DOCTORAL THESIS

Stochastic Functional Differential Equations and Optimal Consumption Problems

Emmet Lawless, BSc

Dissertation submitted for the award of Doctor of Philosophy

School of Mathematical Sciences

Dublin City University

August, 2025

Supervisors:

Professor Paolo Guasoni, **Dublin City University**

Professor John A.D. Appleby, **Dublin City University**

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work, and that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: Emmet Lawless

ID No.: 21267434

Date: 26/08/2025

Acknowledgements

I thank both my examiners Prof. Xuerong Mao and Prof. Brien Nolan for their careful reading of this thesis and their many suggestions and corrections which have lead to its improvement.

I owe a tremendous debt to both my supervisors, Prof. John A.D. Appleby and Prof. Paolo Guasoni.

John, your infectious enthusiasm for mathematics is my greatest source of inspiration. Through tremendous hardship, your much-needed support never wavered. It has been a pleasure and a privilege to do mathematics by your side. Thank you.

Paolo, without your patience and counsel through my seemingly endless mistakes, this thesis would not have been possible. Your guiding principles have instilled in me the utmost respect for what it means to be a mathematician. Thank you.

I give thanks to all the staff at the School of Mathematical Sciences at Dublin City University. In particular Asst. Prof. Martin Friesen.

Martin, although not officially my supervisor, I am truly grateful for your willingness to share your tremendous mathematical knowledge, and your much-needed advice. It has been an honour to learn from you. Thank you.

To my many colleagues in the postgraduate office — Huayuan, Ran, Ciara, Ole, Lorenzo, Lucy, Aditya, Josh, Connor, Brian, Evan, Jack, and Eneko — thank you all for making my years as a doctoral student an unforgettable experience.

Ole, I bid you special thanks for your careful reading, valuable suggestions, and improvements to much of the mathematics in this thesis. Our numerous discussions over coffee and beers have played a key role in my growth as a mathematician.

I extend my deepest gratitude to my parents. Your unwavering support and encouragement have afforded me the opportunity to pursue a life I would never have thought possible. For this, I will be forever grateful.

Dedication

To my parents, David and Brenda.

Contents

List of Figures

Nomenclature

$C(X;Y)$ Continuous functions from X to Y.

$C^\infty(X;Y)$ Continuous and infinitely differentiable functions from X to Y.

$C_c^\infty(X;Y)$ Continuous and infinitely differentiable functions from X to Y with compact support.

$C^{k,\alpha}(X;Y)$ Space of k-times continuously differentiable functions from X to Y whose k^{th} order partial derivatives are Hölder continuous with exponent $\alpha \in (0,1)$.

$L^p(X;Y)$ Measurable functions from X to the normed space Y such that $\int_X \|f\|_Y^p < \infty$.

$L^p_{\text{loc}}(X;Y)$ Measurable functions from X to the normed space Y such that for all compact subsets K of X $\int_K \|f\|_Y^p < \infty$.

$M(J;Y)$ Y- valued, finite, signed, Borel measures on an interval $J \subset \mathbb{R}$.

$BC_0(\mathbb{R}_+;Y)$ Bounded, continuous functions from $\mathbb{R}_+$ to the normed space Y such that $\lim_{t\to\infty}\|f(t)\|_Y = 0$.

$\text{Ces}(\mathbb{R}_+;\mathbb{R})$ Functions in $L^1_{\text{loc}}(\mathbb{R}_+;\mathbb{R})$ such that $\lim_{t\to\infty} \frac{1}{t}\int_0^t f(s)ds$ exists.

$\ell^p(\mathbb{N};Y)$ Y-valued sequences such that $\sum_{n\in\mathbb{N}}\|f(n)\|_Y < \infty$.

$L^p_\eta(X;Y)$ Measurable functions from X to the normed space Y such that $\int_X \|f\|_Y^p\,\eta < \infty$.

X^* The dual space of the Banach Space X

Stochastic Functional Differential Equations and Optimal Consumption Problems

Emmet Lawless

Abstract

The first part of this thesis is concerned with the qualitative behaviour of path dependent stochastic differential equations, specifically equations of functional and Volterra type. We deeply investigate the effects of perturbations on stable linear systems. First, we provide a framework for characterising the solution space of perturbed linear integro-differential Volterra convolution equations and their finite memory counterpart, namely perturbed linear functional differential equations. This improves upon classical admissibility theory which only provides sufficient conditions on the perturbations to ensure solutions are elements of a particular function space.

Next we consider state independent stochastic perturbations and aim to characterise when the sample paths of the solution converge to zero almost surely and when the trajectories are almost surely p-integrable functions in time. We then shift our focus to perturbed equations with multiplicative noise and characterise the asymptotic behaviour of the mean square. Our findings indicate the pointwise behaviour of state independent perturbations is non-dominant and does not necessarily proliferate through to the solution. Rather the main quantity that determines the qualitative behaviour of the solution is a particular functional of the state independent perturbations, namely one must study the integral of the perturbations over compact intervals. It is this insight that leads to sharp characterisations of the qualitative behaviour of solutions.

The second part of this thesis is dedicated to a particular class of stochastic control problems arising in finance. We provide a novel approach to the optimal consumption-investment problem which aims to characterise the value function as the solution to an associated variational problem. The utility of this approach is twofold, it circumvents the need to analytically solve the Hamilton-Jacobi-Bellman equation while also providing a new set of tools to analyse the value function. Sharp estimates on the value function and the optimal policies are obtained via analysing the variational problem directly and employing standard techniques from the theory of elliptic partial differential equations.

Part I

Stochastic Functional Differential Equations

Chapter 1

Introduction and Overview

1.1 Introduction and Motivation

Roughly speaking, a dynamical system can be described as the movement of a particle[1] through an ambient space with respect to a given rule which influences its future evolution over time. The term *influences* was carefully chosen, at a given time if such an evolution rule completely determines the state of the system then we call this a *deterministic* dynamical system. On the contrary it may be the case that the evolution rule does not completely determine the future trajectory, rather at each fixed time, the system depends on a parameter which is unknown a priori and comes from a particular probability distribution, we call this a *random* dynamical system. Both notions give rise to a rich mathematical theory and are commonplace in applications.

One of the most important types of continuous dynamical systems are those described by differential equations. We first describe deterministic systems, consider the non-autonomous system of ordinary differential equations (ODEs):

$$\dot{x}(t) = b(t, x(t)); \quad x(t_0) = x_0 \in \mathbb{R}^d \tag{1.1.1}$$

where $b : \mathbb{T} \times \mathbb{R}^d \to \mathbb{R}^d$, and $\mathbb{T}$ is some connected subset of $\mathbb{R}$ (see [120] for an introductory exposition). In this instance the ambient space is $\mathbb{R}^d$ and the evolution rule is the solution map $x : \mathbb{T} \to \mathbb{R}^d$. When studying ODEs, the solution map depends only on time and takes values in a finite dimensional set, oftentimes one would also like to incorporate spatial dependence into the evolution rule. However this leads to the study of partial differential equations (PDEs) which will not feature in this thesis. An implicit assumption when using ODEs to describe a dynamical system is that the future evolution of the particle depends only on the current position and the starting time. A natural generalisation is to consider an evolution rule which depends also on the path taken by the particle. This

[1] We use the word particle here in an abstract sense which can represent a point in any abstract space.

leads to the following equation,

$$\dot{x}(t) = b(t, x), \tag{1.1.2}$$

where $b : \mathbb{T} \times C(\mathbb{T}; \mathbb{R}^d) \to \mathbb{R}^d$. This is a general path-dependent ODE and is referred to as a functional differential equation (FDE). We mention the following monographs which provide excellent expositions of the general theory of FDEs [58, 72].

The underlying motivation for studying such systems comes, of course, from applications. In many fields one is faced with the task of describing the evolution of some phenomena. An approach which has had an outstanding effect in helping us understand the natural world, is that of mathematical modelling. In this context, modelling is the act of writing down a dynamical system, which either recreates the stylised facts of the phenomena under observation or predicts its evolution (up to some negligible error). However, writing down the exact evolution rule is practically impossible, so instead we study a differential equation which we believe induces an evolution rule which describes the phenomena. The simple question is then, how do we know if the induced evolution rule is correct?

If the differential equation is explicitly solvable, then we can look at the graph of the solution and compare with empirical observations. This is the ideal situation but unfortunately is practically never the case. If explicit solutions are not available, one then turns to numerical schemes. This is a highly effective method for understanding the behaviour of solutions, the drawback lies in the lack of a guarantee that the actual solution exhibits the behaviour we observe from numerical simulations. Thus we seek analytical results regarding the behaviour of solutions of differential equations without access to an explicit representation of the solution, this pursuit is known as analysis of the qualitative behaviour of differential equations. This is the primary focus of Part I of this thesis.

If one wishes to obtain interesting information about the solution map of (1.1.2), then one needs to impose more structure on the path dependence. When the dependence is confined to a compact subset of $\mathbb{T}$, i.e. for any fixed $t \in \mathbb{T}$, the future evolution depends only on the path of x on the interval $[t - \tau, t]$ for some $\tau > 0$, then we call (1.1.2) a delay differential equation (DDE). For example consider the scalar equation:

$$\dot{x}(t) = \delta x(t) + \sum_{j=1}^{n} \alpha_j x(t - \tau_j) + e^{-\beta t} \sin(t) - \int_{t-\tau}^{t} (t-s)^{\gamma-1} e^{-\lambda(t-s)} x(s) ds,$$

where $\delta, \alpha_j \in \mathbb{R}$, $\beta, \lambda, \gamma > 0$ and $\tau_j \in [0, \tau]$. In this equation the path dependence is a combination of point masses and distributed delay according to a gamma distribution. More information on this specific class of equations can be found in [44].

When the evolution at time $t \in \mathbb{T}$ is allowed to depend on the entire path up to time t then we call (1.1.2) a Volterra integro-differential equation (VIDE) or simply a Volterra equation (VE). For example,

$$\dot{x}(t) = Ax(t) + \int_0^t \frac{B\,x(s)}{1+(t-s)^2}ds - \frac{\log(1+t)}{1+t^2},$$

where A, B are $d \times d$ real valued matrices. A thorough account of the theory of such equations can be found in the monograph [53]. We note that in the two examples provided, the right-hand side is linear in x and the explicit dependence on t (besides the kernel) is of an additive form. Such structure is a recurring theme throughout this thesis.

As previously mentioned, the main motivation for the mathematical study of such systems is applications. When building a model, there is a trade-off between model complexity and model accuracy. For instance it is natural to suggest that as one increases the number of factors taken into consideration the accuracy of the model should increase. However there is often a tipping point at which the added complexity makes analysing the model and inferring any reasonable predictions intractable. An extremely effective way around this problem is the addition of noise. The idea is simple, there are too many external factors at play to account for everything in a particular model, so instead we make the assumption that the cumulative behaviour of all these factors acts as a kind of randomness that cannot be perfectly described. In many cases this allows for a significant reduction in model complexity while simultaneously increasing model accuracy. The mathematical objects we use to do this are path-dependent stochastic differential equations (SDEs).

1.2 Path dependent SDEs: General description

In Part I of this thesis, we study the solution maps of both deterministic and stochastic linear path-dependent integral equations. In particular we primarily study the qualitative behaviour of solutions to path-dependent SDEs of the form:

$$\begin{aligned} X(t) &= \psi(0) + \int_0^t b(s, X.)ds + \int_0^t a(s, X.)dB(s), && t \geq 0, \\ X(t) &= \psi(t), && -\tau \leq t < 0. \end{aligned} \tag{1.2.1}$$

where $X.(\omega) := t \mapsto X(t,\omega)$ is an entire path of the process X for each fixed $\omega \in \Omega$ and $\tau > 0$. We shall always employ the implicit assumption of working on a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$, supporting an m-dimensional Brownian motion B and a random variable ψ representing the initial condition of our system. The coefficients are defined as maps $b : \mathbb{R}_+ \times C([-\tau,\infty);\mathbb{R}^d) \times \Omega \to \mathbb{R}^d$ and $a : \mathbb{R}_+ \times C([-\tau,\infty);\mathbb{R}^d) \times \Omega \to \mathbb{R}^{d\times m}$ such that all integrals are well defined. Next we state our notion of solution and provide a well-posedness theorem. Let $A := aa^\top$ and $f_t^* := \sup_{s\leq t} |f(s)|$ be the running maximum.

