpha_0 > 0 \) such that, with

\[
C_0 \overset{\Delta}{=} \frac{\Lambda}{\chi} [\|b\|_\infty T + \|\sigma\|_\infty \sqrt{T}] + 1; \quad C_1 \overset{\Delta}{=} 2C_0[1 + \frac{\Lambda}{\lambda\alpha_0}],
\]

(i) \( c \) is increasing in \((0, M]\) and decreasing in \([-M, 0)\);

(ii) \( c(z) \geq C_0|z|, \; \forall 0 < |z| < \varepsilon_0; \)

(iii) \( c(z_1) + c(z_2) - c(z_1 + z_2) \geq C_1|z_1|, \; \forall 0 < |z_1| < \varepsilon_0, |z_2| \geq \frac{1}{2}|z_1|; \)

(iv) For any \( 0 < |z_1| < \varepsilon_0, |z_3| \geq |z_2| \geq |z_1|, \)

\[
c(z_1) + c(z_2) - c(z_1 + z_2) \geq \alpha_0 \left[ |c(z_1 + z_3) - c(z_3)| \vee |c(\varepsilon_0, |z_1 + z_3|) - c(\varepsilon_0)\right].
\]

**Remark 5.2** (i) Assumptions (iii) and (iv) in (H4) are stronger versions of (v) in (H3).

(ii) If \( c(\cdot) \) satisfies (H3) and \( c(z) \geq c_0 \) for all \( z \neq 0 \), then (H4) (ii)-(iv) hold. In particular, the function \( c(z) = c_0 > 0 \) for any \( z \neq 0 \) satisfies (H4).

(iii) If \( c(z) = c_1|z|^{\alpha}, \; 0 < \alpha < 1 \), then it can be easily checked that (H3) and (H4) hold with

\[
\alpha_0 = \frac{2 - 2^\alpha}{2^\alpha - 1}; \quad \varepsilon_0 = \left[ \frac{c_1(1 - \alpha)^2}{C} \right]^\frac{1}{\alpha n}.
\]

To extend Proposition 5.1 to the general cases without fixed costs, we need an analysis on the numbers of the small jumps. For this purpose let us define the following sets:

\[
A^n_i \overset{\Delta}{=} \{ 0 < |\delta Z^n_i| < \varepsilon_0 \}, \quad B^n_i \overset{\Delta}{=} \{ |\delta Z^n_i| \geq \varepsilon_0 \}, \quad i = 1, \ldots, n; \quad n > 0,
\]

where \( \varepsilon_0 > 0 \) is defined in (H4). The following result is crucial.

**Theorem 5.3** Assume (H1)-(H4). Then for any fixed \( m \),

\[
P\left( \sum_{i=0}^{n} 1_{A^n_i} \geq m \right) \leq \frac{1}{2^{m-1}}, \quad \forall n \geq m.
\]
The proof of Theorem 5.3 depends heavily on the following technical result, whose proof is quite lengthy and will be deferred to the last section in order not to distract the discussion too much.

**Proposition 5.4** Assume (H1)–(H3), and assume (H4) holds with constants \( \varepsilon_0, C_0, \) and \( C_1 \). Let \( A_i^n, B_i^n, i = 1, \cdots, n \) be defined by (5.3). Then for each \( n \) it holds that

(i) For each \( i \), \( P \)-a.s. in \( A_i^n \) one has \( P\{B_{i+1}^n | \mathcal{F}_{\tau_i}\} \leq \frac{C_0}{C_1} < \frac{1}{2} \).

(ii) \( P \)-a.s. in \( A_i^n, Z_{\tau_i}^n = 0. \)

The proof of Theorem 5.3 depends heavily on the following technical result, whose proof will be deferred to the last section in order not to distract the discussion too much.

**Proof of Theorem 5.3.** To simplify notations, in what follows let us denote \( A_i = A_i^n \) and \( B_i = B_i^n, i = 1, \cdots, n \), when the context is clear. Define \( k_0 \triangleq 0 \), and

\[
k_{j+1} \triangleq \inf\{i > k_j : 0 < |\Delta Z_i^n| < \varepsilon_0 \} \wedge (n + 1), \quad j = 0, 1, \cdots, n - 1.
\]

Then \( P\left(\sum_{i=0}^{n} 1_{A_i} \geq m\right) = P(k_m \leq n) \). We claim that, for each \( j < n \),

\[
\{k_{j+1} \leq n\} \subseteq A_{k_j} \cap B_{k_j+1}, \quad P\text{-a.s.} \tag{5.5}
\]

(It is important to note here that in the left side it is \( k_{j+1} \) while in the right side the superscript of \( B \) is \( j+1 \) !) Indeed, we first note that \( \{k_{j+1} \leq n\} \subseteq \{k_j \leq n\} \subseteq A_{k_j} \), and consider the set \( A_{k_j} \setminus B_{k_j+1} \).

Suppose that \( Z_{\tau_{k_j}+1}^n \neq Z_{\tau_{k_j}}^n \) on \( A_{k_j} \setminus B_{k_j+1} \). Then \( 0 < |Z_{\tau_{k_j}+1}^n - Z_{\tau_{k_j}}^n| < \varepsilon_0 \), and by Proposition 5.4-(ii) we must have both \( Z_{\tau_{k_j}}^n = 0 \) and \( Z_{\tau_{k_j}+1}^n = 0 \), \( P \)-a.s., a contradiction. Thus we must have \( Z_{\tau_{k_j}+1}^n = Z_{\tau_{k_j}}^n \) on \( A_{k_j} \setminus B_{k_j+1} \). But then by Definition 2.3 (iii) we know \( \tau_{k_j+1} = T \) and thus \( Z_{\tau_{k_j}}^n = Z_{\tau_{k_j}+1}^n = \cdots = Z_{\tau_n}^n = 0 \). Namely \( k_{j+1} = n + 1 \). In other words, \( A_{k_j} \setminus B_{k_j+1} \subseteq \{k_{j+1} = n + 1\} \). Note that \( \{k_{j+1} \leq n\} \subseteq A_{k_j} \setminus \{k_{j+1} = n + 1\} \), (5.5) follows.

Next, we apply Proposition 5.4-(i) to get, for \( k_j < n \),

\[
P(B_{k_j+1} | \mathcal{F}_{\tau_{k_j}}) < \frac{1}{2}, \quad P\text{-a.s. in } A_{k_j}.
\]

Then it follows that

\[
P\left(\sum_{i=1}^{m} 1_{A_i} \geq m\right) = P(k_m \leq n) \leq P\left(\bigcap_{j=1}^{m-1} [A_{k_j} \cap B_{k_j+1}]\right)
\]

\[
= E\left\{\prod_{j=1}^{m-1} [1_{A_{k_j}} 1_{B_{k_j+1}}]\right\} = E\left\{\prod_{j=1}^{m-2} [1_{A_{k_j}} 1_{B_{k_j+1}}] 1_{A_{k_{m-1}}} E\{1_{B_{k_{m-1}+1}} | \mathcal{F}_{\tau_{k_{m-1}}}\}\right\}
\]

\[
\leq E\left\{\prod_{j=1}^{m-2} [1_{A_{k_j}} 1_{B_{k_j+1}}] 1_{A_{k_{m-1}}} \frac{1}{2}\right\} \leq \frac{1}{2} E\left\{\prod_{j=1}^{m-2} [1_{A_{k_j}} 1_{B_{k_j+1}}]\right\}.
\]

Repeating the argument \( m - 1 \) more times we prove the theorem. \( \blacksquare \)

The following theorem is a generalized version of Proposition 5.1.
Theorem 5.5 Assume Assumptions (H1)–(H4). Then it holds that

\[ E\{N(Z^n)\} \leq C[1 + \frac{1}{c(\varepsilon_0) \wedge c(-\varepsilon_0)}] < \infty, \quad \forall n. \]

