Transaction Tax in a General Equilibrium Model

by

Osian Dafydd Shelley

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy in Mathematics and Statistics

Mathematics and Statistics

May 2023

THE UNIVERSITY OF WARWICK

Contents

Acknowledgments

First and foremost, I would like to thank my supervisors, Dr Martin Herdegen and Dr Gechun Liang, for their expert guidance and diligent proofreading throughout my studies. I am deeply indebted to their generosity of time which has helped me grow as a mathematician. Moreover, I am incredibly appreciative of their support which has extended beyond my studies and has grounded me during challenging times.

I have been fortunate to have made many close friends at Warwick who have inspired and supported me. In particular, I would like to thank Nicolò Paviato, Joseph Jerome, Dimitris Lygkonis, Alice Hodson, Nazem Khan, and Nikolaos Constantinou. Furthermore, I would not have been able to complete my studies without the kindness of Aline Bougard, Francesca Alberigi and Alun Rees; I am immensely grateful. I also profoundly thank Ran Shi for her patience and encouragement, which has calmed and bettered me.

Finally, I thank my family for their boundless love and support, especially my mother and stepfather, Ruth Shelley and Howard Davies, to which I dedicate this thesis.

Declarations

The work in this thesis is my own, except where otherwise stated in the text. I have not submitted the work for any other degree. The content in Chapter 5 and Sections 6.1-6.3 have appeared in preprints [41, 40].

Abstract

In this thesis, we study the effects of a quadratic transaction tax levied against agents in a continuous time, risk-sharing equilibrium model where agents have heterogeneous beliefs about the dynamics of the traded risky asset. The goal of each agent is to choose a trading strategy according to a mean-variance criterion, for which an optimal strategy exists in closed form as the solution to an FBSDE. This tractable setup allows us to analyse the utility loss incurred from taxation, which will be used as a measure in order to determine whether the transaction tax is beneficial.

When agents have homogeneous beliefs about the risky asset, we will show that although the agents cannot benefit en masse, the less risk-averse agents may benefit. Furthermore, when agents have heterogeneous beliefs about the risky asset, we will show that a small transaction tax can benefit the agents (from the planner's perspective) if the their beliefs are sufficiently different.

We also consider theory relating to the vague convergence of real-valued measures. In particular, we comprehensively describe the relationship between the vague convergence of real-valued measures and the pointwise convergence of their distribution functions at continuity points. Using this theory, we extend a classical continuity theorem to the case of real-valued measures and motivate a novel stochastic control problem related to the transaction tax model.

Chapter 1

Introduction

John Maynard Keynes viewed financial transaction taxes as a reasonable measure to curb 'the predominance of speculation' in financial markets [47]. Similar claims about transaction taxes' curative effects have gained relevance during periods of economic turmoil and have consequently been the source of significant economic interest over the past century. Today, financial transaction taxes are most heavily associated with James Tobin, who advocated for their implementation in his 1978 presidential address after the collapse of the Bretton Woods system [70] (many now call such taxes *Tobin taxes*). Since then, the 1987 and 2008 financial crises have motivated economists such as Summers and Summers, Stiglitz, and Krugman to advocate for transaction taxes to reduce speculative trading and to redirect revenue to more socially beneficial investments [68, 67, 50]. More recently, economists such as Jeffery Sachs argued in favour of financial transaction taxes to aid economic recovery after the fallout from COVID. Unfortunately, most public economists base their claims on ad-hoc heuristics, partly due to a lack of general equilibrium models from which one can gather impartial quantitative and qualitative guidance.

In this thesis, we will study the normative effects of a financial transaction tax in a general equilibrium model that allows agents to hold heterogeneous beliefs about the risky asset, which they trade to hedge against fluctuations in their respective endowment streams. As a consequence, the financial market plays a dual role. Firstly, as we will assume that agents have heterogeneous risk aversions, the market admits trading motivated by the need to transfer risk to those more willing to bear them. Secondly, the market enables participants to engage in speculative trading, or gambling, according to their beliefs about the risky asset. This belief discrepancy will motivate trading that may be at odds with the agents' hedging needs, and the interplay between these motivations will characterise whether or not a transaction tax is beneficial.

To study the effects of a transaction tax, we propose a tractable continuous time model where agents hold (local) mean-variance preferences and are penalised with a quadratic tax. It is an extension of the setup introduced by Bouchard et al. [17], studied separately under various guises in the works of Muhle-Karbe et al., Herdegen et al. and Gonon et al. [58, 42, 35]. Importantly, our extension models the heterogeneity in agents' beliefs so that a unique equilibrium still exists, and we achieve explicit formulas for the optimal portfolios.

We measure the effects of the tax by the difference in utility of each agent in the market with and without tax. Whether or not an agent deems the tax beneficial will depend on the sign of this *utility loss*. The planner will consider the tax beneficial if it benefits the agents en masse. In particular, by specifying the beliefs explicitly, we will show that if the planner holds the average belief of the agents, a small transaction tax is beneficial when the beliefs are sufficiently different. In this sense, as Keynes proposed, the transaction tax is a reasonable measure to curb speculation.

1.1 Relevant Literature

Market frictions compound the difficulties in determining equilibrium prices. Thus, it is no surprise that the foundational work on equilibrium pricing relies on frictionless markets; see the works of Sharpe, Black and Scholes, and Cox et al. [66, 13, 23]. When frictions are present, results often rely on numerical methods or particular simplifying assumptions. For example, the works of Heaton and Lucas, Buss and Dumas, and Buss et al. all use a numerical approximation of equilibrium dynamics [39, 20, 19]. In comparison, Lo et al. and Vayanos et al. obtain explicit formulas by focusing on continuous time models with deterministic asset prices [74, 53], while Gârleanu and Pedersen assume there is a single rational agent among noise traders [37]. In our case, the assumption of a quadratic tax enables us to get explicit solutions to our optimisation problem, although we will argue that the resulting phenomena generalise.

Heterogeneous beliefs about the fundamental values of a market are a crucial motive for trading but are an additional hurdle in the presence of transaction costs. When the market is frictionless, there is extensive literature on asset pricing arising from subjective beliefs; see the survey by Scheinkman and Xiong [63]. The work of Muhle-Karbe et al. [57] seems to be the most relevant to our setup, in which they consider the equilibrium price of a traded asset amongst N agents with heterogeneous beliefs in a continuous time model. However, they do not allow agents to have heterogeneous risk aversions, and the paper mainly focuses on the resulting illiquidity. Moreover, although a quadratic transaction cost is present, they do not view it explicitly as a tax nor investigate the beneficial implications of such a penalty.

In contrast, recent work by Davilá [29] explicitly studies the welfare-improving effect of a transaction tax in a one-period equilibrium model where investors can trade speculatively . He models heterogeneous beliefs as disagreements about the parameters of a normally distributed dividend attributed to the traded risky asset. Assuming that a central planner rebates a lump sum to the market participants, he shows that a beneficial proportional transaction tax exists. Our approach differs significantly as we consider a continuous time model where we describe agents' beliefs by idiosyncratic measure changes that alter the risky asset's drift components. In particular, our approach is more in line with the framework seen in the work of Bouchard et al. [17], as we will extend their existence results appropriately.

1.2 Overview

Chapter 2 will introduce the specifics of the model and solve the associated stochastic control problem by means of a coupled FBSDE. Under the assumption of market clearing, we then derive the equilibrium return of the risky asset and define the *utility loss*. This results in a description of when the transaction tax may be beneficial.

Chapter 3 explores the dynamics of the model when agents have homogeneous beliefs. By specifying how the agents' endowments are affected by market shocks, we use custom functions in *Python* and *Sage* to compute and approximate the agents' utility losses when the transaction tax is small. These expressions show that less risk-averse agents can benefit from more favourable returns in the market with tax, explained by the dynamics of the equilibrium portfolios.

Chapter 4 builds upon this setup by letting agents have specific heterogeneous beliefs about the risky asset. We again calculate the utility losses, which results in a concise condition for when a small transaction tax is beneficial. Moreover, we will consider the idiosyncratic perspectives of the agents, resulting in a comprehensive description of the transaction tax's effects.

In Chapter 5, we change tack significantly and highlight the importance of the vague topology on the space of real-valued Radon measures. In particular, we describe this topology's relationship to the notion of weak convergence, seemingly absent from the literature. Moreover, we give novel conditions to describe when vague convergence of measures is equivalent to the convergence of their distribution functions at continuity points.

Chapter 6 uses the abstract theory of Chapter 5 in order to derive a continuity theorem describing the relationship between the convergence of Laplace transforms of real-valued measures and their distribution functions. The continuity theorem is then applied to derive a novel Tauberian condition for an extended version of Karamata's theorem. This motivates a novel stochastic control problem that may relate the model we introduce in Chapter 3 to a long-run average model used in the work of Gonon et al. [35].

Remark 1.2.1. Several lengthy calculations and approximations in Chapters 3 and 4 were derived using custom functions written in the computer language *Sage* [69]. As such, *Jupyter notebooks* containing the calculations are referenced and can be found in the GitHub repository

https://github.com/odshelley/thesis.

For more information on these calculations, please see Appendix D.1.

Chapter 2

The Model

2.1 Model specifications

2.1.1 Preliminaries

Throughout, we fix an interval $\mathscr{T} = [0, T]$ for $T \in (0, \infty)$ ('finite time horizon') or $\mathscr{T} = [0, \infty)$ for $T = \infty$ ('infinite time horizon'). Furthermore, we consider a stochastic basis $(\Omega, \mathcal{F}_\infty, \mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ where $\mathbb{F}$ is the filtration generated by a standard one dimensional Brownian motion $(W_t)_{t\in\mathscr{T}}$ and $\mathcal{F}_\infty := \bigvee_{t\geq 0} \mathcal{F}_t$.

For technical reasons we will enlarge this filtration with respect to the *natural conditions* as opposed to the usual conditions; see Appendix A.1 for a full discussion. For any probability measure $\mathbb{Q}$ on $(\Omega, \mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0})$, we denote its restriction to $\mathcal{F}_t$ by $\mathbb{Q}_t$. We say that a probability measure $\mathbb{Q}$ on $\mathcal{F}_\infty$ is *locally absolutely continuous* with respect to $\mathbb{P}$ if $\mathbb{Q}_t$ is absolutely continuous with respect to $\mathbb{P}_t$, for all $t \geq 0$. In this case we write $\mathbb{Q} \ll_{\text{loc}} \mathbb{P}$. If $\mathbb{Q} \ll_{\text{loc}} \mathbb{P}$ and $\mathbb{P} \ll_{\text{loc}} \mathbb{Q}$ we say that $\mathbb{P}$ and $\mathbb{Q}$ are *locally equivalent* and write $\mathbb{Q} \sim_{\text{loc}} \mathbb{P}$.

Define the space of (locally) equivalent probability measures with respect to $\mathbb{P}$ by

$$\mathcal{P} := \{\mathbb{Q} \in \mathcal{M}_1^+(\Omega, \mathcal{F}_\infty) : \mathbb{Q} \sim_{\text{loc}} \mathbb{P}\},$$

where $\mathcal{M}_1^+(\Omega, \mathcal{F}_\infty)$ is the space of probability measures on $(\Omega, \mathcal{F}_\infty)$. Then let the *natural filtration* of W (defined according to Definition A.1.4) be denoted by $\mathbb{F}^{\mathcal{P}}$. Henceforth, the underlying filtered probability space will be assumed to be $(\Omega, \mathcal{F}_\infty^{\mathcal{P}}, \mathbb{F}^{\mathcal{P}}, \mathcal{P})$.

In order to discuss models with a finite and infinite time horizon simultaneously, for any $\delta \geq 0$ and $p \geq 1$ we define $\mathscr{L}^p_\delta(\mathbb{R}^n)$ to be the space of all $\mathbb{R}^n$-valued progressively-measurable processes $(X_t)_{t\in\mathscr{T}}$ such that

$$\mathbb{E}\left[\int_0^T e^{-\delta t} \left\|X_t\right\|^p \, \mathrm{d}t\right] < \infty,$$

where $\|\cdot\|$ is any norm on $\mathbb{R}^n$. Similarly, an $\mathbb{R}^n$-valued local martingale $(M_t)_{t\in\mathscr{T}}$ belongs to $\mathscr{M}^p_\delta$, $p \geq 1$, if

$$\mathbb{E}\left[\left\|\int_0^T e^{-2\delta s} d[M]_s\right\|^{p/2}\right] < \infty.$$

Here, $\|\cdot\|$ denotes any norm on $\mathbb{R}^n$.

2.1.2 Financial Market

We study a financial market which consists of an exogenously given riskless asset, normalised to one, and a risky asset $(S_t)_{t\in\mathscr{T}}$ with dynamics

$$\mathrm{d}S_t = \mu_t \, \mathrm{d}t + \sigma \, \mathrm{d}W_t. \tag{2.1.1}$$

Initially both the instantaneous returns process $(\mu_t)_{t\in\mathscr{T}} \in \mathscr{L}^4_\delta(\mathbb{R})$ and constant volatility $\sigma \in \mathbb{R}\backslash\{0\}$ are exogenous. We make these assumptions on the volatility for maximal tractability, which becomes crucial in Chapters 3 and 4, as it allows for analytical results; see [42] for an equilibrium model where the volatility is a free parameter. In contrast, we will later endogenise the equilibrium return by matching market participants' demand to a fixed net supply of the risky asset. For simplicity, this will always be a zero net supply.

Remark 2.1.1. One can easily generalise this set-up to where there are d risky assets if we additionally suppose that the infinitesimal covariance matrix $\Sigma := \sigma^\intercal \sigma \in \mathbb{R}^{d\times d}$ is positive definite. We use the $d = 1$ case to avoid cumbersome matrices in our exposition.

2.1.3 Agents

A finite number of agents indexed by $n = 1, \dots, N$, trade in the market to hedge against fluctuations of their random endowment streams $(Y_t^n)_{t \in \mathscr{T}}$. These endowments have dynamics,

$$\mathrm{d}Y_t^n = \zeta_t^n \sigma \, \mathrm{d}W_t, \quad n = 1, \dots, N, \tag{2.1.2}$$

where $\zeta^n \in \mathscr{L}_\delta^4(\mathbb{R})$. One could include a finite variation drift (modelling an absolutely continuous cash flow) or an additional orthogonal component (modelling unhedgeable shocks) to the dynamics of Y^n. However, this would not change the optimiser in the following linear-quadratic goal functional (2.1.7), so we content ourselves by focusing on the current most parsimonious specification.

In addition to the agents' hedging needs, we assume that the agents' trades are motivated by a discrepancy in beliefs about the risky asset. Allowing for idiosyncratic beliefs permits agents to engage in betting or gambling, a form of *non-fundamental trading*. Davilá coined this term in [29] in opposition to trading due to hedging, which he viewed as fundamental.

We model this phenomenon in a similar fashion to Kogan et al. [49], by assuming that agent $n = 1, \dots, N$, views the risky asset under the lens of a (locally) equivalent probability measure that shifts the equilibrium return process according to an application of Girsanov's Theorem (A.1.13). In particular, we assume as in [57, 29] that agents *agree to disagree* in the sense of Aumann [7], such that they do not learn from each other nor the price.

To make this precise, we specify which probability measures we deem admissible. So that we only allow shifts to the returns process that do not blow up the linear quadratic goal functional (2.1.7), we define the set of *pre-admissible beliefs* $\mathcal{B}$ to be the space of all predictable and locally bounded processes $\varepsilon \in \mathscr{L}_\delta^4(\mathbb{R})$ such that

$$\mathcal{E}\left(\varepsilon \bullet W\right) = \exp\left(\int_0^\cdot \varepsilon_s \, \mathrm{d}W_s - \frac{1}{2}\int_0^\cdot \varepsilon_s^2 \, \mathrm{d}s\right) \in \mathscr{M}_\delta^2.$$

According to Theorem A.1.13, each $\varepsilon \in \mathcal{B}$ defines a unique probability measure

$\mathbb{Q} \in \mathcal{P}$ via

$$\mathsf{Z}(\varepsilon)_t := \frac{\mathrm{d}\mathbb{Q}_t}{\mathrm{d}\mathbb{P}_t} := \mathcal{E}\left(\varepsilon \bullet W\right)_t.$$

Let $\mathsf{P} : \mathcal{B} \to \mathcal{P}$ be the mapping that takes each $\varepsilon \in \mathcal{B}$ to its aforementioned unique probability measure.

Definition 2.1.2. We define the set of *admissible beliefs* to be the set of tuples

$$\{(\varepsilon, \mathsf{P}(\varepsilon)) : \varepsilon \in \mathcal{B}\} = \operatorname{graph}(\mathsf{P}).$$

We call both the tuple $(\varepsilon, \mathsf{P}(\varepsilon)) \in \operatorname{graph}(\mathsf{P})$ and ε itself a *belief.*

We now assert that agent $n = 1, \ldots, N$ has a unique belief given by

$$(\varepsilon^n, \mathsf{P}^n) := (\varepsilon^n, \mathsf{P}(\varepsilon^n)) \in \operatorname{graph}(\mathsf{P}).$$

The following proposition describes how agent $n = 1, \ldots, N$ views the risky asset according to their belief.

Proposition 2.1.3. *For any belief* $(\varepsilon, \mathsf{P}(\varepsilon)) \in \operatorname{graph}(\mathsf{P})$*, the risky asset* S *has dynamics*

$$\mathrm{d}S_t = \mu_t^\varepsilon \ \mathrm{d}t + \sigma \ \mathrm{d}W_t^\varepsilon, \tag{2.1.3}$$

where W^ε *is a standard* $\mathsf{P}(\varepsilon)$*-Brownian motion and* $\mu^\varepsilon := \mu + \sigma\varepsilon$*. In particular, agent* $n = 1, \ldots, N$ *views the risky asset as having dynamics*

$$\mathrm{d}S_t = \mu_t^n \ \mathrm{d}t + \sigma \ \mathrm{d}W_t^n, \tag{2.1.4}$$

where $W^n := W^{\varepsilon^n}$ *and* $\mu^n := \mu^{\varepsilon^n}$*.*

Proof. This is a direct consequence of Theorem A.1.13. $\square$

2.1.4 Goal functional

Agents' trading strategies are described by the number of shares $\varphi_t \in \mathscr{L}_\delta^4(\mathbb{R})$ held in the risky asset at time $t \in \mathscr{T}$. When no transaction costs are present, agents choose their strategies in order to satisfy a continuous time analogue of a mean variance criterion with discounting. Specifically, agent $n = 1, \ldots, N$

solves

$$\begin{aligned}
&\underset{\varphi\in\mathscr{L}^4_\delta}{\operatorname{argmax}}\ \mathbb{E}^{\mathsf{P}^n}\left[\int_0^T e^{-\delta t}\left\{\varphi_t\ \mathrm{d}S_t+\mathrm{d}Y^n_t-\frac{\gamma_n}{2}\ \mathrm{d}\left\langle\int_0^\cdot\varphi_s\ \mathrm{d}S_s+Y^n_\cdot\right\rangle_t\right\}\right]\\
&=\underset{\varphi\in\mathscr{L}^4_\delta}{\operatorname{argmax}}\ \mathbb{E}^{\mathsf{P}^n}\left[\int_0^T e^{-\delta t}\left\{\varphi_t\left(\mu^n_t\ \mathrm{d}t+\sigma\ \mathrm{d}W^n_t\right)-\frac{\gamma_n}{2}\sigma^2\left(\varphi_t+\zeta^n_t\right)^2\mathrm{d}t\right\}\right]\\
&=\underset{\varphi\in\mathscr{L}^4_\delta}{\operatorname{argmax}}\ \mathbb{E}^{\mathsf{P}^n}\left[\int_0^T e^{-\delta t}\left\{\varphi_t\mu^n_t-\frac{\gamma_n}{2}\sigma^2\left(\varphi_t+\zeta^n_t\right)^2\right\}\mathrm{d}t\right]. \qquad (2.1.5)
\end{aligned}$$

Remark 2.1.4. The process $-\zeta^n_t$ can be interpreted as agent n's target position in the risky asset. Hence, (2.1.5) asserts that agents trade off expected returns against the tracking error relative to this target.

Here, γ_n and δ are positive constants representing agent n's risk aversion and (common) discount rate, respectively. Without loss of generality we will always assume that

$$\gamma_N=\max\{\gamma_1,\ldots,\gamma_N\}.$$

A positive discount rate allows us to postpone the planning horizon indefinitely to obtain stationary infinite-horizon solutions. As is argued in Remark 2.1.8, we assume that all strategies φ lie in $\mathscr{L}^4_\delta$ to ensure that the problem is well posed. It is important to note that the expectation in (2.1.5) is taken under the probability measure P^n as opposed to $\mathbb{P}$, indicating that agents optimise according to their own belief.

To extend this set up to account for a transaction tax, we generalise the approach taken in [42, 35, 17]. To do so, we need to be more specific about the set of admissible portfolios. Note that we denote the set of $\mathbb{R}$-valued absolutely continuous processes on $\bar{\Omega}:=\Omega\times\mathscr{T}$ by $\mathrm{AC}(\bar{\Omega},\mathbb{R})$.

Definition 2.1.5. We define the set of *admissible portfolios* to be

$$\mathcal{A}:=\left\{\varphi\in\mathscr{L}^4_\delta(\mathbb{R})\cap\mathrm{AC}(\bar{\Omega},\mathbb{R}):\frac{\mathrm{d}}{\mathrm{d}t}\varphi\in\mathscr{L}^4_\delta(\mathbb{R})\text{ and }\varphi_{0_-}=0\right\}. \qquad (2.1.6)$$

Remark 2.1.6. Choosing admissible portfolios to be absolutely continuous is a popular choice in optimal execution literature [22, 2] as it allows us to incorporate penalties on the order flow $\frac{\mathrm{d}}{\mathrm{d}t}\varphi=\dot{\varphi}$. We require that $\dot{\varphi}$ belongs to

$\mathscr{L}^4_\delta$ to avoid infinite transaction costs, but requiring the initial stock position to be zero is a choice made for simplicity.

As in [17, 42, 57, 35], we model our transaction cost by a constant $\lambda > 0$ levied on the square of each agent's order flow. While we make this quadratic specification for maximal tractability, the numerical results in [35] suggest that the qualitative and quantitative characteristics of equilibrium asset prices are robust across different convex functions of the trading rate, as was suspected in [56].

Definition 2.1.7. We define the goal functional of an agent $n = 1, \ldots, N$ under the belief $(\varepsilon, \mathsf{P}(\varepsilon)) \in \operatorname{graph}(\mathsf{P})$ to be the mapping

$$K^\varepsilon_n : \mathcal{A} \times \mathscr{L}^4_\delta \times \mathbb{R}_+ \times \mathbb{R}_+ \times (0, \infty] \to \mathbb{R}$$

where

$$\begin{aligned} &K^\varepsilon_n(\varphi, \mu, \lambda, \delta, T) \\ &:= \mathbb{E}^{\mathsf{P}(\varepsilon)}\left[\int_0^T e^{-\delta t}\left\{\varphi_t \mu^\varepsilon_t - \frac{\gamma_n \sigma^2}{2}\left(\varphi_t + \zeta^n_t\right)^2 - \lambda\left(\dot{\varphi}_t\right)^2\right\}\, \mathrm{d}t\right]. \end{aligned} \tag{2.1.7}$$

For fixed returns process $\mu \in \mathscr{L}^4_\delta$, tax levy $\lambda \in \operatorname{diag}(\mathbb{R}^d_+)$, discount factor $\delta \geq 0$, and time horizon $T \in (0, \infty]$, the optimisation problem of agent $n = 1, \ldots, N$ is

$$\underset{\varphi \in \mathcal{A}}{\operatorname{argmax}}\ K^{\varepsilon^n}_n(\varphi, \mu, \lambda, \delta, T). \tag{2.1.8}$$

When any of the parameters are understood they may be omitted from the notation.

Remark 2.1.8. For (2.1.8) to be well posed, we would like all relevant processes to be appropriately integrable with respect to all admissible beliefs $\mathsf{P}(\varepsilon)$. This is enforced by our choice of admissible portfolios (2.1.6). Indeed, we may use the fact that $\zeta, \mu, \varphi, \dot{\varphi} \in \mathscr{L}^4_\delta$ and $\mathsf{Z}(\varepsilon) \in \mathscr{L}^2_\delta$ along with Hölder's inequality to deduce that

$$\begin{aligned} &\mathbb{E}^{\mathsf{P}(\varepsilon)}\left[\int_0^T e^{-\delta t}\left|\varphi_t \mu^\varepsilon_t - \frac{\gamma_n \sigma^2}{2}\left(\varphi_t + \zeta^n_t\right)^2 - \lambda\left(\dot{\varphi}_t\right)^2\right| \mathrm{d}t\right] \\ &= \mathbb{E}\left[\int_0^T e^{-\delta t}\left|\varphi_t \mu^\varepsilon_t - \frac{\gamma_n \sigma^2}{2}\left(\varphi_t + \zeta^n_t\right)^2 - \lambda\left(\dot{\varphi}_t\right)^2\right| \mathsf{Z}(\varepsilon)_t\, \mathrm{d}t\right] \end{aligned}$$

$$
\begin{aligned}
&\leq \mathbb{E}\left[\int_0^T e^{-\delta t}\left|\varphi_t \mu_t^{\varepsilon} - \frac{\gamma_n \sigma^2}{2}\left(\varphi_t + \zeta_t^n\right)^2 - \lambda\left(\dot{\varphi}_t\right)^2\right|^2 \mathrm{d}t\right]^{\frac{1}{2}} \\
&\qquad\qquad \times \mathbb{E}\left[\int_0^T e^{-\delta t}\left|\mathsf{Z}(\varepsilon)_t\right|^2 \mathrm{d}t\right]^{\frac{1}{2}} \\
&< \infty.
\end{aligned}
$$

As is often found in the literature [29, 39, 42, 17, 35, 57], the transaction cost only depends on each agent's individual trading rate. This rules out the cost modelling a temporary price impact since in this case one would expect the trades to affect the price of S and for each agent's trades to affect the others' execution prices. An example of a partial equilibrium model which allows agents to interact through their common price impact can be found in [16]. Fortunately, our current setup doesn't cause a conflict since we assume that the cost is an exogenous tax.

2.2 Individual and equilibrium optimisers

Following the approach of Muhle-Karbe et al. [57], we asses the effect of the transaction tax on trading in equilibrium by first showing the existence and uniqueness of a solution to the individual optimisation problem (2.1.8). When the agents are trading in equilibrium, these optimal portfolios allow us to find the *equilibrium return* process and *equilibrium portfolios* in turn.

Definition 2.2.1. A process $\nu \in \mathscr{L}_{\delta}^{4}(\mathbb{R})$ is an *equilibrium return* if there exists portfolios $\varphi^n \in \mathcal{A}$ for agents $n = 1, \ldots, N$, such that

(Market clearing) The total demand $\sum_{n=1}^{N} \varphi^n$ matches the zero net supply of the risky asset S at all times;

(Individual optimality) The portfolio φ^n solves agent n's control problem (2.1.8).

In this case φ^n is called the *equilibrium portfolio* for agent n.

2.2.1 Frictionless baseline

Without the transaction tax, agent $n = 1, \dots, N$ must find a portfolio φ that solves

$$\mathbb{E}^{\mathsf{P}^n}\left[\int_0^T e^{-\delta t}\left\{\varphi_t\mu_t^n - \frac{\gamma_n\sigma^2}{2}\left(\varphi_t + \zeta_t^n\right)^2\right\}\,\mathrm{d}t\right] \to \max! \tag{2.2.1}$$

By pointwise optimisation one can readily see that the individual optimiser to (2.2.1) is given by

$$\varphi_t^n := \frac{\mu_t^n}{\sigma^2\gamma_n} - \zeta_t^n. \tag{2.2.2}$$

This is the classical (myopic) Merton portfolio along with the tracking error. In particular, note that it does not depend on the discount factor, nor the time horizon.

Assuming that the portfolios clear the market, it is immediate from (2.2.2) that the *frictionless* equilibrium return is given by

$$\mu_t^0 := \frac{\sum_{k=1}^N\left(\sigma^2\zeta_t^k - \sigma\varepsilon_t^k/\gamma_k\right)}{\sum_{k=1}^N 1/\gamma_k}. \tag{2.2.3}$$

Furthermore, by substituting (2.2.3) into (2.2.2), we see that the *frictionless* equilibrium portfolio for agents $n = 1, \dots, N$ is

$$\bar{\varphi}_t^n := \frac{\sum_{k=1}^N\left(\zeta_t^k - \frac{\varepsilon_t^k}{\sigma\gamma_k}\right)}{\sum_{k=1}^N \gamma_n/\gamma_k} + \frac{\varepsilon_t^n}{\gamma_n\sigma} - \zeta_t^n. \tag{2.2.4}$$

2.2.2 Quadratic costs

We now solve (2.1.8) in the presence of the quadratic tax. Unlike its frictionless counterpart, the goal functional is no longer myopic since current positions affect future choices. As such, the optimal strategies will also depend on the discount factor.

Note that the optimal positions of the agents evolve forward from their initial allocations. In contrast, one must determine the initial optimal trading rates as part of the solution. This naturally leads to a characterisation of the optimal portfolio by a coupled system of FBSDEs. We make this precise by the following lemma, altered from [17, Lemma 4.1]. We include the proof for completeness.

Lemma 2.2.2. *Let $\varphi^n = \frac{\mu^n}{\sigma^2\gamma_n} - \zeta^n$ be the frictionless optimiser from (2.2.2). Then the frictional optimisation problem (2.1.8) for agent $n = 1, \ldots, N$ has a unique solution, characterised by the FBSDE*

$$\mathrm{d}\varphi_t^{\lambda,n} = \dot{\varphi}^{\lambda,n}\,\mathrm{d}t, \quad \varphi_0^{\lambda,n} = 0, \tag{2.2.5}$$

$$\mathrm{d}\dot{\varphi}_t^{\lambda,n} = Z_t^n \mathrm{d}W_t^n + \frac{\gamma_n \sigma^2}{2\lambda}(\varphi_t^{\lambda,n} - \varphi_t{}^n)\,\mathrm{d}t + \delta\dot{\varphi}_t^{\lambda,n}\,\mathrm{d}t. \tag{2.2.6}$$

Here, the processes $(\varphi_t^{\lambda,n})_{t\in\mathscr{T}}$, $(\dot{\varphi}_t^{\lambda,n})_{t\in\mathscr{T}}$ are supposed to be in $\mathscr{L}_\delta^4$, and $Z^n \in L^2(W^n)$, and is determined as part of the solution. If $T < \infty$, the dynamics (2.2.5)-(2.2.6) are complemented by the terminal condition

$$\dot{\varphi}_T^{\lambda,n} = 0. \tag{2.2.7}$$

For $T = \infty$, agent n's individually optimal strategy $\varphi^{\lambda,n}$ has the explicit representation (B.1.4); the corresponding strategy $\dot{\varphi}^{\lambda,n}$ is given in feedback form (B.1.6).

Remark 2.2.3. Corresponding formulas to (B.1.4) and (B.1.6) exist when $T < \infty$, but they will not be used in this thesis.

Proof. For ease of exposition, let K denote the functional $K^{\varepsilon^n}(\cdot, \mu, \lambda, , \delta, T)$. The goal functional (2.1.7) is strictly convex, whence (2.1.8) has a unique solution if and only if

$$\langle K'(\varphi), \vartheta\rangle = 0, \tag{2.2.8}$$

for all absolutely continuous ϑ with $\vartheta_0 = 0$ and $\vartheta, \dot{\vartheta} \in \mathscr{L}_\delta^2$. Here, the Gâteaux derivative of K in the direction $\dot{\vartheta}$ is given by

$$\begin{aligned}
&\langle K'(\varphi), \vartheta\rangle \\
&= \lim_{\rho\downarrow 0} \frac{K(\varphi + \rho\vartheta) - K(\varphi)}{\rho} \\
&= \mathbb{E}^{\mathsf{P}^n}\left[\int_0^T e^{-\delta t}\left\{\left(\mu_t^n - \gamma_n\left(\varphi_s + \zeta_s^n\right)\sigma^2\right)\left(\int_0^t \dot{\vartheta}_s\,\mathrm{d}s\right) - 2\lambda\dot{\varphi}_t\dot{\vartheta}_t\right\}\mathrm{d}t\right].
\end{aligned}$$

By Fubini's theorem, it follows that

$$\int_0^T e^{-\delta t}\left\{\left(\mu_t^n - \gamma_n\left(\varphi_s + \zeta_s^n\right)\sigma^2\right)\left(\int_0^t \dot{\vartheta}_s \,\mathrm{d}s\right)\right\}\mathrm{d}t$$
$$= \int_0^T \left(\int_s^T e^{-\delta t}\left(\mu_t^n - \gamma_n\left(\varphi_s + \zeta_s^n\right)\sigma^2\right)\,\mathrm{d}t\right)\dot{\vartheta}_s \,\mathrm{d}s.$$

Thus, using the tower property of conditional expectation, (2.2.8) can be written as

$$\mathbb{E}^{\mathsf{P}^n}\left[\int_0^T \left(\mathbb{E}^{\mathsf{P}^n}\left[\int_t^T e^{-\delta s}\left(\mu_s^n - \gamma_n\left(\varphi_s + \zeta_s^n\right)\sigma^2\right)\mathrm{d}s\Big|\mathcal{F}_t\right] - 2\lambda e^{-\delta t}\dot{\varphi}_t\right)\dot{\vartheta}_t \,\mathrm{d}t\right]$$
$$= 0.$$

Since this needs to hold for any $\dot{\vartheta}$, (2.1.8) has a (P^n-a.s.) unique solution $\varphi^{\lambda,n}$ if and only if

$$\dot{\varphi}_t^{\lambda,n} = \frac{\gamma_n\sigma^2}{2\lambda}e^{\delta t}\mathbb{E}^{\mathsf{P}^n}\left[\int_t^T e^{-\delta s}\left(\frac{\mu_s^n}{\sigma^2\gamma_n} - \varphi_s^{\lambda,n} - \zeta_s^n\right)\mathrm{d}s\Big|\mathcal{F}_t\right] \tag{2.2.9}$$

has a (P^n-a.s.) unique solution for dt-*a.e.* $t \in \mathscr{T}$.

Let's first assume (2.2.9) has a unique solution $\dot{\varphi}^{\lambda,n}$. Then we may define the square-integrable martingale

$$M_t = \frac{\gamma_n\sigma^2}{2\lambda}\mathbb{E}^{\mathsf{P}^n}\left[\int_0^T e^{-\delta s}\left(\varphi_s^n - \varphi_s^{\lambda,n}\right)\,\mathrm{d}s\Big|\mathcal{F}_t\right], \quad t \in \mathscr{T}.$$

By an integration by parts we may rewrite (2.2.9) as

$$\mathrm{d}\dot{\varphi}_t^{\lambda,n} = e^{\delta t}\mathrm{d}M_t - \frac{\gamma_n\sigma^2}{2\lambda}\left(\varphi_t^n - \varphi_t^{\lambda,n}\right)\mathrm{d}t + \delta\dot{\varphi}_t^{\lambda,n}\,\mathrm{d}t.$$

By the martingale representation theorem, $\int_0^\cdot e^{\delta s}\,\mathrm{d}M_s$ can be expressed as as an integral $Z^n \bullet W^n$ for some $Z^n \in L^2(W^n)$. Together with $\mathrm{d}\varphi_t^{\lambda,n} = \dot{\varphi}^{\lambda,n}\,\mathrm{d}t$, this yields the FBSDE representation (2.2.5)-(2.2.6).

Conversely, assume (2.2.5)-(2.2.6) has a unique solution

$$(\varphi^{\lambda,n}, \dot{\varphi}^{\lambda,n}, Z^n) \in \mathscr{L}_\delta^4 \times \mathscr{L}_\delta^4 \times L^2(W^n),$$

and let $M^n := Z^n \bullet W^n$. For any $t \in \mathbb{R}$, an integration by parts yields

$$e^{-\delta t}\dot{\varphi}_t^{\lambda,n} = \dot{\varphi}_0^{\lambda,n} + \int_0^t e^{-\delta s} \mathrm{d}M_s^n + \frac{\gamma_n \sigma^2}{2\lambda} \int_0^t e^{-\delta s} \left(\varphi_s^{\lambda,n} - \varphi_s^n\right) \mathrm{d}s. \tag{2.2.10}$$

We claim that

$$\dot{\varphi}_0^{\lambda,n} = -\int_0^T e^{-\delta s} \mathrm{d}M_s^n - \int_0^T e^{-\delta s} \frac{\gamma_n \sigma^2}{2\lambda} \left(\varphi_s^{\lambda,n} - \varphi_s^n\right) \mathrm{d}s.$$

If $T < \infty$, this follows from (2.2.10) for $t = T$ together with the terminal condition (2.2.7). When $T = \infty$ we argue that since $\dot{\varphi}^{\lambda,n} \in \mathscr{L}_\delta^2$, there must be a sequence $\{t_k\}_{k\in\mathbb{N}}$ such that $e^{-\delta t_k}\dot{\varphi}_{t_k}^{\lambda,n}$ converges a.s. to zero. Furthermore, Proposition B.1.2, the martingale convergence theorem and $\varphi^{\lambda,n}, \varphi^n \in \mathscr{L}_\delta^4$ show that the right hand side of (2.2.10) converges (along t_k) to

$$\dot{\varphi}_0^{\lambda,n} + \int_0^T e^{-\delta s}\ \mathrm{d}M_s^n + \frac{\gamma_n \sigma^2}{2\lambda} \int_0^T e^{-\delta s} \left(\varphi_s^{\lambda,n} - \varphi_s^n\right)\ \mathrm{d}s,$$

whence

$$\dot{\varphi}_0^{\lambda,n} = -\int_0^T e^{-\delta s}\ \mathrm{d}M_s^n - \frac{\gamma_n \sigma^2}{2\lambda} \int_0^T e^{-\delta s} \left(\varphi_s^{\lambda,n} - \varphi_s^n\right)\ \mathrm{d}s. \tag{2.2.11}$$

Inserting (2.2.11) into (2.2.10), taking conditional expectations and rearranging in turn yields (2.2.9).

Finally, the FBSDE (2.2.5)-(2.2.6) has a unique solution $(\varphi^{\lambda,n}, \dot{\varphi}^{\lambda,n}, Z^{,n})$ by Theorem B.1.1. $\square$

We now assume that agents clear the market in order to deduce the equilibrium return. A crucial requirement is the solution of another system of coupled but linear FBSDE. The following result is proved in the case where $N = 2$ only. For the case where $N > 2$ we manipulate the FBSDE into a form where we expect a solution exists.

Lemma 2.2.4 (Conjecture)**.** *There exists a unique solution*

$$(\varphi^\lambda, \dot{\varphi}^\lambda, Z^\lambda) = \left((\varphi^{\lambda,1}, \dots, \varphi^{\lambda,N-1})^\intercal, (\dot{\varphi}^{\lambda,1}, \dots, \dot{\varphi}^{\lambda,N-1})^\intercal, (Z^{,1}, \dots, Z^{,N-1})^\intercal\right)$$

of the following FBSDE:

$$d\varphi_t^\lambda = \dot{\varphi}_t^\lambda \, dt, \quad \varphi_0^\lambda = 0, \tag{2.2.12}$$

$$d\dot{\varphi}_t^\lambda = Z_t^\lambda dW_t + \left(\Theta^\varphi \varphi_t^\lambda + \Theta^\zeta \zeta_t + \Theta^\varepsilon \varepsilon_t + E_t Z_t^\lambda + \delta \dot{\varphi}_t^\lambda\right) \, dt, \tag{2.2.13}$$

satisfying the terminal condition $\dot{\varphi}_T^\lambda = 0$ if $T < \infty$. Here $(Z_t^\lambda)_{t\in\mathscr{T}}$ is an $\mathbb{R}^{(N-1)\times 1}$-valued process, determined as part of the solution, and has the form

$$Z^\lambda := \left(Z^{,1}, \ldots, Z^{,N-1}\right)^\intercal, \tag{2.2.14}$$

where $Z^n \in L^2_{loc}(W)$. Moreover,

$$\begin{aligned}
\zeta_t &:= \left(\zeta_t^1, \ldots, \zeta_t^N\right)^\intercal \in \mathbb{R}^{N\times 1} \quad \forall t \in \mathscr{T}, \\
\varepsilon_t &:= \left(\varepsilon_t^1, \ldots, \varepsilon_t^{N-1}\right)^\intercal \in \mathbb{R}^{N\times 1} \quad \forall t \in \mathscr{T}, \\
\Theta^\varphi &:= \begin{pmatrix} \left(\frac{\gamma_N-\gamma_1}{N} + \gamma_1\right)\frac{\sigma^2}{2\lambda} & \cdots & \frac{\gamma_N-\gamma_{N-1}}{N}\frac{\sigma^2}{2\lambda} \\ \vdots & \ddots & \vdots \\ \frac{\gamma_N-\gamma_1}{N}\frac{\sigma^2}{2\lambda} & \cdots & \left(\frac{\gamma_N-\gamma_1}{N} + \gamma_{N-1}\right)\frac{\sigma^2}{2\lambda} \end{pmatrix} \in \mathbb{R}^{N-1\times N-1}, \\
\Theta^\zeta &:= \begin{pmatrix} -\left(\frac{\gamma_1}{N} - \gamma_1\right)\frac{\sigma^2}{2\lambda} & \cdots & -\frac{\gamma_{N-1}}{N}\frac{\sigma^2}{2\lambda} & -\frac{\gamma_n}{N}\frac{\sigma^2}{2\lambda} \\ \vdots & \ddots & \vdots & \vdots \\ -\frac{\gamma_1}{N}\frac{\sigma^2}{2\lambda} & \cdots & -\left(\frac{\gamma_{N-1}}{N} - \gamma_{N-1}\right)\frac{\sigma^2}{2\lambda} & -\frac{\gamma_n}{N}\frac{\sigma^2}{2\lambda} \end{pmatrix} \in \mathbb{R}^{N-1\times N}, \\
\Theta^\varepsilon &:= \begin{pmatrix} \left(\frac{1}{N} - 1\right)\frac{\sigma}{2\lambda} & \cdots & \frac{\sigma}{2N\lambda} & \frac{\sigma}{2N\lambda} \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\sigma}{2N\lambda} & \cdots & \left(\frac{1}{N} - 1\right)\frac{\sigma}{2\lambda} & \frac{\sigma}{2N\lambda} \end{pmatrix} \in \mathbb{R}^{N-1\times N}, \\
E_t &:= \begin{pmatrix} \frac{\varepsilon_t^1-\varepsilon_t^N}{N} - \varepsilon_t^1 & \cdots & \frac{\varepsilon_t^{N-1}-\varepsilon_t^N}{N} \\ \vdots & \ddots & \vdots \\ \frac{\varepsilon_t^1-\varepsilon_t^N}{N} & \cdots & \frac{\varepsilon_t^{N-1}-\varepsilon_t^N}{N} - \varepsilon_t^{N-1} \end{pmatrix} \in \mathbb{R}^{N-1\times N-1} \quad \forall t \in \mathscr{T}.
\end{aligned}$$

Proof. When $N = 2$ (2.2.13) becomes

$$\begin{aligned}
d\dot{\varphi}_t^{\lambda,1} = Z_t \, dW_t + \frac{(\gamma_1+\gamma_2)\sigma^2}{4\lambda}\left(\varphi_t^{\lambda,1} - \left[\frac{(\gamma_1\zeta_t^1 - \gamma_2\zeta_t^2)}{(\gamma_1+\gamma_2)} - \frac{(\varepsilon_t^1 - \varepsilon_t^2)}{\sigma(\gamma_1+\gamma_2)}\right]\right) dt \\
- \frac{1}{2}\left(\varepsilon_t^1 + \varepsilon_t^2\right) Z_t \, dt + \delta\dot{\varphi}_t^{\lambda,1} \, dt
\end{aligned}$$

$$= Z_t \, \mathrm{d}W_t + \frac{(\gamma_1+\gamma_2)\sigma^2}{4\lambda}\left(\varphi_t^{\lambda,1} - \bar{\varphi}_t^1\right) \mathrm{d}t - \frac{1}{2}\left(\varepsilon_t^1 + \varepsilon_t^2\right) Z_t \, \mathrm{d}t + \delta\dot{\varphi}_t^{\lambda,1} \, \mathrm{d}t. \tag{2.2.15}$$

By defining the belief

$$(\bar{\varepsilon}, \bar{\mathsf{P}}) := \left(\frac{\varepsilon^1+\varepsilon^2}{2}, \mathsf{P}\left(\frac{\varepsilon^1+\varepsilon^2}{2}\right)\right) \in \mathrm{graph}(\mathsf{P}) \tag{2.2.16}$$

we can re-write (2.2.15) as

$$\mathrm{d}\dot{\varphi}_t^{\lambda,1} = Z_t \, \mathrm{d}W_t^{\bar{\varepsilon}} + \frac{(\gamma_1+\gamma_2)\sigma^2}{4\lambda}\left(\varphi_t^{\lambda,1} - \bar{\varphi}_t^1\right) \mathrm{d}t + \delta\dot{\varphi}_t^{\lambda,1} \, \mathrm{d}t.$$

The result now follows from Theorem B.1.1 under the measure $\bar{\mathsf{P}}$. For the case where $N > 2$, define the exponential matrix process $V : \Omega\times\, [0,\infty) \to \mathbb{R}^{N-1\times N-1}$ by

$$\mathrm{d}V_t = -V_t E_t \, \mathrm{d}W_t, \quad V_0 = 1$$

V will act as an integrating factor that removes Z^λ from the driver of (2.2.13). Indeed, by defining $Y := V\dot{\varphi}^\lambda$ it follows that

$$\mathrm{d}\varphi_t^\lambda = V_t^{-1} Y_t \, \mathrm{d}t, \quad \varphi_0^\lambda = 0, \tag{2.2.17}$$

$$\mathrm{d}Y_t = \tilde{Z}_t \, \mathrm{d}W_t + V_t\left(\Theta^\varphi \varphi_t^\lambda + \Theta^\zeta \zeta_t + \Theta^\varepsilon \varepsilon_t\right) \mathrm{d}t + \delta Y_t \, \mathrm{d}t, \tag{2.2.18}$$

where $\tilde{Z}_t = \left(V_t Z_t^\lambda - V_t E_t Y_t\right)$. The FBSDE (2.2.17)-(2.2.18) has a linear driver independent of $\tilde{Z}$, whence we *expect* a unique solution $(Y, \tilde{Z})$ to exist. A solution to the original system (2.2.12)-(2.2.13) then follows by noting that

$$\dot{\varphi}^\lambda = V^{-1} Y \quad \text{and} \quad Z = V^{-1}\tilde{Z} + EY.$$

□

Remark 2.2.5. (i) We emphasise that the case $N > 2$ is conjectural at present, although we are confident that a unique solution does indeed exists. There is no conflict in Chapters 3 and 4 as we will only be considering the $N = 2$ case.

(ii) A representative agent with belief (2.2.16) holds the 'average' belief of the agents. This average belief crops up in the existence of a unique equilibrium price in [57, Proposition 5.1] and [59, II. B. Remark 3].

We now define a state process for each $n = 1, \dots, N$, that expresses the deviation between the optimal portfolio in the equilibrium market with and without friction. These deviations will drive the upcoming equilibrium returns process and equilibrium portfolios. Moreover, they allow for a succinct analysis of the utility losses defined in Section 2.4.

Definition 2.2.6. We define the state process $\Delta^n \in \mathscr{L}^4_\delta(\mathbb{R})$ for agent $n = 1, \dots, N$, as

$$\Delta^n_t := \bar{\varphi}^{\lambda,n}_t - \bar{\varphi}^n_t, \tag{2.2.19}$$

where $\bar{\varphi}^{\lambda,n}$ is to be characterised in Theorem 2.2.7.