Definition 1.2.1. *We say the triple (X, B, ψ) defined on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ is a solution of (1.2.1) if*

(i) B is a standard $\mathbb{R}^m$-valued $\mathcal{F}_t$-BM, ψ is an $\mathcal{F}_0$-meas. r.v and X is $\mathcal{F}_t$-adapted.

(ii) $\int_0^t |b(s, X_\cdot)| + \operatorname{tr}(A(s, X_\cdot))\, ds < \infty$ for each $t > 0$ a.s..

(iii) $X(t) = \psi(0) + \int_0^t b(s, X_\cdot)ds + \int_0^t a(s, X_\cdot)dB(s)$.

Theorem 1.2.1. *Let $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ be a filtered probability space (satisfying the usual conditions) supporting an $\mathbb{R}^m$-valued Brownian Motion and an $\mathcal{F}_0$-measurable r.v. $\psi \in L^2(\Omega; C([-\tau, 0]; \mathbb{R}^d))$. If for every $T \in (0, \infty)$, there exists a positive constant C_T such that for all $t \in [0, T]$ and all $\phi, \varphi \in C([-\tau, \infty); \mathbb{R}^d)$*

$$|b(t, \phi) - b(t, \varphi)| + |a(t, \phi) - a(t, \varphi)| \leq C_T(\phi - \varphi)^*_t, \tag{1.2.2}$$

then there exists a solution to equation (1.2.1) and pathwise uniqueness holds[2]. Moreover if $p \geq 1$ and $\psi \in L^p(\Omega; C([-\tau, 0]; \mathbb{R}^d))$ then $\mathbb{E}\left[(X)^{*p}_t\right] < +\infty$ for all $t > 0$.

Well-posedness of all stochastic equations appearing in Part I of this thesis is guaranteed by Theorem 1.2.1. The proof is standard and can be found in many texts on stochastic differential equations, see the monographs [36, 51, 70, 85, 100, 108].

We shall often be concerned with equations on the half line $[0, +\infty)$. The fact Theorem 1.2.1 is formulated on compacts poses no issue as $T > 0$ is arbitrary and pathwise uniqueness holds. Thus, we may always consider equations on $[0, +\infty)$. We do not consider systems with infinite memory, hence the parameter τ controls how far back into the past the solution can depend. We do note however, this set up allows for equations whose memory grows as $t \uparrow +\infty$. The well posedness of (1.2.1) has been a topic of intense study over the past fifty years with widely varying hypotheses on the coefficients, this is not the main focus of this thesis so we do not provide a cutting edge account of the existence and uniqueness literature.

Once well posedness has been established it is natural to ask whether we can describe the dynamics of the solution. By this we mean what information can we obtain about the solution map $X : [-\tau, \infty) \times \Omega \to \mathbb{R}^d$. An explicit representation for X in terms of elementary functions and the driving Brownian motion is almost never available which makes inferring properties about the solution map a challenging (and thus mathematically interesting) task. We shall mainly focus on the qualitative behaviour of the solution process X and of certain functionals of X.

[2]This ensures that the solution is strong in the sense that X is adapted to the filtration generated by the driving Brownian motion.

1.3 Overview

Our primary concern is the effect of deterministic, state independent perturbations on the solution map for a specific class of coefficients. To make this precise, consider the coefficient functions,

$$b(t,\varphi) = f(t) + \tilde{b}(t,\varphi); \qquad a(t,\varphi) = \sigma(t) + \tilde{a}(t,\varphi),$$

where f, σ are continuous, $\mathbb{R}^d$ and $\mathbb{R}^{d\times m}$ valued functions, and $\tilde{b}(t,0) = 0$ and $\tilde{a}(t,0) = 0$. We then study the solutions to the equations,

$$\tilde{X}(t) = \psi(0) + \int_0^t \tilde{b}(s,\tilde{X}_\cdot)ds + \int_0^t \tilde{a}(s,\tilde{X}_\cdot)dB(s), \tag{1.3.1}$$

$$X(t) = \psi(0) + \int_0^t f(s) + \tilde{b}(s,X_\cdot)ds + \int_0^t \sigma(s) + \tilde{a}(s,X_\cdot)dB(s). \tag{1.3.2}$$

For illustrative purposes we describe our main goal with a formal argument which is made rigorous in the subsequent chapters. Consider the operator $T_{\tilde{X}}$ which acts on perturbation functions f, σ and returns the solution to the perturbed SDE X, i.e $T_{\tilde{X}}(f,\sigma) = X$. Next we specify our desired co-domain which we call V. We now have,

$$T_{\tilde{X}} : \mathcal{D}(T_{\tilde{X}}) \to V.$$

Part I of this thesis will be dedicated to characterising the domain of the operator $T_{\tilde{X}}$ for various spaces of interest V and different classes of coefficient functions $\tilde{b}, \tilde{a}$. For example, consider the space of processes

$$\mathcal{V}_0^2 := \{X \in L^2(\Omega; C([-\tau,T];\mathbb{R}^d)) \text{ for all } T > 0 : \mathbb{E}[|X(t)|^2] \to 0 \text{ as } t \to \infty\}.$$

Then we have $T_{\tilde{X}} : \mathcal{D}(T_{\tilde{X}}) \to \mathcal{V}_0^2$, hence characterising $\mathcal{D}(T_{\tilde{X}})$ means characterising the perturbation functions that result in the mean square of the solution of (1.3.2) vanishing asymptotically. Of course one needs to make appropriate assumptions on the solution of the unperturbed equation (i.e $\tilde{X}$) in order to make meaningful progress on this question. With our main goal now clear, we specify which classes of coefficients we shall study throughout the following three chapters.

In Chapter 2 we consider deterministic equations (i.e, we set $a = 0$ and fix a constant initial condition), for two types of path dependence, namely

$$b(t,\varphi) = f(t) + \int_{[0,t]} \nu(ds)\varphi(t-s); \qquad a(t,\varphi) = 0,$$

$$b(t,\varphi) = f(t) + \int_{[-\tau,0]} \mu(ds)\varphi(t+s); \qquad a(t,\varphi) = 0,$$

where $f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R})$[3], $\nu \in M(\mathbb{R}_+;\mathbb{R})$ and $\mu \in M([-\tau,0];\mathbb{R})$. Hence we first consider path dependence of Volterra type and then so called functional equations. It is true that one can embed functional equations within the Volterra theory, however in certain cases it can be fruitful to exploit the added structure of a compactly supported kernel. This specialisation still encompasses a vast array of interesting equations but also allows one to prove stronger results. With that being said our main focus is still Volterra equations so we focus our discussion to this case. Consider the equations,

$$\begin{aligned}\dot{\tilde{x}}(t) &= \int_{[0,t]} \nu(ds)\tilde{x}(t-s), && t>0; \quad x(0)=\psi \in \mathbb{R}\backslash\{0\},\\ \dot{x}(t) &= \int_{[0,t]} \nu(ds)x(t-s) + f(t), && t>0; \quad x(0)=\psi \in \mathbb{R}\backslash\{0\}. \end{aligned} \tag{1.3.3}$$

In this case the operator $T_{\tilde{x}}$ can actually be rigorously defined for many interesting spaces by availing of the variation of constants formula,

$$T_{\tilde{x}} : \mathcal{D}(T_{\tilde{x}}) \to V,\ f \mapsto \tilde{x} + \xi(\tilde{x} * f),$$

where $\xi = 1/\psi$. We focus on the cases $V = L^p(\mathbb{R}_+;\mathbb{R}), BC_0(\mathbb{R}_+;\mathbb{R}), \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$. Classical theory tells us in each of these cases $V \subseteq \mathcal{D}(T_{\tilde{x}})$ under the assumption that $\tilde{x} \in L^1(\mathbb{R}_+;\mathbb{R})$. This follows from the fact that the operator defined by convolution with an $L^1(\mathbb{R}_+;\mathbb{R})$ function defines an endomorphism on each of these spaces. Our contribution is that we show the domain is actually strictly larger then the classical theory may suggest, in fact we provide a complete characterisation of $\mathcal{D}(T_{\tilde{x}})$ in each of these three cases. The culmination of all main results from Chapter 2 is presented in the following theorem,

Theorem 1.3.1. *Let $V = BC_0(\mathbb{R}_+;\mathbb{R}), \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$, or $L^p(\mathbb{R}_+;\mathbb{R})$ for $p \geq 1$, and assume $\tilde{x} \in L^1(\mathbb{R}_+;\mathbb{R})$. Then,*

$$\mathcal{D}(T_{\tilde{x}}) = \left\{ f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}) : t \mapsto \int_t^{t+\theta} f(s)ds \in V \text{ for all } \theta \in (0,1] \right\}, \tag{1.3.4}$$

i.e

$$x \in V \iff f \in \left\{ f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}) : t \mapsto \int_t^{t+\theta} f(s)ds \in V \text{ for all } \theta \in (0,1] \right\}. \tag{1.3.5}$$

After stating this result, the following natural questions arise:

1. Is $\left\{ f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}) : t \mapsto \int_t^{t+\theta} f(s)ds \in V \text{ for all } \theta \in (0,1] \right\}$ an enlargement of V?
2. Why consider only $\theta \in (0,1]$?

[3] As we first deal with deterministic equations the continuity assumption on the perturbation function may be dropped and shall only be reinstated when we study stochastic equations.

3. Is the assumption $\tilde{x} \in L^1(\mathbb{R}_+;\mathbb{R})$ necessary?

4. Why is the mapping $t \mapsto \int_t^{t+\theta} f(s)ds$ the correct object to study?

We address these four questions in details throughout Chapter 2. Ones intuition suggests that we have enlarged $\mathcal{D}(T_{\tilde{x}})$, however a priori this is not completely trivial. To satisfy any doubt we provide examples of functions f such that $f \in \mathcal{D}(T_{\tilde{x}})$ but $f \notin V$ (see examples 2.3.1 and 2.4.3). This proves we have in fact strictly enlarged the domain in the sense that $V \subset \mathcal{D}(T_{\tilde{x}})$, i.e V is a proper subset of $\mathcal{D}(T_{\tilde{x}})$.

Regarding the size of the interval of integration, this is in fact a matter of preference. One can prove the following sets are equal:

(i) $\left\{f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}) : t \mapsto \int_t^{t+\theta} f(s)ds \in V \text{ for all } \theta \in (0,1]\right\}$

(ii) $\left\{f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}) : t \mapsto \int_t^{t+\theta} f(s)ds \in V \text{ for all } \theta > 0\right\}$.

The reverse inclusion is immediate. For the forward inclusion fix an arbitrary $T > 0$ and write

$$\int_t^{t+T} f(s)ds = \sum_{j=1}^{N} \int_{t+\frac{(j-1)T}{N}}^{t+\frac{jT}{N}} f(s)ds,$$

where N is chosen so that the interval of integration is smaller than one for each summand. Hence by (i) each summand is an element of V, as V is a linear space we are done. Throughout the thesis we shall work with both definitions depending on our particular needs at the time, for instance in many situations it is advantageous to work with θ taking values in a bounded set. We shall use this equivalence interchangeably throughout the thesis without further reference.

The assumption of $\tilde{x} \in L^1(\mathbb{R}_+;\mathbb{R})$ may seem strong at first but upon some reflection one realises it is actually quite natural indeed. It can even be characterised in terms of a condition on the Laplace transform of the kernel [44, 53]. In a sense the optimal result would be to prove Theorem 1.3.1 under the assumption that $\tilde{x} \in V$, however this would require proving $f, g \in V \implies f * g \in V$. The problem is that this implication is provably false in general, so one needs to make stronger assumptions with the natural candidate being integrability. This is exactly the assumption made in the classical theory [53] and is advantageous as $\tilde{x} \in L^1(\mathbb{R}_+;\mathbb{R}) \implies \tilde{x} \in V$.

