Central European University

Department of Mathematics and its Applications

Asymptotic Arbitrage Strategies for Long-Term Investments in Discrete-Time Financial Markets

Thesis submitted in partial fulfillment of the requirements towards the degree

Doctor of Philosophy (Ph.D)

in Mathematics and its Applications

Research Area: Mathematical Finance

By

Martin Le Doux Mbele Bidima

Research Period: 10/09/2007 - 07/06/2010

Thesis defended before the Exam Committee:

Chair: Prof. Gheorghe Moroşanu, CEU, Budapest, Hungary
Internal Examiner: Prof. Tamás Szabados, Technical University, Budapest
External Examiner: Dr József Gáll, University of Debrecen, Hungary
Supervisor: Dr Miklós Rásonyi, SZTAKI, Budapest - University of Edinburgh, U.K.
Associate Supervisor: Prof. László Gerencsér, SZTAKI, Budapest, Hungary

Grade: ***Ph.D Magna Cum Laude***

CEU, Budapest, July 1, 2010

Dedication

I dedicate this Ph.D Thesis to,

First, in honor and thanksgiving to my Lord God Almighty, His Only Son Jesus-Christ my Lord and Savior, and His Holy Spirit my Comforter and Guide.

Next, in gratitude to the Virgin Mary my Holy Mother, Saint Joseph, Saint Moses, Saint Michael, Saint Barachiel, Saints Martin, Blessed Father César Ditona, Blessed Yvonne Avénerie, Blessed Godefroy Bidima Ndengte, and all Saints.

And, in merit to my wife Danielle Sandrine Mbele, and finally to all my present and future children (in Christ): Bien, Ben, Larissa, Valery, Raïssa, Manuel, Mary, ...

Acknowledgments

I would like to express my gratitude to the Philanthropist George Soros, founder of Central European University, for the three years schorlarship I benefited from his donations, which supported my Ph.D researches I present in this thesis.

Next, I am deeply indebted and grateful to my research Supervisor and associate Supervisor Dr Miklós Rásonyi and Prof. László Gerencsér respectively. Particularly to Miklós for his full availability and sacrifice during 33 months of intense, rigorous, and exceptional Ph.D research supervision. Moreover, I sincerely thank them for their patience, attention and kindness to me during that research period.

Also, I am very grateful to Prof. Gheorghe Moroşanu, the Head of CEU Department of Mathematics and its Applications, for his general research advice, undersdanding and encouragements from which I benefited during my Ph.D studies at CEU.

My special thanks go also to the CEU community and all lecturers, staff members such as Mrs Elvira Kadvany, Ph.D graduates and current research students in the Department of Mathematics and its Applications. A particular thank to Dr Tihomir Gyulov for his helpful discussion on the proof of Proposition 1.1.5 in the first chapter of this thesis.

Next, I am thankful to Prof. Chris Rogers, my former supervisor during my 2004/2005 Master' studies in Mathematical Finance at the University of Cambridge, who kept paying a certain attention to me till the completion of the present Ph.D studies.

I thank Prof. Neil Turok, Prof. Fritz Hahne, founder and former director of the African Institute for Mathematical Sciences (AIMS), South Africa, my lecturers I had there namely Prof. Alan Macfarlane, Prof. Martin Bucher, Prof. Kavilan Moodley, Prof. Ekkehard Kopp, Prof. Ronald Becker, Prof. Alan Beardon, Prof. Sanjoy Mahajan, ...

Also, I am very thankful to Prof. Georges Edward Njock and Dr Célestin Nkuimi, my former Supervisors during my first Master' studies in Pure Mathematics at the University of Yaoundé I, Cameroon, to whom I owe my very first ambition to get a Ph.D, and who kept offering their constant moral support to me until present.

Finally, I sincerely thank my own families' members such as Dr Théophile Mbele, ..., and all non cited people whose various support helped me on my way to getting a Ph.D.

Abstract

In the present thesis I consider models of financial markets where the price process of the risky asset follows a Markov chain taking values in a subinterval of $\mathbb{R}$. In particular, we deal with time-discretizations of stochastic differential equations, a model class often ocurring in practice.

Motivated by recent articles, I investigate the possibility of realizing arbitrage as the time horizon of trading, T, tends to infinity.

Under suitable hypotheses we construct explicit trading strategies which provide linear/exponential growth of wealth as $T \to \infty$ with a probability converging to 1. Using the theory of Large Deviations, we refine this result showing that the probability in question tends to 1 geometrically fast, under suitable hypotheses.

Finally, we consider arbitrage in the sense that the expected utility of investors tends to the maximal achievable utility. I investigate how our previously constructed strategies perform in this sense.

A.M.S Class. Code: 91G80 (Stochastic Control in Finance)

Contents

Introduction

Arbitrage (riskless profit) is the cornerstone concept of modern mathematical finance. Several version of the so-called "fundamental theorem of asset pricing" have been proved over the past two decades, see [6] for an overview. This theorem states that absence of arbitrage is equivalent to the existence of "suitable" pricing functionals for derivative securities (such as options).

It is quite clear that short-term arbitrage shouldn't exist: it would immediately be exploited by traders and hence their activity would move the prices, making the aribtrage opportunities disappear. It can be argued, however, that one may generate long-term riskless profit (mathematically speaking, when the time horizon T tends to infinity), this is indeed observed in most models of financial markets.

The existence and nature of such infinite horizon asymptotic arbitrage opportunities have been studied in a number of papers treating various models, we mention for example [1], [9], [16], [14], [21], [22], [48]. One convenient framework for studying asymptotic arbitrage is the theory of "large financial markets", initiated in [20], then developed in [27], [28], [25], [26], [5], [21], [41], [42], [43] and [48].

In the present thesis I am dealing with Markovian models of financial markets. The specific structure of these allows us to define rather strong forms of arbitrage, peculiar to the Markovian setting, and prove their existence under appropriate hypotheses. We introduce two types of arbitrage: almost sure (trajectoriwise) and utility-based. The first one guarantees that our portfolio grows exponentially outside a set of probability converging to 0 as time goes on. Indeed, we outline the thesis presentation as follows.

First, to make the presentation as more self-contained as possible, in Chapter 1, we collect essential tools of Avanced Probability that we need to handle and use in the sequel chapters of the thesis.

In Chapter 2, after discussing an inspiring "toy model", we consider discretizations of stochastic differential equations and we derive in Section 2.2 the existence of asymptotic arbitrage such that investor's wealth tends linearly to infinity with probability tending to 1 at a geometric rate.

Chapter 3 is the core part in this thesis. We were motivated first by the work of [14] and found all the research developments that we present in the whole thesis while trying to settle questions raised there. Hence we quickly sketch the setting of that [14].

The authors of [14] considered a d-dimensional diffusion process

$$dS_t = \Sigma(S_t)(dW_t + \phi(S_t)dt), \tag{1}$$

on a suitable filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, where $W_t, t \geq 0$ is an N-dimensional standard Brownian motion. S_t is thought to represent the price evolution of d risky assets such as stocks, $\Sigma : \mathbb{R}^d \to \mathbb{R}^{d\times N}$ is the volatility matrix that determines the correlations between the assets, $\phi : \mathbb{R}^d \to \mathbb{R}^N$ is the so-called market price of risk function.

The latter has a straightforward interpretation when $d = N = 1$: it is the stock's rate of return per unit volatility. In other words, the drift $\Sigma\phi$ represesents the rate(s) of return on the stock(s).

We recall their,

Definition 0.0.1. (Definition 1.3 of [14])
They said that S_t has a non-trivial market price of risk if there is $c > 0$ such that

$$\lim_{T\to\infty} \mathbb{P}\left(\frac{1}{T}\int_0^T |\phi(S_t)|^2 dt < c\right) = 0. \tag{2}$$

Intuitively, in such a market even if $T \to \infty$ there is still a nonvanishing "drift normalized by the volatility" (that is, $\phi(S_t)$) which, in some sense, means that market opportunities do not "run dry" as time goes on. As pointed out in [14], (2) holds whenever S_t is "ergodic" (it satisfies a suitable law of large numbers) and has an invariant measure φ such that ϕ is non-zero φ-almost surely.

From this, they derived a first result, Theorem 1.4 of [14], which we restate below.

Theorem 0.0.2.
If S_t has a nontrivial market price of risk then there exists $b > 0$ and for each $\epsilon > 0$ there exists T_ϵ such that for all $T > T_\epsilon$

$$\mathbb{P}(X_T \geq e^{bT}) \geq 1 - \epsilon \tag{3}$$

for some $X_T \geq -e^{-bT}$, where X_T is the outcome of an "admissible" trading strategy on $[0, T]$ starting from 0 initial capital.

Since this result serves only as a motivation for our work we do not provide a definition of "admissibility" here but rather underline the essential content of Theorem 0.0.2: it says that for any tolerance level ϵ one may find T large enough such that an exponentially growing profit can be obtained on $[0,T]$ with an exponentially decreasing potential loss and with a probability of failure below ϵ. This can be considered as a rather strong form of long-term arbitrage.

Nevertheless, there are unsettling features of Theorem 0.0.2: the relationship between ϵ and T_ϵ is not clarified (one may need to wait for a very long time to achieve a desired tolerance level) and the trading strategies are not explicitly given (indeed, the proof is non-constructive).

Next, considering the special case $S_t := \exp(X_t)$, for $t \in [0,\infty]$, where X_t is the Ornstein-Uhlenbeck process

$$dX_t = -\rho X_t dt + \sigma dW_t,\ X_0 = x \in \mathbb{R}, \tag{4}$$

for some constants $\rho > 0$, $\sigma > 0$, the authors of [14] formulated another condition on the market price of risk that is stronger than (2). We recall it as below,

Definition 0.0.3. (Definition 1.3 in [14])

The market price of risk satisfies a large deviation estimate if there are $c_1, c_2 > 0$ *such that*

$$\limsup_{T\to\infty} \frac{1}{T} \log \mathbb{P}\left(\frac{1}{T}\int_0^T |\phi(S_t)|^2 dt \le c_1\right) \le -c_2. \tag{5}$$

It is an appealing conjecture that when (5) holds, a strengthening of Theorem 0.0.2 should hold true: the existence of constants $b_1, b_2, C > 0$ such that

$$\mathbb{P}(X_T < e^{b_1 T}) \le C e^{-b_2 T} \tag{6}$$

for all time T large enough and for suitable outcomes $X_T \ge -e^{-b_1 T}$ of admissible trading strategies. Noting that (6) is equivalent to

$$\mathbb{P}(X_T \ge e^{b_1 T}) \ge 1 - e^{-b_3 T}, \tag{7}$$

for some constant $c_3 > 0$. Such a result would establish an explicit relationship between a preset tolerance level ϵ and the time necessary to reach that tolerance level. In particular, the probability of having a loss (that is, $\mathbb{P}(X_T < 0)$) could be controlled.

At the end of section 3 in [14] the authors sketch how large deviations theory could be applied to show (6) from (5), but they do not carry out this programme: it seems that some of the necessary theoretical tools are still lacking. They only study the concrete case of the Ornstein-Uhlenbeck process, where explicit b_1, b_2 are given, together with an array of convergence rates for related optimization problems.

Hence in Chapter 3, in a discrete-time version of the model (1), we will derive results implying a discrete-time version of Theorem 0.0.2 (see Corollary 3.2.7, Theorem 3.3.6 and

Theorem 3.3.8). Next, still in our discrete-time setting, we will show that (5) implies (6) or (7) (see Theorem 3.2.6 and Theorem 3.3.11). And we provide in Theorem 3.3.12 easily verifiable conditions that guarantee a discrete-time version of (5). An important novelty is that the strategies we use will be explicitly constructed.

Finally in Chapter 4, we introduce and discuss "utility-based" asymptotic arbitrage that uses the concept of von Neumann-Morgenstern utilities (see Chapter 2 of [13]): a concave increasing function U is considered such that $U(x)$ is thought to represent the subjective value of x dollars for a given investor. The monotonicity property is natural: investors prefer more to less. Concavity is related to the risk-averse behaviour which is typical for agents in the market. An optimal investment for an agent with utility U is the available portfolio with (random) payoff X for which the expected utility $\mathbb{E}U(X)$ is maximal.

In the present thesis we do not focus on the construction of optimal strategies but rather on ones that provide (rapidly) increasing expected utilities for the agents, i.e. his/her satisfaction will tend to the highest available utility (finite or infinite, depending on whether U is bounded or not from above) as the time horizon T tends to infinity.

We will see that, surprisingly, almost sure asymptotic arbitrage strategies may perform poorly in the expected utility sense, showing that different performance criteria require different strategies, there is no "robustness" in general.

Chapter 1

Review of Advanced Probability

In this preliminary chapter, I briefly review techniques and results of Large Deviations Theory, related articles, and those of Markov chains Theory which will be used in the main chapters of my present thesis. As a review, I may not reprove all these results, and I assume well known tools from Real Analysis, Basic Probability and Stochastic Calculus.

1.1 Large Deviations Theory

Large Deviations is a theory of rare events, based on the analysis of tails of probability distributions. Its applications lie in various subfields of Probability. In this thesis, we present Control Theory as one of its unusual applications, see however for instance [40].

Before recalling concepts and results of this theory, we begin with a,

1.1.1 Preliminary Convex Analysis

Let d be a positive integer, then,

Definition 1.1.1.

A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ *is convex if,* $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$, *for all* $x, y \in \mathbb{R}^d$ *and all* $\alpha \in (0,1)$.

A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ *is concave if,* $-f$ *is convex in the sense above.*

If f *is finitely valued and the inequality above is strict for all* $x \neq y$ *and all* $\alpha \in (0,1)$, *then we say that* f *is strictly convex and strictly concave respectively.*

The result below shows the use of convex and concave functions in Optimal Control.

Proposition 1.1.2.

If a function $f : \mathbb{R}^d \to (-\infty, +\infty]$ *is convex (or concave), then a local minimum is global. And if* f *is strictly convex (or strictly concave), any such minimum is unique.*

Proof. Let f be convex and x a local minimum, then $f(x) \leq f(z)$ for all z in a neighborhood V of x. For all $y \in \mathbb{R}^d$, we have $\alpha x + (1-\alpha)y \in V$ for any $\alpha < 1$ close to 1. Letting $\alpha \to 1$, we get by convexity that, $f(x) \leq f(y)$. Hence x is a global minimum. Moreover, if f is strictly convex, the unicity is straightforward, as required, ∎

Definition 1.1.3.

Let $f : \mathbb{R}^d \to [-\infty, +\infty)$ be any function which is not necessarily convex. The conjugate (or dual) of f is the function $f^* : \mathbb{R}^d \to [-\infty, \infty]$ defined for all $\theta \in \mathbb{R}^d$ by,

$$f^*(\theta) := \sup_{x \in \mathbb{R}^d} \{\theta \cdot x - f(x)\}, \text{ where } \theta \cdot x \text{ denotes the inner product of } \theta \text{ and } x \text{ in } \mathbb{R}^d.$$

Remark 1.1.4.

As the supremum of affine functions, f^* is always convex even if f is not. f^* is therefore called the convex conjugate of f. Moreover, for $d = 1$, we have,

Proposition 1.1.5.

If a function $f : \mathbb{R} \to \mathbb{R}$ is differentiable, then f^* is strictly convex on its effective domain $\mathcal{D}_{f^*} := \{\theta \in \mathbb{R} : f^*(\theta) < \infty\}$.

Proof. Suppose by contradiction that, there are $\theta_1 < \theta_2 < \theta_3$, and $\alpha \in (0,1)$ such that, $\theta_2 = \alpha\theta_1 + (1-\alpha)\theta_3$ and $f^*(\theta_2) = \alpha f^*(\theta_1) + (1-\alpha)f^*(\theta_3)$. Consider the linear functions $g_i(x) := \theta_i x - f^*(\theta_i)$, $i = 1, 2, 3$, for $x \in \mathbb{R}$. Then we have $f \geq \max\{g_1, g_2, g_3\}$. Next, set $x^* := \frac{f^*(\theta_3) - f^*(\theta_1)}{\theta_3 - \theta_1}$, then $g_1(x^*) = g_3(x^*)$. Since $g_2(x^*) = \alpha g_1(x^*) + (1-\alpha)g_3(x^*)$, then $g_1(x^*) = g_2(x^*) = g_3(x^*)$. On the other hand, there is a sequence $x_n \in \mathbb{R}$ such that $\theta_2 x_n - f(x_n) \to f^*(\theta_2)$. Since $\theta_2 x - f(x) \leq \theta_2 x - g_3(x) \to -\infty$ as $x \to \infty$, and $\theta_2 x - f(x) \leq \theta_2 x - g_1(x) \to -\infty$ as $x \to -\infty$, then x_n is bounded. Let a be an accumulation point of x_n, we may suppose $a = \lim x_n$. Then $\theta_2 a - f(a) = f^*(\theta_2)$ by continuity of f. This implies $f(a) = g_2(a)$. Since $\theta_1 < \theta_2 < \theta_3$ and $g_1(x^*) = g_2(x^*) = g_3(x^*)$, then we get $g_1(x) > g_2(x) > g_3(x)$ for $x < x^*$ and, $g_1(x) < g_2(x) < g_3(x)$ for $x > x^*$. Hence, $\max\{g_1(x), g_3(x)\} > g_2(x)$, for $x \neq x^*$. So $f(x) > g_2(x)$ for all $x \neq x^*$. As we got $f(a) = g_2(a)$, this implies $a = x^*$, and $f(x^*) = g_1(x^*) = g_2(x^*) = g_3(x^*)$. Since $f(x) \geq \max\{g_1(x), g_3(x)\} = g_1(x)$ or $g_3(x)$ as whether $x \leq x^*$ or $x \geq x^*$, then we get $f'(x^*-) \leq g_1'(x^*) = \theta_1 < \theta_3 = g_3'(x^*) \leq f'(x^*+)$, contradicting the differentiability of f, ∎

Finally for this preliminary subsection, we recall without proof, that,

Proposition 1.1.6.

If a function $f : \mathbb{R}^d \to [-\infty, +\infty)$ is differentiable and convex, then,

$$f^*(f'(x)) = f'(x) \cdot x - f(x), \text{ for all } x \in \mathbb{R}^d. \tag{1.1}$$

Proof. Cf. Lemma 2.4 in [15] for details, ∎

1.1.2 An Introduction to Large Deviations

Let $(X_t)_{t\in\mathbb{N}}$ be a sequence of independent and identically distributed ($i.i.d$) random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with values in $\mathbb{R}$, and with finite expectation μ and variance σ^2.

Throughout the whole thesis, X_t which denotes a single random variable, may also denote, when there is no ambiguity, the whole sequence $(X_t)_{t\in\mathbb{N}}$.

For $t \geq 1$, set $S_t := X_1 + \cdots + X_t$. The Strong Law of Large numbers says that,

$$\frac{S_t}{t} \to \mu, \text{ as } t \to \infty \text{ almost surely}^1 \ (a.s.).$$

The Central Limit Theorem gives fluctuations of $\frac{S_t}{t}$ from μ of size $O(1/\sqrt{t})$ by stating that,

$$\lim_{t\to\infty} \mathbb{P}\Big(\frac{\sqrt{t}}{\sigma}\big(\frac{S_t}{t} - \mu\big) \leq c\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c} e^{-\frac{1}{2}x^2} dx,$$

for any real constant c.

The theory of Large Deviations then deals with larger size fluctuations, that is, fluctuations far from the mean μ. Just like extensions of Strong Law of Large Numbers and Central Limit Theorem, there are versions of Large Deviations that apply to various sequences of random variables, not necessarily $i.i.d.$

Before formally discussing the foundation of this theory, we have

Definition 1.1.7.

Let $f : \mathbb{R}^d \to [-\infty, \infty]$ be a function defined at $x_0 \in \mathbb{R}^d$.

i) We say that f is lower semicontinuous at x_0 if, for all $\alpha \in \mathbb{R}$ satisfying $\alpha < f(x_0)$, there exists a neighborhood V of x_0 such that $\alpha < f(x)$ for all $x \in V$.

ii) We say that f is lower semicontinuous if, f is lower semincontinuous at every point x_0 in $\mathbb{R}^d$ as above.

iii) f is upper semicontinuous on $\mathbb{R}^d$ if, $-f$ is lower semincontinuous on $\mathbb{R}^d$.

From this, we recall that,

Proposition 1.1.8.

A function $f : \mathbb{R}^d \to [-\infty, \infty]$ is lower semicontinuous if and only if the level sets $\{x : f(x) \leq \alpha\}$ are all closed, for $\alpha \in \mathbb{R}$.

Proof. Straightforward, ■

Then, we review the foundation, that is, the basic principle and results of this theory in the next subsection.

[1] $a.s.$ means that $\mathbb{P}(\lim_{t\to\infty}(\frac{S_t}{t}) = \mu) = 1$.

1.1.3 LDP in $\mathbb{R}^d$ and the Gärtner-Ellis Theorem

First,

Definition 1.1.9.
A function $I : \mathbb{R}^d \to [-\infty, \infty]$ is a rate function if,
i) $I(x) \geq 0$ for all $x \in \mathbb{R}^d$,
ii) and I is lower semicontinuous.
We say that I is a good rate function if in addition, the level sets are all compact.

Then, we have,

Definition 1.1.10. *Large Deviations Principle (LDP).*
Let $(X_t)_{t\in\mathbb{N}}$ be a sequence of random variables taking values in $\mathbb{R}^d$. We say that the sequence X_t satisfies an LDP (or a large deviations principle) in $\mathbb{R}^d$ with rate function I if, I is a rate function, and if for every measurable set $B \in \mathcal{B}(\mathbb{R}^d)$, we have

$$-\inf_{x\in B^\circ} I(x) \leq \liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}(X_t \in B) \leq \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}(X_t \in B) \leq -\inf_{x\in \bar{B}} I(x), \quad (1.2)$$

where $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel σ-algebra on $\mathbb{R}^d$, B° is the interior of B, and $\bar{B}$ its closure.

We interprete the one-sided *LDP* $\limsup_t \frac{1}{t} \log \mathbb{P}(X_t \in B) \leq -\inf_{x\in\bar{B}} I(x)$ by saying that the probability that a process X_t lies in any Borel set B decays exponentially in time at rate $\inf_{x\in\bar{B}} I(x)$.

In the case X_t is sum of *i.i.d* random variables, we have the Cramér's Theorem below,

Theorem 1.1.11.
Let $(X_t)_{t\geq 1}$ be an i.i.d sequence of random variables in $\mathbb{R}^d$, set $S_t := X_1 + \cdots X_t$. Let $\Lambda(\theta) := \log \mathbb{E}e^{\theta\cdot X_1}$, for $\theta \in \mathbb{R}^d$, be the common log moment generating function of the X_t's, where $\mathbb{E}$ denotes the expectation with respect to the probability measure $\mathbb{P}$.

If Λ is finite in a neighborhood of zero, then the average sequence S_t/t satisfies an LDP in $\mathbb{R}^d$ with good rate function Λ^*, the convex conjugate of Λ.

Proof. For $d = 1$ see Theorem 2.8 in [15] or Theorem 2.2.3 in [8], and for $d > 1$ see Theorem 2.2.30 in [8] for details, ■

In this result, as S_t is sum of *i.i.d* random variables, one observes that for all $\theta \in \mathbb{R}^d$, we have $\Lambda(\theta) = \frac{1}{t} \log \mathbb{E}e^{\theta\cdot S_t}$. But this is not true for any sequence of random variables.

Hence, for a standard generalization of Cramér's Theorem, consider any sequence S_t of random variables in $\mathbb{R}^d$, and consider the limit below, when it exists,

$$\Lambda(\theta) := \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}e^{\theta\cdot S_t}, \text{ for } \theta \in \mathbb{R}^d. \quad (1.3)$$

Recalling first the,

Definition 1.1.12.

A function $f : \mathbb{R}^d \to [-\infty, \infty]$ *is essentially smooth if,*

i) the interior of its effective domain $\mathcal{D}_f$ *is non-empty,*

ii) f *is differentiable in the interior of its such effective domain* $\mathcal{D}_f$*,*

iii) and f *is steep; that is,* $\lim_{t\to\infty} |f'(\theta_t)| = \infty$ *for any sequence* θ_t *converging to a boundary point of the effective domain* $\mathcal{D}_f$*.*

For example, for $d := 1$, if $\mathcal{D}_f = \mathbb{R}$, and f is analytic and not linear, then f is essentially smooth.

Then, we state the following more general result, known as Gärtner-Ellis'

Theorem 1.1.13.