We may now deduce the equilibrium return.

Theorem 2.2.7. *The unique frictional equilibrium return is given by*

$$\begin{aligned} \mu^\lambda_t &= \sum_{n=1}^{N-1} \left\{ \frac{(\gamma_n - \gamma_N)\sigma^2}{N} \bar{\varphi}^{\lambda,n}_t - \frac{2\lambda}{N} Z^n_t (\varepsilon^n_t - \varepsilon^N_t) \right\} + \sum_{n=1}^{N} \left\{ \frac{\gamma_n \sigma^2}{N} \zeta^n_t - \frac{\sigma}{N} \varepsilon^n_t \right\} \\ &= \mu^0_t + \sum_{n=1}^{N-1} \frac{(\gamma_n - \gamma_N)\sigma^2}{N} \Delta^n_t - \sum_{n=1}^{N-1} \frac{2\lambda}{N} Z^n_t \left(\varepsilon^n_t - \varepsilon^N_t \right). \end{aligned} \tag{2.2.20}$$

where Z^n and $\bar{\varphi}^{\lambda,n}$ for $n = 1, \dots, N-1$ are taken from Lemma 2.2.4. Moreover, $\bar{\varphi}^{\lambda,1}, \dots, \bar{\varphi}^{\lambda,N}$ are the individually optimal trading strategies of agents $n = 1, \dots, N-1$ where $\bar{\varphi}^{\lambda,N} = -\sum_{n=1}^{N-1} \bar{\varphi}^{\lambda,n}$.

Proof. Let $\nu \in \mathscr{L}^4_\delta$ be any equilibrium return and let $\vartheta^\lambda = (\vartheta^{\lambda,1}, \dots, \vartheta^{\lambda,N})$ be the corresponding individually optimal trading strategies. Due to market clearing we have

$$\vartheta^{\lambda,N} = -\sum_{n=1}^{N-1} \vartheta^{\lambda,n} \quad \text{and} \quad \sum_{n=1}^{N} \dot{\vartheta}^{\lambda,n} = 0.$$

Together with the FBSDE (2.2.5)-(2.2.6) it follows that

$$0 = \sum_{n=1}^{N} Z^n_t \mathrm{d}W^n_t + \sum_{n=1}^{N} \frac{\gamma_n \sigma^2}{2\lambda} \left\{ \vartheta^{\lambda,n}_t - \left(\frac{(\nu_t + \sigma \varepsilon^n_t)}{\sigma^2 \gamma_n} - \zeta^n_t \right) \right\} \mathrm{d}t + \sum_{n=1}^{N} \delta \dot{\vartheta}^{\lambda,n}_t \, \mathrm{d}t$$

$$= \sum_{n=1}^{N} Z_t^n \mathrm{d}W_t^n + \frac{1}{2\lambda}\left\{ \sum_{n=1}^{N-1} (\gamma_n - \gamma_N)\sigma^2 \vartheta_t^{\lambda,n} - \sum_{n=1}^{N} \left(\nu_t + \sigma\varepsilon_t^n - \gamma_n \sigma^2 \zeta_t^n\right) \right\} \mathrm{d}t.$$

By the definition of the agents' beliefs it follows that

$$Z_t^n \mathrm{d}W_t^n = Z_t^n \left(\mathrm{d}W_t - \varepsilon_t^n \ \mathrm{d}t\right),$$

whence

$$\begin{aligned} 0 = &\left(\sum_{n=1}^{N} Z_t^n\right) \mathrm{d}W_t \\ &+ \frac{1}{2\lambda}\left\{ \sum_{n=1}^{N-1} (\gamma_n - \gamma_N)\sigma^2 \vartheta_t^{\lambda,n} - \sum_{n=1}^{N} \left(\nu_t + (\sigma + 2\lambda Z_t^n)\,\varepsilon_t^n - \gamma_n \sigma^2 \zeta_t^n\right) \right\} \mathrm{d}t. \end{aligned}$$

Since any continuous local martingale of finite variation is constant we have $Z^N = -\sum_{n=1}^{N-1} Z^n$ and so

$$\nu_t = \sum_{n=1}^{N-1} \left\{ \frac{(\gamma_n - \gamma_N)\sigma^2}{N} \vartheta_t^{\lambda,n} - \frac{2\lambda}{N} Z_t^n (\varepsilon_t^n - \varepsilon_t^N) \right\} + \sum_{n=1}^{N} \left\{ \frac{\gamma_n \sigma^2}{N} \zeta_t^n - \frac{\sigma}{N} \varepsilon_t^n \right\}.$$

Plugging this back into agent $n = 1, \ldots, N-1$'s individual optimality condition (2.2.6), we deduce that

$$\begin{aligned} &\mathrm{d}\dot{\vartheta}_t^{\lambda,n} \\ = \ & \frac{1}{2\lambda}\Bigg\{ \gamma_n \sigma^2 \vartheta_t^{\lambda,n} - \sigma \varepsilon_t^n + \gamma_n \sigma^2 \zeta_t^n \\ &- \sum_{n=1}^{N-1} \left(\frac{(\gamma_n - \gamma_N)\sigma^2}{N} \vartheta_t^{\lambda,n} - \frac{2\lambda}{N} Z_t^n (\varepsilon_t^n - \varepsilon_t^N) \right) - \sum_{n=1}^{N} \left(\frac{\gamma_n \sigma^2}{N} \zeta_t^n - \frac{\sigma}{N} \varepsilon_t^n \right) \Bigg\} \mathrm{d}t \\ &+ Z_t^n \left(\mathrm{d}W_t - \varepsilon_t^n \ \mathrm{d}t\right) + \delta \dot{\vartheta}^{\lambda,1} \mathrm{d}t \\ = \ & \frac{\sigma^2}{2\lambda}\Bigg\{ \left[\gamma_n \vartheta_t^{\lambda,n} + \sum_{k=1}^{N-1} \frac{(\gamma_N - \gamma_k)}{N} \vartheta_t^{\lambda,k} \right] + \left[\gamma_n \zeta_t^n - \sum_{k=1}^{N} \frac{\gamma_k}{N} \zeta_t^k \right] \\ &+ \left[\sum_{k=1}^{N} \frac{1}{\sigma N} \varepsilon_t^k - \varepsilon_t^n / \sigma \right] \Bigg\} \mathrm{d}t - \left[Z_t^n \varepsilon_t^n + \sum_{n=1}^{N-1} \frac{Z_t^n}{N} (\varepsilon_t^N - \varepsilon_t^n) \right] \mathrm{d}t \\ &+ Z_t^n \ \mathrm{d}W_t + \delta \dot{\vartheta}^{\lambda,n} \mathrm{d}t. \end{aligned}$$

Hence,

$$\left(\left(\vartheta^{\lambda,1},\ldots,\vartheta^{\lambda,N-1}\right)^{\intercal},\left(\dot{\vartheta}^{\lambda,1},\ldots,\dot{\vartheta}^{\lambda,N-1}\right)^{\intercal},\left(Z^{1},\ldots,Z^{N-1}\right)^{\intercal}\right)$$

solves the FBSDE (2.2.12)-(2.2.13)and therefore coincides with its unique solution from Lemma 2.2.4. Having zero net supply shows $\vartheta_t^{\lambda,N}=\varphi_t^{\lambda,N}$ and we may assert that the equilibrium return coincides with (2.2.20). This establishes that if an equilibrium exists, it has to be of the proposed form.

We must now verify that the proposed equilibrium return and the corresponding strategies indeed form an equilibrium according to definition 2.2.1. Market clearing follows by the definition of $\varphi^{\lambda,N}$, so we must only check that $\varphi^{\lambda,n}$ is the individual optimiser for agent $n=1,\ldots,N$. This can be seen for agent $n=1,\ldots,N-1$ by substituting μ^{λ} into condition (2.2.9) and noting that it coincides with the respective equation in (2.2.13). For agent $n=N$, individual optimality follows from market clearing.

Finally, we use (2.2.4) and (2.2.19) to deduce that

$$\begin{aligned}
\mu_t^{\lambda} &= \sum_{n=1}^{N-1}\left\{\frac{(\gamma_n-\gamma_N)\sigma^2}{N}\varphi_t^{\lambda,n}-\frac{2\lambda}{N}Z_t^n(\varepsilon_t^n-\varepsilon_t^N)\right\}+\sum_{n=1}^{N}\left\{\frac{\gamma_n\sigma^2}{N}\zeta_t^n-\frac{\sigma}{N}\varepsilon_t^n\right\}\\
&=\left[\sum_{n=1}^{N}\frac{\gamma_n\sigma^2}{N}\left(\bar{\varphi}_t^n+\zeta_t^n\right)-\sum_{n=1}^{N}\frac{\sigma}{N}\varepsilon_t^n\right]+\sum_{n=1}^{N-1}\frac{(\gamma_N-\gamma_n)\sigma^2}{N}\Delta_t^n-\sum_{n=1}^{N-1}\frac{2\lambda}{N}Z_t^n\left(\varepsilon_t^n-\varepsilon_t^N\right)\\
&=\mu_t^0+\sum_{n=1}^{N-1}\frac{(\gamma_n-\gamma_N)\sigma^2}{N}\Delta_t^n-\sum_{n=1}^{N-1}\frac{2\lambda}{N}Z_t^n\left(\varepsilon_t^n-\varepsilon_t^N\right).
\end{aligned}$$

□

We end this section by noting that in the market with friction, the equilibrium portfolios can be written in a form similar to their frictionless counterparts as seen in (2.2.2).

Corollary 2.2.8. *The equilibrium portfolio of agent $n=1,\ldots,N$, may be written as*

$$\begin{aligned}
\bar{\varphi}_t^{\lambda,n} &= \frac{\mu_t^{\lambda,n}}{\sigma^2\gamma_n}-\zeta_t^n\\
&\quad+\Delta^n-\frac{1}{\gamma_n}\sum_{k=1}^{N}\frac{\gamma_k}{N}\Delta_t^k+\frac{1}{\gamma_n}\sum_{k=1}^{N}\frac{2\lambda}{\sigma^2 N}Z_t^k\varepsilon_t^k.
\end{aligned}$$

Proof. Using (2.2.20) it follows that

$$\frac{(\mu_t^\lambda + \varepsilon_t^n)}{\sigma^2 \gamma_n} - \zeta_t^n = \bar{\varphi}_t^n + \frac{1}{\gamma_n}\left\{ \sum_{k=1}^{N-1} \frac{(\gamma_k - \gamma_N)}{N} \Delta_t^k - \sum_{k=1}^{N-1} \frac{2\lambda}{\sigma^2 N} Z_t^k \left(\varepsilon_t^k - \varepsilon_t^N\right) \right\}.$$

The result now follows from Definition 2.2.6 and the market clearing condition. □

2.3 Liquidity premium

As a direct consequence of Theorem 2.2.7 it follows that we have a formula for the liquidity premium due to the tax

$$\mu_t^\lambda - \mu_t^0 = \sum_{k=1}^{N-1} \frac{(\gamma_k - \gamma_N)\sigma^2}{N} \Delta_t^k - \sum_{k=1}^{N-1} \frac{2\lambda}{N} Z_t^k \left(\varepsilon_t^k - \varepsilon_t^N\right). \tag{2.3.1}$$

Under homogeneous beliefs, the first term fully encapsulates how the transaction costs change the equilibrium return. In particular, as is shown in [17, 35], if agents have homogeneous risk aversions, we see that the frictionless equilibrium return clears the market with transaction costs. Under heterogeneous beliefs, this is no longer the case, as equation (2.3.1) shows that there is a non-trivial liquidity premium depending on the heterogeneity itself. That one sees the beliefs reflected in the equilibrium pricing is to be expected, as agents with pessimistic individual evaluations will sell to those agents who are more optimistic.

When trading is frictionless, and agents have heterogeneous beliefs, there is extensive literature on asset pricing; see the survey by Scheinkman & Xiong [63] and the numerous references therein. In a market with quadratic costs, the interplay between heterogeneous beliefs and liquidity is studied in the paper by Muhle-Karbe et al. [57]. They derive leading-order asymptotics for small transaction and holding costs. Although not the focus of this thesis, we will shed further light on the dynamics of such liquidity premia in Chapter 3, where we consider concrete examples of the agents' endowment streams and beliefs.

2.4 Utility loss

In order to assess the effects of the transaction tax on the trading of agents $n = 1, \ldots, N$, we analyse the difference between their value function in the equilibrium market with and without friction. This is made precise by the following definition.

Definition 2.4.1. Consider an admissible belief $(\varepsilon, \mathsf{P}(\varepsilon)) \in \operatorname{graph}(\mathsf{P})$.

(i) We define the *utility loss* of agent $n = 1, \ldots, N$ under belief ε as

$$U^{n,\varepsilon} := K_n^\varepsilon(\bar{\varphi}^n, \mu^0, T, \delta, 0) - K_n^\varepsilon(\bar{\varphi}^{\lambda,n}, \mu^\lambda, T, \delta, \lambda).$$

(ii) We define the *direct loss*, *return loss* and *portfolio loss* of agent $n = 1, \ldots, N$ under belief ε as

$$\begin{aligned}
U_{\mathrm{d}}^{n,\varepsilon} &:= \mathbb{E}^{\mathsf{P}(\varepsilon)}\left[\int_0^T e^{-\delta t}\lambda\left(\dot{\bar{\varphi}}_t^{\lambda,n}\right)^2 \mathrm{d}t\right], \\
U_{\mathrm{r}}^{n,\varepsilon} &:= \mathbb{E}^{\mathsf{P}(\varepsilon)}\left[\int_0^T e^{-\delta t}\bar{\varphi}_t^{\lambda,n}\left(\mu_t^0 - \mu_t^\lambda\right) \mathrm{d}t\right], \\
U_{\mathrm{p}}^{n,\varepsilon} &:= \mathbb{E}^{\mathsf{P}(\varepsilon)}\Bigg[\int_0^T e^{-\delta t}\Bigg\{\left(\bar{\varphi}_t^n - \bar{\varphi}_t^{\lambda,n}\right)\left(\mu_t^0 + \sigma\varepsilon_t\right) \\
&\qquad - \frac{\gamma_n\sigma^2}{2}\left[\left(\bar{\varphi}_t^n + \zeta_t^n\right)^2 - \left(\bar{\varphi}_t^{\lambda,n} + \zeta_t^n\right)^2\right]\Bigg\}\mathrm{d}t\Bigg],
\end{aligned}$$

respectively.

(iii) We define the *aggregate utility loss* under belief ε as

$$U^\varepsilon := \sum_{n=1}^N U^{n,\varepsilon}.$$

Remark 2.4.2. (a) Clearly, for any $(\varepsilon, \mathsf{P}(\varepsilon)) \in \operatorname{graph}(\mathsf{P})$

$$U^{n,\varepsilon} = U_{\mathrm{d}}^{n,\varepsilon} + U_{\mathrm{r}}^{n,\varepsilon} + U_{\mathrm{p}}^{n,\varepsilon}, \quad n = 1, \ldots, N.$$

(b) It is clear that the portfolio loss of agent $n = 1, \dots, N$ may be written as

$$U_p^{n,\varepsilon} = K_n^\varepsilon(\bar{\varphi}^n, \mu^0, T, \delta, 0) - K_n^\varepsilon(\bar{\varphi}^{\lambda,n}, \mu^0, T, \delta, 0). \tag{2.4.1}$$

The right hand side of (2.4.1) encapsulates the utility loss due to deviating from the optimal strategy $\bar{\varphi}^n$ to $\bar{\varphi}^{\lambda,n}$. It is often called the *indirect loss.*

2.4.1 Planner's perspective

Suppose that a planner with belief $(\varepsilon, \mathsf{P}(\varepsilon))$ sets the transaction tax λ. We assume that they view the transaction tax as beneficial if the aggregate utility loss under their belief ε is negative, i.e. the transaction tax is beneficial to the agents *en masse.* The following proposition enables us to decipher when this is the case.

Proposition 2.4.3. *Let* $(\varepsilon, \mathsf{P}(\varepsilon)) \in \operatorname{graph}(\mathsf{P})$*. The portfolio loss of agent* $n = 1, \dots, N$ *under under belief* ε *may be expressed as*

$$U_p^{n,\varepsilon} = \mathbb{E}^{\mathsf{P}(\varepsilon)}\left[\int_0^T e^{-\delta t}\Delta_t^n\left(\frac{\gamma_n\sigma^2}{2}\Delta_t + \sigma(\varepsilon_t^n - \varepsilon_t)\right)\mathrm{d}t\right]. \tag{2.4.2}$$

Moreover, the aggregate loss under belief ε *is given by*

$$U^\varepsilon = \mathbb{E}^{\mathsf{P}(\varepsilon)}\left[\int_0^T e^{-\delta t}\sum_{n=1}^N\left\{\Delta_t^n\left(\frac{\gamma_n\sigma^2}{2}\Delta_t^n + \sigma\varepsilon_t^n\right) + \lambda\left(\dot{\bar{\varphi}}_t^{\lambda,n}\right)^2\right\}\mathrm{d}t\right]. \tag{2.4.3}$$

Proof. Using Definition 2.2.6 and (2.2.2) we see that

$$\begin{aligned}
&\left(\bar{\varphi}_t^1 - \bar{\varphi}_t^{\lambda,1}\right)\left(\mu_t^0 + \sigma\varepsilon_t\right) - \frac{\gamma_n\sigma^2}{2}\left[(\bar{\varphi}_t^n + \zeta_t^n)^2 - \left(\bar{\varphi}_t^{\lambda,n} + \zeta_t^n\right)^2\right] \\
&= -\Delta_t^n\left(\mu_t^0 + \sigma\varepsilon_t\right) + \frac{\gamma_n\sigma^2}{2}\left[(\Delta_t^n)^2 + 2\Delta_t^n\left(\frac{\left(\mu_t^0 + \sigma\varepsilon_t^n\right)}{\sigma^2\gamma_n}\right)\right] \\
&= \Delta_t^n\left(\frac{\gamma_n\sigma^2}{2}\Delta_t^n + \sigma(\varepsilon_t^n - \varepsilon_t)\right),
\end{aligned}$$

making (2.4.2) is clear. Due to market, we have $\sum_{n=1}^{N-1}\Delta_t^n = -\Delta_t^N$. This allows us to write

$$\sum_{n=1}^{N} \Delta_t^n \left(\frac{\gamma_n \sigma^2}{2} \Delta_t^n + \sigma(\varepsilon_t^n - \varepsilon_t) \right) = \sum_{n=1}^{N} \Delta_t^n \left(\frac{\gamma_n \sigma^2}{2} \Delta_t^n + \sigma \varepsilon_t^n \right),$$

from which (2.4.3) follows trivially. □

To study when the transaction tax is beneficial from the planner's perspective, we will sometimes make the (naïve) assumption that we rebate the transaction tax as a lump sum to the agents post-trading. We will not specify under what rule the tax is rebated and will not incorporate the knowledge of the rebate into the control problems of the agents. We do this for simplicity, although Davilá argues in [29] that 'since investors are small, they never internalize the impact of their actions on the rebate they receive'. In this case, the object under scrutiny is the *post-rebate aggregate loss* denoted by

$$\tilde{U}^\varepsilon := U^\varepsilon - \sum_{n=1}^{N} U_{\mathrm{d}}^{n,\varepsilon} = \mathbb{E}^{\mathsf{P}(\varepsilon)} \left[\int_0^T e^{-\delta t} \sum_{n=1}^{N} \Delta_t^n \left(\frac{\gamma_n \sigma^2}{2} \Delta_t^n + \sigma \varepsilon_t^n \right) \mathrm{d}t \right]. \qquad (2.4.4)$$

It follows immediately that no matter what belief the planner holds, when agents have homogeneous beliefs about the risky asset

$$\tilde{U}^\varepsilon = \mathbb{E}^{\mathsf{P}(\varepsilon)} \left[\int_0^T e^{-\delta t} \sum_{n=1}^{N} \frac{\gamma_n \sigma^2}{2} \left(\Delta_t^n \right)^2 \, \mathrm{d}t \right] \geq 0,$$

implying that the transaction tax is detrimental.

The story changes when agents have heterogeneous beliefs about the risky asset. We note that the planner's belief only enters (2.4.4) via the local Radon-Nikodym derivative $\mathsf{Z}(\varepsilon)$. Since this is nonnegative, we will focus on the sign of the sum

$$\sum_{n=1}^{N} (-\Delta_t^n) \left(\frac{\gamma_n \sigma^2}{2} \Delta_t^n + \sigma \varepsilon_t^n \right). \qquad (2.4.5)$$

It is clear that in order for (2.4.1) to be negative, at the very least, we need that

$$\frac{1}{N} \sum_{n=1}^{N} (-\Delta_t^n) \, \sigma \varepsilon_t^n > 0,$$

i.e. the sample covariance between the scaled beliefs $\sigma(\varepsilon^1, \ldots, \varepsilon^N)$ and the current deviations $(\bar{\varphi}_t^1 - \bar{\varphi}_t^{\lambda,1}, \dots, \bar{\varphi}_t^N - \bar{\varphi}_t^{\lambda,N})$ must be positive. This will be the case when agents with optimistic (pessimistic) beliefs hold more (less) of the risky asset in the frictionless market. In other words, *the transaction tax can only be beneficial when it punishes false beliefs*, matching the intuition of Keynes [47].

In general, after division by N, we see that (2.4.1) is the sample covariance between the deviations $(\bar{\varphi}_t^1 - \bar{\varphi}_t^{\lambda,1}, \dots, \bar{\varphi}_t^N - \bar{\varphi}_t^{\lambda,N})$ and the vector

$$\left(\sigma\varepsilon_t^1 - \frac{\gamma_1\sigma^2}{2}(\bar{\varphi}_t^1 - \bar{\varphi}_t^{\lambda,1}), \dots, \sigma\varepsilon_t^N - \frac{\gamma_n\sigma^2}{2}(\bar{\varphi}_t^N - \bar{\varphi}_t^{\lambda,N})\right).$$

It is now clear that whether or not the transaction cost is beneficial depends on the interplay between the heterogeneity between beliefs and the agents' hedging needs. In particular, this showcases in a normative manner that the transaction tax dampens the wild beliefs of the agents while impeding their underlying need to hedge.

Note that no matter what convex function of the trading rate we choose to model the transaction tax, the preceding argument still applies as long as an optimal equilibrium portfolio exists. Moreover, since the sum (2.4.1) is independent of the planner's belief, this is a reasonably robust and general observation.

Chapter 3

Utility loss: Homogeneous Beliefs

We will now shed further light on the dynamics of the model introduced in the previous chapter by considering concrete examples. Here we analyse the case where agents have correct beliefs about the risky asset, and the transaction tax is small. In Chapter **??** we incorporate heterogeneous beliefs about the risky assets into the examples. To efficiently compute the relevant utility losses explicitly, we will assume that there are two agents, as is the case in [42, 35]. Furthermore, we will always consider an infinite time horizon ($T = \infty$).

We reiterate that some of the lengthier calculations in this chapter are aided by the use of custom functions in the computer language *Sage*. In particular, *Jupyter notebooks* containing these calculations can be found in the GitHub repository *https://github.com/odshelley/thesis*.

3.1 Explicit formulae

Suppose that the agents have homogeneous beliefs about the risky asset. In particular, we set $\varepsilon^n = 0$ for $n = 1, 2$. According to the discussion in Section 2.4.1, the planner will not deem the transaction tax beneficial. Nevertheless, we can still investigate the explicit forms of the utility losses of both agents to see whether the tax benefits an agent on an individual level.

In order to make such calculations, we suppose that the volatilities ζ^1 and ζ^2 in the dynamics of Y^1 and Y^2, are given by arithmetic Brownian motions (ABMs)

$$\zeta_t^1 := \alpha_1 t + \beta_1 W_t \quad \text{and} \quad \zeta_t^2 = \alpha_2 t + \beta_2 W_t, \tag{3.1.1}$$

for some $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$. This is a mild generalisation of the endowment streams considered in [42, 35], as they do not incorporate a drift component.

3.1.1 Frictionless baseline

From (2.2.4) it is immediate that the frictionless equilibrium portfolios of agents $n = 1, 2$ are ABMs

$$\bar{\varphi}_t^1 = \bar{\alpha} t + \bar{\beta} W_t = -\bar{\varphi}_t^2 \tag{3.1.2}$$

where

$$\bar{\alpha} := \frac{\gamma_2 \alpha_2 - \gamma_1 \alpha_1}{(\gamma_1 + \gamma_2)} \quad \text{and} \quad \bar{\beta} := \frac{\gamma_2 \beta_2 - \gamma_1 \beta_1}{(\gamma_1 + \gamma_2)}.$$

Moreover, by (2.2.3) the frictionless equilibrium return becomes

$$\mu_t^0 = \frac{\gamma_1 \gamma_2 \sigma^2}{(\gamma_1 + \gamma_2)} \left[(\alpha_1 + \alpha_2) t + (\beta_1 + \beta_2) W_t \right]. \tag{3.1.3}$$

3.1.2 Quadratic tax

In the presence of the quadratic transaction tax, we derive expressions for the equilibrium portfolios by utilising Theorem B.1.1(b). To make expressions succinct, we introduce the constants B and C given by

$$B := \frac{(\gamma_1 + \gamma_2)\sigma^2}{4\lambda} \quad \text{and} \quad C := \sqrt{B + \frac{\delta^2}{4}} - \frac{\delta}{2}.$$

Here, B is the tracking speed in the backward component (2.2.13) of the FBSDE seen in Lemma 2.2.4 that characterises $\bar{\varphi}^{\lambda,1}$.

Proposition 3.1.1. *Suppose that the volatilities ζ^1, ζ^2 are described by (3.1.1).*

(a) *The state variable Δ^1 is an Ornstein Uhlenbeck (OU) process with dynam-*

ics

$$\begin{cases} \mathrm{d}\Delta_t^1 &= -C\left(\frac{\bar{\alpha}\delta}{C(C+\delta)} + \Delta_t^1\right)\,\mathrm{d}t - \bar{\beta}\,\mathrm{d}W_t, \\ \Delta_0^1 &= 0. \end{cases} \tag{3.1.4}$$

Moreover,

$$\bar{\varphi}^{\lambda,1} = \bar{\varphi}^1 + \Delta^1 \quad \text{and} \quad \bar{\varphi}^{\lambda,2} = -\bar{\varphi}^{\lambda,1}.$$

(b) *The equilibrium trading rates are given by*

$$\dot{\bar{\varphi}}_t^{\lambda,1} = C\left(\frac{\bar{\alpha}}{(C+\delta)} - \Delta_t^1\right) \quad \text{and} \quad \dot{\bar{\varphi}}_t^{\lambda,2} = -\dot{\bar{\varphi}}_t^{\lambda,1}. \tag{3.1.5}$$

(c) *The equilibrium return is given by*

$$\mu_t^\lambda = \mu_t^0 + \frac{(\gamma_1 - \gamma_2)\sigma^2}{2}\Delta_t^1. \tag{3.1.6}$$

Proof. (a) Since the equilibrium portfolio of agent 1 is characterised by the FBSDE (2.2.12)-(2.2.13) from Lemma 2.2.4, it follows from Theorem B.1.1 that the portfolio is of the form

$$\int_0^t e^{-C(t-s)}\bar{\xi}_s\,\mathrm{d}s \tag{3.1.7}$$

where

$$\bar{\xi}_t = C(C+\delta)\mathbb{E}\left[\int_t^\infty e^{-(C+\delta)(s-t)}\bar{\varphi}_s^1\,\mathrm{d}s\middle|\mathcal{F}_t\right], \quad t \geq 0.$$

Using the conditional Fubini theorem

$$\begin{aligned} &\mathbb{E}\left[\int_t^\infty e^{-(C+\delta)(s-t)}W_s\,\mathrm{d}s\middle|\mathcal{F}_t\right] \\ &= \int_t^\infty e^{-(C+\delta)(s-t)}\mathbb{E}\left[W_s|\mathcal{F}_t\right]\,\mathrm{d}s \\ &= \frac{1}{(C+\delta)}W_t. \end{aligned}$$

Thus, $\bar{\xi}$ is clearly given by

$$\bar{\xi}_t = C\left(\frac{\bar{\alpha}}{(C+\delta)} + \bar{\varphi}^1_t\right). \tag{3.1.8}$$

Substituting (3.1.8) into (3.1.7) we see that an integration by parts gives

$$\begin{aligned}\varphi^{\lambda,1}_t &= -\frac{\delta\bar{\alpha}}{C(C+\delta)}(1-e^{-Ct}) - \bar{\beta}\int_0^t e^{-C(t-s)}\,\mathrm{d}W_s + \bar{\varphi}^1_t \\ &= \Delta^1_t + \bar{\varphi}^1_t. \end{aligned} \tag{3.1.9}$$

(b) It follows directly from (B.1.6), (3.1.8) and (3.1.9) that

$$\dot{\varphi}^{\lambda,1}_t = C\left(\frac{\bar{\alpha}}{(C+\delta)} - \Delta^1_t\right).$$

(c) This follows immediately from (2.2.7). □

Remark 3.1.2. (i) As the discount factor δ tends to zero, the dynamics of Δ^1 given in part (a) become identical to their equivalent in the 'long run average' setup considered in [35, Remark 3.6]. We will discuss this in Section 5.4.

(ii) Since we are primarily focusing on the effects of a small transaction tax, it is useful to note that

$$C = \frac{1}{\sqrt{\lambda}}\left\{\frac{\sqrt{\gamma_1+\gamma_2}|\sigma|}{2} + O(\sqrt{\lambda}\,)\right\} \quad \text{as } \lambda \downarrow 0. \tag{3.1.10}$$

In particular, part (b) tells us that for small transaction tax $\lambda \ll 1$, we have

$$\dot{\varphi}^{\lambda,1}_t \approx -\frac{1}{\sqrt{\lambda}}\left\{\frac{\sqrt{\gamma_1+\gamma_2}|\sigma|}{2}\Delta^1_t\right\}.$$

We are now in a position to explicitly calculate the utility losses introduced in Section 2.4. Before we do so, it is worth noting that we can also calculate the value functions of the agents in the market with and without friction. This allows us to see how they are affected by the parameters of the endowment streams.

Proposition 3.1.3. *When the volatilities ζ^1, ζ^2 are described by (3.1.1),*

$$\begin{aligned}
&\lim_{\lambda\downarrow 0} K^{\lambda,0}(\bar{\varphi}_t^{\lambda,n}) \\
&= -\frac{\gamma_1\gamma_2\sigma^2}{\delta^2}\left\{\frac{(2\,\beta_n\gamma_n + \beta_n\gamma_m - \beta_m\gamma_m)(\beta_1+\beta_2)}{2\,(\gamma_1+\gamma_2)^2}\right\} \\
&\qquad\qquad - \frac{\gamma_1\gamma_2\sigma^2}{\delta^3}\left\{\frac{(2\,\alpha_n\gamma_n + \alpha_n\gamma_m - \alpha_m\gamma_m)(\alpha_1+\alpha_2)}{(\gamma_1+\gamma_2)^2}\right\} \\
&= K^{0,0}(\bar{\varphi}_t^n).
\end{aligned} \tag{3.1.11}$$

for $(n,m) \in \{(1,2),(2,1)\}$.

Proof. This is shown via direct calculation in the notebook `homogeneous`. $\square$

When the agents have the same endowment stream, i.e. $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$, it follows from (3.1.11) that

$$K^{0,0}(\bar{\varphi}_t^n) = -\frac{\gamma_1\gamma_2\beta_1^2\sigma^2}{\delta^2(\gamma_1+\gamma_2)}\gamma_n - \frac{4\gamma_1\gamma_2\alpha_1^2\sigma^2}{\delta^3(\gamma_1+\gamma_2)}\gamma_n, \quad n = 1,2.$$

In this case, we can clearly see in what way the size of the volatilities of the endowments negatively impact the utilities of the agents.

3.2 Utility loss

We now present the main result of this subsection.

Theorem 3.2.1. *Suppose that ζ^1, ζ^2 are described by (3.1.1).*

(a) *The utility loss of the agents is given by*

$$U^{n,0} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{(\beta_1\gamma_1 - \beta_2\gamma_2)^2|\sigma|}{4(\gamma_1+\gamma_2)^{\frac{5}{2}}}(5\gamma_n - \gamma_m)\right\} + O(\lambda), \tag{3.2.1}$$

as $\lambda \downarrow 0$, for $(n,m) \in \{(1,2),(2,1)\}$.

(b) *The direct, return and portfolio losses of the agents are given by*

$$U_d^{n,0} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{(\beta_1\gamma_1 - \beta_2\gamma_2)^2|\sigma|}{4(\gamma_1+\gamma_2)^{\frac{3}{2}}}\right\}, \tag{3.2.2}$$

$$U_{\mathrm{r}}^{n,0} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{(\beta_1\gamma_1 - \beta_2\gamma_2)^2\,|\sigma|}{(\gamma_1+\gamma_2)^{\frac{5}{2}}}(\gamma_n - \gamma_m)\right\} + O(\lambda), \tag{3.2.3}$$

$$U_{\mathrm{p}}^{n,0} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{(\beta_1\gamma_1 - \beta_2\gamma_2)^2\,|\sigma|}{2(\gamma_1+\gamma_2)^{\frac{5}{2}}}\gamma^n\right\} + O(\lambda), \tag{3.2.4}$$

as $\lambda \downarrow 0$*, for* $(n, m) \in \{(1, 2), (2, 1)\}$.

(c) *The aggregate utility loss is given by*

$$U^0 = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{(\beta_1\gamma_1 - \beta_2\gamma_2)^2\,|\sigma|}{\delta(\gamma_1+\gamma_2)^{\frac{3}{2}}}\right\} + O(\lambda), \tag{3.2.5}$$

as $\lambda \downarrow 0$.

Proof. Note that part (a) and (c) follow directly from part (b), so we need only derive the expressions for the direct, return and portfolio losses. We approach this by rewriting each quantity in a form which can be computed using a symbolic algebra package.

Proposition 3.1.1(b) lets us write the direct loss as

$$U_{\mathrm{d}}^{1,0} = \lambda C^2 \mathbb{E}\left[\int_0^\infty e^{-\delta t}\left(\frac{\bar{\alpha}}{(C+\delta)} - \Delta_t^1\right)^2 dt\right]. \tag{3.2.6}$$

Furthermore, using (2.3.1), it follows that

$$U_{\mathrm{r}}^{1,0} = \frac{(\gamma_1-\gamma_2)\sigma^2}{2}\mathbb{E}\left[\int_0^\infty e^{-\delta t}(\bar{\varphi}_t^1 + \Delta_t^1)\Delta_t^1 \, \mathrm{d}t\right]. \tag{3.2.7}$$

Finally, using Proposition 2.4.3 we see that

$$U_{\mathrm{p}}^{1,0} = \frac{\gamma_1\sigma^2}{2}\mathbb{E}\left[\int_0^\infty e^{-\delta t}(\Delta_t^1)^2 dt\right]. \tag{3.2.8}$$

By using market clearing we get analogous expressions for the losses of agent 2. We complete the proof via direct calculation, found in the notebook `homogeneous`, which utilises (3.2.6), (3.2.7) and (3.2.8). ☐

Remark 3.2.2. Note that the drift parameters α_1 and α_2 from (3.1.1) are absent from the expressions as they are absorbed into the second order effects when transaction costs are small.

From the discussion in Section 2.4.1, we know that the planner will find the transaction tax detrimental since the aggregate loss is always nonnegative when the agents have correct beliefs. However, (3.2.1) tells us that agent $n = 1$ can benefit from the introduction of a small transaction tax, as long as her risk aversion parameter γ_1 is sufficiently small when compared to γ_2. This happens without a redistributive scheme, so it may seem odd. However, one can explain the phenomena by looking at the dynamics of the optimal portfolios.

From Lemma 2.2.2 and Theorem B.1.1, it follows that the dynamics of the individual optimal portfolios of the agents are given by

$$\mathrm{d}\bar{\varphi}_t^{\lambda,n} = C_n \left(\bar{\xi}_t^n - \bar{\varphi}_t^{\lambda,n} \right) \, \mathrm{d}t \quad n = 1, 2,$$

where

$$\bar{\xi}_t^n := (C_n + \delta)\mathbb{E}\left[\int_t^\infty e^{-(C_n+\delta)(s-t)} \xi_s^n \, \mathrm{d}s \middle| \mathcal{F}_t \right], \tag{3.2.9}$$

$$\xi_t^n := \frac{\mu_t^\lambda}{\gamma_n \sigma^2} - \zeta_t^n, \tag{3.2.10}$$

and

$$C_n := \sqrt{\frac{\gamma_n \sigma^2}{2\lambda} + \frac{\delta^2}{4}} - \frac{\delta}{2}.$$

Thus, we can view determining the optimal portfolio of agent n as a tracking problem, an idea formulated formally by Cai, Rosenbaum and Tankov [21]. The agent would like to keep her portfolio as close to the optimal frictionless portfolio as possible, but the transaction tax acts as an intervention cost for position adjustment, causing the optimal portfolios to become sluggish. Thus, agents trade towards the deviation portfolio (3.2.9), which is an average of the future values of the optimal frictionless portfolio (3.2.10), computed using an exponential discounting kernel. This matches the rhetoric of Gârleanu and Pedersen, who argue that an agent trades 'in front of the target' in the presence of transaction costs [37]. Indeed, due to this it is no surprise that the deviation process Δ^1 is mean-reverting. As the tax λ tends to zero the trading speed C_n tends to infinity, whence $\bar{\xi}^n$ approaches $\bar{\varphi}^n$, matching the small-cost asymptotics seen in [56]. This also aligns with Corollary 2.2.8, which shows that $\bar{\varphi}^{\lambda,n}$ is a combination of ξ^n along with a deviation term that tends to zero with the transaction tax.

In equilibrium, market clearing intertwines the agents' optimal portfolios. The fact that $C_n \propto \gamma_n$ incentivises agent 2 to track her target faster than agent 1, thus offering her favourable trades. This encapsulates that, due to her larger risk aversion, the hedging needs of agent 2 dominate that of agent 1. Hence, one does expect the return loss of agent $n = 1$ (2.2.20) to be strictly negative. However, it is notable that this effect can dominate the losses incurred directly from taxation and indirectly from the change in equilibrium portfolios.

This phenomenon is also related to the liquidity premium. Suppose we have an arbitrary amount of agents in the market. Then as is remarked in [17], when agents have homogeneous beliefs, (2.3.1) is the sample covariance between the vector of risk aversions $(\gamma_1, \ldots, \gamma_N)$ and the deviations $(\bar{\varphi}^{\Lambda,1} - \bar{\varphi}^1, \ldots, \bar{\varphi}^{\Lambda,N} - \bar{\varphi}^N)$. Thus, a positive liquidity premium occurs when the more risk-averse agents hold more risky assets in the market with friction. Since their need to hedge against their endowment fluctuations is stronger than that of the less risk-averse agents, they tend to be net buyers, resulting in a positive liquidity premium and benefitting the less risk-averse agents.

Remark 3.2.3 (Analogue)**.** Consider Merton's optimal investment problem, solved in [54], which studies optimal consumption and investment decisions for an investor who has available a bank account paying a fixed rate of interest and a stock whose price is a log-normal diffusion. In this model, transactions between bank and stock are costless and instant. For utility functions in the CRRA class, the optimal proportion of wealth to invest in a stock is constant, and one should consume at a rate proportional to total wealth. Thus, if the investor acts optimally, the portfolio holdings always lie on a line in the plane where the x-axis denotes holdings in the bank, and the y-axis denotes holdings in the stock. One calls this line the '*Merton line*'.

Applying the same strategy when proportional transaction costs are present results in immediate penury since constant trading is necessary to hold the portfolio on the Merton line; note that trading volume is proportional to the total variation of a portfolio, and prices follow diffusions that have infinite variation, which leads to the absurd conclusion that trading volume is infinite over any time interval. In this case, there must be some *no-trade zone*, or wedge around the line in the plane mentioned above, where the portfolio is optimal enough to make trading worthless. In this model, an agent's optimal portfolio lies within this wedge, where trade occurs at the boundary, allowing for wide portfolio oscillations. This was first shown explicitly by Davis and Norman in [25] as the solution to a singular stochastic control problem. Guasoni and Muhle-Kabre show in [36] that if we optimise the portfolio with respect to its long-run average, then the no-trading region reoccurs, and the width of this wedge is inversely proportional to the risk aversion of the agent.

In the above discussion, one can view the target (3.2.9) as the analogue of the Merton line. The no-trade zone ceases to exist in our model, but trading is slower the closer we are to the target, which gives a similar notion. Moreover, the tracking speed is proportional to the agent's risk aversion, similar to the relationship between risk aversion and wedge width when one has proportional costs.

3.2.1 Idiosyncratic tax

For the sake of completeness, we check the effect of allowing the planner to give idiosyncratic penalties to the agents. In particular, let's assume that the goal functional of agent $n = 1, 2$, is

$$K^0(\varphi, \lambda_n) = \mathbb{E}\left[\int_0^\infty e^{-\delta t}\left\{\varphi_t\mu_t - \frac{\gamma_n\sigma^2}{2}\left(\varphi_t + \zeta_t^n\right)^2 + \lambda_n\left(\dot{\varphi}_t\right)^2\right\}\,\mathrm{d}t\right], \quad (3.2.11)$$

where $\lambda_1 := \lambda \in (0, \infty)$ and $\lambda_2 := k\lambda$ for some $k \in \mathbb{R}_+$.

It is a straightforward generalisation of Theorem 2.2.7 to deduce an equilibrium return in this market.

Theorem 3.2.4. *The unique frictional equilibrium return is given by*

$$\begin{aligned}\nu_t^\lambda &= \frac{\sigma^2}{(k+1)}\left\{\left(k\gamma_1 - \gamma_2\right)\bar{\varphi}_t^{\lambda,1} + k\gamma_1\zeta_t^1 + \gamma_2\zeta_t^2\right\}\\ &= \mu_t^0 + \frac{(k\gamma_1 - \gamma_2)}{(k+1)}\sigma^2\Delta_t^1, \end{aligned} \quad (3.2.12)$$

where the corresponding individually optimal trading strategy $\bar{\varphi}_t^{\lambda,1}$ of agent $n = 1$ is given as the solution to the FBSDE

$$\mathrm{d}\bar{\varphi}_t^{\lambda,1} = \dot{\bar{\varphi}}_t^{\lambda,1}\,\mathrm{d}t, \quad \bar{\varphi}_0^{\lambda,1} = 0,$$
$$\mathrm{d}\dot{\bar{\varphi}}_t^{\lambda,1} = Z_t \mathrm{d}W_t + \left[\frac{(\gamma_1+\gamma_2)}{(1+k)}\frac{\sigma^2}{2\lambda}\right]\left(\bar{\varphi}_t^{\lambda,1} - \bar{\varphi}_t^1\right)\mathrm{d}t + \delta\dot{\bar{\varphi}}_t^{\lambda,1}\,\mathrm{d}t.$$

The corresponding individually optimal trading strategy of agent $n = 2$ is then given by $\bar{\varphi}_t^{\lambda,2} = -\bar{\varphi}_t^{\lambda,1}$.

Proof. The proof is analogous to the proof of Theorem 2.2.7. □

Remark 3.2.5. When the agents have homogeneous beliefs and have idiosyncratic taxes, from (3.2.12) we see that the liquidity premium becomes

$$\nu_t^\lambda - \mu_t^0 = \frac{(k\gamma_1 - \gamma_2)}{(k+1)}\sigma^2\Delta_t^1.$$

Unlike the market with a blanket tax for all agents, the frictionless equilibrium return no longer clears the market when agents have homogeneous risk aversions. This is caused by the heterogeneity in taxes, which causes the tracking speeds of the agents to differ, despite the homogeneous risk aversions.

Theorem 3.2.6. *Suppose that ζ^1, ζ^2 are described by (3.1.1).*

(a) *The utility loss of the agents is given by*

$$U^{1,0} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{\sqrt{2}(\beta_1\gamma_1 - \beta_2\gamma_2)^2|\sigma|}{4(\gamma_1+\gamma_2)^{\frac{5}{2}}(k+1)^{\frac{1}{2}}}((2+3k)\gamma_1 - \gamma_2)\right\} + O(\lambda), \tag{3.2.13}$$
$$U^{2,0} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{\sqrt{2}(\beta_1\gamma_1 - \beta_2\gamma_2)^2|\sigma|}{4(\gamma_1+\gamma_2)^{\frac{5}{2}}(k+1)^{\frac{1}{2}}}((2k+3)\gamma_2 - k\gamma_1)\right\} + O(\lambda), \tag{3.2.14}$$

as $\lambda \downarrow 0$.

(b) *The direct, return and portfolio losses of the agents are given by*

$$U_d^{1,0} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{\sqrt{2}(\beta_1\gamma_1 - \beta_2\gamma_2)^2|\sigma|}{4(\gamma_1+\gamma_2)^{\frac{3}{2}}(k+1)^{\frac{1}{2}}}\right\} + O(\lambda), \quad U_d^{2,0} = kU_d^{1,0}, \tag{3.2.15}$$

$$U_{\mathrm{r}}^{n,0} = (-1)^n \frac{\sqrt{\lambda}}{\delta} \left\{ \frac{\sqrt{2}(\beta_1\gamma_1 - \beta_2\gamma_2)^2 |\sigma|}{(\gamma_1+\gamma_2)^{\frac{5}{2}}(k+1)^{\frac{1}{2}}} (\gamma_2 - k\gamma_1) \right\} + O(\lambda), \quad (3.2.16)$$

$$U_{\mathrm{p}}^{n,0} = \frac{\sqrt{\lambda}}{\delta} \left\{ \frac{\sqrt{2}(\beta_1\gamma_1 - \beta_2\gamma_2)^2 |\sigma|}{4(\gamma_1+\gamma_2)^{\frac{5}{2}}(k+1)^{-\frac{1}{2}}} \gamma^n \right\} + O(\lambda), \quad (3.2.17)$$

as $\lambda \downarrow 0$, for $n = 1, 2$.

(c) The aggregate utility loss is given by

$$U^0 = \frac{\sqrt{\lambda}}{\delta} \left\{ \frac{\sqrt{2(k+1)}(\beta_1\gamma_1 - \beta_2\gamma_2)^2 |\sigma|}{2(\gamma_1+\gamma_2)^{\frac{3}{2}}} \right\} + O(\lambda), \quad (3.2.18)$$

as $\lambda \downarrow 0$.

Proof. The proof is analogous to the proof of Theorem 3.2.1. □

It is clear from (3.2.14) that due to the larger risk aversion of agent $n = 2$,

$$U^{2,0} \geq \frac{\sqrt{\lambda}}{\delta} \left\{ \frac{(\beta_1 - \beta_2)^2 |\sigma| \sqrt{\gamma_1}}{16} \left[\frac{(k+3)}{(k+1)^{\frac{1}{2}}} \right] \right\} + O(\lambda), \quad \text{as } \lambda \downarrow 0.$$

This means that the more risk averse agent does not benefit from a small transaction tax. This statement holds even in the extreme case where only the less risk averse agent is taxed, in which case

$$U^{2,0} \geq \frac{\sqrt{\lambda}}{\delta} \left\{ \frac{3(\beta_1 - \beta_2)^2 |\sigma| \sqrt{\gamma_1}}{16} \right\} + O(\lambda) \quad \text{as } \lambda \downarrow 0.$$

The phenomena occurs due to the transaction tax making the agents trading sluggish. In particular, in equilibrium this makes it more difficult for the more risk averse agent to hedge, thus making them worse off even if they are not taxed themselves.