The importance of the mapping $t \mapsto \int_t^{t+\theta} f(s)ds$ cannot be understated, it arises from a technical lemma which acts as the backbone of all results from Part I of this thesis. It is in this regard that one can consider the following lemma the most important result of Part I.

Lemma 1.3.1. *Let $f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R})$ and $V = BC_0(\mathbb{R}_+;\mathbb{R}), \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ or $L^p(\mathbb{R}_+;\mathbb{R})$ for $p \geq 1$. Then the following are equivalent,* *(i)* $t \mapsto \int_t^{t+\theta} f(s)ds \in V$ *for all* $\theta \in (0,1]$.

(ii) $f = f_1 + f_2$ *such that* $f_1 \in V$ *and* $f_2 \in L^1_{loc}(\mathbb{R}_+;\mathbb{R})$ *with* $t \mapsto \int_0^t f_2(s)ds =: f_3 \in V$.

To the best of our knowledge this lemma was first proved in the case $V = BC_0(\mathbb{R}_+;\mathbb{R})$ by Gripenberg, Londen and Staffans [53, Lemma 15.9.2]. This decomposition is what allows one to consider perturbation functions $f \notin V$. In general assuming only $\tilde{x} \in L^1(\mathbb{R}_+;\mathbb{R})$, it is not necessarily the case that $\tilde{x} * f \in V$, however with the decomposition from Lemma 1.3.1 we have (after integration by parts) $\tilde{x} * f = \tilde{x} * f_1 + f_3 + \dot{\tilde{x}} * f_3$. Noting $\tilde{x} \in L^1(\mathbb{R}_+;\mathbb{R}) \implies \dot{\tilde{x}} \in L^1(\mathbb{R}_+;\mathbb{R})$, we can conclude that $\tilde{x} * f \in V$ as required.

In Chapter 3 we consider stochastic perturbations, namely we consider coefficient functions,

$$\begin{aligned}
b(t,\varphi) &= f(t) + \int_{[0,t]} \nu(ds)\varphi(t-s); \qquad & a(t,\varphi) &= \sigma(t),\\
b(t,\varphi) &= f(t) + \int_{[-\tau,0]} \mu(ds)\varphi(t+s); \qquad & a(t,\varphi) &= \sigma(t),
\end{aligned}$$

where $f \in C(\mathbb{R}_+;\mathbb{R}^d)$, $\sigma \in C(\mathbb{R}_+;\mathbb{R}^{d\times m})$, $\nu \in M(\mathbb{R}_+;\mathbb{R}^{d\times m})$ and $\mu \in M([-\tau,0];\mathbb{R}^{d\times m})$. Notably in Chapter 3 we consider equations in finite dimensions. As a special case this generalises some of the results from Chapter 2 to multidimensions. The two main equations now under consideration are the following Volterra equations,

$$\begin{aligned}
\dot{x}(t) &= \int_{[0,t]} \nu(ds)x(t-s), & t>0;\quad x(0) = \psi \in \mathbb{R}\backslash\{0\},\\
dX(t) &= \left(\int_{[0,t]} \nu(ds)X(t-s) + f(t)\right)dt + \sigma(t)dB(t), \quad & t>0;\quad x(0) = \psi \in \mathbb{R}\backslash\{0\}.
\end{aligned} \tag{1.3.6}$$

In this instance we can once again rigorously define our perturbation operator by making use of the variation of constants formula,

$$T_x|_U : \mathcal{D}(T_x|_U) \to \mathcal{V},\ (f,\sigma) \mapsto x + \xi(x*f) + \xi(x*\sigma)_B,$$

where $U = C(\mathbb{R}_+;\mathbb{R}^d) \times C(\mathbb{R}_+;\mathbb{R}^{d\times m})$, $\xi = 1/\psi$ and $(x*\sigma)_B := \int_0^\cdot x(t-s)\sigma(s)dB(s)$. We use calligraphic $\mathcal{V}$ to emphasize our co-domain is a space of processes and not deterministic functions. In Chapter 3 we are primarily interested in the following spaces of processes which are maps from $\Omega \to C(\mathbb{R}_+;\mathbb{R}^d)$,

$$\begin{aligned}
\mathcal{LV}^p &:= \left\{X \in L^p(\Omega; C([0,T];\mathbb{R}^d)) \text{ for all } T>0 : \int_0^\infty \mathbb{E}[|X(t)|^p]dt < +\infty\right\},\\
\mathcal{L}^p &:= \left\{X \in L^p(\Omega; C([0,T];\mathbb{R}^d)) \text{ for all } T>0 : \int_0^\infty |X(t)|^p dt < +\infty \quad a.s\right\}.
\end{aligned}$$

One of our main results from Chapter 3 states that the domain of $T_x|_U$ is the same for both spaces $\mathcal{LV}^p$ and $\mathcal{L}^p$. By applying Fubini's theorem one can see that $\mathcal{LV}^p \subseteq \mathcal{L}^p$, however the reverse inclusion is false in general. Once again we provide a clean characterisation for the domain of $T_x|_U$,

Theorem 1.3.2. *Let $\mathcal{V} = \mathcal{LV}^p$ or $\mathcal{V}^p$ and assume $x \in L^1(\mathbb{R}_+;\mathbb{R})$. Then,*

(i) For $p \in [2,\infty)$, $\mathcal{D}(T_x|_U)$ is exactly the set of pairs $(f,\sigma) \in U$ such that,

$$t \mapsto \int_t^{t+\theta} f(s)ds \in L^p(\mathbb{R}_+;\mathbb{R}^d)\,\forall\,\theta \in (0,1]; \quad t \mapsto \int_t^{t+1} \sigma_{ij}^2(s)ds \in L^{\frac{p}{2}}(\mathbb{R}_+;\mathbb{R}),$$

for $1 \le i \le d$ and $1 \le j \le m$.

(ii) For $p \in [1,2)$, $\mathcal{D}(T_x|_U)$ is exactly the set of pairs $(f,\sigma) \in U$ such that,

$$t \mapsto \int_t^{t+\theta} f(s)ds \in L^p(\mathbb{R}_+;\mathbb{R}^d)\,\forall\,\theta \in (0,1]; \quad n \mapsto \int_n^{n+1} \sigma_{ij}^2(s)ds \in \ell^{\frac{p}{2}}(\mathbb{N};\mathbb{R}),$$

for $1 \le i \le d$ and $1 \le j \le m$.

The boundary of $p = 2$ is not surprising and arises as a technicality due to Itô's isometry. One should note that the conditions on the perturbation functions are of the same type as in the deterministic case. As in the deterministic case we address the assumption of $x \in L^1(\mathbb{R}_+;\mathbb{R}^d)$ and also conduct analysis on the asymptotic behaviour of the paths in an almost sure sense. In order to prove our main result we require a lemma which deals with convergence of certain random series. As a by product we obtain analogous ℓ^p summability results for the discrete version of our continuous Volterra equation.

The approach taken in Chapter 3 is somewhat different then that of Chapter 2. By this we mean the explicit representation of the operator $T_x|_U$ which is provided by the variation of constants formula does not play a major role. Instead we take an indirect route by studying an associated Markovian SDE and leveraging the results from Chapter 2. The main use of the deterministic results lies in considering the process $Z(t) = X(t) - Y(t)$ where $Y(t)$ is the solution to an SDE, namely: $dY(t) = (f(t) - Y(t))dt + \sigma(t)dB(t)$. With this observation one can show that for each fixed ω, Z satisfies an equation of type (2.1.1), see (Chapter 3, Lemma 3.3.1). Thus the study of the Volterra equation reduces to the study of a Markovian SDE and the deterministic equation (1.3.3).

In Chapter 4 we devote our attention on multiplicative noise equations with compactly supported kernels. Consider the coefficient functions,

$$b(t,\varphi) = f(t) + \int_{[-\tau,0]} \nu(ds)\varphi(t+s); \qquad a(t,\varphi) = \sigma(t) + \int_{[-\tau,0]} \mu(ds)\varphi(t+s),$$

where $f, \sigma \in C(\mathbb{R}_+)$ and $\nu, \mu \in M([-\tau, 0]; \mathbb{R})$. Notice we are now in a scalar regime, and are focused on the following equations for $t > 0$:

$$\begin{aligned} d\tilde{X}(t) &= \left(\int_{[-\tau,0]} \nu(ds)\tilde{X}(t-s)\right)dt + \left(\int_{[-\tau,0]} \mu(ds)\tilde{X}(t+s)\right)dB(t), && t>0, \\ dX(t) &= \left(f(t) + \int_{[-\tau,0]} \nu(ds)X(t-s)\right)dt + \left(\sigma(t) + \int_{[-\tau,0]} \mu(ds)X(t+s)\right)dB(t), && t>0, \end{aligned} \tag{1.3.7}$$

and $\tilde{X}(t) = X(t) = \psi(t)$ for $t < 0$ where $\psi \in L^2(\Omega; C([-\tau, 0]; \mathbb{R}))$. A variation of constants formula which gives the solution of equation (1.3.7) explicitly in terms of the solution to the unperturbed is (to the best of our knowledge) not available in the literature, hence our perturbation operator is no longer well defined in this regime. Nevertheless our goal for Chapter 4 is of a similar flavour as previous analysis, in particular we characterise when the following three types phenomena occur,

$$\lim_{t\to\infty} \mathbb{E}\left[X^2(t)\right] = 0; \qquad \int_0^\infty \mathbb{E}[X^2(t)]dt < +\infty; \qquad \mathbb{E}\left[X^2(t)\right] \le Ce^{-\alpha t},$$

for all initial conditions $\psi \in L^2(\Omega, C([-\tau, 0]; \mathbb{R}))$. In each case we find exponential asymptotic stability of the solution of the unperturbed equation a necessary condition, meaning the perturbation functions cannot have a regularising effect. This is consistent with our analysis in previous chapters when considering compactly supported kernels. A complete characterisation of the mean square behaviour of the unperturbed equation is provided in [16]. We state one of our three main theorems from Chapter 4 to illustrate the connectedness of all results throughout Part I of this thesis.

Theorem 1.3.3. *Let X be the solution to equation (1.3.7) and x solve the unperturbed deterministic equation*[4] *. Then the following conditions **(A)** and **(B)** are equivalent:*

(A) *(i)* $x \in L^1(\mathbb{R}_+)$

(ii) $\int_0^\infty \left(\int_{[-\tau,0]} \mu(ds)x(t+s)\right)^2 dt < 1$

(iii) $t \mapsto \int_t^{t+\theta} f(s)ds \in L^2(\mathbb{R}_+)$ *for all* $\theta > 0$,

(iv) $t \mapsto \int_t^{t+1} \sigma^2(s)ds \in L^1(\mathbb{R}_+)$.

(B) $\mathbb{E}[X^2(\cdot)] \in L^1(\mathbb{R}_+)$, *for all* $\psi \in L^2(\Omega; C([-\tau, 0]; \mathbb{R}))$.

One can see this recovers the main result from Chapter 3 in the special case of a compactly supported drift kernel, the diffusion kernel as the zero measure and $p = 2$. The proof techniques in Chapter 4 generalise ideas from [16] to perturbed equations, we show the squared mean of a particular functional of X obeys a deterministic renewal type equation from which the dynamics of the mean square of X itself, can be inferred.

[4] The equation with $f = \sigma = 0$ and μ equal to the zero measure.

Chapter 2

Deterministic Equations

The material in this chapter is based on the following articles:

1. J. A. D. Appleby and E. Lawless. Solution space characterisation of perturbed linear Volterra integrodifferential convolution equations: The L^p case. *Applied Mathematics Letters*, 146:108825, (2023) [12].

2. J. A. D. Appleby and E. Lawless. Solution space characterisation of perturbed linear functional and integrodifferential Volterra convolution equations: Cesàro limits. *Springer Proceedings in Mathematics and Statistics,* to appear, (2025) [15].