Let S_t *be any sequence of random variables in* $\mathbb{R}^d$*. Suppose for each* $\theta \in \mathbb{R}^d$*, that the limit* $\Lambda(\theta)$ *in* (1.3) *exists as an extended real number.*

If Λ *is finite in a neighborhood of* $\theta = 0$*, and is essentially smooth and lower semi-continuous, then the sequence of random variables* S_t/t *satisfies an LDP in* $\mathbb{R}^d$ *with good convex rate function* Λ^**.*

Proof. Cf. Theorem 2.3.6 in [8], ■

Finally for this first section, we also recall the,

1.1.4 Law of Large Numbers for Martingale Differences

In this subsection, we review some useful results by stating a law of large number for typical discrete-time stochastic processes.

Let $(M_t)_{t\geq 0}$ be a discrete-time stochastic process on a probability space $(\Omega, \mathbb{P}, \mathcal{F})$, equipped with a filtration $\mathbb{F} := (\mathcal{F}_t)_{t\geq 0}$. Assume that the process M_t is adapted to this filtration; that is, M_t is $\mathcal{F}_t$-measurable for all time t. Then,

Definition 1.1.14.

We say that the process M_t *is a* $\mathbb{P}$*-martingale with respect to* $\mathbb{F}$ *if,*

i) M_t *is* $\mathbb{P}$*-integrable; that is,* $\mathbb{E}(|M_t|) < \infty$ *for all time* t*,*

ii) and $\mathbb{E}(M_{t+1}|\mathcal{F}_t) = M_t$ *for all time* $t \geq 0$*.*

And we say that M_t *is a martingale difference sequence if it is integrable and if we have* $\mathbb{E}(M_{t+1}|\mathcal{F}_t) = 0$*, for all time* t*.*

For examples, if N_t is a martingale with respect to $\mathcal{F}_t$, then $M_t := N_t - N_{t-1}$ is a martingale difference sequence. And if N_t is an integrable adapted process to a filtration $\mathcal{F}_t$, then $M_t := N_t - \mathbb{E}(N_t|\mathcal{F}_{t-1})$ is also a martingale difference sequence.

We state the almost sure convergence result for martingales below,

Proposition 1.1.15.

Let M_t be a martingale with respect to a filtration $\mathcal{F}_t$, and $Y_t := M_t - M_{t-1}$ the corresponding martingale difference sequence.

If $\sum_{t=1}^{\infty} \mathbb{E}(Y_t^2|\mathcal{F}_{t-1}) < \infty$ a.s., then M_t converges to a finite limit M_∞ almost surely.

Proof. It is an immediate corollary of Theorem 2.15 in [18], ∎

Next, before proving the version of law of large numbers as announced, we recall the following Kronecker's

Lemma 1.1.16.

Let x_n be a real sequence such that the series $\sum_{i=1}^{\infty} x_i$ converges. For any sequence of real numbers $a_n \to \infty$, we have

$$\frac{1}{a_n}\sum_{i=1}^{n} a_i x_i \to 0, \text{ as } n \to \infty.$$

Proof. See p. 31, in [18], ∎

Hence, we get,

Theorem 1.1.17.

Consider the martingale difference sequence $M_t := N_t - \mathbb{E}(N_t|\mathcal{F}_{t-1})$, where N_t is an integrable adapted process to some filtration $\mathcal{F}_t$.

If there is a constant $K < \infty$ such that $\mathbb{E}M_t^2 \leq K$ for all time $t \geq 1$, then,

$$\frac{1}{t}\sum_{i=1}^{t} M_i \to 0, \text{ as } t \to \infty \text{ almost surely.} \tag{1.4}$$

Proof. For $t \geq 1$, set $Y_t := M_t/t$, and consider $X_t := \sum_{i=1}^{t} Y_i$. Clearly X_t is a martingale with respect to the filtration $\mathcal{F}_t$. By Monotone Convergence Theorem and the tower property of conditional expectation, it follows that

$$\mathbb{E}\big(\sum_{t=1}^{\infty}\mathbb{E}(Y_t^2|\mathcal{F}_{t-1})\big) = \sum_{t=1}^{\infty}\mathbb{E}\big(\mathbb{E}(Y_t^2|\mathcal{F}_{t-1})\big) = \sum_{t=1}^{\infty}\mathbb{E}Y_t^2 \leq K\sum_{t=1}^{\infty}\frac{1}{t^2} < \infty.$$

So, the random variable $\sum_{t=1}^{\infty}\mathbb{E}(Y_t^2|\mathcal{F}_{t-1})$ has finite first moment, hence it is finite almost surely. By Proposition 1.1.15 above, it follows that the martingale X_t converges almost surely. Now, applying Kronecker's Lemma 1.1.16 above, with the choice $a_n := n$ and $x_n := M_n$ for all integer $n \geq 1$, we get that,

$$\frac{1}{t}\sum_{i=1}^{t} iY_i = \frac{1}{t}\sum_{i=1}^{t} M_i \to 0, \text{ as } t \to \infty \text{ a.s.},$$

showing the theorem, ∎

1.2 Theory of Markov Chains

In this second section, I do not intend to give an extensive account of Markov processes, which are widely studied in the literature and are classified according to whether the state space is countable or uncountable, and whether the time index set is countable or uncountable. For the purpose of my research works presented in this thesis, I confine myself in reviewing useful techniques and results of discrete-time Markov processes in uncountable state spaces $S \subseteq \mathbb{R}$ as presented in textbooks such as [36].

1.2.1 Preliminary Definitions and Concepts

Consider a discrete-time stochastic process $(X_t)_{t\in\mathbb{N}}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and valued in an interval $S \subseteq \mathbb{R}$. Let $\mathcal{B}(S)$ denote the Borel σ-algebra on S, and η a probability distribution of the random variable X_0.

Definition 1.2.1.

i) We say that X_t is a Markov process (or Markov chain) in the state space S, with initial distribution η if, for all time $t \geq 1$, for all Borel set $A \in \mathcal{B}(S)$ and all states $x_0, x_1, ..., x_t := x \in S$, the Markov property below is satisfied,

$$\mathbb{P}\big(X_{t+1} \in A | X_t = x, ..., X_1 = x_1, X_0 = x_0\big) = \mathbb{P}\big(X_{t+1} \in A | X_t = x\big). \tag{1.5}$$

ii) A Markov chain X_t is stationary or time-homogeneous if for all time t, for $x \in S$ and all $A \in \mathcal{B}(S)$, we have $\mathbb{P}\big(X_{t+1} \in A | X_t = x\big) = \mathbb{P}\big(X_t \in A | X_{t-1} = x\big)$.

We interprete (1.5) by saying that the probability of a future behavior of the process depends only on its current state and not on its past behavior. In other words, a Markov chain is a *memoryless process* in the sense that it forgets its past when evolving.

An immediate consequence of this definition is,

Proposition 1.2.2.

Let X_t be a Markov chain in S. There exists a regular version of the probabilities $\{P(x, A) : x \in S, A \in \mathcal{B}(S)\}$, such that for all time t,

i) If $x \in S$ is fixed, then the map $P(x, \cdot)$ defined by $P(x, A) := \mathbb{P}\big(X_{t+1} \in A | X_t = x\big)$, for $A \in \mathcal{B}(S)$, is a probability measure on $\mathcal{B}(S)$.

ii) For $A \in \mathcal{B}(S)$ fixed, the map $x \mapsto P(x, A)$ is a measurable function on S.

Proof. This follows from the fact that S is a Borel subset of a Polish space[2] and hence these regular versions exist, see [7], ■

[2] A Polish space is a separable completely metrizable topoligical space; that is, a space homeomorphic to a complete metric space that has a countable dense subset.

In vertue of the Markov property (1.5), the whole dynamic of the process X_t can be determined by choosing an initial distribution η and specifying the probabilities $P(x, A)$ of reaching any Borel set from one state in a single transition.

Hence,

Definition 1.2.3.

Let X_t be a Markov chain.

i) The operator $P := \{P(x, A) : x \in S, A \in \mathcal{B}(S)\}$, also denoted by $P(x, A)$, is called the one-step transition probability kernel of the Markov chain X_t.

ii) More generally, the t-step transition probability kernel of the Markov chain X_t is the operator $P^t \equiv P^t(x, A) := \mathbb{P}(X_t \in A|X_0 = x)$, for $t \geq 1$ and $P^0(x, A) := \delta_x(A)$, where $\delta_x(A)$ is the Dirac measure on $\mathcal{B}(S)$ concentrated at x.

From any state, if one needs to reach any Borel set by choosing an arbitrary number of intermediary transitions, this is possible through the celebrated Chapman-Kolmogorov equations below,

Theorem 1.2.4.

If X_t is a Markov chain in the state space S and with t-step transition kernel P^t, then for all $0 \leq s \leq t$, we have

$$P^t(x, A) := \int_S P^s(x, dy)P^{t-s}(y, A), \text{ for all } x \in S, A \in \mathcal{B}(S). \quad (1.6)$$

Proof. Cf. Theorem 3.4.2 in [36] for details, ■

The result below gives an equivalent definition of Markov chains in the form we will handle them in the next chapters.

Theorem 1.2.5.

A discrete-time stochastic process X_t is a Markov chain with state space S if and only if, starting from an initial state X_0 with some distribution η, the process evolves in time according to a stochastic recursion of the form,

$$X_{t+1} = f(X_t, \varepsilon_{t+1}), \text{ for } t \geq 0, \quad (1.7)$$

where $(\varepsilon_t)_t$ is a "driving" sequence of i.i.d random variables independent from X_0, valued in some measurable space $S' \subseteq \mathbb{R}$, and $f : S \times S' \to S$ is a measurable function.

Proof. The first implication follows from Exercise 1, p. 211 in [2]. For the reverse implication, since the state space S is Polish, see p. 228 in the same book, ■

1.2.2 ψ-Irreducibility and Cyclic Behavior

For countable state space Markov chains, the concept of irreducibility requires that a chain moves from any state $x \in C$ to another state $y \in C$ with positive transition probability $P^t(x, y)$, where C is a subset of the state space S called, an irreducible subclass.

In our present framework of uncountable state space $S \subseteq \mathbb{R}$, given any measure φ on $\mathcal{B}(S)$, we look at whether from any state x, the chain ever reachs, at a future time t, any positive φ-measure Borel set A with positive probability $P^t(x, A)$. More formally,

Let X_t be a Markov chain in the state space S. For $x \in S$ and $A \in \mathcal{B}(S)$, define the *first return time*[3] on A, $\tau_A := \min\{t \geq 1 : X_t \in A\}$, and the *return time* probabilities $L(x, A) := \mathbb{P}_x(\tau_A < \infty) = \mathbb{P}_x(X_t \text{ ever enters } A)$, where $\mathbb{P}_x(\cdot) := \mathbb{P}(\,\cdot\,|X_0 = x)$. Then,

Definition 1.2.6.

We say that the Markov chain X_t is φ-irreducible if there exists a measure φ on $\mathcal{B}(S)$ such that, if $\varphi(A) > 0$, then $L(x, A) > 0$ for all $A \in \mathcal{B}(S)$ and all $x \in S$.

Proposition 1.2.7.

Let X_t be a Markov chain in the state space S. The following conditions are equivalent:

i) X_t is φ-irreducible,

ii) For all $x \in S$ and all $A \in \mathcal{B}(S)$, if $\varphi(A) > 0$, then there exists some time $t > 0$, possibly depending on x and A, such that $P^t(x, A) > 0$.

Proof. See Proposition 4.2.1 in [36], ∎

Proposition 1.2.8.

If a Markov chain X_t is φ-irreducible for some measure φ, then there exists a probability measure ψ on $\mathcal{B}(S)$ such that,

i) X_t is ψ-irreducible,

ii) ψ is maximal in the sense that, for any other measure φ', the chain is φ'-irreducible if and only if ψ dominates φ'.

Proof. Also, see Propisition 4.2.2 in [36], ∎

Hence, when we say that X_t is ψ-irreducible, we mean that it is ϕ-irreducible for some measure ϕ, hence it is ψ-irreducible for the maximal measure ψ.

Next, we review the concepts of smallness and petitness as below. For any Markov chain X_t with probability kernel P, let $a := \{a(t)\}$ be a probability measure on $\mathbb{N}$. Define the *sampled chain* $X_{a,t}$ whose probability kernel is $P_a(x, A) := \sum_{t=0}^{\infty} P^t(x, A)a(t)$, for all $x \in S$ and all $A \in \mathcal{B}(S)$. Then, we have the next,

[3] τ_A is actually a *stopping time* with respect to the natural filtration $\mathcal{F}_t := \sigma(X_s, s \leq t)$ of X_t; that is, it is a random variable $\tau_A : \Omega \to \mathbb{N}$ satisfying $\{\tau_A \leq t\} \in \mathcal{F}_t$ for all time t.

Definition 1.2.9.

Let X_t be a Markov chain in the state space S, with transition kernel P.

i) A set $C \in \mathcal{B}(S)$ is called a small set for X_t if there exists a positive time $n > 0$, and a non-trivial measure ν_n on $\mathcal{B}(S)$, such that for all $x \in C$, $A \in \mathcal{B}(S)$, we have $P^n(x, A) \geq \nu_n(A)$. When this holds, we say that C is ν_n-small for the chain X_t.

ii) A set $C \in \mathcal{B}(S)$ is said ν_a-petite for X_t if there exist a probability measure a on $\mathbb{N}$ and a non-trivial measure ν_a on $\mathcal{B}(S)$ such that the sampled chain $X_{a,t}$ satisfies the bound $P_a(x, B) \geq \nu_a(B)$ for all $x \in C$ and all $B \in \mathcal{B}(S)$.

The result below guarantees existence of small sets.

Proposition 1.2.10.

Let X_t be a ψ-irreducible Markov chain. Then there exists a countable collection C_n of small sets in $\mathcal{B}(S)$ such that the state space splits as,

$$S = \bigcup_{n=0}^{\infty} C_n.$$

Proof. See Proposition 5.2.4 in [36], ∎

In practice, the use of small sets can be understood as follows. If C is a small set, and if $\nu_n(C) > 0$, then for all $x \in C$, we have $P^n(x, C) > 0$. This means, if the chain starts in C, then there is a positive probability that the chain will return to C at time $t = n$.

Moreover, given a small set for an irreducible Markov chain, one gets in the result below a decomposition (up to a null set) of the whole state space S into a cycle of subsets reachable from each other in a single transition with probability one. Indeed,

Let C be a fixed ν_M-small set for a ψ-irreducible chain X_t, for some M. Define the set

$$E_C := \{t \geq 1 : \text{ the set C is } \nu_t - \text{small, with } \nu_t = \alpha_t \nu_M \text{ for some } \alpha_t > 0\}, \tag{1.8}$$

and let $\mathcal{B}^+(S) := \{A \in \mathcal{B}(S) : \psi(A) > 0\}$. Then we have,

Theorem 1.2.11.

Let X_t be a ψ-irreducible Markov chain and $C \in \mathcal{B}^+(S)$ a ν_M-small set for X_t.

If $d := gcd(E_C)$ is the greatest common divisor of the set E_C, then there exist disjoint sets $D_1, ..., D_d \in \mathcal{B}(S)$ (an "d-cycle") such that,

i) for all $x \in D_i$, we have $P(x, D_{i+1}) = 1$, for i=0,...,d-1 (mod d),

ii) the set $N := \left(\bigcup_{i=1}^{d} D_i\right)^c$ is ψ-null, that is, $\psi(N) = 0$.

The d-cycle of sets $\{D_i\}$ is maximal in the sense that, for any other collection $\{d', D'_k, k = 1, ..., d'\}$ satisfying i) and ii), we have d' deviding d; whilst if $d = d'$, then by reordering the indices if necessary, $D'_i = D_i$ ψ-a.e.

Proof. See Theorem 5.4.4 in [36] for details, ■

This yields the following

Definition 1.2.12.

Let X_t be a ψ-irreducible Markov chain. Then,

i) the largest time d for which an d-cycle occurs for X_t is called the period of X_t,

ii) if $d = 1$, then we say that the chain X_t is aperiodic.

iii) When there exists a ν_1-small set C with $\nu_1(C) > 0$, then we say that the chain X_t is strongly aperiodic.

Finally in this subsection, we illustrate the connection between the concepts of smallness and petiteness in the

Proposition 1.2.13.

Let X_t be any Markov chain in the state space S. Then,

i) If $C \in \mathcal{B}(S)$ is ν_n-small for some $n \geq 1$, then C is ν_{δ_n}-petite, where δ_n is the Dirac measure on $\mathbb{N}$ concentrated at n. But conversely,

ii) If the chain X_t is ψ-irreducible and aperiodic, then every petite set is small.

Proof. *i)* is straightforward. For *ii)*, see Theorem 5.5.7 in [36] for details, ■

1.2.3 Invariance and Ergodicity of ψ-Irreducible Chains

Given a Markov chain X_t, the t-step transition probability kernel P^t may converge in various senses to a "stable measure" φ, that is, a measure which is preserved under the action of $P(x, A)$. Such a measure φ is said *invariant.* In this last subsection, we review the useful modes of convergence known in [36] as ergodicity, geometrical ergodicity and uniform ergodicity.

Definition 1.2.14.

Let X_t be a Markov chain in the state space S, with transition probability kernel P. And consider any σ-finite measure φ on $\mathcal{B}(S)$.

We say that φ is an invariant (or stationary, or limitting) measure for X_t if,

$$\varphi(A) = \int_S P(x, A)\varphi(dx), \text{ for all } A \in \mathcal{B}(S). \tag{1.9}$$

Next, for any set $A \in \mathcal{B}(S)$, consider the *occupation time* ι_A, that is, the *number of visits by the chain X_t to A after time zero*, which is defined by $\iota_A := \sum_{t=1}^{\infty} \mathbf{1}_{\{X_t \in A\}}$, where $\mathbf{1}_B$ denotes the indicator function on any set B. And from any state $x \in S$, consider the *expected number of such visits* defined by $U(x, A) := \sum_{t=1}^{\infty} P^t(x, A) = \mathbb{E}_x(\iota_A)$.

Definition 1.2.15.

i) A Markov chain X_t is said recurrent if it is ψ-irreducible and $U(x, A) \equiv \infty$ for all $x \in S$ and all $A \in \mathcal{B}^+(S)$.

ii) A positive chain is a ψ-irreducible chain X_t having an invariant probability measure.

Proposition 1.2.16.

i) Every positive chain X_t is recurrent.

ii) If a Markov chain X_t is recurrent, then it admits a unique (up to constant multiples) invariant (probability) measure φ equivalent to ψ.

Proof. Cf. Proposition 10.1.1 and Theorem 10.4.9 in [36], ■

Next, we define the concept of ergodicity as follows. For any signed measure ν on $\mathcal{B}(S)$, define the *total variation norm*

$$\|\nu\| := \sup_{f:|f|\leq 1} |\nu(f)|,$$

where $\nu(f) := \int_S f(x)\nu(dx)$, and f runs over the set of all $\mathbb{R}$-valued measurable functions on S. For any such $f : S \to \mathbb{R}$, define $P^t(x, f) := \int_S f(y)P^t(x, dy)$, $x \in S$, $t \geq 1$. Then,

Definition 1.2.17.

i) A Markov chain X_t is said ergodic if there exists a (probability) measure φ on $\mathcal{B}(S)$ such that

$$\lim_{t\to\infty} \|P^t(x, \cdot) - \varphi\| = 2 \lim_{t\to\infty} \sup_{A\in\mathcal{B}(S)} |P^t(x, A) - \varphi(A)| = 0, \text{ for all } x \in S.$$

ii) A Markov chain X_t is geometrically ergodic if there is a (probability) measure φ on $\mathcal{B}(S)$, and for some constants $r > 1$, $R < \infty$ we have,

$$\|P^t(x, \cdot) - \varphi\| \leq Rr^{-t}, \text{ for all } x \in S \text{ and for all time } t.$$

iii) A chain X_t is uniformly ergodic if there is a (probability) measure φ such that,

$$\sup_{x\in S} \|P^t(x, \cdot) - \varphi\| \to 0, \text{ as } t \to \infty.$$

Remark 1.2.18.

i) Although non-trivial, uniform ergodicity implies geometric ergodicity (see Theorem 16.0.1 in [36]), which clearly implies ergodicity.

ii) If a Markov chain X_t is ergodic, then from *i*) of Definition 1.2.17 above we have $\lim_{t\to\infty} P^t(x, A) = \varphi(A)$ for all $x \in S$ and all $A \in \mathcal{B}(S)$. It follows by Chapman-Kolmogorov Theorem 1.2.4, that the measure φ satisfies the invariance property (1.9).

iii) By uniqueness of limits for sequences of real numbers, *i*), *ii*) or *iii*) of the same definition implies that such an invariant measure is necessarily unique. ■

One may hence understand the stationarity (or limitting) property in Definition 1.2.12 above as follows. If a Markov chain X_t is ergodic with invariant measure φ, the convergence $\lim_t \sup_A |P^t(x,A) - \varphi(A)| = 0$ means that, after the chain has been in operation for a long duration of time, the probability of finding it in any set $A \in \mathcal{B}(S)$ is approximately $\varphi(A)$ no matter the state in which the chain began at time zero.

The result below characterizes geometrical and uniform ergodicities respectively. Moreover it enables in practice, to get ergodicity and hence existence of a unique invariant measure, by checking appropriately condition ii) or iv). Indeed,

Theorem 1.2.19.
Let X_t be any ψ-irreducible Markov chain in the state space S.
The following conditions two are equivalent:
i) X_t is geometrically ergodic.
ii) The chain is aperiodic and satisfies the following drift condition: there are a small set C, a function $V : S \to [1,\infty]$, and constants $\delta > 0$, $b < \infty$ such that

$$PV(x) \leq (1-\delta)V(x) + b\mathbf{1}_C(x), \text{ for all } x \in S, \tag{1.10}$$

where $PV(x) \equiv P(x,V) := \int_S V(y)P(x,dy)$.
The two conditions below are also equivalent:
iii) X_t is uniformly ergodic.
iv) The whole state space S is ν_n-small for some n.

Proof. See Theorem 15.0.1 and Theorem 16.0.2 in [36], and Proposition 1.2.13, ■

As a conclusion of this preliminary chapter, let us notice that the drift condition (1.10) above, known as the geometric condition ($V4$) on p. 376 in [36] or on p. 11 in [30], is weaker than other known (stronger) drift conditions such as ($DV4$) on p. 12, or ($DV3+$)(i) with $V = W$ on p. 4, in [30]. We state this latter here, as it will be important in the sequel:

(DV3+)(i): there are functions $V, W : \mathbb{R} \to [1,\infty)$, a small set C for the chain X_t and constants $\delta > 0$, $b < \infty$ such that the following inequality holds,

$$\log\left(e^{-V} P\, e^{V}\right)(x) \leq -\delta W(x) + b\mathbf{1}_C(x), \text{ for all} x \in \mathbb{R}, \tag{1.11}$$

where $P\, e^V(x) := \int e^{V(y)} P(x,dy)$, similarly to $PV(x)$ defined above.

It is this drift condition, also mentioned as LDP condition imposed by Donsker and Varadhan in [10], which I will be checking in the next chapters, in order to apply LDP theory and ergodic results from [30]. We stress that ($DV3+$)(i) implies (1.10) for ψ-irreducible and aperiodic chains, see Proposition 2.1 of [29].

Chapter 2

Asymptotic Linear Arbitrage in Markovian Financial Markets

In this first main chapter of my present thesis, I discuss the asymptotic behavior of the wealth of an economic agent investing in a stock within two different Markovian settings. I introduce a new concept of arbitrage opportunities producing an asymptotic linear growth for the investor's wealth and prove that (under suitable assumptions) one can produce this kind of arbitrage in both model classes. In the first setting it is assumed that the price process evolves in a compact interval and strong, less-realistic hypotheses are imposed. This serves rather as a motivating "toy model" to explain the basic ideas underlying the results of the whole thesis. The second setting is more realistic: I consider discretizations of stochastic differential equations. The arguments will be based on Large Deviations techniques for Markov processes. For these purposes, let us discuss first the,

2.1 Markovian Modeling and Introduction to ALA

Consider a Markovian financial market consisting of a discrete-time Markov chain $(X_t)_{t\in\mathbb{N}}$ evolving in the state space S where $S \subseteq \mathbb{R}$ is assumed to be an interval. The process X_t represents the (discounted) prices of some risky asset such as stock[1]. "Discounted" means, we assume the existence of a bank account (or risk-free bond[2]) and, for simplicity, we assume that its interest rate is 0; that is, its price is $B_t = 1$, for all time t.