3.2.2 Naïve rebate

One might wonder whether or not the utility losses of the agents can be covered by a rebate funded by the tax itself. For small transaction costs, this is clearly not possible en-masse, since (3.2.2) and (3.2.5) show that

$$\frac{1}{2}U^0 = U_{\mathrm{d}}^{1,0} + U_{\mathrm{d}}^{2,0} + O(\lambda),$$

as $\lambda \downarrow 0$. Furthermore, a rebate can never cover the losses of the more risk averse agent, since

$$U^{2,0} - \left(U_{\mathrm{d}}^{1,0} + U_{\mathrm{d}}^{2,0}\right) + O(\lambda) = \frac{3(\beta_1\gamma_1 - \beta_2\gamma_2)^2\,|\sigma|}{4(\gamma_1+\gamma_2)^{\frac{5}{2}}}(\gamma_2 - \gamma_1) \geq 0 \qquad (3.2.19)$$

as $\lambda \downarrow 0$.

It is often argued (e.g. by Kenneth Arrow in [6]) that the absolute risk aversion of a typical individual falls as wealth rises. Thus, one might deduce from (3.2.19) that when the market is in equilibrium and participants engage in fundamental trading, no matter how one offers a rebate, a transaction tax primarily penalises the poorer agents. Therefore, for completeness, we check whether there exist idiosyncratic taxes that would enable us to give the more risk-averse agent a rebate that covers their losses due to the tax.

We consider the quantity

$$R := \frac{U_d^{1,0} + U_d^{2,0}}{U^{2,0}},$$

noting that a rebate which covers agent 2's losses is possible only when $R \geq 1$. As in Section 3.2.1, we consider a scenario where agent 1 and 2 are penalised by tax levies $\lambda_1 := \lambda \in (0,\infty)$ and $\lambda_2 = k\lambda_1$, respectively. Furthermore, we define $k_\gamma \in (1,\infty)$ to be the constant such that $\gamma_2 = k_\gamma\gamma_1$. From Theorem 3.2.6, it is immediate that

$$R \approx \frac{(k_\gamma+1)(k+1)}{k(2\,k_\gamma - 1) + 3\,k_\gamma} =: \tilde{R}$$

for $\lambda_1 \ll 1$. Furthermore,

$$\frac{\mathrm{d}\tilde{R}}{\mathrm{d}k} = \frac{(k_\gamma+1)^2}{(2\,k_\gamma k + 3k_\gamma - k)^2} > 0.$$

Thus, assuming we give agent 2 the total accrued tax as a rebate, it is (almost paradoxically) more beneficial to tax them at a higher rate than agent 1. This is because a larger tax impedes agent 2's hedging needs and thus lowers the need to offer more favourable trades to agent 1 due to their mismatch in risk aversion. We may conclude that a small transaction tax will never be favoured by the more risk averse agent.

Chapter 4

Utility loss: Heterogeneous Beliefs

We now extend the set up so that agents $n = 1, 2$ have heterogeneous beliefs about the risky asset, modelled by the beliefs $(\varepsilon^n, \mathsf{P}(\varepsilon)) \in \mathrm{graph}(\mathsf{P}(\varepsilon))$ where $(\varepsilon_t^n)_{t\geq 0}$ is a zero mean reverting OU process with respect to a $\mathsf{P}(\varepsilon^n)$-Brownian motion W^n:

$$\begin{cases} \mathrm{d}\varepsilon_t^n = -\theta\varepsilon_t^n \, \mathrm{d}t + \sigma_n \, \mathrm{d}W_t^n \\ \varepsilon_0^n = 0, \end{cases} \tag{4.0.1}$$

for $\sigma_1, \sigma_2 \in \mathbb{R}$ and $\theta \in (0, \infty)$.

For ease of calculation, we will drop the drift terms from the shock volatilities, i.e. we set

$$\zeta_t^n := \beta_n W_t, \quad n = 1, 2. \tag{4.0.2}$$

This is a minor adjustment, as the drift terms do not enter any of the approximations seen in Theorem 3.2.1, and thus will be irrelevant when making comparisons between the markets with and without heterogeneous beliefs.

To simplify later expressions, we introduce the constants

$$\kappa_n := \theta + \sigma_n \quad n = 1, 2,$$

and

$$\kappa := \frac{1}{2}(\kappa_1 + \kappa_2) = \theta + \frac{1}{2}(\sigma_1 + \sigma_2).$$

Moreover, in order for the expressions to be appropriately integrable, we will always assume that

$$\theta > |\sigma_1| \vee |\sigma_2| . \tag{4.0.3}$$

In particular, this ensures that κ_1, κ_2 and $\kappa \in (0, \infty)$.

4.1 Explicit formulae

Since ε^1 and ε^2 are OU processes with respect to different Brownian motions, it will be useful to view them under common measures. Recall that the *average belief* of the agents is denoted by $\bar{\varepsilon} := \frac{1}{2}(\varepsilon^1 + \varepsilon^2)$ and $\bar{\mathsf{P}} := \mathsf{P}(\bar{\varepsilon})$.

Proposition 4.1.1. *Let $\varepsilon^1, \varepsilon^2$ be defined by (4.0.1).*

(a) *The process ε^n is given by*

$$\begin{aligned} \varepsilon_t^n &= \sigma_n \int_0^t e^{-\kappa_n(t-s)}\, \mathrm{d}W_s && (4.1.1)\\ &= (\sigma_n - \sigma_m) \int_0^t e^{-\kappa_n(t-s)} \mathrm{d}W_t^m + \sigma_m \int_0^t e^{-\theta(t-s)}\, \mathrm{d}W_s^m, && (4.1.2) \end{aligned}$$

for $(n, m) \in \{(1,2),(2,1)\}$.

(b) *The difference $\varepsilon^1 - \varepsilon^2$ is given by*

$$\varepsilon_t^1 - \varepsilon_t^2 = (\sigma_1 - \sigma_2) \int_0^t e^{-\kappa(t-s)}\, \mathrm{d}W_s^{\bar{\varepsilon}}. \tag{4.1.3}$$

(c) *When $\sigma_1 = -\sigma_2 = \sigma_0$, the sum $\varepsilon^1 + \varepsilon^2$ is given by*

$$\varepsilon_t^1 + \varepsilon_t^2 = -2\sigma_0^2 \int_0^t e^{-\theta(t-s)}(t-s)\, \mathrm{d}W_s^{\bar{\varepsilon}}. \tag{4.1.4}$$

Otherwise,

$$\begin{aligned} \varepsilon_t^1 + \varepsilon_t^2 = 4\frac{\sigma_1\sigma_2}{(\sigma_1+\sigma_2)} \int_0^t e^{-\theta(t-s)}\, \mathrm{d}W_s^{\bar{\varepsilon}} \\ + \frac{(\sigma_1-\sigma_2)^2}{(\sigma_1+\sigma_2)} \int_0^t e^{-\kappa(t-s)}\, \mathrm{d}W_s^{\bar{\varepsilon}}. && (4.1.5) \end{aligned}$$

Proof. (a) From the dynamics (4.0.1) and Theorem A.1.13 it follows that

$$\begin{aligned}
\mathrm{d}\varepsilon_t^1 &= -\theta\varepsilon_t^1 \,\mathrm{d}t + \sigma_1 \left(\mathrm{d}W_t - \varepsilon_t^1 \,\mathrm{d}t\right) \\
&= -(\theta + \sigma_1)\varepsilon_t^1 \,\mathrm{d}t + \sigma_1 \,\mathrm{d}W_t,
\end{aligned}$$

resulting in (4.1.1). Moreover,

$$\begin{aligned}
\mathrm{d}\varepsilon_t^1 &= -(\theta + \sigma_1)\varepsilon_t^1 \,\mathrm{d}t + \sigma_1 \left(\mathrm{d}W_t^2 + \varepsilon_t^2 \,\mathrm{d}t\right) \\
&= -(\theta + \sigma_1)\varepsilon_t^1 \,\mathrm{d}t + \sigma_1\varepsilon_t^2 \,\mathrm{d}t + \sigma_1 \mathrm{d}W_t^2.
\end{aligned}$$

Using the integrating factor $e^{\kappa_1 t}$ and Proposition C.1.1, it follows that

$$\begin{aligned}
\varepsilon_t^1 &= \sigma_1 \int_0^t e^{-\kappa_1(t-s)}\varepsilon_s^2 \,\mathrm{d}s + \sigma_1 \int_0^t e^{-\kappa_1(t-s)}\mathrm{d}W_t^2 \\
&= \sigma_1\sigma_2 \int_0^t e^{-\kappa_1(t-s)} \int_0^s e^{-\theta(s-r)} \,\mathrm{d}W_r^2 \,\mathrm{d}s + \sigma_1 \int_0^t e^{-\kappa_1(t-s)}\mathrm{d}W_t^2 \\
&= \frac{\sigma_1\sigma_2}{\theta - \kappa_1} \left(\int_0^t e^{-\kappa_1(t-s)} \,\mathrm{d}W_s^2 - \int_0^t e^{-\theta(t-s)} \,\mathrm{d}W_s^2 \right) + \sigma_1 \int_0^t e^{-\kappa_1(t-s)}\mathrm{d}W_t^2 \\
&= (\sigma_1 - \sigma_2) \int_0^t e^{-\kappa_1(t-s)}\mathrm{d}W_t^2 + \sigma_2 \int_0^t e^{-\theta(t-s)} \,\mathrm{d}W_s^2.
\end{aligned}$$

By the same argument we get an analogous expression for ε^2.

(b) From the dynamics (4.0.1), the definition of the belief $(\bar{\varepsilon}, \bar{\mathsf{P}})$, and Theorem A.1.13, it follows that

$$\mathrm{d}\varepsilon_t^n = -\theta\varepsilon_t^n \,\mathrm{d}t + \sigma_1 \left(\mathrm{d}W_t^{\bar{\varepsilon}} + \frac{(-1)^n}{2}(\varepsilon_t^1 - \varepsilon_t^2) \,\mathrm{d}t \right) \qquad n = 1, 2. \tag{4.1.6}$$

Hence,

$$\mathrm{d}(\varepsilon_t^1 - \varepsilon_t^2) = -\left(\theta + \frac{\sigma_1 + \sigma_2}{2}\right)(\varepsilon_t^1 - \varepsilon_t^2) \,\mathrm{d}t + (\sigma_1 - \sigma_2) \,\mathrm{d}W_t^{\bar{\varepsilon}}.$$

The expression (4.1.3) is clear by noting that $\varepsilon_0^1 = \varepsilon_0^2 = 0$.

(c) Using the integrating factor $e^{\theta t}$, it follows from (4.1.6) that

$$
\begin{aligned}
\varepsilon_t^1 &= -\frac{\sigma_1}{2}\int_0^t e^{-\theta(t-s)}(\varepsilon_s^1-\varepsilon_s^2)\ \mathrm{d}s + \sigma_1\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}} \\
&= -\frac{\sigma_1(\sigma_1-\sigma_2)}{2}\int_0^t e^{-\theta(t-s)}\left(\int_0^s e^{-\kappa(s-r)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right)\mathrm{d}s \\
&+ \sigma_1\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}.
\end{aligned} \tag{4.1.7}
$$

It follows from Proposition C.1.7 that when $\sigma_1 \neq -\sigma_2$,

$$
\begin{aligned}
&e^{-\theta t}\int_0^t e^{-(\kappa-\theta)s}\left(\int_0^s e^{\kappa r}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right)\mathrm{d}s \\
&= -\frac{e^{-\theta t}}{\theta-\kappa}\left(\int_0^t e^{\theta s}\ \mathrm{d}W_s^{\bar{\varepsilon}} - e^{\theta t}\int_0^t e^{-\kappa(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right) \\
&= \frac{2}{(\sigma_1+\sigma_2)}\left(\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}} - \int_0^t e^{-\kappa(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right).
\end{aligned} \tag{4.1.8}
$$

Substituting into (4.1.8) into (4.1.7), we see that

$$
\begin{aligned}
\varepsilon_t^1 &= -\frac{\sigma_1(\sigma_1-\sigma_2)}{(\sigma_1+\sigma_2)}\left(\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}} - \int_0^t e^{-\kappa(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right) \\
&+ \sigma_1\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}} \\
&= \sigma_1\left\{\left(\frac{2\sigma_2}{\sigma_1+\sigma_2}\right)\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}} + \left(\frac{\sigma_1-\sigma_2}{\sigma_1+\sigma_2}\right)\int_0^t e^{-\kappa(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right\}.
\end{aligned}
$$

Similarly,

$$
\varepsilon_t^2 = \sigma_2\left\{\left(\frac{2\sigma_1}{\sigma_1+\sigma_2}\right)\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}} + \left(\frac{\sigma_2-\sigma_1}{\sigma_1+\sigma_2}\right)\int_0^t e^{-\kappa(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right\},
$$

whence

$$
\varepsilon_t^1+\varepsilon_t^2 = 4\frac{\sigma_1\sigma_2}{(\sigma_1+\sigma_2)}\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}} + \frac{(\sigma_1-\sigma_2)^2}{(\sigma_1+\sigma_2)}\int_0^t e^{-\kappa(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}.
$$

When $\sigma_1 = -\sigma_2 = \sigma_0$, it follows from (4.1.6) and Proposition C.1.7

$$
\varepsilon_t^n = -\sigma_0^2\int_0^t\left(\int_0^s e^{-\theta(t-r)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right)\mathrm{d}s - (-1)^n\sigma_0\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}
$$

$$= -\sigma_0^2 \int_0^t e^{-\theta(t-s)}(t-s)\ \mathrm{d}W_s^{\bar{\varepsilon}} - (-1)^n \sigma_0 \int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}} \quad n = 1, 2,$$

and so

$$\varepsilon_t^1 + \varepsilon_t^2 = -2\sigma_0^2 \int_0^t e^{-\theta(t-s)}(t-s)\ \mathrm{d}W_s^{\bar{\varepsilon}}.$$

□

Remark 4.1.2. From part (a) and our assumption (4.0.3) we see that ε^1 and ε^2 are OU processes with respect to $\mathbb{P}$. Thus, it is clear that $\varepsilon^1, \varepsilon^2 \in \mathscr{L}_\delta^4$ for any $\delta > 0$, whence $\varepsilon^1, \varepsilon^2$ are pre-admissible beliefs as defined in Section 2.1.3.

The following proposition is for the purpose of calculation.

Proposition 4.1.3. *Let ζ^1, ζ^2 and $\varepsilon^1, \varepsilon^2$ be described by (4.0.1) and (4.0.2), respectively.*

(a) *For $n = 1, 2$*

$$\frac{\gamma_2 \zeta_t^2 - \gamma_1 \zeta_t^1}{(\gamma_1 + \gamma_2)} = \bar{\beta} \left\{ W_t^n - \sigma_n \int_0^t e^{-\theta(t-s)}(t-s)\ \mathrm{d}W_s^n \right\}. \tag{4.1.9}$$

(b) *When $\sigma_1 + \sigma_2 \neq 0$*

$$\begin{aligned}\frac{\gamma_2 \zeta_t^2 - \gamma_1 \zeta_t^1}{(\gamma_1 + \gamma_2)} = \bar{\beta} \Bigg\{ W_t^{\bar{\varepsilon}} &- \frac{2\sigma_1\sigma_2}{(\sigma_1 + \sigma_2)} \int_0^t e^{-\theta(t-s)}(t-s)\ \mathrm{d}W_s^{\bar{\varepsilon}} \\ &- \frac{1}{2} \frac{(\sigma_1 - \sigma_2)^2}{(\sigma_1 + \sigma_2)} \int_0^t e^{-\kappa(t-s)}(t-s)\ \mathrm{d}W_s^{\bar{\varepsilon}} \Bigg\}.\end{aligned} \tag{4.1.10}$$

(c) *When $\sigma_1 = -\sigma_2 = \sigma_0$*

$$\frac{\gamma_2 \zeta_t^2 - \gamma_1 \zeta_t^1}{(\gamma_1 + \gamma_2)} = \bar{\beta} \left\{ W_t^{\bar{\varepsilon}} + \sigma_0^2 \int_0^t e^{-\theta(t-s)}(t-s) W_s^{\bar{\varepsilon}}\ \mathrm{d}s \right\}. \tag{4.1.11}$$

Proof. (a) By the definition of the belief $(\varepsilon^n, \mathsf{P}^n)$ it follows that

$$\frac{\gamma_2 \zeta_t^2 - \gamma_1 \zeta_t^1}{(\gamma_1 + \gamma_2)} = \frac{\gamma_2 \beta_2 - \gamma_1 \beta_1}{(\gamma_1 + \gamma_2)} \left(W_t^n + \int_0^t \varepsilon_s^n\ \mathrm{d}s \right). \tag{4.1.12}$$

Using Proposition C.1.1, it follows that

$$\int_0^t \varepsilon_s^n \, \mathrm{d}s = \sigma_n \int_0^t e^{-\theta(t-s)}(t-s) \, \mathrm{d}W_s^n.$$

By substituting back into (4.1.12) we get (4.1.9).

(b) By the definition of the belief $(\bar{\varepsilon}, \bar{\mathsf{P}})$ it follows that

$$\frac{\gamma_2\zeta_t^2 - \gamma_1\zeta_t^1}{(\gamma_1+\gamma_2)} = \frac{\gamma_2\beta_2 - \gamma_1\beta_1}{(\gamma_1+\gamma_2)} \left(W_t^{\bar{\varepsilon}} - \frac{1}{2}\int_0^t (\varepsilon_s^1 + \varepsilon_s^2) \, \mathrm{d}s \right). \tag{4.1.13}$$

Thus, using Propositions 4.1.1(c) and C.1.7, it follows that

$$\begin{aligned} &\int_0^t (\varepsilon_s^1 + \varepsilon_s^2) \, \mathrm{d}s \\ &= \frac{4\sigma_1\sigma_2}{(\sigma_1+\sigma_2)} \int_0^t e^{-\theta(t-s)}(t-s) \, \mathrm{d}W_s^{\bar{\varepsilon}} + \frac{(\sigma_1-\sigma_2)^2}{(\sigma_1+\sigma_2)} \int_0^t e^{-\kappa(t-s)}(t-s) \, \mathrm{d}W_s^{\bar{\varepsilon}}. \end{aligned}$$

By substituting into (4.1.13) we get (4.1.10).

(c) By Propositions 4.1.1(c) and C.1.8, it follows that

$$\int_0^t (\varepsilon_s^1 + \varepsilon_s^2) \, \mathrm{d}s = -2\sigma_0^2 \int_0^t e^{-\theta(t-s)}(t-s) W_s^{\bar{\varepsilon}} \, \mathrm{d}s.$$

By substituting into (4.1.13) we get (4.1.11). □

4.1.1 Frictionless baselines

From (2.2.4) it is clear that the equilibrium portfolio in the frictionless market is given by

$$\bar{\varphi}_t^n = \frac{\gamma_2\zeta_t^2 - \gamma_1\zeta_t^1}{(\gamma_1+\gamma_2)} + \frac{\varepsilon_t^1 - \varepsilon_t^2}{\sigma(\gamma_1+\gamma_2)} \quad \text{and } n = 1, 2.$$

The exact expressions for $\bar{\varphi}^1$ and $\bar{\varphi}^2$ under the measures $\mathbb{P}, \bar{\mathsf{P}}, \mathsf{P}^1$ and P^2 are clear from Propositions (4.1.1) and (4.1.3). Similarly, one can get the exact expressions for the frictionless equilibrium return since (2.2.3) tells us that

$$\mu_t^0 = \frac{\gamma_1\gamma_2\sigma^2}{(\gamma_1+\gamma_2)}\left[\zeta_t^1 + \zeta_t^2 - \frac{\gamma_2\varepsilon_t^1 + \gamma_1\varepsilon_t^2}{\gamma_1\gamma_2\sigma}\right] \quad .$$

4.1.2 Quadratic costs

We now derive expressions for the equilibrium portfolios, the equilibrium trading rates and equilibrium return in the market with friction. This will allow us to express the utility losses in the next subsection. For notational purposes we define

$$\mathsf{C}(x) := \frac{C(C+\delta)}{(C+\delta+x)(C-x)} \tag{4.1.14}$$

for any $x \in \mathbb{R}$ for which $\mathsf{C}(x)$ is well defined. Note that C depends on λ and from (3.1.10)

$$\mathsf{C}(x) \to 1 \quad \text{as} \quad \lambda \downarrow 0.$$

Proposition 4.1.4. *Suppose that $\varepsilon^1, \varepsilon^2$ and ζ^1, ζ^2 are described by (4.0.1) and (4.0.2), respectively.*

(a) *The equilibrium portfolios are given by*

$$\bar{\varphi}_t^{\lambda,1} = \bar{\varphi}_t^1 + \tilde{\Delta}_t^1 + O(\lambda) \quad \text{and} \quad \bar{\varphi}_t^{\lambda,2} = -\bar{\varphi}_t^{\lambda,1},$$

as $\lambda \downarrow 0$ for all $t \geq 0$, where

$$\tilde{\Delta}^1 := -\frac{[(\sigma_1-\sigma_2) - \sigma(\gamma_1\beta_1 - \gamma_2\beta_2)]}{\sigma(\gamma_1+\gamma_2)}\int_0^{\cdot} e^{-C(t-s)}\, \mathrm{d}W_s^{\bar{\varepsilon}}. \tag{4.1.15}$$

(b) *The equilibrium trading rates are given by*

$$\dot{\bar{\varphi}}_t^{\lambda,1} = -\sqrt{\frac{(\gamma_1+\gamma_2)\sigma^2}{4\lambda}}\tilde{\Delta}_t^1 + O(1) \quad \text{and} \quad \dot{\bar{\varphi}}_t^{\lambda,2} = -\dot{\bar{\varphi}}_t^{\lambda,1},$$

as $\lambda \downarrow 0$ for all $t \geq 0$.

(c) *The equilibrium return is given by*

$$\mu_t^\lambda = \mu_t^0 + \frac{\sigma^2(\gamma_1-\gamma_2)}{2}\Delta_t^1 - \lambda Z(\varepsilon_t^1 - \varepsilon_t^2)$$

where

$$Z = \sqrt{\frac{(\gamma_1+\gamma_2)\sigma^2}{4\lambda}} \left\{ \frac{[(\sigma_1-\sigma_2) - \sigma(\gamma_1\beta_1 - \gamma_2\beta_2)]}{\sigma(\gamma_1+\gamma_2)} \right\} + O(1)$$

as $\lambda \downarrow 0$.

Proof. (a) We only calculate the equilibrium portfolio of agent 1, since agent 2's portfolio is determined by market clearing. According to Theorem 2.2.20 and Theorem B.1.1

$$\varphi_t^{\lambda,1} = \int_0^t e^{-C(t-s)} \xi_s^\varepsilon \, \mathrm{d}s \tag{4.1.16}$$

$$+ \int_0^t e^{-C(t-s)} \xi_s^\zeta \, \mathrm{d}s, \tag{4.1.17}$$

where

$$\xi_t^\varepsilon := C(C+\delta)\mathbb{E}^{\bar{\mathsf{P}}}\left[\int_t^\infty e^{-(C+\delta)(s-t)} \frac{\varepsilon_s^1 - \varepsilon_s^2}{\sigma(\gamma_1+\gamma_2)} \, \mathrm{d}s \middle| \mathcal{F}_t \right],$$

$$\xi_t^\zeta := C(C+\delta)\mathbb{E}^{\bar{\mathsf{P}}}\left[\int_t^\infty e^{-(C+\delta)(s-t)} \frac{\gamma_2\zeta_s^2 - \gamma_1\zeta_s^2}{(\gamma_1+\gamma_2)} \, \mathrm{d}s \middle| \mathcal{F}_t \right].$$

We will compute (4.1.16) and (4.1.17) in turn.

First, note that by Proposition C.1.3

$$\begin{aligned}
&\mathbb{E}^{\bar{\mathsf{P}}}\left[\int_t^\infty e^{-(C+\delta)(s-t)} \frac{\varepsilon_s^1 - \varepsilon_s^2}{\sigma(\gamma_1+\gamma_2)} \, \mathrm{d}s \middle| \mathcal{F}_t \right] \\
&= \frac{(\sigma_1-\sigma_2)}{\sigma(\gamma_1+\gamma_2)} e^{(C+\delta)t} \mathbb{E}\left[\int_t^\infty e^{-(C+\delta)s} \int_0^s e^{-\kappa(s-r)} \, \mathrm{d}W_r^{\bar{\varepsilon}} \, \mathrm{d}s \middle| \mathcal{F}_t \right] \\
&= \frac{(\sigma_1-\sigma_2)}{\sigma(\gamma_1+\gamma_2)} \frac{1}{(C+\delta+\kappa)} \int_0^t e^{-\kappa(t-s)} \, \mathrm{d}W_s^{\bar{\varepsilon}}.
\end{aligned} \tag{4.1.18}$$

Thus, using Proposition C.1.1

$$\begin{aligned}
&\int_0^t e^{-C(t-s)} \xi_s^\varepsilon \, \mathrm{d}s \\
&= \mathsf{C}(\kappa) \left\{ \frac{\varepsilon_t^1 - \varepsilon_t^2}{\sigma(\gamma_1+\gamma_2)} - \frac{(\sigma_1-\sigma_2)}{\sigma(\gamma_1+\gamma_2)} \int_0^t e^{-C(t-s)} \mathrm{d}W_s^{\bar{\varepsilon}} \right\} \\
&= \left\{ \frac{\varepsilon_t^1 - \varepsilon_t^2}{\sigma^2(\gamma_1+\gamma_2)} - \frac{(\sigma_1-\sigma_2)}{\sigma(\gamma_1+\gamma_2)} \int_0^t e^{-C(t-s)} \mathrm{d}W_s^{\bar{\varepsilon}} \right\} + o(1), \quad \lambda \downarrow 0.
\end{aligned}$$

Next we calculate (4.1.17). Suppose that $\sigma_1+\sigma_2\neq 0$. By Propositions 4.1.3(a) and C.1.12, it follows that

$$\begin{aligned}
&\mathbb{E}^{\bar{\mathsf{P}}}\left[\int_t^\infty e^{-(C+\delta)(s-t)}\frac{\gamma_2\zeta_s^2-\gamma_1\zeta_s^2}{(\gamma_1+\gamma_2)}\mathrm{d}s\Big|\mathcal{F}_t\right]\\
&=\frac{\bar{\beta}}{(C+\delta)}W_t^{\bar{\varepsilon}}\\
&-\bar{\beta}\frac{2\sigma_1\sigma_2}{(\sigma_1+\sigma_2)}\frac{1}{(C+\delta+\theta)}\left(\int_0^t e^{-\theta(t-s)}(t-s)\ \mathrm{d}W_s^{\bar{\varepsilon}}+\frac{1}{(C+\delta+\theta)}\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right)\\
&-\bar{\beta}\frac{1}{2}\frac{(\sigma_1-\sigma_2)^2}{(\sigma_1+\sigma_2)}\frac{1}{(C+\delta+\kappa)}\left(\int_0^t e^{-\kappa(t-s)}(t-s)\ \mathrm{d}W_s^{\bar{\varepsilon}}+\frac{1}{(C+\delta+\kappa)}\int_0^t e^{-\kappa(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right).
\end{aligned}\tag{4.1.19}$$

Hence, using Propositions C.1.4 and C.1.10

$$\begin{aligned}
&\int_0^t e^{-C(t-s)}\xi_s^\zeta\ \mathrm{d}s\\
&=\bar{\beta}\left\{W_t^{\bar{\varepsilon}}-\int_0^t e^{-C(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right\}\\
&-\bar{\beta}\frac{2\sigma_1\sigma_2}{(\sigma_1+\sigma_2)}\mathsf{C}(\theta)\Bigg\{\int_0^t e^{-\theta(t-s)}(t-s)\ \mathrm{d}W_s^{\bar{\varepsilon}}\\
&\qquad\qquad+\frac{2\theta+\delta}{(C-\theta)(C+\delta+\theta)}\left(\int_0^t e^{-C(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}-\int_0^t e^{-\theta(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right)\Bigg\}\\
&-\bar{\beta}\frac{1}{2}\frac{(\sigma_1-\sigma_2)^2}{(\sigma_1+\sigma_2)}\mathsf{C}(\kappa)\Bigg\{\int_0^t e^{-\kappa(t-s)}(t-s)\ \mathrm{d}W_s^{\bar{\varepsilon}}\\
&\qquad\qquad+\frac{2\kappa+\delta}{(C-\kappa)(C+\delta+\kappa)}\left(\int_0^t e^{-C(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}-\int_0^t e^{-\kappa(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}\right)\Bigg\}.
\end{aligned}$$

From (3.1.10) it now follows that

$$\int_0^t e^{-C(t-s)}\xi_s^\zeta\ \mathrm{d}s=\frac{\gamma_2\zeta_t^2-\gamma_1\zeta_t^2}{(\gamma_1+\gamma_2)}-\bar{\beta}\int_0^t e^{-C(t-s)}\ \mathrm{d}W_s^{\bar{\varepsilon}}+O(\lambda)\qquad \lambda\downarrow 0.$$

It remains to calculate (4.1.17) in the case where $\sigma_1=-\sigma_2=\sigma_0$. By Propositions 4.1.3(b) and C.1.14 it follows that

$$\begin{aligned}
&\mathbb{E}^{\bar{\mathsf{P}}}\left[\int_t^\infty e^{-(C+\delta)(s-t)}\frac{\gamma_2\zeta_s^2-\gamma_1\zeta_s^2}{\sigma(\gamma_1+\gamma_2)}\mathrm{d}s\Big|\mathcal{F}_t\right]\\
&=\frac{\bar{\beta}}{(C+\delta)}W_t^{\bar{\varepsilon}}
\end{aligned}$$

$$+\frac{\bar{\beta}\sigma_0^2}{(C+\delta+\theta)^2}\left((C+\delta+\theta)\int_0^t e^{-\theta(t-r)}(t-r)W_r^{\bar{\varepsilon}}\,\mathrm{d}r+\int_0^t e^{-\theta(t-r)}W_r^{\bar{\varepsilon}}\,\mathrm{d}r-\frac{W_s^{\bar{\varepsilon}}}{(C+\delta)}\right). \tag{4.1.20}$$

Thus, by Propositions C.1.9 and C.1.8 it follows that

$$\begin{aligned}
&\sigma_0^2\int_0^t e^{-C(t-s)}\int_0^s e^{-\theta(s-r)}(s-r)W_r^{\bar{\varepsilon}}\,\mathrm{d}r\,\mathrm{d}s\\
&=\frac{\sigma_0^2}{(C-\theta)}\left(\int_0^t e^{-\theta(t-s)}(t-s)W_s^{\bar{\varepsilon}}\,\mathrm{d}s+\frac{1}{(C-\theta)}\left(\int_0^t e^{-C(t-s)}W_r^{\bar{\varepsilon}}\,\mathrm{d}r-\int_0^t e^{-\theta(t-s)}W_s^{\bar{\varepsilon}}\,\mathrm{d}s\right)\right)\\
&=-\frac{1}{(C-\theta)}\left(\frac{1}{2}\int_0^t(\varepsilon_s^1+\varepsilon_s^2)\,\mathrm{d}s\right)+\frac{\sigma_0^2}{(C-\theta)^2}\left(\int_0^t e^{-C(t-s)}W_r^{\bar{\varepsilon}}\,\mathrm{d}r-\int_0^t e^{-\theta(t-s)}W_s^{\bar{\varepsilon}}\,\mathrm{d}s\right).
\end{aligned}$$

Moreover,

$$\sigma_0^2\int_0^t e^{-C(t-s)}\int_0^t e^{-\theta(s-r)}W_r^{\bar{\varepsilon}}\,\mathrm{d}r\,\mathrm{d}s=\frac{\sigma_0^2}{(C-\theta)}\left(\int_0^t e^{-\theta(t-s)}W_s^{\bar{\varepsilon}}\,\mathrm{d}s-\int_0^t e^{-C(t-s)}W_s^{\bar{\varepsilon}}\,\mathrm{d}s\right).$$

Combining this together yields

$$\begin{aligned}
&\int_0^t e^{-C(t-s)}\xi_s^{\zeta}\,\mathrm{d}s\\
&=\bar{\beta}\left(W_s^{\bar{\varepsilon}}-\int_0^t e^{-C(t-s)}\,\mathrm{d}W_s^{\bar{\varepsilon}}\right)\\
&-\bar{\beta}\mathsf{C}(\theta)\Bigg\{\frac{1}{2}\int_0^t(\varepsilon_s^1+\varepsilon_s^2)\,\mathrm{d}s\\
&+\frac{(2\theta+\delta)\sigma_0^2}{(C+\delta+\theta)(C-\theta)}\left(\int_0^t e^{-\theta(t-s)}W_s^{\bar{\varepsilon}}\,\mathrm{d}s-\int_0^t e^{-C(t-s)}W_s^{\bar{\varepsilon}}\,\mathrm{d}s\right)\\
&+\frac{(C-\theta)\sigma_0^2}{C(C+\delta)(C+\delta+\theta)}\left(W_s^{\bar{\varepsilon}}-\int_0^t e^{-C(t-s)}\,\mathrm{d}W_s^{\bar{\varepsilon}}\right)\Bigg\}\\
&=\bar{\beta}\left(W_s^{\bar{\varepsilon}}-\frac{\mathsf{C}(\theta)}{2}\int_0^t(\varepsilon_s^1+\varepsilon_s^2)\,\mathrm{d}s\right)-\bar{\beta}\int_0^t e^{-C(t-s)}\,\mathrm{d}W_s^{\bar{\varepsilon}}\\
&-\bar{\beta}\sigma_0^2\mathsf{C}(\theta)\left(\theta_1\int_0^t e^{-\theta(t-s)}\,\mathrm{d}W_s^{\bar{\varepsilon}}+\theta_2\int_0^t e^{-C(t-s)}\,\mathrm{d}W_s^{\bar{\varepsilon}}+\theta_3W_t^{\bar{\varepsilon}}\right),
\end{aligned}$$

where

$$\begin{aligned}
\theta_1&:=-\frac{(2\theta+\delta)}{\theta(C-\theta)(C+\delta+\theta)}=O(\lambda),\\
\theta_2&:=\frac{1}{C(C+\delta+\theta)}\left(\frac{(2\theta+\delta)}{(C-\theta)}-\frac{(C-\theta)}{(C+\delta)}\right)=O(\lambda),
\end{aligned}$$

$$\theta_3 := \frac{\delta(C+\delta)+\theta(3C+2\delta-\theta)}{(C+\delta)(C+\delta+\theta)C\theta} = O(\lambda),$$

as $\lambda \downarrow 0$. From (3.1.10) it now follows that

$$\int_0^t e^{-C(t-s)}\xi_s^\zeta \,\mathrm{d}s = \frac{\gamma_2\zeta_t^2-\gamma_1\zeta_t^2}{(\gamma_1+\gamma_2)} - \bar{\beta}\int_0^t e^{-C(t-s)}\,\mathrm{d}W_s^{\bar{\varepsilon}} + O(\lambda), \qquad \lambda \downarrow 0.$$

(b) As in part (a), we only calculate the trading rate of agent 1 since agent $2's$ trading rate is determined by market clearing. From Theorem 2.2.7 and Theorem B.1.1, it follows that

$$\dot{\bar{\varphi}}^{\lambda,1} = \left(\left[\xi^\varepsilon + \xi^\zeta\right] - C\bar{\varphi}^{\lambda,1}\right). \tag{4.1.21}$$

Using equations (4.1.18), (4.1.19), and (4.1.20), we see that

$$\frac{\xi^\varepsilon}{C} = (C+\delta)\mathbb{E}^{\bar{\mathrm{P}}}\left[\int_t^\infty e^{-(C+\delta)(s-t)}\frac{\varepsilon_t^1-\varepsilon_t^2}{\sigma(\gamma_1+\gamma_2)}\mathrm{d}s\Big|\mathcal{F}_t\right] = \frac{\varepsilon_t^1-\varepsilon_t^2}{\sigma(\gamma_1+\gamma_2)} + o(1),$$
$$\frac{\xi^\zeta}{C} = (C+\delta)\mathbb{E}^{\bar{\mathrm{P}}}\left[\int_t^\infty e^{-(C+\delta)(s-t)}\frac{\gamma_2\zeta_s^2-\gamma_1\zeta_s^2}{(\gamma_1+\gamma_2)}\mathrm{d}s\Big|\mathcal{F}_t\right] = \frac{\gamma_2\zeta_t^2-\gamma_1\zeta_t^2}{(\gamma_1+\gamma_2)} + O\left(\lambda^{\frac{1}{2}}\right),$$

as $\lambda \downarrow 0$. By substituting into (4.1.21) it follows that

$$\dot{\bar{\varphi}}_t^{\lambda,1} = -\sqrt{\frac{(\gamma_1+\gamma_2)\sigma^2}{4\lambda}}\tilde{\Delta}_t^1 + O(1), \quad \lambda \downarrow 0.$$

(c) It follows from part (a) and Theorem 2.2.20 that

$$\mu_t^\lambda = \mu_t^0 + \frac{\sigma^2(\gamma_1-\gamma_2)}{2}\Delta_t^1 - \lambda Z_t(\varepsilon_t^1-\varepsilon_t^2) \tag{4.1.22}$$

From Theorem B.1.1 we see that Z is derived from the martingale representation theorem with respect to the square integrable martingale $(M_t)_{t\geq 0}$ defined by

$$\mathrm{d}M_t = e^{(C+\delta)t}\,\mathrm{d}\bar{M}_t,$$

where

$$\bar{M}_t = \mathbb{E}^{\bar{\mathsf{P}}}\left[\int_0^\infty e^{-(C+\delta)s} B\left\{\frac{\gamma_2\zeta_s^2 - \gamma_1\zeta_s^1}{(\gamma_1+\gamma_2)} + \frac{\varepsilon_s^1 - \varepsilon_s^2}{\sigma(\gamma_1+\gamma_2)}\right\} \mathrm{d}s \middle| \mathcal{F}_t\right]. \tag{4.1.23}$$

We will compute $\bar{M}$ in parts in order to derive an expression for Z.

First, from Propositions 4.1.1 and C.1.2 it follows that

$$\begin{aligned}&\mathbb{E}^{\bar{\mathsf{P}}}\left[\int_0^\infty e^{-(C+\delta)s}\frac{\varepsilon_s^1 - \varepsilon_s^2}{\sigma(\gamma_1+\gamma_2)} \mathrm{d}s \middle| \mathcal{F}_t\right] \\ &= \frac{(\sigma_1-\sigma_2)}{(C+\delta+\kappa)}\frac{1}{\sigma(\gamma_1+\gamma_2)}\int_0^t e^{-(C+\delta)s} \mathrm{d}W_s^{\bar{\varepsilon}}.\end{aligned} \tag{4.1.24}$$

Next, suppose that $\sigma_1+\sigma_2 \neq 0$. Then Propositions 4.1.3 and C.1.11 show that

$$\begin{aligned}&\mathbb{E}^{\bar{\mathsf{P}}}\left[\int_0^\infty e^{-(C+\delta)s}\frac{\gamma_2\zeta_s^2 - \gamma_1\zeta_s^1}{(\gamma_1+\gamma_2)} \mathrm{d}s \middle| \mathcal{F}_t\right] \\ &= \bar{\beta}\Bigg(\frac{1}{(C+\delta)} - \frac{1}{(C+\delta+\theta)^2}\frac{2\sigma_1\sigma_2}{(\sigma_1+\sigma_2)} \\ &- \frac{1}{(C+\delta+\kappa)^2}\frac{1}{2}\frac{(\sigma_1-\sigma_2)^2}{(\sigma_1+\sigma_2)}\Bigg)\int_0^t e^{-(C+\delta)s} \mathrm{d}W_s^{\bar{\varepsilon}}.\end{aligned} \tag{4.1.25}$$

For the case where $\sigma_1 = -\sigma_2 = \sigma_0$, Propositions 4.1.3 and C.1.13 give us

$$\begin{aligned}&\mathbb{E}^{\bar{\mathsf{P}}}\left[\int_0^\infty e^{-(C+\delta)s}\frac{\gamma_2\zeta_s^2 - \gamma_1\zeta_s^1}{(\gamma_1+\gamma_2)} \mathrm{d}s \middle| \mathcal{F}_t\right] \\ &= \frac{\bar{\beta}}{(C+\delta)}\left(1 + \frac{\sigma_0^2}{(C+\delta+\theta)^2}\right)\int_0^t e^{-(C+\delta)s} \mathrm{d}W_s^{\bar{\varepsilon}}.\end{aligned} \tag{4.1.26}$$

By substituting (4.1.24), (4.1.25) and (4.1.20) into (4.1.23) we get

$$\bar{M}_t = \chi\int_0^t e^{-(C+\delta)s} \mathrm{d}W_s^{\bar{\varepsilon}},$$

where

$$\chi := \begin{cases}\dfrac{\sigma(\sigma_1-\sigma_2)}{4\lambda(C+\delta+\kappa)}\end{cases}$$

$$+\frac{\sigma^2(\beta_2\gamma_2-\beta_1\gamma_1)}{4\lambda}\left(\frac{1}{(C+\delta)}-\frac{1}{(C+\delta+\theta)^2}\frac{2\sigma_1\sigma_2}{(\sigma_1+\sigma_2)}-\frac{1}{(C+\delta+\kappa)^2}\frac{1}{2}\frac{(\sigma_1-\sigma_2)^2}{(\sigma_1+\sigma_2)}\right)\mathbb{1}_A$$
$$+\frac{\sigma^2(\beta_2\gamma_2-\beta_1\gamma_1)}{4\lambda}\left(\frac{1}{(C+\delta)}+\frac{\sigma_0^2}{(C+\delta)(C+\delta+\theta)^2}\right)\mathbb{1}_{A^c}\Bigg\},$$

and $A := \{\sigma_1+\sigma_2\neq 0\}$. Thus,

$$\mathrm{d}M_t = \chi\,\mathrm{d}W_t^{\bar{\varepsilon}},$$

from which it is clear that $Z_t = \chi$ for all $t \geq 0$. In particular, due to (3.1.10),

$$Z_t = \sqrt{\frac{(\gamma_1+\gamma_2)\sigma^2}{4\lambda}}\left\{\frac{[(\sigma_1-\sigma_2)-\sigma(\gamma_1\beta_1-\gamma_2\beta_2)]}{\sigma(\gamma_1+\gamma_2)}\right\}+O(1)$$

as $\lambda\downarrow 0$. □

Remark 4.1.5. (i) Despite the initially complicated formulas, the equilibrium portfolios, trading rates and returns processes are of the same form as their counterparts in Proposition 3.1.1 when $\lambda \ll 1$.

(ii) In the proof we derived the expressions in full generality before making the approximation for small λ, so we actually prove a slightly stronger statement. However, the approximations are all we need for the calculations in Theorem 4.2.1.

4.2 Utility loss

4.2.1 Planner's perspective

We now calculate the aggregate utility loss when transaction costs are small.

Theorem 4.2.1. *Suppose ζ^1,ζ^2 and $\varepsilon^1,\varepsilon^2$ are described by (4.0.2) and (4.0.1), respectively.*

(a) *The aggregate utility loss under the belief $(\bar{\varepsilon},\bar{\mathsf{P}})$ is*

$$U^{\bar{\varepsilon}} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{[\sigma^2(\gamma_1\beta_1-\gamma_2\beta_2)^2-(\sigma_1-\sigma_2)^2]}{(\gamma_1+\gamma_2)^{\frac{3}{2}}\,|\sigma|}\right\}+O(\lambda),$$

as $\lambda\downarrow 0$.

(b) *The post rebate aggregate utility loss under the belief* $(\bar{\varepsilon}, \bar{\mathsf{P}})$ *is*

$$\tilde{U}^{\bar{\varepsilon}} = U^{\bar{\varepsilon}} - \frac{\sqrt{\lambda}}{\delta}\left\{\frac{\left(\sigma[\beta_1\gamma_1 - \beta_2\gamma_2] - [\sigma_1 - \sigma_2]\right)^2}{2(\gamma_1+\gamma_2)^{\frac{3}{2}}\,|\sigma|}\right\} + O(\lambda),$$

as $\lambda \downarrow 0$.

Proof. From Proposition 2.4.3 and market clearing we know that the aggregate utility loss is of the form

$$\begin{aligned} U^{\bar{\varepsilon}} &= \mathbb{E}^{\bar{\mathsf{P}}}\left[\int_0^\infty e^{-\delta t}\sum_{n=1}^{2}\left\{(\Delta_t^n)\left(\frac{\gamma^n\sigma^2}{2}\Delta_t^n + \sigma\varepsilon_t^n\right) + \lambda\left(\dot{\bar{\varphi}}_t^{\lambda,n}\right)^2\right\}\mathrm{d}t\right] \\ &= \mathbb{E}^{\bar{\mathsf{P}}}\left[\int_0^\infty e^{-\delta t}\left\{\frac{(\gamma_1+\gamma_2)\,\sigma^2}{2}\left(\Delta_t^1\right)^2 + \sigma(\varepsilon_t^1 - \varepsilon_t^2)\Delta_t^1 + 2\lambda\left(\dot{\bar{\varphi}}_t^{\lambda,1}\right)^2\right\}\right]. \end{aligned} \tag{4.2.1}$$

Using Proposition 3.1.1, we see that (4.2.1) becomes

$$U^{\bar{\varepsilon}} = \mathbb{E}^{\bar{\mathsf{P}}}\left[\int_0^\infty e^{-\delta t}\left\{(\gamma_1+\gamma_2)\,\sigma^2\left(\tilde{\Delta}_t^1\right)^2 + \sigma(\varepsilon_t^1 - \varepsilon_t^2)\tilde{\Delta}_t^1\right\}\right] + O(\lambda)$$

as $\lambda \downarrow 0$. Similarly, one can show that the post rebate utility loss is given by

$$\tilde{U}^{\bar{\varepsilon}} = \mathbb{E}^{\bar{\mathsf{P}}}\left[\int_0^\infty e^{-\delta t}\left\{\frac{(\gamma_1+\gamma_2)\,\sigma^2}{2}\left(\tilde{\Delta}_t^1\right)^2 + \sigma(\varepsilon_t^1 - \varepsilon_t^2)\tilde{\Delta}_t^1\right\}\right] + O(\lambda),$$

as $\lambda \downarrow 0$. We complete the proof via direct calculation, found in the notebook `heterogeneous`. $\square$

Let's assume that the planner, in the interest of the agents, takes the *average belief* $(\bar{\varepsilon}, \bar{\mathsf{P}})$ in order to decide whether or not the transaction tax is beneficial. Assuming that the transaction tax is small, Theorem 4.2.1(a) makes it clear that the tax is beneficial when

$$|\sigma_1 - \sigma_2| - |\sigma(\gamma_1\beta_1 - \gamma_2\beta_2)| > 0. \tag{4.2.2}$$

Condition (4.2.2) encapsulates that a transaction tax penalises agents by impeding their ability to hedge effectively while rewarding them by dampening the effect of their heterogeneous beliefs. Note that it is not the heterogeneity between an agent's belief and the belief $(0, \mathbb{P})$ that is important but the heterogeneity between the agents' beliefs in general. Thus, (4.2.2) suggests that a small transaction tax is beneficial when all trading is non-fundamental, and there is heterogeneity in the speculative beliefs. Moreover, it shows that the tax is detrimental when all speculative beliefs are homogeneous, as in Section 3.1. Note that the effect of (4.2.2) occurs in proportion to the square root $\sqrt{\lambda}$ of the transaction tax, in line with the small-cost analysis seen in [57].