2.1 Introduction

In recent years there has been a surge of interest, from both the mathematical community and those in the applied sciences, in evolution equations of Volterra type. Traditional applications are found in that of population dynamics and modelling of hereditary systems [41] while stochastic extensions of the classical theory have recently attracted the attention of the finance community: see Abi Jaber [1] and the references therein. Analysis of the asymptotic behaviour of both deterministic and stochastic Volterra type integro-differential equations has attracted a lot of attention in recent decades. For deterministic linear equations of non-convolution type, recent works such as [29] investigate stability of the zero solution of systems of perturbed Volterra integro-differential equations while for stochastic equations asymptotic mean square stability is addressed in [5]. Consider the scalar equation,

$$\dot{x}(t) = \int_{[0,t]} \nu(ds)x(t-s) + f(t), \quad t \geq 0, \tag{2.1.1}$$

where ν is a finite signed Borel measure and $x(0) = \xi \in \mathbb{R}$ is a given initial condition. Such convolution equations are in general well understood, to wit, many monographs both classic and modern (such as [30, 37, 38, 53]) have been produced providing an in depth study to which the reader is referred for more details. Although the field has progressed tremendously, there still remains some unanswered questions about classical perturbed deterministic equations, which we aim to address in this thesis. The underlying resolvent is crucial in studying (2.1.1). Its dynamics are given by,

$$\dot{r}(t) = \int_{[0,t]} \nu(ds) r(t-s), \quad t > 0; \quad r(0) = 1. \tag{2.1.2}$$

As shown in [53, Theorem 3.9] it is necessary to assume $r \in L^1(\mathbb{R}_+;\mathbb{R})$ in order to obtain strong admissibility results, thus in this chapter we also employ such an assumption. Recall this is equivalent to $z - \hat{\nu}(z) \neq 0$ for $\mathrm{Re}(z) \geq 0$ where $z \in \mathbb{C}$ and $\hat{\nu}$ denotes the Laplace transform of the measure ν [53, Theorem 3.3.5]. Although this condition on the Laplace transform characterises when our key assumption is true, it can be difficult to check unless one imposes specific structure on the kernel. For example the case wherein $\nu(ds) = a\delta(ds) + b(s)ds$ has received due study [65] (here $\delta(ds)$ is a Dirac measure). In this special case the finite measure condition is exactly $b \in L^1(\mathbb{R}_+;\mathbb{R})$. Assume b is continuous, of one sign and

$$\int_0^\infty t\,|b(t)|dt < +\infty.$$

If $\int_0^\infty b(t)dt > 0$ and $a + \int_0^\infty b(t)dt < 0$ then $z - a - \int_0^\infty e^{-zt}b(t)dt \neq 0$ for $\mathrm{Re}(z) \geq 0$ where $z \in \mathbb{C}$ and thus the resolvent is integrable as required [65, Theorem A]. In order for the zeros of $z - \hat{\nu}(z)$ to lie in the left half complex plane one needs some negativity from the measure ν. For example in [65] the author also proves if $a + \int_0^\infty b(t)dt > 0$ then there exists a $z \in \mathbb{C}$ such that $z - a - \int_0^\infty e^{-zt}b(t)dt = 0$ and $\mathrm{Re}(z) \geq 0$.

If we consider the unperturbed version of (2.1.1), i.e with $f = 0$, the solution is simply $x(t) = r(t)\xi$, hence the assumption that $r \in L^1(\mathbb{R}_+;\mathbb{R})$ endows the solution with many desirable properties. For example solutions will vanish at infinity, be bounded, $p-$integrable ($p \geq 1$), absolutely continuous etc.. A larger list of properties the solution will possess is collected in Gripenberg, Londen and Staffons [53, Theorem 3.9]. In this thesis we will often use the implication $r \in L^1(\mathbb{R}_+;\mathbb{R}) \implies r \in BC_0(\mathbb{R}_+;\mathbb{R}), L^p(\mathbb{R}_+;\mathbb{R})$, a proof of this fact is provided at the end of Section 2.2. It is very natural to ask, can we characterise when such properties are *preserved* after a perturbation is introduced, or in other words can we give necessary and sufficient conditions on the forcing function f to ensure the solution of equation (2.1.1) is an element of a particular function space V. Surprisingly the classical theory does not provide an answer to this question. However this topic is deeply related to the question of so-called *admissibilty* of (Volterra) operators and of pairs of function spaces, see Pulyaev and Tsalyuk [102, 103, 119]. For perturbed Volterra equations, sufficient conditions for solutions to lie in a space V, when forcing functions lie in a space W have been extensively discussed in Miller [93, Ch. 5] and Corduneanu [37, Ch.2].

A precise answer to our question shall be provided in this chapter. For a given function space V the key for obtaining such a result is a decomposition lemma[1] which states (for some locally integrable function f) if $\int_{\cdot}^{\cdot+\theta} f(s)ds \in V$ for each $\theta \in (0,1]$ then $f = f_1 + f_2$ where $f_1 \in V$ and $\int_0^{\cdot} f_2(s)ds \in V$. Such a result was proven for $V = BC_0(\mathbb{R}_+;\mathbb{R})$ (space of bounded continuous functions that vanish at infinity) in [53, Lemma 15.9.2]. In this chapter this takes the form of Lemmas 2.3.2 and 2.4.2 which deal with the cases where $V = L^p(\mathbb{R}_+;\mathbb{R})$ or $V = \text{Ces}(\mathbb{R}_+;\mathbb{R})$ respectively (see Definition 2.2.1). It is the author's opinion that the results provided in this chapter are by no means restricted to the particular spaces listed above and that the framework developed can be extended to a wide variety of function spaces including, but not restricted to, those spaces found in [53, Theorem 3.9]. It is in this regard that this chapter should be understood as strong evidence that the methods developed can be used to characterise non-standard behaviour in the solution of equation (2.1.1).

The key point outlined throughout this chapter is that if one wishes to obtain such characterisation results for solutions to (2.1.1), it is not the pointwise behaviour of the function f that matters but rather its integral over compact intervals:

$$\int_t^{t+\theta} f(s)ds \quad \text{for } \theta \in (0,1]. \tag{2.1.3}$$

To the best of our knowledge consideration of this quantity in the context of the asymptotic behaviour of perturbed dynamical systems, first appeared in papers of Strauss and Yorke [117, 116]. In many situations it is the case that the forcing function may not be well behaved in a pointwise sense but the smoothing effect of integration drastically improves the situation. For instance in section 3.5 of Chapter 3 we provide an example of a continuous function, in which $\limsup_{t\to\infty} f(t) = +\infty$, and yet,

$$\lim_{t\to\infty} \int_t^{t+\theta} f(s)ds = 0 \text{ for each } \theta \in (0,1].$$

We show that given such a perturbation function f, solutions of (2.1.1) obey $x(t) \to 0$ as $t \to \infty$, despite the fact that the perturbation function has a diverging $\limsup$. In fact a characterisation of when $x \in BC_0([0,\infty),\mathbb{R})$ has already been indirectly provided by the literature. The result is as follows: under the assumption that x solves (2.1.1) and $r \in L^1(\mathbb{R}_+)$ then the following are equivalent:

(A) $x(t;\xi) \in BC_0([0,\infty),\mathbb{R})$ for all $\xi \in \mathbb{R}$,

(B) $t \mapsto \int_t^{t+\theta} f(s)ds \in BC_0([0,\infty),\mathbb{R})$ for each $\theta \in (0,1]$.

[1] We also require the operation of convolution with L^1 functions to map V into itself and for $r \in V$, but these conditions are known from the classical theory.

That (**B**) $\implies$ (**A**) is a consequence of Theorem 11.4.3 in [53] and (**A**) $\implies$ (**B**) is an obvious consequence of equation (2.3.4) below.

The theory for Volterra integro-differential equations is intricately linked to the theory of Volterra integral equations, however we make the observation that this phenomenon cannot occur with integral equations. By integral equation we refer to the following:

$$x(t) = \int_0^t k(t-s)x(s)ds + f(t), \quad t \geq 0. \tag{2.1.4}$$

We highlight that the classical theory for (2.1.4) is in fact sharp and in this case it is indeed the pointwise behaviour of the perturbation term that governs rather then the interval average (2.1.3). Thus if one considered the integral equation (2.1.4) with $k \in L^1(\mathbb{R}_+;\mathbb{R})$ and a perturbation function f as described above, the solution of (2.1.4) would have a diverging $\limsup$ while the solution of the integro-differential equation (with the same kernel k) would converge to zero. We also remark that this "*robustness*" of integro-differential equations is also a phenomenon found in so called functional differential equations wherein the system has only a finite memory.

Our first main result is a characterisation of when solutions to the Volterra Integro-differential equation (2.1.1) are p-integrable for $p \geq 1$. We show that solutions of (2.1.1) are in $L^p(\mathbb{R}_+;\mathbb{R})$ if and only if

$$\left\| \int_{\cdot}^{\cdot+\theta} f(s)ds \right\|_{L^p(\mathbb{R}_+)} < +\infty, \quad \text{for each } \theta \in (0,1]. \tag{2.1.5}$$

This "interval average" condition greatly expands the class of perturbing functions which preserve the property of p-integrability. An explicit example is given of a function exhibiting unbounded oscillatory behaviour which satisfies (2.1.5) but is not in $L^p(\mathbb{R}_+;\mathbb{R})$ for any $p \geq 1$. We once again illustrate this "*robustness*" phenomenon of solutions to (2.1.1) by selecting a non standard, yet interesting type of convergence.

We are concerned with when solutions of (2.1.1) and (2.1.4) admit a Cesàro type limit. Assuming $r \in L^1(\mathbb{R}_+,\mathbb{R})$, we show

$$\frac{1}{t}\int_0^t \int_s^{s+\theta} f(u)du\,ds \longrightarrow C\theta \text{ as } t \to \infty \quad \text{for each } \theta \in (0,1], \tag{2.1.6}$$

is equivalent to

$$\frac{1}{t}\int_0^t x(s)ds \longrightarrow \frac{-C}{\nu(\mathbb{R}_+)} \in \mathbb{R} \ \text{ as } \ t \to \infty, \tag{2.1.7}$$

where $C \in \mathbb{R}$ and x is the solution of (2.1.1). Additionally we show that if you alter condition (2.1.6) by imposing that f itself must posses a Cesàro limit then this imposes necessary restrictions on the Cesàro limit of the derivative of x. We show,

$$\frac{1}{t}\int_0^t f(s)ds \longrightarrow C \in \mathbb{R} \text{ as } t \to \infty \tag{2.1.8}$$

is equivalent to

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t x(s)ds = \frac{-C}{\nu(\mathbb{R}_+)}; \quad \lim_{t\to\infty}\frac{1}{t}\int_0^t \dot{x}(s)ds = 0. \tag{2.1.9}$$

In the latter case we note the condition on the Cesàro limit of the derivative cannot be dropped. This highlights the usual situation that conditions imposed directly on the forcing function f are in general *sufficient* to obtain information about the solution, but not *necessary*. However this is once again not the case for solutions to the integral equation (2.1.4), in fact condition (2.1.8) is equivalent to solutions of (2.1.4) having a Cesàro limit which in this case is equal to $C(1 - 1/\nu(\mathbb{R}_+))$. To the best of our knowledge there has been no analysis of the time average of solutions to (2.1.1) however we mention analogous equations in discrete time have received some attention [17, 18]. Despite this apparent lack of literature there has been much work in the context of admissibility theory for so called Marcinkiewicz spaces which are intimately linked to Cesàro limits of functions. For a brief modern introduction to such spaces we recommend the book by Corduneanu [39, pp.41-48] or for a more comprehensive study [27], while for classical admissibility results regarding integral equations we mention [21].