Proof. Let \( \varepsilon_0 \) be that in (H4). Denote

\[ N_1(Z^n) \triangleq \sum_{i=0}^{n} 1_{A_i}, \quad N_2(Z^n) \triangleq \sum_{i=0}^{n} 1_{B_i}. \]

Then \( E\{N(Z^n)\} = E\{N_1(Z^n)\} + E\{N_2(Z^n)\} \). Note that Theorem 5.3 implies that

\[ E\{N_1(Z^n)\} = \sum_{m=1}^{n} P(N_1(Z^n) \geq m) \leq \sum_{m=1}^{n} \frac{1}{2^{m-1}} \leq 2, \quad (5.6) \]

we need only find \( E\{N_2(Z^n)\} \). But since \( N_2 \) contains only big jumps, we can estimate \( E(N_2(Z^n)) \) along the lines as Proposition 5.1. Indeed, note that

\[ E\{U(y + \int_t^T Z^n_s dX_s) - V^n(t, x, y, z)\} \]

\[ = E\left\{U(y + \int_t^T Z^n_s dX_s) - U(y + \int_t^T Z^n_s dX_s - \sum_i c(\Delta Z^n_i))\right\} \]

\[ \geq \lambda E\left\{\sum_i c(\Delta Z^n_i)\right\} \geq \lambda E\left\{\sum_i (c(\varepsilon_0) \wedge c(-\varepsilon_0)) 1_{B_i}\right\} \]

\[ = \lambda [c(\varepsilon_0) \wedge c(-\varepsilon_0)] E\{N_2(Z^n)\}. \]

On the other hand, since

\[ E\{U(y + \int_t^T Z^n_s dX_s) - V^n(t, x, y, z)\} \leq E\{U(y + \int_t^T Z^n_s dX_s) - V^1(t, x, y, z)\} \]

\[ = E\left\{U(y + \int_t^T Z^n_s dX_s) - U(y + \int_t^T z dX_s - c(-z))\right\} \]

\[ \leq \Lambda E\left\{ |\int_t^T (Z^n_s - z) dX_s| + c(-z)\right\} \leq CA, \]

we obtain that

\[ E\{N_2(Z^n)\} \leq \frac{CA}{\lambda [c(\varepsilon_0) \wedge c(-\varepsilon_0)]}. \]

This, together with (5.6), proves the theorem. \( \blacksquare \)

As a direct consequence of Theorem 5.5, we can now prove the second main result of this section.

Theorem 5.6 Assume (H1)–(H3). Assume further that either \( c(z) \geq c_0 > 0 \), for all \( z \neq 0 \) or (H4) holds. Then there exists a generic constant \( C > 0 \), such that for any \( z_1, z_2 \) with the same sign, and for all \( n \), it holds that

\[ |V^n(t, x, y, z_1) - V^n(t, x, y, z_2)| \leq C[|\delta z| + \rho(|\delta z|)]. \quad (5.7) \]

23
Furthermore, one has

\[ V(t, x, y, z) - V^n(t, x, y, z) \leq \frac{C}{n}. \]  

(5.8)

Proof. Without loss of generality, assume \( z_1 > z_2 > 0 \). We first recall from (3.23) in Theorem 3.5 that the following inequality holds:

\[ V^n(t, x, y, z_2) - V^n(t, x, y, z_1) \leq C[|\delta z| + \rho(|\delta z|)]. \]

We need only check the other half of the inequality. To this end, let \( Z^1 \) be the optimal strategy of \( V^n(t, x, y, z_1) \), and as in Theorem 3.5 we define \( Z^2 \in Z(z_2) \) that is “dominated” by \( Z^1 \). We note that, for \( i > N(Z^1) \), \( Z^1_{\tau_i} = Z^1_{\tau_{i-1}} \), which implies that \( Z^2_{\tau_i} = Z^2_{\tau_{i-1}} \). Then

\[
\begin{align*}
V^n(t, x, y, z_1) - V^n(t, x, y, z_2) &\leq E\{U(Y^n_{t,x,y,Z^1}) - U(Y^n_{t,x,y,Z^2})\} \\
&\leq \Lambda E\left\{\int_t^T \delta Z_s dX_s + \sum_{i=0}^{N(Z^1)} c(\Delta Z^1_{\tau_i} - c(\Delta Z^2_{\tau_i}))\right\} \\
&\leq C|\delta z| + CE\left\{\sum_{i=0}^{N(Z^1)} \rho(|\Delta Z^1_{\tau_i} - \Delta Z^2_{\tau_i}|)\right\} \\
&\leq C|\delta z| + C\rho(|\delta z|)E\{N(Z^1)\} \leq C[|\delta z| + \rho(|\delta z|)],
\end{align*}
\]

where the last inequality is due to Theorems 5.1 and 5.5. This proves (5.7).

To prove (5.8), we denote, for any \( m > n \), \( Z^m = \sum_{i=0}^{n} Z^m_{\tau_{i-1}} \delta z \) be the optimal strategy of \( V^m(t, x, y, z) \). Define \( Z^m_{s,n} \equiv Z^m_s 1_{s < \tau_n} \). Then \( Z^m_{s,n} \in Z^n(t, z) \), and

\[
\begin{align*}
Y^n_{t} - Y^n_{T} &\leq \left[ \int_{\tau_n}^T Z^n_s dX_s + c(-Z^n_{\tau_{n-1}}) \right] 1_{\{\tau_n < T\}} \\
&\leq \left[ \int_{\tau_n}^T Z^n_s dX_s + c(-Z^n_{\tau_{n-1}}) \right] 1_{\{\tau_n < T\}}.
\end{align*}
\]

Note that \( \{\tau_n < T\} = \{N(Z^m) > n\} \), it follows that

\[
\begin{align*}
V^m(t, x, y, z) - V^n(t, x, y, z) &\leq E\left\{U(Y^n_{t}) - U(Y^n_{T})\right\} \\
&= E\left\{U'(\cdot)Y^n_{t} - Y^n_{T}\right\} \leq E\{U'(|\cdot|)\left[ \int_{\tau_n}^T Z^n_s dX_s + c(-Z^n_{\tau_{n-1}}) \right] 1_{\{\tau_n < T\}}\} \\
&\leq CE\left\{ E\{\int_{\tau_n}^T Z^n_s dX_s\} 1_{\{\tau_n < T\}}\right\} \\
&\leq C\{\tau_n < T\} = CP\{N(Z^m) > n\} \leq \frac{C}{n} E\{N(Z^m)\} \leq \frac{C}{n}.
\end{align*}
\]

24
Sending $m \to \infty$ and applying Theorem 3.1, we obtain the result.

As the direct consequences of Theorems 3.2, 3.6, 4.2, and 5.6 we have

**Theorem 5.7** Assume $(H1)$–$(H3)$, and assume either $c(z) \geq c_0$, $z \neq 0$ or $(H4)$. Then, with

\[ \delta \xi = \xi_1 - \xi_2, \xi = x, y, z, \text{ and } t, \text{ respectively,} \]

(i) \[ |V(t, x_1, y, z) - V(t, x_2, y, z)| \leq C|\delta x|. \]
(ii) $\lambda \delta y \leq V(t, x, y_1, z) - V(t, x, y_2, z) \leq \Lambda \delta y$, $\delta y \triangleq y_1 - y_2 \geq 0$.
(iii) \[ |V(t_1, x, y, z) - V(t_2, x, y, z)| \leq C|\delta t|^{\frac{1}{2}}. \]
(iv) \[ |V(t, x, y, z_1) - V(t, x, y, z_2)| \leq C[|\delta z| + \rho(|\delta z|)], \forall z_1, z_2 \text{ such that } z_1 z_2 > 0. \]
(v) \[ V(t, x, y, 0+) \leq V(t, x, y, 0), V(t, x, y, 0-) \leq V(t, x, y, 0). \]

6 The Optimal Strategy $Z^*$

In this section we construct the optimal impulse control for the original problem. We should note that by virtue of Theorems 5.1 and 5.5, one can easily show that under our assumptions $Z^n$ should converge in distribution. But this does not seem to be helpful for our construction of the optimal strategy. In fact, in general we will have to extend the probability space, and it is not clear whether the limit process will have the desired adaptedness that is essential in our application.