The effect of a naïve rebate on the aggregate utility loss is described succinctly by Theorem 4.2.1(b). It is pleasing that the qualitative effect of the transaction tax is largely unchanged by the (unrealistic) rebate. Thus, when there is purely speculative trading, one can advocate for a transaction tax in favour of the agents whilst accruing the tax as capital for other means.

Despite the planner's belief being absent from the integrand seen in (2.4.3), we need to take the expectation with respect to $\bar{\mathsf{P}}$ in order for condition (4.2.2) to remain intact. Fortunately, if we let the planner hold the 'correct' belief $(0, \mathbb{P})$, the phenomena persist for many parameter combinations.

Corollary 4.2.2. *Suppose that $\varepsilon^1, \varepsilon^2$ and ζ^1, ζ^2 are described by (4.0.1) and (4.0.2), respectively. Then the aggregate utility loss under the belief $(0, \mathbb{P})$ is*

$$
\begin{aligned}
&U^0 - U^{\bar{\varepsilon}} + O(\lambda) \\
&= \frac{\sqrt{\lambda}}{\delta}\left\{\frac{2\,(\beta_1\gamma_1\sigma - \beta_2\gamma_2\sigma - \sigma_1 + \sigma_2)(2\,\sigma_1\sigma_2 - \delta\sigma_1 - \delta\sigma_2 - 2\,\sigma_1\theta - 2\,\sigma_2\theta)(\sigma_1 - \sigma_2)}{(\delta - 2\,\sigma_1 + 2\,\theta)(\delta - 2\,\sigma_2 + 2\,\theta)(\gamma_1 + \gamma_2)^{\frac{3}{2}}\,|\sigma|}\right\},
\end{aligned}
$$

as $\lambda \downarrow 0$.

Proof. To proceed we rewrite the formulae for $\varepsilon^1, \varepsilon^2$ and $\tilde{\Delta}^1$ such that their stochastic components are with respect to the $\mathbb{P}$-Brownian motion W as opposed to the $\bar{\mathsf{P}}$-Brownian motion $W^{\bar{\varepsilon}}$. This has already been done for the agents' beliefs in Proposition 4.1.1(a). Thus, we need only find an appropriate expression for the process $\tilde{\Delta}^1$.

It follows by the definition of the belief $(\bar{\varepsilon}, \bar{\mathsf{P}})$, Theorem A.1.13, and Propositions 4.1.1(a) and C.1.1 that

$$
\begin{aligned}
&-\left[\frac{[(\sigma_1 - \sigma_2) - \sigma(\gamma_1\beta_1 - \gamma_2\beta_2)]}{\sigma(\gamma_1 + \gamma_2)}\right]^{-1} \tilde{\Delta}^1_t \\
&= \int_0^t e^{-C(t-s)}\,\mathrm{d}W^{\bar{\varepsilon}}_s
\end{aligned}
$$

$$
\begin{aligned}
&= \int_0^t e^{-C(t-s)} \left(\mathrm{d}W_s - \int_0^t \left(\varepsilon_s^1 + \varepsilon_s^2\right) \, \mathrm{d}s \right) \\
&= \int_0^t e^{-C(t-s)} \mathrm{d}W_s \\
&- \sigma_1 \int_0^t e^{-C(t-s)} \int_0^s e^{-\kappa_1(s-r)} \, \mathrm{d}W_r \, \mathrm{d}s - \sigma_2 \int_0^t e^{-C(t-s)} \int_0^s e^{-\kappa_2(s-r)} \, \mathrm{d}W_r \, \mathrm{d}s \\
&= \int_0^t e^{-C(t-s)} \mathrm{d}W_s \\
&\qquad + \frac{\sigma_1}{(C-\kappa_1)} \left(\int_0^t e^{-C(t-s)} \, \mathrm{d}W_s - \int_0^t e^{-\kappa_1(t-s)} \, \mathrm{d}W_s \right) \\
&\qquad\qquad + \frac{\sigma_2}{(C-\kappa_2)} \left(\int_0^t e^{-C(t-s)} \, \mathrm{d}W_s - \int_0^t e^{-\kappa_2(t-s)} \, \mathrm{d}W_s \right). \qquad (4.2.3)
\end{aligned}
$$

We are now in a position to conclude the proof by calculating U^0 in the same manner as we calculated $U^{\bar{\varepsilon}}$. The calculations can be found in the notebook `heterogeneous`. □

Corollary 4.2.2 shows that when $0 < \sqrt{\lambda} < \delta \ll 1$, the difference between the aggregate utility loss under the beliefs $(0, \mathbb{P})$ and $(\bar{\varepsilon}, \bar{\mathsf{P}})$ is approximately

$$
\frac{\sqrt{\lambda}}{\delta} \left\{ \frac{(\beta_1\gamma_1\sigma - \beta_2\gamma_2\sigma - [\sigma_1 - \sigma_2])(\sigma_1\sigma_2 - \sigma_1\theta - \sigma_2\theta)(\sigma_1 - \sigma_2)}{(\gamma_1 + \gamma_2)^{\frac{3}{2}} (\theta - \sigma_1)(\theta - \sigma_2) \, |\sigma|} \right\}. \qquad (4.2.4)
$$

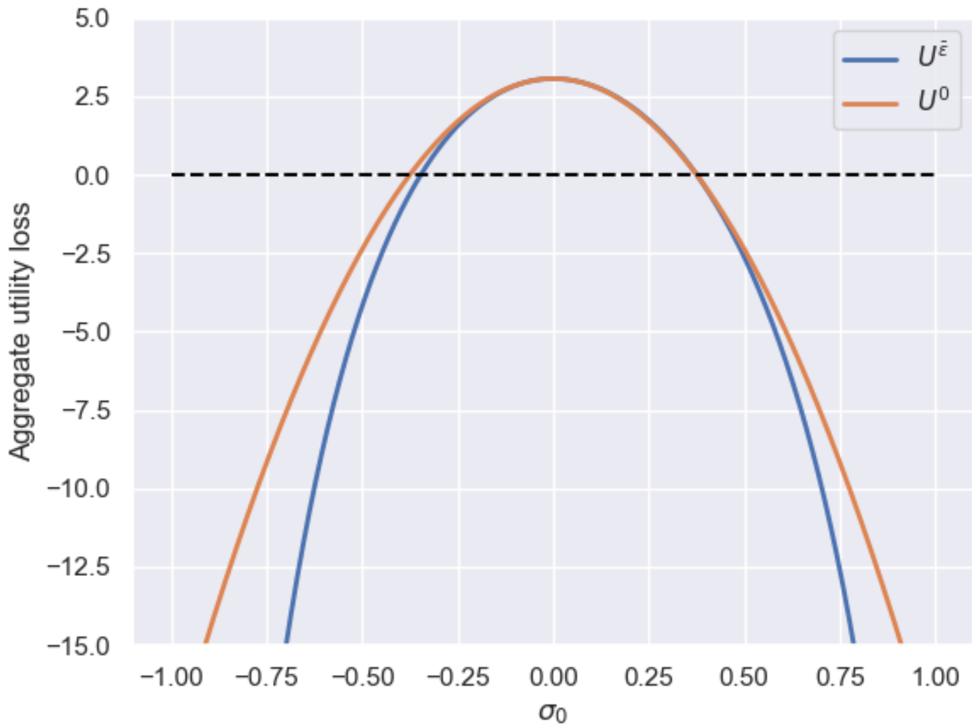
In this case, we see that if the mean reversion speed θ seen in the dynamics of the beliefs (4.1.1) is large enough such that

$$
\theta \gg 1 \vee \sigma_1 \vee \sigma_2, \qquad (4.2.5)
$$

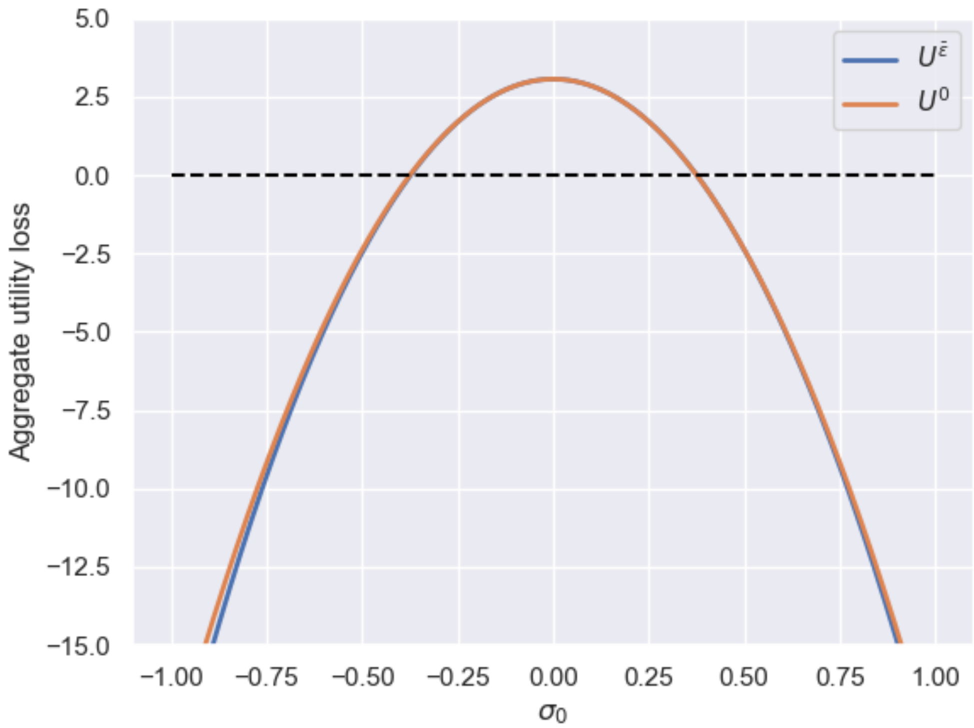
then (4.2.4) will be decreasing as θ grows. In particular, this means that the aggregate loss under the beliefs $(0, \mathbb{P})$ and $(\bar{\varepsilon}, \bar{\mathsf{P}})$ will be similar when (4.2.5) holds. This is no surprise since the covariance function of the beliefs tends to zero as in θ increases.

Despite not having an ideal formula such as (4.2.2), plots of U^0 show that a small positive transaction tax is beneficial for all parameter combinations where the agents have sufficiently heterogeneous beliefs, i.e. for large enough $|\sigma_1 - \sigma_2|$. The notebook `heterogeneous` contains the function *aggre-*

gate_utility_loss, which creates an interactive plot in order to test such combinations; figures 4.1a and 4.1b are examples of such plots. Importantly, these plots match the intuition of Section 2.4.1 and the condition (4.2.2).



(a) The aggregate utility loss under the average belief as a function of the heterogeneity in beliefs.



(b) The aggregate utility loss under the average belief as a function of the heterogeneity in beliefs.

Figure 4.1: We assume that $\sigma_1 = -\sigma_2 = \sigma_0$, so that $2\sigma_0$ measures the heterogeneity in beliefs. The plots show the aggregate utility loss under the beliefs $(\bar{\varepsilon}, \bar{\mathsf{P}})$ and $(0, \mathbb{P})$ against σ_0.
(a) Here we plot under the parameter combination

$$\theta = 1.2; \ \delta = 0.1; \ \sigma = \gamma_2 = 1; \ \beta_1 = -\beta_2 = \gamma_1 = 0.5.$$

(b) We plot the same as in part (a) except $\theta = 5$. As is suggested by Corollary 4.2.2, the difference between U^0 and $U^{\bar{\varepsilon}}$ is smaller than in part (a).

We finish this subsection by noting that, for small transaction costs, the direct utility loss has the same expression under all the relevant beliefs we have considered so far.

Lemma 4.2.3. *Suppose ζ^1, ζ^2 and $\varepsilon^1, \varepsilon^2$ are described by (4.0.2) and (4.0.1). Then*

$$U_{\mathrm{d}}^{n,\varepsilon} = \frac{\sqrt{\lambda}}{\delta} \left\{ \frac{(\beta_1\gamma_1\sigma - \beta_2\gamma_2\sigma - \sigma_1 + \sigma_2)^2}{(\gamma_1+\gamma_2)^{\frac{5}{2}} \sigma^2 |\sigma|} \right\} + O(\lambda)$$

as $\lambda \downarrow 0$, for $n = 1, 2$ and any belief $(\varepsilon, \mathsf{P}(\varepsilon)) \in \{(0, \mathbb{P}), (\bar{\varepsilon}, \bar{\mathsf{P}}), (\varepsilon^1, \mathsf{P}^1), (\varepsilon^2, \mathsf{P}^2)\}$.

Proof. Let $(n, m) \in \{(1, 2), (2, 1)\}$. It follows from Theorem A.1.13 and Propositions 3.1.1(a) and C.1.1 that

$$\begin{aligned}
&- \left[\frac{[(\sigma_1 - \sigma_2) - \sigma(\gamma_1\beta_1 - \gamma_2\beta_2)]}{\sigma(\gamma_1+\gamma_2)} \right]^{-1} \tilde{\Delta}_t^1 \\
&= \int_0^t e^{-C(t-s)} \, \mathrm{d}W_s^{\bar{\varepsilon}} \\
&= \int_0^t e^{-C(t-s)} \, \mathrm{d}W_s^n + \frac{1}{2} \int_0^t e^{-C(t-s)} (\varepsilon_s^1 - \varepsilon_s^2) \, \mathrm{d}s \\
&= \int_0^t e^{-C(t-s)} \, \mathrm{d}W_s^n + \frac{(\sigma_1 - \sigma_2)}{2} \int_0^t e^{-C(t-s)} \int_0^s e^{-\kappa_m(s-r)} \, \mathrm{d}W_r^1 \, \mathrm{d}s \\
&= \int_0^t e^{-C(t-s)} \, \mathrm{d}W_s^n - \frac{1}{2} \frac{(\sigma_1 - \sigma_2)}{(C - \kappa_m)} \left\{ \int_0^t e^{-C(t-s)} \, \mathrm{d}W_s^n - \int_0^s e^{-\kappa_m(t-s)} \, \mathrm{d}W_s^n \right\} \\
&= \left(1 - \frac{1}{2} \frac{(\sigma_1 - \sigma_2)}{(C - \kappa_m)} \right) \int_0^t e^{-C(t-s)} \, \mathrm{d}W_s^n + \frac{1}{2} \frac{(\sigma_1 - \sigma_2)}{(C - \kappa_m)} \int_0^s e^{-\kappa_m(t-s)} \, \mathrm{d}W_s^n.
\end{aligned} \tag{4.2.6}$$

Noting that

$$\mathbb{E}\left[\int_0^\infty e^{-\delta t} \left(\int_0^t e^{-A(t-s)} \, \mathrm{d}W_s \right) \left(\int_0^t e^{-B(t-s)} \, \mathrm{d}W_s \right) \, \mathrm{d}t \right] = \frac{1}{(A+B)\delta + \delta^2},$$

it follows from (4.2.3), (4.2.6) and Proposition 4.1.4 that

$$\begin{aligned}
&\lambda \mathbb{E}^{\mathbb{Q}} \left[\int_0^\infty e^{-\delta t} \left(\dot{\varphi}_t^{\lambda,1} \right)^2 \, \mathrm{d}t \right] \\
&= \mathbb{E}^{\mathbb{Q}} \left[\int_0^\infty e^{-\delta t} \left(\tilde{\Delta}_t^1 \right)^2 \, \mathrm{d}t \right] + O(\lambda) \\
&= \frac{\sqrt{\lambda}}{\delta} \left\{ \frac{(\beta_1\gamma_1\sigma - \beta_2\gamma_2\sigma - \sigma_1 + \sigma_2)^2}{(\gamma_1+\gamma_2)^{\frac{5}{2}} \sigma^2 |\sigma|} \right\} + O(\lambda)
\end{aligned} \tag{4.2.7}$$

as $\lambda \downarrow 0$ for $\mathbb{Q} \in \{\mathbb{P}, \bar{\mathsf{P}}, \mathsf{P}^1, \mathsf{P}^2\}$. $\square$

4.2.2 Idiosyncratic perspective

We now check how the agents perceive the transaction tax under their own idiosyncratic beliefs.

Proposition 4.2.4. *Suppose ζ^1, ζ^2 and $\varepsilon^1, \varepsilon^2$ are described by (4.0.2) and (4.0.1), respectively.*

(a) *The portfolio and direct losses are given by*

$$U_{\mathrm{p}}^{n,\varepsilon^n} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{(\beta_1\gamma_1\sigma - \beta_2\gamma_2\sigma - \sigma_1 + \sigma_2)^2\gamma_n}{2\,(\gamma_1+\gamma_2)^{\frac{5}{2}}\,|\sigma|}\right\} + O(\lambda)$$
$$U_{\mathrm{d}}^{n,\varepsilon^n} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{(\beta_1\gamma_1\sigma - \beta_2\gamma_2\sigma - \sigma_1 + \sigma_2)^2}{4\,(\gamma_1+\gamma_2)^{\frac{3}{2}}\,|\sigma|}\right\} + O(\lambda)$$

as $\lambda \downarrow 0$, for $n = 1, 2$.

(b) *When $\beta_1 = \beta_2 = 0$, the return loss is given by*

$$U_{\mathrm{r}}^{n,\varepsilon^n} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{[\gamma_n(\sigma_n - \sigma_m) + (\kappa_n+\delta)(\gamma_n - \gamma_m) + \gamma_n\kappa_n - \gamma_m\kappa_m](\sigma_1-\sigma_2)^2}{2(\delta+2\kappa_n)(\gamma_1+\gamma_2)^{\frac{5}{2}}\,|\sigma|}\right\} + O(\lambda)$$

as $\lambda \downarrow 0$, for $(n,m) \in \{(1,2),(2,1)\}$.

(c) *When $\beta_1 = \beta_2 = 0$, the utility loss is given by*

$$U^{n,\varepsilon^n} = \frac{\sqrt{\lambda}}{\delta}\left\{\frac{[\delta(5\,\gamma_n - \gamma_m) + 2(5\,\gamma_n\kappa_n - \gamma_m\kappa_m) + 2\,\gamma_n(\sigma_n - \sigma_m)](\sigma_1-\sigma_2)^2}{4\,(\delta+\kappa_n)(\gamma_1+\gamma_2)^{\frac{5}{2}}\,|\sigma|}\right\} + O(\lambda)$$

as $\lambda \downarrow 0$, for $(n,m) \in \{(1,2),(2,1)\}$.

Proof. The calculations are explicit in the notebook `heterogeneous`. $\square$

Under the agents' own beliefs, the portfolio and direct losses of each agent have succinct forms for small λ, as they are multiples of

$$\mathbb{E}\left[\int_0^\infty e^{-\delta t}\left(\tilde{\Delta}_t^1\right)^2 \, \mathrm{d}t\right].$$

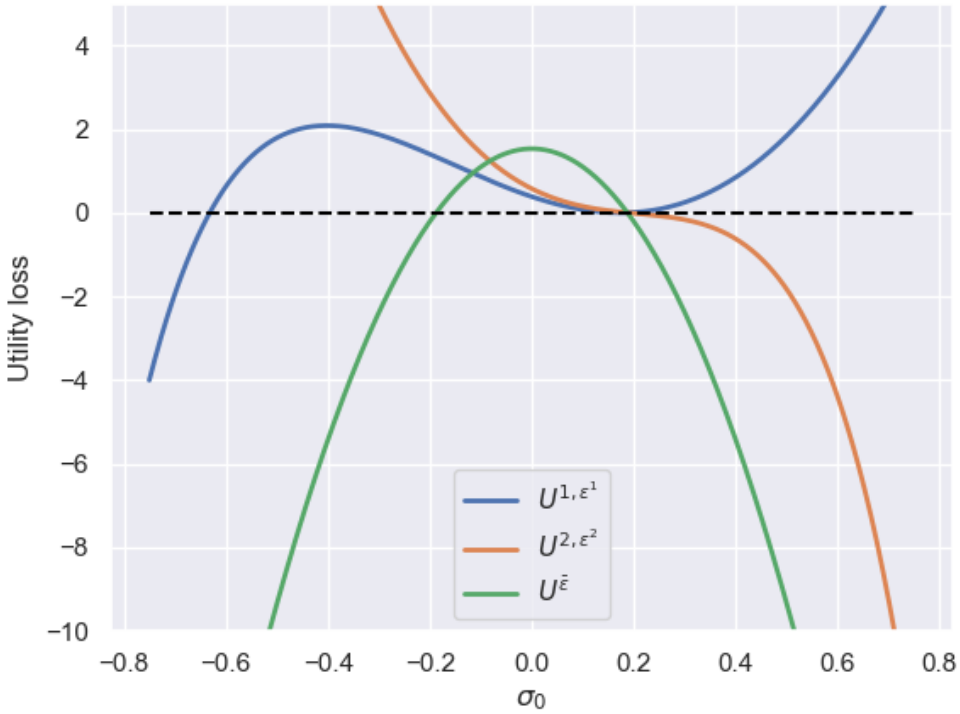
Since they are positive, the return losses will determine whether or not the tax is deemed beneficial by an individual agent, similar to the case with homogeneous beliefs.

When the transaction tax is small, expressions for the return losses are not succinct unless one supposes that the agents have purely speculative trading motives (i.e. $\beta_1 = \beta_2 = 0$). In this case, Proposition 4.2.4(b) shows that when the heterogeneity between the agent's beliefs is small, the return loss is given by

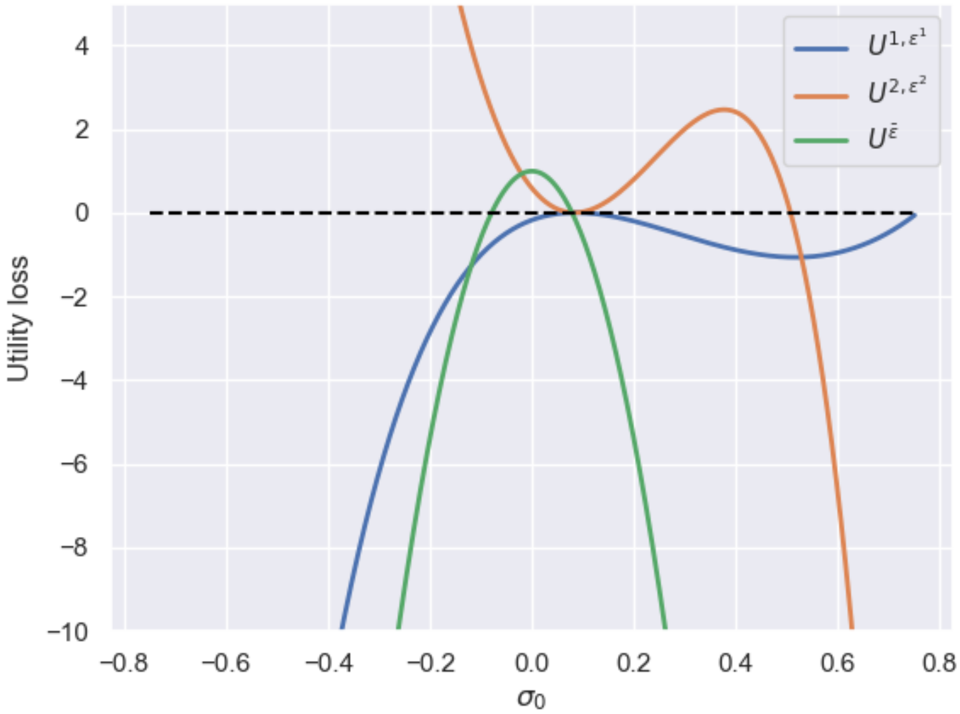
$$U_r^{n,\varepsilon^n} \approx \frac{\sqrt{\lambda}}{\delta}\left\{\frac{[(\kappa_n+\delta)(\gamma_n-\gamma_m)+\gamma_n\kappa_n-\gamma_m\kappa_m](\sigma_1-\sigma_2)^2}{2(\delta+2\kappa_n)(\gamma_1+\gamma_2)^{\frac{5}{2}}\,|\sigma|}\right\}+O(\lambda),$$

as $\lambda \downarrow 0$ for $(n,m) \in \{(1,2),(2,1)\}$. In particular, this shows that when the transaction tax is small and $\gamma_1 \ll \gamma_2$, we have $U_r^{1,\varepsilon^1} < 0$ and thus agent $n = 1$ is offered favourable returns. Indeed, by the same argument, one can see from Proposition 4.2.4(c) that if there is a large enough discrepancy between the risk aversions, then U^{1,ε^1} is negative. Thus, we see similar phenomena to those seen in the homogeneous beliefs case, despite the endowments being absent.

Although the expressions are complex, we can easily plot U^{1,ε^1} and U^{2,ε^2} in full generality for a small transaction tax, using the function *total_loss* found in the notebook `heterogeneous`. Figures 4.2a and 4.2b give examples of such plots. Note that Figure 4.2b shows that it is possible for $U^{\bar{\varepsilon}}$, U^{1,ε^1} and U^{2,ε^2} to all be negative when the transaction tax is small. Although there is seemingly little economic interpretation for the parameter combination, the planner and the agents would deem the transaction tax beneficial in this case.



(a) The idiosyncratic utiliy losses and the aggregate utility loss (under the average belief) as functions of the heterogeneity in the agents' beliefs.



(b) The idiosyncratic utiliy losses and the aggregate utility loss (under the average belief) as functions of the heterogeneity in the agents' beliefs.

Figure 4.2: We assume that $\sigma_1 = -\sigma_2 = \sigma_0$. The plots are examples of the aggregate utility loss $U^{\bar{\varepsilon}}$ and the idiosyncratic utility losses U^{1,ε^1} and U^{2,ε^2} when $\lambda \ll 1$, as functions of σ_0, for specific parameter combinations. Note that if we changed the sign of σ, we would get the same plots reflected in the vertical axis at zero.
(a) Here we plot under the parameter combination

$$\theta = 1.2;\ \delta = 0.1;\ \sigma = \gamma_2 = 1;\ \beta_1 = -\beta_2 = \gamma_1 = 0.5.$$

(b) We plot the same as in part (a) except $\gamma_1 = 0.03$ and $\gamma_2 = 0.6$. The plot shows that $U^{\bar{\varepsilon}}$, U^{1,ε^1} and U^{2,ε^2} are all negative when $0.6 \lesssim \sigma_0 \lesssim 0.8$.

4.3 Optimal transaction tax

Proposition 4.1.4 shows that when agents have heterogeneous beliefs, their optimal strategies behave analogously to their counterparts in the homogeneous case. Notably, the strategies' behaviour is understandable. Thus, keeping the transaction tax small from a planner's perspective makes sense to avoid introducing odd behaviour. Moreover, if the transaction tax is small and the planner takes the average belief of the agents, then (4.2.2) gives a coherent condition as to when it is beneficial.

This begs the question of whether a transaction tax exists which is small enough to keep the agent's behaviour sensible but not so small that we forgo additional benefits to the agents. Since we have considered the effects of the transaction tax at the order of $\lambda^{\frac{1}{2}}$, we will say that these effects diminish once the effects of order λ dominate. To investigate when this happens, we begin by assuming that $\beta_1 = \beta_2 = 0$, since in this case, the transaction tax is always beneficial for small enough λ. To proceed, we generalise Theorem 4.2.1 to include higher order terms in λ in our expansion of $U^{\bar{\varepsilon}}$.

Corollary 4.3.1. *Suppose $\zeta^1 = \zeta^2 = 0$ and $\varepsilon^1, \varepsilon^2$ are described by (4.0.1). Then the aggregate utility loss under the belief $(\bar{\varepsilon}, \bar{\mathsf{P}})$ is given by*

$$U^{\bar{\varepsilon}} = \bar{U} + O\left(\lambda^{\frac{3}{2}}\right)$$

as $\lambda \downarrow 0$, where

$$\bar{U} = -\frac{\sqrt{\lambda}}{\delta}\left\{\frac{(\sigma_1 - \sigma_2)^2}{(\gamma_1 + \gamma_2)^{\frac{3}{2}}\,|\sigma|}\right\} + \frac{\lambda}{\delta}\left\{\frac{(3\,\delta + 2(\kappa_1 + \kappa_2))(\sigma_1 - \sigma_2)^2}{(\gamma_1 + \gamma_2)^2\sigma^2}\right\}.$$

Proof. In the proof of Proposition 4.1.4 we get a general expression for $\varphi^{\lambda,1}$ before approximating for small λ. We can thus include higher order terms (in λ) into our calculations. Thus, the proof is analogous to the proof of Theorem 4.2.1. The calculations are explicit in the notebook `optimal_tax`. □

Using Corollary 4.3.1, we can see clearly that the second order effects (i.e. of order λ) begin to dominate after the minimum of $\bar{U}$. This takes place when the transaction tax is equal to

$$\lambda^* := \frac{(\gamma_1 + \gamma_2)\sigma^2}{4\,(3\,\delta + 2\,\sigma_1 + 2\,\sigma_2 + 4\,\theta)^2},$$

resulting in the minimum

$$-\frac{1}{\delta}\left\{\frac{(\sigma_1-\sigma_2)^2}{4\,(3\,\delta+2\,\sigma_1+2\,\sigma_2+4\,\theta)(\gamma_1+\gamma_2)}\right\}.$$

Despite only being optimal in a loose sense, λ^* behaves as expected with respect to the other parameters. For instance,

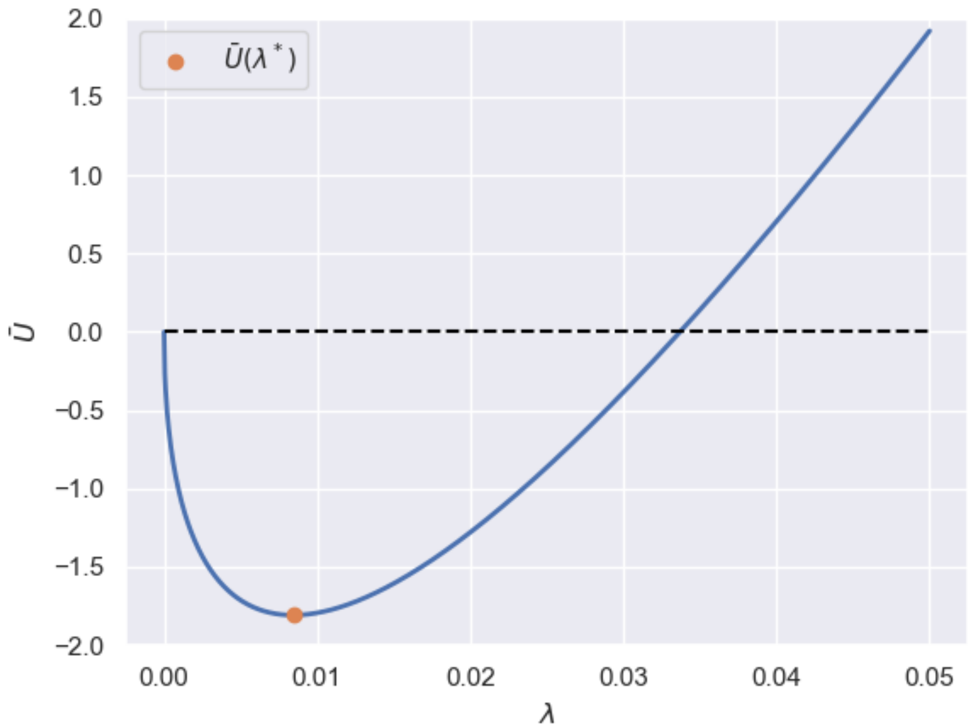
$$\lambda^* \propto (\gamma_1+\gamma_2). \tag{4.3.1}$$

This is explained by looking at the goal functional (2.1.7), from which we see that without the endowment stream present, the risk aversion parameters become holding cost penalties. Indeed, this is the model and rationale taken in [57]. Thus, agent n has a more significant motive to offload the risky asset the larger the parameter γ_n. Since we assume that agents have heterogeneous beliefs, this trading is detrimental from the planner's perspective, whence (4.3.1) is to be expected. Furthermore, one can see that

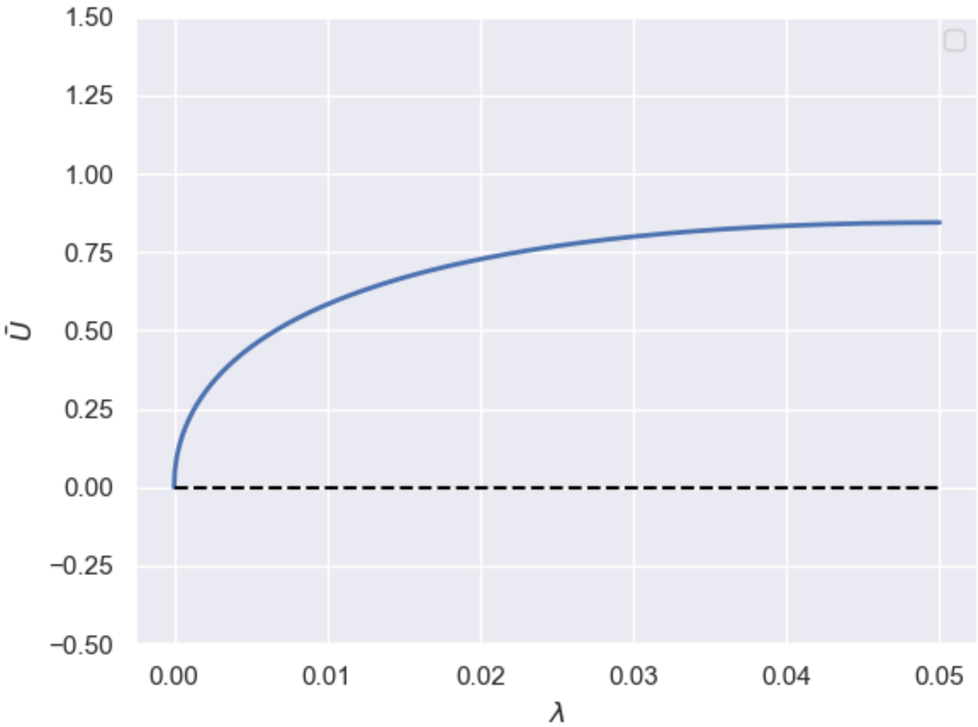
$$\lambda^* \propto \frac{1}{\theta}.$$

This encapsulates that as θ increases, the processes ε_1 and ε_2 converge to zero, eliminating the heterogeneity in beliefs and making the transaction tax unnecessary.

One can also compute a similar result to Corollary 4.3.1 where β_1 and β_2 are not necessarily zero. In this case, we do not get a pleasant expression for $\bar{U}$, but we can easily make plots, such as Figures 4.3a and 4.3b. Such plots suggest that $\bar{U}$ has local a minimum with respect to λ when condition (4.2.2) is satisfied. This warrants an effort to search for a more general result that characterises optimal taxation where we do not specify ζ^1, ζ^2 and $\varepsilon^1, \varepsilon^2$. However, one may have to include a naïve rebate in order to do so, as we will not necessarily get the practical dynamics for $\dot{\bar{\varphi}}_t^{\lambda,1}$ seen in Proposition 4.1.4, which gave us a valuable form for the direct utility losses and allowed us to calculate $U^{\bar{\varepsilon}}$ so quickly.



(a) $\bar{U}$ plotted as a function of λ in a scenario where agents have speculative beliefs and $\beta_1 = \beta_2 = 0$.



(b) $\bar{U}$ plotted as a function of λ in a scenario where the agents' hedging needs are dominant.

Figure 4.3: Both figures plot $\bar{U}$ as a function of λ, and are computed using the function *optimal* in the notebook `optimal_tax`.
(a) Here we plot under the parameter combination

$$\theta = 1.5;\ \delta = 0.1;\ \gamma_2 = \gamma_1 = 0.5;\ \beta_1 = -\beta_2 = 0.25;\ \sigma = \sigma_1 = -\sigma_2 = 1.$$

Note that heterogeneity in beliefs dominate the hedging needs of the agents according to condition (4.2.2) and thus a tax exists that minimises $\bar{U}$.
(b) Here we plot under the parameter combination

$$\theta = 1.5;\ \delta = 0.1;\ \gamma_2 = \gamma_1 = 0.5;\ \sigma_1 = -\sigma_2 = 0.25;\ \sigma = \beta_1 = -\beta_2 = 1.$$

Unlike part (a), the hedging needs of the agents dominate the heterogeneity in their beliefs, whence a transaction tax is detrimental.

Chapter 5

Vague Convergence

In this chapter, we depart from the economic model introduced in Chapter 2 and consider abstract theory about the convergence of real-valued measures. We describe the relationship between this content and the economic model in Chapter 6.

The so-called weak topology on the space of measures is used frequently in the many sub-fields of analysis and probability theory [30, 11, 15], but there are times when one needs a strictly different topology, such as the vague topology. Indeed, it is a subtle point that vague convergence is at the heart of König's approach to the proof of Karamata's Tauberian theorem (see e.g. Feller [32, XIII.5, Theorem 1]), a seminal result in probability theory. The proof relies on the equivalence between the vague convergence of finite positive measures and the pointwise convergence of their distribution functions (at continuity points of the limiting measure). As it turns out, the exact relationship between vague convergence and the convergence of their distribution functions is seemingly absent from the literature when the measures in question are real-valued. We fill this gap in our preprint [41], and include the arguments in this chapter.

We aim to make the content self-contained, so we will briefly introduce the theory of real-valued Radon measures and showcase why they are a vital element of modern analysis. We will then concentrate on notions of convergence of Radon measures, highlighting the equivalences and non-equivalences between vague and weak convergence. Finally, we will pinpoint the relationship between the vague convergence of Radon measures and their distribution functions, allowing us to derive a Tauberian condition for the real-valued version of Karamata's theorem in Chapter 6.

5.1 Convergence of measures

For any topological space (Ω, τ), we denote the Borel σ-algebra generated by the open sets of τ by $\mathscr{B}(\Omega, \tau)$. When understood, we will omit τ from the notation.

For a signed measure μ on $(\Omega, \mathscr{B}(\Omega))$, we denote its Hahn-Jordan decomposition by $\mu = \mu^+ - \mu^-$, and its associated variation measure by $|\mu| := \mu^+ + \mu^-$. The *total variation* of a signed measure μ is denoted by $\|\mu\| := |\mu|(\Omega)$, and we say that μ is *finite* if $\|\mu\| < \infty$.

For a topological space Ω, we let $C(\Omega)$ be the space of all continuous $\mathbb{R}$-valued functions on Ω, $C_b(\Omega)$ the subspace of all $f \in C(\Omega)$ such that f is bounded, $C_0(\Omega)$ the subspace of all $f \in C(\Omega)$ such that for any $\varepsilon > 0$, there exists a compact set $K_\varepsilon \in \mathscr{B}(\Omega)$ with $|f| < \varepsilon$ on K_ε^c, and $C_c(\Omega)$ the subspace of all $f \in C(\Omega)$ such that f has compact support. We have the inclusions $C_c(\Omega) \subseteq C_0(\Omega) \subseteq C_b(\Omega) \subseteq C(\Omega)$, and they are all equal to each other when the underlying space is compact.

5.1.1 Real-valued measures

The theory of Radon measures has its roots in the early 20th century via the works of Johann Radon, John von Neumann, and Stefan Banach. Radon was a Czech mathematician who introduced the concept of Radon measures in his 1913 paper [51], which set the foundation for this theory. Neumann and Banach significantly contributed to the theory, especially in the fields of functional analysis and measure theory, which paved the way for the development of modern integration theory, including the approach put forward by Bourbaki in the volume *Intégration* [18].

Definition 5.1.1. Let Ω be a topological Hausdorff space and let μ be a measure on $(\Omega, \mathscr{B}(\Omega))$.

(i) We say that μ is *inner regular* at $B \in \mathscr{B}(\Omega)$ if

$$\mu(B) = \sup\{\mu(F) : F \subseteq B,\ F \text{ closed }\}.$$

Moreover, we say that μ is inner regular if it is inner regular at all $B \in \mathscr{B}(\Omega)$.

(ii) We say that μ is *tight* at $B \in \mathscr{B}(\Omega)$ if

$$\mu(B) = \sup\{\mu(K) : K \subseteq B,\ K \text{ compact }\}.$$

Moreover, we say that μ is tight if it is tight at all $B \in \mathscr{B}(\Omega)$.

(iii) We say that μ is *outer regular* at $B \in \mathscr{B}(\Omega)$ if

$$\mu(B) = \inf\{\mu(G) : G \supseteq B,\ G \text{ open }\}.$$

Moreover, we say that μ is outer regular if it is outer regular at all $B \in \mathscr{B}(\Omega)$.

(iv) We say that μ is regular if it is inner and outer regular.

(v) We say that μ is a Borel measure if it is locally finite.

(vi) We say that μ is a Radon[1] measure if it is a tight Borel measure.

Some authors say that a measure is inner regular when they mean tight, despite there being a distinction between the two notions; see [14, Example 7.1.6]. Fortunately, this distinction does not matter when the underlying space is Polish. Thus, as is standard knowledge, all probability measures on $\mathbb{R}$ are both Radon and regular.

Theorem 5.1.2. *Let Ω be a metric space. Then every Borel measure μ on Ω is regular. If Ω is complete and separable, then the measure μ is Radon.*

Proof. See [14, Theorem 7.1.7]. □

Although one usually thinks of measures as being positive, in the sequel we will be dealing with real-valued Radon measures.

[1]A Radon measure is sometimes referred to as a tight measure.

Definition 5.1.3. Let (Ω, τ) be a a topological Hausdorff space. A set function $\mu : \mathscr{B}(\Omega) \to \mathbb{R} \cup \{\infty\}$ is called a *real-valued Radon measure* if there exist mutually singular Radon measures μ^+ and μ^- such that for any $A \in \mathscr{B}(\Omega)$,

$$\mu(A) := \begin{cases} \mu^+(A) - \mu^-(A) & \text{if both } \mu^+(A), \mu^-(A) < \infty, \\ +\infty & \text{otherwise.} \end{cases}$$

We set $|\mu| := \mu^+ + \mu^-$ and $\|\mu\| := |\mu|(\Omega)$. We say that μ is a *finite real-valued Radon measure*if $\|\mu\| < \infty$. We denote the space of real-valued Radon measures on $(\Omega, \mathscr{B}(\Omega))$ by $\mathcal{M}^*(\Omega)$, the set of all finite real-valued Radon measures on $(\Omega, \mathscr{B}(\Omega))$ by $\mathcal{M}(\Omega)$, the set of all finite positive Radon measures by $\mathcal{M}^+(\Omega)$ and the set of all probability measures on $(\Omega, \mathscr{B}(\Omega))$ by $\mathcal{M}_1^+(\Omega)$, so $\mathcal{M}^* \supset \mathcal{M} \supset \mathcal{M}^+ \supset \mathcal{M}_1^+$.

Remark 5.1.4. (a) For a measurable space $(\Omega, \mathcal{F})$, one defines a *signed measure* as a σ-additive set function $\mu : \mathcal{F} \to [-\infty, \infty]$ that maps the empty-set to zero. In which case μ can only ever attain one of the values in $\{-\infty, \infty\}$. According to this definition, signed measures can exist such that $|\mu|$ is Radon, and μ is not a real-valued measure. However, we may be satisfied that μ will equal a unique real-valued measure on compact sets, so we do not incur much trouble. Indeed, authors such as Berg [9, Chapter 2.2] merely define signed Radon measures as the difference between Radon measures, which largely agrees with Definition 5.1.3

(b) It is important to note that elements of $\mathcal{M}(\Omega)$ are signed measures in the traditional sense.

We momentarily restrict ourselves the case where Ω is a *locally compact Hausdorff* (LCH) space and make two additional definitions in order for us to state the famous *Riesz Representation Theorem.*

Definition 5.1.5. Let Ω be an LCH space.

(i) For a compact subset $K \subset \Omega$, we let $C^{(K)}(\Omega)$ be the space of all continuous functions whose support is contained in K.

(ii) We say that a linear form F on $C_c(\Omega)$ is *locally-continuous* if for any compact set $K \subset \Omega$, the restriction $F|_{C^{(K)}(\Omega)}$ is continuous. We denote the space of all locally-continuous linear mappings on $C_c(\Omega)$ by $(C_c(\Omega))'_{\text{loc}}$.

Remark 5.1.6. Let Ω be a LCH space.

(i) For any compact set $K \subset \Omega$, $C^{(K)}(\Omega)$ is a Banach space with respect to the supremum norm and

$$C_c(\Omega) = \bigcup_{K \subset \Omega : K \text{ compact}} C^{(K)}(\Omega).$$

(ii) The elements of $(C_c(\Omega))'_{\text{loc}}$ are sometimes called *relatively bounded* linear forms [18, Chapter III, §1.5].

(iii) The continuous dual space $(C_c(\Omega))'$ is clearly a subset of $(C_c(\Omega))'_{\text{loc}}$. They are equal when the underlying space Ω is compact.

Definition 5.1.7. Let $\mu \in \mathcal{M}^*(\Omega)$. For any subset $V \subset C(\Omega)$ we let

$$V_\mu := \left\{ f \in V : \int_\Omega |f| \ \mathrm{d}\,|\mu| < \infty \right\}.$$

We then define the mapping $I_\mu : V_\mu \to \mathbb{R}$ by

$$I_\mu(f) := \int_\Omega f \ \mathrm{d}\mu.$$

When understood, we will not mention the domain of I_μ.

Theorem 5.1.8 (Riesz-Markov-Kakutani Representation Theorem)**.** *Let Ω be a LCH space.*

(a) *The mapping $\mu \mapsto I_\mu$, where $I_\mu : C_0(\Omega) \to \mathbb{R}$, defines an isometric isomorphism from $\mathcal{M}(\Omega)$ to $(C_0(\Omega))^*$.*

(b) *The mapping $\mu \mapsto I_\mu$, where $I_\mu : C_c(\Omega) \to \mathbb{R}$, defines an isomorphism from $\mathcal{M}^*(\Omega)$ to $(C_c(\Omega))^*_{\text{loc}}$.*

Proof. For (a) see [33, Theorem 7.17] and for (b) see [9, Chapter 2, Theorem 2.5]. □

Remark 5.1.9. The stated iteration of the Riesz representation theorem is actually a corollary to a more general intergal representation theorem due to Pollard and Toposøe [60]. Although not included here, the more general theorem also implies the abstract extension theorem of Daniell as a corollary.

The Riesz representation theorem gives the study of measures an excellent topological framework which led several authors [33, 18, 71] to restrict the definition of Radon measures to LCH spaces. For example, this approach is taken by Bourbaki, who defines Radon measures explicitly as the continuous linear functionals on the space $C_c(\Omega)$ [18, Chapter 3, Definition 1]. Unfortunately, a restriction to LCH spaces is not satisfactory in probability theory, partly because no infinite-dimensional topological vector space can be locally compact [62, Theorem 1.22]. Instead, one usually considers the class of Polish spaces, which provide 'the simplest and more interesting' case when studying the convergence of random processes (Prokhorov [61, Introduction]). In this case, one can utilize the theory of Radon measures on arbitrary Hausdorff spaces as developed by Schwarz [65] and Toposøe [71].

5.1.2 Weak and vague convergence

To allow for a unified discourse, unless otherwise stated, we will assume that Ω is a metrisable space. We now come to the key definition of this chapter.

Definition 5.1.10.

(a) Let $\mu_1, \mu_2, \dots, \mu \in \mathcal{M}^*(\Omega)$. We say that μ_n converges *vaguely* to μ if $I_{\mu_n}(f) \to I_\mu(f)$ for all $f \in C_c(\Omega)$, and we write

$$\operatorname*{v-lim}_{n\to\infty} \mu_n = \mu.$$

(b) Let $\mu_1, \mu_2, \dots, \mu \in \mathcal{M}(\Omega)$. We say that μ_n converges *weakly*[2] if $I_{\mu_n}(f) \to I_\mu(f)$ for all $f \in C_b(\Omega)$, and we write

$$\operatorname*{w-lim}_{n\to\infty} \mu_n = \mu.$$

We also note the following straightforward result that sheds light on the relationship between parts (a) and (b) in Theorem 5.1.8. It follows directly from the Stone-Weierstraß Theorem (Theorem E.1.4) and the triangle inequality.