2.2 Mathematical Preliminaries

Let $p \in [1, \infty)$ and $f : \mathbb{R}_+ \to \mathbb{R}$. We say $f \in L^p(\mathbb{R}_+;\mathbb{R})$ if $\int_{\mathbb{R}_+} |f(s)|^p ds < +\infty$ and $f \in L^p_{loc}(\mathbb{R}_+;\mathbb{R})$ if $\int_K |f(s)|^p ds < +\infty$, where $K \subset \mathbb{R}_+$ is compact and both integrals are understood in the Lebesgue sense. If $M(\mathbb{R}_+;\mathbb{R})$ is the space of finite signed Borel measures on $\mathbb{R}_+$, and $\nu \in M(\mathbb{R}_+;\mathbb{R})$, we consider the halfline Volterra equation given by

$$\dot{x}(t) = \int_{[0,t]} x(t-s)\nu(ds) + f(t), \quad t \geq 0; \quad x(0) = \xi \in \mathbb{R}. \tag{2.2.1}$$

We say x is a solution of (2.2.1) on an interval $(0, T]$ whenever x is locally absolutely continuous, satisfies the initial condition and obeys the dynamics in (2.2.1) for almost all $t \in (0, T]$. We stipulate throughout, and *without further reference*, that $f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R})$, which guarantees the existence of a unique solution x (see Gripenberg et al.[53, Theorem 3.3.3]). In particular, the solution satisfies a variation of constants formula

$$x(t) = r(t)\xi + \int_0^t r(t-s)f(s)ds, \quad t \geq 0, \tag{2.2.2}$$

where r is the so–called differential resolvent of ν, which is the unique absolutely continuous function from $\mathbb{R}_+$ to $\mathbb{R}$, satisfying

$$\dot{r}(t) = \int_{[0,t]} r(t-s)\nu(ds), \quad t > 0; \quad r(0) = 1. \tag{2.2.3}$$

If μ is a finite measure on $[0,\infty)$, and $f : [0,\infty) \to \mathbb{R}$ then their convolution is

$$(f * \mu)(t) := \int_{[0,t]} f(t-s)\mu(ds), \quad t \geq 0.$$

The convolution of two functions is defined analogously. Given a measure $\mu \in M(\mathbb{R}_+;\mathbb{R})$, we denote $|\mu|$ as the set function which takes each Borel set $E \subseteq \mathbb{R}_+$ and assigns the total variation of μ on E. Recall the total variation of a measure μ on a set E is given by $|\mu|(E) := \sup_\pi \sum_{j=1}^n |\mu(E_j)|$ where π denotes the collection of finite partitions of the set E. For more details on these definitions see [53, Section 3.5]. With a slight abuse of terminology we shall call $|\mu|$ the total variation measure of μ. The following standard estimate [53, Theorem 3.5.6] will be heavily utilised:

$$|(f * \mu)(t)| \leq \int_{[0,t]} |f(t-s)||\mu|(ds).$$

The reader should keep in mind that all results presented for solutions of equation (2.2.1) can be emulated for the finite memory problem, which is a type of functional differential equation. In this instance we would fix a constant $\tau > 0$ and consider

$$\dot{x}(t) = \int_{[-\tau,0]} x(t+u)\mu(du) + f(t), \quad t \geq 0; \quad x(t) = \psi(t), \quad t \leq 0 \tag{2.2.4}$$

where $\mu \in M([-\tau,0];\mathbb{R})$ and $\psi \in C([-\tau,0];\mathbb{R})$. With the following convention one can essentially use identical proofs for results regarding equations (2.2.1) and (2.2.4). For any subset of the real line, write $-E := \{x \in \mathbb{R} : -x \in E\}$. If $\mu \in M([-\tau,0];\mathbb{R})$, we can construct a $\tilde{\mu} \in M([0,\infty);\mathbb{R})$ by writing

$$\begin{aligned} &\tilde{\mu}(E) = \mu(-E), \quad \text{for any Borel set } E \subseteq [0,\tau], \\ &\tilde{\mu}(E) = 0, \quad \text{for any Borel set } E \text{ with } E \cap [0,\tau] = \emptyset. \end{aligned} \tag{2.2.5}$$

Let $g : \mathbb{R} \to \mathbb{R}$ be such that $g(t) = 0$ for all $t < 0$. With this construction we have

$$\int_{[-\tau,0]} g(t+s)\mu(ds) = \int_{[0,t]} g(t-s)\tilde{\mu}(ds) = (g * \tilde{\mu})(t), \quad t \geq 0.$$

With this notation the resolvent of equation (2.2.4) obeys,

$$\dot{r}_\tau(t) = \int_{[0,t]} \tilde{\mu}(ds) r_\tau(t-s), \quad t>0; \quad r_\tau(0)=1; \quad r_\tau(t)=0 \quad t<0. \tag{2.2.6}$$

Here we use the notation r and r_τ to distinguish between the resolvent of the Volterra equation (2.2.1) and the functional equation (2.2.4) respectively. If we define $F(t) := \int_{[-\tau,0]} \left(\int_s^0 r_\tau(t+s-u)\psi(u)du\right)\mu(ds)$, the solution of (2.2.4) has representation

$$x(t,\psi) = r_\tau(t)\psi(0) + F(t) + \int_0^t r_\tau(t-s)f(s)ds, \tag{2.2.7}$$

for $t \geq 0$. The assumption $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R})$ ensures that $F = O(e^{-\alpha t})$ for some $\alpha > 0$, this follows from Theorem I.5.4 and Corollary I.5.5 in [44] paired with Lemma 2.1 in [16], we postpone the discussion of this equivalence until Chapter 4 where it is heavily utilised. Thus this extra term which does not appear in the solution for the Volterra equation (2.2.2) will not have any contribution towards the Cesàro mean, for which we now provide a precise definition.

Definition 2.2.1. *We define the space of Cesàro functions as follows:*

$$\mathrm{Ces}(\mathbb{R}_+;\mathbb{R}) := \left\{ f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}) : \lim_{t\to\infty} \frac{1}{t}\int_0^t f(s)ds = C \text{ for some } C \in \mathbb{R} \right\}. \tag{2.2.8}$$

Additionally we say f admits a Cesàro limit if $f \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$.

Remark 2.2.1. *Clearly all integrable functions admit a trivial Cesàro limit. It is in this regard that the extra term (F) appearing in (2.2.7) does not provide any additional difficulties. To see this consider $y(t) = x(t,\psi) - F(t)$, then y is the solution of the Volterra integro-differential equation,*

$$\dot{y}(t) = \int_{[0,t]} y(t-s)\tilde{\mu}(ds) + f(t), \quad t \geq 0; \quad y(0) = \psi(0).$$

Thus we can apply the theory developed for equation (2.2.1) and noting as $F(t)$ is integrable, the Cesàro limits of $y(t)$ and $x(t,\psi)$ agree.

As mentioned we shall always employ the assumption that the resolvent is integrable which implies the resolvent is p-integrable for any $p \geq 1$ and vanishes at infinity. We state a lemma to this effect in the Volterra case, for functional equations the result is given in [16].

Lemma 2.2.1. *Let r be the solution of equation 2.2.3 and assume $r \in L^1(\mathbb{R}_+;\mathbb{R})$. Then $r \in BC_0(\mathbb{R}_+;\mathbb{R}), L^p(\mathbb{R}_+;\mathbb{R})$ for any $p \geq 1$.*

Proof of Lemma 2.2.1. Recall $\dot{r}(t) = (\nu * r)(t)$. As $\mu \in M(\mathbb{R}_+;\mathbb{R})$ this implies $\dot{r} \in L^1(\mathbb{R}_+;\mathbb{R})$. Using the absolute continuity of r on compacts we can write

$$r(T) = 1 + \int_0^T \dot{r}(t)dt. \tag{2.2.9}$$

Now $\dot{r} \in L^1(\mathbb{R}_+;\mathbb{R}) \implies \lim_{T\to\infty} r(T) = L$ exists, but as $r \in L^1(\mathbb{R};\mathbb{R})$ (by assumption) we must have $L = 0$. Hence $r \in BC_0(\mathbb{R}_+;\mathbb{R})$ is proven. Next fix $k > 0$ such that $|r(t)| < 1$ for all $t \geq k$. Now for $p \geq 1$ and $T > k$,

$$\begin{aligned}\int_0^T |r(t)|^p dt &= \int_0^k |r(t)|^p dt + \int_k^T |r(t)|^p dt \\ &\leq \int_0^k |r(t)|^p dt + \int_k^T |r(t)| dt.\end{aligned}$$

Sending $T \to \infty$ and using the fact $r \in L^1(\mathbb{R}_+;\mathbb{R})$, yields $r \in L^p(\mathbb{R}_+;\mathbb{R})$ as required. □

The exact value of the integral of r and r_τ are needed in many subsequent proofs so we include this as a lemma.

Lemma 2.2.2. *Let r and r_τ be the solutions to equations (2.2.3) and (2.2.6) respectively and assume $r, r_\tau \in L^1(\mathbb{R}_+;\mathbb{R})$. Then,*

$$\begin{aligned}\int_0^\infty r(s)ds &= \frac{-1}{\nu(\mathbb{R}_+)}; \quad \int_0^\infty \dot{r}(s)ds = -1; \\ \int_0^\infty r_\tau(s)ds &= \frac{-1}{\mu([-\tau,0])}; \quad \int_0^\infty \dot{r}_\tau(s)ds = -1.\end{aligned}$$

Proof of Lemma 2.2.2. We only provide the proof for r as if one employs the convention outlined by (2.2.5) then the result for r_τ follows immediately. We note the assumption $r \in L^1(\mathbb{R}_+;\mathbb{R})$ guarantees $r(t) \to 0$ as $t \to \infty$. Thus integrating equation (2.2.3) over the interval $[0,T]$ we see, $r(T) = 1 + \int_0^T \int_{[0,t]} r(t-s)\nu(ds)dt$. Sending $T \to \infty$ and applying Fubini's theorem we obtain,

$$-1 = \int_0^\infty \int_{[0,t]} r(t-s)\nu(ds)dt = \int_{[0,\infty]} \int_s^\infty r(t-s)dt\,\nu(ds) = \int_0^\infty r(s)ds\nu(\mathbb{R}_+).$$

To show the second assertion we integrate (2.2.3) over $\mathbb{R}_+$ and apply the above calculation, i.e $\int_0^\infty \dot{r}(t)dt = \int_0^\infty \int_{[0,t]} r(t-s)\nu(ds)dt = -1$. □

2.3 Results for $V = L^p$

Before we present the main result of this section we prepare two lemmas, the first of which is required in the proof of the second.

Lemma 2.3.1. *Suppose* $\int_{\cdot}^{\cdot+\theta} f(s)ds \in L^p(\mathbb{R}_+;\mathbb{R})$ *for all* $\theta \in (0,1]$. *Then,*

$$\left\| \int_{\cdot}^{\cdot+\theta} f(s)ds \right\|_{L^p(\mathbb{R}_+)} < B, \qquad \text{for all } \theta \in (0,1]. \tag{2.3.1}$$

Note that the upper bound B is independent of the parameter θ. This is required if one wishes to bound $|\int_{\cdot}^{\cdot+\theta} f(s)ds|^p$ while integrating over θ. Before proving Lemma 2.3.1 we remark that its reverse implication is obviously true and will be used in proofs of later results.