To construct the optimal portfolio $Z^*$ directly, let us first note that if $V$ is the value function, then one can easily check that the following identity holds

\[ \tilde{V}(t, x, y, z) = V(t, x, y, z), \quad \forall (t, x, y, z), \quad (6.1) \]

where $\tilde{V}$ is defined by (4.3). However, in light of the construction of the optimal strategy $Z^n$, we know that the function $\tilde{V}$ should play the role of an “obstacle” that will trigger the jumps, as it is usually the case in impulse control literature. A more careful analysis of the function $\tilde{V}$ and the identity (6.1) is in order.

To this end let us consider the following set

\[ O(z) \triangleq \{ (t, x, y) : V(t, x, y, z) > V(t, x, y - c(\tilde{z} - z), \tilde{z}) \forall \tilde{z} \neq z \}. \quad (6.2) \]

Intuitively, the set $O(z)$ should define an “inaction region”, since a change of position (on $z$) would decrease the value function. Furthermore, following the standard impulse control theory one would expect that $O(z)$ is an open set and the trade will take place when $(t, x, y) \in \partial O(z)$. This is indeed the case when $c(z) \geq c_0 > 0$ for $z \neq 0$. However, unfortunately in our more general case we only have the following result.
Lemma 6.1 Assume (H1)–(H4). Define
\[ O_n(z) \triangleq \{ (t, x, y) : V(t, x, y, z) > V(t, x, y - c(\tilde{z} - z), \tilde{z}), \forall |\tilde{z} - z| \geq \frac{1}{n} \}. \] (6.3)

Then \( O_n(z) \) is open, for all \( n \), and \( O(z) = \bigcap_n O_n(z) \).

Proof. Denote
\[ V_n(t, x, y, z) \triangleq \sup_{||\tilde{z} - z|| \geq \frac{1}{n}} V(t, x, y - c(\tilde{z} - z), \tilde{z}). \] (6.4)

Apply Theorem 5.7 and follow the proof of Lemma 4.1, we know \( V_n \) is continuous in \((t, x, y)\) and there exists a Borel measurable function \( \psi_n \) such that \( |\psi_n(t, x, y, z) - z| \geq \frac{1}{n} \) and \( V(t, x, y - c(\psi_n(t, x, y, z) - z), \psi_n(t, x, y, z)) = V_n(t, x, y, z) \).

This implies that
\[ O_n(z) = \{ (t, x, y) : V(t, x, y, z) > V_n(t, x, y, z) \} \]
and thus \( O_n(z) \) is open. That \( O(z) = \bigcap_{n=1}^{\infty} O_n(z) \) is obvious. The proof is complete. \( \blacksquare \)

We remark that Lemma 6.1 does not imply that the set \( O(z) \) is an open set. Therefore, if we follow the scheme in the previous sections to define, for given \((t, x, y, z)\), \( X_s \triangleq X_{t,x}^s \), \( Y_s \triangleq y + z[X_s - x] \), and
\[ \tau \triangleq \inf\{ s \geq t : (s, X_s, Y_s) \notin O(z) \} \land T. \] (6.5)

Then it is possible that \( P(\tau = t) > 0 \) and/or \( P(\tau, X_\tau, Y_\tau) \in O(z) \) > 0. In either case the construction procedure will fail. The following Theorem is therefore essential.

Theorem 6.2 Assume (H1)–(H4). Define, for each \((t, x, y, z)\) and \( n > 0 \),
\[ \tau^n \triangleq \inf\{ s \geq t : (s, X_s, Y_s) \notin O_n(z) \} \land T. \] (6.6)

Then
(i) \( \tau^n \) are decreasing stopping times and, if \( \tau^n < T \), then \( (\tau^n, X_{\tau^n}, Y_{\tau^n}) \notin O_n(z) \).
(ii) \( \tau^n \nmid \tau \) and thus \( \tau \) is also a stopping time.
(iii) \( P(\tau^n > \tau, \forall n) = 0 \) and thus, \( P-a.s., (\tau, X_\tau, Y_\tau) \notin O(z) \) when \( \tau < T \).
(iv) \( V(t, x, y, z) = E\{ V(\tau, X_\tau, Y_\tau, z) \} \).

The proof of Theorem 6.2 will depend heavily on an important, albeit technical, lemma that characterizes the possible behavior of the small jumps under our basic assumptions on the transaction cost function. The proof of this lemma is again rather tedious, and we defer it to the last section.
Lemma 6.3 Assume (H1)-(H4). Suppose that for given \((t, x, y, z)\), \(\tilde{z}\) is such that \(0 < |\tilde{z} - z| < \varepsilon_0\) and \(V(t, x, y, z) = V(t, x, y - c(\tilde{z} - z), \tilde{z})\), where \(\varepsilon_0\) is that in (H4), then \(\tilde{z} = 0\).

[Proof of Theorem 6.2] (i) That \(\tau^n\)'s are decreasing stopping times is obvious by definition. Also, since each \(\mathcal{O}_n(z)\) is an open set, thanks to Lemma 6.1, it follows immediately that \((\tau^n, X_{\tau^n}, Y_{\tau^n}) \notin \mathcal{O}_n(z)\), whenever \(\tau^n < T\).

(ii) Denote \(\tau^\infty = \lim_{n \to \infty} \tau^n\). Since \(\mathcal{O}_n \supseteq \mathcal{O}\), we have \(\tau^n \geq \tau\) for any \(n\) and thus \(\tau^\infty \geq \tau\), P-a.s. Moreover, if \(\tau(\omega) < T\), then for any \(\varepsilon > 0\), there exists \(s < \tau(\omega) + \varepsilon\) such that \((s, X_s, Y_s) \notin \mathcal{O}(z)\). Since \(\mathcal{O}(z) = \bigcap_n \mathcal{O}_n(z)\), there exists \(n(\omega)\) such that \((s, X_s(\omega), Y_s(\omega)) \notin \mathcal{O}_n(z)\). Thus \(\tau^n(\omega) \leq s < \tau(\omega) + \varepsilon\) and therefore \(\tau^\infty(\omega) < \tau(\omega) + \varepsilon\). Since \(\varepsilon\) is arbitrary, we get \(\tau^\infty \leq \tau\), and hence \(\tau^\infty = \tau\). Note that the result is trivial on \(\{\tau = T\}\), the claim follows.

(iii) Choose \(n_0\) such that \(n_0 > \max\{\frac{1}{\varepsilon_0}, \frac{1}{|z|}, 1_{\{z \neq 0\}}\}\), and note that \(\{\tau^n > \tau, \forall n\} \subset \{\tau < T\}\). On \(\{\tau < T\}\), for \(n \geq n_0\) large enough, by (ii) we have \(\tau^n < T\) and thus there exists \(Z_{\tau^n}\) such that \(|Z_{\tau^n} - z| \geq \frac{1}{n_0}\) and

\[
V(\tau^n, X_{\tau^n}, Y_{\tau^n}, z) = V(\tau^n, X_{\tau^n}, Y_{\tau^n} - c(Z_{\tau^n} - z), Z_{\tau^n}).
\]

By Lemma 6.3, either \(Z_{\tau^n} = 0\) or \(|Z_{\tau^n} - z| \geq \varepsilon_0\). If \(z = 0\), then \(Z_{\tau^n} \neq 0\) and thus \(|Z_{\tau^n} - z| \geq \varepsilon_0 \geq \frac{1}{n_0}\). If \(z \neq 0\), then either \(|Z_{\tau^n} - z| = |z| \geq \frac{1}{n_0}\) or \(|Z_{\tau^n} - z| \geq \varepsilon_0 \geq \frac{1}{n_0}\). So in all the cases we have \(|Z_{\tau^n} - z| \geq \frac{1}{n_0}\). This implies that \(\tau^n = \tau^{n_0}\) for all \(n\) large enough. Therefore, \(\tau = \tau^{n_0}\) and thus \((\tau, X_\tau, Y_\tau) \notin \mathcal{O}(z)\).