We assume that the Markov chain X_t starts from some constant $X_0 \in S$, and is

[1]An asset is any possession that yields value in an exchange. And a stock is any ownership in a company indicated by shares and that yields value in an exchange.

[2]A bond is a security (i.e., a piece of paper) promising the holder an interest payment in the future. Here we assume that it is riskless and will not default.

an adapted process on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} := (\mathcal{F}_t)_t$ is the natural filtration of X_t: $\mathcal{F}_t := \sigma(X_s, s \leq t)$ models the history of the stock prices up to (and including) each time t.

In the whole chapter, λ, λ_2 denote the Lebesgue measure on $\mathbb{R}, \mathbb{R}^2$ respectively.

A trading strategy in this model is, as usual, a discrete-time stochastic process $(\pi_t)_{t\in\mathbb{N}}$, where π_t denotes the number of units of the stock an economic agent holds at time t. The investment decision for time t is assumed to be taken before the price X_t is revealed, hence we assume π_t to be *predictable*, which means that π_t is $\mathcal{F}_{t-1}$-measurable for all $t \geq 1$.

Due to the Markovian structure of the prices process X_t, it is reasonable and natural to restrict ourselves to the following class of predictable strategies.

Definition 2.1.1.

A Markovian strategy in this stock prices model is any trading strategy π_t of the form $\pi_t := \pi(X_{t-1})$, for all time $t \geq 0$, where $\pi : S \to \mathbb{R}$ is a measurable function.

This means, the amount to invest in the stock at time t depends on the only knowledge of the previous price X_{t-1} of such a stock.

Next, given any such strategy π_t, we model the corresponding wealth V_t^π of an investor to allocate in the stock as a process obeying the following stochastic difference equation,

$$\textbf{Model } I\textbf{:} \quad \begin{cases} V_t^\pi = V_{t-1}^\pi + \pi(X_{t-1})(X_t - X_{t-1}) \text{ for all time } t \geq 1, \\ V_0 = v \in \mathbb{R}_+ \text{ is the investor's initial capital.} \end{cases} \tag{2.1}$$

We notice that $V_t^\pi = v + \sum_{n=1}^t Z_n^\pi$ where $Z_n^\pi := \pi(X_{n-1})(X_n - X_{n-1})$ is the wealth increment at time n. This is the discrete-time version of the stochastic integral modeling the wealth of an investor in the time horizon $[0, t]$; see Definition 1.1 of [14].

Then, my main purpose in this chapter is to discuss the asymptotic behavior of the wealth V_t^π under a new concept of asymptotic arbitrage strategies. Indeed, motivated by similar (but slightly different) concepts in [14], [21], I introduce this concept as follows.

In classical Arbitrage Theory, on a finite time horizon $[0, T]$ where T is fixed, we know that a trading strategy π_t is an arbitrage if $V_0 = 0$ and $V_T^\pi > 0$ *a.s.* with $\mathbb{P}(V_T^\pi > 0) > 0$. This means that an arbitrageur gets a gain with no initial risk (at time $t = 0$). It is a principle generally accepted in the literature that such opportunities should not exist in reasonable models of an economy. The standard argument is that when an arbitrage would occur, all investors rush to exploit it and their activity moves the prices and makes the arbitrage disappear. Indeed, in most models used in practice there is absence of arbitrage in the previous sense.

However, it may still be the case that at the end of each finite trading period (at each time horizon T) the wealth grows linearly (or even exponentially, see Chapter 3) with strictly positive probability; that is: $\mathbb{P}(V_T^\pi \geq cT) > 0$ for some real constant $c > 0$. If we are fortunate, this probability may tend to 1 as $T \to \infty$. When this is the case, we may naturally interpret it by saying that the strategy π_t produces a long-term or *asymptotic* linear arbitrage. It has been observed, see for example [9] and [14], that most models used in practice are arbitrage-free on finite intervals $[0, T]$ but produce riskless profit in the limit as $T \to \infty$.

Knowing $\mathbb{P}(V_T^\pi \geq cT) \to 1$ is not enough for real-life applications as the convergence may be too slow and one has to wait indefinitely long for realizing the desired profit with a desired probability (close to 1). It would thus be important to control the probability of failing to achieve such a linear arbitrage in the long-run by requiring that, it decays exponentially as: $\mathbb{P}(V_T^\pi < cT) \leq e^{-c'T}$ as time T gets large, for another constant $c' > 0$.

Hence, we formally define this new concept as below,

Definition 2.1.2.

Let π_t be any Markovian strategy in the wealth Model I. We say that π_t produces an asymptotic linear arbitrage (ALA) with geometrically decaying probability (GDP) of failure if, starting with $V_0 = 0$, there are real constants $b > 0$ and $c > 0$ such that,

$$\mathbb{P}(V_t^\pi \geq bt) \geq 1 - e^{-ct}, \text{ for large time } t. \tag{2.2}$$

This means that, if a strategy π_t is an *ALA* with *GDP* of failure, then outside a set whose probability decreases geometrically fast to 0, the wealth V_t^π of an investor taking such a strategy grows linearly as t goes to infinity.

To investigate such strategies, first we consider a less realistic case, serving as a motivation and starting point, in the section below.

2.2 ALA for Stock Prices in a Compact State Space

We assume that the state space of the Markov chain X_t is a non-empty compact interval S, and $\lambda(S) > 0$. I will apply the classical Gärtner-Ellis *LDP* Theorem to derive existence of *ALA* in the wealth Model *I*. But first, we set the,

2.2.1 Structural Assumptions on the Stock Process

Let $\mathcal{B}(S)$ denote, as usual, the Borel σ-algebra on S. For $x \in S$ and $A \in \mathcal{B}(S)$, we assume that the one-step transition probability kernel $P(x, A) := \mathbb{P}(X_{t+1} \in A | X_t = x)$, $t \geq 0$, of the Markov chain X_t has a positive density $p(x, \cdot) : S \to \mathbb{R}_+$ with respect to the

Lebesgue measure λ. Denote again $P^t(x, A) := \mathbb{P}(X_t \in A | X_0 = x)$ the t-step transition probability kernel of the chain X_t.

Next, we impose the following structural conditions:

(A_1) The kernel density $p(x, \cdot)$ is uniformly positive and bounded, that is, there are constants $c, d \in \mathbb{R}$ such that $0 < c \leq p(x, y) \leq d < \infty$, for all $x, y \in S$.

(A_2) The Markovian strategies π_t are (uniformly) bounded; that is, the π's are bounded functions.

Then, first we have,

Proposition 2.2.1.

i) The t-step transition probability kernel $P^t(x, A)$ has density $p^t(x, \cdot) : S \to \mathbb{R}_+$ with respect to the Lebesgue measure λ.

ii) For all $t \geq 1$, the law of X_t also has density $p_t : S \to \mathbb{R}_+$ with respect to λ.

Proof. We prove $i)$ by induction. Indeed, for $t = 1$, $P^1(x, A) = P(x, A)$ has density $p(x, \cdot)$ by hypothesis. Suppose for $t > 1$ that $P^t(x, A)$ has density, say $p^t(x, \cdot)$, then by Chapman-Kolmogorov Theorem 1.2.4, we have

$$P^{t+1}(x, A) = \int_S P(x, dy) P^t(y, A) = \int_S P(x, dy) \int_A p^t(y, u) \lambda(du),$$

by induction hypothesis.

So if $\lambda(A) = 0$, then $P^{t+1}(x, A) = 0$, which means P^{t+1} is dominated by the Lebesgue measure λ. Hence by Radon-Nikodym Theorem, P^{t+1} also has a density $p^{t+1}(x, \cdot)$. We therefore conclude that for all $t \geq 1$, $P^t(x, A)$ has a density $p^t(x, \cdot)$.

For $ii)$, we derive it from $i)$. Indeed, for all $t \geq 1$, and all $A \in \mathcal{B}(S)$ we have,

$$\begin{aligned} P(X_t \in A) &= P^t(X_0, A) \\ &= \int_A p^t(X_0, y) \lambda(dy). \end{aligned} \tag{2.3}$$

Hence $p_t(y) := p^t(X_0, y)$; $y \in S$, is the density of X_t, as required, ■

Next, we have,

Proposition 2.2.2.

The Markov chain X_t is ψ-irreducible and aperiodic.

Proof. First, to get the irreducibility, we have to show that if $A \in \mathcal{B}(S)$ such that $\lambda(A) > 0$, then there is an integer $t \geq 1$ such that $P^t(x, A) > 0$ for all $x \in S$. Indeed, set $t := 1$, then we have

$$\begin{aligned} P(x, A) &= \int_A p(x, y) \lambda(dy) \\ &\geq \int_A c \lambda(dy) \text{ by Assumption } (A_2) \\ &= c \lambda(A). \end{aligned} \tag{2.4}$$

Since $\lambda(A) > 0$ and $c > 0$, it follows that $P(x, A) > 0$ and, by Proposition 1.2.7 hence by Proposition 1.2.8, that the chain X_t is ψ-irreducible.

For the aperiodicity property, in equations (2.4) above, setting $\nu_1 := c\lambda$, we obtain that the whole compact state space S is a ν_1-small set for the chain X_t. So we have $1 \in E_S := \{t \geq 1 :$ S is ν_t-small with $\nu_t = \delta_t \nu_1$, for some $\delta_t > 0\}$. Which implies that $d := g.c.d(E_S) = 1$. Moreover since $\lambda(S) > 0$, that is $S \in \mathcal{B}^+(S)$, we get by Theorem 1.2.11 and Definition 1.2.12, that the Markov chain X_t is aperiodic, as required ■

After getting the setup and these initial results, let us move to the key part of this section, leading to the first main result of the present thesis.

2.2.2 The Asymptotic Linear Arbitrage Theorem

First, we state and prove the following,

Lemma 2.2.3.
There is a unique invariant measure φ of the chain X_t, having a stationary positive density $\phi : S \to \mathbb{R}_+$ with respect to λ, such that the following limit holds,

$$\lim_{t\to\infty} \mathbb{P}(X_t \in A) = \varphi(A) = \int_A \phi(x)\lambda(dx), \text{ for all } A \in \mathcal{B}(S). \tag{2.5}$$

Proof. We proved in Proposition 2.2.2 above that the whole compact state space S is ν_1-small for the chain X_t, hence by Theorem 1.2.19, the Markov chain X_t is uniformly ergodic, hence ergodic. So, there is a unique invariant measure φ for the chain X_t such that $\|P^t(x, \cdot) - \varphi\| \to 0$ as $t \to \infty$ for all $x \in S$. In particular for the initial constant $X_0 \in S$, we obtain that

$$\sup_{f:|f|\leq 1} |P^t(X_0, f) - \varphi(f)| \to 0 \text{ as } t \to \infty,$$

where f runs over the set of real measurable functions on S. In other words, we have

$$\sup_{f:|f|\leq 1} \left| \int_S f(y)P^t(x, dy) - \int_S f(y)\varphi(dy) \right| \to 0 \text{ as } t \to \infty.$$

Setting $f := \mathbf{1}_A$ for any $A \in \mathcal{B}(S)$, we have in particular that $|P^t(x, A) - \varphi(A)| \to 0$ as $t \to \infty$. Since $P^t(x, A) = \mathbb{P}(X_t \in A)$, hence $\mathbb{P}(X_t \in A) \to \varphi(A)$, as $t \to \infty$.

To show that φ has a density, we have by the invariance property that for all $A \in \mathcal{B}(S)$,

$$\varphi(A) = \int_S P(x, A)\varphi(dx) = \int_S \int_A p(x, y)\lambda(dy)\varphi(dx) = \int_A \Big(\int_S p(x, y)\varphi(dx) \Big) \lambda(dy),$$

by Fubini Theorem. Hence, φ has density $\phi(y) := \int_S p(x, y)\varphi(dx)$, as required, ■

From this lemma, we get,

Proposition 2.2.4.

Let π_t be any Markovian strategy in the wealth Model I. Then, there exists $z_\pi \in \mathbb{R}$ such that the sequence of expected wealth increments $\mathbb{E}(Z_t^\pi)$ converges to z_π.

We call this real number z_π, the asymptotic expectation of the wealth increment Z_t^π.

Proof. We know by Proposition 2.2.1 that for all time t, X_t has density p_t. So for all $A, B \in \mathcal{B}(S)$, we have for $t \geq 1$,

$$\begin{aligned}\mathbb{P}(X_{t-1} \in A, X_t \in B) &= \int_A \mathbb{P}(X_t \in B | X_{t-1} = x) p_{t-1}(x)\lambda(dx) \\ &= \int_A \int_B p(x,y)\lambda(dy)p_{t-1}(x)\lambda(dx) \\ &= \int_A \int_B p(x,y)p_{t-1}(x)\lambda_2(dx,dy),\end{aligned}$$

This means that for $t \geq 1$, (X_{t-1}, X_t) has density $p(x,y)p_{t-1}(x)$, for $x, y \in S$. Next by Lemma 2.2.3 above, since $\pi(x)(y-x)p(x,y)$ is bounded on S^2 (and is measurable), we get,

$$\begin{aligned}\mathbb{E}(Z_t^\pi) &= \int_{S^2} \pi(x)(y-x)p(x,y)p_{t-1}(x)\lambda_2(dx,dy) \\ &\to \int_{S^2} \pi(x)(y-x)p(x,y)\phi(x)\lambda_2(dx,dy) \text{ as } t \to \infty.\end{aligned}$$

The later integral finite since $p(x,\cdot)$ and ϕ are probability densities, and $\pi(x)(y-x)$ is bounded on S^2. It is now enough to take $z_\pi := \int_{S^2} \pi(x)(y-x)p(x,y)\phi(x)\lambda_2(dx,dy)$, ■

Next, we derive the key *LDP* result below, whose arguments follow from [17] and [23].

Proposition 2.2.5.

Let π_t be any Markovian strategy in the wealth Model I such that $\{x : \pi(x) \neq 0\}$ has positive Lebesgue measure. Then, there is a positive analytic function $\beta(\theta)$, $\theta \in \mathbb{R}$ such that the average wealth $(V_t^\pi - v)/t$ satisfies an LDP with good convex rate function Λ^*; that is, the convex conjugate function of $\Lambda(\theta) := \log\big(\beta(\theta)\big)$.

Proof. For $\theta \in \mathbb{R}$, consider the scaled kernels $K_\theta(x,y) := e^{\theta\alpha(x,y)}p(x,y)$, where $\alpha(x,y) := \pi(x)(y-x)$, for all $x, y \in S$. Since, by Assumption (A_2), $\alpha(X_{n-1}, X_n)$ is bounded for all n, it follows by (A_2) again and by (A_1) that K_θ satisfies the conditions of Theorem 10.1 in [17], for all θ. So K_θ has a positive eigenvalue[3] $\beta(\theta)$. It hence follows by Theorem 1 in [23] that $\lim_{t\to\infty} \big(\mathbb{E}\big(e^{\theta(V_t^\pi - v)}\big)\big)^{1/t} = \beta(\theta)$, and that, $\beta(\theta)$ is analytic in

[3] As defined in [17], there are two functions $f, g \neq 0$ on S, the left and right eigenfunctions associated to $\beta(\theta)$, such that $\beta(\theta)f(y) = \int_S f(x)K_\theta(x,y)\lambda(dx)$ and $\beta(\theta)g(x) = \int_S K_\theta(x,y)g(y)\lambda(dy)$, for all $x, y \in S$.

θ. This implies by continuity of Logarithm that $\frac{1}{t}\log\mathbb{E}\big(e^{\theta(V_t^\pi - v)}\big) \to \log(\beta(\theta))$ as $t\to\infty$. Set $\Lambda(\theta) := \log\big(\beta(\theta)\big)$, for all $\theta\in\mathbb{R}$. First we consider the case where the asymptotic variance is nonzero; that is,

$$\Lambda''(0) = \beta''(0) - z_\pi^2 = \lim_{t\to\infty}(1/t)\mathrm{var}[V_t^\pi - v] > 0.$$

Then Λ satisfies the conditions of Gärtner-Ellis Theorem 1.1.13 (see the remark following Definition 1.1.12). Hence $(V_t^\pi - v)/t$ satisfies a large deviations principle in $\mathbb{R}$ with good convex rate function Λ^*.

One can check, as in Proposition 2.2.4, that

$$\Lambda''(0) = \int_{S^2}\pi^2(x)(y-x)^2 p(x,y)\phi(x)\lambda_2(dx,dy) - \left(\int_{S^2}\pi(x)(y-x)p(x,y)\phi(x)\lambda_2(dx,dy)\right)^2$$

and this can be 0 only if $\pi(x)(y-x)$ is λ_2-$a.e$ constant which happens only if $\pi(x)=0$ λ-$a.e.$, a case we exclude in the statement of this Proposition. As we required, ■

At last, before stating the first main result in this thesis, we prove first the following technical,

Lemma 2.2.6.

For every Markovian strategy π_t as in Proposition 2.2.5, the corresponding asymptotic expectation z_π is the unique minimizer of the convex rate function Λ^*. Moreover, we have $\Lambda^*(x) > 0$ for all $x\neq z_\pi$.

Proof. In the proof of Proposition 2.2.5 above, we obtained the following limit, $\lim_{t\to\infty}\big(\mathbb{E}\big(e^{\theta(V_t^\pi - v)}\big)\big)^{1/t} = \beta(\theta)$. Setting $\theta := 0$, then we get that $\beta(0) = 1$. Thus, $\Lambda(0) = \log(\beta(0)) = 0$. So, for all $x\in\mathbb{R}$, we have $\Lambda^*(x) \geq 0\times x - \Lambda(0) = 0$. Hence in particular we have $\Lambda^*(z_\pi)\geq 0$. Conversely, let us also show that $\Lambda^*(z_\pi)\leq 0$ and conclude that $\Lambda^*(z_\pi) = 0 \leq \Lambda^*(x)$ for all $x\in\mathbb{R}$. Indeed, for all $\theta\in\mathbb{R}$, we have,

$$\begin{aligned}
\theta z_\pi - \Lambda(\theta) &= \theta z_\pi + \lim_{t\to\infty}\tfrac{1}{t}\Big(-\log\mathbb{E}\big(e^{\theta\sum_{n=1}^t Z_n^\pi}\big)\Big)\\
&\leq \theta z_\pi + \lim_{t\to\infty}\tfrac{1}{t}\big(\mathbb{E}(-\theta\textstyle\sum_{n=1}^t Z_n^\pi)\big) \text{ by Jensen-inequality}\\
&= \theta z_\pi - \theta\lim_{t\to\infty}\tfrac{1}{t}\big(\textstyle\sum_{n=1}^t\mathbb{E}(Z_n^\pi)\big)\\
&= \theta z_\pi - \theta z_\pi \text{ since } \lim_{n\to\infty}\mathbb{E}(Z_n^\pi) = z_\pi\\
&= \theta(z_\pi - z_\pi)\\
&= 0.
\end{aligned}$$

Taking the supremum over all $\theta\in\mathbb{R}$ we get that $\Lambda^*(z_\pi)\leq 0$.

Hence, we have proved that $\Lambda^*(z_\pi) = 0 \leq \Lambda^*(x)$ for all $x\in\mathbb{R}$. This implies that z_π is a global minimun for Λ^*.

On the other hand, β is analytic hence differentiable on $\mathbb{R}$; and since $\beta(\theta) > 0$ for all $\theta \in \mathbb{R}$, it follows that $\Lambda = \log \beta$ is also differentiable on $\mathbb{R}$. Thus, by Proposition 1.1.5, Λ^* is strictly convex on its effective domain. We conclude by Proposition 1.1.2 that z_π is the unique minimizer of Λ^*.

Moreover, let $x_0 \neq z_\pi$ such that $\Lambda^*(x_0) \leq 0$, then $\Lambda^*(x_0) \leq \Lambda^*(x)$ for all $x \in \mathbb{R}$. This means, x_0 is a different global minimum for Λ^*, contradicting the unicity of z_π. This completes the proof, as required, ∎

Finally we state and prove the first main result as below,

Theorem 2.2.7.

For every Markovian strategy π_t in Model I such that $\lambda(\{x : \pi(x) \neq 0\}) > 0$, and arbitrarily small $\epsilon > 0$, the wealth process V_t^π satisfies the following estimate,

$$\mathbb{P}\big(V_t^\pi \geq v + (z_\pi - \epsilon)t\big) \geq 1 - e^{-t\Lambda^*(z_\pi - \epsilon)} \text{ for large time } t. \tag{2.6}$$

Proof. By Proposition 2.2.5, $(V_t^\pi - v)/t$ satisfies an *LDP* with good rate function Λ^*, so for any arbitrary small $\epsilon > 0$ we have from Gärtner-Ellis Theorem 1.1.13 that,

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}\Big(\frac{V_t^\pi - v}{t} < z_\pi - \epsilon\Big) \leq -\inf_{x\in(-\infty, z_\pi - \epsilon]} \Lambda^*(x).$$

In the proof of Lemma 2.2.6, we obtained that Λ^* is strictly convex, so it is nonincreasing on $(-\infty, z_\pi]$. It follows by this lemma that,

$$\inf_{x\in(-\infty, z_\pi - \epsilon]} \Lambda^*(x) = \Lambda^*(z_\pi - \epsilon) > 0.$$

Hence, $\mathbb{P}\big(V_t^\pi \geq v + (z_\pi - \epsilon)t\big) \geq 1 - e^{-t\Lambda^*(z_\pi - \epsilon)}$ for large time t. As we required, ∎

In this result, one may not get a straight linear growth of the wealth V_t^π in the long-run, if $z_\pi = 0$ for all strategies π_t. In the result below, using Martingale Theory, we show that in the wealth Model I, there is always Markovian strategy π_t with $z_\pi \neq 0$, hence there is always *ALA* with *GDP* of failure. Indeed,

Proposition 2.2.8.

In the wealth Model I,

i) If there is a Markovian strategy π_t with $\lambda(\{x : \pi(x) \neq 0\}) > 0$, such that $z_\pi \neq 0$, then π_t is an ALA with GDP of failure.

ii) There is no Markovian strategy π_t such that $z_\pi \neq 0$ if and only if, for λ-almost all $x \in S$, the Markov chain X_t starting from $X_0 = x$, with transition density $p(x, \cdot)$, is a martingale with respect to the natural filtration $\mathcal{F}_t$. However,

iii) Under assumption (A_1), X_t cannot be a martingale for almost all $X_0 = x$. Hence under the condition of Theorem 2.2.7, there is always ALA with GDP of failure.

Proof. *i*) Let π_t be a Markovian strategy such that $z_\pi \neq 0$. Then if $z_\pi > 0$, we choose ϵ small enough such that $z_\pi - \epsilon > 0$, hence we get an asymptotic linear arbitrage by (2.6). Similarly if $z_\pi < 0$, we choose the "opposite" strategy $-\pi$ for which $z_{-\pi} = -z_\pi$ which is strictly positive. So, with a similar choice of ϵ, one also gets an *ALA* with *GDP* of failure.

ii) Let π_t be any Markovian strategy. If X_t is a martingale with respect to $\mathcal{F}_t$ for λ-*a.e.* starting point x, then for all time t, $\mathbb{E}(X_t|\mathcal{F}_{t-1}) = X_{t-1}$. This holds whatever the law of X_{t-1} is. By a property of Conditional Expection, we get

$$\mathbb{E}\big(\pi(X_{t-1})(X_t - X_{t-1})|X_{t-1}\big) = 0.$$

Hence $\mathbb{E}(Z_t) = 0$ for all time t, implying that $z_\pi = 0$.

Conversely, suppose that for some $A \in \mathcal{B}(S)$ with $\lambda(A) > 0$ and for all $x \in A$ we have for example,

$$\mathbb{E}(X_1 - X_0|X_0 = x) = \int_S p(x,y)(y-x)\lambda(dy) > 0.$$

Then consider the Markovian strategy $\pi(x) := \mathbf{1}_A(x)$ for all $x \in S$. From the proof of Proposition 2.2.4, we have

$$\begin{aligned} z_\pi &= \int_{S^2} \pi(x)(y-x)p(x,y)\phi(x)\lambda_2(dx,dy) \\ &= \int_A \int_S (y-x)p(x,y)\lambda(dy)\phi(x)\lambda(dx) > 0. \end{aligned} \tag{2.7}$$

Since $\int_S (y-x)p(x,y)\lambda(dy) > 0$, $\lambda(A) > 0$ and ϕ is positive on S, it follows that $z_\pi > 0$.

iii) Finally, without loss of generality, we may suppose that the state space is $S = [0,1]$. If X_t were a martingale for almost all $X_0 = x$ then there would be a sequence $x_n \to 1$ such that

$$\mathbb{E}[X_1|X_0 = x_n] = x_n \to 1, \ n \to \infty.$$

On the other hand, let $M > 1$ be an upper bound for $p(x,y)$,

$$\mathbb{E}[X_1|X_0 = x_n] = \int_{[0,1]} yp(x_n,y)dy \leq \int_{[1-1/M,1]} yMdy < 1,$$

a contradiction. We may hence conclude, as required. ∎

Although the result above is new in its nature, it can only be used under the restrictive conditions (A_1) and (A_2), where one models a stock's evolution within a chosen bounded interval. This limits the scope of its applications since, in practice, stock prices in financial modeling are usually specified by stochastic difference/differential equations.