Proof of Lemma 2.3.1. By hypothesis, for all $\theta \in (0,1]$ we have that,

$$\varphi(\theta) := \int_0^\infty \left| \int_t^{t+\theta} f(s)ds \right|^p dt < +\infty.$$

Moreover, since f is locally integrable, it follows that the integrand is indeed a well–defined measurable function and moreover the non-negativity ensures that φ is also measurable; this holds by e.g. Theorem 8.8 in Rudin [109]. Thus by hypothesis for each $m \in \mathbb{N}$, the set $Q_m = \{\theta \in [0,1] : \varphi(\theta) \le m\}$ is well–defined and measurable. Notice also that $Q_m \subseteq Q_{m+1}$ and that the hypothesis ensures that $\bigcup_{m=1}^\infty Q_m = [0,1]$. This ensures that there exists at least one $m' \in \mathbb{N}$ such that the set $Q_{m'}$ has positive Lebesgue measure. In light of this we fix $m = m'$. Then applying Lemma 15.9.3 in Gripenberg et al. [53] to the set Q_m ensures the set $Q_m - Q_m := \{\theta - \theta' : \theta, \theta' \in Q_m\}$, contains an interval $(-\epsilon, \epsilon)$. As both θ,$\theta' \in Q_m$ we have that $\varphi(\theta) \le m$ and $\varphi(\theta') \le m$. Thus without loss of generality we take $\theta > \theta'$ and by extending f to be equal to zero on the interval $[-1, 0)$ we can say for $t \ge 0$,

$$\left| \int_t^{t+\theta-\theta'} f(s)ds \right| \le \left| \int_{t-\theta'}^{t+\theta-\theta'} f(s)ds \right| + \left| \int_{t-\theta'}^{t} f(s)ds \right|.$$

Using the inequality $(a+b)^p \le 2^{p-1}(a^p+b^p)$ for $a, b \ge 0$, integrating both sides and letting $g(t) := |\int_0^t f(s)ds|^p$ gives

$$\begin{aligned}
\frac{1}{2^{p-1}} \int_0^\infty \left| \int_t^{t+\theta-\theta'} f(s)ds \right|^p dt &\le \int_0^\infty \left| \int_{t-\theta'}^{t+\theta-\theta'} f(s)ds \right|^p dt + \int_0^\infty \left| \int_{t-\theta'}^{t} f(s)ds \right|^p dt \\
&= \int_{-\theta'}^0 g(\tau+\theta)d\tau + \varphi(\theta) + \int_{-\theta'}^0 g(\tau+\theta')d\tau + \varphi(\theta') \\
&\le \sup_{-\theta' \le \tau \le 0} g(\tau+\theta) + \sup_{-\theta' \le \tau \le 0} g(\tau+\theta') + 2m \\
&\le 2\left(\sup_{0 \le t \le 1} g(t) + m \right).
\end{aligned}$$

Hence, for all $T \in (-\epsilon, \epsilon)$ we have

$$\int_0^\infty \left| \int_t^{t+T} f(s)ds \right|^p dt \leq 2^p \left(\sup_{0 \leq t \leq 1} \left| \int_0^t f(s)ds \right|^p + m \right) =: B_1.$$

Note that B_1 is independent of T, and take any $\theta \in (0,1]$. There is a minimal $N = N(\epsilon) \in \mathbb{N}$ such that $N\epsilon \geq 2$, and $\theta \in (n\frac{\epsilon}{2}, (n+1)\frac{\epsilon}{2}]$ for exactly one $n \in \{0, \ldots, (N-1)\}$. Write next

$$\int_t^{t+\theta} f(s)ds = \sum_{j=1}^N \int_{t+\frac{(j-1)\theta}{N}}^{t+\frac{j\theta}{N}} f(s)ds = \sum_{j=1}^N \int_{t+(j-1)\theta^*}^{t+j\theta^*} f(s)ds,$$

where $\theta^* := \frac{\theta}{N} \leq \frac{n+1}{2N}\epsilon \leq \frac{\epsilon}{2} < \epsilon$. Thus

$$\left| \int_t^{t+\theta} f(s)ds \right|^p \leq N^{p-1} \sum_{j=1}^N \left| \int_{t+(j-1)\theta^*}^{t+(j-1)\theta^*+\theta^*} f(s)ds \right|^p,$$

and since $\theta^* \in (0,\epsilon)$, for any $\theta \in (0,1]$ we have

$$\int_0^\infty \left| \int_t^{t+\theta} f(s)ds \right|^p dt \leq N \cdot B_1 =: B_p.$$

Since B_p depends only on N and B_1, and both are θ–independent, the proof is complete. $\square$

Both Lemmas 2.3.1 and 2.3.2 are adapted from a decomposition result for a function f in which $\int_t^{t+\theta} f(s)ds \to 0$ as $t \to \infty$ for all $\theta \in (0,1]$. This result can be found as Lemma 15.9.2 in [53].

Lemma 2.3.2. *Suppose* $\| \int_\cdot^{\cdot+\theta} f(s)ds \|_{L^p(\mathbb{R}_+)} < +\infty$ *for all* $\theta \in (0,1]$. *Then* $f = f_1 + f_2$ *where* $f_1 \in L^p(\mathbb{R}_+;\mathbb{R})$ *and* f_2 *is such that* $\int_0^t f_2(s)ds \in L^p(\mathbb{R}_+;\mathbb{R})$

Proof of Lemma 2.3.2. Take $f_1(t) = \int_{t-1}^t f(s)ds$, $t \geq 0$. Here we consider an extended version of f where $f(t) \equiv 0$, for all $t < 0$. Then f_1 is continuous and by supposition $f_1 \in L^p(\mathbb{R}_+;\mathbb{R})$. Next we let $f_2 = f - f_1$ for $t \geq 0$. Then for $t \geq 1$ we have

$$\int_0^t f_2(s)ds = \int_0^t f(s)ds - \int_0^t \int_{s-1}^s f(u)duds = \int_0^1 \int_{t+v-1}^t f(s)dsdv, \tag{2.3.2}$$

the last identity following from a routine but tedious set of calculations, involving repeated application of Fubini's theorem which are detailed in the appendix. By Jensen's inequality, we get

$$\left| \int_0^t f_2(s)ds \right|^p \leq \int_0^1 \left| \int_{t+v-1}^t f(u)du \right|^p dv.$$

Integrating over t on both sides and using Fubini's theorem, it follows that,

$$\int_1^\infty \left| \int_0^t f_2(s)ds \right|^p dt \le \int_0^1 \left(\int_0^\infty \left| \int_t^{t+v} f(u)du \right|^p dt \right) dv \le \int_0^1 B dv = B,$$

where the constant B is taken from Lemma 2.3.1. If $t \in [0,1]$ then the fact that f_2 is locally bounded ensures $\int_0^1 \left| \int_0^t f_2(s)ds \right|^p dt$ is finite. □

Remark 2.3.1. *Note the decomposition of f in Lemma 2.3.2 is certainly not unique. One could define $f_1 = \int_{t-\theta}^t f(s)ds$ for any $\theta \in (0,1]$ resulting in an infinite number of potential decompositions.*

We are now in a position to state and prove our main result. By Lemma 2.3.1, the condition (**A**) below is equivalent to (2.3.1), but the latter condition is harder to check than (**A**).

Theorem 2.3.1. *Let x be the solution of (2.2.1) and suppose $r \in L^1(\mathbb{R}_+;\mathbb{R})$. Then for $p \ge 1$, the following are equivalent:*

*(**A**)* $\left\| \int_\cdot^{\cdot+\theta} f(s)ds \right\|_{L^p(\mathbb{R}_+;\mathbb{R})} < +\infty,$ *for each $\theta \in (0,1]$;*

*(**B**) $x(\cdot,\xi) \in L^p(\mathbb{R}_+;\mathbb{R})$ for all $\xi \in \mathbb{R}$.*

Proof of Theorem 2.3.1. First we show (**A**) $\implies$ (**B**). Recall that x obeys a variation of constants formula given by $x(t,\xi) = r(t)\xi + (r * f)(t)$ for $t \ge 0$, where $*$ denotes convolution of functions on $\mathbb{R}_+$. As $\nu \in M(\mathbb{R}_+;\mathbb{R})$ and $r \in L^1(\mathbb{R}_+;\mathbb{R})$ we have that $r' \in L^1(\mathbb{R}_+;\mathbb{R})$ and also $r(t) \to 0$, as $t \to \infty$. These facts, along with r being absolutely continuous and well–behaved at both zero and infinity, ensure that $r \in L^p(\mathbb{R}_+;\mathbb{R})$ for all $p \ge 1$. Thus $r(\cdot)\xi \in L^p(\mathbb{R}_+;\mathbb{R})$, so we need only focus on $x_1(t) := (r * f)(t)$. Defining $f_3(t) := \int_0^t f_2(s)ds$, writing $f = f_1 + f_2$ and integrating by parts we get

$$x_1(t) = (r * f_1)(t) + f_3(t) + (r' * f_3)(t), \quad t \ge 0. \tag{2.3.3}$$

As $r, r' \in L^1(\mathbb{R}_+)$ and $f_1, f_3 \in L^p(\mathbb{R}_+;\mathbb{R})$ (by Lemma 2.3.2 and (**A**)), an application of Theorem 2.2.2 in [53] ensures $x_1 \in L^p(\mathbb{R}_+;\mathbb{R})$ and the claim is proven.
To show (**B**) $\implies$ (**A**), take $\theta \in (0,1]$ and integrate (2.2.1) over the interval $[t, t+\theta]$ to get

$$x(t+\theta) - x(t) - \int_t^{t+\theta} (\nu * x)(s)ds = \int_t^{t+\theta} f(s)ds. \tag{2.3.4}$$

The first two terms on the left are in $L^p(\mathbb{R}_+;\mathbb{R})$. For the third, with $\frac{1}{p} + \frac{1}{q} = 1$, by Hölder's inequality we get

$$\int_0^\infty \left| \int_t^{t+\theta} (\nu * x)(s) ds \right|^p dt \leq \theta^{\frac{p}{q}} \int_0^\infty \int_t^{t+\theta} |(\nu * x)(s)|^p \, dsdt$$
$$\leq \int_0^\infty \int_{\max(s-\theta,0)}^s |(\nu * x)(s)|^p \, dtds \leq \int_0^\infty |(\nu * x)(s)|^p \, ds.$$

The last term is finite by virtue of Theorem 3.6.1 in Gripenberg et al. [53]. Thus, the left-hand side of (2.3.4) is in $L^p(\mathbb{R}_+;\mathbb{R})$, and therefore the righthand side is also, as was required. $\square$

The three key identities in extending this result to other spaces are (2.3.2), (2.3.3) and for converse results, (2.3.4). These identities hold, regardless of conditions on f, and show how the solution x can be related to f purely through the interval average $\int_\cdot^{\cdot+\theta} f(s)ds$. The reader should observe that the results in this section can easily be extended to multi-dimensional equations; this arises as a special case in Chapter 3. It is worth noting however that a much stronger result is true for functional equations, in this case the resolvent being integrable is no longer an assumption but is in fact a necessary condition. For completeness we state the precise Theorem below but postpone the proof until Chapter 3.

Theorem 2.3.2. *Let x be the solution of (2.2.4). Then for $p \geq 1$, the following are equivalent:*

*(**A**) $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R})$ and $\left\| \int_\cdot^{\cdot+\theta} f(s)ds \right\|_{L^p(\mathbb{R}_+;\mathbb{R})} < +\infty$, for each $\theta \in (0,1]$;*

*(**B**) $x(\cdot,\psi) \in L^p([-\tau,\infty);\mathbb{R})$ for each $\psi \in C([-\tau,0];\mathbb{R})$.*

2.3.1 Example

We consider the equation (2.2.1) with

$$f(t) = e^{\alpha t} \sin(e^{\beta t}), \quad t \geq 0, \tag{2.3.5}$$

where $0 < \alpha < \beta$. In this case f exhibits high frequency and rapidly growing oscillations. Write $T = e^{\beta t}$, $A = e^{\beta\theta}$. Note that $\epsilon := 1 - \alpha/\beta \in (0,1)$. Then

$$\int_t^{t+\theta} f(s) \, ds = \frac{1}{\beta} \int_T^{AT} u^{-\epsilon} \sin(u) \, du.$$

Integrating the right hand side by parts we see that it is $O(T^{-\epsilon})$ as $T \to \infty$, where we use the conventional "big O" Landau notation. Therefore

$$\int_t^{t+\theta} f(s) \, ds = O(e^{-(\beta-\alpha)t}), \quad t \to \infty,$$

for each $\theta > 0$. This exponential decay ensures $\int_{\cdot}^{\cdot+\theta} f(s)ds \in L^p(\mathbb{R}_+;\mathbb{R})$ for any choices of $p \geq 1$ and $\theta > 0$. Thus Theorem 2.3.1 tells us that $x \in L^p(\mathbb{R}_+;\mathbb{R})$ despite the fact that $\int_0^\infty |f(s)|^p ds = \infty$ for any choice of $p \geq 1$.