(iv) We first note that, taking \(\tau\) as the first trading time, then we should have

\[
E\{V(\tau, X_\tau, Y_\tau, z)\} = \sup\{E\{U(Y^{1,2,3}_{1,2,3})\} : Z \in Z_t(z), Z_s = z \text{ for } \forall s \leq t\}.
\]

It then follows that \(E\{V(\tau, X_\tau, Y_\tau, z)\} \leq V(t, x, y, z)\).

On the other hand, let us fix \((t, x, y, z)\), and note that \(F\) is quasi-left continuous. Then we can choose a sequence of stopping times \(\tau_m \uparrow \tau\) such that \(\tau_m < \tau\) whenever \(\tau > t\). We claim that

\[
V(t, x, y, z) \leq E\{V(\tau_m, X_{\tau_m}, Y_{\tau_m}, z)\}. \quad (6.7)
\]

If this is true, then by sending \(m \to \infty\) we prove the theorem.

To prove (6.7), we recall (6.4) and choose \(n_0\) as in (iii). On the set \(\{\tau > t\}\) and for \(t \leq s < \tau\), denote

\[
I_s \triangleq V(s, X_s, Y_s, z) - V_{n_0}(s, X_s, Y_s, z).
\]

By the proof of Lemma 6.1 we have \(I_s > 0\). Since \(I\) is continuous in \(s\), we get

\[
I^m \triangleq \inf_{s \leq \tau_m} I_s > 0. \quad (6.8)
\]
For any \( n \geq n_0 \), let \( Z^n \) be the optimal portfolio of \( V^n(t,x,y,z) \) and let \( \tau^n_1 \) be the first jump time of \( Z^n \). By Proposition 5.4 (ii) and following similar arguments as in (iii), we have \(|Z^n_{\tau^n_1} - z| \geq \frac{1}{n_0} \) on \( \{ \tau^n_1 < T \} \). Then, for any \( m \), on \( \{ \tau^n_1 < \tau_m \} \subset \{ \tau^n_1 < T \} \), using (5.8) we have

\[
I^n_m \leq I^n_{\tau^n_1} = V(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1}, z) - V_{n_0}(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1}, z) \\
\leq V(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1}, z) - V(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1} - c(Z^n_{\tau^n_1} - z), Z^n_{\tau^n_1}) \\
\leq V(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1}, z) - V^{n-1}(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1} - c(Z^n_{\tau^n_1} - z), Z^n_{\tau^n_1}) \\
= V(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1}, z) - V(\tau^n_1, X_{\tau^n_1}, Y_{\tau^n_1}, z) \leq \frac{C}{n}.
\]

This, together with (6.8), implies that

\[
\lim_{n \to \infty} P(\tau^n_1 < \tau_m) = 0. \quad (6.9)
\]

Next, recall the proof of Theorem 4.2 that \( \tau^n_1 \) is a solution to an optimal stopping problem (4.7), and thus (cf. e.g., [8]), \( V^n(s,X_s,Y_s,z) \) is a martingale for \( t \leq \tau^n_1 \). Therefore

\[
V^n(t,x,y,z) = E\{V^n(\tau^n_1 \land \tau_m, X_{\tau^n_1 \land \tau_m}, Y_{\tau^n_1 \land \tau_m}, z)\}
\]

\[
= E\{V^n(\tau_m, X_{\tau_m}, Y_{\tau_m}, z)\} + \sum_{\tau_1 < \tau_m} E\{V^n(\tau_1, X_{\tau_1}, Y_{\tau_1}, z) - V(\tau_1, X_{\tau_1}, Y_{\tau_1}, z)\} 1_{\{\tau_1 < \tau_m\}}
\]

Applying Proposition 3.1 we then have

\[
V^n(t,x,y,z) \leq E\{V(\tau_m, X_{\tau_m}, Y_{\tau_m}, z)\} + C[1 + \sup_{t \leq s \leq T} |Y_s|] 1_{\{\tau^n_1 < \tau_m\}}.
\]

Sending \( n \to \infty \) and by (6.9) we obtain (6.7), whence the theorem.

We are now ready to construct the optimal strategy. Let \((t,x,y,z)\) be given. We construct an impulse control \(\{(\tau^n_s, Z^n_s)\}_{s=0}^\infty\) in the following steps.

- **Set** \( \tau^n_0 \triangleq t \). If \((t,x,y) \in O(z)\), then we define \( Z^n_{\tau^n_0} = Z^n_{\tau^n_0-} \triangleq z; X_s \triangleq X_s^{t,x}; \) and \( Y_s \triangleq y + z[X_s - x] \). Whereas if \((t,x,y) \notin O(z)\), namely there exists \( \tilde{z} \neq z \) such that \( V(t,x,y,z) = V(t,x,y - c(\tilde{z} - z), \tilde{z}) \), then we make an initial jump so that \( Z^n_{\tau^n_0} = \tilde{z} \).

- **Define** \( \tau^n_i \triangleq \inf\{s \geq t : (s,X_s,Y_s) \notin O(z)\} \land T \). Then \( \tau^n_i \) is a stopping time and \((\tau^n_i, X_{\tau^n_i}, Y_{\tau^n_i}) \notin O(z)\). Set \( Y^n_s \triangleq Y_s, Z^n_s \triangleq z, \) for \( s \in [t, \tau^n_i) \). Note that if \( \tau^n_i < T \), then by the definition of \( O(z) \) there exists \( Z^n_{\tau^n_i} \neq z \) such that

\[
V(\tau^n_i, X_{\tau^n_i}, Y^n_{\tau^n_i}, z) = V(\tau^n_i, X_{\tau^n_i}, Y^n_{\tau^n_i} - c((Z^n_{\tau^n_i} - z)), Z^n_{\tau^n_i})
\]
\* Assume we have defined $\tau^*_i$, $Y^*_s$ for $s < \tau^*_i$ and $Z^*_s$ for $s \leq \tau^*_i$. For $s \geq \tau^*_i$, set
\[
Y^*_s = Y^*_{\tau^*_i -} - c(Z^*_{\tau^*_i -} - Z^*_{\tau^*_i}) + Z^*_{\tau^*_i}[X_s - X_{\tau^*_i}].
\]
Define
\[
\tau^*_{i+1} \triangleq \inf\{s > \tau^*_i : (s, X_s, Y_s) \notin O(Z^*_{\tau^*_i})\} \land T.
\]
Similarly $\tau^*_{i+1}$ is a stopping time, and we can define $Y^*_s \triangleq Y^*_s$ and $Z^*_s \triangleq Z^*_{\tau^*_i}$, for
$s \in [\tau^*_i, \tau^*_{i+1})$. Again, if $\tau^*_{i+1} < T$, we choose $Z^*_{\tau^*_i} \neq Z^*_{\tau^*_{i+1}}$ such that
\[
V(\tau^*_{i+1}, X_{\tau^*_{i+1}}, Y^*_{\tau^*_{i+1}}, Z^*_{\tau^*_i}) = V(\tau^*_i, X_{\tau^*_i}, Y^*_{\tau^*_i}, Z^*_{\tau^*_i} - c(Z^*_{\tau^*_i} - Z^*_{\tau^*_i}), Z^*_{\tau^*_{i+1}}).
\]

If $\tau^*_{i+1} = T$, then we set $Z^*_{\tau^*_i} \triangleq 0$ and $Y^*_{\tau^*_{i+1}} \triangleq Y^*_{\tau^*_i} - c(-Z^*_{\tau^*_i})$.