This observation forces us to move to a more realistic class of models in the following, last part of this chapter.

2.3 ALA for Stock Prices in a General State Space

In this section, using more advanced (and more recent) tools from Large Deviations Theory, I prove again the existence of ALA in the wealth Model I under a more satisfactory set of Markovian modeling conditions. The proofs heavily rely on the ergodic results for functions of Markov chains presented in the article [30]. For that, let us set out and get the,

2.3.1 New Modeling Conditions and Preliminary Results

We relax the strong condition (A_1) of Section 2.2, and we now assume that the stock prices are modeled by a stochastic difference equation evolving (possibly) in the whole real line as,

$$X_{t+1} = X_t + \mu(X_t) + \sigma(X_t)\varepsilon_{t+1}, \text{ for all } t \in \mathbb{N}, \tag{2.8}$$

where μ, $\sigma : \mathbb{R} \to \mathbb{R}$ are given measurable functions, the so-called drift and volatility of the stock, and (ε_t) is an $i.i.d$ sequence of random variables in $\mathbb{R}$, with common strictly positive density γ with respect to the Lebesgue measure λ on $\mathbb{R}$. X_0 is assumed constant. It is clear by Theorem 1.2.5, that X_t is a Markov chain in the whole state space $S = \mathbb{R}$.

We notice that the process evolution (2.8) can be thought as the time-discretization of a stochastic differential equation. Similar models were considered in the asymptotic arbitrage context in the article [14], but in continuous time. Note that, in particular, if $\mu(x) := -\alpha x$ with $0 < \alpha < 1$ and $\sigma(x) := 1$, for all $x \in \mathbb{R}$, then we get the usual discrete-time Ornstein-Uhlenbeck process (or $AR(1)$ process).

Let $\mathcal{B}(\mathbb{R})$ denote the Borel σ-algebra on $\mathbb{R}$. We assume that the chain X_t starts from some constant $X_0 = x$ in $\mathbb{R}$. Next, we set the following conditions:

(A_2) We keep the boundedness assumption on Markovian strategies π_t in the wealth Model I.

(A_3) We suppose that the drift μ is locally bounded; that is, bounded on each compact, and the volatility σ is bounded away from zero on each compact.

(A_4) We impose the bounded volatility and mean-reverting drift conditions below,

$$(i)\ \exists M > 0 \text{ such that } \sigma(x) < M \text{ for all } x, \text{ and } (ii) \limsup_{|x|\to\infty} \frac{|x + \mu(x)|}{|x|} < 1 \tag{2.9}$$

(A_5) Next, we assume the following integrability property for the law of the ε_t,

$$\exists\, \kappa > 0 \text{ such that } \mathbb{E}\left(e^{\kappa\varepsilon^2}\right) =: I < \infty \tag{2.10}$$

where the distribution of ε is the same as that of the $\varepsilon_i, i \in \mathbb{N}$. We also assume that γ is $(a.s.)$ bounded away from 0 on compacts and that it is $(a.s.)$ bounded on each compact.

We remark that (A_4) implies, in particular, that $\mu(x)$ has at most linear growth.

Observing the dynamics of the investor's wealth process in equation (2.1) in the wealth Model I, we express it in the form $V_t^\pi = V_0 + \sum_{n=1}^t g(\Phi_n)$, for all time $t \geq 1$, where $\Phi_n := (X_{n-1}, X_n)$ is the process of the two consecutive values of the stock prices process, and g is the function defined on $\mathbb{R}^2$ by $g(x,y) := \pi(x)(y-x)$.

Let $P(x, A)$, with $x \in \mathbb{R}$ and $A \in \mathcal{B}(\mathbb{R})$, be the usual transition probability kernel of the chain X_t, and $P^t(x, A)$ its t-step transition kernel. Then, we prove the large set of technical initial results below. Some of them consist of checking suitable conditions for results in the paper [30], which we extensively apply to derive ours. In most cases this is done first for the one-dimensional chain X_t, then for the two-dimensional chain Φ_t we are more interested in. Indeed, we have,

Proposition 2.3.1.
The Markov chain X_t is ψ-irreducible.

Proof. We have to prove that if $A \in \mathcal{B}(\mathbb{R})$ such tha $\lambda(A) > 0$, then, there is an integer $t \geq 1$ such that $P^t(x, A) > 0$ for all $x \in \mathbb{R}$. Indeed, for $t := 1$, we have

$$\begin{aligned} P(x,A) := \mathbb{P}\big(X_{t+1} \in A \mid X_t = x\big) &= \mathbb{P}\big(x + \mu(x) + \sigma(x)\varepsilon_{t+1} \in A\big) \\ &= \textstyle\int_{(A-x-\mu(x))/\sigma(x)} \gamma(y)\lambda(dy) \end{aligned}$$

Note that in Assumption (A_3), the assumption "σ is bounded away from zero on each compact" clearly implies that σ is strictly positive everywhere. So if $\lambda(A) > 0$, then for every $x \in \mathbb{R}$, $\lambda\big((A - x - \mu(x))/\sigma(x)\big) = \lambda(A)/\sigma(x)$ by the translation invariance property of λ, which is strictly positive. Since γ is strictly positive, we conclude that the later integral is also strictly positive. It follows by Proposition 1.2.7, that the chain X_t is λ-irreducible, and then by Proposition 1.2.8, that X_t is ψ-irreducible. As required, ■

Proposition 2.3.2.
All compact sets in $\mathbb{R}$ are ν_1-small sets for the chain X_t.

Proof. Let C be any compact subset in $\mathbb{R}$, then C is included in some closed interval $[-b, b]$, $b \in \mathbb{R}$. For all $x \in C$ and for all $A \in \mathcal{B}(\mathbb{R})$, we got from the preceeding proof that,

$$P(x,A) = \int_{(A-x-\mu(x))/\sigma(x)} \gamma(y)\lambda(dy) \tag{2.11}$$

Since by assumption, μ and σ are respectively bounded and bounded away from zero on the compact C, then, there are strictly positive constants a, c_1, c_2 such that $|\mu(x)| < a$, and $0 < c_1 < \sigma(x) < c_2$, for all $x \in C$. So, if $x \in C$, then we have $\big(C - x - \mu(x)\big)/\sigma(x) \subseteq$

$[(-2b-a)/c_1,(2b+a)/c_1] =: B$. This implies that $\bigcup_{x\in C}\big((C-x-\mu(x))/\sigma(x)\big) \subseteq B$. B is bounded, so $\gamma(x) \geq c'$ for some $c' > 0$ for $x \in B$.

Now, if $A \subseteq C$, then $(A-x-\mu(x))/\sigma(x) \subseteq B$, for all $x \in C$. So we have from (2.11) that $P(x,A) \geq c'\lambda(A)$.

Suppose now that, A is any Borel set, then we have

$$\begin{aligned} P(x,A) &\geq P(x, A\cap C) \\ &\geq c'\lambda(A\cap C) \text{ from the preceeding case} \\ &=: \nu_1(A), \text{ where } \nu_1 := c'\lambda \mathbf{1}_C. \end{aligned}$$

Hence, we conclude from Definition 1.2.9, that the compact set C is a ν_1-small set for the chain X_t, as required, ■

Proposition 2.3.3.
The Markov chain X_t is aperiodic.

Proof. Consider any compact set C in $\mathbb{R}$ such that $\lambda(C) > 0$. Then, since the chain X_t is ψ-irreducible, it follows by Proposition 1.2.8, that $\psi(C) > 0$; that is, $C \in \mathcal{B}^+(\mathbb{R})$. From the proof of the preceeding proposition, C is a ν_1-small set for the chain X_t. So, we obtain that $1 \in E_C := \{t \geq 1;$ C is ν_t-small with $\nu_t = \delta_t\nu_1$, for some $\delta_t > 0\}$. Hence $d := g.c.d(E_C) = 1$. Applying Theorem 1.2.11, we conclude using Definition 1.2.12, that the irreducible chain X_t is aperiodic, as required, ■

Proposition 2.3.4.
The process $\Phi_t := (X_{t-1}, X_t)$ is also a Markov chain in the state space $\mathbb{R}^2$.

Proof. Using Theorem 1.2.5 which also holds as in [2] even for a general Polish state space, in particular for $\mathbb{R}^2$, let us show that Φ_t is of the form $\Phi_{t+1} = \Gamma_t + \Sigma_t \cdot \mathscr{E}_{t+1}$, where $\mathscr{E}_t$ is a sequence of $i.i.d$ random variables in $\mathbb{R}^2$ and $\Gamma_t = \Gamma(\Phi_t)$, $\Sigma_t = \Sigma(\Phi_t)$ for some $\Gamma : \mathbb{R}^2 \to \mathbb{R}^2$ and $\Sigma : \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$. Indeed, using (2.8), we have for all time $t \geq 1$,

$$\begin{aligned} \Phi_{t+1} &= (X_t, X_{t+1}) \\ &= \big(X_t, X_t + \mu(X_t) + \sigma(X_t)\varepsilon_{t+1}\big) \\ &= \big(X_t, X_t + \mu(X_t)\big) + \mathrm{diag}\big(0, \sigma(X_t)\big)\big(0, \varepsilon_{t+1}\big) \\ &=: \Gamma_t + \Sigma_t \cdot \mathscr{E}_{t+1} \end{aligned}$$

where $\Gamma_t := (X_t, X_t + \mu(X_t))$, $\Sigma_t := \mathrm{diag}(0, \sigma(X_t))$ and $\mathscr{E}_{t+1} := (0, \varepsilon_{t+1})$. Because the ε_t's are $i.i.d$, the $\mathscr{E}_t$'s are also $i.i.d$. This shows that, the next state Φ_{t+1} of the process is generated from the previous state Φ_t, plus an independent noise $\mathscr{E}_{t+1}$. Which means that Φ_t is a Markov chain in $\mathbb{R}^2$, as required, ■

Proposition 2.3.5.
The Markov chain Φ_t is ψ-irreducible.

Proof. Let Q denote the transition probability kernel of the chain Φ_t, and λ_2 denote again the Lebesgue measure on $\mathbb{R}^2$. By the assumptions on the ε_i, for all $y \in \mathbb{R}$ the random variable $y + \mu(y) + \sigma(y)\varepsilon_1$ has a λ-*a.e.* positive density, $p_1(w)$. By the same argument, for all $w \in \mathbb{R}$ the random variable $y + \mu(y + \mu(y) + \sigma(y)w) + \sigma(y + \mu(y) + \sigma(y)w)\varepsilon_2$ has a λ-*a.e.* positive density $p_2(w, u)$ which can be chosen jointly measurable in (w, u). Hence, by independence of $\varepsilon_1, \varepsilon_2$, when $\Phi_0 = (x, y)$, the density of

$$\Phi_2 = (y + \mu(y) + \sigma(y)\varepsilon_1, y + \mu(y + \mu(y) + \sigma(y)\varepsilon_1) + \sigma(y + \mu(y) + \sigma(y)\varepsilon_1)\varepsilon_2)$$

with respect to λ_2 equals $p_1(u)p_2(w, u)$ and this is λ_2-*a.e.* positive. In particular, for all $A \subset \mathbb{R}^2$ with $\lambda_2(A) > 0$ we have

$$\mathbb{P}(\Phi_2 \in A) = \int_A p_1(w)p_2(w, u)\lambda_2(dw, du) > 0,$$

showing λ_2-irreducibility and hence ψ-irreducibility of Φ_t by Propositions 1.2.7, 1.2.8, ■

Proposition 2.3.6.
If C_1 and C_2 are two compact subsets in $\mathbb{R}$, then the compact rectangle $C := C_1 \times C_2$ is a ν_2-small set for the chain Φ_t.

Proof. The argument of Proposition 2.3.2 shows that $p_1(w) \leq c_1$ and $p_2(w, u) \geq c_2$ for all $(w, u) \in C_1 \times C_2$ for suitable $c_1, c_2 > 0$. This implies that C is a ν_2-small set for the chain Φ_t, with $\nu_2 := \mathbf{1}_{C_1 \times C_2} c_1 c_2 \lambda_2$. ■

Proposition 2.3.7.
The Markov chain Φ_t is also aperiodic.

Proof. One can easily extend the previous argument to show that, for compact intervals C_1 and C_2 in $\mathbb{R}$ such that $\lambda_2(C) > 0$ where $C := C_1 \times C_2$; we have that C is a ν_3-small set (actually, a ν_k-small set for all $k \geq 2$). Then, it follows that the ν_2-small set C belongs to $\mathcal{B}^+(\mathbb{R}^2)$. Next we have $d := g.c.d\big(E_C\big) = 1$ since $2, 3 \in E_C$. Applying Theorem 1.2.11 and Definition 1.2.12, we obtain that the ψ-irreducible chain Φ_t is aperiodic, as required, ■

Next, in order to check the remaining conditions to be satisfied by the chain Φ_t in [30], we prove first the following,

Lemma 2.3.8.

Let ε be a random variable in $\mathbb{R}$ satisfying (2.10) *in Condition* (A_5). *Then for every real number $a > 0$ large enough, we have,*

$$\mathbb{E}\big(e^{a|\varepsilon|}\big) \leq e^{ca^2} \text{ for some fixed constant } c > 0.$$

Proof. Set $\xi := |\varepsilon|$. Then we have

$$\begin{aligned} \mathbb{P}\big(e^{a\xi} > x\big) &= \mathbb{P}\Big(\exp\Big(\kappa\Big[\tfrac{\log(e^{a\xi})}{a}\Big]^2\Big) > \exp\Big(\kappa\Big[\tfrac{\log x}{a}\Big]^2\Big)\Big) \\ &\leq I\exp\Big(-\kappa\big(\log(x)/a\big)^2\Big) \text{ by Markov Inequality} \\ &= I(1/x)^{(\kappa/a^2)\log x}, \end{aligned}$$

recall I from (A_5).

Because the exponent $(\kappa/a^2)\log x > 2$ if and only if $x > e^{2a^2/\kappa}$, so we have,

$$\mathbb{E}\big(e^{a\xi}\big) = \int_0^\infty \mathbb{P}\big(e^{a\xi} > x\big)dx \leq e^{2a^2/\kappa} + I\int_{\exp(2a^2/\kappa)}^\infty 1/x^2 dx$$

.

Since the last integral is less than $\int_1^\infty 1/x^2dx$ which is finite, we conclude the proof by taking for example $c = 1 + (2/\kappa)$, for a large enough. ∎

After this, we move forward the recipe of preliminary results by showing,

Proposition 2.3.9.

The Markov chain X_t satisfies the drift condition $(DV3+)\,(i)$ *of [30] restated below.*

Proof. As recalled in the concluding remark of Chapter 1, this condition says that: *there are functions $V, W : \mathbb{R} \to [1, \infty)$, a small set C for the chain X_t and constants $\delta > 0$, $b < \infty$ such that the following inequality holds,*

$$\log\big(e^{-V} P\, e^V\big)(x) \leq -\delta W(x) + b\mathbf{1}_C(x), \text{ for all } x \in \mathbb{R}, \tag{2.12}$$

where $P\,e^V(x) := \int e^{V(y)}P(x, dy)$

This is equivalent to requiring that,

$$P\,e^V(x) \leq e^{V(x)-\delta W(x)+b\mathbf{1}_C(x)} \text{ for all } x \in \mathbb{R}, \tag{2.13}$$

for some such functions and constants.

Define $V(x) = W(x) := 1 + qx^2$, $x \in \mathbb{R}$, where $q > 0$ is a small number to be chosen later. And consider some compact set $C := [-K, K]$ for a large positive constant K. This is a small set for the chain X_t by Proposition 2.3.2.

Since $Pe^V(x) = \mathbb{E}\big(e^{V(X_1)} \mid X_0 = x\big) = \mathbb{E}\big(e^{V(x+\mu(x)+\sigma(x)\varepsilon)}\big)$, it follows from (2.13) that we need to show that,

$$\mathbb{E}\big(e^{1+q(x+\mu(x))^2+2q(x+\mu(x))\sigma(x)\varepsilon+q\sigma^2(x)\varepsilon^2}\big) \leq e^{(1-\delta)V(x)+b\mathbf{1}_C(x)} \text{ for all } x \in \mathbb{R} \tag{2.14}$$

To get this, it is sufficient to prove the following two conditions:

Condition 1: for $|x|$ large enough such that x lies outside C, we have

$$\mathbb{E}\big(e^{1+q(x+\mu(x))^2+2q(x+\mu(x))\sigma(x)\varepsilon+q\sigma^2(x)\varepsilon^2}\big) \leq e^{(1-\delta)V(x)} \tag{2.15}$$

Condition 2: for small $|x|$ that is, x in the suitable compact $C = [-K, K]$, we have

$$\sup_{x\in C} \mathbb{E}\big(e^{1+q(x+\mu(x))^2+2q(x+\mu(x))\sigma(x)\varepsilon+q\sigma^2(x)\varepsilon^2}\big) < G(K) \tag{2.16}$$

for some positive constant $G(K) < \infty$ and then take $b := \log G(K)$ with any $\delta \leq 1$.

For that, we have,

Proof of Condition 1. Using Condition $(A_4)\,(ii)$, for $|x|$ large enough, there is a small $\delta > 0$ such that $(x + \mu(x))^2 \leq (1 - 4\delta)x^2$. And since $1 \leq \delta(1 + qx^2)$ for $|x|$ large, it follows that $e^{1+q(x+\mu(x))^2} \leq e^{(1-3\delta)(1+qx^2)}$.

Moreover, if we choose q using Condition $(A_4)\,(i)$, such that $q\sigma^2(x) < \kappa/2$ for all x, then it is enough to show that,

$$\mathbb{E}\big(e^{2q|x+\mu(x)|M|\varepsilon|+(\kappa/2)\varepsilon^2}\big) \leq e^{2\delta qx^2}$$

By Cauchy-Schwarz Inequality, this requires to prove that,

$$\sqrt{\mathbb{E}\big(e^{4q|x+\mu(x)|M|\varepsilon|}\big)}\sqrt{\mathbb{E}\big(e^{\kappa\varepsilon^2}\big)} \leq e^{2\delta qx^2} \tag{2.17}$$

By (2.10), the second term on the left-hand side of (2.17) is the constant $\sqrt{I}$. This is smaller than $e^{\delta qx^2}$ which tends to infinity as $|x| \to \infty$. So, since again by Condition $(A_4)\,(ii)$, $4q|x + \mu(x)|M \leq 4qM|x|$ for $|x|$ large, it follows that we finally have to show that,

$$\sqrt{\mathbb{E}\big(e^{4qM|x||\varepsilon|}\big)} \leq e^{\delta qx^2} \text{ for large } |x|$$

or equivalently that,

$$\mathbb{E}\big(e^{4qM|x||\varepsilon|}\big) \leq e^{2\delta qx^2} \text{ for large } |x| \tag{2.18}$$

But by Lemma 2.3.8, the left-hand side of (2.18) is smaller than $e^{16cq^2M^2|x|^2}$ for some fixed constant $c > 0$. Hence, in addition to the first requirement on q, if one chooses q small enough such that $16q^2M^2c < 2\delta q$ that is, $q\sigma^2(x) < \kappa/2$ for all x and $8qcM^2 < \delta$. Hence the statement of the condition follows.

Proof of Condition 2. By assumption (A_3), μ is bounded above on the compact $C = [-K, K]$ by some positive constant A. Since μ is bounded on C, the function $x \mapsto (x + \mu(x))^2$ is also bounded on C. We assume it bounded above on that C by some positive constant B. So, with the later choice of q, we have the following estimate by Cauchy-Schwarz Inequality and by (2.10),

$$\begin{aligned} \mathbb{E}\Big(e^{1+q(x+\mu(x))^2+2q(x+\mu(x))\sigma(x)\varepsilon+q\sigma^2(x)\varepsilon^2}\Big) &\leq \mathbb{E}\Big(e^{1+qB+2q(K+A)M|\varepsilon|+(\kappa/2)\varepsilon^2}\Big) \\ &\leq e^{(1+qB)}\sqrt{\mathbb{E}\Big(e^{4q(K+A)M|\varepsilon|}\Big)}\sqrt{\mathbb{E}\Big(e^{\kappa\varepsilon^2}\Big)} \\ &= e^{(1+qB)}\sqrt{I}\sqrt{\mathbb{E}\Big(e^{4q(K+A)M|\varepsilon|}\Big)} \end{aligned} \tag{2.19}$$

We then choose K large enough such that $4q(K + A)M$ is also large, and we get by Lemma 2.3.8 that for all $x \in C = [-K, K]$,

$$\mathbb{E}\Big(e^{1+q(x+\mu(x))^2+2q(x+\mu(x))\sigma(x)\varepsilon+q\sigma^2(x)\varepsilon^2}\Big) \leq e^{(1+qB)}\sqrt{I}\sqrt{e^{16c'q^2(K+A)^2M^2}},$$

for a fixed constant $c' > 0$. This holds for all $x \in C$, hence Condition 2 holds true by taking the supremum over C of the left-hand side of this later inequality. This completes the proof of the whole result, as required, ■

Proposition 2.3.10.
The Markov chain Φ_t also satisfies the drift condition $(DV3+)\,(i)$ of [30].

Proof. Similar to (2.13), for apropriate functions $V, W : \mathbb{R}^2 \to [1, \infty)$, a small set C in $\mathbb{R}^2$ and some constants $\delta > 0$, $b < \infty$, we need to show that,

$$Qe^V(x, y) \leq e^{V(x,y)-\delta W(x,y)+b\mathbf{1}_C(x,y)} \text{ for all } x, y \in \mathbb{R}. \tag{2.20}$$

Define here $V(x, y) = W(x, y) := 1 + q(x^2 + y^2)$, where $q > 0$ is again a small number to be chosen later. Consider a compact rectangle $C := [-K, K] \times [-K, K]$ where $K > 0$ is also an appropriate large constant. Observing that the chain Φ_t starts at time $t = 1$ since the chain X_t starts at time $t = 0$, we have,

$$Qe^V(x, y) = \mathbb{E}\Big(e^{V(\Phi_2)} \mid \Phi_1 = (x, y)\Big) = \mathbb{E}\Big(e^{V\big(x+\mu(x)+\sigma(x)\varepsilon,\, y+\mu(y)+\sigma(y)\varepsilon\big)}\Big) \tag{2.21}$$

With the choice of V and W, it turns out one needs to show that,

$$\mathbb{E}\Big(e^{1+q[(x+\mu(x))^2+(y+\mu(y))^2]+2q[(x+\mu(x))\sigma(x)+(y+\mu(y))\sigma(y)]\varepsilon+q[\sigma^2(x)+\sigma^2(y)]\varepsilon^2}\Big) \leq e^{\Delta}, \tag{2.22}$$

for all $x, y \in \mathbb{R}$, where $\Delta := (1 - \delta)(1 + q(x^2 + y^2)) + b\mathbf{1}_C(x, y)$.