2.4 Results for $V = \textbf{Ces}(\mathbb{R}_+;\mathbb{R})$

2.4.1 Cesàro functions

In this section we discuss some facts about members of the function space $\text{Ces}(\mathbb{R}_+;\mathbb{R})$. The main results of this section rely on a decomposition of the forcing function f. The general result has the following flavour, let V represent some function space of interest. If $\int_{\cdot}^{\cdot+\theta} f(s)ds \in V$ for each $\theta \in (0,1]$ then $f = f_1 + f_2$ such that $f_1 \in V$ and $\int_0^{\cdot} f_2(s)ds \in V$. Successfully improving the classical admissibility theory is completely determined by whether or not one can prove such a decomposition result. This is accomplished by Lemmas 2.4.1 and 2.4.2 below for $V = \text{Ces}(\mathbb{R}_+;\mathbb{R})$.

Lemma 2.4.1. *Assume $f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R})$. Suppose for all $\theta \in (0,1]$, $t \mapsto \int_t^{t+\theta} f(s)ds \in$ $\text{Ces}(\mathbb{R}_+;\mathbb{R})$, i.e*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \int_s^{s+\theta} f(u)duds = C(\theta) \in \mathbb{R}.$$

Then for any fixed θ,

$$\sup_{\substack{t\in[1,\infty) \\ T\in[0,\theta]}} \left| \frac{1}{t}\int_0^t \int_s^{s+T} f(u)duds - C(T)\right| < \infty.$$

Proof of Lemma 2.4.1. Introduce for some $m \in \mathbb{N}$, the set,

$$Q_m := \left\{ T \in [0,\theta] \colon \left| \frac{1}{t}\int_0^t \int_s^{s+T} f(u)duds - C(T)\right| \leq 1, \text{ for all } t \geq m \right\}.$$

First we show this set is measurable. The mapping $t \mapsto \left| \frac{1}{t}\int_0^t \int_s^{s+T} f(u)duds - C(T)\right|$ is continuous thus,

$$\begin{aligned} Q_m &= \left\{ T \in [0,\theta] \colon \left| \frac{1}{t}\int_0^t \int_s^{s+T} f(u)duds - C(T)\right| \leq 1, \text{ for all } t \in \mathbb{Q}\cap[m,\infty) \right\} \\ &= \bigcap_{t\in\mathbb{Q}\cap[m,\infty)} \left\{ T \in [0,\theta] \colon \left| \frac{1}{t}\int_0^t \int_s^{s+T} f(u)duds - C(T)\right| \leq 1 \right\}. \end{aligned}$$

Now for each t introduce, $F_t : [0,\theta] \to \mathbb{R}_+$, $T \mapsto \left| \frac{1}{t}\int_0^t \int_s^{s+T} f(u)duds - C(T)\right|$. Then we can write,

$$Q_m = \bigcap_{t\in\mathbb{Q}\cap[m,\infty)} F_t(T)^{-1}([0,1]).$$

Thus to show Q_m is measurable we need only show that $F_t(T)$ is a measurable function, but this is immediate as $C(T)$ is the pointwise limit of measurable functions and thus measurable, and the mapping $T \mapsto \frac{1}{t}\int_0^t \int_s^{s+T} f(u)duds$ is continuous, thus $F_t(T)$ is a measurable function and so Q_m is a well defined measurable set. Additionally it follows from our supposition that $\bigcup_{m=1}^{\infty} Q_m = [0,\theta]$, thus there must exists an $m' \in \mathbb{N}$ such that $Q_{m'}$ has non zero Lebesgue measure. Fix such an m', then by Lemma 15.9.3 in Gripenberg et. al. [53], it must by the case that the set $Q_{m'} - Q_{m'} = \{T - T' : T, T' \in Q_{m'}\}$ contains an open interval $(-\varepsilon, \varepsilon)$. Without loss of generality take $T, T' \in Q_{m'}$ such that $T > T'$, then we have

$$
\begin{aligned}
&\Big|\frac{1}{t}\int_0^t \int_s^{s+T-T'} f(u)duds - C(T-T')\Big| \\
&\quad = \left|\frac{1}{t}\int_0^t \int_s^{s+T-T'} f(u)duds - C(T) + C(T')\right| \\
&\quad \leq \left|\frac{1}{t}\int_0^t \int_{s-T'}^{s+T-T'} f(u)duds - C(T)\right| + \left|\frac{1}{t}\int_0^t \int_{s-T'}^{s} f(u)duds - C(T')\right| \\
&\quad \leq 2
\end{aligned}
$$

for all $t \geq m' + \theta$. To see why the first equality holds consider

$$
\frac{1}{t}\int_0^t \int_s^{s+T-T'} f(u)duds = \frac{1}{t}\int_0^t \int_{s-T'}^{s+T-T'} f(u)duds - \frac{1}{t}\int_0^t \int_{s-T'}^{s} f(u)duds.
$$

Taking limits on both sides yields

$$
C(T-T') = C(T) - C(T'). \tag{2.4.1}
$$

Note one can extend this argument via induction to show the limit is an additive function of the length of the interval of integration in the inner integral i.e. $C\left(\sum_{j=1}^{n} t_j\right) = \sum_{j=1}^{n} C(t_j)$ for a finite positive sequence t_j. Now as the above bound holds for all $T \in Q_{m'} - Q_{m'}$, we have in particular for all $T \in [0, \varepsilon)$,

$$
\Big|\frac{1}{t}\int_0^t \int_s^{s+T} f(u)duds - C(T)\Big| \leq 2, \qquad \text{for all } t \geq m' + \theta.
$$

Let $T \in [0, \theta]$ and fix $N \in \mathbb{N}$ such that $\frac{T}{N} \leq \frac{\theta}{N} < \varepsilon$. Then using the additive property of the limit, we have

$$\begin{aligned}\left|\frac{1}{t}\int_0^t\int_s^{s+T} f(u)duds - C(T)\right| &= \left|\frac{1}{t}\int_0^t\sum_{j=1}^N\int_{s+\frac{(j-1)T}{N}}^{s+\frac{jT}{N}} f(u)duds - \sum_{j=1}^N C\left(\frac{jT}{N}-\frac{(j-1)T}{N}\right)\right|\\ &\leq \sum_{j=1}^N\left|\frac{1}{t}\int_0^t\int_{s+\frac{(j-1)T}{N}}^{s+\frac{jT}{N}} f(u)duds - C\left(\frac{jT}{N}-\frac{(j-1)T}{N}\right)\right|\\ &\leq 2N,\end{aligned}$$

for all $t \geq m' + \theta$ where the last inequality follows as $\frac{jT}{N} - \frac{(j-1)T}{N} = \frac{T}{N} \leq \frac{\theta}{N} < \varepsilon$. Thus as N is independent of both T and t, we have

$$\sup_{\substack{t\in[1,\infty)\\ T\in[0,\theta]}}\left|\frac{1}{t}\int_0^t\int_s^{s+T} f(u)duds - C(T)\right| < \infty.$$

□

Remark 2.4.1. *The identity (2.4.1) tells us that when the mapping $t \mapsto \int_t^{t+\theta} f(s)ds \in$ Ces$(\mathbb{R}_+;\mathbb{R})$, the limit is an additive function of the parameter θ. This along with measurability ensures the limit has the form $C\theta$ for some $C \in \mathbb{R}$, [28, Theorem 1.1.8]. In the sequel we shall always write the limit in this explicit form.*

Lemma 2.4.2. *Let $f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R})$. For each $\theta \in (0,1]$, the following are equivalent,*

(i)

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t\int_s^{s+\theta} f(u)duds = C\theta.$$

(ii) $f = f_1 + f_2$, where $f_1 \in C(\mathbb{R}_+;\mathbb{R})$ and $f_2 \in L^1_{loc}(\mathbb{R}_+;\mathbb{R})$ such that,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t f_1(u)du = C, \qquad \lim_{t\to\infty}\frac{1}{t}\int_0^t\int_0^s f_2(u)duds = \frac{C\theta}{2}.$$

Proof of Lemma 2.4.2. We extend f such that $f(t) = 0$ for $t < 0$. Fix a $\theta \in (0,1]$ and set $f_1(t) = \frac{1}{\theta}\int_{t-\theta}^t f(u)du$. Thus by hypothesis we immediately obtain the first limit while the local integrability of f ensures f_1 is continuous on the positive real half line. Next we set $f_2 = f - f_1$ and integrate to obtain:

$$\int_0^t f_2(s)ds = \int_0^t f(s)ds - \int_0^t\frac{1}{\theta}\int_{s-\theta}^s f(u)duds = \frac{1}{\theta}\int_0^\theta\int_{t-v}^t f(s)dsdv. \tag{2.4.2}$$

A proof of this identity is provided in the proof section at the end of this chapter. Now we calculate the Cesàro limit of $\int_0^t f_2(s)ds$.

$$\frac{1}{t}\int_0^t\int_0^s f_2(u)duds = \frac{1}{\theta}\int_0^\theta\frac{1}{t}\int_0^t\int_{s-v}^s f(u)dudsdv.$$

Subtracting $\frac{C}{\theta}\int_0^\theta v\,dv$ from both sides, taking absolute values and applying (2.4.2) yields,

$$\left|\frac{1}{t}\int_0^t\int_0^s f_2(u)duds - \frac{C}{\theta}\int_0^\theta vdv\right| \le \frac{1}{\theta}\int_0^\theta \left|\frac{1}{t}\int_0^t\int_{s-v}^s f(u)duds - vC\right| dv.$$

By Lemma 2.4.1, the integrand on the right hand side is uniformly bounded, thus we can invoke Arzelà's dominated convergence theorem and pass to the limit inside the integral. Hence,

$$\lim_{t\to\infty}\left|\frac{1}{t}\int_0^t\int_0^s f_2(u)duds - \frac{C}{\theta}\int_0^\theta v\,dv\right| = 0,$$

as required. For the converse, consider

$$\begin{aligned}\frac{1}{t}\int_0^t\int_{s-\theta}^s f(u)duds = \frac{1}{t}\int_0^t\int_{s-\theta}^s f_1(u)duds + \frac{1}{t}\int_0^t\int_0^s f_2(u)duds \\ - \frac{1}{t}\int_0^t\int_0^{s-\theta} f_2(u)duds.\end{aligned}$$

The two terms involving f_2 will cancel when we pass to the limit, so we need only focus on the f_1 term. Consider for some $k > 0$ and $t > k$,

$$\begin{aligned}\frac{1}{t}\int_0^t\int_{s-\theta}^s f_1(u)duds &= \frac{1}{t}\int_0^t\int_0^s f_1(u)duds - \frac{1}{t}\int_0^t\int_0^{s-\theta} f_1(u)duds \\ &= \frac{1}{t}\int_0^k\int_0^s f_1(u)duds + \frac{1}{t}\int_k^t\int_0^s f_1(u)duds \\ &\quad - \frac{1}{t}\int_0^k\int_0^{s-\theta} f_1(u)duds - \frac{1}{t}\int_k^t\int_0^{s-\theta} f_1(u)duds.\end{aligned}$$

The first and third terms will vanish as $t \to \infty$ so we focus on the second and fourth.

$$\begin{aligned}&\frac{1}{t}\int_k^t\int_0^s f_1(u)duds - \frac{1}{t}\int_k^t\int_0^{s-\theta} f_1(u)duds \\ &\qquad = \frac{1}{t}\int_k^t s\left[\frac{\int_0^s f_1(u)du}{s} - C\right]ds + \frac{C}{t}\int_k^t s\,ds \\ &\qquad\quad - \frac{1}{t}\int_k^t (s-\theta)\left[\frac{\int_0^{s-\theta} f_1(u)du}{s-\theta} - C\right]ds - \frac{C}{t}\int_k^t (s-\theta)\,ds.\end{aligned}$$

Combining the second and fourth integrals yields,

$$\theta C\frac{(t-k)}{t} \quad\longrightarrow\quad \theta C \quad \text{as } t\to\infty.$$