Our main result is the following theorem. Since its proof depends on a group of lemmas that will be proved in section 7, we shall therefore defer it to the last section as well.

**Theorem 6.4** Assume (H1)–(H3), and assume either $c(z) \geq c_0 > 0$, for all $z \neq 0$, or (H4) is in force. Then

(i) $Z^* \in Z_t(z)$ and $E\{N(Z^*)\} < \infty$ satisfies the estimate in Theorem 5.1 or 5.5. In particular, $P(\tau^*_n < T, \forall n) = 0$.

(ii) $Z^*$ is an optimal impulse control. That is, $V(t, x, y, z) = E\{U(Y^*_T)\}$.

7 Proofs of Proposition 5.4, Lemma 6.3, and Theorem 6.4

In this section we collect the proofs of several technical results used in the previous sections. We note that these results are instrumental in the construction of the piecewise constant optimal strategy, and some of these results are of interest in their own right. As a matter of fact, many of these results can be considered as the necessary conditions of the optimality.

7.1 Proof of Proposition 5.4

We split the proof into several lemmas. To begin with, we fix $(t_0, x_0, y_0, z_0)$ and $n$, and let $Z^n$ be the optimal strategy of $V^n(t_0, x_0, y_0, z_0)$. Recall (5.3) (suppressing superscript “$n$”):
\[
A_i \triangleq \{0 < |\Delta Z^n_{\tau^*_i}| < \varepsilon_0\}, \quad B_i \triangleq \{|\Delta Z^n_{\tau^*_i}| \geq \varepsilon_0\}, \quad i = 1, \cdots, n.
\]

Throughout this section we assume that (H1)–(H4) are all in force. Keep in mind that our purpose is to show that on the set of small jumps (the set $A_i$’s) the jump will only happen when it jumps to 0.

29
Lemma 7.1 P-a.s. on A₁, either \([0 ∨ Z^n_{\tau_1}] ≤ Z^n_{\tau_{i-1}}\) or \(Z^n_{\tau_{i-1}} ≤ |Z^n_{\tau_1} ∧ 0|\).

Proof. Suppose that the lemma is not true. Then for some \(i > 0\), may assume without loss of generality that \(i = 1\), such that \(Z^n_{\tau_{i-1}} = Z^n_{\tau_0} = z_0 ≥ 0\), and \(P\{\{Z^n_{\tau_1} > z_0\} ∩ A_1\} > 0\).

Denote \(D_1 \triangleq \{Z^n_{\tau_1} > z_0\} ∩ A_1\) for simplicity (hence \(P(D_1) > 0\)).

We now define \(\tilde{Z}^n\) as follows. First, let \(k \triangleq \inf\{i ≥ 1 : Z^n_{\tau_i} ≤ 0\}\). Since \(Z^n_{\tau_0} = 0\), we have \(k ≤ n\). Now, set \(\tilde{Z}^n_{\tau_0} \triangleq z_0\), and \(\tilde{Z}^n_{\tau_i} \triangleq Z^n_{\tau_i} 1_{D_1} + Z^n_{\tau_i} 1_{D_2}\). For \(i ≥ 2\), we define

\[
\tilde{Z}^n_{\tau_i} \triangleq \begin{cases} 
\{[Z^n_{\tau_{i-1}} - Z^n_{\tau_i} + z_0] ∨ 0\} 1_{D_1} + Z^n_{\tau_i} 1_{D_2} & i < k; \\
Z^n_{\tau_i} & i ≥ k.
\end{cases}
\]

Then \(\tilde{Z}^n ∈ Z^n_i(z_0)\). We shall prove that \(\tilde{Z}^n\) will perform strictly better than \(Z^n\), which will contradict the optimality of \(Z^n\) and hence prove the lemma. To this end, denote \(δZ^n \triangleq \tilde{Z}^n - Z^n\). Then,

\[
\delta Y_T^n = Y_T^n - Y_T^n = \int_{\tau_1}^T \delta Z^n_s dX_s + \sum_{i=0}^n [c(\Delta Z^n_{\tau_i}) - c(\Delta \tilde{Z}^n_{\tau_i})].
\]

By definition of \(\tilde{Z}^n\) it is clear that \(\delta Y_T^n = 0\) on \(D_1\). On \(D_1\), first note that \(|δZ^n_{\tau_1}| ≤ Z^n_{\tau_1} - z_0\) for all \(i\). Further, for \(i > k\), one has \(Δ\tilde{Z}^n_{\tau_i} = ΔZ^n_{\tau_i}\); and for \(i ≤ k\), one can check that either \(0 ≤ Δ\tilde{Z}^n_{\tau_i} ≤ ΔZ^n_{\tau_i}\) or \(ΔZ^n_{\tau_i} ≤ Δ\tilde{Z}^n_{\tau_i} ≤ 0\). It then follows from (H4)-(i) that \(c(ΔZ^n_{\tau_i}) ≥ c(Δ\tilde{Z}^n_{\tau_i})\). Moreover, note that when \(i = 1\),

\[
c(ΔZ^n_{\tau_1}) - c(Δ\tilde{Z}^n_{\tau_1}) = c(Z^n_{\tau_1} - z) > C_0|Z^n_{\tau_1} - z_0|,
\]

thanks to Assumption (H4)-(ii). Thus, on \(D_1\),

\[
\delta Y_T^n ≥ \int_{\tau_1}^T \delta Z^n_s dX_s + c(Z^n_{\tau_1} - z) > \int_{\tau_1}^T \delta Z^n_s dX_s + C_0|Z^n_{\tau_1} - z_0|.
\]

Therefore, we have

\[
E\{U(\tilde{Y}_T^n) - U(Y_T^n)\} = E\{U'(\cdot)δY_T^n\} = E\{U'(\cdot) ΔY_T^n 1_{D_1}\}
\geq E\left\{ [λC_0|Z^n_{\tau_1} - z_0| - Λ| \int_{\tau_1}^T \delta Z^n_s dX_s |] 1_{D_1}\right\}
\geq E\left\{ [λC_0|Z^n_{\tau_1} - z_0| - ΛE_{τ_1}\{ | \int_{\tau_1}^T \delta Z^n_s dX_s |\}] 1_{D_1}\right\}
\]

Moreover, we see that on \(D_1\) it holds that

\[
ΛE_{τ_1}\{ | \int_{\tau_1}^T \delta Z^n_s dX_s |\} ≤ ΛE_{τ_1}\{ | \int_{\tau_1}^T |δZ^n_s b(s, X_s)| ds + | \int_{\tau_1}^T |δZ^n_s σ(s, X_s)| dW_s |\}
\leq ΛE_{τ_1}\{ \int_{τ_1}^T |δZ^n_s b(s, X_s)| ds \} + ΛE_{τ_1}\{ \int_{τ_1}^T |δZ^n_s σ(s, X_s)|^2 ds \} \frac{1}{2}
\leq Λ[\|b\|_∞ T + \|σ\|_∞ \sqrt{T}] |Z^n_{\tau_1} - z_0| = λ(C_0 - 1)|Z^n_{\tau_1} - z_0|.
\]

30
Thus $E\{U(Y_{T_2}^{Z_n}) - U(Y_{T_2}^{Z_n})\} \geq \lambda E\{[Z_{n_1}^n - z_0]1_{A_1}\} > 0$, a contradiction, proving the lemma.

Lemma 7.2 For any $\tilde{A}_i \subset A_i$, if $P(\tilde{A}_i) > 0$, then $P(\tilde{D}_{i+1}) > 0$, where

$$\tilde{D}_{i+1} \triangleq \{-1 \leq \frac{\Delta Z_{n_1}^n \Delta n_{i+1}}{\Delta Z_{n_1}^n} \leq \frac{1}{2}\} \cap \tilde{A}_i. \quad (7.3)$$

Consequently, P.a.s. in $A_i$, it holds that $|Z_{n_1}^n| \leq |\Delta Z_{n_1}^n|$.