[2] Weak convergence is sometimes referred to as narrow convergence; see [14, Section 8.1].

Proposition 5.1.11. *Let Ω be locally compact and $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}(\Omega)$ with* $\sup_{n\in\mathbb{N}} \|\mu_n\| < \infty$. *Then*

$$I_{\mu_n}(f) \to I_\mu(f) \text{ for all } f \in C_0(\Omega) \tag{5.1.1}$$

if and only if

$$I_{\mu_n}(f) \to I_\mu(f) \text{ for all } f \in C_c(\Omega).$$

Given that one can find a variety of definitions for vague convergence in the extant literature, some remarks on our definition are in order.

Remark 5.1.12. (a) Our definition of vague convergence is the most common one found in the literature; see e.g. Berg et al [9, Chapter 2], Dieudonné and Macdonald [27, Section XIII.4], Kallenberg [45, Chapter 5] or Klenke [48, Section 13.2].

(b) In a setting where Ω is locally compact and motivated by Theorem 5.1.8, vague convergence is defined for test functions in $C_0(\Omega)$ (rather than in $C_c(\Omega)$) by Folland [33, Section 7.3]. However, in light of Proposition 5.1.11, this stronger definition coincides with our definition if the sequence of measures is uniformly bounded.

(c) When Ω is a Polish space (i.e., complete and separable), the vague topology on $\mathcal{M}^+(\Omega)$ (which characterises vague convergence) has alternatively been defined to be generated by the family of mappings $\mu \mapsto I_\mu(f)$ where f are nonnegative continuous functions with metric bounded support. This is the approach taken by Kallenberg [44, Section 4.1] and Daley and Vere-Jones [24, Section A2.6]. Basrak and Planinić [8] show that this definition coincides with our definition using the theory of boundedness due to Hue [43]. Moreover, [8] show explicitly that these vague topologies make $\mathcal{M}^+(\Omega)$ a Polish space in its own right. In particular, this latter fact convinces us that our definition is the most natural one.

(d) There is often some confusion about how weak and vague convergence of measures (as defined in Definition 5.1.10) relate to the topological notions of weak and weak* convergence. We give a complete comparison in Appendix E.1, which further bolsters our definition of vague convergence.

5.1.3 Relationship between vague and weak convergence

We first revisit the direct relationship between weak and vague convergence for signed measures. As a warm-up, we recall that vague convergence allows for a loss of mass in the limit, while weak convergence does not.

Example 5.1.13. Let μ be the zero measure and $\{\mu_n\} \subset \mathcal{M}(\mathbb{R})$ be such that $\mu_n := \delta_n - \delta_{-n}$, where for $x \in \mathbb{R}$, δ_x denotes the Dirac measure at x. Then v-$\lim_{n\to\infty}\mu_n = \mu$ since for any $f \in C_c(\mathbb{R})$,

$$\lim_{n\to\infty} I_{\mu_n}(f) = \lim_{n\to\infty} (f(n) - f(-n)) = 0 = I_\mu(f).$$

Moreover, it holds that $\lim_{n\to\infty} \mu_n(\mathbb{R}) = \mu(\mathbb{R})$, i.e. the *signed mass* is preserved.

Now take $f \in C_b(\mathbb{R})$ such that

$$f(x) = \begin{cases} x & \text{for } x \in (-1,1), \\ \operatorname{sign}(x) & \text{otherwise,} \end{cases}$$

Thus, we do not have w-$\lim_{n\to\infty} = \mu$ since

$$2 = \lim_{n\to\infty} I_{\mu_n}(f) \neq \lim_{n\to\infty} I_\mu(f) = 0.$$

Intuitively, what goes wrong in Example 5.1.13 is that mass is "sent to infinity". The precise condition that avoids this is *tightness*.

Definition 5.1.14. A sequence $\{\mu_n\} \subset \mathcal{M}(\Omega)$ is called *tight* if for any $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset \Omega$ such that

$$\sup_{n\in\mathbb{N}} |\mu_n| (K_\varepsilon^c) \leq \varepsilon. \tag{5.1.2}$$

Remark 5.1.15. Since each $\mu \in \mathcal{M}(\Omega)$ is tight by definition, we can replace (5.1.2) by

$$\limsup_{n\to\infty} |\mu_n| (K_\varepsilon^c) \leq \varepsilon. \tag{5.1.3}$$

Tightness is exactly the condition that lifts vague to weak convergence for positive measures. This remains true for signed measures. The proof of the next result follows from Prokhorov's theorem for signed measures.[3]

Theorem 5.1.16 (Prokhorov's Theorem)**.** *Let Ω be a metrisable space and $\mathbf{M} \subset \mathcal{M}(\Omega)$ nonempty.*

(a) *If $\mathbf{M}$ is uniformly bounded and tight, then $\mathbf{M}$ is weakly relatively sequentially compact.*

(b) *If the space Ω is Polish and $\mathbf{M}$ is weakly relatively sequentially compact, then $\mathbf{M}$ is uniformly bounded and tight.*

Proof. (a) Take any $\{\mu_n\} \subset \mathbf{M}$. Since $\mathbf{M}$ is a uniformly bounded and tight sequence, both $\{\mu_n^+\}$ and $\{\mu_n^-\}$ are uniformly bounded and tight. By [48, Theorem 13.29], it follows that there exists a subsequence $\{n_k\}$ such that $\text{w-lim}_{k\to\infty} \mu_{n_k}^+ = \nu$, for some positive measure $\nu \in \mathcal{M}(\Omega)$. Similarly, there exists a subsequence $\{n_{k_l}\} \subset \{n_k\}$ such that $\text{w-lim}_{l\to\infty} \mu_{n_{k_l}}^- = \eta$, for some positive measure $\mathcal{M}(\Omega)$. Thus it follows that $\text{w-lim}_{l\to\infty} \mu_{n_{k_l}} = (\nu - \eta) \in \mathcal{M}(\Omega)$.

(b) See [14, Theorem 8.6.2]. □

Proposition 5.1.17. *Let $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}(\Omega)$.*

(a) *If $\text{v-lim}_{n\to\infty} \mu_n = \mu$ and $\{\mu_n\}$ is tight, then $\text{w-}\lim_{n\to\infty} \mu_n = \mu$.*

(b) *If $\text{w-}\lim_{n\to\infty} \mu_n = \mu$, then $\text{v-lim}_{n\to\infty} \mu_n = \mu$. If in addition Ω is Polish (i.e., complete and separable), then $\{\mu_n\}$ is tight.*

Proof. (a) The statement follows from Theorem 5.1.16(a) as soon as we show that $\{\mu_n\}$ is uniformly bounded. Since we can isometrically embed $\mathcal{M}(\Omega)$ into the norm dual of $C_b(\Omega)$ via the mapping I_μ, the Banach-Steinhaus theorem lets us assert that $\{\mu_n\}$ is bounded if and only if

$$\sup_{n\in\mathbb{N}} |I_{\mu_n}(f)| < \infty \text{ for every } f \in C_b(\Omega). \tag{5.1.4}$$

To show (5.1.4) we fix an $f \in C_b(\Omega)$ and take any $\varepsilon > 0$. Let $K_\varepsilon \subset \Omega$ be compact set such that (5.1.3) holds and find any $g \in C_c(\Omega)$ such that $g \equiv f$ on K. Then, by hypothesis and (5.1.3)

$$\limsup_{n\to\infty} |I_{\mu_n}(f)| \leq \limsup_{n\to\infty} |I_{\mu_n}(f-g)| + |I_\mu(g)|$$

[3]A direct proof of Proposition 5.1.17(a) follows also from a generalisation of [45, Lemma 5.20].

$$
\begin{aligned}
&\leq \|f-g\|_\infty \sup_{n\in\mathbb{N}} |\mu_n|\,(K_\varepsilon^c) + |I_\mu(g)| \\
&\leq \|f-g\|_\infty\, \varepsilon + |I_\mu(g)| < \infty.
\end{aligned}
$$

The claim (5.1.4) is now clear.
(b) This is a direct consequence of Theorem 5.1.16(b). □

If Ω is locally compact, the heuristic that vague convergence ignores mass "being sent to infinity" leads us to note that vague convergence in $\mathcal{M}(\Omega)$ (without loss of signed mass) can be viewed as weak convergence in $\mathcal{M}(\Omega_\infty)$, where Ω_∞ denotes the one-point compactification of Ω; see Definition E.1.2. To this end, note that a measure $\mu \in \mathcal{M}(\Omega)$ can be canonically extended to a measure $\mu^\infty \in \mathcal{M}(\Omega_\infty)$ by setting $\mu^\infty(A) := \mu(A)$ for $A \in \mathscr{B}(\Omega)$ and $|\mu^\infty|\,(\{\infty\}) := 0$. We then have the following result, which follows directly from Proposition 5.1.11 and Theorem E.1.3.

Proposition 5.1.18. *Let Ω be locally compact and $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}(\Omega)$ with $\sup_{n\in\mathbb{N}} \|\mu_n\| < \infty$. Denote by μ_n^∞ and μ^∞ the canonical extension of μ_n and μ, respectively. Then $\operatorname{v-lim}_{n\to\infty} \mu_n = \mu$ and $\mu_n(\Omega) \to \mu(\Omega)$ if and only if $\operatorname{w-lim}_{n\to\infty} \mu_n^\infty = \mu^\infty$.*

Remark 5.1.19. For *signed* measures, weak convergence in $\mathcal{M}(\Omega_\infty)$ is strictly weaker than weak convergence in $\mathcal{M}(\Omega)$. Indeed, Example 5.1.13 gives an example of $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}(\Omega)$ with $\sup_{n\in\mathbb{N}} \|\mu_n\| < \infty$ such that $\text{v-lim}_{n\to\infty} \mu_n = \mu$ and $\mu_n(\Omega) \to \mu(\Omega)$ (and hence $\operatorname{w-lim}_{n\to\infty} \mu_n^\infty = \mu^\infty$), but $\text{w-lim}_{n\to\infty} \mu_n \neq \mu$.

We next investigate under which conditions vague convergence implies the convergence of the positive and negative parts in the Hahn–Jordan decomposition. In order to do so, we need a Portmanteau theorem and related results.

Theorem 5.1.20 (Vague Portmanteau Theorem for positive measures)**.** *Let Ω be a locally compact metrisable space and $\{\mu_n\} \cup \{\mu\} \in \mathcal{M}^+(\Omega)$. Then the following are equivalent:*

(a) $\text{v-lim}_{n\to\infty} \mu_n = \mu$.

(b) *For any compact set $K \subset \Omega$,*

$$\limsup_{n \to \infty} \mu_n(K) \leq \mu(K)$$

and for any open set $\Theta \subset \Omega$,

$$\liminf_{n \to \infty} \mu_n(\Theta) \geq \mu(\Theta).$$

(c) *For any set $A \subset \Omega$ such that $A \subset K$ for some compact set K and $\mu(\partial A) = 0$,*

$$\lim_{n \to \infty} \mu_n(A) = \mu(A).$$

Proof. See [64, Theorem 21.15] □

One part of the direction "(a) ⇒ (b)" in the vague Portmanteau theorem extends to signed measures. This result is often attributed to Varadarajan in [73], although the cited paper only contains a proof for positive measures. A proof for the general case will need a consequence of Urysohn's lemma.

Lemma 5.1.21. *Let Ω be a normal space and suppose that $\Theta \subset \Omega$ is open and $\mu \in \mathcal{M}(\Omega)$. If $|\mu|(\Theta) > \varepsilon$, then there exists $f \in C(\Omega)$ such that $|f| \leq 1, f \equiv 0$ on Θ^c and*

$$\int_\Theta f \, \mathrm{d}\mu > \varepsilon. \tag{5.1.5}$$

Proof. Let $P \cup N$ be the Hahn-Jordan decomposition of Ω with respect to μ, and define

$$V^+ := P \cap \Theta, \quad V^- := N \cap \Theta$$

Moreover, define

$$\delta := \frac{1}{3}(|\mu|(\Theta) - \varepsilon) > 0$$

Since μ is tight, we choose compact sets $K^\pm \subset V^\pm$such that $|\mu| \left(V^\pm \backslash K^\pm\right) < \frac{\delta}{2}$. Furthermore, noting that Ω is normal, we choose disjoint open neighbourhoods $\Theta^+$and $\Theta^-$of $K^+$and K^-, respectively, which are contained in Θ. Using Lemma E.1.1, we thus find $f^+, f^- \in C(\Omega)$ such that $|f^\pm| \leq 1$ and

$$f^\pm(x) = \begin{cases} 1 & \text{for } x \in K^\pm, \\ 0 & \text{on } x \in \left(\Theta^\pm\right)^c. \end{cases}$$

We now define $f \in C(\Omega)$ such that $f := f^+ - f^-$. By construction, we have

$$\operatorname{supp}\left(f^+\right) \cap \operatorname{supp}\left(f^-\right) \subset \Theta^+ \cap \Theta^- = \emptyset,$$

whence $|f| \leq 1$. Furthermore,

$$\begin{aligned}
\int_\Theta f \, \mathrm{d}\mu &= \int_{K^+} f \, \mathrm{d}\mu + \int_{K^-} f \, \mathrm{d}\mu + \int_{\Theta \backslash (K^+ \cup K^-)} f \, \mathrm{d}\mu \\
&= \int_{K^+} f^+ \mathrm{d}\mu - \int_{K^-} f^- \mathrm{d}\mu + \int_{\Theta \backslash (K^+ \cup K^-)} f \, \mathrm{d}\mu \\
&\geq \mu^+\left(K^+\right) + \mu^-\left(K^-\right) - 2|\mu|\left(\Theta \backslash \left(K^+ \cup K^-\right)\right) \\
&= |\mu|\left(K^+\right) + |\mu|\left(K^-\right) - 2\left(|\mu|(\Theta) - |\mu|\left(K^+ \cup K^-\right)\right) \\
&= 3\left(|\mu|\left(K^+\right) + |\mu|\left(K^-\right)\right) - 2|\mu|(\Theta) \\
&> 3\left(|\mu|\left(V^+\right) + |\mu|\left(V^-\right)\right) - 2|\mu|(\Theta) - 3\delta \\
&= \varepsilon.
\end{aligned}$$

$\square$

Corollary 5.1.22. *Let Ω be a locally compact normal space, $\Theta \subset \Omega$ be open, and $\mu \in \mathcal{M}(\Omega)$. Then for any $\varepsilon > 0$, there exists $f \in C_c(\Omega)$ such that $|f| \leq 1, f \equiv 0$ on Θ^c and*

$$\int_\Theta f \, \mathrm{d}\mu > |\mu|(\Theta) - \varepsilon$$

Proof. We will only consider the case where $|\mu|(\Theta) > \varepsilon$. Since μ is tight there exists compact set $K_\varepsilon \subset \Theta$ such that $|\mu|\left(\Theta \backslash K_\varepsilon\right) \leq \frac{1}{2}\varepsilon$. Moreover, since Ω is a locally compact Hausdorff space, by [1, Corollary 2.69] there exists an open set E with compact closure $\bar{E}$ such that

$$K_\varepsilon \subset E \subset \bar{E} \subset \Theta$$

Noting that $|\mu|(\Theta \backslash E) \leq |\mu|\left(\Theta \backslash K_\varepsilon\right)$, it follows that

$$|\mu|(E) - \frac{1}{2}\varepsilon \geq |\mu|(\Theta) - \varepsilon > 0$$

Thus, by Lemma 5.1.21, there exists $f \in C(\Omega)$ and $|f| \leq 1, f \equiv 0$ on E^c such

that
$$\int_\Theta f \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu > |\mu|(E) - \frac{1}{2}\varepsilon \geq |\mu|(\Theta) - \varepsilon$$
Since $\overline{\mathrm{supp}(f)} \subset \bar{E}$ it follows that $f \in C_c(\Omega)$. $\square$

We can now state a version of "(a) $\Rightarrow$ (b)" in Theorem 5.1.20 for signed measures.

Theorem 5.1.23. *Let Ω be a locally compact normal Hausdorff space. Let $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}(\Omega)$ and assume that* $\text{v-lim}_{n\to\infty} \mu_n = \mu$. *Then for any open set $\Theta \subset \Omega$,*
$$|\mu|(\Theta) \leq \liminf_{n\to\infty} |\mu_n|(\Theta). \tag{5.1.6}$$
In particular, $\|\mu\| \leq \liminf_{n\to\infty} \|\mu_n\|$.

Proof. Let $\Theta \subset \Omega$ be open and $\varepsilon > 0$. Since μ is tight and Ω is normal and locally compact, Corollary 5.1.22 tells us that there exists $f \in C_c(\Omega)$ such that $|f| \leq 1$, $\mathrm{supp}(f) \subset \Theta$ and
$$\int f \, \mathrm{d}\mu \geq |\mu|(\Theta) - \varepsilon.$$
Then by vague convergence of $\{\mu_n\}$,
$$|\mu|(\Theta) - \varepsilon \leq \int f \, \mathrm{d}\mu = \lim_{n\to\infty} \int f \, \mathrm{d}\mu_n \leq \liminf_{n\to\infty} \int |f| \, \mathrm{d}|\mu_n| \leq \liminf_{n\to\infty} |\mu_n|(\Theta)$$
Now the result follows by letting $\varepsilon \downarrow 0$. $\square$

We are now in a position to show that the necessary and sufficient extra condition for the convergence of the separate parts of the Hahn-Jordan decomposition is that no mass is lost on compact sets.

Proposition 5.1.24. *Let Ω be locally compact and $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}(\Omega)$. Then* $\text{v-lim}_{n\to\infty} \mu_n^\pm = \mu^\pm$ *if and only if* $\text{v-lim}_{n\to\infty} \mu_n = \mu$ *and*
$$\limsup_{n\to\infty} |\mu_n|(K) \leq |\mu|(K). \tag{5.1.7}$$
for every compact set $K \subset \Omega$.

Proof. First, suppose that $\text{v-lim}_{n\to\infty}\mu_n^\pm = \mu^\pm$. Then clearly $\text{v-lim}_{n\to\infty}\mu_n = \mu$, and (5.1.7) is satisfied due to the Portmanteau Theorem in the form of Theorem 5.1.20(b).

Conversely, suppose that $\text{v-lim}_{n\to\infty}\mu_n = \mu$ and (5.1.7) is satisfied. By Theorem 5.1.23, for every open set $\Theta \subset \Omega$,

$$\liminf_{n\to\infty} |\mu_n|(\Theta) \geq |\mu|(\Theta).$$

Thus, Theorem 5.1.20(b) gives $\text{v-lim}_{n\to\infty}|\mu_n| = |\mu|$. Now $\text{v-lim}_{n\to\infty}\mu_n{}^\pm = \mu^\pm$ follows by noting that

$$\mu_n^+ = \frac{1}{2}(|\mu_n| + \mu_n) \quad \text{and} \quad \mu_n^- = \frac{1}{2}(|\mu_n| - \mu_n). \qquad \square$$

Note that Condition (5.1.7) does not restrict "total mass being lost at infinity". By imposing an additional restriction to mitigate this possibility, we can strengthen Proposition 5.1.24 to deduce that $\text{w-lim}_{n\to\infty}\mu_n{}^\pm = \mu_n{}^\pm$.

Proposition 5.1.25. *Let Ω be locally compact and $\{\mu_n\}\cup\{\mu\} \subset \mathcal{M}(\Omega)$. Then* $\text{w-lim}_{n\to\infty}\mu_n^\pm = \mu^\pm$ *if and only if* $\text{v-lim}_{n\to\infty}\mu_n = \mu$ *and*

$$\limsup_{n\to\infty} \|\mu_n\| \leq \|\mu\|. \tag{5.1.8}$$

Proof. First, suppose that $\text{w-lim}_{n\to\infty}\mu_n^\pm = \mu^\pm$. Then $\text{w-lim}_{n\to\infty}\mu_n = \mu$ and $\text{w-lim}_{n\to\infty}|\mu_n| = |\mu|$. This implies in particular that $\text{v-lim}_{n\to\infty}\mu_n = \mu$ and

$$\lim_{n\to\infty}\|\mu_n\| = \lim_{n\to\infty}\int_\Omega \mathrm{d}|\mu_n| = \int_\Omega \mathrm{d}|\mu| = \|\mu\|. \tag{5.1.9}$$

Conversely, suppose that $\text{v-lim}_{n\to\infty}\mu_n = \mu$ and (5.1.8) is satisfied. By Propositions 5.1.24 and 5.1.17, it suffices to show that (5.1.7) is satisfied and the sequence $\{\mu_n\}$ is tight.

First, we establish (5.1.7). Seeking a contradiction, suppose there exists a compact set $K \subset \Omega$ such that

$$\limsup_{n\to\infty} |\mu_n|(K) > |\mu|(K). \tag{5.1.10}$$

Since K^c is open, it follows from Theorem 5.1.23 that

$$\liminf_{n\to\infty} |\mu_n|(K^c) \geq |\mu|(K^c). \tag{5.1.11}$$

Adding (5.1.10) and (5.1.11), it follows that

$$\limsup_{n\to\infty} \|\mu_n\| = \limsup_{n\to\infty} |\mu_n|(\Omega) \geq \limsup_{n\to\infty} |\mu_n|(K) + \liminf_{n\to\infty} |\mu_n|(K^c) > |\mu|(\Omega) = \|\mu\|,$$

and we arrive at a contradiction to (5.1.8).

Next, we show that the sequence $\{\mu_n\}$ is tight. Let $\varepsilon > 0$. By tightness of μ, there exists a compact set $K \subset \Omega$ such that $|\mu|(K^c) \leq \varepsilon$. By local compactness of Ω, there exists an open set $L \supset K$ such that its closure $\overline{L} =: K_\varepsilon$ is compact. Using (5.1.8) and Theorem 5.1.23, we obtain

$$\begin{aligned}
\limsup_{n\to\infty} |\mu_n|(K_\varepsilon^c) &= \limsup_{n\to\infty} (\|\mu_n\| - |\mu_n|(K_\varepsilon)) \\
&\leq \limsup_{n\to\infty} (\|\mu_n\| - |\mu_n|(L)) \\
&\leq \|\mu\| - \liminf_{n\to\infty} |\mu_n|(L) \leq \|\mu\| - |\mu|(L) \\
&\leq \|\mu\| - |\mu|(K) = \mu(K^c) \leq \varepsilon.
\end{aligned}$$

□

Table 5.1: Ω is a (Polish*, locally compact**) metrisable space and $\{\mu_n\}\cup\{\mu\} \subset \mathcal{M}(\Omega)$.

Condition(s) A		**Condition(s) B**
$\text{v-lim}_{n\to\infty} \mu_n = \mu$, and $\forall \varepsilon > 0$, $\exists$ compact set K_ε such that $\limsup_{n\to\infty} |\mu_n|(K_\varepsilon^c) \leq \varepsilon$	$\Rightarrow$ $\overset{\star}{\Leftarrow}$	$\text{w-lim}_{n\to\infty} \mu_n = \mu$
$\text{v-lim}_{n\to\infty} \mu_n = \mu$, and $\forall$ compact $K \subset \Omega$ $\limsup_{n\to\infty} |\mu_n|(K) \leq |\mu|(K)$	$\overset{\star\star}{\Leftrightarrow}$	$\text{v-lim}_{n\to\infty} {\mu_n}^\pm = \mu^\pm$
$\text{v-lim}_{n\to\infty} \mu_n = \mu$, and $\limsup_{n\to\infty} \|\mu_n\| \leq \|\mu\|$	$\overset{\star\star}{\Leftrightarrow}$	$\text{w-lim}_{n\to\infty} {\mu_n}^\pm = \mu^\pm$

To summarise, starting from vague convergence $\text{v-lim}_{n\to\infty} \mu_n = \mu$, Proposition 5.1.17 tells us that we get $\text{w-lim}_{n\to\infty} \mu_n = \mu$ if mass is not "lost at infinity". Proposition 5.1.24 asserts that if mass is not "lost on compact sets", then we get $\text{v-lim}_{n\to\infty} \mu_n^\pm = \mu^\pm$. Finally, Proposition 5.1.25 tells us that if mass is not "lost globally", then we even get $\text{w-lim}_{n\to\infty} \mu_n^{\pm} = \mu^{\pm}$. These results are summarised in Table 5.1.

5.2 Distribution functions of real-valued measures

In this section, we study the particular case that $\Omega = \mathbb{R}$ (with the usual order topology) and link vague convergence on $\mathbb{R}$ to the pointwise convergence of their distribution functions (at continuity points of the limiting measure). To this end, we first need to introduce other pieces of notation.

5.2.1 Functions of bounded variation

First we recall the definition of a function of bounded variation.

Definition 5.2.1. Let $I \subset \mathbb{R}$ be an interval and let Π_I be the set of all partitions

$$\pi = \{x_0^\pi < x_1^\pi < \cdots < x_n^\pi\} \subset I.$$

The *total variation* of a function $F : I \to \mathbb{R}$ is defined by

$$\mathrm{Var}_F := \sup_{\pi\in\Pi_I} \left\{ \sum_{i=1}^{n} \left| u(x_i^\pi) - u(x_{i-1}^\pi) \right| \right\}$$

The function F is said to have finite variation if $\mathrm{Var}_F < \infty$ and the space of all functions of bounded variation on I is denoted by $\mathrm{BV}(I)$. We denote by $\mathrm{BV}_{\mathrm{loc}}(I)$ the space of all functions $F : I \to \mathbb{R}$ such that $F|_{[a,b]} \in \mathrm{BV}([a,b])$ for any $[a,b] \subset I$.

For $F \in \mathrm{BV}(I)$ and $x \in I$, denote the total variation of F on $(-\infty, x] \cap I$ by $\mathbf{V}_F(x)$ and set $F^\uparrow(x) := \frac{1}{2}(\mathbf{V}_F(x) + F(x))$ and $F^\downarrow(x) := \frac{1}{2}(\mathbf{V}_F(x) - F(x))$. Note that $\mathbf{V}_F, F^\uparrow, F^\downarrow : \mathbb{R} \to [0,\infty]$ are nondecreasing functions.

Definition 5.2.2. For any $\alpha \in \mathbb{R}$ and $\mu \in \mathcal{M}^*(\mathbb{R})$, the *distribution function* *of* μ*, centred at* α*,* is the function $F_\mu^{(\alpha)} \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$ defined by

$$F_\mu^{(\alpha)}(x) := \begin{cases} \mu((\alpha, x]) & \text{if } x > \alpha, \\ 0 & \text{if } x = \alpha, \\ -\mu((x, \alpha]) & \text{if } x < \alpha. \end{cases}$$

Note that $F_\mu^{(\alpha)}$ is right-continuous, and for any $a \le b$ with $a, b \in \mathbb{R}$,

$$F_\mu^{(\alpha)}(b) - F_\mu^{(\alpha)}(a) = \mu((a, b]). \tag{5.2.1}$$

The relationship (5.2.1) between distribution functions and real-valued measures is bijective, which follows from the following result; for a proof see [52, Theorem 5.13].

Theorem 5.2.3. *Let* $F \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R})$ *be right-continuous. Then there exists a unique* $\mu_F \in \mathcal{M}^*(\mathbb{R})$ *such that*

$$\mu_F((a, b]) = F(b) - F(a)$$

for all $a \le b$ *with* $a, b \in \mathbb{R}$*. Moreover,* $|\mu_F| = \mu_{\boldsymbol{V}_F}$*.*

Until the next chapter, we will only focus on functions of strictly finite variation. In which case, let $[-\infty, \infty]$ be the (affine) extended real line (with the order topology). Any $\mu \in \mathcal{M}(\mathbb{R})$ can canonically be extended to $\mathcal{M}([-\infty, \infty])$ by setting $|\mu|(\{\pm\infty\}) := 0$. Similarly, for $\alpha \in \mathbb{R}$, $F_\mu^{(\alpha)}$ can canonically be extended to $[-\infty, \infty]$ by setting $F_\mu^{(\alpha)}(\pm\infty) := \lim_{x \to \pm\infty} F_\mu^{(\alpha)}(x)$. Finally, we can define $F^{(-\infty)}, F^{(+\infty)} \in \mathrm{BV}(\mathbb{R})$ by

$$F_\mu^{(-\infty)}(x) := \mu((-\infty, x]) \quad \text{and} \quad F_\mu^{(+\infty)}(x) := -\mu((x, \infty)), \quad x \in \mathbb{R},$$

respectively, which again can canonically be extended to $[-\infty, \infty]$. Note that $F_\mu^{(-\infty)}$ is usually called the distribution function of μ and denoted by F_μ.

Last but not least, we say that $x \in \mathbb{R}$ is a *continuity point* of $\mu \in \mathcal{M}(\mathbb{R})$ if $\mu(\{x\}) = 0$.

5.2.2 Relationship between the convergence of distribution functions and vague convergence

We start our discussion on the relationship between the convergence of distribution functions and vague convergence by recalling the key result for *positive* measures. This type of result is essentially known – at least under stronger conditions, see e.g. [33, Proposition 7.19]. It will follow as a corollary of our main result, Theorem 5.2.11 below.

Theorem 5.2.4. *Let $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}^+(\mathbb{R})$ and $\alpha \in \mathbb{R}$ be a continuity point of μ. Then the following are equivalent:*

(a) $F_{\mu_n}^{(\alpha)} \to F_{\mu}^{(\alpha)}$ *at the continuity points of μ.*

(b) $\text{v-lim}_{n\to\infty} \mu_n = \mu$.

Moreover, if $\alpha = -\infty$ or $\alpha = +\infty$, the equivalence remains true if we require in addition that $\lim_{K\downarrow -\infty} [\limsup_{n\to\infty} \mu_n((-\infty, K])] = 0$ when $\alpha = -\infty$, or $\lim_{K\uparrow\infty} [\limsup_{n\to\infty} \mu_n((K, \infty))] = 0$ when $\alpha = +\infty$.

Remark 5.2.5. (a) The assumption that α is a continuity point of μ in Theorem 5.2.4 is necessary. Indeed, let $\mu_n := \delta_{1/n}$ and $\mu := \delta_0$. Then $\text{v-lim}_{n\to\infty} \mu_n = \mu$ but

$$F_{\mu_n}^{(0)}(x) = 0 \not\to -1 = F_{\mu}^{(0)}(x), \quad x < 0.$$

(b) As a sanity check, one notes that if $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}_1^+(\mathbb{R})$ are probability measures, whence $\limsup_{n\to\infty} \|\mu_n\| = \|\mu\| = 1$, then Theorem 5.2.4 together with Proposition 5.1.25 shows that $\text{w-lim}_{n\to\infty} \mu_n = \mu$ if and only if $F_{\mu_n}^{(-\infty)} \to F_{\mu}^{(-\infty)}$ at all continuity points of μ. This is often shown as a consequence of Portmanteau's theorem for weak convergence.

Both implications "(a) $\Rightarrow$ (b)" and "(b) $\Rightarrow$ (a)" in Theorem 5.2.4 are false for signed measures. The first counterexample shows that $F_{\mu_n}^{(\alpha)} \to F_{\mu}^{(\alpha)}$ at the continuity points of μ does *not* imply that $\text{v-lim}_{n\to\infty} \mu_n = \mu$. It relies on $\{F_{\mu_n}^{(\alpha)}\}$ being unbounded on a compact set.

Example 5.2.6. Let $F_n : \mathbb{R} \to \mathbb{R}$ be supported on $[0, 2/n]$ and linear between the points $\{0, 1/n, 2/n\}$ such that $F(0) := 0 =: F(2/n)$ and $F(1/n) := 2^n$;

see Figure 5.1 for a clear visualisation. For $n \in \mathbb{N}$, let $\mu_n := \mu_{F_n}$ according to Theorem 5.2.3 and denote by μ the zero measure. Then for any $x \in \mathbb{R}$, we have $F^{(0)}_{\mu_n}(x) = F^{(0)}_n(x) \to F^{(0)}_\mu(x)$.

Now take $f \in C_c(\mathbb{R})$ such that

$$f(x) := \begin{cases} (1+x) & \text{for } x \in [-1,0), \\ (1-x) & \text{for } x \in [0,1], \\ 0 & \text{for } x \in [-1,1]^c. \end{cases}$$

Then for $n \geq 2$.

$$I_{\mu_n}(f) = 2^n \left\{ \int_0^{1/n} (1-x) \, \mathrm{d}x - \int_{1/n}^{2/n} (1-x) \, \mathrm{d}x \right\} = \frac{2^{n+1}}{n^2}.$$

Thus, $I_{\mu_n}(f) \not\to I_\mu(f) = 0$.

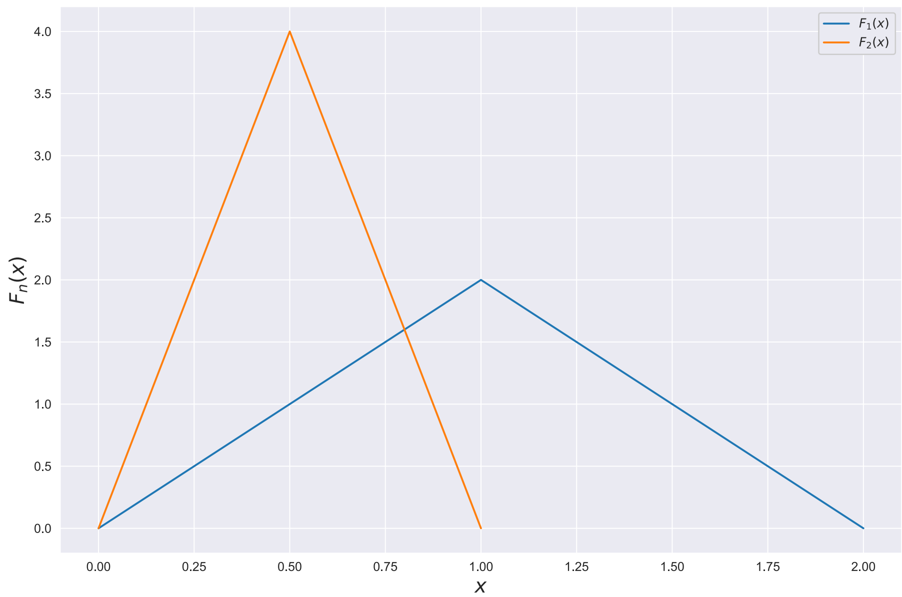


Figure 5.1: A visualisation of F_1 and F_2 defined in Example 5.2.6.

The next counterexample shows that $\text{v-lim}_{n\to\infty} \mu_n = \mu$ does not imply $F^{(\alpha)}_{\mu_n} \to F^{(\alpha)}_\mu$ at the continuity points of μ since mass can be lost locally. This happens when the positive and negative parts of the singular decompositions of $\{\mu_n\}$ cancel in the limit.

Example 5.2.7. Let $\mu_n := \delta_0 - \delta_{1/n}$, and let μ be the zero measure. Then it is straightforward to check that $\text{v-lim}_{n\to\infty} \mu_n = \mu$ (even $\text{w-lim}_{n\to\infty} \mu_n = \mu$). However, we do not have $F^{(0)}_{\mu_n} \to F^{(0)}_{\mu}$ at all continuity points of μ. Indeed, fix $x > 0$. Then for $n \geq \frac{1}{x}$,

$$F^{(0)}_{\mu_n}(x) = \delta_0((0,x]) - \delta_{1/n}((0,x]) = -1,$$

so

$$-1 = \lim_{n\to\infty} F^{(0)}_{\mu_n}(x) \neq F^{(0)}_{\mu}(x) = 0.$$

Thus, in order to ensure that the distribution functions converge at continuity points, one must ensure that mass is preserved locally. This motivates the following definition.

Definition 5.2.8. Let $\{\mu_n\} \subset \mathcal{M}(\mathbb{R})$.

1. We say that the sequence $\{\mu_n\}$ *has no mass at a point* $x \in \mathbb{R}$, if for any $\varepsilon > 0$, there exists an open neighbourhood $N_{x,\varepsilon}$ of x, such that

$$\limsup_{n\to\infty} |\mu_n| (N_{x,\varepsilon}) \leq \varepsilon.$$

   We say that the sequence $\{\mu_n\}$ has no mass at $+\infty$ (resp. $-\infty$), when the family of canonical extensions of $\{\mu_n\}$ has no mass at $+\infty$ (resp. $-\infty$).

2. We say that the sequence $\{\mu_n\}$ is *right equi-continuous at* $x \in \mathbb{R}$, if for any $\varepsilon > 0$, there exists a $h > 0$ such that for all $\delta < h$

$$\limsup_{n\to\infty} |\mu_n((x, x+\delta])| \leq \varepsilon.$$

Remark 5.2.9. Definition 5.2.8 implies that the family $\{\mu_n\} \subset \mathcal{M}(\mathbb{R})$ is tight if and only if it has no mass at $+\infty$ and $-\infty$.

If the family $\{\mu_n\} \subset \mathcal{M}(\mathbb{R})$ has no mass at a point $x \in \mathbb{R}$, then it is clearly right equi-continuous at x. The converse implication does not hold, as shown by the following example.

Example 5.2.10. For $n \in \mathbb{N}$, let $F_n : \mathbb{R} \to \mathbb{R}$ be supported on $[-2^{-n}, 2^{-n}]$ and linear between the points $\{k2^{-2n} : k \in \{-2^n, \dots, 2^n\}\}$ such that

$$F_n\left(k2^{-2n}\right) := (k \bmod(2))2^{-n}, \quad k \in \{-2^n, \dots, 2^n\};$$

see Figure 5.2 for a clear visualisation. Set $\mu_n := \mu_{F_n}$ and let μ be the zero measure. Then $\{\mu_n\} \subset \mathcal{M}(\mathbb{R})$ satisfies the properties:

(i) μ_n is supported on $[-2^{-n}, 2^{-n}]$,

(ii) $|\mu_n|([-2^{-n}, 2^{-n}]) = 2$,

(iii) $\left|F_{\mu_n}^{(0)}(x)\right| \leq 2^{-n}$ for all $x \in \mathbb{R}$.

In particular, property (ii) shows that $\{\mu_n\}$ does not satisfy the no-mass condition at $x = 0$, while property (iii) shows that $\{\mu_n\}$ is right equi-continuous everywhere.

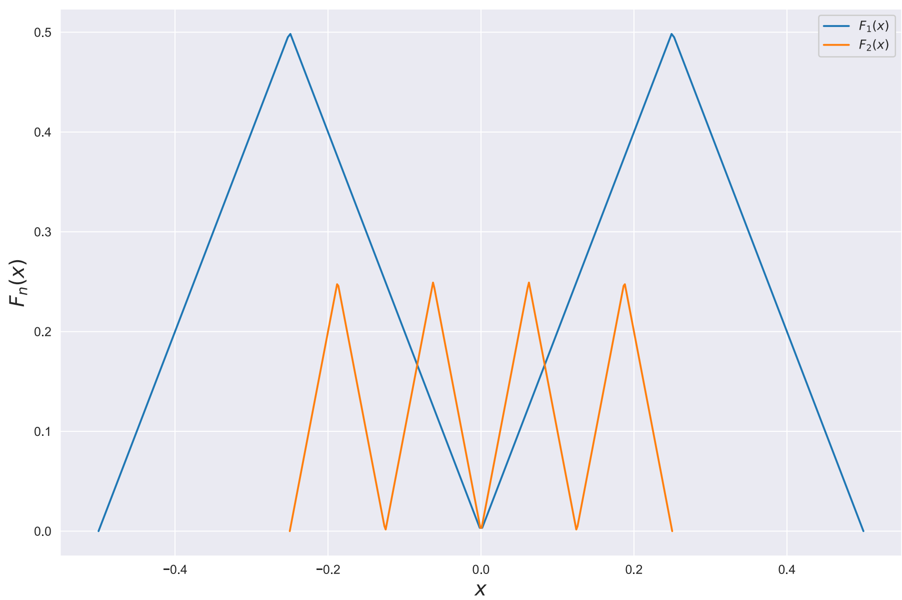


Figure 5.2: A visualisation of F_1 and F_2 defined in Example 5.2.10.

For $\{\mu_n\} \subset \mathcal{M}(\mathbb{R})$, the preceding discussion leads us to a clear characterisation of vague and weak convergence of $\{\mu_n\}$ from the convergence of F_{μ_n}, and vice versa.

Theorem 5.2.11. *Let $\alpha \in \mathbb{R}$ and $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}(\mathbb{R})$ be such that $\{\mu_n\}$ are bounded on compact sets.*

(a) *If $F^{(\alpha)}_{\mu_n}(x) \to F^{(\alpha)}_{\mu}(x)$ at all continuity points of μ, then $\text{v-lim}_{n\to\infty} \mu_n = \mu$.*

(b) *If $\text{v-lim}_{n\to\infty} \mu_n = \mu$ and the sequence $\{\mu_n\}$ is right equi-continuous at the continuity points of μ, then $F^{(\alpha)}_{\mu_n} \to F^{(\alpha)}_{\mu}$ at the continuity points of μ.*

Moreover, if $\alpha = -\infty$ or $\alpha = +\infty$, both parts remain true if we require in addition in (a) *that $\{\mu_n\}$ is bounded on compact neighbourhoods of α (in the extended order topology) and in* (b) *that $\{\mu_n\}$ has no mass at α (for the canonical extensions of $\{\mu_n\}$).*

Proof. We only establish the result for $\alpha \in \mathbb{R}$. The extension of the proof to $\alpha \in \{-\infty, \infty\}$ is straightforward.

(a) First, let $f \in \mathcal{C} := C^1(\mathbb{R}) \cap C_c(\mathbb{R})$. Then f is supported by a compact interval $K \subset \mathbb{R}$, and we may assume without loss of generality that $\alpha \in K$. Then $\{F^{(\alpha)}_{\mu_n}\}$ is bounded on K since

$$\left|F^{(\alpha)}_{\mu_n}(x)\right| \leq \sup_{n\in\mathbb{N}} |\mu_n| (K) < \infty, \quad x \in K.$$

Moreover, $F^{(\alpha)}_{\mu_n} \to F^{(\alpha)}_{\mu}$ a.e. by the fact that μ has only countably many atoms. Therefore, an integration by parts and the dominated convergence theorem gives

$$I_{\mu_n}(f) = -\int_K f'(x) F^{(\alpha)}_{\mu_n}(x)\, \mathrm{d}x \to -\int_K f'(x) F^{(\alpha)}_{\mu}(x)\, \mathrm{d}x = I_\mu(f). \tag{5.2.2}$$

Next, let $f \in C_c(\mathbb{R}) \subset C_0(\mathbb{R})$ and $\varepsilon > 0$. Since $\mathcal{C}$ is a subalgebra of $C_0(\mathbb{R})$ that separates points and vanishes nowhere, it is dense in $C_0(\mathbb{R})$ by Theorem E.1.4. Thus, there exists $g \in \mathcal{C}$ such that $\|f - g\|_\infty < \varepsilon$. Then f and g are both supported by some compact interval L. Hence, the triangle inequality and (5.2.2) give

$$\begin{aligned}\limsup_{n\to\infty} |I_{\mu_n}(f) - I_\mu(f)| &\leq \limsup_{n\to\infty} \left(|I_{\mu_n}(f-g)| + |I_{\mu_n}(f) - I_\mu(f)| + |I_\mu(f-g)|\right) \\ &\leq \left(\sup_{n\in\mathbb{N}} |\mu_n| (L) + \|\mu\|\right) \varepsilon.\end{aligned}$$

Using that $\{\mu_n\}$ is bounded on compact sets and taking $\varepsilon \downarrow 0$ establishes the claim.

(b) Fix $T > 0$ and define $\mu_n^{(T)} := \mu_{n|_{[-T,T]}}$. By hypothesis, the family $\{\mu_n^{(T)}\}$ is uniformly bounded. Hence, by Helly's selection theorem [52, Theorem 2.35], it holds that along a subsequence $\{n_k\}$,

$$F_{\mu_{n_k}}^{(\alpha)} \to F \in \mathrm{BV}([-T,T]).$$

According to Theorem 5.2.3 there exists a measure $\mu^{(T)} \in \mathcal{M}(\mathbb{R})$, supported on $[-T,T]$, such that $\lim_{y\downarrow\downarrow x} F(y) = F_{\mu^{(T)}}^{(\alpha)}(x)$. Since the sequence $\{\mu_n\}$ is right equi-continuous at all continuity points of μ, $F = F_{\mu^{(T)}}^{(\alpha)}$ at all such points.

Taking any $f \in C_c([-T,T])$, it follows from part (a)

$$I_\mu(f) = \lim_{k\to\infty} I_{\mu_{n_k}}(f) = \lim_{k\to\infty} I_{\mu_{n_k}^{(T)}}(f) = I_{\mu^{(T)}}(f).$$

Thus, $\mu_{|_{[-T,T]}} = \mu^{(T)}$. By the subsequence criterion, it follows that $F_{\mu_n}^{(\alpha)}(x) \to F_\mu^{(\alpha)}(x)$ for all $x \in [-T,T]$ that are continuity points of μ. Since T was arbitrary, the proof is complete. $\square$

In part (b) of the previous theorem, one needs that the family $\{\mu_n\} \subset \mathcal{M}(\mathbb{R})$ is bounded on compact sets in order to utilise Helly's selection theorem. Assuming that the family $\{\mu_n\}$ has no mass at a continuity point of μ, one can relax this assumption by using a different approach in the proof.

Theorem 5.2.12. *Let $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}(\mathbb{R})$ and α a continuity point of μ. If* $\text{v-}\lim_{n\to\infty} \mu_n = \mu$ *and $\{\mu_n\}$ has no mass at a continuity points x of μ, then $F_{\mu_n}^{(\alpha)}(x) \to F_\mu^{(\alpha)}(x)$.*

Proof. Let $t \in \mathbb{R}$ be a continuity point of μ. The case when $t = \alpha$ is trivial, so we may assume without loss of generality that $t > \alpha$, since $F_\mu^{(\alpha)}(t) = -F_\mu^{(t)}(\alpha)$.

For $\delta > 0$, define the cut-off function $\rho_\delta \in C_c(\mathbb{R})$ by

$$\rho_\delta(x) = \begin{cases} 0 & \text{if } x \notin (\alpha-\delta, t+\delta) \\ \frac{1}{\delta}(x+\delta-\alpha), & \text{if } x \in (\alpha-\delta, \alpha), \\ 1 & \text{if } x \in [\alpha, t], \\ \frac{1}{\delta}(t+\delta-x), & \text{if } x \in (t, t+\delta), \end{cases}$$

and for $x \in \mathbb{R}$, the open ball around x of radius δ by $B_\delta(x)$. Then

$$\begin{aligned}
&\limsup_{n\to\infty} \left(\left|F_{\mu_n}^{(\alpha)}(t) - F_\mu^{(\alpha)}(t)\right|\right) \\
\leq\ & \limsup_{n\to\infty} \Bigg(\left|\int (\mathbb{1}_{(\alpha,t]} - \rho)(x)\mu_n(\mathrm{d}x)\right| + \left|\int \rho(x)\mu_n(\mathrm{d}x) - \int \rho(x)\mu(\mathrm{d}x)\right| \\
&\quad + \left|\int (\mathbb{1}_{(\alpha,t]} - \rho)(x)\mu(\mathrm{d}x)\right| \Bigg) \\
\leq\ & \limsup_{n\to\infty} \Big(|\mu_n|\left((\alpha-\delta,\alpha]\right) + |\mu_n|\left((t,t+\delta)\right) + |\mu|\left((\alpha-\delta,\alpha]\right) + |\mu|\left((t,t+\delta)\right)\Big) \\
\leq\ & \limsup_{n\to\infty} |\mu_n|\left(B_\delta(\alpha)\right) + \limsup_{n\to\infty} |\mu_n|\left(B_\delta(t)\right) + |\mu|\left((\alpha-\delta,\alpha]\right) + |\mu|\left((t,t+\delta)\right).
\end{aligned} \tag{5.2.3}$$

Now the result follows by taking $\delta \to 0$, noting that the first two terms on the right had side of (5.2.3) vanish by the fact that $\{\mu_n\}$ has no mass at t and α, whereas the last two terms on the right had side of (5.2.3) vanish by σ-continuity of μ and the fact that α is a continuity point of μ. □

We proceed to prove Theorem 5.2.4, which is in fact a corollary to Theorem 5.2.11.