Similar to what we did in (2.15) and (2.16) with the chain X_t, the proof will be completed if we prove Condition 3 and Condition 4 below,

Condition 3: For $|x|$ and $|y|$ large enough such that (x, y) lies outside C, we have,

$$\mathbb{E}\Big(e^{1+q[(x+\mu(x))^2+(y+\mu(y))^2]+2q[(x+\mu(x))\sigma(x)+(y+\mu(y))\sigma(y)]\varepsilon+q[\sigma^2(x)+\sigma^2(y)]\varepsilon^2}\Big) \leq e^{(1-\delta)V(x,y)} \quad (2.23)$$

Indeed, again by (2.9) (ii), there is $\delta > 0$ small enough such that,

$$e^{1+q[(x+\mu(x))^2+(y+\mu(y))^2]} \leq e^{(1-3\delta)(1+q(x^2+y^2))} \text{ for } |x|, |y| \text{ large enough.}$$

Using (A_4) (i), let choose first q small enough such that $q(\sigma^2(x)+\sigma^2(y)) < \kappa/2$ for all x, y. Then, showing (2.23) is sufficient to prove the inequality,

$$\mathbb{E}\Big(e^{2q(|x+\mu(x)|+|y+\mu(y)|)M|\varepsilon|+(\kappa/2)\epsilon^2}\Big) \leq e^{2\delta q(x^2+y^2)} \text{ for large } |x|, |y|,$$

which, by Cauchy-Schwarz Inequality, we can get by showing that,

$$\sqrt{\mathbb{E}\Big(e^{4qM(|x+\mu(x)|+|y+\mu(y)|)|\varepsilon|}\Big)}\sqrt{\mathbb{E}\Big(e^{\kappa\varepsilon^2}\Big)} \leq e^{2\delta q(x^2+y^2)} \text{ for large } |x|, |y| \quad (2.24)$$

Here, by (2.10) the second factor in the left-hand side of this inequality is just $\sqrt{I}$ which is finite, and so less than $e^{\delta q(x^2+y^2)}$ which tends to $+\infty$ when $|x|, |y| \to +\infty$. Therefore, getting Condition 3 remains to show that,

$$\sqrt{\mathbb{E}\Big(e^{4qM(|x+\mu(x)|+|y+\mu(y)|)|\varepsilon|}\Big)} \leq e^{\delta q(x^2+y^2)} \text{ for large } |x|, |y|$$

or equivalently that,

$$\mathbb{E}\Big(e^{4qM(|x+\mu(x)|+|y+\mu(y)|)|\varepsilon|}\Big) \leq e^{2\delta q(x^2+y^2)} \text{ for large } |x|, |y|$$

Finally, using (2.9) (ii), it follows again by Cauchy-Schwarz Inequality that it is enough to prove both

$$\mathbb{E}\Big(e^{8qM|x|\varepsilon|}\Big) \leq e^{2\delta qx^2} \text{ and } \mathbb{E}\Big(e^{8qM|y|\varepsilon|}\Big) \leq e^{2\delta qy^2} \text{ for large } |x|, |y| \quad (2.25)$$

Which follows by Lemma 2.3.8 for some constants $c_1, c_2 > 0$ and choosing also q small enough such that $64q^2M^2c_1 < 2\delta q$ and $64q^2M^2c_2 < 2\delta q$.

Condition 4: For $|x|, |y|$ small such that (x, y) belongs to the suitable compact $C = [-K, K] \times [-K, K]$, we shall have,

$$\sup_{(x,y)\in C} \mathbb{E}\Big(e^{1+q[(x+\mu(x))^2+(y+\mu(y))^2]+2q[(x+\mu(x))\sigma(x)+(y+\mu(y))\sigma(y)]\varepsilon+q[\sigma^2(x)+\sigma^2(y)]\varepsilon^2}\Big) \leq H(K), \quad (2.26)$$

for some positive constant $H(K) < \infty$ and then take $b := \log H(K)$ with any $\delta \leq 1$.

Indeed, we use again arguments similar to those used in Condition 2. By hypothesis, suppose μ is bounded above on $[-K, K]$ by a positive constant A. Since the function

$x \mapsto (x+\mu(x))^2$ is also bounded on $[-K, K]$, we also assume it bounded above on $[-K, K]$ by a positive constant B. And with the choice of q such that $q(\sigma^2(x) + \sigma^2(y)) < \kappa/2$ for all x, y, we have by Cauchy-Schwarz Inequality and by (2.10) that,

$$\begin{aligned}\mathbb{E}\big(e^{Y}\big) &\leq e^{(1+2qB)}\mathbb{E}\big(e^{4q(K+A)M|\varepsilon|+(\kappa/2)\varepsilon^2}\big)\\ &\leq e^{(1+2qB)}\sqrt{\mathbb{E}\big(e^{8q(K+A)M|\varepsilon|}\big)}\sqrt{\mathbb{E}\big(e^{\kappa\varepsilon^2}\big)}\\ &= e^{(1+2qB)}\sqrt{I}\sqrt{\mathbb{E}\big(e^{8q(K+A)M|\varepsilon|}\big)}\end{aligned}$$

where $Y(x,y) := 1 + q[(x+\mu(x))^2 + (y+\mu(y))^2] + 2q[(x+\mu(x))\sigma(x) + (y+\mu(y))\sigma(y)]\varepsilon + q[\sigma^2(x) + \sigma^2(y)]\varepsilon^2$, for all $(x, y) \in C$.

Let choose K large such that $8q(K + A)M$ is also large, we get by Lemma 2.3.8 that for all $(x, y) \in C$,

$$\mathbb{E}\big(e^{Y}\big) \leq e^{(1+2qB)}\sqrt{I}\sqrt{e^{64cq^2(K+A)^2M^2}}, \tag{2.27}$$

for some constant $c > 0$.

This holds for all $(x, y) \in C$. So Condition 4 also holds by taking the supremum over C of the left-hand side of (2.27). Hence the whole result follows, as required ■

An immediate consequence of this is,

Corollary 2.3.11.
The Markov chain Φ_t has an invariant probability measure $\nu \sim \lambda_2$.

Proof. We have proved in the preceding proposition that the chain Φ_t satisfies the drift condition $(DV3+)\,(i)$ on page 6 of [30] with $V = W$. This means, we have verified Condition $(DV4)$ on page 12 of the same article. But by Proposition 2.1 of that article, $(DV4)$ implies the drift condition $(V4)$ on that page 12; which is the inequality (1.10) in Theorem 1.2.19 of our previous chapter. Since the chain Φ_t is ψ-irreducible and aperiodic, it follows by this latter theorem that, Φ_t is geometrically ergodic, hence ergodic. We then conclude by Remark 1.2.18, $ii)$ that Φ_t has a unique invariant probability measure, say ν.

Furthermore, from the proof of Proposition 2.3.5, $\mathbb{P}(\Phi_2 \in \cdot|\Phi_0 = (x, y))$ is λ_2-absolutely continuous for each (x, y), we easily get $\nu \ll \lambda_2$. Finally, since the chain Φ_t is λ_2-irreducible with ν as its invariant probability measure, it follows from Definition 1.2.15 $ii)$ and Proposition 1.2.16 that $\nu \sim \lambda_2$, showing the result, ■

Next, we have,

Proposition 2.3.12.
The Markov chain X_t satisfies Condition $(DV3+)\,(ii)$ of [30] restated below.

Proof. As stated in [30], this conditions says that:

For functions $V, W : \mathbb{R} \to \mathbb{R}$*, there is a time* $t_0 > 0$ *such that for all* $r < \|W\|_\infty = \infty$*, there exists a measure* β_r *on* $\mathcal{B}(\mathbb{R})$ *such that we have both*

$$\beta_r(e^V) < \infty \text{ and } P_x\big(X_{t_0} \in A, \tau_{C^c_W(r)} > t_0\big) \le \beta_r(A), \forall\, x \in C_W(r), A \in \mathcal{B}(\mathbb{R}), \qquad (2.28)$$

where $\tau_{C^c_W(r)}$ *is the first return time to* $C^c_W(r)$*, defined in Subsection 1.2.2.*

To show this, consider the same choice of $V(x) = W(x) = 1 + qx^2$, $x \in \mathbb{R}$ for Condition $(DV3+)\,(i)$ that we proved in Proposition 2.3.9. Set $t_0 = 1$ and let $r < \infty$, then we have $C_W(r) = \{x \in \mathbb{R} : 1 + qx^2 \le r\}$.

If $0 \le r < 1$, then $C_W(r) = \emptyset$, and $\tau^c_W(r) = \tau_{\mathbb{R}}$; hence (2.28) holds trivially.

Suppose now that $r \ge 1$, then we have $C_W(r) = \Big[-\sqrt{\frac{r-1}{q}}, \sqrt{\frac{r-1}{q}}\Big]$, which is compact. By Assumption (A_3), σ is bounded from below and μ is bounded on every compact.

Let us consider then the bounded set

$$H := \bigcup_{x \in C_W(r)} \Big(\frac{C_W(r) - x - \mu(x)}{\sigma(x)}\Big),$$

then, by assumption (A_5), γ is bounded from above on H by some constant $D_r > 0$. For all $x \in C_W(r)$ and all $A \in \mathcal{B}(\mathbb{R})$, $A \subset C_W(r)$ we have,

$$\begin{aligned} P_x\big(X_1 \in A, \tau_{C^c_W(r)} > 1\big) &= \mathbb{P}\big(X_1 \in A, X_1 \in C_W(r) \mid X_0 = x\big) \\ &= \mathbb{P}\big(X_1 \in A \mid X_0 = x\big) \\ &= \mathbb{P}\big(x + \mu(x) + \sigma(x)\varepsilon_1 \in A\big) \\ &\le \int_{\frac{A - x - \mu(x)}{\sigma(x)}} \gamma(y)\lambda(dy) \\ &\le D_r \lambda(A). \end{aligned} \qquad (2.29)$$

Hence we define the required measure β_r by

$$\beta_r(A) := D_r \lambda(A \cap C_W(r)), \text{ for all } A \in \mathcal{B}(\mathbb{R}).$$

Now, to complete the proof, it remains to show with this choice of β_r, that $\beta_r(e^V) < \infty$. Indeed, e^V is locally bounded and β_r has compact support, ∎

As the last in this set of preliminary results, we have,

Proposition 2.3.13.
The Markov chain Φ_t *also verifies Condition* $(DV3+)\,(ii)$ *of [30].*

Proof. It follows exactly like Proposition 2.3.12, ∎

Now, we proceed to the key part of this section, the last in this first main chapter.

2.3.2 ALA Theorem under more General LDP Conditions

In the previous propositions, we have checked that the Markov chain Φ_t satisfies all sufficient conditions for the use of results in [30]. In this subsection, we begin with proving specific results attached with the wealth Model I itself using those ergodic results in [30]. Then, finally we state and prove the another ALA theorem using again the Gärtner-Ellis LDP theorem for the new modeling conditions that we set in Subection 2.3.1.

For that, recall first that for any bounded Markovian strategy $\pi_t = \pi(X_{t-1})$, the wealth process in Model I is expressed as $V_t^\pi = V_0 + \sum_{n=1}^t g(\Phi_n)$ at each time $t \geq 1$, where g is the function defined on $\mathbb{R}^2$ by $g(x,y) := \pi(x)(y-x)$, and $\Phi_n = (X_{n-1}, X_n)$ is also a Markov chain starting at time $t=0$; assuming X_{-1} and X_0 are given fixed constants.

Next, to use Gärtner-Ellis Theorem, we need to insure first that the average sum $(V_t^\pi - V_0)/t$ satisfies an LDP; that is, the limit $\Lambda_g(\theta) := \lim_{t\to\infty} \frac{1}{t} \log \mathbb{E}(e^{\theta \sum_{n=1}^t g(\Phi_n)})$ for each $\theta \in \mathbb{R}$, exists with Λ_g satisfying the remaining conditions in Gärtner-Ellis Theorem.

By now we have established that Φ_t is ψ-irreducible, aperiodic and satisfies the (DV3+) condition of [30]. Under the conditions established, all results I cite from [30] do hold, hence we will simply refer to them in the proofs below. Indeed,

For all $\theta \in \mathbb{R}$, we observe that $\theta \sum_{n=1}^t g(\Phi_n) = \sum_{n=1}^t G_\theta(\Phi_n)$ where $G_\theta = \theta g$. Next, consider the two functions $V(x,y) = W(x,y) := 1 + q(x^2+y^2)$ in Propositions 2.3.10 and 2.3.12 such that the Markov chain Φ_t verifies Condition $(DV3+)$ in [30] with the unbounded W. And define $W_0(x,y) := 1 + q(|x|+|y|)$, for $x,y \in \mathbb{R}$. We see immediately that if $W(x,y) > r$ and $r \to \infty$, then $W(x,y) \to \infty$ faster than $W_0(x,y)$. Hence,

$$\lim_{r\to\infty} \sup_{x,y\in\mathbb{R}} \Big(\frac{W_0(x,y)}{W(x,y)} \mathbf{1}_{W(x,y)>r} \Big) = 0. \tag{2.30}$$

So Condition (6) on p. 7, in [30] is satisfied. We now consider the Banach space $L_\infty^{W_0}$ defined in that paper as $L_\infty^{W_0} := \{h : \mathbb{R}^2 \to \mathbb{C} : \sup_{x,y} \frac{|h(x,y)|}{W_0(x,y)} < \infty\}$. We equip this space with the norm $\|h\|_{W_0} := \sup_{x,y} |h(x,y)|/W_0(x,y)$.

Then, we have,

Lemma 2.3.14.
For all $\theta \in \mathbb{R}$, the function G_θ belongs to the space $L_\infty^{W_0}$.

Proof. It is enough to show this for $\theta = 1$. Indeed, by Assumption (A_2), π is bounded, so for some constant $c > 0$, we have $|\pi(x)| \leq c$ for all $x \in \mathbb{R}$. If follows that $|G_1(x,y)| \leq c|y-x|$ for all $x,y \in \mathbb{R}$. Since clearly $|y-x| \leq 1 + |x| + |y|$, then we obtain that $|G_1(x,y)| \leq c(1+|x|+|y|)$, for all $x,y \in \mathbb{R}$. Hence, taking the supremum over $(x,y) \in \mathbb{R}^2$, we get $\sup_{x,y} |G_1(x,y)|/W_0(x,y) < \infty$; that is $G_1 \in L_\infty^{W_0}$, as required, ∎

Next, consider the sequence of non-linear operators $\Gamma_t : L_\infty^{W_0} \to L_\infty^V$ defined as in [30], by setting for all $F \in L_\infty^{W_0}$ and all $(x,y) \in \mathbb{R}^2$,

$$\Gamma_t(F)(x,y) := \frac{1}{t}\log \mathbb{E}_{x,y}\Big(\exp\big(\sum_{n=1}^{t} F(\Phi_n)\big)\Big). \tag{2.31}$$

where $\mathbb{E}_{x,y}$ means that we have started the chain from $\Phi_0 := (x,y)$ and we compute the expectation accordingly. Then we get,

Proposition 2.3.15.

Let π_t be any bounded Markovian strategy in the wealth Model I with $\lambda(\{x : \pi(x) \neq 0\}) > 0$. Then there is an analytic function $\Lambda_g(\theta) := \lim_{t\to\infty} \frac{1}{t}\log \mathbb{E}_{(X_{-1},X_0)}\big(e^{\theta(V_t^\pi - V_0)}\big)$, defined for all $\theta \in \mathbb{R}$, such that the average sum $(V_t^\pi - V_0)/t$ satisfies an LDP with good convex rate function Λ_g^*.

Proof. Proposition 3.6, [30] says, there exists a non-linear operator $\Gamma : L_\infty^{W_0} \to L_\infty^V$ such that the following uniform convergence holds over balls in $L_\infty^{W_0}$,

$$\sup_{\|F-F_0\|_{W_0}\leq\delta} \|\Gamma_t(F) - \Gamma(F)\|_V \to 0 \textit{ as } t \to \infty,$$

for each F_0 and each $\delta > 0$. For every $\theta \in \mathbb{R}$, set $F := G_\theta = \theta g$ and $F_0 := 0$. Since V_t^π depends on g, it follows that for all $\theta \in \mathbb{R}$, the limit

$$\begin{aligned}\Lambda_g(\theta) := \Gamma(G_\theta)(X_{-1}, X_0) &= \lim_{t\to\infty} \tfrac{1}{t}\log \mathbb{E}_{(X_{-1},X_0)}\Big(\exp\big(\textstyle\sum_{n=1}^{t} \theta g(\Phi_n)\big)\Big)\\ &= \lim_{t\to\infty} \tfrac{1}{t}\log \mathbb{E}_{(X_{-1},X_0)}\big(e^{\theta(V_t^\pi - V_0)}\big),\end{aligned} \tag{2.32}$$

exists in $\mathbb{R}$. Moreover, by Proposition 4.3 (ii) in [30], Λ_g is an analytic function of θ.

Again from (ii) of Proposition 4.3, [30], we deduce the second-order Taylor expansion about zero as, $\Lambda_g(\theta) = \Lambda_g(0) + \theta\nu(g) + \frac{1}{2}\theta^2 v_g + O(\theta^3)$ for all $\theta \in \mathbb{R}$, where, ν is the invariant measure of Φ_t in Corollary 2.3.11, the expectation $\nu(g) := \int_{\mathbb{R}^2} g(x,y)\nu(dx,dy)$ is finite, and $v_g := \lim_{t\to\infty} \mathbb{E}_\nu \sum_{n=1}^{t} \big(g(\Phi_n) - \nu(g)\big)^2$ is the asymptotic variance given in (37), p. 24 of [30].

As in Proposition 2.2.5, one may check that $v_g = 0$ implies $\pi(x) = 0$ for λ-almost all x, hence $v_g \neq 0$ under our assumptions and hence $\Lambda_g(\theta)$ is essentially smooth.

So, applying Gärtner-Ellis Theorem 1.1.13, we conclude that $(V_t^\pi - V_0)/t$ satisfies a large deviations principle with good convex rate function Λ_g^*. As we required, ■

Remark 2.3.16.

Hence we obtain from the Taylor expansion above that, $\Lambda_g'(0) = \nu(g)$. ■

Next, we have the following useful result,

Proposition 2.3.17.
$\nu(g)$ is the unique minimizer of Λ_g^*; and we have $\Lambda_g^*(x) > 0$ for all $x \neq \nu(g)$.

Proof. Using (2.32), we see that $\Lambda_g(0) = 0$. And from the preceding remark, we have $\Lambda_g'(0) = \nu(g)$, so we get by Proposition 1.1.6 that $\Lambda_g^*(\nu(g)) = \nu(g) \times 0 - \Lambda_g(0) = 0$. On the other hand, we always have $\Lambda_g^*(x) \geq 0 \times x - \Lambda_g(0) = 0$ for all $x \in \mathbb{R}$. It follows that $\nu(g)$ is a global minimum for Λ_g^*. Since by Proposition 2.3.15, Λ_g is analytic, hence, from Proposition 1.1.5, Λ_g^* is strictly convex on its effective domain (which is, in fact, $\mathbb{R}$). Applying Proposition 1.1.2, we obtain that $\nu(g)$ is the only minimum for Λ_g^*.

Finally, by unicity of $\nu(g)$, it is immediate that $\Lambda_g^*(x) > 0$ for all $x \neq \nu(g)$, ■

Before proceeding to the main result, recall the stock prices governed by equation (2.8) in the form: $X_{t+1} - X_t = \mu(X_t) + \sigma(X_t)\varepsilon_{t+1} = \sigma(X_t)\big(\mu(X_t)/\sigma(X_t) + \varepsilon_{t+1}\big)$. Hence,

Definition 2.3.18.
The market price of risk for the stock prices X_t, is the function $\rho(x) := \mu(x)/\sigma(x)$, defined for all $x \in \mathbb{R}$.

Indeed, since $\mu(X_t)$ represents the average one-step return of the stock price while $\sigma(X_t)$ measures the one-step deviation of this stock price as driven by the random "noise" ε_t, then $\rho(X_t)$ represents the one-step return of the stock per unit volatility.

Next, let $m := \mathbb{E}(\varepsilon)$, recalling that ε is a random variable having the same distribition as the ε_t's. Then, as the final and key assumption to the *ALA* theorem, we suppose that the market price of risk function ρ satisfies the following *risk-condition* below,

$$(RC): \quad \text{the set } R_m := \{x \in \mathbb{R} \mid \rho(x) \neq -m\} \ \text{ satisfies } \lambda(R_m) > 0. \tag{2.33}$$

We interpret the set R_m as representing all states of the stock prices X_t whose market price of risk is different from the value m, the expectation of the driving noise ε_t. Then,

Lemma 2.3.19.
Suppose that the market price of risk function ρ satisties the risk-condition (RC) in (2.33) *above. Then there is a bounded Markovian strategy π^0 such that,*

$$\nu(g) = \mathbb{E}\big(\pi^0(\tilde{X}_0)(\tilde{X}_1 - \tilde{X}_0)\big) > 0, \tag{2.34}$$

where $(\tilde{X}_0, \tilde{X}_1)$ has distribution ν, the invariant probability measure of Φ_t.

Proof. Set $R_m^+ := \{x \in \mathbb{R} \mid \rho(x) > -m\}$, $R_m^- := \{x \in \mathbb{R} \mid \rho(x) < -m\}$. Since ν is a probability measure on $\mathcal{B}(\mathbb{R}^2)$, then it is well known that, there is a pair of random variables $(\tilde{X}_0, \tilde{X}_1)$ on Ω, valued in $\mathbb{R}$, and with distribution ν. Since $g(x,y) = \pi(x)(y-x)$

for all $(x, y) \in \mathbb{R}^2$, and $\nu(g) = \int_{\mathbb{R}^2} g(x, y)\nu(dx, dy)$ by definition (see Remark 2.3.16 *i*)), it follows that $\nu(g) = \mathbb{E}\big(\pi(\tilde{X}_0)(\tilde{X}_1 - \tilde{X}_0)\big)$.

Next, for all $x \in \mathbb{R}$, we have

$$\begin{aligned}
\mathbb{E}\big(\tilde{X}_1 \mid \tilde{X}_0 = x\big) &= \mathbb{E}\big(x + \mu(x) + \sigma(x)\varepsilon_0 \mid \tilde{X}_0 = x\big) \\
&= x + \mu(x) + \sigma(x)\mathbb{E}(\varepsilon_0 \mid \tilde{X}_0 = x) \\
&= x + \mu(x) + \sigma(x)\mathbb{E}(\varepsilon_0) \\
&= x + \sigma(x)(\rho(x) + m).
\end{aligned}$$

with ε_0 independent of $\tilde{X}_0$ and of the same law as the ε_t. So, if $x \in R_m$, then, as $\sigma > 0$, we have $\mathbb{E}(\tilde{X}_1 \mid \tilde{X}_0 = x) \neq x$.

Consider now the bounded Markovian strategy $\pi^0(x) := \mathbf{1}_{R_m^+}(x) - \mathbf{1}_{R_m^-}$, that is, we invest all our money in the stock whenever its market price of risk remains above $-m$, we sell the stock short when the market price of risk is below $-m$, otherwise we put everything into the bank account. By Corollary 2.3.11, ν has a λ_2-*a.e.* positive density with respect to λ_2, hence its $\tilde{X}_0$-marginal, denoted by η, has a λ-*a.e.* positive density $\ell(x)$. Therefore,

$$\begin{aligned}
\nu(g) &= \int_{\mathbb{R}} \mathbb{E}\big(\pi^0(x)(\tilde{X}_1 - x) \mid \tilde{X}_0 = x\big)\eta(dx) \\
&= \int_{R_m} \mathbb{E}\big(\tilde{X}_1 - x \mid \tilde{X}_0 = x\big)\ell(x)\lambda(dx) \\
&= \int_{R_m} \operatorname{sgn}(\mathbb{E}\big(\tilde{X}_1 - x \mid \tilde{X}_0 = x\big))\mathbb{E}\big(\tilde{X}_1 - x \mid \tilde{X}_0 = x\big)\ell(x)\lambda(dx) > 0.
\end{aligned}$$

We conclude that $\nu(g) > 0$, as required ∎

We finally derive the *ALA* result below,

Theorem 2.3.20.

Suppose that the market price of risk function ρ satisfies the risk-condition (RC) as in (2.33).

Then the Markovian strategy π^0 produces an ALA with GDP of failure; that is,

$$\mathbb{P}\big(V_t^{\pi^0} \geq V_0 + \nu(g)t/2\big) \geq 1 - e^{-t\Lambda_g^*(\nu(g)/2)} \text{ for large time } t. \tag{2.35}$$

Proof. Consider the Markovian strategy π^0 above. Proposition 2.3.15 says that $(V_t^{\pi^0} - V_0)/t$ satisfies an *LDP* with good rate function Λ_g^*. By the lemma above $\nu(g) > 0$, by Proposition 2.3.17, $\nu(g)$ is the unique minimizer of Λ_g^*, and by strict convexity, Λ_g^* is decreasing on $(-\infty, \nu(g)]$. Hence applying Gärtner-Ellis Theorem 1.1.13, we get,

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}\Big(\frac{V_t^{\pi^0} - V_0}{t} < \nu(g)/2\Big) \leq - \inf_{x\in(-\infty,\nu(g)/2]} \Lambda_g^*(x) = -\Lambda_g^*(\nu(g)/2).$$

This clearly implies that $\mathbb{P}\big(V_t^{\pi^0} \geq V_0 + \nu(g)t/2\big) \geq 1 - e^{-t\Lambda_g^*(\nu(g)/2)}$ for large time t.

To complete the proof, it remains to check that $\Lambda_g^*(\nu(g)/2) > 0$, this follows again by Proposition 2.3.17 since $\nu(g)/2 \neq \nu(g)$, ∎

Remark 2.3.21.