Proof. Again we assume $i = 1$, $z_0 \geq 0$ and we will prove by contradiction. By Lemma 7.1, we have $Z_{n_1}^n < z_0$ in $A_1$, and thus also in $\tilde{A}_1$. Suppose that the result is not true, namely $P(\tilde{D}_2) = 0$. Then, with possibly an exception of a null set, one has

$$\tilde{A}_1 \subseteq \tilde{D}_{21} \cup \tilde{D}_{22} \triangleq (\{\Delta Z_{n_1}^n > z_0 - Z_{n_1}^n\} \cap A_1) \cup (\{\Delta Z_{n_1}^n < \frac{1}{2}[Z_{n_1}^n - z_0]\} \cap A_1).$$

Slightly different from the previous lemma, we now define $	ilde{Z}_{n_0}^n \triangleq z_0$; $\tilde{Z}_{n_1}^n \triangleq Z_{n_1}^n 1_{\tilde{A}_1} + z_0 1_{\tilde{A}_1}$; and $\tilde{Z}_{n_1}^n \triangleq Z_{n_1}^n (z_0)$, and

$$\delta Y_{T_2}^n = [z_0 - Z_{n_1}^n]X_{T_2} - X_{\tau_{n_1}} + c(Z_{n_1}^n - z_0) + c(Z_{n_1}^n - Z_{n_1}^n) - c(Z_{n_1}^n - z_0)]1_{\tilde{A}_1}.$$ 

Note that, on $\tilde{D}_{21}$, $Z_{n_1}^n > z_0 > Z_{n_1}^n$. Then (H4)-(i) and (ii) yield that

$$c(Z_{n_1}^n - z_0) + c(Z_{n_1}^n - Z_{n_1}^n) - c(Z_{n_1}^n - z_0) \geq c(Z_{n_1}^n - z_0) \geq C_0|Z_{n_1}^n - z_0|.$$ 

On the set $\tilde{D}_{22}$, however, one has $Z_{n_1}^n - Z_{n_1}^n < \frac{1}{2}[Z_{n_1}^n - z_0] < 0$. Then by (H4)-(iii) we have

$$c(Z_{n_1}^n - z_0) + c(Z_{n_1}^n - Z_{n_1}^n) - c(Z_{n_1}^n - z_0) \geq C_1|Z_{n_1}^n - z_0| \geq C_0|Z_{n_1}^n - z_0|.$$ 

So, P.a.s. in $\tilde{A}_1$,

$$\delta Y_{T_2}^n \geq [z_0 - Z_{n_1}^n]X_{T_2} - X_{\tau_{n_1}} + C_0|Z_{n_1}^n - z_0|.$$ 

Thus, following similar arguments as in Lemma 7.1, we have

$$E\{U(Y_{T_2}^{Z_n}) - U(Y_{T_2}^{Z_n})\} \geq E\left\{\lambda C_0|Z_{n_1}^n - z_0| - \lambda|z_0 - Z_{n_1}^n||X_{T_2} - X_{\tau_{n_1}}|\right\}1_{\tilde{A}_1} \geq \lambda E\{|Z_{n_1}^n - z_0|1_{\tilde{A}_1}\} > 0,$$ 

a contradiction. Hence $P(\tilde{D}_2) > 0$ must hold.

To prove the last assertion we again assume $i = 1$ and $z_0 \geq 0$, and that the result is not true. That is, denoting $\tilde{D}_1 \triangleq \{|Z_{n_1}^n| > |Z_{n_1}^n - z_0|\} \cap A_1$, one has $P(\tilde{D}_1) > 0$. Now, denote

$$\tilde{D}_{i+1} \triangleq \{-1 \leq \frac{\Delta Z_{n_1}^n \Delta n_{i+1}}{\Delta Z_{n_1}^n} \leq \frac{1}{2}\} \cap \tilde{D}_i, \quad i = 1, \ldots, n - 1.$$
We shall prove by induction that that $\hat{D}_i \subset A_i$ and $Z^n_{i+1} \geq Z^n_i \geq \frac{1}{2}Z^n_{i-1}$ on $\hat{D}_i$, for $i = 1, \ldots, n$. Indeed, for $i = 1$, by definition $\hat{D}_1 \subset A_1$. Moreover, Lemma 7.1 tells us that $Z^n_1 < z_0$ on $\hat{D}_1$. If $Z^n_1 \leq 0$, then obviously $|Z^n_1| \leq |Z^n_1 - z_0|$. If $Z^n_1 > 0$ in $\hat{D}_1$, then $Z^n_1 > z_0 - Z^n_1$ and hence $Z^n_1 > \frac{1}{2}z_0$ on $\hat{D}_1$. Namely the claim holds for $i = 1$.

Assume now that for all $i \leq j$, the claim holds. In particular, this implies that $Z^n_{\tau_j} > \frac{1}{\tau_j} z_0 \geq 0$ on $\hat{D}_j$, we show that the claim is true for $i = j + 1$. Note that on $\hat{D}_{j+1}$, one has $|\Delta Z^n_{\tau_j+1}| \leq |\Delta Z^n_{\tau_j}| < \varepsilon_0$. Since $Z^n_{\tau_j} \neq 0$ on $\hat{D}_{j+1} \subset \hat{D}_j$, by Definition 2.3 (iii) we know $Z^n_{\tau_j+1} \neq Z^n_{\tau_j}$. Thus $\hat{D}_{j+1} \subset A_{j+1}$. Moreover, since $\delta Z^n_{\tau_j} < 0$, we have $\Delta Z^n_{\tau_j+1} \geq \frac{1}{\tau_j} \Delta Z^n_{\tau_j}$ on $\hat{D}_{j+1}$. Thus by inducational hypothesis we have

$$Z^n_{\tau_{j+1}} \geq \frac{3}{2} Z^n_{\tau_j} - \frac{1}{2} Z^n_{\tau_{j-1}} \geq \frac{1}{2} Z^n_{\tau_j}, \quad \text{on } \hat{D}_{j+1}.$$ 

That is, the claim is true for $i = j + 1$, and hence it is true for all $i$.

Finally, by applying the same argument repeatedly we have $P(\hat{D}_n) > 0$. But the claim tells us that $Z^n_{\tau_n} > \frac{1}{\tau_n} z_0 \geq 0$ on $\hat{D}_n$. This is impossible since $Z^n_{\tau_n} = 0$ must hold almost surely by definition of $Z^n_0(z_0)$. The proof is now complete.

[Proof of Proposition 5.4] (i) We follow the arguments in Lemma 7.2. Assume $i = 1$, $z_0 \geq 0$, and that the result is not true. Then $P(D_1) > 0$, where

$$D_1 \triangleq \{ P\{B_1|F_{\tau_1}\} > \frac{C_0}{C_1} \} \bigcap A_1.$$ 

As before, we define $\hat{Z}^n_{\tau_0} \triangleq z_0$; $\hat{Z}^n_{\tau_1} \triangleq Z^n_{\tau_1} 1_{D_1} + z_0 1_{D_1}$, and $\hat{Z}^n_{\tau_i} \triangleq Z^n_{\tau_i}$, for $i \geq 2$. Then $\hat{Z}^n \in Z^n_{\hat{B}}(z_0)$, and

$$\delta Y^n_T = \left[ (z_0 - Z^n_{\tau_1})|X_{\tau_2} - X_{\tau_1}| + c(Z^n_{\tau_1} - z_0) + c(Z^n_{\tau_2} - Z^n_{\tau_1}) - c(Z^n_{\tau_2} - z_0) \right] 1_{D_1}.$$ 