Proof of Theorem 5.2.4. We only establish the result for $\alpha \in \mathbb{R}$. The extension of the proof to $\alpha \in \{-\infty, \infty\}$ is straightforward.

"(a) $\Rightarrow$ (b)". By Theorem 5.2.11(a), it suffices to show that $\{\mu_n\}$ are bounded on compact sets. So let $K \subset \mathbb{R}$ be a compact set. Then there exists continuity points $b_1, b_2 \in \mathbb{R}$ of μ such that $K \subset (b_1, b_2]$. By hypothesis, $\lim_{n\to\infty} F_{\mu_n}^{(\alpha)}(b) = F_\mu^{(\alpha)}(b)$ for $b \in \{b_1, b_2\}$. Moreover, $\mu_n((b_1,b_2]) = F_{\mu_n}^{(\alpha)}(b_2) - F_{\mu_n}^{(\alpha)}(b_1)$ for each $n \in \mathbb{N}$. Thus, by positivity of $\{\mu_n\}$,

$$\begin{aligned}
\limsup_{n\to\infty} \mu_n(K) &\leq \lim_{n\to\infty} \mu_n((b_1,b_2]) \\
&= \lim_{n\to\infty} F_{\mu_n}^{(\alpha)}(b_2) - \lim_{n\to\infty} F_{\mu_n}^{(\alpha)}(b_1) \\
&= F_\mu^{(\alpha)}(b_2) - F_\mu^{(\alpha)}(b_1) < \infty.
\end{aligned}$$

"(b) $\Rightarrow$ (a)". By Theorem 5.2.12, it suffices to sow that $\{\mu_n\}$ has no mass at the continuity points of μ. So let $x \in \mathbb{R}$ be a continuity point of μ and fix $\varepsilon > 0$. For $\delta > 0$, denote by $B_\delta(x)$ the open ball around x of radius δ and by

$\overline{B_\delta(x)}$ its closure. By σ-continuity of μ, for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\mu(\overline{B_\delta(x)}) \le \varepsilon$. Thus, by Theorem 5.1.23(b),

$$\limsup_{n\to\infty} \mu_n \left(B_\delta(x)\right) \le \limsup_{n\to\infty} \mu_n \left(\overline{B_\delta(x)}\right) \le \mu \left(\overline{B_\delta(x)}\right) \le \varepsilon. \qquad \square$$

Remark 5.2.13. The direction "(a) $\Rightarrow$ (b)" in Theorem 5.2.4 (for $\alpha \in \mathbb{R}$) follows also directly from "(a) $\Rightarrow$ (c)" in the vague Portmanteau Theorem; see Theorem 5.1.20.

Unlike Theorem 5.2.4, Theorem 5.2.11(b) requires an extra condition not in part (a). Fortunately, the assumption that $\{\mu_n\}$ has no mass at *any* point is sufficient to establish a proper equivalence result. Note that this slightly stronger assumption is equivalent to the original assumption in the important case that μ does not have any atoms.

Theorem 5.2.14. *Let $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}(\mathbb{R})$ and $\alpha \in \mathbb{R}$. Suppose $\{\mu_n\}$ does not have mass on any point of $\mathbb{R}$. Then the following are equivalent:*

1. $F_{\mu_n}^{(\alpha)} \to F_{\mu}^{(\alpha)}$ *at the continuity points of μ.*

2. $\text{v-}\lim_{n\to\infty} \mu_n = \mu$.

Moreover, if $\alpha = -\infty$ or $\alpha = +\infty$, the result remains to true under the additional assumption that $\{\mu_n\}$ has no mass at α (for the canonical extensions of μ_n).

Proof. We only establish the result for $\alpha \in \mathbb{R}$. The extension of the proof to $\alpha \in \{-\infty, \infty\}$ is straightforward.

By Theorem 5.2.11, it suffices to show that the assumption that $\{\mu_n\}$ has no mass on any point of $\mathbb{R}$ implies that $\{\mu_n\}$ is bounded on compact sets. So let $K \subset \mathbb{R}$ be a compact set.

By hypothesis, for each $x \in \mathbb{R}$, the exists an open neighbourhood N_x of x such that $\limsup_{n\to\infty} |\mu_n|(N_x) \le 1$. Moreover, by compactness, there exists $x_1, \ldots, x_J \in \mathbb{R}$ such that $K \subset \bigcup_{j=1}^J N_{x_j}$. It follows that

$$\limsup_{n\to\infty} |\mu_n|(K) \le \limsup_{n\to\infty} |\mu_n| \left(\bigcup_{j=1}^J N_{x_j}\right) \le \sum_{j=1}^J \limsup_{n\to\infty} |\mu_n|(N_{x_j}) \le J < \infty.$$

$\square$

We end this section by noting that the assumption that $\{\mu_n\}$ has no mass at any point of $\mathbb{R}$ is not enough to conclude from $\text{v-lim}_{n\to\infty} \mu_n = \mu$ that $\text{v-lim}_{n\to\infty} |\mu_n| = |\mu|$.

Example 5.2.15. For $n \in \mathbb{N}$, let $F_n : \mathbb{R} \to \mathbb{R}$ be supported on $[-1, 1]$ and linear between the points $\{k2^{-n} : k \in \{0, \dots, 2^n\}\}$ such that

$$F_n\left(k2^{-n}\right) := (k \bmod(2))2^{-n}, \quad k \in \{-2^n, \dots, 2^n\};$$

see Figure 5.3 for a clear visualisation. Set $\mu_n := \mu_{F_n}$ and let μ be the zero measure. Note that $|\mu_n| = |\mu_1|$ for each $n \in \mathbb{N}$. Hence it follows trivially that $\text{v-lim}_{n\to\infty} |\mu_n| = |\mu_1|$. However, using that $\|\mu_n\| = 2$ and $|F_n^{(0)}| \leq 2^{-n}$ for each $n \in \mathbb{N}$, it follows from Theorem 5.2.12 that $\text{v-lim}_{n\to\infty} \mu_n = \mu$. It remains to show that $\{\mu_n\}$ has no mass at any point of $\mathbb{R}$. So fix $x \in \mathbb{R}$ and let $\varepsilon > 0$ be given. Let $N_{x,\varepsilon}$ be the open ball around x of radius $\varepsilon/2$. Then

$$\limsup_{n\to\infty} |\mu_n| \left(N_{x,\varepsilon}\right) = \varepsilon. \tag{5.2.4}$$

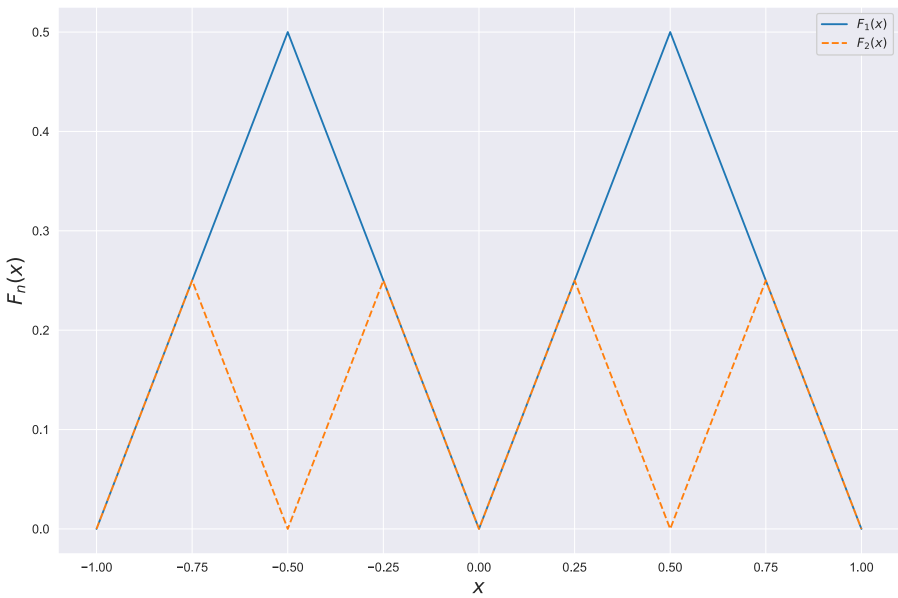


Figure 5.3: A visualisation of F_1 and F_2 defined in Example 5.2.15.

Chapter 6

Long Term Behaviour

It is well known that Laplace transforms of positive measures on $\mathbb{R}_+ := [0, \infty)$ converge if and only if their distribution functions converge at continuity points of the limiting measure; see e.g. [32, XIII.1, Theorem 2a]. Using the results of the previous chapter, we will extend this so-called continuity theorem to the case of real-valued Radon measures; this is the content of our preprint [40]. This will show us how far we can extend the famous Karamata Tauberian theorem to work for real-valued measures using König's approach alone and compare it to the approach used by Bingham et al.; see [51, 12]. We will apply the extended Tauberian theorem in a novel stochastic control problem.

6.1 Laplace Transforms

Throughout this chapter, we will take the underlying space to be $\mathbb{R}_+$. Moreover, we will consider measures in $\mathcal{M}^*$ whose Laplace transform is well defined on all of $\mathbb{R}_{++} := (0, \infty)$. Thus, we set

$$\mathcal{M}_\Psi := \left\{ \mu \in \mathcal{M}^* : \int_{\mathbb{R}_+} e^{-\lambda x} \left| \mu \right| (\mathrm{d}x) < \infty \text{ for all } \lambda > 0 \right\}.$$

For $\mu \in \mathcal{M}_\Psi$, we define its *Laplace transform* $\Psi_\mu \in \mathrm{BV}_{\mathrm{loc}}(\mathbb{R}_{++})$ by

$$\Psi_\mu(\lambda) := \int_{\mathbb{R}_+} e^{-\lambda x} \mu(\mathrm{d}x), \quad \lambda > 0.$$

It will be important that Laplace transforms characterise elements of $\mathcal{M}_\Psi$.

Proposition 6.1.1 (Characterisation via Laplace Transforms)**.** *Any $\mu \in \mathcal{M}_\Psi$ is uniquely determined by its Laplace transform Ψ_μ.*

Proof. First assume that $\mu \in \mathcal{M}$. Let $\mathcal{C}' := \{e^{-\lambda x} : \lambda > 0\}$ and $\mathcal{C}$ be the set of finite linear combinations of elements in $\mathcal{C}'$. Then $\mathcal{C}$ is a sub-algebra of $C_0(\mathbb{R}_+)$ that separates points and vanishes nowhere. Theorems E.1.11 and E.1.4 and the observation that a dense subset (for the topology of uniform convergence) of a separating family is again a separating family yields that $\mathcal{C}$ is a separating family for $\mathcal{M}$. Since the elements of $\mathcal{C}'$ correspond to Ψ_μ for different values of $\lambda > 0$, the result holds for all $\mu \in \mathcal{M}$.

Next, let $\mu \in \mathcal{M}_\Psi$. Fix $\varepsilon > 0$ and define $\mu^{(\varepsilon)} \in \mathcal{M}(\mathbb{R}_+)$ via $\mu^{(\varepsilon)}(\mathrm{d}x) := e^{-\varepsilon x}\mu(\mathrm{d}x)$. Then for $\lambda > 0$,

$$\Psi_\mu(\lambda + \varepsilon) = \Psi_{\mu^{(\varepsilon)}}(\lambda). \tag{6.1.1}$$

Since $\mu^{(\varepsilon)} \in \mathcal{M}$, it is uniquely determined by $\Psi_{\mu^{(\varepsilon)}}$. The latter is in turn uniquely determined by Ψ_μ by (6.1.1). Since μ is uniquely determined by $\mu^{(\varepsilon)}$ and ε, the result holds for all $\mu \in \mathcal{M}_\Psi$. □

Note that any $\mu \in \mathcal{M}_\Psi$ can be viewed as a measure on $\mathbb{R}$ where $\mu|_{(-\infty,0)} \equiv 0$. In order to be consistent with the previous chapter, we fix some $\eta < 0$ such that for any $\mu \in \mathcal{M}_\Psi$ the distribution function of μ is defined as $F_\mu := F_\mu^{(\eta)}$. In particular, for each $\mu \in \mathcal{M}_\Psi$ we view F_μ as an element of $\mathrm{BV}_{\mathrm{loc}}(\mathbb{R}_+)$ where

$$F_\mu(x) = \mu([0, x]).$$

6.2 Continuity Theorem

In this section, we state and prove a continuity theorem for real-valued Radon measures. This extends the classical continuity theorem from Feller [32, XIII.1, Theorem 2a]. They key tool will be vague convergence.

Theorem 6.2.1 (Continuity Theorem)**.** *Let $\{\mu_n\} \cup \{\mu\} \subset \mathcal{M}_\Psi$ be such for any $\lambda > 0$,*

$$\limsup_{n\to\infty} \Psi_{|\mu_n|}(\lambda) < \infty. \tag{6.2.1}$$

(a) *Suppose* $\Psi_{\mu_n}(\lambda) \to \Psi_\mu(\lambda)$ *for all* $\lambda > 0$. *If* $\{\mu_n\}$ *is right-equicontinuous at a continuity point* $x \in \mathbb{R}_+$ *of* μ *(see Definition 5.2.8), then* $F_{\mu_n}(x) \to F_\mu(x)$.

(b) *If* $F_{\mu_n} \to F_\mu$ *a.e., then* $\Psi_{\mu_n} \to \Psi_\mu$.

Proof. For $\nu \in \mathcal{M}^*$ and $\varepsilon > 0$, define $\nu^{(\varepsilon)} \in \mathcal{M}$ by $\nu^{(\varepsilon)}(\mathrm{d}x) := e^{-\varepsilon x}\nu(\mathrm{d}x)$. Then for each $\varepsilon > 0$, (6.2.1) and the fact that $\Psi_{|\mu_n|}(\varepsilon) < \infty$ for each $n \in \mathbb{N}$ gives

$$\sup_{n\in\mathbb{N}} \left\|\mu_n^{(\varepsilon)}\right\| = \sup_{n\in\mathbb{N}} \Psi_{|\mu_n|}(\varepsilon) < \infty. \tag{6.2.2}$$

Moreover, if in addition $\text{v-lim}_{n\to\infty}\, \mu_n^{(\varepsilon)} = \nu$ for some $\nu \in \mathcal{M}$, then for each $\lambda > 0$, (6.2.2), Proposition 5.1.11 and the fact that $\exp(-\lambda\cdot) \in C_0(\mathbb{R}_+)$ gives

$$\lim_{n\to\infty} \Psi_{\mu_n^{(\varepsilon)}}(\lambda) = \lim_{n\to\infty} \int_0^\infty e^{-\lambda t}\mu_n^{(\varepsilon)}(\mathrm{d}t) = \int_0^\infty e^{-\lambda t}\nu(\mathrm{d}t) = \Psi_\nu(\lambda) \tag{6.2.3}$$

(a) By a simple generalisation of Theorem 5.2.11, it suffices to show that

$$\operatorname*{v-lim}_{n\to\infty} \mu_n = \mu.$$

To establish the latter, it suffices to show that for each $\varepsilon > 0$,

$$\operatorname*{v-lim}_{n\to\infty} \mu_n^{(\varepsilon)} = \mu^{(\varepsilon)}, \tag{6.2.4}$$

Indeed, fix $\varepsilon > 0$ and $f \in C_c(\mathbb{R}_+)$. Then the vague convergence of $\{\mu_n^{(\varepsilon)}\}$ and the fact that $f\exp(-\varepsilon\cdot) \in C_c(\mathbb{R}_+)$ give

$$\int_{\mathbb{R}_+} f\ \mathrm{d}\mu_n = \int_{\mathbb{R}_+} f(x)e^{\varepsilon x}\mu_n^{(\varepsilon)}(\mathrm{d}x) \to \int_{\mathbb{R}_+} f(x)e^{\varepsilon x}\mu^{(\varepsilon)}(\mathrm{d}x) = \int_{\mathbb{R}_+} f\ \mathrm{d}\mu.$$

To establish (6.2.4), fix $\varepsilon > 0$. By the subsequence criterion, it suffices to show that for every subsequence $\{n_k\}$, there exists a further subsequence $\{n_{k_\ell}\}$ such that

$$\operatorname*{v-lim}_{\ell\to\infty} \mu_{n_{k_\ell}}^{(\varepsilon)} = \mu^{(\varepsilon)}. \tag{6.2.5}$$

So let $\{n_k\}$ be a subsequence. Then (6.2.1) gives

$$\sup_{k\in\mathbb{N}} \left\|\mu_{n_k}^{(\varepsilon)}\right\| = \sup_{k\in\mathbb{N}} \Psi_{|\mu_{n_k}|}(\varepsilon) < \infty.$$

Thus, by Theorem 5.1.8 we can view $\{\mu_{n_k}^{(\varepsilon)}\}$ as a bounded family in $(C_0(\mathbb{R}_+))^*$. Hence, Theorem E.1.9 implies that $\{\mu_{n_k}^{(\varepsilon)}\}$ is $\sigma((C_0(\mathbb{R}_+))^*, C_0(\mathbb{R}_+))$-compact. Furthermore, as a consequence of Theorem E.1.4, $C_0(\mathbb{R}_+)$ is separable, whence the sequence is sequentially compact. Thus, there exists a subsequence $\{n_{k_\ell}\}$ and $\tilde{\mu} \in \mathcal{M}$ such that

$$\text{v-}\lim_{\ell\to\infty} \mu_{n_{k_\ell}}^{(\varepsilon)} = \tilde{\mu} \tag{6.2.6}$$

We proceed to show that $\tilde{\mu} = \mu^{(\varepsilon)}$. Since by Proposition 6.1.1 each element in $\mathcal{M}_\Psi$ is uniquely characterised by its Laplace transform, it suffices to show that $\Psi_{\tilde{\mu}}(\lambda) = \Psi_{\mu^{(\varepsilon)}}(\lambda)$ for each $\lambda > 0$. So fix $\lambda > 0$. Then the hypothesis together with (6.2.3) (applied to the subsequence $\{n_{k_\ell}\}$) give

$$\Psi_{\mu^{(\varepsilon)}}(\lambda) = \Psi_\mu(\lambda+\varepsilon) = \lim_{\ell\to\infty} \Psi_{\mu_{n_{k_\ell}}}(\lambda+\varepsilon) = \lim_{\ell\to\infty} \Psi_{\mu_{n_{k_\ell}}^{(\varepsilon)}}(\lambda) = \Psi_{\tilde{\mu}}(\lambda).$$

(b) It suffices to show that for each $\varepsilon > 0$,

$$F_{\mu_n^{(\varepsilon)}} \to F_{\mu^{(\varepsilon)}} \text{ a.e.} \tag{6.2.7}$$

Indeed, Theorem 5.2.2(a) shows that $\text{v-}\lim_{n\to\infty}\mu_n^{(\varepsilon)} = \mu^{(\varepsilon)}$, and this in turn together with (6.2.3) yields for $\lambda > 0$,

$$\lim_{n\to\infty} \Psi_{\mu_n}(\lambda+\varepsilon) = \lim_{n\to\infty} \Psi_{\mu_n^{(\varepsilon)}}(\lambda) = \Psi_{\mu^{(\varepsilon)}}(\lambda) = \Psi_\mu(\lambda+\varepsilon).$$

Since ε is arbitrary, the claim follows.

To establish (6.2.7), fix $\varepsilon > 0$ and let $t > 0$ be such that $\lim_{n\to\infty} F_{\mu_n}(t) = F_\mu(t)$. An integration by parts gives

$$F_{\mu_n^{(\varepsilon)}}(t) = \int_{[0,t]} e^{-\varepsilon x}\mu(\mathrm{d}x) = e^{-\varepsilon t} F_{\mu_n}(t) + \varepsilon \int_0^t e^{-\varepsilon x} F_{\mu_n}(x)\,\mathrm{d}x.$$

Using that

$$\sup_{x\in[0,t]} \sup_{n\in\mathbb{N}} F_{|\mu_n|}(x) \le e^t \sup_{n\in\mathbb{N}} \Psi_{|\mu_n|}(1) < \infty$$

by (6.2.1) for $\lambda = 1$, the hypothesis, dominated convergence and an integration by parts give

$$\lim_{n\to\infty} F_{\mu_n^{(\varepsilon)}}(t) = e^{-\varepsilon t}F_\mu(t) + \varepsilon \int_0^t e^{-\varepsilon x} F_\mu(x)\,\mathrm{d}x = F_{\mu^{(\varepsilon)}}(t). \qquad \square$$

Remark 6.2.2. (i) One can also prove part (a) by using Helly's selection theorem [52, Theorem 2.35] to find a convergent subsequence of the family $\{F_{\mu_n^{(\varepsilon)}}\}$. The limiting function will be of bounded variation, and right continuous due to the right-equicontinuity condition. Thus, it will be the distribution function of a measure. The remainder of the proof is identical.

(ii) When the measures are positive, condition (6.2.1) is clearly satisfied under the hypothesis of part (a). Moreover, the right-equicontinuity condition of part (a) is no longer needed due to Theorem 5.2.4.

(iii) Under the assumption $\Psi_{\mu_n}(\lambda) \to \Psi_\mu(\lambda)$ for all $\lambda > 0$, (6.2.1) is trivially satisfied when the measures are positive. Otherwise, a sufficient condition for (6.2.1) is that there exists $\delta \in [0,1)$ such that either $\Psi_{\mu_n^-}(\lambda) < \delta\Psi_{\mu_n^+}(\lambda)$ for each $\lambda > 0$ or $\Psi_{\mu_n^+}(\lambda) < \delta\Psi_{\mu_n^-}(\lambda)$ for each $\lambda > 0$. We only establish the first case. Then

$$\begin{aligned}\limsup_{n\to\infty} \Psi_{|\mu_n|}(\lambda) \le \limsup_{n\to\infty} (1+\delta)\Psi_{\mu_n^+}(\lambda) &\le \left(\frac{1+\delta}{1-\delta}\right) \limsup_{n\to\infty} \Psi_{\mu_n}(\lambda) \\ &= \left(\frac{1+\delta}{1-\delta}\right) \Psi_\mu(\lambda) < \infty.\end{aligned}$$

The following example illustrates that the right-equicontinuity condition is indeed needed for Theorem 6.2.1(a).

Example 6.2.3. Let $\{\mu_n\} \cup \{\mu\} \in \mathcal{M}_\Psi$ be defined by $\mu_n := \delta_x - \delta_{x+\frac{1}{n}}$ for for some $x > 0$, and $\mu \equiv 0$. Note that $\{\mu_n\}$ is not right-equicontinuous at x. Indeed, for any $\delta > 0$

$$|\mu_n((x,\delta])| = 1 \quad \text{for} \quad n \ge \frac{1}{\delta}.$$

It is straightforward to check that

$$\limsup_{n\to\infty} \Psi_{|\mu_n|}(\lambda) = \limsup_{n\to\infty} \left(e^{-\lambda x} + e^{-\lambda\left(x+\frac{1}{n}\right)}\right) = 2e^{-\lambda x} < \infty, \quad \lambda > 0$$

and

$$\Psi_{\mu_n}(\lambda) = e^{-\lambda x} - e^{-\lambda\left(x+\frac{1}{n}\right)} \to 0 = \Psi_{\mu}(\lambda), \quad \lambda > 0.$$

However, x is a continuity point of μ and

$$F_{\mu_n}(x) = 1 \not\to 0 = F_{\mu}(x).$$

6.3 Application: Karamata's Tauberian Theorem

For this section it will be useful to recall the definition of a regularly varying function.

Definition 6.3.1. We say that a function $f : \mathbb{R}_+ \to \mathbb{R}$ is *regularly varying at infinity with exponent* $\rho \in \mathbb{R}$ if there exists some $a > 0$ such that $f|_{[a,\infty)}$ or $-f|_{[a,\infty)} \in \mathbb{R}_{++}$ and

$$\frac{f(\lambda x)}{f(x)} \to \lambda^{\rho} \quad \forall \lambda > 0. \tag{6.3.1}$$

Regular variation at zero is defined analogously. We call a regularly varying function with exponent $\rho = 0$ *slowly varying*.

Examples of slowly varying functions include the constant functions and $\log(\cdot)$, while polynomials of order ρ are the easiest examples of regularly varying functions with exponent ρ. In fact, the following theorem shows that all regularly varying functions behave much like polynomials.

Theorem 6.3.2 (Representation Theorem)**.** *A function $f : \mathbb{R}_+ \to \mathbb{R}$ is regularly varying with exponent $\rho \in \mathbb{R}$ if and only if there exists a slowly varying function l and some $a > 0$ such that $f(x) = x^{\rho} l(x)$ for all $x \geq a$ where*

$$l(x) = c(x) \exp\left\{\int_a^x \frac{\varepsilon(u)}{u} \,\mathrm{d}u\right\},$$

c and ε are bounded and measurable such that

$$c(x) \to \bar{c} \in \mathbb{R}\backslash\{0\} \quad \text{and} \quad \varepsilon(x) \to 0,$$

as $x \to \infty$.

In the study of regular variation, Karamata's Tauberian Theorem for Laplace-Stieltjes transforms is a classical result; see [31, Theorem 1], [12, Theorem 1.7.1], [51, Theorem 1]. It relates regular variation of a positive monotone functions F at infinity to the regular variation of its Laplace transforms Ψ_{μ_F} at zero. Due to the relationship between positive monotone functions and positive Radon measures, the theorem can be also stated for measures.

Theorem 6.3.3. *Let $\mu \in \mathcal{M}_\Psi$ be a positive measure and $\rho \geq 0$. The limit statements*

$$\lim_{\tau \downarrow 0} \frac{\Psi_\mu(\tau\lambda)}{\Psi_\mu(\tau)} = \frac{1}{\lambda^\rho}, \quad \lambda > 0, \tag{6.3.2}$$

and

$$\lim_{t \uparrow \infty} \frac{F_\mu(tx)}{F_\mu(t)} = x^\rho, \quad x > 0. \tag{6.3.3}$$

imply each other. In either case, we also have

$$\Psi_\mu(t^{-1}) \sim F_\mu(t)\Gamma(\rho+1) \quad \text{as } t \to \infty. \tag{6.3.4}$$

Remark 6.3.4. The theorem may be phrased as 'the Laplace transform of a positive measure is regularly varying at zero if and only if its distribution function is regularly varying at infinity'.

According to Feller, Theorem 6.3.3 has a 'glorious history' [32, Section XIII.5], even though modern books on probability theory often omit the theorem. The two implications are usually separated, where Eq. (6.3.3) $\Rightarrow$ Eq. (6.3.2) is called an *Abelian* theorem, while Eq. (6.3.2) $\Rightarrow$ Eq. (6.3.3) is called a *Tauberian* theorem. Looking at its origins, we see that the Tauberian implication caused the most difficulty. It was first proved via lengthy calculations in 1914 by Hardy and Littlewood in their famous paper [38]. Karamata simplified their proof in [46], subsequently introducing the present-day class of regularly varying functions.

6.3.1 Karamata's Tauberian theorem for real-valued measures

Karamata's theorem can be extended to functions of local bounded variation, and hence real-valued measures. While there are some results, see e.g. [12, Section 4.0, 5], they do not seem to be well known. They always require some additional conditions in the Tauberian direction, usually referred to as *Tauberian conditions*. The latter are needed to account for the lack of monotonicity of F_μ in the proof of Theorem 6.3.3.

We proceed to apply our continuity theorem to obtain a version of Karamata's Tauberian theorem for real-valued measures.

Theorem 6.3.5 (Karamata's Tauberian theorem for real-valued measures). *Let $\mu \in \mathcal{M}_\Psi$ and $\rho \geq 0$.*

(a) *Suppose that*

$$\liminf_{\tau \downarrow 0} \frac{|\Psi_\mu(\tau)|}{\Psi_{|\mu|}(\tau)} > 0 \tag{6.3.5}$$

as well as

$$\limsup_{h \downarrow 0} \limsup_{\tau \downarrow 0} \left| \frac{F_\mu(\tau^{-1}(x+h)) - F_\mu(\tau^{-1}x)}{\Psi_\mu(\tau)} \right| = 0, \quad x > 0. \tag{6.3.6}$$

Then (6.3.2) implies (6.3.3) and (6.3.4) holds.

(b) *(6.3.3) implies (6.3.2) and (6.3.4) holds.*

Moreover, in either case, we have the asymptotic relationship (6.3.4).

Remark 6.3.6. The limit statements (6.3.2) and (6.3.3) show that $\Psi_\mu(\tau)$ and $F_\mu(t)$ are non-zero for sufficiently small τ and sufficiently large t, respectively. In particular (6.3.2) implies that Ψ_μ has one sign near the origin, and (6.3.3) implies that F_μ has one sign near infinity.

Proofs of Theorem 6.3.5. Throughout the proof, let $\{\tau_n\} \subset \mathbb{R}_{++}$ be an arbitrary null sequence and define the sequence $\{t_n\} \subset (0, \infty)$ by $t_n := \tau_n^{-1}$.

(a) By (6.3.5), we may assume without loss of generality that $\Psi_\mu(\tau_n) \neq 0$ for all $n \in \mathbb{N}$. Thus, for each $n \in \mathbb{N}$, we may define $\nu_n \in \mathcal{M}_\Psi$ by

$$\nu_n(\mathrm{d}x) := \frac{\mu(\mathrm{d}(t_n x))}{\Psi_\mu(\tau_n)}.$$

Moreover, define the measure $\nu \in \mathcal{M}_\Psi$ by

$$\nu(\mathrm{d}x) := \begin{cases} \frac{x^{\rho-1}}{\Gamma(\rho)}\,\mathrm{d}x & \text{if } \rho > 0, \\ \delta_0(\mathrm{d}x) & \text{if } \rho = 0. \end{cases}$$

Then (6.3.2) gives

$$\lim_{n\to\infty} \Psi_{\nu_n}(\lambda) = \lim_{n\to\infty} \frac{\Psi_\mu(\tau_n\lambda)}{\Psi_\mu(\tau_n)} = \lambda^{-\rho} = \Psi_\nu(\lambda).$$

This together with (6.3.5) in turn yields

$$\limsup_{n\to\infty} \Psi_{|\nu_n|}(\lambda) = \limsup_{n\to\infty} \frac{\Psi_{|\mu|}(\tau_n\lambda)}{\Psi_\mu(\tau_n)} = \lambda^{-\rho} \limsup_{n\to\infty} \frac{\Psi_{|\mu|}(\tau_n\lambda)}{\Psi_\mu(\tau_n\lambda)} < \infty, \quad \lambda > 0.$$

Moreover, $\{\nu_n\}$ is right-equicontinuous at all points in $\mathbb{R}_{++}$. Indeed, let $x \in (0,\infty)$. Then (6.3.6) gives

$$\limsup_{h\downarrow 0} \limsup_{n\to\infty} |\nu_n((x, x+h])| = \limsup_{h\downarrow 0} \limsup_{\tau\downarrow 0} \frac{|F_\mu(\tau^{-1}(x+h)) - F_\mu(\tau^{-1}x)|}{\Psi_\mu(\tau)} \tag{6.3.7}$$

$$= 0.$$

It now follows from Theorem 6.2.1(a) that $F_{\nu_n} \to F_\nu$ on $\mathbb{R}_{++}$. Recalling the definition of ν_n, this implies that

$$\lim_{n\to\infty} \frac{F_\mu(t_n x)}{\Psi_\mu(\tau_n)} = \lim_{n\to\infty} F_{\nu_n}(x) = F_\nu(x) = \frac{x^\rho}{\Gamma(\rho+1)}, \quad x > 0. \tag{6.3.8}$$

Finally, (6.3.8) for $x = 1$ gives (6.3.4), and then combining (6.3.8) and (6.3.4) yields (6.3.3) via

$$\lim_{n\to\infty} \frac{F_\mu(t_n x)}{F_\mu(t_n)} = \lim_{n\to\infty} \frac{F_\mu(t_n x)}{\Psi_\mu(\tau_n)} \frac{\Psi_\mu(\tau_n)}{F_\mu(t_n)} = \frac{x^\rho}{\Gamma(\rho+1)}\Gamma(\rho+1) = x^\rho, \quad x > 0.$$

(b) By Remark, without loss of generality we choose $X > 0$ such that

$\mathbb{1}_{[X,\infty)}F_\mu$ is strictly positive. Define the positive measure $\xi \in \mathcal{M}_\Psi$ via

$$F_\xi(x) := \int_0^x \mathbb{1}_{[X,\infty)}(t)F_\mu(t)\ \mathrm{d}t. \tag{6.3.9}$$

By Theorem 6.3.2, (6.3.3) implies there exists a slowly varying function l such that $F_\mu(x) = x^\rho l(x)$. In particular, $\mathbb{1}_{[X,\infty)}(x)F_\mu(x) \sim x^\rho l(x)$, and so using [12, Proposition 1.5.8] it follows

$$F_\xi(x) \sim \frac{x^{\rho+1}l(x)}{(\rho+1)} \qquad \text{as } x \to \infty.$$

Since ξ is a positive measure, Theorem 6.3.3 lets us infer that $\Psi_\xi(\tau) \sim F_\xi(\tau^{-1})\Gamma(\rho+2)$ as $\tau \to 0$, whence

$$\Psi_\xi(\tau) \sim \Gamma(\rho+1)l(1/\tau)\tau^{-(\rho+1)} \qquad \text{as } \tau \to 0. \tag{6.3.10}$$

Noting that $\Psi_\xi(\lambda) = \frac{1}{\lambda}\int_X^\infty e^{-\lambda x}\mu(\mathrm{d}x)$, (6.3.10) implies

$$\Psi_\mu(\tau) \sim \Gamma(\rho+1)l(1/\tau)\tau^{-\rho} \qquad \text{as } \tau \to 0.$$

Equations (6.3.2) and (6.3.4) follow immediately. $\square$

We apply Theorem 6.3.5 in the following example.

Example 6.3.7. Define $\mu \in \mathcal{M}_\Psi$ via the density $f_\mu(x) := x\left(\frac{1}{2} + \cos(x)\right)$. Then one can readily check that

$$\Psi_\mu(\tau) = \frac{3\tau^4+1}{2(\tau^3+\tau)^2} = \frac{1}{2\tau^2} + o(\tau^{-1}), \qquad \tau \downarrow 0.$$

Thus, Ψ_μ is regularly varying with exponent $\rho = 2$. Noting that

$$\liminf_{\tau\downarrow 0} \frac{|\Psi_\mu(\tau)|}{\Psi_{|\mu|}(\tau)} \geq \liminf_{\tau\downarrow 0} \frac{\frac{1}{2\tau^2}(1+\tau o(\tau^{-1}))}{\frac{3}{2\tau^2}} = \frac{1}{3} > 0,$$

and

$$\limsup_{h\downarrow 0}\limsup_{\tau\downarrow 0}\left|\frac{F_\mu(\tau^{-1}(x+h)) - F_\mu(\tau^{-1}x)}{\Psi_\mu(\tau)}\right|$$

$$= \limsup_{h\downarrow 0} \limsup_{\tau\downarrow 0} \left| \frac{xh + h^2/2 + \tau^2 o(\tau^{-1})}{1 + \tau^2 o(\tau^{-1})} \right|$$
$$= 0,$$

we see that both (6.3.5) and (6.3.6) are satisfied. Hence, by Theorem 6.3.5 it follows that

$$\Psi_\mu(\mu)(\tau) \sim \Gamma(2+1) F_\mu(\tau^{-1}), \quad \tau \downarrow 0.$$

The method of proof of Theorem 6.3.5(a) is due to König [51] and Feller [31], which relies entirely on Theorem 6.2.1 for positive measures. In the case of real-valued measures, one needs the additional right-equicontinuity condition on the measures due to Example 5.2.7. As a byproduct, we get somewhat awkward looking Tauberian conditions (6.3.5) and (6.3.6). One can get an alternative condition by changing the approach.

Indeed, suppose we have $\mu \in \mathcal{M}_\Psi$ such that its Laplace transform is regularly varying with exponent ρ. Then as in part (b) of the proof of Theorem 6.3.6, we define the positive measure $\xi \in \mathcal{M}_\Psi$ via

$$F_\xi(x) := \int_0^x \mathbb{1}_{[X,\infty)}(t) F_\mu(t)\, \mathrm{d}t. \tag{6.3.11}$$

Then

$$\Psi_\xi(\lambda) = \frac{1}{\lambda} \int_X^\infty e^{-\lambda x} \mu(\mathrm{d}x) \sim \Psi_\mu(\lambda)\lambda^{-1},$$

whence Ψ_ξ is regularly varying at zero with exponent $-(\rho+1)$. By Theorem 6.3.3 it follows that F_ξ is regularly varying at infinity with exponent $(\rho+1)$. Now one can focus on extracting whether or not F_μ is regularly varying with exponent ρ from the integral relationship (6.3.11). It turns out that Bingham et al. show that this holds in the affirmative under the following Tauberian condition

$$\lim_{\lambda\downarrow 1} \liminf_{x\to\infty} \inf_{t\in[1,\lambda]} \frac{F_\mu(tx) - F_\mu(x)}{x^\rho l(x)} \geq 0, \tag{6.3.12}$$

where l is taken from Theorem 6.3.2 with respect to F_ξ; see [12, Theorem 1.7.5]. Thus, by only considering the relationship between F_ξ and F_μ we can bypass using Theorem 6.2.1 for measures which are not positive, and replace conditions (6.3.5) and (6.3.6) with (6.3.12) in Theorem 6.3.5. This results in the following proposition.

Proposition 6.3.8. *Let $\mu \in \mathcal{M}_\Psi$ and $\rho \geq 0$. Suppose that F_μ satisfies (6.3.12). Then (6.3.2) implies (6.3.3) and (6.3.4) holds.*

6.4 Ergodic-type Stochastic Control

We now aim to show how the content of the previous section gives rise to a novel stochastic control problem. Note that we intend the following formulation to be motivational and explained using a toy setup.

Consider a stochastic basis $(\Omega, \mathbb{F} := \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$, which we associate with a d-dimensional Brownian motion W and a space $\boldsymbol{A}$ of admissible processes (or controls) that are at minimum $\mathcal{F}$-progressively measurable and take values in a set $A \subseteq \mathbb{R}^k$. In general, we apply other conditions to the elements of $\boldsymbol{A}$ depending on further specifications of the control problem. For any $\alpha \in \boldsymbol{A}$ we consider an $\mathbb{R}^d$-valued state process $\{X_t^{\alpha,x}\}_{t\geq 0}$ with diffusion dynamics

$$\mathrm{d}X_t^{\alpha,x} = \mu(X_t^{\alpha,x}, \alpha_t)\ \mathrm{d}t + \sigma(X_t^{\alpha,x}, \alpha_t)\ \mathrm{d}W_t, \quad X_0 = x. \tag{6.4.1}$$

It is implicitly assumed that both $\mu : \mathbb{R}^d \times A \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \times A \to \mathbb{R}^d$ allow for a unique strong solution to (6.4.1). Finally, for the above system we consider a cost functional

$$J : \mathbb{R}^d \times A \times (0, \infty] \times [0, \infty) \to \mathbb{R}$$

defined by

$$J(x, \alpha, T, \delta) := \mathbb{E}\left[\int_0^T e^{-\delta t} f(X_t^{\alpha,x}, \alpha_t)\ \mathrm{d}t\right].$$

We will be concerning ourselves with an ergodic-type control problem, that is a control problem that only depends on the asymptotic behavior of the state. A common example is the so called *long run average* problem [4], that involves finding $\alpha^* \in \boldsymbol{A}$ that maximises

$$\limsup_{T\uparrow\infty}\left[\frac{1}{T}J(x, \alpha, T, 0)\right]. \tag{6.4.2}$$

One often solves this problem via the solution of the *discounted infinite time horizon problem*[1]

$$u(x, \delta) := \sup_{\alpha \in \boldsymbol{A}} J(x, \alpha, \infty, \delta). \tag{6.4.3}$$

Indeed, assuming (6.4.3) is well defined, a common technique to solving the long run average problem is to find an appropriate null sequence $\{\delta_n\}$ such that $\delta_n u(x, \delta_n)$ converges to a constant λ independent of x, and argue that it is the solution of (6.4.2). This is often dubbed the *vanishing discount paradigm* (see [3, Section 2.7]) and is used in both analytical and BSDE approaches to the problem [34, 4, 5].

In all cases, one enforces restrictions on the state processes in order to find a solution. For instance, the BSDE approach to solving (6.4.2) is associated with the theory of ergodic backward stochastic differential equations (EBSDE), detailed in the paper by Furhman, Hu, and Tessitore [34]. For their setup to make sense, one requires a so-called *dissipative* condition on the drift component of (6.4.1). This enforces the state process to be *ergodic*, giving it beneficial asymptotic characteristics; for a detailed discussion on ergodic processes, please see [3, Section 1.5].

What is not included in the current literature is that in some cases one can find an asymptotic relationships between the discounted infinite time horizon problem (6.4.3) and a weighted control problem of the type

$$\sup_{\alpha \in \boldsymbol{A}} \limsup_{T \uparrow \infty} \left[\frac{1}{T^\rho} J(x, \alpha, T, 0) \right], \tag{6.4.4}$$

for some $\rho > 0$ where ρ need not be 1. Indeed, let us consider the setup from Chapter 2 in the case where we have $N = 2$ agents, a single risky asset, a finite time horizon T, a tax levy $\lambda > 0$, and volatilities ζ_t^n given by $\beta_n W_t$ for $n = 1, 2$. Furthermore, let us consider a new set of admissible portfolios $\tilde{\mathcal{A}}$ which consists of all portfolios $\varphi \in \mathrm{AC}(\Omega \times [0, T], \mathbb{R})$ with $\varphi, \dot{\varphi} \in \mathscr{L}_0^2$ satisfying the transversality condition

$$\lim_{T \to \infty} \frac{1}{T^2} \mathbb{E}\left[\varphi_T^2\right] = 0.$$

[1]The set of admissible controls may in fact differ and depend on δ.

In the paper [35], they find a long run equilibrium return and optimal strategies according to the following definition.

Definition 6.4.1. $\mu \in \mathscr{L}_0^2$ is a (long-run) equilibrium return if there exist portfolios $\bar{\eta}^1, \bar{\eta}^2 \in \tilde{\mathcal{A}}$ for agents 1 and 2 such that:

(Market clearing) The total demand $\sum_{n=1}^{2} \bar{\eta}^n$ matches the zero net supply of the risky asset S at all times;

(Individual optimality) The portfolio $\bar{\eta}^n$ is optimal for the long-run version of agent n 's control problem in that[2]

$$\limsup_{T\to\infty} \frac{1}{T} \left[K_n^0(\varphi, \lambda, 0, T) - K_n^0(\bar{\eta}^n, \lambda, 0, T)\right] \leq 0, \tag{6.4.5}$$

for all competing admissible trading rates $\varphi \in \tilde{\mathcal{A}}$.

The condition (6.4.5) is, to the best of our knowledge, unique to the paper at hand. It is *not* the traditional long run average problem as expressed in (6.4.2), and its form is a fix to amend for the fact that are admissible portfolios (such as $\varphi \equiv 0$) where

$$\limsup_{T\to\infty} \left[\frac{1}{T} K_n^0(\varphi, \lambda, 0, T)\right] = \infty.$$

Despite the seemingly artificial requirement (6.4.5), we get the following asymptotic relationship between the equilibrium strategies from Definition 2.2.1 and Definition 6.4.5

$$\lim_{\delta\downarrow 0} \left[\delta^2 K_n^0(\bar{\varphi}^{\lambda,n}, \lambda, \infty, 0)\right] = \Gamma(2+1) \lim_{T\to\infty} \left[\frac{1}{T^2} K_n^0(\bar{\eta}^n, \lambda, 0, T)\right], \tag{6.4.6}$$

which is shown by explicit calculation in the notebook `long_term_behaviour`. It is even true that $\bar{\varphi}^{\lambda,n}$, as derived in Proposition 3.1.1, converges pointwise to $\bar{\eta}^n$ for $n = 1, 2$, as $\delta \downarrow 0$. This suggests that under certain conditions the vanishing discount technique may extend in order to solve the novel control problem (6.4.4) where the underlying state process is not restricted to be ergodic. Furthermore, it suggests that it is reasonable to replace the condition

[2] Recall that K_n^0 is defined in (2.1.7).

(6.4.5) with finding a portfolio φ that maximises

$$\limsup_{T\to\infty}\left[\frac{1}{T^2}K_n^0(\varphi,\lambda,0,T)\right].$$

6.4.1 Goal functional relationship

In order to investigate the problem further, we ask under what conditions does the goal functional $J(x,\alpha,T,0)$ display an asymptotic relationship with $J(x,\alpha,\infty,\delta)$ similar to that seen in (6.4.6). Considering Karamata's Tauberian theorem, we note that (6.3.4) bears a clear resemblance to (6.4.6), motivating the following observation. Assuming that both $J(x,\alpha,T,0)$ and $J(x,\alpha,\infty,\delta)$ are well defined for all T and δ in $\mathbb{R}_{++}$, by an application of Fubini's theorem we see that

$$J(x,\alpha,T,0)=\int_0^T \mathbb{E}[f(X_t^{\alpha,x},\alpha_t)]\;\mathrm{d}t =: F_{\mu^\alpha}(T), \tag{6.4.7}$$

$$J(x,\alpha,\infty,\delta)=\int_0^\infty e^{-\delta t}\mathbb{E}[f(X_t^{\alpha,x},\alpha_t)]\;\mathrm{d}t =: \Psi_{\mu^\alpha}(\delta), \tag{6.4.8}$$

where μ^α is an element of $\mathcal{M}_\Psi$ and has the density $f_{\mu^\alpha} := \mathbb{E}[f(X_\cdot^{\alpha,x},\alpha_\cdot)]$ with respect to the Lebesgue measure. It is now clear that if a control α is selected such that the measure μ^α is positive, and if F_{μ^α} (resp. Ψ_{μ^α}) is regularly varying with exponent $\rho\geq 0$ at infinity (resp. zero), then according to Theorem 6.3.3

$$\lim_{\delta\downarrow 0}\left[\delta^\rho J(x,\alpha,\infty,\delta)\right]=\lim_{T\uparrow\infty}\left[\frac{\Gamma(\rho+1)}{T^\rho}J(x,\alpha,T,0)\right].$$

In general we have the following proposition.

Proposition 6.4.2. *Suppose that F_{μ^α} is regularly varying at infinity with exponent $\rho\geq 0$. Then there exists a slowly varying function l such that*

$$\lim_{\delta\downarrow 0}\left[\frac{\delta^\rho}{l(\delta^{-1})}J(x,\alpha,\infty,\delta)\right]=\lim_{T\uparrow\infty}\left[\frac{\Gamma(\rho+1)}{T^\rho l(T)}J(x,\alpha,T,0)\right]. \tag{6.4.9}$$

Similarly, if Ψ_{μ^α} is regularly varying at zero with exponent $\rho\geq 0$ and either (6.3.12) or (6.3.5) and (6.3.6) are satisfied, then there exists a slowly varying function l such that (6.4.9) holds.