Compare to the previous *ALA* result, we observe in the theorem that the bounded Markovian strategy producing the *ALA* is known explicitly; which is of economic interest to investors.

Next, we end this section and hence this chapter by discussing a practical,

Example 2.3.22. The Ornstein-Uhlenbeck process.

Consider the (discrete-time) auto-regressive process $AR(1)$,

$$X_{t+1} = \alpha X_t + \varepsilon_{t+1}, \text{ for all time } t \geq 1, \tag{2.36}$$

where $|\alpha| < 1$, X_0 are constants and ε_t are standard $i.i.d$ normal $\mathcal{N}(0,1)$.

Here, the drift and volatility functions are identified as $\mu(x) = (\alpha - 1)x$ and $\sigma(x) = 1$, for all $x \in \mathbb{R}$, and are clearly measurable. Also, the market price of risk function is $\rho(x) = (\alpha - 1)x$. All the assumptions of the present section trivially hold. Here $m = 0$ and hence $R_m = \mathbb{R} \setminus \{0\}$, obviously $\lambda(R_m) > 0$. It follows that there is *ALA* with *GDP* of failure for this model of financial market.

One may construct other μ, σ which satisfy our conditions and hence the correspod- ing models admit asymptotic linear arbitrage with geometrically decaying probability of failure, ∎

Chapter 3

Asymptotic Exponential Arbitrage in Markovian Financial Markets

The results of this Chapter are directly inspired by those of [14] which were reviewed in the Introduction. Now I'll discuss a new concept of asymptotic exponential arbitrage within a new wealth model (Model II) under two different sets of conditions. First, keeping the conditions imposed on the process X_t in Subsection 2.3.1, we show existence of asymptotic exponential arbitrage with GDP of failure in Model II. Next, under different conditions (neither stronger, nor weaker than the preceding ones), using a suitable LDP result in [32] we prove two more results on asymptotic exponential arbitrage, with no GDP of failure and with GDP of failure respectively. Before that, we begin with,

3.1 Log-Markovian Modeling and Definition of AEA

We now consider a log-Markovian financial model for which, the stock prices evolution is expressed in the exponential form

$$S_t := \exp(X_t), \text{ for all time } t \in \mathbb{N}, \tag{3.1}$$

that is, where

$$\log(S_t) := X_t = X_{t-1} + \mu(X_{t-1}) + \sigma(X_{t-1})\varepsilon_t \tag{3.2}$$

is the discrete-time $\mathbb{R}$-valued Markov chain evolving as in (2.8), and X_0 is assumed to be a constant. Hence it is then an adapted process on the probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ of Section 2.1, where $\mathbb{F}$ is the natural filtration of X_t. We consider again an accompanying riskless bond normalized to $B_t = 1$ at all t. λ again denotes the Lebesgue measure on $\mathbb{R}$ and $P(x, A)$, with $x \in \mathbb{R}$, $A \in \mathcal{B}(\mathbb{R})$, is the same one-step transition probability kernel of the chain X_t.

We make some natural restrictions on trading that are absent in [14]. We assume that investors are prohibited from short selling the stock and from borrowing from the bank account. This means that at each time t, they invest a proportion $\pi_t \in [0,1]$ of their overall wealth into the stock while the rest remains in the bank account. Again, we assume that the interest of the latter is set to zero.

Formally, trading strategies are now $(\mathcal{F}_t)_{t\geq 0}$-predictable $[0,1]$-valued processes $\pi_t, t \geq 1$ (that is, π_t is $\mathcal{F}_{t-1}$-measurable). π_t represents the proportion of wealth allocated to the risky asset at time t. This has to be chosen before the price S_t is revealed, that's why predictability is imposed on the strategy. Again, due to the Markovian structure on X_t and hence on S_t, we are mostly considering Markovian strategies; that is strategies where $\pi_t = \pi(X_{t-1})$ for all time $t \geq 1$, for some measurable $\pi : \mathbb{R} \to [0,1]$.

Next, given any such strategy π_t, the corresponding wealth of an investor is therefore modeled as a process V_t^{π} obeying the dynamics,

$$\textbf{Model } II\textbf{:}\ \ V_t^{\pi} = V_{t-1}^{\pi}\big((1-\pi_t) + \pi_t(S_t/S_{t-1})\big), \text{ for all time } t \geq 1, \tag{3.3}$$

and $V_0^{\pi} = V_0 > 0$ is an investor's initial capital.

Then, similar to the introduction to asymptotic linear arbitrage (ALA) discussed in Section 2.1, we define two types of asymptotic exponential arbitrage as below,

Definition 3.1.1.

A Markovian strategy π_t is an asymptotic exponential arbitrage (AEA) in Model II if there is a constant $b > 0$ such that, for all $\epsilon > 0$, there is $t_\epsilon \in \mathbb{N}$ satisfying

$$\mathbb{P}(V_t^{\pi} \geq e^{bt}) \geq 1 - \epsilon, \ \textit{for all time } t \geq t_\epsilon. \tag{3.4}$$

We point out that Definition 3.1.1 is seemingly quite different from the conclusion of Theorem 0.0.2 of [14] that we recalled in the thesis introduction. Indeed, AEA is about the existence of a single trading strategy π_t producing arbitrage in the long-run while Theorem 0.0.2 (translated into our setting) gives for each ϵ and $t > t_\epsilon$ (possibly different) $\pi_t(\epsilon, t)$ satisfying both (3.4) and a geometrically decreasing (in t) loss bound on V_t^{π}.

It turns out, however, that AEA implies this latter kind of arbitrage. Indeed,

Proposition 3.1.2.

If there is AEA in Model II, then for each $\epsilon > 0$, there exist t_ϵ and trading strategies $\pi_t(\epsilon, t), t \geq 1$, $t \geq t_\epsilon$ satisfying $V_t^{\pi(\epsilon,t)} \geq V_0 - e^{-bt/2}$ and

$$\mathbb{P}(V_t^{\pi(\epsilon,t)} \geq e^{bt/2}) \geq 1 - \epsilon, \ \textit{for all time } t \geq t_\epsilon. \tag{3.5}$$

Proof. We may and will assume $V_0 = 1$ for the portfolio realizing AEA as well as for the portfolio we are about to construct. Fix $\epsilon > 0$, take π_t and t_ϵ as in Definition 3.1.1, fix also $t \geq t_\epsilon$ and define recursively

$$\tilde{\pi}_t := \frac{V^{\pi}_{t-1} e^{-bt/2} \pi_t}{V^{\tilde{\pi}}_{t-1}}, \quad t \geq 1.$$

One can check that $V^{\tilde{\pi}}_t = V^{\pi}_t e^{-bt/2} + 1 - e^{-bt/2}$, hence $V^{\tilde{\pi}}_t \geq 1 - e^{-bt/2}$ indeed holds and we also have

$$\mathbb{P}(V^{\tilde{\pi}}_t \geq e^{bt/2}) \geq 1 - \epsilon, \tag{3.6}$$

showing (3.5) for $\pi(\epsilon, t) := \tilde{\pi}$, as required, ■

Next, we have more importantly,

Definition 3.1.3.

We say that a Markovian strategy π_t in Model II realizes an asymptotic exponential arbitrage (AEA) with geometrically decaying probability (GDP) of failure if there are constants $b > 0$, and $c > 0$ such that,

$$\mathbb{P}(V^{\pi}_t \geq e^{bt}) \geq 1 - e^{-ct}, \ \textit{for all time } t \geq 1. \tag{3.7}$$

Similarly to Definition 2.1.2 on ALA, we interprete this by saying that, an investor may achieve exponential growth of his/her wealth in the long-run while controlling at a geometric rate the probability of failing to achieve this.

The main goal in this chapter is to prove existence of explicit AEA strategies in this wealth model under two different sets of conditions on the modeling process X_t. First,

3.2 AEA Theorems under Previous μ, σ, ε -Conditions

In this section, we assume that the Markov chain X_t evolving as in (3.2) still satisfies the conditions (A_3), (A_4), and (A_5), imposed in Subsection 2.3.1. This implies that all results we proved in that subsection under these assumptions are still valid here for the two Markov chains X_t and $\Phi_t := (X_{t-1}, X_t)$. Our purpose here is to use these results, combined again with those in [30], and apply the Gärtner-Ellis LDP theorem in order to get a first AEA theorem in the Model II.

For that, we use the same technique as in Subsection 2.3.2 by observing that, for any relative Markovian strategy π_t, the wealth in Model II can be expressed in the form

$$V^{\pi}_t = V_0 \exp\big(\sum_{n=1}^{t} f(\Phi_n)\big) = V_0 \exp\Big(t \frac{\sum_{n=1}^{t} f(\Phi_n)}{t}\Big), \tag{3.8}$$

where the function f is defined as, $f(x,y) := \log\big((1-\pi(x)) + \pi(x)\exp(y-x)\big)$ for $x, y \in \mathbb{R}$, and $\Phi_t = (X_{t-1}, X_t)$ is the preceding Markov chain starting at time zero, assuming given two initial fixed constants X_{-1} and X_0. Hence we need to show that the sequence $\log(V_t^\pi/V_0) = \sum_{n=1}^t f(\Phi_n)$ satisfies the conditions of Theorem 1.1.13 with any strategy π_t. To get this, it is sufficient to show that for all $\theta \in \mathbb{R}$, the limit $\Lambda_f(\theta) := \lim_{t\to\infty} \frac{1}{t}\log\mathbb{E}(e^{\theta\sum_{n=1}^t f(\Phi_n)})$ exists and is analytic in θ.

We use again the recipe of Subsection 2.3.1 and the results of the article [30] as follows. For all $\theta \in \mathbb{R}$, we observe that $\theta\sum_{n=1}^t f(\Phi_n) = \sum_{n=1}^t F_\theta(\Phi_n)$ where $F_\theta := \theta f$. We consider again the two functions $V(x,y) = W(x,y) := 1 + q(x^2+y^2)$, $x, y \in \mathbb{R}$, for the suitable $q > 0$ as in Propositions 2.3.10 and 2.3.12 such that Condition $(DV3+)$ of [30] is satisfied. And we also take again the function $W_0(x,y) := 1 + q(|x|+|y|)$, for $x, y \in \mathbb{R}$, such that (2.30) (Condition (6) of the same paper) is met. Hence, considering the same Banach space $\big(L_\infty^{W_0}, \|\cdot\|_{W_0}\big)$ as in Subsection 2.3.2, we get

Lemma 3.2.1.
For all $\theta \in \mathbb{R}$, the function F_θ belongs to the space $L_\infty^{W_0}$.

Proof. Similarly to Lemma 2.3.14, it is sufficient to show it for $\theta = 1$. Indeed, for all $x, y \in \mathbb{R}$, since $\pi(x) \in [0,1]$, we have $1 - \pi(x) + \pi(x)\exp(y-x) \leq 1 + \exp(y-x)$. It follows that $F_1(x,y) \leq |x| + |y| + 1$.

On the other hand, for $0 \leq a \leq 1/2$, we have $1 - a + a\exp(y-x) \geq 1/2$. And for $a > 1/2$, we have $1 - a + a\exp(y-x) \geq (1/2)\exp(y-x)$. To sum up, we get that $F_1(x,y) \geq \log(1/2) - |x| - |y|$.

We conclude that $|F_1(x,y)| \leq c(1 + |x| + |y|)$, for some constant $c > 0$. Taking the supremum over all $(x,y) \in \mathbb{R}^2$, we get that $\sup_{(x,y)} |F_1(x,y)|/W_0(x,y) < \infty$, hence the claim follows as required, ■

Next, recalling the sequence of operators defined in (2.31), we obtain,

Proposition 3.2.2.
Let π_t be a relative Markovian strategy in the wealth Model II such that $\lambda(\{x : \pi(x) \neq 0\}) > 0$. Then, there is an analytic function $\Lambda_f(\theta) := \lim_{t\to\infty} \frac{1}{t}\log\mathbb{E}_{(X_{-1},X_0)}\big(e^{\theta\sum_{n=1}^t f(\Phi_n)}\big)$, for all $\theta \in \mathbb{R}$, such that the average sum $\frac{1}{t}\log(V_t^\pi/V_0) = \frac{1}{t}\sum_{n=1}^t f(\Phi_n)$ satisfies a large deviations principle with good convex rate function Λ_f^*.

Proof. Just like Proposition 2.3.15. ■

Remark 3.2.3.
Similar to Remark 2.3.16, we get also from Proposition 4.3, [30] that, $\nu(f) < \infty$ and $\Lambda_f'(0) = \nu(f)$, where ν is the invariant measure of the chain Φ_t in Corollary 2.3.11, ■

This leads to the following useful,

Proposition 3.2.4.
If $\lambda(\{x : \pi(x) \neq 0\}) > 0$, then $\nu(f)$ is the unique minimizer of Λ_f^*, and we have $\Lambda_f^*(x) > 0$ for all $x \neq \nu(f)$.

Proof. Similar to the proof of Proposition 2.3.17, ■

Consider the function ρ as in Definition 2.3.18. This time it is more appropriate to call it log-market price of risk function. As in our present setting no short-selling is allowed, when $\rho < -m$ there is no hope to realize profit. Hence we slightly modify the (RC) condition in (2.33), and define

$$R_m^+ := \{x \in \mathbb{R} : \rho(x) > -m\}. \tag{3.9}$$

We say that (RC_+) holds if $\lambda(R_m^+) > 0$. We remark that this is the case wherever ρ is lower semicontinuous and R_m^+ is nonempty. Hence,

Lemma 3.2.5.
If the log-market price of risk function ρ satisfies the risk-condition $(RC+)$, then there is a Markovian strategy π^0 such that

$$\nu(f) = \mathbb{E}\Big(\log\big((1 - \pi^0(\tilde{X}_0)) + \pi^0(\tilde{X}_0)\exp(\tilde{X}_1 - \tilde{X}_0)\big)\Big) > 0, \tag{3.10}$$

where $(\tilde{X}_0, \tilde{X}_1)$ has distribution ν, the invariant measure of Φ_t in Corollary 2.3.11.

Proof. Similar to the proof of Lemma 2.3.19, next, if $x \in R_m^+$, then we easily obtain that $\mathbb{E}(\tilde{X}_1 \mid \tilde{X}_0 = x) > x$. Consider the explicitly defined relative Markovian strategy $\pi^0(x) := \mathbf{1}_{R_m^+}(x)$ for $x \in \mathbb{R}$, consisting again of investing all current wealth in the stock if the log-market price of risk for that stock is above $-m$, and putting everything in the bank account otherwise. Then we get,

$$\begin{aligned} \nu(f) &= \int_{\mathbb{R}} \mathbb{E}\big(\log((1-\pi^0(x)) + \pi^0(x)\exp(X_1 - x)) \mid X_0 = x\big)\eta(dx) \\ &\geq \int_{R_m^+} \mathbb{E}\big(\log\exp(X_1 - x) \mid X_0 = x\big)\eta(dx) \\ &= \int_{R_m^+} \mathbb{E}\big(X_1 - x \mid X_0 = x\big)\ell(x)\lambda(dx) > 0, \end{aligned} \tag{3.11}$$

where η and ℓ are as in the proof of Lemma 2.3.19, as we required, ■

Given this lemma, we derive the AEA with GDP result as below,

Theorem 3.2.6.
Suppose that the log-market price of risk function ρ satisfies the risk-condition (RC_+). Then the Markovian strategy π_t^0 produces an AEA with GDP of failure, indeed,

$$\mathbb{P}\big(V_t^{\pi^0} \geq e^{\log(V_0) + \nu(f)t/2}\big) \geq 1 - e^{-t\Lambda^*(\nu(f)/2)} \text{ for large time } t. \tag{3.12}$$

Proof. By Proposition 3.2.2, the sequence $\frac{1}{t}\log(V_t^{\pi^0}/V_0)$ satisfies a large deviations principle with good rate function Λ_f^*. Since $\nu(f) > 0$ by the lemma above, $\nu(f)$ is the unique minimizer of Λ_f^* by Proposition 3.2.4, and by strict convexity Λ_f^* is decreasing on $(-\infty, \nu(f)]$, then Gärtner-Ellis Theorem 1.1.13 implies that,

$$\limsup_{t\to\infty} \frac{1}{t}\log \mathbb{P}\Big(\frac{\Sigma_t}{t} < \nu(f)/2\Big) \leq -\inf_{\{x\in(-\infty,\nu(f)/2]\}} \Lambda^*(x) = -\Lambda^*(\nu(f)/2).$$

On the other hand, $\nu(f)/2 \neq \nu(f)$, implies by Proposition 3.2.4 that $\Lambda_f^*(\nu(f)/2) > 0$. Hence, we conclude from all these that the inequality,

$$\mathbb{P}\big(V_t^{\pi^0} \geq e^{\log(V_0)+\nu(f)t/2}\big) \geq 1 - e^{-t\Lambda^*(\nu(f)/2)}, \text{ for large time } t,$$

achieving the proof, as required, ■

This gives below the precise analogue of Theorem 0.0.2 in the present context,

Corollary 3.2.7.

Under the conditions of Theorem 3.2.6, for all $\epsilon > 0$, there exist $b > 0$, $t_\epsilon \in \mathbb{N}$ such that for all $t \geq t_\epsilon$ there are trading strategies $\pi_t(\epsilon, t), t \geq 1$ satisfying $V_t^{\pi(\epsilon,t)} \geq V_0 - e^{-bt/2}$ and

$$\mathbb{P}(V_t^{\pi(\epsilon,t)} \geq e^{bt/2}) \geq 1 - \epsilon. \tag{3.13}$$

Proof. This is clear from Proposition 3.1.2, ■

3.3 AEA Theorems under Different μ, σ, ε -Conditions

This section is based on our works in [35] submitted recently. Our purpose is to establish *AEA* (respectively, *AEA* with *GDP* of failure), using techiniques different from the ergodic results in [30], that we handled in Chapter 2 and Section 3.2 of the present chapter. For that, we keep wealth Model *II* of the preceding section, but we modify the conditions on μ, σ and ϵ there by assuming, instead of A_3, A_4, A_5, that,

(B_1) The drift and the volatility are bounded and the latter is nonzero; that is,

$$\exists\, N, M \in \mathbb{R} : \ |\mu(x)| < N \text{ and } 0 < \sigma(x) < M \text{ for all } x \in \mathbb{R}. \tag{3.14}$$

(B_2) The ε_t's, still *i.i.d.* with the same law as an $\mathbb{R}$-valued random variable ε, are assumed independent of X_t and have in absolute value all exponential moments; that is,

$$\text{for all time } t \text{ and for all } \kappa \in \mathbb{R}, \ \mathbb{E}(e^{\kappa|\varepsilon|}) < \infty. \tag{3.15}$$

Next, given any strategy π_t in the Model II, recall the corresponding wealth of an investor expressed in (3.8) as,

$$V_t^\pi = V_0 \exp\Big(\sum_{n=1}^{t} \log\big((1-\pi_n+\pi_n e^{X_n-X_{n-1}})\big)\Big), \text{ for all time } t \geq 1. \tag{3.16}$$

Again, we will mostly deal with Markovian strategies $\pi_t := \pi(X_{t-1})$ for some measurable $\pi : \mathbb{R} \to [0,1]$.

Then, we display the new investigating technique as follows. For a fixed $x \in \mathbb{R}$, set $Y := x + \mu(x) + \sigma(x)\varepsilon$, that is, at each one-step period of time, the random variable Y plays the role of the log-price of stock X_t in (3.1) conditional to $X_{t-1} = x$. Hence any Markovian strategy $\pi(X_{t-1})$ in this model becomes $\pi(x)$ at that time t. Since x is assumed fixed, then for simplicity we omit for the moment the dependence of π on x and we use π to denote the number $\pi(x)$ running over the interval $[0,1]$.

Next, we discuss the behavior of the function $v_x(\pi) := \mathbb{E}\log\big(1-\pi+\pi\exp(Y-x)\big)$ through its random integrand $L_x(\pi) := \log\big(1-\pi+\pi\exp(Y-x)\big)$ for $\pi \in [0,1]$. For that, since L_x is clearly twice almost surely differentiable, we have for all $\pi \in [0,1]$,

$$L_x'(\pi) = \frac{-1+e^{Y-x}}{1-\pi+\pi e^{Y-x}} \text{ and } L_x''(\pi) = \frac{-\big(e^{Y-x}-1\big)^2}{\big(1-\pi+\pi e^{Y-x}\big)^2}. \tag{3.17}$$

Hence we obtain first,

Proposition 3.3.1.
There are measurable functions J_1 and J_2 such that,

$$\max_\pi |L_x'(\pi)| < J_1 \textit{ and } \max_\pi |L_x''(\pi)| < J_2 \textit{ a.s., with } \mathbb{E}(J_i) =: D_i < \infty,\ i = 1,2.$$

Proof. We have $|L_x'(\pi)| \leq \frac{|e^{Y-x}-1|}{1/2}$ if $\pi < 1/2$ and $|L_x'(\pi)| \leq \frac{|e^{Y-x}-1|}{\frac{1}{2}e^{Y-x}}$ for $\pi \geq 1/2$. It follows that $|L_x'(\pi)| \leq 2\max\{e^{x-Y}, e^{Y-x}\} + 1$ a.s., for all $\pi \in [0,1]$. Since we have $\mathbb{E}(e^{\pm(Y-x)}) = \mathbb{E}(e^{\pm(\mu(x)+\sigma(x)\varepsilon)}) < \infty$ by assumptions (B_1) and (B_2), then if follows by setting $J_1 := 2\max\{e^{x-Y}, e^{Y-x}\} + 2$, that $\max_\pi |L_x'(\pi)| < J_1$ *a.s.* with $\mathbb{E}J_1 < \infty$.

Similarly, because $|L_x''(\pi)| = |L_x'(\pi)|^2 \leq (2\max\{e^{x-Y}, e^{Y-x}\}+1)^2$ for all $\pi \in [0,1]$, we also get $\max_\pi |L_x''(\pi)| < (2\max\{e^{x-Y}, e^{Y-x}\}+1)^2 + 1 =: J_2$ *a.s.* Moreover we clearly have $\mathbb{E}J_2 < \infty$ again by assumptions (3.14) and (3.15). This completes the proof, ∎

An immediate consequence of this proposition is,

Corollary 3.3.2.
The derivatives $v_x'(\pi) = \mathbb{E}L_x'(\pi)$ and $v_x''(\pi) = \mathbb{E}L_x''(\pi)$ exist for all $\pi \in [0,1]$.

Proof. Let $\pi \in [0,1]$, then by Lagrange Mean Value Theorem, for each small $h > 0$, there is $\xi(h) \in (\pi, \pi + h)$ such that $\frac{L_x(\pi+h)-L_x(\pi)}{h} = L'_x(\xi(h))$. This implies by the almost sure differentiability of L_x that $\frac{L_x(\pi+h)-L_x(\pi)}{h} \to L'_x(\pi)$ *a.s.* as $h \to 0$.

On the other hand we have $|\frac{L_x(\pi+h)-L_x(\pi)}{h}| = |L'_x(\xi(h))| \leq J_1$ which is integrable by Proposition 3.3.1. It follows by Lebesgue Dominated Convergence Theorem that,

$$\frac{v_x(\pi+h)-v_x(\pi)}{h} = \mathbb{E}\Big(\frac{L_x(\pi+h)-L_x(\pi)}{h}\Big) \to \mathbb{E}L'_x(\pi) \text{ as } h \to 0.$$

Hence v_x is differentiable and $v'_x(\pi) = \mathbb{E}L'_x(\pi)$ for all $\pi \in [0,1]$.

Similar arguments also give the existence of the second derivative, as required ■

Next, the key quantity for our asymptotic exponential arbitrage investigation is the function

$$r(x) := v'_x(0) = \mathbb{E}\big(e^{Y-x} - 1\big) = \mathbb{E}\big(e^{\mu(x)+\sigma(x)\varepsilon} - 1\big) \text{ for all } x \in \mathbb{R}, \tag{3.18}$$

which is clearly measurable (even continuous) in x.

Remark 3.3.3.

i) The function r has the following economic interpretation. When the current log-price of the stock is $X_t = x$, then the expected value of the return on one unit of stock (at the next time step) is $\mathbb{E}(S_{t+1}/S_t) = \mathbb{E}\big(e^{Y-x}\big) = \mathbb{E}\big(e^{\mu(x)+\sigma(x)\varepsilon}\big)$. Recalling that the price of the risk-free bond (or bank account) is assumed to be 1 all the time and thus has constant 1 expected rate of return. Hence $r(x)$ shows how much the (one-step) future return on one unit of stock exceeds the return on one unit of bond when the current log-price is x.