On $D_1 \setminus B_2$, we use (H4)-(v) to get $c(Z^n_{\tau_1} - z_0) + c(Z^n_{\tau_2} - Z^n_{\tau_1}) - c(Z^n_{\tau_2} - z_0) \geq 0$. On $D_1 \cap B_2$, we have $|Z^n_{\tau_2} - Z^n_{\tau_1}| \geq \varepsilon_0 \geq |Z^n_{\tau_1} - z_0|$. Thus (H4)-(iii) tells us that

$$c(Z^n_{\tau_1} - z_0) + c(Z^n_{\tau_2} - Z^n_{\tau_1}) - c(Z^n_{\tau_2} - z_0) \geq C_1 |Z^n_{\tau_1} - z_0|.$$ 

Combining above we conclude that

$$E\{U(Y^n_T) - U(Y^n_T)\}$$

$$\geq E\left\{ C_1 |Z^n_{\tau_1} - z_0| 1_{D_1} \bigcap B_2 - \Lambda |Z^n_{\tau_1} - z_0| 1_{E_1} \{ |X_{\tau_2} - X_{\tau_1}| \} 1_{D_1} \right\} + E\left\{ C_1 |Z^n_{\tau_1} - z_0| 1_{B_2} (1 - \Lambda) |Z^n_{\tau_1} - z_0| 1_{D_1} \right\}$$

$$\geq \Lambda E\left\{ |C_1 |Z^n_{\tau_1} - z_0| 1_{D_1} - (C_0 - 1) |Z^n_{\tau_1} - z_0| \right\} 1_{D_1} \right\} = \Lambda E\left\{ |Z^n_{\tau_1} - z_0| 1_{D_1} \right\} > 0.$$ 

32
This proves the part (i).

We shall prove part (ii) by backward induction on $i$. Since $Z_{r_n}^n = 0$, the result is true for $i = n$. Without loss of generality we assume it is true for $i = 2$ and will prove it for $i = 1$. Assume $z_0 \geq 0$. If it is not true for $i = 1$, then $P(\tilde{D}_1) > 0$ where

$$\tilde{D}_1 \triangleq \{Z_{r_1}^n \neq 0\} \cap A_1.$$ 

We now define $\tilde{Z}_{\tau_0}^n \triangleq z_0$; $\tilde{Z}_{\tau_1}^n \triangleq Z_{r_1}^n \mathbf{1}_{\tilde{D}_1}$; and $\tilde{Z}_{\tau_i}^n \triangleq Z_{r_i}^n$, for $i \geq 2$. Then $\tilde{Z}^n \in Z_{t_0}^n(z_0)$, and

$$\delta Y_T^n = \left[-Z_{\tau_1}^n [X_{\tau_2} - X_{\tau_1}] + c(Z_{\tau_2}^n - z_0) + c(Z_{\tau_2}^n - Z_{\tau_1}^n) - c(-z_0) - c(Z_{\tau_2}^n)\right] \mathbf{1}_{\tilde{D}_1}.$$ 

On $\tilde{D}_1 \cap B_2^0$, by inductive hypothesis we have $Z_{\tau_2}^n = 0$. Then by (H3)-(v) we have

$$c(Z_{\tau_2}^n - z_0) + c(Z_{\tau_2}^n - Z_{\tau_1}^n) - c(-z_0) - c(Z_{\tau_2}^n) = c(Z_{\tau_1}^n - z_0) + c(-Z_{\tau_1}^n) - c(-z_0) \geq 0.$$ 

On $\tilde{D}_1 \cap B_2$, we have $|Z_{\tau_2}^n - Z_{\tau_1}^n| \geq |Z_{\tau_2}^n - z_0| \geq |Z_{\tau_1}^n|$, thanks to Lemma 7.2. By (H4)-(iv) we have

$$c(-Z_{\tau_1}^n) + c(Z_{\tau_1}^n - z_0) - c(-z_0) \geq \alpha_0 [c(Z_{\tau_2}^n) - c(Z_{\tau_2}^n - Z_{\tau_1}^n)];$$

and thus

$$c(Z_{\tau_1}^n - z_0) - c(-z_0) + c(Z_{\tau_2}^n - Z_{\tau_1}^n) - c(Z_{\tau_2}^n) \geq -\frac{2}{\alpha_0} [c(-Z_{\tau_1}^n) + c(Z_{\tau_1}^n - z_0) - c(-z_0)].$$

It follows that

$$E\{U(Y_{\tau_1}^n) - U(Y_{\tau_2}^n)\} = E\{U'()|Y_{\tau_1}^n - Y_{\tau_2}^n\}$$

$$\geq E\left\{U'()\left[ -Z_{\tau_1}^n [X_{\tau_2} - X_{\tau_1}] \mathbf{1}_{\tilde{D}_1} + [c(-Z_{\tau_1}^n) + c(Z_{\tau_1}^n - z_0) - c(-z_0)] [\mathbf{1}_{\tilde{D}_1 \cap B_2^0} - \frac{2}{\alpha_0} \mathbf{1}_{\tilde{D}_1 \cap B_2}] \right]\right\}$$

$$\geq E\left\{ - \Lambda|Z_{\tau_1}^n| |X_{\tau_2} - X_{\tau_1}| \mathbf{1}_{\tilde{D}_1} + [c(-Z_{\tau_1}^n) + c(Z_{\tau_1}^n - z_0) - c(-z_0)] [\Lambda \mathbf{1}_{\tilde{D}_1 \cap B_2^0} - \frac{2\Lambda}{\alpha_0} E_{\tau_1} \{\mathbf{1}_{B_2^0}\}] \right\}$$

$$\geq E\left\{ - \Lambda|Z_{\tau_1}^n| E_{\tau_1} \{|X_{\tau_2} - X_{\tau_1}|\} + [c(-Z_{\tau_1}^n) + c(Z_{\tau_1}^n - z_0) - c(-z_0)] [\Lambda E_{\tau_1} \{\mathbf{1}_{B_2^0}\} - \frac{2\Lambda}{\alpha_0} E_{\tau_1} \{\mathbf{1}_{B_2}\}] \right\} \mathbf{1}_{\tilde{D}_1}.$$ 

By part (i), we know that $P$-a.s. on $\tilde{D}_1 \subset A_1$, $P\{B_2 \cap \mathcal{F}_{\tau_1}\} \leq \frac{C_0}{C_1}$ and $P\{B_2^0 \cap \mathcal{F}_{\tau_1}\} \geq 1 - \frac{C_0}{C_1}$. Then, using (H4)-(iii) and recalling that $C_1 > 2C_0[1 + \frac{\Lambda}{\alpha_0}]$,

$$E\{U(Y_{\tau_1}^n) - U(Y_{\tau_2}^n)\}$$
\[
\geq \lambda E \left\{ - (C_0 - 1)|Z^n_\tau| + |c(-Z^n_\tau) + c(Z^n_\tau - z_0) - c(-z_0)|\left[1 - \frac{C_0}{C_1} - \frac{2\Lambda}{\lambda \alpha_0 C_1}\right] 1_{\delta_1} \right\}
\]
\[
\geq \lambda E \left\{ - (C_0 - 1)|Z^n_\tau| + C_1|Z^n_\tau|\left[1 - \frac{C_0}{C_1} - \frac{2\Lambda}{\lambda \alpha_0 C_1}\right] 1_{\delta_1} \right\} = \lambda E \left\{ |Z^n_\tau| 1_{\delta_1} \right\} > 0.
\]

This contradicts the assumption that \( Z^n \) is the optimal strategy. \( \blacksquare \)

### 7.2 Proofs of Lemma 6.3 and Theorem 6.4

**[Proof of Lemma 6.3.]** We follow the proof of Proposition 5.4. For each \( n \), let \( Z^n \in Z^n_T(\tilde{z}) \) be the optimal portfolio of \( V^n(t, x, y - c(\tilde{z} - z), \tilde{z}) \). We shall prove several claims by contradiction. In each case, we show that if the claim is not true, then we can construct some \( \tilde{Z}^n \in Z^n_T(z) \) such that

\[
E \left\{ U(Y^{t,x,y,\tilde{Z}^n}_T) - U(Y^{t,x,y-c(\tilde{z}-z),Z^n}_T) \right\} \geq c(z, \tilde{z}) > 0, \tag{7.6}
\]

where \( c(z, \tilde{z}) \) is some constant independent of \( n \). Since this leads to that

\[
V^n(t, x, y, z) - V^n(t, x, y - c(\tilde{z} - z), \tilde{z}) \geq c(z, \tilde{z}).
\]

Sending \( n \to \infty \) we obtain the contradiction.