Suppose that one has an optimal control α^* for the discounted infinite time horizon control problem that is independent of the discount factor δ and such that α^* lies in the admissible set $\boldsymbol{A}$ associated to (6.4.4). Furthermore, suppose that $\Psi_{\mu^{\alpha^*}}(\delta) = u(x, \delta)$ is regularly varying at zero with exponent $\rho \geq 0$ and either (6.3.12) or (6.3.5) and (6.3.6) are satisfied. Then from Proposition 6.4.2 there exists a slowly varying function l such that

$$\begin{aligned}
\sup_{\alpha \in \boldsymbol{A}} \limsup_{T \uparrow \infty} \left[\frac{1}{l(T) T^{\rho}} J(x, \alpha, T, 0) \right] &\geq \limsup_{T \uparrow \infty} \left[\frac{1}{l(T) T^{\rho}} J(x, \alpha^*, T, 0) \right] \\
&= \limsup_{\delta \to 0} \left[\frac{\delta^{\rho}}{\Gamma(\rho+1) l(\delta^{-1})} J(x, \alpha^*, \infty, \delta) \right] \\
&= \lim_{\delta \to 0} \left[\frac{\delta^{\rho}}{\Gamma(\rho+1) l(\delta^{-1})} u(x, \delta) \right].
\end{aligned}$$

In particular, if one can take $l \equiv 1$, then we get a lower bound for the control problem (6.4.4). To apply a similar method to prove the opposite inequality is more difficult as the value function (6.4.4) is independent of T and the goal functional (6.4.7) is not necessarily regularly varying for an arbitrary control.

In general, when attempting to show either inequality, one will not be able to assume that admissible controls, and thus the integrands in (6.4.7) and (6.4.8), are independent of the time horizon or discount factor. This fundamentally complicates the setup. However, with stronger assumptions, one sees that a solution to the discounted infinite time horizon control problem (6.4.3) leads to a solution of (6.4.4) as shown by the following theorem.

Theorem 6.4.3. *Fix $\rho > 0$. Suppose that for all $\alpha \in \boldsymbol{A}$*

$$\lim_{T \to \infty} \frac{J(x, \alpha, T, 0)}{T^{\rho}} \tag{6.4.10}$$

exists and lies in $\mathbb{R}_{++}$. If an optimal control $\alpha^* \in \boldsymbol{A}$ exists for (6.4.3), then α^* is an optimal control for (6.4.4) and

$$\lim_{\delta \downarrow 0} \left[\frac{\delta^{\rho}}{\Gamma(\rho+1)} J(x, \alpha^*, \delta, \infty) \right] = \sup_{\alpha \in \boldsymbol{A}} \limsup_{T \uparrow \infty} \left[\frac{1}{T^{\rho}} J(x, \alpha, T, 0) \right].$$

Proof. By Proposition 6.4.2

$$\begin{aligned}
\sup_{\alpha \in \boldsymbol{A}} \limsup_{T \uparrow \infty} \left[\frac{1}{T^\rho} J(x, \alpha, T, 0) \right] &= \sup_{\alpha \in \boldsymbol{A}} \lim_{T \uparrow \infty} \left[\frac{1}{T^\rho} J(x, \alpha, T, 0) \right] \\
&= \sup_{\alpha \in \boldsymbol{A}} \lim_{\delta \downarrow 0} \left[\frac{\delta^\rho}{\Gamma(\rho + 1)} J(x, \alpha, \infty, \delta) \right] \\
&\leq \lim_{\delta \downarrow 0} \left[\frac{\delta^\rho}{\Gamma(\rho + 1)} J(x, \alpha^*, \infty, \delta) \right] \\
&= \lim_{T \uparrow \infty} \left[\frac{1}{T^\rho} J(x, \alpha^*, T, 0) \right] \\
&\leq \sup_{\alpha \in \boldsymbol{A}} \limsup_{T \uparrow \infty} \left[\frac{1}{T^\rho} J(x, \alpha, T, 0) \right].
\end{aligned}$$

□

The previous observations warrant further exploration into this new class of control problems, possibly leading to a new class of BSDE. In particular, this may allow us to rigorously explain the relationship between the control problems introduced in Definitions 2.2.1 and 6.4.5. For now, we showcase a definite class of state processes and specific goal functionals that satisfy (6.4.9).

Example 6.4.4. Let W be a Brownian motion on $(\Omega, \mathbb{F}, \mathbb{P})$. Moreover, suppose that $f : \mathbb{R}_+ \to \mathbb{R}_+$ is defined such that $f(x) = x^\rho l(x)$ for some $\rho > 0$ and a slowly varying function l. Then, fixing $\lambda > 0$, it follows that

$$\frac{\mathbb{E}\left[f\left(|W_{\lambda t}|\right)\right]}{\mathbb{E}\left[f\left(|W_t|\right)\right]} = \lambda^{\rho/2} \frac{\mathbb{E}\left[|Z|^p l(|Z|\sqrt{\lambda t})\right]}{\mathbb{E}\left[|Z|^p l(|Z|\sqrt{t})\right]},$$

where $Z \sim N(0,1)$. By Theorem 6.3.2, l must be of the form

$$l(x) = c(x) \exp\left\{ \int_a^x \frac{\epsilon(x)}{u} \, \mathrm{d}u \right\}$$

for $x \geq a > 0$, where c and ϵ are both bounded and measurable such that $c(x) \to \bar{c} \in (0, \infty)$, and $\epsilon(x) \to 0$ as $x \uparrow \infty$. Supposing that ϵ is bounded by $b > 0$, it follows that for any $z \geq \frac{1}{\sqrt{\lambda}}$ and $\sqrt{t} > a > 0$ there exists some constant $M > 0$ such that

$$
\begin{aligned}
\frac{l(\sqrt{\lambda t}z)}{l(\sqrt{t})} &\le M \exp\left\{\int_{\sqrt{t}}^{\sqrt{\lambda t}z} \frac{\epsilon(u)}{u}\,\mathrm{d}u\right\} \\
&\le M \exp\left\{b \log\left(\frac{\sqrt{\lambda t}z}{\sqrt{t}}\right)\right\} \\
&= M\lambda^{\frac{b}{2}} z^b
\end{aligned} \tag{6.4.11}
$$

Fix $\varepsilon > 0$. Then one may find $T(\varepsilon) > \frac{1}{\sqrt{\lambda}} \vee a \vee 1$ such that

$$
\left| \int_0^\infty z^{\rho+b}\varphi(z)\mathrm{d}z - \int_0^{T(\varepsilon)} z^{\rho+b}\varphi(z)\mathrm{d}z \right| \le \varepsilon, \tag{6.4.12}
$$

and

$$
\left| \int_0^\infty z^{\rho}\varphi(z)\mathrm{d}z - \int_0^{\sqrt{\lambda}T(\varepsilon)} z^{\rho}\varphi(z)\mathrm{d}z \right| \le \varepsilon, \tag{6.4.13}
$$

where φ is the pdf of a standard normal random variable. Furthermore, one may use the uniform convergence theorem for regularly varying functions [12, Theorem 1.5.2] to find $t(T(\varepsilon)) > T(\varepsilon)^2$ such that

$$
\lambda^{\frac{\rho}{2}} w^\rho \left| \frac{l(\sqrt{t\lambda}w)}{l(\sqrt{t})} - 1 \right| \le \varepsilon
$$

for all $w \in (0, T(\varepsilon)]$ when $t > t(T(\varepsilon))$. Then, using (6.4.11) and (6.4.12), we see that for large enough t

$$
\begin{aligned}
\frac{\mathbb{E}\left[|Z|^p l(|Z|\sqrt{\lambda t})\right]}{2l(\sqrt{t})} &= \int_0^{T(\varepsilon)} z^\rho \frac{l(\sqrt{\lambda t}z)}{l(\sqrt{t})}\varphi(z)\mathrm{d}z + \int_{T(\varepsilon)}^\infty z^\rho \frac{l(\sqrt{\lambda t}z)}{l(\sqrt{t})}\varphi(z)\mathrm{d}z \\
&= \int_0^{T(\varepsilon)} \lambda^{-\frac{\rho}{2}} \left(\lambda^{\frac{\rho}{2}} z^\rho + \varepsilon\right)\varphi(z)\mathrm{d}z + \int_{T(\varepsilon)}^\infty z^\rho \frac{l(\sqrt{\lambda t}z)}{l(\sqrt{t})}\varphi(z)\mathrm{d}z \\
&\le \int_0^\infty z^\rho\varphi(z)\mathrm{d}z + \varepsilon\left(\lambda^{-\frac{\rho}{2}} + M\lambda^{\frac{b}{2}}\right).
\end{aligned}
$$

Similarly, using (6.4.13), we see that for large enough t

$$\begin{aligned}
\frac{\mathbb{E}\left[|Z|^p l(|Z|\sqrt{t})\right]}{2l(\sqrt{t})} &= \int_0^{\sqrt{\lambda}T(\varepsilon)} z^\rho \frac{l(\sqrt{t}z)}{l(\sqrt{t})}\varphi(z)\mathrm{d}z + \int_{\sqrt{\lambda}T(\varepsilon)}^\infty z^\rho \frac{l(\sqrt{\lambda t}z)}{l(\sqrt{t})}\varphi(z)\mathrm{d}z \\
&\geq \int_0^{\sqrt{\lambda}T(\varepsilon)} (z^\rho - \varepsilon)\varphi(z)\mathrm{d}z \\
&\geq \int_0^\infty z^\rho \varphi(z)\mathrm{d}z - 2\varepsilon.
\end{aligned}$$

Thus, it follows that

$$\begin{aligned}
\limsup_{t\to\infty} \frac{\mathbb{E}\left[f\left(|W_{\lambda t}|\right)\right]}{\mathbb{E}\left[f\left(|W_t|\right)\right]} &\leq \lambda^{\rho/2} \limsup_{\varepsilon\to 0} \left\{ \frac{\int_0^\infty z^\rho \varphi(z)\mathrm{d}z + \varepsilon\left(\lambda^{-\frac{\rho}{2}} + M\lambda^{\frac{b}{2}}\right)}{\int_0^\infty z^\rho \varphi(z)\mathrm{d}z - 2\varepsilon} \right\} \\
&= \lambda^{\rho/2}.
\end{aligned}$$

By an analogous argument, we also have

$$\liminf_{t\to\infty} \frac{\mathbb{E}\left[f\left(|W_{\lambda t}|\right)\right]}{\mathbb{E}\left[f\left(|W_t|\right)\right]} \geq \lambda^{\rho/2}$$

whence

$$\lim_{t\to\infty} \frac{\mathbb{E}\left[f\left(|W_{\lambda t}|\right)\right]}{\mathbb{E}\left[f\left(|W_t|\right)\right]} = \lambda^{\rho/2}$$

Thus, if we take our state process (independent of any control) to be $X_t^\alpha = W_t + x$, then one clearly has that

$$J(x,\alpha,T,\delta) = \mathbb{E}\left[\int_0^T e^{-\delta t} f(X_t^\alpha)\ \mathrm{d}t\right].$$

is regularly varying at infinity with exponent $\frac{\rho+2}{2}$ and the relationship (6.4.9) holds with $l \equiv 1$.

Appendix A

A.1 Agents' Beliefs: Technicalities

This section contains a series of technical definitions in order to state a general version of the Girsanov Theorem, which is needed in Chapter 2. The definitions and results are taken from [10, Sections 1.3 & 3.9].

A.1.1 The Natural Conditions

For a measure μ on the measurable space $(\Omega, \mathcal{F})$, we let μ^* denotes its outer-measure. Recall that we say that $N \subset \Omega$ is μ-negligible if $\mu^*(N) = 0$.

Definition A.1.1. Consider a stochastic basis $(\Omega, \mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ We close the filtration at ∞ with three objects,

1. the algebra of sets
$$\mathcal{A}_\infty := \bigcup_{0\leq t<\infty} \mathcal{F}_t.$$

2. the σ-algebra
$$\mathcal{F}_\infty := \bigvee_{0\leq t<\infty} \mathcal{F}_t,$$

3. and the universal completion $\mathcal{F}^*_\infty$ of $\mathcal{F}_\infty$, i.e. the collection of all sets in $\mathcal{F}_\infty$ that are μ-measurable for every $\mu \in \mathcal{M}(\Omega, \mathcal{F}_\infty)$.

We will say that the filtration $\mathbb{F}$ is universally complete if $\mathcal{F}_t$ is universally complete for any $t < \infty$.

Definition A.1.2. Consider a stochastic basis $(\Omega, \mathbb{F}, \mathbb{P})$.

(i) A subset A of Ω is $\mathbb{P}$-*nearly empty* if there exists a family $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{A}_\infty$ such that $A \subset \bigcup_{n\in\mathbb{N}} A_n$ and $\bigcup_{n\in\mathbb{N}} A_n$ is $\mathbb{P}$-negligible.

(ii) A property $\mathscr{P}$ of the point $\omega \in \Omega$ is said to hold *nearly* if the set N of points of Ω were it does not hold is nearly empty.

(iii) Two processes X and Y are *indistinguishable* if the set $\{X_. \neq Y_.\} \subseteq \Omega$ is nearly empty.

One can think of a nearly empty set as being a set such that someone can measure it and assert that it is negligible in finite time or a countable union of such sets. If one must wait an infinite amount of time (i.e. check whether $N \in \mathcal{F}_\infty$) to check if N is negligible, then N is not nearly empty even though it may be negligible.

With the current definition of indistinguishability, there may exist measurable indistinguishable processes X and Y, where one is adapted, and the other is not. This becomes impossible by demanding that the filtration we use is regular.

Definition A.1.3. Let $\mathfrak{P}$ be a family of probability measures on the filtration $(\Omega, \mathbb{F})$.

(i) For any $\mathbb{P} \in \mathfrak{P}$, set

$$\begin{aligned}\mathcal{F}_t^{\mathbb{P}} &:= \{A \subset \Omega : \exists A_{\mathbb{P}} \in \mathcal{F}_t \text{ such that } A \triangle A_{\mathbb{P}} \text{ is } \mathbb{P}\text{-nearly empty}\} \\ &= \sigma\left(\mathcal{F}_t \cup \{N : N \text{ is } \mathbb{P}\text{-nearly empty}\}\right).\end{aligned}$$

We call $\{\mathcal{F}_t^{\mathbb{P}}\}_{t\geq 0}$ the $\mathbb{P}$-regularisation of $\mathbb{F}$. The filtration $\mathbb{F}^{\mathfrak{P}}$ composed of the σ-algebras

$$\mathcal{F}_t^{\mathfrak{P}} := \bigcap_{\mathbb{P}\in\mathfrak{P}} \mathcal{F}_t^{\mathbb{P}},$$

is the $\mathfrak{P}$-regularisation of $\mathcal{F}$.

(ii) The filtered space $(\Omega, \mathbb{F}, \mathfrak{P})$ is regular if $\mathbb{F} = \mathbb{F}^{\mathfrak{P}}$ and we say that $\mathbb{F}$ is $\mathfrak{P}$-regular, or regular when $\mathfrak{P}$ is understood.

We are now in a position to define the *natural conditions* for a filtration.

Definition A.1.4. Let $\mathfrak{P}$ be a collection of probability measures on the filtration $(\Omega, \mathbb{F})$.

(i) The *natural enlargement* of $\mathbb{F}$ is the filtration $\mathbb{F}_+^{\mathfrak{P}}$ obtained by regularising the right continuous version of $\mathbb{F}$.

(ii) Suppose that X is a process. Then the natural enlargement of the raw filtration $\mathbb{F}^0[X]$ is called the *natural filtration* of X and is denoted by $\mathbb{F}[X]$. If $\mathfrak{P}$ must be mentioned then we write $\mathbb{F}^{\mathfrak{P}}[X]$.

(iii) A filtered space is said to satisfy the *natural conditions* if it equals its natural enlargement.

If we were to complete $\mathbb{F}$ with respect to the *usual conditions*, then we would enlarge it by taking its right continuous version and throw all $\mathbb{P}$-negligible sets of $\mathcal{F}_\infty$ into $\mathcal{F}_{t+}$ for every $t < \infty$. Thus, a probability measure that is absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}_0$ is automatically absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}_\infty$. Failure to observe this has led to erroneous versions of Girsanov's theorem.

Definition A.1.4 furnishes advantages such as path regularity of integrators and a plentiful supply of stopping times, despite not generally containing every negligible set of $\mathcal{F}_\infty$. For instance, the Debut theorem holds for any progressively measurable set.

We end this subsection by defining a local version of the notion of equivalence for probability measures.

Definition A.1.5. Let $\mathfrak{P}$ be a family of probability measures on the filtration $(\Omega, \mathbb{F})$.

(i) For any $\mathbb{P} \in \mathfrak{P}$ we let $\mathbb{P}_t$ denote its restriction to $\mathcal{F}_t$ for any $t < \infty$.

(ii) Let $\mathbb{P} \in \mathfrak{P}$. A probability measure $\mathbb{Q}$ on $\mathcal{F}_\infty$ is called *locally absolutely continuous* with respect to $\mathbb{P}$ if $\mathbb{Q}_t$ is absolutely continuous with respect to $\mathbb{P}_t$. In this case we write $\mathbb{Q} \ll_{\text{loc}} \mathbb{P}$.

(iii) If $\mathbb{Q} \ll_{\text{loc}} \mathbb{P}$ and $\mathbb{P} \ll_{\text{loc}} \mathbb{Q}$, we say that $\mathbb{P}$ and $\mathbb{Q}$ are *locally equivalent* and write $\mathbb{Q} \sim_{\text{loc}} \mathbb{P}$.

Remark A.1.6. If $\mathbb{Q} \ll_{\text{loc}} \mathbb{P}$, then all $\mathbb{P}$-nearly empty sets are $\mathbb{Q}$-nearly empty.

A.1.2 The Girsanov Theorem

We now consider how much a martingale under $\mathbb{Q} \sim \mathbb{P}$ deviates from being a $\mathbb{P}$-martingale. We will assume throughout that the underlying filtration satisfies the natural conditions under $\mathbb{P}$, and hence $\mathbb{Q}$.

Note that $\mathbb{Q}$ is not necessarily absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}_\infty$, whence we cannot define the standard density process of $\mathbb{Q}$ with respect to $\mathbb{P}$. However, due to local absolute continuity and the Radon-Nikodym theorem, there exists the random variables $Z_t^{\mathbb{Q}} := \frac{\mathrm{d}\mathbb{Q}_t}{\mathrm{d}\mathbb{P}_t}$ and $Z_t^{\mathbb{P}} := \frac{\mathrm{d}\mathbb{P}_t}{\mathrm{d}\mathbb{Q}_t}$ for each $t \geq 0$ such that $Z_t^{\mathbb{Q}}$ is a $\mathbb{P}$-martingale and $Z_t^{\mathbb{P}}$ is a $\mathbb{Q}$-martingale. Moreover, we can choose both $Z_t^{\mathbb{Q}}$ and $Z_t^{\mathbb{P}}$ to be right-continuous, strictly positive such that $Z_t^{\mathbb{Q}} Z_t^{\mathbb{P}} \equiv 1$. Note that $\mathbb{Q}$ is absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}_\infty$ if and only if $Z^{\mathbb{Q}}$ is uniformly $\mathbb{P}$-integrable.

Lemma A.1.7 (Girsanov-Meyer)**.** *Suppose M is a local $\mathbb{Q}$-martingale. Then $MZ^{\mathbb{Q}}$ is a local $\mathbb{P}$-martingale, and*

$$M = \left(M_0 - Z^{\mathbb{P}}_{\cdot\_} \bullet \left[M, Z^{\mathbb{Q}}\right]\right) + \left(Z^{\mathbb{P}}_{\cdot\_} \bullet \left(MZ^{\mathbb{Q}}\right) - (MZ^{\mathbb{P}})_{\cdot\_} \bullet Z^{\mathbb{Q}}\right).$$

Reversing the roles of $\mathbb{P}$ and $\mathbb{Q}$ we see that if M is a local $\mathbb{P}$-martingale then

$$\begin{aligned} M - Z^{\mathbb{P}}_{\cdot\_} \bullet \left[M, Z^{\mathbb{Q}}\right] &= M + Z^{Q}_{\cdot\_} \bullet \left[M, Z^{\mathbb{P}}\right] \\ &= M_0 + Z^{\mathbb{Q}}_{\cdot\_} \bullet \left(MZ^{\mathbb{P}}\right) - \left(MZ^{\mathbb{Q}}\right)_{\cdot\_} \bullet Z^{\mathbb{P}}. \end{aligned} \tag{A.1.1}$$

Every one of the processes in (A.1.1) is a local $\mathbb{Q}$-martingale.

Now let's suppose we have a standard d-dimensional Brownian motion with respect to the measured filtration $(\Omega, \mathbb{F}, \mathbb{P})$ and let h be a d-dimensional locally bounded $\mathbb{F}$-predictable process. Note that $M := \sum_{i=1}^d h_i \bullet W_i$ is a locally bounded local martingale and so is its Doléans-Dade exponential

$$Z_t := \exp\left(M_t - \frac{1}{2}\int_0^t |h_s|^2 \, \mathrm{d}s\right) = 1 + \int_0^t Z_s \, \mathrm{d}M_s. \tag{A.1.2}$$

Here Z is a strictly positive supermartingale and is a martingale if and only if $\mathbb{E}[Z_t] = 1$ for all $t > 0$.

Proposition A.1.8. *Let W be a standard d-dimensional Brownian motion and let h be a d-dimensional locally bounded $\mathbb{F}$-predictable process. Suppose that Z is defined by (A.1.2) and define*

$$W^{\mathbb{Q}} := W - H, \qquad \text{where } H_{\cdot} := \int_0^{\cdot} h_s \, \mathrm{d}s = [M, W]_{\cdot} \, .$$

If Z is a uniformly integrable martingale, then the probability measure $\mathbb{Q}$ defined by $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} := Z_\infty$ is absolutely continuous with respect to $\mathbb{P}$ on $\mathcal{F}_\infty$, and $W^{\mathbb{Q}}$ is a standard $\mathbb{Q}$-Brownian motion.

In particular, if there is a Lebesgue square integrable function η on $[0, \infty)$ such that $|h_t(\omega)| \leq \eta_t$ for all t and all $\omega \in \Omega$, then Z is uniformly integrable and moreover $\mathbb{P}$ and $\mathbb{Q}$ are absolutely continuous on $\mathcal{F}_\infty$.

Proof. First note that since Z is a uniformly integrable martingale, there indeed exists a limit Z_∞ in L^1 and a.s.-convergence. Clearly, the probability measure $\mathbb{Q}$ is absolutely continuous with respect ot $\mathbb{P}$ and locally equivalent to $\mathbb{P}$ on $\mathcal{F}_\infty$. Note that

$$Z^{\mathbb{P}} \bullet [Z, W] = Z^{\mathbb{P}} \bullet [Z \bullet M, W] = Z^{\mathbb{P}} Z \bullet [M, W] = H,$$

so by Lemma A.1.7 it follows that $W^{\mathbb{Q}}$ is a vector of $\mathbb{Q}$-martingales. Since it has the same bracket as a Brownian Motion, it must be a $\mathbb{Q}$-Brownian motion by the Levy Characterisation.

The final statement of the theorem follows from Novikov's condition. $\square$

Example A.1.9. Note that Proposition A.1.8 is rather restrictive. For example it does not cover the simple shift $W_t^{\mathbb{Q}} = W_t + t$ as in this case $Z_t = \exp\left(W_t - \frac{t}{2}\right)$, which is not uniformly integrable.

It is the case that for each t, there exists a probability measure $\mathbb{Q}_t$ on $\mathcal{F}_t$ equivalent to $\mathbb{P}_t$. The pairs $(\mathcal{F}_t, \mathbb{Q}_t)$ are consistent in as much that for every $s < t$, $\mathbb{Q}_t$ restricted to $\mathcal{F}_s$ equals $\mathbb{Q}_s$. Therefore, there exists a unique pre-measure $\mathbb{Q}$ on $\mathcal{A}_\infty$, called the *projective limit*, such that for any $s < t$ we have

$$\mathbb{Q}(A) = \mathbb{Q}_t(A) = \mathbb{Q}_s(A) \qquad \text{for } A \in \mathcal{F}_s.$$

However, even if $\mathbb{Q}$ is σ-additive on $\mathcal{F}_\infty$, $\mathbb{Q}$ cannot be absolutely continuous with respect to $\mathbb{P}$. Namely, since $\lim_{t\to\infty} \frac{W_t}{t} = 0$ $\mathbb{P}$-a.s., the set

$$\left\{ \lim_{t\to\infty} \frac{W_t}{t} = -1 \right\}$$

is $\mathbb{P}$-negligible, despite the fact that $\lim_{t\to\infty} \frac{W_t^{\mathbb{Q}}}{t} = 0$ $\mathbb{Q}$-a.s.

We now introduce a definition in order to force the projective limit discussed in Example A.1.9 to be σ-additive on $\mathcal{F}_\infty$ by an application of Kolmogorov's extension theorem; see [10, Proposition 3.9.18] for further details.

Definition A.1.10. (i) The filtered space $(\Omega, \mathbb{F})$ is *full* if whenever $(\mathcal{F}_t, \mathbb{P}_t)$ is a consistent family of probabilities on $\mathcal{F}$, then there exists a σ-additive probability $\mathbb{P}$ on $\mathcal{F}_\infty$ whose restriction to $\mathcal{F}_t$ is $\mathbb{P}_t$ for each $t \geq 0$.

(ii) The filtered space $(\Omega, \mathbb{F}, \mathbb{P})$ is *full* if whenever $(\mathcal{F}_t, \tilde{\mathbb{P}}_t)$ is a consistent family of probabilities with $\tilde{\mathbb{P}}_t \ll_{\text{loc}} \mathbb{P}$ on $\mathcal{F}_t$, then there exists a σ-additive probability $\tilde{\mathbb{P}}$ on $\mathcal{F}_\infty$, whose restriction to $\mathcal{F}_t$ is $\tilde{\mathbb{P}}_t$, for each $t \geq 0$.

Proposition A.1.11. *Let X be a Polish space. Then $X^{[0,\infty)}$ equipped with its raw filtration is full. Moreover, the* càdlàg *and continuous path spaces, equipped with their raw filtrations, are full.*

The following proposition tells us that we can discard inconsequential nearly empty sets from Ω and proceed to the natural enlargement without obliterating the fullness property.

Proposition A.1.12. (a) *Suppose that $(\Omega, \mathbb{F})$ is full, and let $N := \bigcup_{n\in\mathbb{N}} A_n$ for some $\{A_n\} \subset \mathcal{A}_\infty$. Set $\Omega' = \Omega \backslash N$ and define*

$$\mathcal{F}'_t := \{A \cap \Omega' : A \in \mathcal{F}_t\}.$$

Then $(\Omega', \mathbb{F}')$ is full. Similarly, if the filtration $(\Omega, \mathbb{F}, \mathbb{P})$ is full and the $\mathbb{P}$-nearly empty set N is removed from Ω, then the filtered space induced on $\Omega' = \Omega \backslash N$ is full.

(b) *If the filtered space $(\Omega, \mathbb{F}, \mathbb{P})$ is full then so is its natural enlargement. In particular, the natural filtration on the canonical path space is full.*

We are now in a position to state Girsanov's Theorem in its general form.

Theorem A.1.13 (Girsanov's Theorem)**.** *Assume that W is a d-dimensional standard Brownian motion on the full filtered space $(\Omega, \mathbb{F}, \mathbb{P})$, and let h be a locally bounded predictable process. If the Doléans-Dade exponential Z of the local martingale $M := \sum_{n=1}^{d} h_i \bullet W_i$ is a martingale, then there is a unique σ-additive probability $\mathbb{Q}$ on $\mathcal{F}_\infty$ such that $Z_t = \frac{\mathrm{d}\mathbb{Q}_t}{\mathrm{d}\mathbb{P}_t}$ at all finite instants t, and*

$$W^{\mathbb{Q}} := W - [M, W] = W - \int_0^{\cdot} h_s \, \mathrm{d}s$$

is a standard Brownian motion under $\mathbb{Q}$.

Remark A.1.14. We reiterate that Example A.1.9 shows that $\mathbb{Q}$ need not be equivalent to $\mathbb{P}$ on $\mathcal{F}_\infty$.

Appendix B

B.1 Existence and uniqueness of linear FBSDE

In this section we provide results from [17, Appendix A] that assert the existence and uniqueness of a solution to the following FBSDE on the filtered space $(\Omega, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} = \{\mathcal{F}_t\}_{t\geq 0}$ is the naturally augmented raw filtration with respect to a $\mathbb{P}$-Brownian motion W:

$$\mathrm{d}\varphi_t = \dot{\varphi}_t \, \mathrm{d}t, \quad \varphi_0 = 0 \quad t \in \mathscr{T}, \tag{B.1.1}$$

$$\mathrm{d}\dot{\varphi}_t = Z_t \, \mathrm{d}W_t + B(\varphi_t - \xi_t) \, \mathrm{d}t + \delta\dot{\varphi}_t, \tag{B.1.2}$$

where $\mathscr{T}$ is either $[0, T]$ or $[0, \infty)$, $B \in \mathbb{R}^{\ell\times\ell}$ has only positive eigenvalues, $\delta \geq 0$, and $\xi \in \mathscr{L}^4_\delta(\mathbb{R}^\ell)$. If $\mathscr{T} = [0, T]$ for some $T < \infty$, then (B.1.1)-(B.1.2) is complemented by the terminal condition

$$\dot{\varphi}_T = 0. \tag{B.1.3}$$

If $\mathscr{T} = [0, \infty)$ we implicitly assume that $\delta > 0$ and the terminal condition is replaced by the transversality condition implicit in $\varphi, \dot{\varphi} \in \mathscr{L}^4_\delta(\mathbb{R}^\ell)$. A solution to (B.1.1)-(B.1.2) is a triple

$$(\varphi, \dot{\varphi}, Z) \in \mathscr{L}^4_\delta(\mathbb{R}^\ell) \times \mathscr{L}^4_\delta(\mathbb{R}^\ell) \times \mathscr{L}^4_\delta(\mathbb{R}^\ell).$$

The following theorem combines [17, Theorem A.2] and [17, Theorem

A.4]. For convenience, we set

$$C := \sqrt{B + \frac{\delta^2}{4} I_\ell} - \frac{\delta}{2} I_\ell \in \mathbb{R}^{\ell \times \ell}.$$

Theorem B.1.1. (a) *Suppose that $T < \infty$ and the matrix $B + \frac{\delta^2}{4} I_\ell$ has only positive eigenvalues. Then the FBSDE (B.1.1)-(B.1.2) with terminal condition (B.1.3) has a unique solution.*

(b) *Suppose that $T = \infty$, $\delta > 0$, and the matrix $B + \frac{\delta^2}{4} I_\ell$ has only positive eigenvalues. Then the unique solution to the FBSDE (B.1.1)-(B.1.2) is given by*

$$\varphi_t = \int_0^t e^{-C(t-s)} \bar{\xi}_s \, \mathrm{d}s \tag{B.1.4}$$

where

$$\bar{\xi}_t = C(C + \delta I_\ell) \mathbb{E} \left[\int_t^\infty e^{-(C+\delta I_\ell)(s-t)} \xi_s \, \mathrm{d}s \middle| \mathcal{F}_t \right]. \tag{B.1.5}$$

Moreover,

$$\dot{\varphi}_t = \bar{\xi}_t - C \varphi_t \tag{B.1.6}$$

and Z is derived from the martingale representation theorem with respect to the square integrable martingale M that has dynamics

$$\mathrm{d}M_t = e^{(C+\delta)t} \, \mathrm{d}\bar{M}_t,$$

where

$$\bar{M}_t := B \mathbb{E} \left[\int_0^t e^{-(C+\delta)s} \xi_s \, \mathrm{d}s \middle| \mathcal{F}_t \right].$$

The following is a technical condition required in Chapter 2, taken from [17, Proposition A.1]

Proposition B.1.2. *Let $T = \infty$. If $(\varphi, \dot{\varphi}, M) \in \mathscr{L}^4_\delta(\mathbb{R}^\ell) \times \mathscr{L}^4_\delta(\mathbb{R}^\ell) \times \mathscr{L}^4_\delta(\mathbb{R}^\ell)$ is a solution to the FBSDE (B.1.1)-(B.1.2), then $M \in \mathscr{M}^2_\delta$.*

Appendix C

In Chapters 3 and 4 we often need to compute fairly complex integrals and conditional expectations. Fortunately, many of these integrals and expectations are of the same form, which we list and simplify here.

C.1 Generalised Calculations

Proposition C.1.1. *Assume that* $A, B \in \mathbb{R}$ *such that* $A + B \neq 0$. *Then*

$$\begin{aligned}\int_0^t e^{-(A+B)s}\int_0^s e^{Br}\mathrm{d}W_r\mathrm{d}s &= \frac{A}{A+B}\int_0^t e^{-As}W_s\mathrm{d}s + \frac{B}{A+B}e^{-(A+B)t}\int_0^t e^{Bs}W_s\mathrm{d}s\\ &= \frac{1}{A+B}\left(\int_0^t e^{-As}\mathrm{d}W_s - e^{-At}\int_0^t e^{-B(t-s)}\mathrm{d}W_s\right).\end{aligned}$$

Proof. By repeated use of integration by parts and Fubini's theorem, it follows that

$$\begin{aligned}&\int_0^t e^{-(A+B)s}\int_0^s e^{Br}\mathrm{d}W_r\mathrm{d}s\\ &= \int_0^t e^{-(A+B)s}\left(e^{Bs}W_s - B\int_0^s e^{Br}W_r\mathrm{d}r\right)\mathrm{d}s\\ &= \int_0^t e^{-As}W_s\mathrm{d}s - B\int_0^t e^{-(A+B)s}\int_0^s e^{Br}W_r\mathrm{d}r\mathrm{d}s\\ &= \int_0^t e^{As}W_s\mathrm{d}s - B\int_0^t e^{Br}W_r\int_r^t e^{(A+B)s}\mathrm{d}s\mathrm{d}r\\ &= \int_0^t e^{-As}W_s\mathrm{d}s + \frac{B}{A+B}\int_0^t e^{Br}W_r\left(e^{-(A+B)t} - e^{-(A+B)r}\right)\mathrm{d}r\\ &= \frac{A}{A+B}\int_0^t e^{-As}W_s\mathrm{d}s + \frac{B}{A+B}e^{-(A+B)t}\int_0^t e^{Br}W_r\mathrm{d}r \qquad \text{(C.1.1)}\end{aligned}$$

$$= -\frac{1}{A+B}e^{-At}W_t - \int_0^t e^{-As}\mathrm{d}W_s + \frac{1}{A+B}e^{-(A+B)t}e^{Bt}W_t - \int_0^t e^{Br}\mathrm{d}W_r$$
$$= \frac{1}{A+B}\left(\int_0^t e^{-As}\mathrm{d}W_s - e^{-At}\int_0^t e^{-B(t-r)}\mathrm{d}W_r\right) \tag{C.1.2}$$

□

Proposition C.1.2. *Assume that $A, B \in (0,\infty)$. Then*

$$\mathbb{E}\left[\int_0^\infty e^{-(A+B)s}\int_0^s e^{Ar}\mathrm{d}W_r\ \mathrm{d}s\Big|\mathcal{F}_t\right] = \frac{1}{A+B}\int_0^t e^{-As}\ \mathrm{d}W_s$$
$$= \frac{1}{A+B}\left(e^{-At}W_t + A\int_0^t e^{-As}W_s\ \mathrm{d}s\right).$$

Proof. Using Proposition C.1.1 and the fact that $\int_0^t e^{-As}\ \mathrm{d}W_s$ is an L^2-bounded martingale, it follows that

$$\mathbb{E}\left[\int_0^\infty e^{-(A+B)s}\left(\int_0^s e^{Br}\mathrm{d}W_r\right)\mathrm{d}s\Big|\mathcal{F}_t\right] = \frac{1}{A+B}\mathbb{E}\left[\int_0^\infty e^{-As}\ \mathrm{d}W_s\Big|\mathcal{F}_t\right]$$
$$= \frac{1}{A+B}\int_0^t e^{-As}\ \mathrm{d}W_s.$$

□

Proposition C.1.3. *Assume that $A, B \in (0,\infty)$. Then*

$$\mathbb{E}\left[\int_t^\infty e^{-(A+B)s}\left(\int_0^s e^{Br}\mathrm{d}W_r\right)\mathrm{d}s\Big|\mathcal{F}_t\right] = \frac{e^{-At}}{A+B}\int_0^t e^{-B(t-s)}\mathrm{d}W_s.$$

Proof. Using Proposition C.1.1 and Proposition C.1.2 it follows that

$$\mathbb{E}\left[\int_t^\infty e^{-(A+B)s}\left(\int_0^s e^{Br}\mathrm{d}W_r\right)\mathrm{d}s\Big|\mathcal{F}_t\right]$$
$$= \frac{1}{A+B}\int_0^t e^{-As}\ \mathrm{d}W_s - \frac{1}{A+B}\left(\int_0^t e^{-As}\mathrm{d}W_s - e^{-At}\int_0^t e^{-B(t-s)}\mathrm{d}W_s\right)$$
$$= \frac{e^{-At}}{A+B}\int_0^t e^{-B(t-r)}\mathrm{d}W_t.$$

□

Proposition C.1.4. *Assume that $A, B \in \mathbb{R}$ such that $A + B \neq 0$. Then*

$$\begin{aligned}
&\int_0^t e^{-(A+B)s} \int_0^s e^{Br} W_r \, \mathrm{d}r \, \mathrm{d}s \\
&= -\frac{e^{-At}}{AB} W_t + \frac{1}{A+B} \left(\frac{e^{-At}}{B} \int_0^t e^{-B(t-s)} \, \mathrm{d}W_s + \frac{1}{A} \int_0^t e^{-As} \, \mathrm{d}W_s \right).
\end{aligned}$$

Proof. By repeated use of integration by parts and Fubini's theorem, it follows that

$$\begin{aligned}
&\int_0^t e^{-(A+B)s} \int_0^s e^{Br} W_r \, \mathrm{d}r \, \mathrm{d}s \\
&= -\frac{1}{A+B} \int_0^t e^{Br} W_r \left(e^{-(A+B)t} - e^{-(A+B)r} \right) \mathrm{d}r \\
&= -\frac{1}{A+B} \left(e^{-(A+B)t} \int_0^t e^{Br} W_r \mathrm{d}r - \int_0^t e^{-As} W_s \mathrm{d}s \right) \\
&= -\frac{1}{A+B} \left(\frac{e^{-At}}{AB} (A+B) W_t - \frac{e^{-At}}{B} \int_0^t e^{-B(t-s)} \, \mathrm{d}W_s - \frac{1}{A} \int_0^t e^{-As} \, \mathrm{d}W_s \right) \\
&= -\frac{e^{-At}}{AB} W_t + \frac{1}{A+B} \left(\frac{e^{-At}}{B} \int_0^t e^{-B(t-s)} \, \mathrm{d}W_s + \frac{1}{A} \int_0^t e^{-As} \, \mathrm{d}W_s \right).
\end{aligned}$$

□

Proposition C.1.5. *Assume that $A, B \in (0, \infty)$. Then*

$$\begin{aligned}
\mathbb{E}\left[\int_0^\infty e^{-(A+B)s} \int_0^s e^{Br} W_r \, \mathrm{d}r \, \mathrm{d}s \Big| \mathcal{F}_t \right] &= \frac{1}{A(A+B)} \int_0^t e^{-As} \, \mathrm{d}W_s \\
&= \frac{1}{A(A+B)} \left(e^{-At} W_t + A \int_0^t e^{-As} W_s \, \mathrm{d}s \right).
\end{aligned}$$

Proof. Using Proposition C.1.4 and the fact that $\int_0^t e^{-As} \, \mathrm{d}W_s$ is an L^2-bounded martingale, it follows that

$$\begin{aligned}
\mathbb{E}\left[\int_0^\infty e^{-(A+B)s} \left(\int_0^s e^{Br} W_r \, \mathrm{d}r \right) \mathrm{d}s \Big| \mathcal{F}_t \right] &= \frac{1}{A(A+B)} \mathbb{E}\left[\int_0^\infty e^{-As} \, \mathrm{d}W_s \Big| \mathcal{F}_t \right] \\
&= \frac{1}{A(A+B)} \int_0^t e^{-As} \, \mathrm{d}W_s.
\end{aligned}$$

$\square$

Proposition C.1.6. *Assume that $A, B \in (0,\infty)$. Then*

$$\mathbb{E}\left[\int_t^\infty e^{-(A+B)s}\left(\int_0^s e^{Br}W_r \, \mathrm{d}r\right)\mathrm{d}s\Big|\mathcal{F}_t\right]$$
$$= \frac{e^{-At}}{AB}W_t - \frac{e^{-At}}{B(A+B)}\int_0^t e^{-B(t-s)} \, \mathrm{d}W_s.$$

Proof. Using Proposition C.1.4 and Proposition C.1.5 it follows that

$$\begin{aligned}
&\mathbb{E}\left[\int_t^\infty e^{-(A+B)s}\left(\int_0^s e^{Br}W_r \, \mathrm{d}r\right)\mathrm{d}s\Big|\mathcal{F}_t\right] \\
= \; &\frac{1}{A(A+B)}\int_0^t e^{-As} \, \mathrm{d}W_s + \frac{e^{-At}}{AB}W_t \\
&- \frac{1}{A+B}\left(\frac{e^{-At}}{B}\int_0^t e^{-B(t-s)} \, \mathrm{d}W_s + \frac{1}{A}\int_0^t e^{-As} \, \mathrm{d}W_s\right) \\
&= \frac{e^{-At}}{AB}W_t - \frac{e^{-At}}{B(A+B)}\int_0^t e^{-B(t-s)} \, \mathrm{d}W_s.
\end{aligned}$$

$\square$

Proposition C.1.7. *Assume that $A \in \mathbb{R}\backslash\{0\}$. Then*

$$\begin{aligned}
\int_0^t\int_0^s e^{-A(t-s)} \, \mathrm{d}W_r \, \mathrm{d}s &= \int_0^t e^{-A(t-s)}W_s \, \mathrm{d}s - A\int_0^t e^{-A(t-s)}(t-s)W_s \, \mathrm{d}s \\
&= \int_0^t e^{-A(t-s)}(t-s) \, \mathrm{d}W_s.
\end{aligned}$$

Proof. By repeated use of Fubini's theorem and integration by parts it follows that

$$\begin{aligned}
&e^{-At}\int_0^t\int_0^s e^{Ar}\mathrm{d}W_r \, \mathrm{d}s \\
&= e^{-At}\int_0^t\left(e^{As}W_s - A\int_0^s e^{A_r}W_r \, \mathrm{d}r\right)\mathrm{d}s \\
&= e^{-At}\int_0^t e^{As}W_s \, \mathrm{d}s - Ae^{-At}\int_0^t\int_0^s e^{A_r}W_r \, \mathrm{d}r \, \mathrm{d}s \\
&= e^{-At}\int_0^t e^{As}W_s \, \mathrm{d}s - Ae^{-At}\int_0^t e^{A_r}W_r\left(\int_r^t \, \mathrm{d}s\right) \, \mathrm{d}r
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^t e^{-A(t-s)} W_s \, \mathrm{d}s - A \int_0^t e^{-A(t-s)} (t-s) W_s \, \mathrm{d}s \\
&= (1 - At)\, e^{-At} \int_0^t e^{As} W_s \, \mathrm{d}s + \frac{1}{A} e^{-At} \left((At-1) e^{At} W_t - \int_0^t (As-1) e^{As} \, \mathrm{d}W_s \right) \\
&= \frac{1}{A} (1 - At)\, e^{-At} \left(e^{At} W_t - \int_0^t e^{As} \mathrm{d}W_s \right) \\
&+ \frac{1}{A} e^{-At} \left((At-1) e^{At} W_t - \int_0^t (As-1) e^{As} \, \mathrm{d}W_s \right) \\
&= \frac{1}{A} (At-1)\, e^{-At} \int_0^t e^{As} \mathrm{d}W_s - \frac{1}{A} e^{-At} \int_0^t (As-1) e^{As} \, \mathrm{d}W_s \\
&= \int_0^t e^{-A(t-s)} (t-s) \, \mathrm{d}W_s.
\end{aligned}
$$

□

Proposition C.1.8. *Assume that $A \in \mathbb{R}\backslash\{0\}$. Then*

$$
\begin{aligned}
&\int_0^t \int_0^s e^{-A(s-r)} (s-r) \, \mathrm{d}W_r \, \mathrm{d}s \\
&= \int_0^t e^{-A(t-s)} (t-s) W_s \, \mathrm{d}s \\
&= \frac{1}{A^2} W_t - \frac{1}{A} \int_0^t e^{-A(t-s)} (t-s) \, \mathrm{d}W_s - \frac{1}{A^2} \int_0^t e^{-A(t-s)} \, \mathrm{d}W_s.
\end{aligned}
$$

Proof. From Proposition C.1.7 we see that

$$
\begin{aligned}
&\int_0^t \int_0^s e^{-A(s-r)} (s-r) \mathrm{d}W_r \, \mathrm{d}s \\
&= \int_0^t \left(\int_0^s e^{-A(s-r)} W_r \, \mathrm{d}r - A \int_0^s e^{-A(s-r)} (s-r) W_s \, \mathrm{d}r \right) \mathrm{d}s \\
&= \int_0^t \left(e^{-As} \int_0^s e^{Ar} W_r \, \mathrm{d}r - A e^{-As} s \int_0^s e^{A_r} W_r \, \mathrm{d}r + A e^{-As} \int_0^s r e^{A_r} W_r \, \mathrm{d}r \right) \mathrm{d}s.
\end{aligned}
\tag{C.1.3}
$$

We look at each of the three double integrals in turn.

$$
\begin{aligned}
\int_0^t e^{-As} \int_0^t e^{Ar} W_r \, \mathrm{d}s &= \int_0^t e^{Ar} W_r \int_r^t e^{-As} \, \mathrm{d}s \, \mathrm{d}r \\
&= \frac{1}{A} \int_0^t e^{Ar} W_r \left(e^{-Ar} - e^{-At} \right) \mathrm{d}r
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{A}\int_0^t W_r \, \mathrm{d}r - \frac{1}{A}e^{-At}\int_0^t e^{Ar}W_r \, \mathrm{d}r, && \text{(C.1.4)}\\
\int_0^t Ae^{-As}s\int_0^s e^{A_r}W_r \, \mathrm{d}r \, \mathrm{d}s &= A\int_0^t e^{A_r}W_r\int_r^t e^{-As}s \, \mathrm{d}s \, \mathrm{d}t\\
&= \frac{1}{A}\int_0^t e^{Ar}W_r\left((Ar+1)e^{-Ar}-(At+1)e^{-At}\right)\mathrm{d}r\\
&= \int_0^t rW_r \, \mathrm{d}r + \frac{1}{A}\int_0^t W_r \mathrm{d}r\\
&- e^{-At}\left(t\int_0^t e^{Ar}W_r \, \mathrm{d}r + \frac{1}{A}\int_0^t e^{Ar}W_r \, \mathrm{d}r\right), && \text{(C.1.5)}\\
\int_0^t Ae^{-As}\int_0^s e^{Ar}rW_r \, \mathrm{d}r \, \mathrm{d}s &= A\int_0^t e^{Ar}rW_r\int_r^t e^{-As} \, \mathrm{d}s \, \mathrm{d}r\\
&= \int_0^t e^{Ar}rW_r\left(e^{-Ar}-e^{-At}\right)\mathrm{d}r\\
&= \int_0^t rW_r\mathrm{d}r - e^{-At}\int_0^t e^{Ar}rW_r \, \mathrm{d}r. && \text{(C.1.6)}
\end{aligned}
$$