Therefore, a natural investment strategy consists in buying stock at time t only if $r(X_t) > 0$; that is, if the expected return of the stock is better than that of the bond.

ii) Moreover, by Jensen's inequiality, for $x \in \mathbb{R}$, then $r(x) \geq e^{\sigma(x)(\rho(x)+m)} - 1$, where $m = \mathbb{E}\varepsilon$ and $\rho = \mu/\sigma$ is the log-market price of risk as in the preceding section. Hence, since $\sigma > 0$, a sufficient condition to have $r(X_t) > 0$ is that, $\rho(X_t) > -m$; that is the log-price X_t of the stock belong to the "risk-condition" set R^+_m defined in (3.9).

So, this remark shows that, it is reasonable and natural to investigate AEA with GDP of failure in the wealth Model II, tracking the condition that $r(X_t) > 0$ for some stock prices X_t. For this purpose, we state first the following,

Lemma 3.3.4.

There is a measurable function $s : \mathbb{R} \to \mathbb{R}$ such that, for any $x \in \mathbb{R}$, if $r(x) > 0$, then $v_x(s(x)) \geq r^2(x)/4D$, where $D := D_2 = \mathbb{E}J_2$ as in Proposition 3.3.1.

Proof. Define $s(x) := \max\{0, r(x)/2D\}$, for all $x \in \mathbb{R}$. s is clearly measurable. We notice that $r(x) \leq D$, thus $s(x) \leq 1/2$. Next, let $x \in \mathbb{R}$ such that $r(x) > 0$, then, by

Proposition 3.3.1, we have $|v''_x(\pi)| \leq D$ for all $\pi \in [0,1]$. It follows by the Mean Value Theorem that for all $y \in [0,1]$, we have $|v'_x(y) - v'_x(0)| \leq Dy$. Since $v'_x(0) = r(x) > 0$, we get $v'_x(y) \geq r(x)/2$ for all $y \in [0, s(x)]$. Furthermore, for all $a > 0$, we have $v_x(a) = v_x(0) + \int_0^a v'_x(y)dy$. So, noting $v_x(0) = 0$, we conclude that $v_x(s(x)) \geq s(x)r(x)/2 \geq r^2(x)/4D$, showing the lemma, ∎

Next, consider the natural filtration $\mathcal{F}_t := \sigma(X_s, s \leq t)$, $t \geq 0$, of the log-stock prices process X_t. From (3.3), the following equality holds, for all time $t \geq 1$,

$$\frac{1}{t}\log V_t^\pi = \frac{1}{t}\log V_0 + \frac{1}{t}\sum_{i=1}^{t} \log\big(1 - \pi(X_{i-1}) + \pi(X_{i-1})e^{X_i - X_{i-1}}\big) =$$

$$\frac{1}{t}\log V_0 + \frac{1}{t}\sum_{i=1}^{t} M_i + \frac{1}{t}\sum_{i=1}^{t}\mathbb{E}\big(\log\big(1 - \pi(X_{i-1}) + \pi(X_{i-1})e^{X_i - X_{i-1}}\big)|\mathcal{F}_{i-1}\big), \qquad (3.19)$$

where $M_i := \log\big(1-\pi(X_{i-1})+\pi(X_{i-1})e^{X_i-X_{i-1}}\big) - \mathbb{E}\big(\log\big(1-\pi(X_{i-1})+\pi(X_{i-1})e^{X_i-X_{i-1}}\big)|\mathcal{F}_{i-1}\big)$.

From this, we get,

Proposition 3.3.5.
For every relative Markovian strategy π_t in the wealth Model II,
i) The process M_t is a martingale difference with respect to the filtration $\mathcal{F}_t$.
ii) And the average sequence $\sum_{i=1}^t M_i/t$ is a martingale converging to 0 almost surely.

Proof. *i*) is straightforward from the Tower Property of Conditional Expection and using Definition 1.1.14.

For *ii*), using Theorem 1.1.17, it remains to show that there is a constant $K < \infty$ such that $\mathbb{E}M_t^2 \leq K$, for all time $t \geq 1$. By the tower property of Conditional Expection, it is enough to show that $\mathbb{E}(M_t^2|\mathcal{F}_{t-1}) \leq K$ *a.s.* for all $t \geq 1$. For that, set $B_t := \mathbb{E}(A_t|\mathcal{F}_{t-1})$, where $A_t := \log\big(1-\pi(X_{t-1})+\pi(X_{t-1})e^{X_t-X_{t-1}}\big)$. Since $(a-b)^2 \leq 2(a^2+b^2)$ for all $a, b \in \mathbb{R}$, then we have $\mathbb{E}(M_t^2|\mathcal{F}_{t-1}) \leq 2\mathbb{E}(A_t^2|\mathcal{F}_{t-1}) + 2\mathbb{E}(B_t^2|\mathcal{F}_{t-1})$. By Jensen Inequality and Tower Property, we have $\mathbb{E}(B_t^2|\mathcal{F}_{t-1}) \leq \mathbb{E}\big(\mathbb{E}(A_t^2|\mathcal{F}_{t-1})|\mathcal{F}_{t-1}\big) = \mathbb{E}(A_t^2|\mathcal{F}_{t-1})$. This implies that $\mathbb{E}(M_t^2|\mathcal{F}_{t-1}) \leq 4\mathbb{E}(A_t^2|\mathcal{F}_{t-1})$.

Now, similarly to the proof of Lemma 3.2.1, since $\pi(X_{t-1}) \in [0,1]$, then we have

$$\min\{\frac{1}{2}e^{X_t - X_{t-1}}, 1/2\} \leq A_t \leq 1 + e^{X_t - X_{t-1}}.$$

Noting that $X_t - X_{t-1} = \mu(X_{t-1}) + \sigma(X_{t-1})\varepsilon$, hence $|A_t| \leq c|\mu(X_{t-1}) + \sigma(X_{t-1})\varepsilon| + 1$ for some constant $c > 0$. This implies by (3.14) in Assumption (B_1) that, for some constants K_1, K_2, we have $|A_t| \leq K_1 + K_2|\varepsilon|$. And since $K_1 + K_2|\varepsilon| \leq e^{K_1 + K_2|\varepsilon|}$, we get by Assumption (B_2) that $\mathbb{E}(A_t^2|\mathcal{F}_{t-1}) \leq K_3$ *a.s.* for some constant $K_3 < \infty$. And the result follows from this, as required, ∎

We resume the dependence notation of π on the variable x or on X_{t-1} at each time t. Then in vertue of Remark 3.3.3 and Lemma 3.3.4, consider now the relative Markovian strategy

$$\pi^a(X_{t-1}) := s(X_{t-1}), \text{ defined for all time } t \geq 1. \tag{3.20}$$

Next, we assume that the function r satisfies the following estimate,

$$\exists\, c > 0 \text{ such that } \lim_{t\to\infty} \mathbb{P}\Big(\frac{1}{t}\sum_{i=1}^{t} r^2(X_{i-1})\mathbf{1}_{\{r(X_{i-1})>0\}} < c\Big) = 0, \tag{3.21}$$

which will be regarded in Remark 3.3.10 as a discrete-time analogue of the market price of risk estimate recalled in (2), in the thesis introduction. Hence we obtain the following first result,

Theorem 3.3.6.

Suppose that r satisfies the estimate (3.21), and consider the trading strategy π^a above. Then there is a constant $b > 0$, such that for all $\epsilon > 0$, there is a time $T_\epsilon > 0$ satisfying,

$$\mathbb{P}(V_t^{\pi^a} \geq e^{bt}) \geq 1 - \epsilon, \ \textit{for all time } t \geq T_\epsilon, \tag{3.22}$$

that is, there is AEA.

Proof. The proof goes technically as follows. In the equality (3.19), the first term $\frac{1}{t}\log V_0$ goes to 0 as $t \to \infty$. By Proposition 3.3.5, the second term $\frac{1}{t}\sum_{i=1}^{t} M_i$ converges to 0 almost surely, hence in probability; that is, for all $\epsilon > 0$,

$$\lim_{t\to\infty} \mathbb{P}\Big(\big|\frac{1}{t}\sum_{i=1}^{t} M_i\big| \geq \epsilon\Big) = 0, \text{ and so } \lim_{t\to\infty} \mathbb{P}\Big(\frac{1}{t}\sum_{i=1}^{t} M_i \geq \epsilon\Big) = 0. \tag{3.23}$$

Next, the estimate the third term in (3.19) as below. Using Lemma 3.3.4 and the Markov property of the log-stock prices process X_t, we have for all $i = 1, ...t$,

$$\begin{aligned} &\mathbb{E}\big(\log(1-\pi^a(X_{i-1}) + \pi^a(X_{i-1})e^{X_i - X_{i-1}})|\mathcal{F}_{i-1}\big) = \\ &\mathbb{E}\big(\log(1-\pi^a(X_{i-1}) + \pi^a(X_{i-1})e^{X_i - X_{i-1}})|X_{i-1}\big) \quad = \quad v_{X_{i-1}}(\pi^a_i) \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \geq \quad \tfrac{r^2(X_{i-1})}{4C}\mathbf{1}_{\{r(X_{i-1})>0\}}, \end{aligned} \tag{3.24}$$

applying Lemma 3.3.4. Hence,

$$\frac{1}{t}\sum_{i=1}^{t}\mathbb{E}\big(\log(1-\pi^a(X_{i-1})+\pi^a(X_{i-1})e^{X_i-X_{i-1}})|\mathcal{F}_{i-1}\big) \geq \frac{1}{t}\sum_{i=1}^{t}\frac{r^2(X_{i-1})}{4D}\mathbf{1}_{\{r(X_{i-1})>0\}}, \tag{3.25}$$

for all time $t \geq 1$. Using (3.23) and recalling $\lim_{t\to\infty} \frac{1}{t} \log V_0 = 0$, this implies by the estimate (3.21) that,

$$\lim_{t\to\infty} \mathbb{P}\Big(\frac{1}{t} \log V_t^{\pi^a} \geq \frac{c}{4D}\Big) = 1.$$

Taking $b := c/4D$, the result follows, as required, ■

Example 3.3.7.

When $\varepsilon \sim \mathcal{N}(0,1)$; the standard normal random variable, we have

$$r(x) = e^{\mu(x)+\frac{\sigma^2(x)}{2}} - 1 \geq \mu(x) + \frac{\sigma^2(x)}{2}$$

whenever this latter is ≥ 0 (using $e^u \geq 1+u$ for $u \geq 0$). It follows that if for some $c > 0$

$$\lim_{T\to\infty} \mathbb{P}\left(\frac{1}{T}\sum_{i=1}^{T}\left(\mu(X_{i-1}) + \frac{\sigma^2(X_{i-1})}{2}\right)^2 1_{\{\mu(X_{i-1})+\frac{\sigma^2(X_{i-1})}{2}>0\}} < c\right) = 0, \tag{3.26}$$

one has AEA.

Further, we sharpen this case in the result below,

Theorem 3.3.8.

Assume the conditions on μ, σ, namely $\sigma > 0$, and assume ε is standard Gaussian. If for some $c > 0$,

$$\lim_{T\to\infty} \mathbb{P}\left(\frac{1}{T}\sum_{i=1}^{T}\left(\frac{\mu(X_{i-1})}{\sigma(X_{i-1})} + \frac{\sigma(X_{i-1})}{2}\right)^2 1_{\{\frac{\mu(X_{i-1})}{\sigma(X_{i-1})}+\frac{\sigma(X_{i-1})}{2}>0\}} < c\right) = 0. \tag{3.27}$$

Then there is AEA.

Note that, as σ is bounded, (3.27) is a weaker condition than (3.26).

Lemma 3.3.9.

If ε_1 is standard Gaussian then $|u_x''(\pi)| \leq G\sigma^2(x)$ for some $G > 0$, for all $x \in \mathbb{R}$ and for all $0 \leq \pi \leq 1/2$.

Proof. For $0 \leq \pi \leq 1/2$ we have $|u_x''(\pi)| \leq 4\mathbb{E}[e^{\mu(x)+\sigma(x)\varepsilon_1} - 1]^2$, as directly verifiable. One can compute

$$\mathbb{E}[e^{\mu(x)+\sigma(x)\varepsilon} - 1]^2 = e^{2\mu(x)+2\sigma^2(x)} - 2e^{\mu(x)+\sigma^2(x)/2} + 1 =$$
$$(e^{2\mu(x)+2\sigma^2(x)} - 1) - 2(e^{\mu(x)+\sigma^2(x)/2} - 1).$$

Fix $0 \leq m \leq 2K$ where K is a bound for both $|\mu(x)|$ and $|\sigma(x)|$. We consider a Taylor-expansion of $e^{m+s} - 1$ in $0 \leq s < 1$:

$$e^{m+s} - 1 = m + s + R(s)$$

where the remainder term $R(s)$ satisfies

$$|R(s)| \leq \frac{s^2}{2} \sup_{0 \leq t \leq 1} e^{m+t}.$$

Hence

$$|R(s)| \leq V s^2 \leq V s,$$

for some constant $V := (1/2)e^{2K+1} < \infty$ and for $0 \leq s < 1$.

It follows that for $0 \leq \sigma(x) < 1$,

$$\mathbb{E}[e^{\mu(x)+\sigma(x)\varepsilon} - 1]^2 \leq |2\mu(x) + 2\sigma^2(x) - 2(\mu(x) + \sigma^2(x)/2)| + V\sigma^2(x) = \sigma^2(x) + V\sigma^2(x).$$

If $\sigma(x) \geq 1$ then

$$\mathbb{E}[e^{\mu(x)+\sigma(x)\varepsilon} - 1]^2 \leq \mathbb{E}[e^{K+K|\varepsilon|} + 1]^2 =: H < \infty$$

by (B_2). Obviously, $H \leq H\sigma^2(x)$ for $\sigma(x) \geq 1$.

It follows that, for all x,

$$\mathbb{E}[e^{\mu(x)+\sigma(x)\varepsilon} - 1]^2 \leq \max\{1 + V, H\}\sigma^2(x),$$

showing the Lemma.

Proof (of Theorem 3.3.8). Using Lemma 3.3.9 we may repeat the same proof as for Theorem 3.3.6, but defining $s(x) := \min\{\max\{r(x)/(2G\sigma^2(x)), 0\}, 1/2\}$. We get that

$$\lim_{T\to\infty} \mathbb{P}\left(\frac{1}{T}\sum_{i=1}^{T} \frac{r^2(X_{i-1})}{\sigma^2(X_{i-1})} 1_{\{r(X_{i-1})>0\}} < c\right) = 0$$

for some $c > 0$ implies AEA. As

$$r(x) \geq \mu(x) + \frac{\sigma^2(x)}{2}$$

whenever $r(x) \geq 0$, so (3.27) indeed implies AEA and we may conclude, ■

Remark 3.3.10.

Let us summarize what we have discussed so far in the present section. We should compare Theorem 3.3.8 to the results of [14] that we recalled in the introduction, in particular to Theorem 0.0.2. First notice in Theorem 3.3.8 that, the particular estimate

(3.27) may be regarded as a discrete-time analogue of (2). Next, since Brownian motion has Gaussian increments, when ε is Gaussian, (3.2) can be regarded as a standard discretization of the stochastic differential equation for $\log S_t$ where S_t is positive and satisfies (1).

To make a reasonable comparison we should consider a typical case of (1), where $S_t = \exp(X_t), t \in [0, \infty)$ for some X_t satisfying

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t.$$

Ito's formula gives us

$$dS_t = S_t\mu(\log S_t)dt + S_t\sigma(\log S_t)dW_t + \frac{1}{2}S_t\sigma^2(\log S_t)dt.$$

From this we get that the market price of risk is

$$\phi(S_t) = \frac{\mu(\log S_t)}{\sigma(\log S_t)} + \frac{\sigma(\log S_t)}{2}.$$

We can write ϕ as a function of X_t and get

$$\phi(X_t) = \frac{\mu(X_t)}{\sigma(X_t)} + \frac{\sigma(X_t)}{2},$$

hence market price of risk estimate (2) takes the form:

$$\lim_{T\to\infty} \mathbb{P}\left(\frac{1}{T}\int_0^T \left(\frac{\mu(X_t)}{\sigma(X_t)} + \frac{\sigma(X_t)}{2}\right)^2 dt < c\right) = 0. \tag{3.28}$$

Now the analogy with (3.27) is straightforward, we only need to account for the indicators $1_{\{\frac{\mu(X_{i-1})}{\sigma(X_{i-1})}+\frac{\sigma(X_{i-1})}{2}>0\}}$; which comes from the prohibition of short-selling, ■

Further, in the statement of the *AEA* Theorem 3.3.6 and its special case Theorem 3.3.8, there is no relationship between ϵ and the time t, an investor using the trading strategy π_t^a may wait very long before reaching the time threshold t_ϵ from which s/he may then perform an exponential growth in his/her wealth. And s/he cannot control efficiently the probability of failing to perform such a wealth growth. Hence, we seek in the present μ, σ, ε-conditions, a new *AEA* result where the probability of failing to produce such an exponential growth in the wealth depends on time t and decays geometrically fast to 0; that is, as in Theorem 3.2.6 of the preceding section.

In order to achieve this goal, we construct our own *large deviations estimate* by assuming that,

$$\exists\, c_1 > 0, c_2 > 0 \;:\; \limsup_{t\to\infty} \frac{1}{t}\log\left(\mathbb{P}\Big(\frac{1}{t}\sum_{i=1}^{t} r^2(X_{i-1})\mathbf{1}_{\{r(X_{i-1})>0\}} < c_1\Big)\right) < -c_2. \tag{3.29}$$

Then we obtain the required main result below,

Theorem 3.3.11.

If the function r satisfied the LDP estimate (3.29) above, then the Markovian strategy π_t^a generates in the wealth Model II an AEA with GDP of failure.

Proof. We use a different technique by applying an LDP result for martingale differences in [32] as follows. Reconsider the martingale difference $M_t = A_t + \mathbb{E}(A_t|\mathcal{F}_{t-1})$ where $A_t := \log\big(1 - \pi(X_{t-1}) + \pi(X_{t-1})e^{X_t - X_{t-1}}\big)$ as in the proof of Proposition 3.3.5. And let us show that there is a constant $K < 0$ such that $\mathbb{E}(e^{|M_t|}|\mathcal{F}_{t-1}) \leq K$ *a.s.* for all $t \geq 1$.

Indeed, we have

$$\begin{aligned}
\mathbb{E}(e^{M_t}|\mathcal{F}_{t-1}) &= \mathbb{E}(e^{A_t - \mathbb{E}(A_t|\mathcal{F}_{t-1})}|\mathcal{F}_{t-1}) \\
&= e^{-\mathbb{E}(A_t|\mathcal{F}_{t-1})}\mathbb{E}(e^{A_t}|\mathcal{F}_{t-1}) \text{ by } \mathcal{F}_{t-1}\text{-measurability} \\
&\leq \mathbb{E}(e^{-A_t}|\mathcal{F}_{t-1})\mathbb{E}(e^{A_t}|\mathcal{F}_{t-1}) \text{ by Jensen Inequality} \\
&\leq \mathbb{E}(e^{|A_t|}|\mathcal{F}_{t-1})\mathbb{E}(e^{|A_t|}|\mathcal{F}_{t-1}) \\
&= \big(\mathbb{E}(e^{|A_t|}|\mathcal{F}_{t-1})\big)^2 \\
&= \big(\mathbb{E}(e^{|A_t|}|X_{t-1}\big)^2 \text{ by Markov Property.}
\end{aligned}$$

In that proof of Proposition 3.3.5, we obtained that $|A_t| \leq K_1 + K_2\varepsilon$ for constants K_1, K_2. It follows by Assumption (B_2) that $\mathbb{E}(e^{|A_t|}|X_{t-1}\big) < \infty$ *a.s*, hence $\mathbb{E}(e^{M_t}|\mathcal{F}_{t-1}) < \infty$ *a.s.* Similarly, we also get $\mathbb{E}(e^{-M_t}|\mathcal{F}_{t-1}) < \infty$ *a.s.* Hence we have $\mathbb{E}(e^{|M_t|}|\mathcal{F}_{t-1}) \leq K$ *a.s.* for some constant $K < \infty$. It follows by Theorem 1.1 in [32] that, for some constant $c_3 > 0$, we have

$$\mathbb{P}\Big(\Big|\frac{\sum_{i=1}^t M_i}{t}\Big| \geq \frac{c_1}{4D}\Big) \leq e^{-c_3 t}, \text{ for large time } t. \tag{3.30}$$

Using the LDP estimate (3.29) and again the inequality (3.24), then setting $c := \min\{c_2, c_3\}$ and $b := c_1/4D$, we obtain from (3.19) and by (3.30) above that,

$$\limsup_{t\to\infty} \frac{1}{t}\log\mathbb{P}\Big(\frac{1}{t}\log V_t^{\pi^a} - \frac{1}{t}\log V_0 \leq b\Big) \leq -c. \tag{3.31}$$

Hence,

$$\mathbb{P}\big(V_t^{\pi^a} \geq e^{\log V_0 + bt}\big) \geq 1 - e^{ct}, \text{ for large time } t; \tag{3.32}$$

which shows that the trading strategy π_t^a yields an AEA with GDP of failure, ■

Next, we now give easily verifiable sufficient conditions for (3.29).

Theorem 3.3.12.

In addition to conditions (B_1) and (B_2), let us assume that the law of ε is absolutely continuous with respect to the Lebesgue measure with a density $\gamma(u), u \in \mathbb{R}$ that is bounded away from 0 on compacts. Assume further again that $\sigma(x)$ is bounded away from 0 on *compacts and* $\{x \in \mathbb{R} : r(x) > 0\}$ *has positive Lebesgue-measure. If there is a measurable function* $V : \mathbb{R} \to [1, \infty)$ *such that for all* $x \in \mathbb{R}$,

$$\mathbb{E}[V(X_1)|X_0 = x] \leq (1-\delta)V(x)1_{\{x \notin C\}} + b1_{\{x \in \Gamma\}} \tag{3.33}$$

for a bounded interval $C := [c, d]$, $c < d$ *and for some* $0 < \delta < 1$, $b > 0$,

then (3.29) *holds true and hence there is AEA with GDP of failure.*

Lemma 3.3.13.

Under the conditions of Theorem 3.3.12, the Markov chain X_t *is* λ*-irreducible and aperiodic; intervals* $[c, d]$ *with* $c < d$ *are small sets and* X_t *is geometrically ergodic.*

Proof. Irreducibility and aperiodicity follows just like in Chapter 2 together with the fact that compact sets are small. The drift condition (3.33) implies geometric ergodicity, see Theorem 1.2.19, since C is a small set. ∎

Proof. (of Theorem 3.3.12). One can show as in Corollary 2.3.11 that the chain X_t also has an invariant probability measure $\nu_1 \sim \lambda$. Define $F(u) := r^2(u)1_{\{r(u)>0\}}$, this is bounded and measurable. Then $\nu_1 \sim \lambda$ implies that $z = \int_{\mathbb{R}} F(u)\nu_1(du) > 0$. The chain X_t is Lebesgue-irreducible, aperiodic and geometrically ergodic by Lemma 3.3.13 above. One may always assume that $\int_{\mathbb{R}} V^2(x)\nu_1(dx) < \infty$, see Theorem 14.0.1 and Lemma 15.2.9 of [36]. Hence

$$v^2 := \lim_{t\to\infty} \frac{1}{t} \text{var}[F(X_0) + \ldots + F(X_{t-1})]$$

is well defined, see p. 317 of [29]. If $v^2 = 0$ then (i) of Proposition 2.4 in [29] shows that F is Lebesgue $a.s.$ constant. In this case Theorem 3.3.12 follows trivially. Hence we may and will assume $v^2 > 0$.