Without loss of generality we assume \( z \geq 0 \). The key observation is that we may also view \( Z^n \) as a portfolio starting from \((t, x, y, z)\), with an initial jump from \( z \) to \( \tilde{z} \).

**Claim 1.** \( \tilde{z} < z \). Indeed, if \( \tilde{z} > z \), then we follow the proof of Lemma 7.1. For fixed \( n \), let \( k = \inf \{ i : Z^n_i \leq 0 \} \), and define \( \hat{Z}^n_i \overset{\Delta}{=} [Z^n_i - \tilde{z} + z] \vee 0 \) for \( i < k \), and \( \hat{Z}^n_i \overset{\Delta}{=} Z^n_i \) for \( i \geq k \). Then \( \hat{Z}^n \in Z^n_T(z) \) and, following exactly the same arguments there we get

\[
E \left\{ U(Y^{t,x,y,\hat{Z}^n}_T) - U(Y^{t,x,y-c(\tilde{z}-z),Z^n}_T) \right\} \geq \lambda(\tilde{z} - z) > 0.
\]

Namely (7.6) holds and we have a contradiction.

**Claim 2.** \( P(-1 \leq \frac{Z^n_t - \tilde{z}}{\tilde{z} - z} \leq \frac{1}{2}) > 0 \). Indeed, if \( P(-1 \leq \frac{Z^n_t - \tilde{z}}{\tilde{z} - z} \leq \frac{1}{2}) = 0 \), then we define \( \hat{Z}^n_{t_0} = z, \hat{Z}_i^n = Z^n_{t_i} \), for all \( i \geq 1 \). Then \( \hat{Z}^n \in Z^n_T(z) \), and similar to (7.4) we again have

\[
E \left\{ U(Y^{t,x,y,\hat{Z}^n}_T) - U(Y^{t,x,y-c(\tilde{z}-z),Z^n}_T) \right\} \geq \lambda(z - \tilde{z}) > 0.
\]

**Claim 3.** On \( \{-1 \leq \frac{Z^n_t - \tilde{z}}{\tilde{z} - z} \leq \frac{1}{2}\} \), we have \( |Z^n_{\tau_1} - \tilde{z}| \leq |\tilde{z} - z| < \varepsilon_0 \). Indeed, if \( Z^n_{\tau_1} = \tilde{z} \), by Definition 2.3 (iii) we get \( \tau_1 = T \) and thus \( \tilde{z} = 0 \). If \( Z^n_{\tau_1} \neq \tilde{z} \), recalling that \( Z^n \) is the optimal portfolio of \( V^n(t, x, y - c(\tilde{z} - z), \tilde{z}) \), by Proposition 5.4 we get \( Z^n_{\tau_1} = 0 \). Then \( \{-1 \leq \frac{Z^n_t - \tilde{z}}{\tilde{z} - z} \leq \frac{1}{2}\} \). In both cases we get \( |\tilde{z}| \leq z - \tilde{z} \).

---

34
Claim 4. \( P(|Z^n_0 - \bar{z}| \geq \varepsilon_0) \leq \frac{C_0}{C_1}. \) Indeed, if \( P(|Z^n_0 - \bar{z}| \geq \varepsilon_0) > \frac{C_0}{C_1}, \) then we define \( \tilde{Z}^n_0 = z, \) and \( \tilde{Z}^n_{\tau_i} = Z^*_{\tau_i}, \) for \( i \geq 1. \) Similar to (7.5) once again we have
\[
E\left\{ U(Y^t,x,y,\tilde{Z}^n) - U(Y^t,x,y-c(\bar{z}-z),Z^n) \right\} \geq \lambda(z - \bar{z}) > 0.
\]

Finally, recall Definition 2.1. Consider the strategy starting from \( z \) and jumping to 0 at the initial time \( t, \) and then following \( Z^n, \) we have the terminal wealth
\[
\tilde{Y}^n = y + \int_{\tau_1}^{T} Z^n_s dX_s - c(-z) - c(Z^n_1) - \sum_{i=2}^{n} c(Z^n_{\tau_i} - Z^n_{\tau_{i-1}}).
\]
Note that \( Y^t,x,y-c(\bar{z}-z),Z^n = y - c(\bar{z} - z) + \bar{z}[X_{\tau_1} - x] + \int_{\tau_1}^{T} Z^n_s dX_s - c(Z^n_1 - \bar{z}) - \sum_{i=2}^{n} c(Z^n_{\tau_i} - Z^n_{\tau_{i-1}}), \)
we have
\[
\tilde{Y}^n - Y^t,x,y-c(\bar{z}-z),Z^n = c(\bar{z} - z) + c(Z^n_{\tau_1} - \bar{z}) - c(-z) - c(Z^n_1) - \bar{z}[X_{\tau_1} - x].
\]

Similar to (7.6) we can prove
\[
V(t,x,y) - V^n(t,x,y-c(\bar{z}-z),\bar{z}) \geq E\left\{ U(\tilde{Y}^n) - U(Y^t,x,y-c(\bar{z}-z),Z^n) \right\} \geq \lambda|\bar{z}|.
\]
Send \( n \to \infty \) and noting that \( V(t,x,y,z) = V(t,x,y-c(\bar{z}-z),\bar{z}), \) we must have \( \bar{z} = 0. \)

[Proof of Theorem 6.4.] (i) If \( c(z) \geq c_0 > 0 \) for all \( z \neq 0, \) then following the arguments in Theorem 5.1 one can easily prove that \( E\{N(Z^*)\} \leq \frac{C_0}{C_1}. \)

Now assume (H4) holds. Following the proof of Theorem 6.3 it is readily seen that, for any \( i \) and \( P \)-almost surely on \( \{0 < |Z^*_{\tau_i} - Z^*_{\tau_{i-1}}| < \varepsilon_0\}, \) it holds that
\[
Z^*_{\tau_i} = 0 \quad \text{and} \quad P(|Z^*_{\tau_{i+1}} - Z^*_{\tau_i}| \geq \varepsilon_0|\mathcal{F}_{\tau_i}) \leq \frac{C_0}{C_1}.
\]
Then following the proof of Theorem 5.3 we get
\[
P\left( \sum_{i=0}^{n} 1_{0 < |Z^*_{\tau_i} - Z^*_{\tau_{i-1}}| < \varepsilon_0} \geq m \right) \leq \frac{1}{2^{m-1}}, \quad \forall n \geq m.
\]
Similar to Theorem 5.5 one can then prove that \( E\left\{ \sum_{i=0}^{\infty} 1_{|Z^*_{\tau_i} - Z^*_{\tau_{i-1}}| > \varepsilon_0} \right\} < \infty. \) This implies that \( P(\tau^*_i < T, \forall i) = 0 \) and \( E(N(Z^*)) < \infty. \)

(ii) Applying Lemma 6.2 repeatedly we have
\[
V(t,x,y,z) = E\left\{ V(t_n,x_{\tau_n},Y^*_{\tau_n},Z^*_{\tau_n}) \right\}, \quad \forall n.
\]
Now by (i), we conclude that \( \tau^*_n = T \) and \( Z^*_{\tau_n} = 0 \) for \( n \) large enough. Sending \( n \to \infty \) we obtain that \( V(t,x,y,z) = E\{U(Y^*_T)\}. \) This means that \( Z^* \) is an optimal portfolio for \( V(t,x,y,z). \)
References