Thus, substituting (C.1.4), (C.1.5) and (C.1.6) into (C.1.3) we see that

$$
\begin{aligned}
&\int_0^t\int_0^s e^{-A(s-r)}(s-r) \, \mathrm{d}W_r \, \mathrm{d}s\\
&= \int_0^t e^{-A(t-s)}(t-s)W_s \, \mathrm{d}s\\
&= \frac{1}{A}e^{-At}t\left(e^{At}W_t - \int_0^t e^{As} \, \mathrm{d}W_s\right)\\
&- \frac{1}{A^2}e^{-At}\left((At-1)e^{At}W_t - \int_0^t (As-1)e^{As} \, \mathrm{d}W_s\right)\\
&= \frac{1}{A^2}W_t - \frac{1}{A}\int_0^t e^{-A(t-s)}(t-s) \, \mathrm{d}W_s - \frac{1}{A^2}\int_0^t e^{-A(t-s)} \, \mathrm{d}W_s.
\end{aligned}
$$

□

Proposition C.1.9. *Assume that $A, B \in \mathbb{R}\backslash\{0\}$ such that $A+B \neq 0$. Then*

$$
\begin{aligned}
&\int_0^t e^{-As}\int_0^s\int_0^r (r-q)e^{-B(r-q)}\mathrm{d}W_q \, \mathrm{d}r \, \mathrm{d}s\\
&= \int_0^t e^{-(A+B)s}\int_0^s e^{Br}(s-r)W_r \, \mathrm{d}r \, \mathrm{d}s
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{(A+B)^2}\int_0^t e^{-Ar}W_r \,\mathrm{d}r - \frac{e^{-At}}{(A+B)}\int_0^t e^{-B(t-r)}(t-r)W_r \,\mathrm{d}r \\
&- \frac{e^{-At}}{(A+B)^2}\int_0^t e^{-B(t-r)}W_r \,\mathrm{d}r \\
&= -\frac{1}{AB^2}e^{-At}W_t + \frac{1}{B(A+B)}e^{-At}\int_0^t e^{-B(t-s)}(t-s)\,\mathrm{d}W_s \\
&+ \frac{A+2B}{B^2(A+B)^2}e^{-At}\int_0^t e^{-B(t-s)}\,\mathrm{d}W_s + \frac{1}{A(A+B)^2}\int_0^t e^{-As}\mathrm{d}W_s.
\end{aligned}
$$

Proof. Using Proposition C.1.8 we see that

$$
\begin{aligned}
&\int_0^t e^{-As}\int_0^s\int_0^r (r-q)e^{-B(r-q)}\mathrm{d}W_q \,\mathrm{d}r\,\mathrm{d}s \\
&= \int_0^t e^{-(A+B)s}\int_0^s e^{Br}(s-r)W_r \,\mathrm{d}r\,\mathrm{d}s
\end{aligned} \tag{C.1.7}
$$

We look at each of the double integrals in turn, simplifying them utilising Fubini's theorem and integration by parts repeatedly.

$$
\begin{aligned}
&\int_0^t e^{-(A+B)s}s\int_0^s e^{Br}W_r \,\mathrm{d}r\,\mathrm{d}s \\
&= \frac{1}{(A+B)^2}\int_0^t e^{Br}W_r\left(((A+B)r+1)e^{-(A+B)r} - ((A+B)t+1)e^{-(A+B)t}\right)\mathrm{d}r \\
&= \frac{1}{(A+B)}\int_0^t e^{-Ar}rW_r \,\mathrm{d}r + \frac{1}{(A+B)^2}\int_0^t e^{-Ar}W_r \,\mathrm{d}r \\
&- \frac{e^{-At}t}{(A+B)}\int_0^t e^{-B(t-r)}W_r \,\mathrm{d}r - \frac{e^{-At}}{(A+B)^2}\int_0^t e^{-B(t-r)}W_r \,\mathrm{d}r,
\end{aligned} \tag{C.1.8}
$$

$$
\begin{aligned}
&\int_0^t e^{-(A+B)s}\int_0^s e^{Br}rW_r \,\mathrm{d}r\,\mathrm{d}s \\
&= \frac{1}{(A+B)}\int_0^t e^{Br}rW_r\left(e^{-(A+B)r} - e^{-(A+B)t}\right)\mathrm{d}r \\
&= \frac{1}{(A+B)}\int_0^t e^{-Ar}rW_r \,\mathrm{d}r - \frac{e^{-At}}{(A+B)}\int_0^t e^{-B(t-r)}rW_r \,\mathrm{d}r.
\end{aligned} \tag{C.1.9}
$$

Substituting (C.1.9) and (C.1.8) into (C.1.7) we see that

$$
\int_0^t e^{-As}\int_0^s\int_0^r (r-q)e^{-B(r-q)}\mathrm{d}W_q \,\mathrm{d}r\,\mathrm{d}s
$$

$$
\begin{aligned}
&= \frac{1}{(A+B)^2}\int_0^t e^{-Ar}W_r \,\mathrm{d}r - \frac{e^{-At}}{(A+B)}\int_0^t e^{-B(t-r)}(t-r)W_r \,\mathrm{d}r \\
&- \frac{e^{-At}}{(A+B)^2}\int_0^t e^{-B(t-r)}W_r \,\mathrm{d}r \\
&= -\frac{1}{A(A+B)^2}\left(e^{-At}W_t - \int_0^t e^{-Ar}\mathrm{d}W_r\right) \\
&- \frac{e^{-At}}{B(A+B)}\left(\frac{1}{B}W_t - \int_0^t e^{-B(t-r)}(t-r)\mathrm{d}W_r - \frac{1}{B}\int_0^t e^{-B(t-r)}\,\mathrm{d}W_r\right) \\
&- \frac{e^{-(A+B)t}}{B(A+B)^2}\left(e^{Bt}W_t - \int_0^t e^{Br}\,\mathrm{d}W_r\right) \\
&= -\frac{1}{AB^2}e^{-At}W_t + \frac{1}{B(A+B)}e^{-At}\int_0^t e^{-B(t-s)}(t-s)\,\mathrm{d}W_s \\
&+ \frac{A+2B}{B^2(A+B)^2}e^{-At}\int_0^t e^{-B(t-s)}\,\mathrm{d}W_s + \frac{1}{A(A+B)^2}\int_0^t e^{-As}\mathrm{d}W_s.
\end{aligned}
$$

□

Proposition C.1.10. *Assume that $A, B \in (0,\infty)$. Then*

$$
\begin{aligned}
&\int_0^t e^{-As}\int_0^s e^{-B(s-r)}(s-r)\,\mathrm{d}W_r\,\mathrm{d}s \\
&= \frac{1}{(A+B)^2}\int_0^t e^{-As}\,\mathrm{d}W_s - \frac{1}{(A+B)}e^{-At}\int_0^t e^{-B(t-s)}(t-s)\,\mathrm{d}W_s \\
&- \frac{1}{(A+B)^2}e^{-At}\int_0^t e^{-B(t-s)}\,\mathrm{d}W_s.
\end{aligned}
$$

Proof. By Proposition C.1.7, Proposition C.1.4 and Proposition C.1.9

$$
\begin{aligned}
&\int_0^t e^{-As}\int_0^s e^{-B(s-r)}(s-r)\,\mathrm{d}W_r\,\mathrm{d}s \\
&= \int_0^t e^{-As}\left(\int_0^s e^{-B(s-r)}W_r\,\mathrm{d}r - B\int_0^t e^{-B(s-r)}(s-r)W_r\,\mathrm{d}r\right)\,\mathrm{d}s \\
&= \left\{-\frac{e^{-At}}{AB}W_t + \frac{1}{A+B}\left(\frac{e^{-At}}{B}\int_0^t e^{-B(t-s)}\,\mathrm{d}W_s + \frac{1}{A}\int_0^t e^{-As}\,\mathrm{d}W_s\right)\right\} \\
&+ \left\{\frac{1}{AB}e^{-At}W_t - \frac{1}{(A+B)}e^{-At}\int_0^t e^{-B(t-s)}(t-s)\,\mathrm{d}W_s\right. \\
&\left.- \frac{A+2B}{B(A+B)^2}e^{-At}\int_0^t e^{-B(t-s)}\,\mathrm{d}W_s - \frac{B}{A(A+B)^2}\int_0^t e^{-As}\mathrm{d}W_s\right\}
\end{aligned}
$$

$$= \frac{1}{(A+B)^2}\int_0^t e^{-As}\,\mathrm{d}W_s - \frac{1}{(A+B)}e^{-At}\int_0^t e^{-B(t-s)}(t-s)\,\mathrm{d}W_s$$
$$- \frac{1}{(A+B)^2}e^{-At}\int_0^t e^{-B(t-s)}\,\mathrm{d}W_s.$$

□

Proposition C.1.11. *Assume that $A, B \in (0,\infty)$. Then*

$$\mathbb{E}\left[\int_0^\infty e^{-As}\int_0^s e^{-B(s-r)}(s-r)\,\mathrm{d}W_r\,\mathrm{d}s\,\middle|\mathcal{F}_t\right] = \frac{1}{(A+B)^2}\int_0^t e^{-As}\,\mathrm{d}W_s.$$

Proof. Using Proposition (C.1.10) and the fact that $\int_0^\cdot e^{-As}\,\mathrm{d}W_s$ is an L^2-bounded martingale, it follows that

$$\mathbb{E}\left[\int_0^\infty e^{-As}\int_0^s e^{-B(s-r)}(s-r)\,\mathrm{d}W_r\,\mathrm{d}s\,\middle|\mathcal{F}_t\right]$$
$$= \frac{1}{(A+B)^2}\mathbb{E}\left[\int_0^\infty e^{-As}\,\mathrm{d}W_s\,\middle|\mathcal{F}_t\right]$$
$$= \frac{1}{(A+B)^2}\int_0^t e^{-As}\,\mathrm{d}W_s.$$

□

Proposition C.1.12. *Assume that $A, B \in (0,\infty)$. Then*

$$\mathbb{E}\left[\int_t^\infty e^{-As}\int_0^s e^{-B(s-r)}(s-r)\,\mathrm{d}W_r\,\mathrm{d}s\,\middle|\mathcal{F}_t\right]$$
$$= \frac{e^{-At}}{(A+B)}\left\{\int_0^t e^{-B(t-s)}(t-s)\,\mathrm{d}W_s + \frac{1}{(A+B)}\int_0^t e^{-B(t-s)}\,\mathrm{d}W_s\right\}.$$

Proof. Using Proposition (C.1.10) and Proposition C.1.11, it follows that

$$\mathbb{E}\left[\int_0^\infty e^{-As}\int_0^s e^{-B(s-r)}(s-r)\,\mathrm{d}W_r\,\mathrm{d}s\,\middle|\mathcal{F}_t\right]$$
$$= \frac{e^{-At}}{(A+B)}\left\{\int_0^t e^{-B(t-s)}(t-s)\,\mathrm{d}W_s + \frac{1}{(A+B)}\int_0^t e^{-B(t-s)}\,\mathrm{d}W_s\right\}.$$

□

Proposition C.1.13. *Assume that $A, B \in (0,\infty)$. Then*

$$\mathbb{E}\left[\int_0^\infty e^{-As}\int_0^s\int_0^r e^{-B(r-q)}(r-q)\ \mathrm{d}W_q\ \mathrm{d}r\ \mathrm{d}s\ \Big|\mathcal{F}_t\right] = \frac{1}{A(A+B)^2}\int_0^t e^{-Ar}\ \mathrm{d}W_r.$$

Proof. Using Proposition C.1.9 and the fact that $\int_0^\cdot e^{-As}\mathrm{d}W_s$ is an L^2-bounded martingale, it follows that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^\infty e^{-As}\int_0^s\int_0^r e^{-B(r-q)}(r-q)\ \mathrm{d}W_q\ \mathrm{d}r\ \mathrm{d}s\ \Big|\mathcal{F}_t\right]\\
&= \frac{1}{A(A+B)^2}\mathbb{E}\left[\int_0^\infty e^{-As}\mathrm{d}W_s|\mathcal{F}_t\right] \qquad \text{(C.1.10)}\\
&= \frac{1}{A(A+B)^2}\int_0^t e^{-As}\mathrm{d}W_s. \qquad \text{(C.1.11)}
\end{aligned}$$

□

Proposition C.1.14. *Assume that $A, B \in (0,\infty)$. Then*

$$\begin{aligned}
&\mathbb{E}\left[\int_t^\infty e^{-As}\int_0^s\int_0^r (r-q)e^{-B(r-q)}\mathrm{d}W_q\ \mathrm{d}r\ \mathrm{d}s\ \Big|\mathcal{F}_t\right]\\
&= \frac{e^{-At}}{(A+B)^2}\left((A+B)\int_0^t e^{-B(t-r)}(t-r)W_r\ \mathrm{d}r + \int_0^t e^{-B(t-r)}W_r\ \mathrm{d}r - \frac{W_t}{A}\right)
\end{aligned}$$

Proof. Using Propositon C.1.9 and Proposition C.1.13 it follows that

$$\begin{aligned}
&\mathbb{E}\left[\int_t^\infty e^{-As}\int_0^s\int_0^r (r-q)e^{-B(r-q)}\mathrm{d}W_q\ \mathrm{d}r\ \mathrm{d}s\ \Big|\mathcal{F}_t\right]\\
&= \frac{1}{A(A+B)^2}\int_0^t e^{-As}\mathrm{d}W_s + \frac{1}{(A+B)^2}\int_0^t e^{-Ar}W_r\ \mathrm{d}r\\
&+ \frac{e^{-At}}{(A+B)}\int_0^t e^{-B(t-r)}(t-r)W_r\ \mathrm{d}r + \frac{e^{-At}}{(A+B)^2}\int_0^t e^{-B(t-r)}W_r\ \mathrm{d}r\\
&= \frac{e^{-At}}{(A+B)^2}\left((A+B)\int_0^t e^{-B(t-r)}(t-r)W_r\ \mathrm{d}r + \int_0^t e^{-B(t-r)}W_r\ \mathrm{d}r - \frac{W_t}{A}\right)\\
&= \frac{e^{-At}}{B}\Big(\frac{1}{AB}W_t - \frac{A+2B}{B(A+B)^2}\int_0^t e^{-B(t-s)}\ \mathrm{d}W_s\\
&\qquad - \frac{1}{(A+B)}\int_0^t e^{-B(t-s)}(t-s)\ \mathrm{d}W_s\Big).
\end{aligned}$$

□

Appendix D

D.1 Symbolic Algebra

We compute the main results in Chapters 3 and 4 with a computer algebra package. Originally we used the package *SymPy* in *Python* but later switched to *Sage* due to performance-related issues [72, 69, 55]. All of the code used in these chapters can be found in the GitHub repository

https://github.com/odshelley/thesis.

The calculations for which we use a symbolic algebra package are all linear combinations of expectations of the form

$$\mathbb{E}\left[\int_0^\infty e^{-\delta t} X_t^1 X_t^2 \, \mathrm{d}t\right],$$

where both X^1 and X^2 are either an ABM, an OU process, or something similar. Hence, we write functions to calculate such expectations for the combinations. For instance, suppose that X^1 is an ABM, and X^2 is an OU process, such that

$$X_t^1 = x_0^1 + \alpha t + \beta W_t$$

and

$$X_t^2 = x_0^2 e^{-Ct} + \mu\left(1 - e^{-Ct}\right) + \sigma \int_0^t e^{-C(t-s)},$$

for some $x_0^1, x_0^2, \alpha, \beta, \mu, \sigma \in \mathbb{R}$ and $C \in \mathbb{R}_{++}$. Then a straightforward (albeit tedious) calculation shows that

$$\mathbb{E}\left[\int_0^\infty e^{-\delta t} X_t^1 X_t^2 \,\mathrm{d}t\right]$$
$$= \frac{\delta^3 x_0^1 x_0^2 + \delta^2 C x_0^1 x_0^2 + \delta^2 C x_0^1 \mu + \delta C^2 x_0^1 \mu + \delta^2 \alpha x_0^2 + 2\,\delta C \alpha \mu + C^2 \alpha \mu + \delta^2 \beta \sigma + \delta C \beta \sigma}{(\delta + C)^2 \delta^2}.$$

The above formula is the output of the function *abm_times_ou* evaluated at $(x_0^1, \alpha, \beta, x_0^2, \mu, \sigma, C)$ found in the notebook `homogeneous`:

```
def abm_times_ou( Z11, Z12, Z13, Z21, Z22, Z23 , c):
    integral = definite_integral( exp( c * s ) * s, s , 0 , t )

    term1 = Z11*Z21*exp( - c * t )
    term2 = Z11*Z22*( 1 - exp( - c * t ) )
    term3 = Z12*Z21*exp( -c * t ) * t
    term4 = Z12*Z22*( 1 - exp( -c * t ) ) * t
    term5 = Z13*Z23*( t - c * exp( - c * t ) * integral )

    summand = term1 + term2 + term3 + term4 + term5
    return definite_integral(exp(- delta * t ) * ( summand ), t, 0,infinity)
```

Example D.1.1. The following code calculates the return loss seen in Theorem 3.2.6 before approximating for small transaction cost λ.

```
bph_t = ( gamma2 * alpha2 - gamma1 * alpha1 ) / ( ( gamma1 + gamma2 ) )
bph_W = ( gamma2 * beta2 - gamma1 * beta1 ) / ( ( gamma1 + gamma2 ) )

bmu_t = ( gamma1 * gamma2 * sigma**2 / ( gamma1 + gamma2 ) ) * ( alpha1 + alpha2 )
bmu_W = ( gamma1 * gamma2 * sigma**2 / ( gamma1 + gamma2 ) ) * ( beta1 + beta2 )

dlta_mu = bph_t * delta / ( C * ( C + delta ) )
dlta_W = - bph_W

terms = []

terms.append(          abm_times_abm( 0 , theta1 , theta2 , 0 , theta5 , theta6 ) )
terms.append( theta7 * abm_times_ou(  0 , theta1 , theta2 , 0 , theta3 , theta3 , C ) )
terms.append(          abm_times_ou(  0 , theta5 , theta6 , 0 , theta3 , theta4 , C ) )
terms.append( theta7 * ou_times_ou(   0 , theta3 , theta4 , 0 , theta3 , theta4 , C , C ) )

utility_return = sum(terms)

utility_return = utility_return.subs(
    theta1 == bph_t
).subs(
    theta2 == bph_W
).subs(
    theta3 == dlta_mu
).subs(
    theta4 == dlta_W
).subs(
    theta5 == bmu_t
).subs(
    theta6 == bmu_W
).subs(
    theta7 == ( k * gamma1 - gamma2 ) * sigma**2 / (k+1)
).factor()
```

Appendix E

E.1 Auxiliary results from Functional Analysis and Measure Theory

In this appendix, we collect some key results from Functional Analysis and Measure Theory that are referenced in Chapter 5.

E.1.1 Urysohn's Lemma

For the convenience of the reader we recall Urysohn's lemma as stated in Aliprantis [1, Lemma 2.46].

Lemma E.1.1. *Let Ω be a topological space. The following statements are equivalent.*

1. *The space Ω is normal.*

2. *Every pair of nonempty disjoint closed subsets of Ω can be separated by a continuous function.*

3. *If C is a closed subset of Ω and $f : C \to [0,1]$, then there is a continuous extension $\hat{f} : C \to [0,1]$ of f satisfying*

$$\sup_{x\in\Omega} \hat{f}(x) = \sup_{x\in C} f(x).$$

E.1.2 Locally compact spaces

Here we reference some results and concepts relating to locally compact Hausdorff spaces.

Definition E.1.2. Let Ω be a non-compact locally compact Hausdorff space with topology τ. Set $\Omega_\infty := \Omega \cup \{\infty\}$, where $\infty \notin \Omega$, and let

$$\tau_\infty := \tau \cup \{\Omega_\infty \backslash K : K \subset \Omega \text{ is compact}\}$$

Then Ω_∞ (with the topology τ_∞) is called the *one-point* compactification of Ω.

The one-point compactification of a non-compact locally compact Hausdorff space has nice properties; see [33, Proposition 4.36] for a proof.

Theorem E.1.3. *Let Ω be a non-compact locally compact Hausdorff space. Then Ω_∞ is a compact Hausdorff space and Ω is an open dense subset of Ω_∞. Moreover, $f \in C(\Omega)$ extends continuously to $f_\infty \in C(\Omega_\infty)$ if and only if $f = f_0 + c$ where $f_0 \in C_0(\Omega)$ and c is a constant. In this case, the extension satisfies $f_\infty(\infty) = c$.*

Next, we recall the famous Stone-Weierstraß Theorem. To this end, recall that a subset $\mathcal{C} \subset C_0(\Omega)$ *vanishes nowhere* if for all $x \in \Omega$, there exists some $f \in \mathcal{C}$ such that $f(x) \neq 0$, and it *separates points* if for each $x, y \in \Omega$ with $x \neq y$, there exists $f \in \mathcal{C}$ such that $f(x) \neq f(y)$.

Theorem E.1.4 (Stone-Weierstraß Theorem)**.** *Let Ω be a locally compact Hausdorff space and $\mathcal{C}$ be a subalgebra of $C_0(\Omega)$. Then $\mathcal{C}$ is dense in $C_0(\Omega)$ (for the topology of uniform convergence) if and only if it separates points and vanishes nowhere.*

Proof. See [26]. □

E.1.3 Dual spaces

We now consider the topological notions of weak and weak* convergence. This allows us to compare the topological notions of weak convergence of measures with the notions of weak and vague convergence as stated in Definition 5.1.10. Along the way we will be able to state the Banach Alaoglu theorem.

Definition E.1.5. Let Ω and Ω' both be vector spaces.

(a) We call the space of all linear forms on Ω its *algebraic dual* and denote it by Ω^*. For a topology τ on Ω, we call the space of all τ-continuous linear forms on Ω the *topological dual* of Ω with respect to τ and denote it by $(\Omega, \tau)'$. When the topology is understood it will be omitted from the notation.

(b) We say that the pair $\langle \Omega, \Omega' \rangle$ is a *dual pair* if it has an associated bilinear form $\langle \cdot, \cdot \rangle : \Omega \times \Omega' \to \mathbb{R}$ that separates points of Ω and Ω', i.e.

- if $\langle \omega, \omega' \rangle = 0$ for all $\omega' \in \Omega'$, then $\omega = 0$;
- if $\langle \omega, \omega' \rangle = 0$ for all $\omega \in \Omega$, then $\omega' = 0$.

We call $\langle \cdot, \cdot \rangle$ the *duality* of the dual pair.

(c) If we associate a norm $\|\cdot\|$ to Ω, then we say Ω' is its *norm dual* if it is the space $\mathrm{L}(\Omega, \mathbb{R})$ of all continuous linear forms from Ω to $\mathbb{R}$. In which case, the operator norm on Ω' is also called the *dual norm*. In particular, $\langle \Omega, \Omega' \rangle$ form a dual pair.

Suppose we have a dual pair $\langle \Omega, \Omega' \rangle$. We define the *weak topology* on Ω to be the topology $\sigma(\Omega, \Omega')$ generated by the family of seminorms on $\{p_{\omega'} : \omega' \in \Omega'\}$ where

$$p_{\omega'}(\omega) := |\langle \omega, \omega' \rangle| \quad \forall \omega \in \Omega.$$

Thus, $(\Omega, \sigma(\Omega, \Omega'))$ is a locally convex topological vector space. Note that $\omega_\alpha \xrightarrow{w} \omega$ if and only if $\langle \omega_\alpha, \omega' \rangle \to \langle \omega, \omega' \rangle$ in $\mathbb{R}$ for all $\omega' \in \Omega'$. Similarly, the *weak* topology* $\sigma(\Omega', \Omega)$ on Ω' is generated by $\{p_\omega : \omega \in \Omega\}$, and $\omega'_\alpha \xrightarrow{w^*} \omega'$ if and only if $\langle \omega, \omega'_\alpha \rangle \to \langle \omega, \omega' \rangle$ in $\mathbb{R}$ for all $\omega \in \Omega$.

It is key that the spaces in a dual pair are each others continuous topological duals [1, Theorem 5.93].

Theorem E.1.6. *Let $\langle \Omega, \Omega' \rangle$ be a dual pair. Then the topological dual of the topological vector space $(\Omega, \sigma(\Omega, \Omega'))$ is Ω'. That is, for any $\sigma(\Omega, \Omega')$-continuous linear form F, there exists a unique $\omega' \in \Omega'$ such that $F(\omega) = \langle \omega, \omega' \rangle$. Similarly we have $(\Omega', \sigma(\Omega', \Omega))' = \Omega$.*

Due to Theorem E.1.6 and the fact that for an infinite dimensional space Ω there are numerous subspaces Ω' of the algebraic dual Ω^* that separate points, we have some elbow room in our choice of a dual. The following definition narrows our focus.

Definition E.1.7. Let $\langle \Omega, \Omega' \rangle$ be a dual pair. A locally convex topology τ on a Ω is called *consistent* with $\langle \Omega, \Omega' \rangle$ if $(\Omega, \tau)' = \Omega'$. Consistent topologies on Ω' are defined similarly.

Remark E.1.8. The remarkable Mackey-Arens Theorem describes exactly which locally convex topologies on Ω are consistent with $\langle \Omega, \Omega' \rangle$ [1, Theorem 5.112].

We can now state the Banach-Alaoglu theorem [1, Theorem 5.105].

Theorem E.1.9. *Let $\langle \Omega, \Omega' \rangle$ be a dual pair and V any neighbourhood of zero with respect to a locally convex topology τ on Ω that is consistent with $\langle \Omega, \Omega' \rangle$. Then*

$$V^\circ := \{\omega' \in \Omega' : |\langle \omega, \omega' \rangle| \le 1 \ \forall \omega \in V\} \tag{E.1.1}$$

is weak compact. In particular, the closed unit ball of the norm dual of a normed space is weak* compact.*

E.1.4 Comparing weak notions of convergence

As in Section 5.1.2 we now assume that Ω is a metrisable topological space, to which we associate a metric d.

Definition E.1.10. Let $\mathcal{F} \subset \mathcal{M}(\Omega)$. We say that a family $\mathcal{C}$ of measurable maps $\Omega \to \mathbb{R}$ is a separating family for $\mathcal{F}$ if, for any two measures $\mu, \nu \in \mathcal{F}$,

$$\left(\int f \, \mathrm{d}\mu = \int f \, \mathrm{d}\nu \quad \forall f \in \mathcal{C} \cap L^1(\mu) \cap L^1(\nu) \right) \Rightarrow \mu = \nu.$$

Useful separating classes for $\mathcal{M}(\Omega)$ are introduced in the next theorem. For any $K \subset \Omega$ and $\varepsilon > 0$, define $\rho_{K,\varepsilon} : \Omega \to \mathbb{R}$ to be

$$\rho_{K,\varepsilon} := 1 - \left[\varepsilon^{-1} d(x, K) \wedge 1\right].$$

Note that $\rho_{K,\varepsilon}$ is a Lipschitz continuous function such that

$$\rho_{K,\varepsilon} = \begin{cases} 1, & x \in K \\ 0, & d(x,K) \geq \varepsilon \end{cases}$$

Theorem E.1.11.

(a) $\mathrm{Lip}_1(\Omega, [0,1])$ *is separating for* $\mathcal{M}(\Omega)$.

(b) *If* Ω *is additionally locally compact, then* $C_c(\Omega) \cap \mathrm{Lip}_1(\Omega, [0,1])$ *is a separating class for* $\mathcal{M}(\Omega)$.

Proof. Assume that $\mu_1, \mu_2 \in \mathcal{M}(\Omega)$ such that $\int f \, \mathrm{d}\mu_1 = \int f \, \mathrm{d}\mu_2$ for all $f \in \mathrm{Lip}_1(\Omega; [0,1]) \cap L^1(\mu_1) \cap L^1(\mu_2)$. Since both μ_1 and μ_2 are tight, we need only show that $\mu_1(K) = \mu_2(K)$ for K compact. Indeed, suppose this is the case, and let $A \in \mathscr{B}(\Omega)$. Then by tightness, for any $\varepsilon > 0$, there exists a compact set K_ε such that $|\mu_i|(A \backslash K_\varepsilon) < \varepsilon$ for $i = 1, 2$, whence

$$\begin{aligned} |\mu_1(A) - \mu_2(A)| &= |(\mu_1(K_\varepsilon) - \mu_2(K_\varepsilon)) + \mu_1(A\backslash K_\varepsilon) - \mu_2(A \backslash K_\varepsilon)| \\ &\leq |\mu_1|(A\backslash K_\varepsilon) + |\mu_2|(A\backslash K_\varepsilon) \\ &< \varepsilon. \end{aligned}$$

Take $K \in \mathscr{B}(\Omega)$ compact. Since μ_1, μ_2 are Radon measures, for each $x \in K$ there exists some open set U_x such that $x \in U_x$, $|\mu_1|(U_x) < \infty$ and $|\mu_2|(U_x) < \infty$. Since K is compact we can find finitely many points $x_1, \ldots, x_n$ such that $K \subset \bigcup_{k=1}^n U_{x_k} =: U$. By construction we have $|\mu_i|(U) < \infty$, whence $\mathbb{1}_U \in L^1(\mu_i)$ for $i = 1, 2$, respectively. Since U^c is closed and $K \cap U^c = \emptyset$, we have $\delta := d(U^c, K) > 0$. By definition, $\mathbb{1}_K \leq \rho_{K,\varepsilon} \leq \mathbb{1}_U \in L^1(\mu_i)$ if $\varepsilon \in (0, \delta)$. Note that $\rho_{K,\varepsilon} \to \mathbb{1}_K$, so by the dominated convergence theorem

$$\begin{aligned} \mu_i(K) = \int_\Omega \mathbb{1}_k \, \mathrm{d}\mu_i &= \int_\Omega \mathbb{1}_k \, \mathrm{d}\mu_i^+ - \int_\Omega \mathbb{1}_k \, \mathrm{d}\mu_i^- \\ &= \lim_{\varepsilon \downarrow 0} \left(\int_\Omega \rho_{K,\varepsilon} \, \mathrm{d}\mu_i^+ - \int_\Omega \rho_{K,\varepsilon} \, \mathrm{d}\mu_i^- \right) \\ &= \lim_{\varepsilon \downarrow 0} \int_\Omega \rho_{K,\varepsilon} \, \mathrm{d}\mu_i. \end{aligned}$$

However, $\varepsilon\rho_{K,\varepsilon} \in \mathrm{Lip}_1(\Omega; [0,1])$ for all $\varepsilon > 0$, hence by assumption

$$\int \rho_{K,\varepsilon}\, \mathrm{d}\mu_1 = \varepsilon^{-1} \int \varepsilon\rho_{K,\varepsilon}\, \mathrm{d}\mu_1 = \varepsilon^{-1} \int \varepsilon\rho_{K,\varepsilon}\, \mathrm{d}\mu_2 = \int \rho_{K,\varepsilon}\, \mathrm{d}\mu_2,$$

whence $\mu_1(K) = \mu_2(K)$. Finally, if Ω is locally compact, then we can choose the neighbourhoods U_x to be relatively compact. Hence U is relatively compact, whence $\rho_{K,\varepsilon}$ has compact support. $\square$

According to to Theorem E.1.10, $C_b(\Omega)$ is a separating class for $\mathcal{M}(\Omega)$. In particular, this means that $\langle \mathcal{M}(\Omega), C_b(\Omega)\rangle$ forms a dual pair under the duality

$$\langle \mu, f\rangle := I_\mu(f). \tag{E.1.2}$$

This lets us define the (weak) topology $\tau_{\mathrm{w}} := \sigma(\mathcal{M}(\Omega), C_b(\Omega))$. We note that the mapping $\mu \mapsto I_\mu(f)$ is τ_{w}-continuous for each $f \in C_b(\Omega)$. Thus, a net $\{\mu_\alpha\} \subset \mathcal{M}(\Omega)$ satisfies $\mu_\alpha \xrightarrow{\tau_{\mathrm{w}}} \mu$ if and only if $I_{\mu_\alpha}(f) \to I_\mu(f)$ for all $f \in C_b(\Omega)$. The following theorem gives us the key to comparing τ_{w}-convergence with the weak convergence of Definition 5.1.10.

Theorem E.1.12. *Ω is Polish if and only if $(\mathcal{M}(\Omega), \tau_{\mathrm{w}})$ is Polish.*

Proof. See [1, Theorem 15.15] $\square$

Note that in a metric space, we can replace the convergence of nets with the convergence of sequences. Thus, Theorem E.1.12 shows us that when Ω is Polish for $\{\mu_\alpha\} \cup \{\mu\} \subset \mathcal{M}(\Omega)$ we have $\mu_\alpha \xrightarrow{\tau_{\mathrm{w}}} \mu$ if and only if for each subsequence $\{n\} \subset \alpha$ we have $\text{w-lim}_{\to\infty} \mu_n = \mu$, so the two concepts of weak convergence coincide.

We now investigate in what sense this is weak* convergence. To do so, we include the integral representation theorem for $(C_b(\Omega))'$ found in the work of Dunford and Schwarz [28, Chapter IV.6, Theorem 2].

Definition E.1.13. For a Hausdorff pace Ω we define the space $rba(\Omega)$ to be the linear space of regular bounded additive set functions defined on $\mathscr{B}(\Omega)$. Associated to $rba(\Omega)$ is the total variation norm.

Theorem E.1.14. *Let Ω be a normal topological space. Then the dual norm of $C_b(\Omega)$ is isometrically isomorphic to $rba(\Omega)$ via the mapping $\mu \mapsto \int_\Omega \cdot\, \mathrm{d}\mu$.*

Thus, the weak topology $\sigma(rba(\Omega), C_b(\Omega))$ is equivalent to the weak* topology $\sigma((C_b(\Omega))', C_b(\Omega))$. It is clear that $\mathcal{M}(\Omega) \subseteq rba(\Omega)$, so $\{\mu_n\} \subset \mathcal{M}(\Omega)$ is weakly convergent in the sense of Definiton 5.1.10 if and only if $\{I_{\mu_n}\} \subset (C_b(\Omega))'$ is weak* convergent.

If the metrisable space is additionally locally compact, then Theorem E.1.11 shows that $C_c(\Omega)$ is a separating class for $\mathcal{M}(\Omega)$. In particular, we can define the dual pairing $\langle \mathcal{M}(\Omega), C_c(\Omega)\rangle$ by the duality (E.1.2) and let $\tau_{\mathrm{v}} := \sigma(\mathcal{M}(\Omega), C_c(\Omega))$, which we call the *vague topology* on $\mathcal{M}(\Omega)$.

It holds again that if Ω is Polish, then $(\mathcal{M}(\Omega), \tau_v)$ is Polish. Thus it is clear that for $\{\mu_\alpha\} \cup \{\mu\} \subset \mathcal{M}(\Omega)$ we have $\mu_\alpha \xrightarrow{\tau_{\mathrm{v}}} \mu$ if and only if for each subsequence $\{n\} \subset \alpha$ we have $\text{v-lim}_{\to\infty} \mu_n = \mu$. Furthermore, by Theorem 5.1.8(a) it follows that $(C_0(\Omega))'$ is isometrically isomorphic to $\mathcal{M}(\Omega)$. In particular, this means that $\{\mu_n\} \subset \mathcal{M}(\Omega)$ is vaguely convergent in the sense of Definiton 5.1.10 if and only if $\{I_{\mu_n}\} \subset (C_0(\Omega))'$ is weak* convergent. Of course, if Ω is compact then $\mathcal{M}(\Omega) = (C_c(\Omega))'$ and the notions of vague convergence and weak* convergence coincide.

Bibliography

[1] Charalambos D. Aliprantis and Kim C. Border, *Infinite dimensional analysis: A hitchhiker's guide*, Springer, 1999.

[2] R. F. Almgren, *Optimal execution with nonlinear impact functions and trading-enhanced risk*, Appl. Math. Finance **10** (2003), no. 1, 1–18.

[3] A. Arapostathis, V. S. Borkar, and M. K. Ghosh, *Ergodic control of diffusion processes*, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 2011.

[4] M. Arisawa, *Ergodic problem for the Hamilton-Jacobi-Bellman equation. I. Existence of the ergodic attractor*, Ann. Inst. H. Poincaré Anal. Non Linéaire **14** (1997), no. 4, 415–438.

[5] ———, *Ergodic problem for the Hamilton-Jacobi-Bellman equation. II*, Ann. Inst. H. Poincaré Anal. Non Linéaire **15** (1998), no. 1, 1–24.

[6] K. J. Arrow, *Collected papers of Kenneth J. Arrow, Volume 3: Individual choice under certainty and uncertainty*, Belknap Press, Cambridge, MA, 1984.

[7] R. J. Aumann, *Agreeing to disagree*, Ann. Stat. **4** (1976), no. 6, 1236–1239.

[8] B Basrak and H. Planinić, *A note on vague convergence of measures*, Statist. Probab. Lett. **153** (2019), 180–186.

[9] C. Berg, J. P.-R. Christensen, and P. Ressel, *Harmonic analysis on semigroups*, Graduate Texts in Mathematics, Springer New York, New York, NY, 1984.

[10] K. Bichteler, *Stochastic integration with jumps*, Encyclopedia of mathematics and its applications, Cambridge University Press, Cambridge, 2002.

[11] P. Billingsley, *Convergence of probability measures*, Wiley, 1968.

[12] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation*, Encyclopedia of mathematics and its applications, vol. 27, Cambridge university press, Cambridge, 1989.

[13] F Black and M. Scholes, *The pricing of options and corporate liabilities*, J. Pol. Econ. **81** (1973), no. 3, 637–654.

[14] V. I. Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007.

[15] ______, *Weak convergence of measures*, American mathematical society, 2018.

[16] M. Bonelli, A. Landier, G. Simon, and D. Thesmar, *The capacity of trading strategies.*

[17] B. Bouchard, M. Fukasawa, M. Herdegen, and J. Muhle-Karbe, *Equilibrium returns with transaction costs*, Finance Stoch. **22** (2018), no. 3, 569–601.

[18] N. Bourbaki, *Elements of mathematics: Integration I*, Springer, Berlin, Heidelberg, 2004.

[19] A. Buss, B. Dumas, R. Uppal, and G. Vilkov, *The intended and unintended consequences of financial-market regulations: A general-equilibrium analysis*, J. Monet. Econ. **81** (2016), 25–43.

[20] Adrian Buss and B. Dumas, *The dynamic properties of financial-market equilibrium with trading fees*, J. FInance **74** (2019), no. 2, 795–844.

[21] J. Cai, M. Rosenbaum, and P. Tankov, *Asymptotic optimal tracking: feedback strategies*, Stochastics **89** (2017), no. 6-7, 943–966.

[22] N. Chriss and R. F. Almgren, *Optimal execution of portfolio transactions*, J. Risk. **3** (2001), 5–40.

[23] J. C. Cox, J. E. Ingersoll, and S. A. Ross, *A theory of the term structure of interest rates*, Econometrica **53** (1985), no. 2, 385–407.

[24] D. J. Daley and D. Vere-Jones, *An introduction to the theory of point processes: Vol. I*, 2nd ed., Probability and its applications, Springer, New York, 2003.

[25] M. H. A. Davis and A. R. Norman, *Portfolio selection with transaction costs*, Math. Oper. Res. **15** (1990), no. 4, 676–713.

[26] Louis de Branges, *The Stone-Weierstrass theorem*, Proc. Amer. Math. Soc. **10** (1959), no. 5, 822–824.

[27] J. Dieudonné, *Treatise on analysis*, Pure and applied mathematics, vol. 10-II, Academic press, New York-London, 1970, Translated from the French by I. G. Macdonald.

[28] N. Dunford and J. T. Schwartz, *Linear operators, Part 1: General theory*, 1 ed., vol. 1, Wiley, New York, April 1988.

[29] E. Dávila, *Optimal financial transaction taxes*, J. Finance **78** (2023), no. 1, 5–61.

[30] L. C. Evans and R. F. Garzepy, *Measure theory and fine properties of functions*, Textbooks in mathematics, CRC Press, Boca Raton, FL, 2015.

[31] W. Feller, *On the classical Tauberian theorems*, Arch. Math **14** (1963), no. 1, 317–322.

[32] ______, *An introduction to probability theory and its applications. Vol. II.*, second ed., John Wiley & Sons, Inc., New York-London-Sydney, 1971.

[33] G. B. Folland, *Real analysis: Modern techniques and their applications, 2nd edition*, 2nd ed., Pure and applied mathematics, Wiley, New York, 1999.

[34] M. Fuhrman, Y. Hu, and G. Tessitore, *Ergodic BSDEs and optimal ergodic control in Banach spaces*, SIAM J. Control Optim. **48** (2009), no. 3, 1542–1566.

[35] L. Gonon, J. Muhle-Karbe, and X. Shi, *Asset pricing with general transaction costs: Theory and numerics*, Math. Finance **31** (2021), no. 2, 595–648.

[36] P. Guasoni and J. Muhle-Karbe, *Portfolio choice with transaction costs: a user's guide*, Paris-Princeton lectures on mathematical finance, Lecture Notes in Math., vol. 2081, Springer, Cham, 2013, pp. 169–201.

[37] N. Gârleanu and L.-H. Pedersen, *Dynamic trading with predictable returns and transaction costs*, J. Finance **68** (2013), no. 6, 2309–2340.

[38] G. H. Hardy and J. E. Littlewood, *Tauberian theorems concerning power series and Dirichlet's series whose coefficients are positive*, Proc. Lond. Math. Soc. (2) **13** (1914), no. 1, 174–191.

[39] J. Heaton and D. J. Lucas, *Evaluating the effects of incomplete markets on risk sharing and asset pricing*, J. Political Econ. **104** (1996), no. 3, 443–87.

[40] M. Herdegen, G. Liang, and O. Shelley, *A continuity theorem for generalised signed measures with an application to Karamata's Tauberian theorem*, 2023.

[41] ———, *Vague and weak convergence of signed measures*, 2023, Preprint arXiv:2205.13207.

[42] M. Herdegen, J. Muhle-Karbe, and D. Possamaï, *Equilibrium asset pricing with transaction costs*, Finance Stoch. **25** (2021), no. 2, 231–275.

[43] S.-T. Hu, *Introduction to general topology*, Holden-Day, San Francisco, Calif.-London-Amsterdam, 1966.

[44] O. Kallenberg, *Random measures, theory and applications*, Probability theory and stochastic modelling, vol. 77, Springer, Cham, 2017.

[45] ———, *Foundations of modern probability*, 3rd ed., Probability theory and stochastic modelling, vol. 99, Springer, Cham, 2021.

[46] J. Karamata, *Sur un mode de croissance régulière des fonctions*, Inst. de arte graf. Ardealul, 1930.

[47] J. M. Keynes, *General theory of employment, interest and money*, Atlantic Publishers & Dist, 2016.

[48] A. Klenke, *Probability theory: A comprehensive course*, 2nd ed., Universitext, Springer, London, 2014.

[49] L. Kogan, S. A. Ross, J. Wang, and M. M. Westerfield, *The price impact and survival of irrational traders*, J. FInance **61** (2006), no. 1, 195–229.

[50] P. Krugman, *Opinion | Taxing the speculators*, The New York Times (2009), Section A, Page 39.

[51] H. König, *Neuer beweis eines klassischen Tauber-Satzes*, Arch. Math **11** (1960), no. 1, 278–279.

[52] G. Leoni, *A first course in Sobolev spaces*, 2nd ed., Graduate studies in mathematics, vol. 181, American mathematical society, Providence, R.I, 2017.

[53] A. W. Lo, H. Mamaysky, and J. Wang, *Asset prices and trading volume under fixed transactions costs*, J. Pol. Econ. **112** (2004), no. 5, 1054–1090.

[54] R. C. Merton, *Optimum consumption and portfolio rules in a continuous-time model*, J. Econ. Theory **3** (1971), no. 4, 373–413.

[55] Aaron Meurer, Christopher P. Smith, Mateusz Paprocki, Ondřej Čertík, Sergey B. Kirpichev, Matthew Rocklin, AMiT Kumar, Sergiu Ivanov, Jason K. Moore, Sartaj Singh, Thilina Rathnayake, Sean Vig, Brian E. Granger, Richard P. Muller, Francesco Bonazzi, Harsh Gupta, Shivam Vats, Fredrik Johansson, Fabian Pedregosa, Matthew J. Curry, Andy R. Terrel, Štěpán Roučka, Ashutosh Saboo, Isuru Fernando, Sumith Kulal, Robert Cimrman, and Anthony Scopatz, *Sympy: symbolic computing in python*, PeerJ Comput. Sci. **3** (2017), e103.

[56] L. Moreau, J. Muhle-Karbe, and H. M. Soner, *Trading with small price impact*, Math. Finance **27** (2017), no. 2, 350–400.

[57] J. Muhle-Karbe, M. Nutz, and X. Tan, *Asset pricing with heterogeneous beliefs and illiquidity*, Math. Finance **30** (2020), no. 4, 1392–1421.

[58] J. Muhle-Karbe, X. Shi, and C. Yang, *An equilibrium model for the cross section of liquidity premia*, Math. Oper. Res. (2022).

[59] M. Nutz and J. A. Scheinkman, *Shorting in speculative markets*, J. FInance **75** (2020), no. 2, 995–1036.

[60] D. Pollard and F. Topsøe, *A unified approach to Riesz type representation theorems*, Stud. Math. **54** (1975), no. 2, 173–190.

[61] Y. V. Prokhorov, *Convergence of random processes and limit theorems in probability theory*, Theory Probab. Appl. **1** (1956), no. 2, 157–214.

[62] W. Rudin, *Functional analysis*, 2 ed., McGraw-Hill, 1991.

[63] J. Scheinkman and W. Xiong, *Heterogeneous beliefs, speculation and trading in financial markets*, Paris-Princeton lectures on mathematical finance 2003, Lecture notes in mathematics, Springer, Berlin, Heidelberg, 2004, pp. 217–250.

[64] R. L. Schilling, *Measures, integrals and martingales*, 2nd ed., Cambridge University Press, Cambridge, 2017.

[65] L. Schwarz, *Radon measures on arbitrary topological spaces and cylindrical measures*, Oxford University press for the Tata institute of fundamental research, London, 1973.

[66] W. F. Sharpe, *Capital asset prices: A theory of market equilibrium under conditions of risk*, J. Finance **19** (1964), no. 3, 425–442.

[67] Joseph E. Stiglitz, *Using tax policy to curb speculative short-term trading*, Journal of Financial Services Research **3** (1989), no. 2, 101–115 (en).

[68] L. H. Summers and V. P. Summers, *When financial markets work too well: A cautious case for a securities transactions tax*, J. Finan. Serv. Res. **3** (1989), no. 2, 261–286.

[69] The Sage Developers, *Sage math, the sage mathematics software system*, 2022, DOI 10.5281/zenodo.6259615.

[70] J. Tobin, *A proposal for international monetary reform*, East. Econ. J. **4** (1978), no. 3/4, 153–159.

[71] F. Topsøe, *Radon measures, some basic constructions*, Lecture notes in mathematics, Springer, Berlin, Heidelberg, 1983.

[72] Guido Van Rossum and Fred L Drake Jr, *Python tutorial*, Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands, 1995.

[73] V. S. Varadarajan, *Measures on topological spaces*, American mathematical society translations: Series 2, vol. 48, American mathematical society, 1965.

[74] D. Vayanos and J-L. Vila, *Equilibrium interest rate and liquidity premium with transaction costs*, Econ. Theory **13** (1999), no. 3, 509–539, Publisher: Springer.