Theorem 4.1 and $P4$ on page 343 from [29] show that there is $\theta > 0$ and an analytic function $\Lambda(\alpha)$, $\alpha \in (z-\theta, z+\theta)$. such that

$$\lim_{t\to\infty} \frac{1}{t} \ln \mathbb{E}e^{\alpha(F(X_0)+\ldots+F(X_{t-1}))} = \Lambda(\alpha)$$

and $\Lambda''(\alpha) = \rho^2 > 0$. We may assume that θ is so small that $\Lambda''(\alpha) > 0$ for $\alpha \in (z-\theta, z+\theta)$, hence $I(\beta) := (\Lambda')^{-1}(\beta)$ is well-defined for $\beta \in (\Lambda'(z-\theta), \Lambda'(z+\theta)) =: (\underline{b}, \overline{b})$. Then the Legendre-transform

$$\Lambda^*(\beta) := \sup_{\alpha \in (z-\theta, z+\theta)} [\beta\alpha - \Lambda(\alpha)]$$

can be written as $\Lambda^*(\beta) = \beta I(\beta) - \Lambda(I(\beta))$ for $\beta \in (\underline{b}, \overline{b})$ and one may check that $(\Lambda^*)''(\beta) = 1/\Lambda''(I(\beta)) > 0$ for $\beta \in (\underline{b}, \overline{b})$ showing the strict convexity of Λ^*. As easily seen, $\Lambda^*(\beta) \geq 0$ for all $\beta \in (\underline{b}, \overline{b})$ and $\Lambda^*(z) = 0$ hence for all $\kappa \in (z-\theta, z)$, $\Lambda^*(\kappa) > 0$.

Theorem 4.1 of [29] and the Gärtner-Ellis Theorem 1.1.13 guarantee that the following large deviation principle holds:

$$\mathbb{P}\left(\frac{\sum_{i=1}^{t} F(X_{i-1})}{t} < \kappa\right) \leq ce^{-t\Lambda^*(\kappa)},$$

for some $c > 0$. This shows that (3.29) holds true and then Theorem 3.3.11 allows us to conclude, ∎

We end this chapter by showing that the log-stock prices process X_t satisfies (3.33) provided that the drift $\mu(x)$ is "mean-reverting enough". Indeed,

Proposition 3.3.14.
If there are constants $N_+, N_- > 0$ such that

$$\mu(x) \leq -M \text{ for } x \geq N_+ \text{ and } \mu(x) \geq M \text{ for } x \leq -N_-,$$

then there exists $M > 0$, depending on σ, ε_1 such that X_t satisfies (3.33).

Proof. Let K_σ, K_μ denote bounds for $|\sigma|, |\mu|$, respectively. Let us take the Lyapunov function $V(x) := e^{|x|}$ and note

$$\mathbb{E}[V(X_1)|X_0 = x] \leq e^{|x+\mu(x)|}L_1 = e^{x+\mu(x)}L_1$$

for $x \geq K_\mu$ with $L_1 := \mathbb{E}e^{K_\sigma|\varepsilon_1|}$. Similarly,

$$\mathbb{E}[V(X_1)|X_0 = x] \geq e^{|x+\mu(x)|}L_2 = e^{x+\mu(x)}L_2$$

for $x \leq -K_\mu$ with $L_2 = \mathbb{E}e^{-K_\sigma|\varepsilon_1|}$. Let $M := 1 + \max\{\ln L_2, -\ln L_1\}$, take N_-, N_+ as in the hypothesis. Define

$$C := [\min\{-N_-, -K_\mu\}, \max\{N_+, K_\mu\}].$$

We can see that for $x \notin C$,

$$\mathbb{E}[V(X_1)|X_0 = x] \leq (1 - \delta)V(x)$$

for some $\delta > 0$. It is clear that for all $x \in C$,

$$\mathbb{E}[V(X_1)|X_0 = x] \leq c$$

for a suitable $c > 0$, hence (3.33) holds, as required, ∎

We remark that the condition of the above Proposition is much weaker than (A_4) (ii) of the previous Chapter. This shows that, though we put more stringent conditions on μ, σ in the present section than in section 3.2, in exchange we may relax the mean-reverting condition imposed there.

Chapter 4

Utility-Based AEA Strategies in Discrete-Time Financial Markets

In this last chapter, after reviewing the concept of expected utility, I discuss the link between the previously constructed AEA strategies and the corresponding expected utility performance in the long-run for suitable subclass of investors' utility functions. Indeed,

4.1 Introductory Review of Expected Utility

In economic theory agents are assumed to act according to their preferences. Preferences express agents' attitude towards risk. One way of representing these preferences in a quantitative way is to use utility functions. Such a utility U is defined on (a subinterval of) the real line and $U(x)$ is interpreted as the subjective value of holding x dollars for the given agent. In other words, $U(x)$ expresses a level of satisfaction for an agent holding x dollars in an investement.

The most widely used utility functions go back to von Neumann and Morgenstern, these are the ones we are dealing with here. Formally, as discussed in the textbook [13], we have,

Definition 4.1.1.
A function $U : (0, \infty) \to \mathbb{R}$ *is called a utility function if it is strictly increasing and strictly concave.*

As one can see fron from Theorem 10.1 of [46], concave functions are contituous.

We interpret this definition as follows. Utility functions are assumed increasing because investors usually prefer more money than less. Let us assume that an agent pursues strategy π_t (in the wealth model II from Chapter 3). The concavity of U expresses the fact that investors are "risk-averse"; that is, by Jensen's inequality, $U(\mathbb{E}V_t^\pi) \geq \mathbb{E}U(V_t^\pi)$. This means that their satisfaction $U(\mathbb{E}V_t^\pi)$ from the deterministic amount $\mathbb{E}V_t^\pi$ (their expected future wealth) is higher than the expected value $\mathbb{E}U(V_t^\pi)$ of their satisfaction from the (uncertain) random amount V_t^π. Hence they assume a risk-averse attitude, by preferring, in a certain sense, deterministic to the uncertain. Another interpretation of concavity is that $U'(x)$ expresses how a "small" amount added to x increases the agent's satisfaction and that this increases as x decreases; that is, the investor becomes more and more sensitive to losses. This again corresponds to a risk-averse behaviour.

Let a risk averse investor with initial capital $V_0 = x \in \mathbb{R}$ express her/his preference over a risky investment in the market in term of a utility criterion U. Then,

Definition 4.1.2.

An optimal investment problem or utility maximization problem for this investor on a finite time horizon $[0,T]$ *consists of finding an optimal strategy* $(\pi_t^*)_{0\leq t\leq T}$ *that maximizes the expected utility* $\mathbb{E}U(V_T^{\pi_t})$ *of her/his terminal wealth over all strategies admissible in some sense; that is, s/he seeks both the maximal expected utility,*

$$u(x) = \sup_{\pi_t} \mathbb{E}_x U(V_T^{\pi_t}), \tag{4.1}$$

and a trading strategy $(\pi_t^*)^*_{0\leq t\leq T}$ *such that* $u(x) = \mathbb{E}U(V_T^{\pi_t^*})$ *for all initial capital* $x \in \mathbb{R}$.

Under appropriate settings finite horizon optimal investment problems are well discussed in the literature, see for intance [44]. This study depends more on the choice of the risk-aversion utility function.

As presented in [13], there are two important classes of risk aversion utility functions. The class of Constant Absolute Risk Aversion ($CARA$) utilities $U(x) := 1 - e^{-\alpha x}$, $x \in \mathbb{R}$, with $\alpha > 0$, defined on the whole real line and the class of Hyperbolic Absolute Risk Aversion ($HARA$) utilities $U(x) = \log x$, $U(x) = x^\alpha$, with $0 < \alpha < 1$, and $U(x) = -x^\alpha$, with $\alpha < 0$, for all $x \in (0,\infty)$. We concentrate on these latter. α is a risk-aversion parameter here; the larger $-\alpha$ is, the more afraid the agents become of losses.

In this last chapter, we do not intend to solve the utility maximization problem (4.1), but instead, we analyse the relationship between asymptotic exponential arbitrage discussed in the previous wealth Model II and utility maximization problems (4.1). More precisely, using asymptotic exponential arbitrage strategies, will the expected utility of the investor tend to $U(\infty)$ (the maximal utilty, which can be finite or ∞)? How fast this convergence will take place?

Finally, there is another related question: investors are thought to trade in such a way that they maximize their expected utility on the given trading period $[0,T]$. Pursuing such a trading strategy, will they have asymptotic exponential arbitrage in the sense of the previous chapter (that is, AEA or AEA with GDP of failure)? In the main section below, we provide some answers to theses questions for the above subclass of power utilities and under suitable assumptions.

4.2 AEA versus HARA Expected Utility

As the unique main section of this chapter, we consider trading in the wealth Model II of Chapter 3. All modeling objets, the log-stock prices process X_t, Markovian strategies π_t and the corresponding wealth process V_t^π are still assumed relative to the same filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ of the preceding chapters.

Next, condider first the subclass of power utility functions $U(x) := x^\alpha$, with $0 < \alpha < 1$, for $x \in (0, \infty)$. Then we derive the following result,

Proposition 4.2.1.

If a trading strategy π_t realizes an AEA as in Definition 3.1.1, in the wealth Model II, then there is a constant $b > 0$ such that,

$$\mathbb{E}U(V_t^\pi) \geq e^{\alpha bt}, \text{ for large enough time } t; \tag{4.2}$$

that is, the expected utility of the corresponding wealth grows exponentially fast.

Proof. By definition, there are constants $b > 0$ and t_0 such that $\mathbb{P}(V_t^\pi \geq e^{a+bt}) \geq 1/2$ for $t \geq t_0$.

We have $\mathbb{E}U(V_t^\pi) \geq \mathbb{E}U(e^{bt})\mathbf{1}_{\{V_t^\pi \geq e^{bt}\}} = U(e^{bt})\mathbb{P}(V_t^\pi \geq e^{bt}) \geq (1/2)e^{\alpha bt} \geq e^{\alpha b't}$ for any $b' < b$ and for large time t. Hence we have an exponential growth in the expected utility, as we required, ■

More generally, if we do not stick to having a convergence rate, consider the following larger class of utility functions $U : (0, \infty) \to \mathbb{R}$ satisfying $U(0) := \lim_{x \to 0} U(x)$ is finite. Then one may prove the following easy statement for AEA strategies,

Proposition 4.2.2.

If a trading strategy π_t is an AEA in the wealth Model II in the sense of Definition 3.1.1, then

$$\mathbb{E}U(V_t^\pi) \to U(\infty), \text{ as } t \to \infty; \tag{4.3}$$

meaning that, the expected utility of the corresponding wealth converges to the maximal utility.

Proof. If (4.3) did not hold, there would be a subsequence such that $V_{t_k}^{\pi} \to \infty$ almost surely, $k \to \infty$ and $\lim_{k\to\infty} \mathbb{E}U(V_{t_k}^{\pi}) \to G < U(\infty)$. This would contradict Fatou Lemma, since $U(V_{t_k}^{\pi}) \geq U(0) > -\infty$ and hence $\liminf_{k\to\infty} \mathbb{E}U(V_{t_k}^{\pi}) \geq U(\infty)$. Hence the result follows, as required, ■

Consider now that investor trading in the Model II choosing their utility in the the second subclass of power utility functions $U(x) := -x^{\alpha}$, with $\alpha < 0$, for all $x \in (0, \infty)$. These functions express larger risk-aversion and are thought to be more realistic. We derive the first key result of the chapter as below. We remark that, despite of the short proof, this Theorem relies on all the heavy machinery of the paper [30] as well as on our preliminary work in Section 2.3 and it is, in fact, highly non-trivial. Indeed,

Theorem 4.2.3.

Suppose the log-stock prices process X_t in (2.8) *satisfies all the conditions of Section 3.2 in the preceding chapter. Let π_t be any Markovian strategy in Model II. Then there is $\alpha_0 < 0$ such that for any risk-aversion coefficient $0 > \alpha > \alpha_0$ in the subclass above, the expected utility of the corresponding wealth converges to* 0 *at an exponential rate; that is, with the power utility $U(x) := -x^{\alpha}$, we have,*

$$|\mathbb{E}U(V_t^{\pi})| \leq Ke^{-ct}, \text{ for large time } t, \tag{4.4}$$

for some constants $K = K(\alpha), c = c(\alpha) > 0$.

Proof. Under the Assumptions of section 3.2 and using the notation there, $\Lambda_f(0) = 0$, $\Lambda_f'(0) = \nu(f) > 0$, and Λ' being continuous, there exists $\alpha_0 < 0$ such that $\Lambda(\alpha) < 0$ for $\alpha \in (\alpha_0, 0)$. Theorem 3.1 of [30] implies that for some (positive) constant c_{α},

$$\frac{-\mathbb{E}e^{\alpha(f(\Phi_1)+\ldots+f(\Phi_n))}}{e^{n\Lambda(\alpha)}} = \frac{\mathbb{E}U(V_n^{\pi})/V_0^{\alpha}}{e^{n\Lambda(\alpha)}} \to c_{\alpha},\ n \to \infty, \tag{4.5}$$

showing the statement, ■

It seems that in general we should not expect more than Theorem 4.2.3 (i.e. the same result for all $\alpha < 0$). To see this, we construct an example as below, such that there is *AEA* with *GDP* of failure but for some $\alpha < 0$, we have $\mathbb{E}U(V_t^{\pi}) \to -\infty$.

Example 4.2.4.

Consider the log-stock prices process X_t governed by the equation $X_{t+1} = X_t + \varepsilon_{t+1}$, $t \in \mathbb{N}$, with $X_0 = 0$, where ε_t are *i.i.d* random variables in $\mathbb{R}$ with common distribution chosen such that $\mathbb{E}e^{-\varepsilon_1} > 1$ and $\mathbb{E}\varepsilon_1 > 0$. For example $\varepsilon_1 \sim \mathcal{N}(1/4, 1)$ will do. We identify the drift and volatility as $\mu \equiv 0$ and $\sigma \equiv 1$.

Choose the trading strategy $\pi_t \equiv 1$ for all time t and let let $V_0 = 1$. Then we have $V_t := \exp(\varepsilon_1 + \cdots + \varepsilon_t)$ for all time $t \geq 1$. As $1/5 < 1/4 = \mathbb{E}\varepsilon_1$, by Theorem 1.1.11 , for each $\epsilon > 0$, there is $c, t_0 > 0$ such that for all $t \geq t_0$, we have $\mathbb{P}(V_t \geq e^{t/5}) \geq 1 - e^{-ct}$. Hence there is *AEA* with *GDP* of failure.

However, for $\alpha = -1$, we have by independence

$$\mathbb{E}U(V_t) = \mathbb{E}(-V_t^{-1}) = -\mathbb{E}\exp\{-(\varepsilon_1 + \cdots + \varepsilon_t)\} = -(\mathbb{E}e^{-\varepsilon_1})^t \to -\infty$$

as $t \to \infty$, surprisingly ! ■

Finally we investigate what happens if a risk-averse agent produces expected utility tending to $0 = U(\infty)$ exponentially fast as $t \to \infty$. It turns out that his/her strategy produces *AEA* with *GDP* of failure; that is arbitrage in the almost sure sense. This is a kind of converse to Theorem 4.2.3 above, inspired by Proposition 2.2 of [14]. Indeed,

Proposition 4.2.5.
Consider the power utility $U(x) = -x^\alpha$ for some $\alpha < 0$.
Let π_t be a trading strategy in the wealth Model II such that $|\mathbb{E}U(V_t^\pi)| \leq Ke^{-ct}$ for large enough time t, for some constants $c, K > 0$. Then π_t gives an AEA with geometrically decaying probability of failure.

Proof. We may assume $K = 1$, then we need to find constants $b > 0$, $c' > 0$ such that $\mathbb{P}(V_t^\pi \geq e^{bt}) \geq 1 - e^{-c't}$ for large time t.

Let any $b > 0$ such that $c + \alpha b > 0$, then we have

$$\begin{aligned} \mathbb{P}(V_t^\pi < e^{bt}) &\leq \mathbb{P}\big(|U(V_t^\pi)| \geq |U(e^{bt})|\big) \\ &\leq \frac{\mathbb{E}|U(V_t^\pi)|}{|U(e^{bt})|} \text{ by Markov Inequality.} \end{aligned} \tag{4.6}$$

But $\mathbb{E}|U(V_t^\pi)| = |\mathbb{E}U(V_t^\pi)| \leq e^{-ct}$ and also since $|U(e^{bt})| = |-e^{\alpha bt}| = e^{\alpha bt}$, then we have $\mathbb{P}(V_t^\pi < e^{bt}) \leq e^{-ct}e^{-\alpha bt} = e^{-(c+\alpha b)t}$. Which implies that $\mathbb{P}(V_t^\pi \geq e^{bt}) \geq 1 - e^{-(c+\alpha b)t}$ for large t. Hence taking $c' := c + \alpha b$, the result follows as required, ■

To summarize, if a *HARA* utility maximizer with $\alpha < 0$ achieves a utility that converges exponentially fast to 0, then his/her strategy provides *AEA* with *GDP* of failure, too. Conversely, under the stringent conditions of Section 3.2, one is able to construct strategies producing *AEA* with *GDP* of failure which also give a utility tending to 0 exponentially fast *for α large enough* (that is, for not too risk-averse investors).

References

[1] P.H. Algoet and T.M. Cover; Asymptotic Optimality and Asymptotic Equipartition Properties of Log-Optimum Investment, *Ann. Probab.*, Vol. 16, 876-898, 1988.

[2] R.N. Bhattacharya and E.C. Waymire; Stochastic Processes with Applications, John Wiley & Sons, Inc., New York, 1990.

[3] J.M. Borwein and A.S. Lewis; Convex Analysis and Nonlinear Optimization. Theory and examples, Springer, 2000.

[4] S. Boyd and L. Vandenberghe; Convex Optimization, Cambridge University Press, 2004.

[5] M. De Donno; A Note on Completeness in Large Financial Markets, *Math. Finance*, Vol. 14 (2004), No. 2, 295-315.

[6] F. Delbaen and W. Schachermayer; The Mathematics of Arbitrage, Springer Finance, 2006.

[7] C. Dellacherie and P.A. Meyer; Probabilities and Potential, *North-Holland Mathematics Studies*, Vol. 29, North-Holland Publishing Co., Amsterdam, 1978.

[8] A. Dembo and O. Zeitouni; Large Deviations Techniques and Applications, Second Edition, Springer, 1998.

[9] N. Dokuchaev; Mean-Reverting Market Models: Speculative Opportunities and Non-Arbitrage, *Appl. Math. Finance*, Vol. 14, 319-337, 2007.

[10] M.D. Donsker and S.R.S. Varadhan; Asymptotic Evaluation of Certain Markov Process Expectations for Large Time, IV, *Comm. Pure Appl. Math.*, Vol. 36(2), 183-212, 1983.

[11] J.L. Doob; Stochastic Processes, Wiley, 1953.

[12] J. Elton; A Law of Large Numbers for Identically Distributed Martingale Differences, *Ann. Probab.*, Vol. 9, 405-412, 1981.

[13] H. Föllmer and A. Schied; Stochastic Finance: An Introduction in Discrete Time, 2nd Edition, Walter de Gruyter, 2004.

[14] H. Föllmer and W. Schachermayer; Asymptotic Arbitrage and Large Deviations, *Math. Financ. Econ.*, Vol. 1, 213-249, 2007.

[15] A. Ganesh, N. O'Connell and D. Wischik; Big Queues, *Lecture Notes in Mathematics*, Springer, 2004.

[16] L. Györfi and I. Vajda; Growth Optimal Portfolio Selection Stratefies with Transaction Costs, Inc: Algorithmic Learning Theory, ed. Y. Freund, L. Györfi, G. Turán, Th. Zeugmann, *Lecture Notes in Computer Science*, Vol. 5254, 108-122, Springer, 2010.

[17] T.E. Harris; The Theory of Branching Processes, Springer, 1963.

[18] P. Hall and C.C. Heyde; Martingale Limit Theory and its Application, Academic Press, 1980.

[19] O. Hernández-Lerma and J.B. Lassere; Discrete-Time Markov Control Processes: Basic Optimality Criteria, Springer, 1996.

[20] Y.M. Kabanov and D.O. Kramkov; Large Financial Markets: Asymptotic Arbitrage and Contiguity, *Theory Probab. Appl.*, Vol. 39, (1994), No. 1, 182-187 (1995).

[21] Y.M. Kabanov and D.O. Kramkov; Asymptotic Arbitrage in Large Financial Markets, *Finance Stochast.*, Vol. 2, 143-172, 1998.

[22] I. Karatzas and C. Kardaras; The Numéraire Portfolio in Semimartingale Financial Models, *Finance Stoch.*, Vol. 11, 447-493, 2007.

[23] G. Kim and H.T. David; Large Deviations of Functions of Markovian Transitions and Mathematical Programming Duality, *Ann. Probab.*, Vol. 7, 874-881, 1979.

[24] F.C. Klebaner; Introduction to Stochastic Calculus with Applications, Second Edition, Imperial College Press, 2005.

[25] I. Klein; A Fundamental Theorem of Asset Pricing for Large Financial Markets, *Math. Finance*, Vol. 10 (2000), No. 4, 443-458.

[26] I. Klein; Free Lunch for Large Financial Markets with Continuous Price Processes, *Ann. Appl. Probab.*, Vol. 13 (2003), No. 4, 1494-1503.

[27] I. Klein and W. Schachermayer; Asymptotic Arbitrage in Non-complete Large Financial Markets, *Theory Probab. Appl.*, Vol. 41 (1996), No. 4, 780-788 (1997).

[28] I. Klein and W. Schachermayer; A Quantitative and a Dual Version of the Halmos-Savage Theorem with Applications to Mathematical Finance, *Ann. Probab.*, Vol. 24 (1996), No. 2, 867-881.

[29] I. Kontoyiannis and S. Meyn; Spectral Theory and Limit Theorems for Geometrically Ergodic Markov Processes, *Ann. Appl. Probab.*, Vol. 13, 304-362, 2003.

[30] I. Kontoyiannis and S. Meyn; Large Deviations Asymptotics and the Spectral Theory of Multiplicative Regular Processes, *Electronic. Journ. Probab.*, Vol. 10, 61-123, 2005.

[31] Y. Li; A Martingale Inequality and Large Deviations, *Statistics* & *Probability Letters*, Vol. 62, 317-321, 2003.

[32] Q. Liu and F. Watbled; Exponential Inequalities for Martingales and Asymptotic Properties of the Free Energy of Directed Polymers in a Random Environment, *Stoch. Proc. Appl.*, Vol. 119, 3101-3132, 2009.

[33] R.C. Merton; Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Model, *Rev. Econom. Statist.*, Vol. 51, 247-257, 1969.

[34] S.P. Meyn; Large Deviation Asymptotics and Control Variates for Simulating Large Functions, *Ann. Appl. Probab.*, Vol. 16, 310-339, 2006.

[35] M.L.D. Mbele Bidima and M. Rásonyi; On Long-term Arbitrage Opportunities in Markovian Models of Financial Markets, *Ann. Operations Research*, submitted on May 15, 2010.

[36] S. Meyn and R. Tweedie; Markov Chains and Stochastic Stability, Second Edition, Cambridge University Press, 2009.

[37] P. Ney and E. Nummelin; Markov Additive Processes I. Eigenvalue Properties and Limit Theorems, *Ann. Probab.*, Vol. 15, 561-592, 1987.

[38] P. Ney and E. Nummelin; Markov Additive Processes II. Large Deviations, *Ann. Probab.*, Vol. 15, 593-609, 1987.

[39] J. R. Norris; Markov Chains, *Cambridge Series in Statistical and Probabilistic Mathematics*, Cambridge University Press, 1999.

[40] H. Pham; A Large Deviations Approach to Optimal Long-Term Investment, *Finance Stochast.*, Vol. 7, 169-195, 2003.

[41] M. Rásonyi; Equivalent Martingale Measures for Large Financial Markets in Discrete Time, *Math. Methods Oper. Res.*, Vol. 58 (2003), No. 3, 401-415.

[42] M. Rásonyi; Arbitrage Pricing Theory and Risk-Neutral Measures, *Decis. Econ. Finance*, Vol. 27 (2004), No. 2, 109-123.

[43] M. Rásonyi; A Note on Arbitrage in Term Structure, *Decis. Econ. Finance*, Vol. 31 (2008), No. 1, 73-79.

[44] M. Rásonyi and L. Stettner; On Utility Maximization in Discrete-Time Financial Market Models, *Ann. Appl. Probab.*, Vol. 15, 1367-1395, 2005.

[45] D. Revuz and M. Yor; Continuous Martingales and Brownian Motion, Third Edition, Springer, 1999.

[46] R. T. Rockafellar; Convex analysis. Princeton University Press, 1970.

[47] L.C.G. Rogers and D. Willians; Diffusions, Markov Processes and Martingales, Vol. 1 and 2, Cambridge University Press, 2006.

[48] D.B. Rokhlin; Asymptotic Arbitrage and Numériare Portfolio in Large Financial Markets, *Finance Stoch.*, Vol. 12, 173-194, 2008.

[49] S. Ross; A First Course in Probability, Eighth Edition, Pearson Prentice Hall, 2009.

[50] S.E. Shreve; Stochastic Calculus for Finance II: Continuous-Time Models, Springer Finance, 2004.