Thus to complete the proof we show the terms involving f_1 vanish. We have

$$\frac{1}{t}\int_k^t s\left[\frac{\int_0^s f_1(u)du}{s} - C\right]ds - \frac{1}{t}\int_k^t (s-\theta)\left[\frac{\int_0^{s-\theta} f_1(u)du}{s-\theta} - C\right]ds.$$

Letting $u = s - \theta$ in the second integral and combining the two yields,

$$\frac{1}{t}\int_{t-\theta}^t s\left[\frac{\int_0^s f_1(u)du}{s} - C\right]ds - \frac{1}{t}\int_{k-\theta}^k s\left[\frac{\int_0^s f_1(u)du}{s} - C\right]ds.$$

The second term will vanish and the first term can be estimated by,

$$\left|\frac{1}{t}\int_{t-\theta}^t s\left[\frac{\int_0^s f_1(u)du}{s} - C\right]ds\right| \leq \int_{t-\theta}^t \left|\frac{\int_0^s f_1(u)du}{s} - C\right| ds.$$

The integrand on the right hand side converging to zero forces the entire integral to vanish and thus the theorem is proven. □

Remark 2.4.2. *One should be careful when applying identity (2.4.2). We defined $f_1 \coloneqq \frac{1}{\theta}\int_{t-\theta}^t f(s)ds$ with an extended version of f where $f(t) = 0$ for $t < 0$. This is equivalent to defining f_1 piecewise, i.e $f_1(t) = \frac{1}{\theta}\int_0^t f(s)ds$ for $t \in [0,\theta]$ and $f_1(t) = \frac{1}{\theta}\int_{t-\theta}^t f(s)ds$ for $t > \theta$. This of course leads to a piecewise definition of $f_2 \coloneqq f - f_1$, the reader should always keep this in mind as at first glance it may seem identity (2.4.2) is not correct for simple examples such as when $f(t) = c$ for $c \in \mathbb{R}\backslash\{0\}$. However with the consideration explained above one can calculate the explicit integrals and see the identity does indeed hold.*

With this fundamental decomposition at hand we state one more crucial lemma, whose proof is relegated to the final section of the chapter, before providing our main result

Lemma 2.4.3. *Let $\nu \in M(\mathbb{R}_+)$ and $g \in L^1(\mathbb{R}_+)$. Suppose $f \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ with Cesàro limit, $C \in \mathbb{R}$. Then,*

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t (f*\nu)(s)ds = C\nu(\mathbb{R}_+); \quad \lim_{t\to\infty}\frac{1}{t}\int_0^t (f*g)(s)ds = C\int_0^\infty g(s)ds.$$

2.4.2 Main Results

We can now state and prove the main result of this section.

Theorem 2.4.1. *Let x be the solution of equation (2.2.1) and assume $r \in L^1(\mathbb{R}_+;\mathbb{R})$. Then the following statements are true,*

(i) $\int_\cdot^{\cdot+\theta} f(s)ds \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ *for all* $\theta \in (0,1] \iff x \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$.

(ii) $f \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R}) \iff x \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ *and* $\dot{x} \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$.

Moreover in each case the limits are given explicitly, that is as $t \to \infty$,

(i)

$$\frac{1}{t}\int_0^t \int_s^{s+\theta} f(u)duds \longrightarrow C\theta; \qquad \frac{1}{t}\int_0^t x(s)ds \longrightarrow \frac{-C}{\nu(\mathbb{R}_+)}.$$

(ii)

$$\frac{1}{t}\int_0^t f(s)ds \longrightarrow C; \qquad \frac{1}{t}\int_0^t x(s)ds \longrightarrow \frac{-C}{\nu(\mathbb{R}_+)}; \qquad \frac{1}{t}\int_0^t \dot{x}(s)ds \longrightarrow 0.$$

where $C \in \mathbb{R}$.

Proof of Theorem 2.4.1. First we prove (i). Similarly to the proof of Theorem 2.3.1, using the variation of constants formula (2.2.2), Lemma 2.4.2 and integrating by parts,

$$x(t) = r(t)\xi + (r * f_1)(t) + f_3(t) + (\dot{r} * f_3)(t), \quad t \geq 0,$$

where $f_3(t) := \int_0^t f_2(s)ds$. Thus we have,

$$\frac{1}{t}\int_0^t x(s)ds = \frac{1}{t}\int_0^t r(s)\xi ds + \frac{1}{t}\int_0^t (r * f_1)(s)ds + \frac{1}{t}\int_0^t f_3(s)ds + \frac{1}{t}\int_0^t (\dot{r} * f_3)(s)ds.$$

Sending $t \to \infty$ and applying Lemmas 2.2.2, 2.4.2, and 2.4.3 we see,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t x(s)ds = \frac{-C}{\nu(\mathbb{R}_+)}.$$

For the converse we consider an extended solution so that $x(t) = 0$ for all $t < 0$, then we integrate (2.2.2) over the interval $[t-\theta, t]$ which results in,

$$\int_{t-\theta}^t f(u)du = x(t) - x(t-\theta) - \int_{t-\theta}^t \int_{[0,s]} x(s-u)\nu(du)ds.$$

Introducing the notation $X(t) := (x * \nu)(t)$ we obtain,

$$\frac{1}{t}\int_0^t \int_{s-\theta}^s f(u)duds = \frac{1}{t}\int_0^t x(s)ds - \frac{1}{t}\int_0^t x(s-\theta)ds - \frac{1}{t}\int_0^t \int_{s-\theta}^s X(u)duds.$$

When passing to the limit the first two terms on the right hand side will cancel, thus we need only focus on the third. Now recall by Lemma 2.4.3,

$$\frac{1}{t}\int_0^t X(s)ds \longrightarrow \frac{-C}{\nu(\mathbb{R}_+)}\nu(\mathbb{R}_+) = -C,$$

as $t \to \infty$. Thus we may follow the proof of the converse of Lemma 2.4.2 with X in place of f_1 to obtain,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t\int_{s-\theta}^s X(u)duds=-\theta C.$$

But this implies,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t\int_{s-\theta}^s f(u)duds=\theta C,$$

as required. For the forward implication of (ii), recall equation (2.2.2), $x(t)=r(t)\xi+(r*f)(t)$. As $r\in L^1(\mathbb{R}_+;\mathbb{R})$ its Cesàro limit will be zero, thus,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t x(s)ds=\lim_{t\to\infty}\frac{1}{t}\int_0^t(r*f)(s)ds=\frac{-C}{\nu(\mathbb{R}_+)},$$

where the last equality follows from Lemma 2.2.2 and 2.4.3. We now use this to show the Cesàro limit for $\dot{x}$. By (2.2.1) we obtain,

$$\frac{1}{t}\int_0^t\dot{x}(s)ds=\frac{1}{t}\int_0^t(x*\nu)(s)ds+\frac{1}{t}\int_0^t f(s)ds.$$

Passing to the limit and applying Lemma 2.4.3 and the limit just proven for x yields,

$$\frac{1}{t}\int_0^t\dot{x}(s)ds\to\frac{-C}{\nu(\mathbb{R}_+)}\nu(\mathbb{R}_+)+C=0,\qquad\text{as } t\to\infty.$$

For the converse, rearrange equation (2.2.1) to obtain $f(t)=\dot{x}(t)-\int_{[0,t]}x(t-s)\nu(ds)$. By supposition the Cesàro limit of $\dot{x}$ is zero and Lemma 2.4.3 tells us the Cesàro limit of the convolution is $\frac{C}{\nu(\mathbb{R}_+)}\nu(\mathbb{R}_+)=C$, as required. $\square$

The assumptions that $r,r_\tau\in L^1(\mathbb{R}_+;\mathbb{R})$ are known to be sharp in order to obtain admissibility results for solutions to equations (2.2.1) and (2.2.4). We provide a converse result for the finite memory problem which shows that $r_\tau\in L^1(\mathbb{R}_+;\mathbb{R})$ is a necessary condition in order for solutions of (2.2.4) to admit a Cesàro limit. Throughout the proof we make heavy use of the semi-explicit asymptotic expansion of the resolvent. This approach is much more involved in the Volterra case due to the lack of information about the general form of the Volterra resolvent. It is in this regard that we only consider the finite memory equation for such converse results.

Proposition 2.4.1. *Let x be the solution of (2.2.4). If for each $\psi\in C([-\tau,0];\mathbb{R})$ and $f\in L^1_{loc}(\mathbb{R}_+;\mathbb{R})$ such that $t\mapsto\int_t^{t+\theta}f(s)ds\in\text{Ces}(\mathbb{R}_+;\mathbb{R})$ for all $\theta\in(0,1]$, we have $x(\cdot,\psi)\in\text{Ces}(\mathbb{R}_+;\mathbb{R})$, then $r_\tau\in L^1(\mathbb{R}_+;\mathbb{R})$.*

For the readers convenience, before we provide a proof of Theorem 2.4.1 we recall some common notation in the field of functional equations. Define

$$x_0(t,\psi):=r_\tau(t)\psi(0)+\int_{[-\tau,0]}\left(\int_s^0 r_\tau(t+s-u)\psi(u)du\right)\mu(ds),\tag{2.4.3}$$

so we can rewrite (2.2.7) as $x(t,\psi) = x_0(t,\psi) + (r_\tau * f)(t)$. Thus x_0 solves equation (2.2.4) with the perturbation term switched off. The roots of the transcendental equation, $h(\lambda) = \lambda - \int_{[-\tau,0]} e^{-\lambda s}\mu(ds)$, govern the asymptotic behaviour of r_τ. Let $\Lambda = \{\lambda \in \mathbb{C} : h(\lambda) = 0\}$ and define $v_0(\mu) := \sup\{\mathrm{Re}(\lambda) : \lambda \in \Lambda\}$. The cardinality of $\Lambda' = \{\lambda \in \Lambda : \mathrm{Re}(\lambda) = v_0(\mu)\}$ is finite, thus r_τ admits the expansion,

$$e^{-v_0(\mu)t} r_\tau(t) = \sum_{\lambda_j \in \Lambda'} \{p_j(t)\cos(\beta_j t) + q_j(t)\sin(\beta_j t)\} + g(t), \tag{2.4.4}$$

where p_j and q_j are polynomials of degree $m_j - 1$ (where m_j is the multiplicity of λ_j), $\beta_j = \mathrm{Im}(\lambda_j)$ and $g(t) = o(e^{-\varepsilon t})$ for some $\varepsilon > 0$. For proofs of all facts recalled above the reader may consult [44, Chapter I].

Proof of Theorem 2.4.1. We aim to prove $v_0(\mu) < 0$ which yields automatically $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R})$ due to the expansion (2.4.4). We proceed by contradiction, thus assume $v_0(\mu) > 0$. Setting $f = 0$ yields $x_0(\cdot,\psi) \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ for all $\psi \in C([-\tau,0];\mathbb{R})$. Hence select $\lambda \in \Lambda'$ and set $\psi = \mathrm{Re}(e^{\lambda t})$ which yields the solution, $x_0(t;\psi) = \mathrm{Re}(e^{\lambda t})$ for all $t \in [-\tau,\infty)$. Let $\lambda = v_0(\mu) + i\beta$, then

$$\int_0^t \mathrm{Re}(e^{\lambda s})ds = \int_0^t e^{v_0(\mu)s}\cos(\beta s)ds = \frac{e^{v_0(\mu)t}\left(\beta\sin(v_0(\mu)t) + v_0(\mu)\cos(\beta t)\right)}{|\lambda|^2},$$

which upon dividing by t will not tend to a limit, contradicting $x_0(\cdot,\psi) \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$. Thus we must have $v_0(\mu) \le 0$. Now assume $v_0(\mu) = 0$, in which we distinguish two cases. First assume there exists a $\lambda \in \Lambda'$ with multiplicity strictly great than one. In this case we fix $\psi = t\cos(\mathrm{Im}(\lambda)t)$ which again yields a solution $x_0(t;\psi) = t\cos(\mathrm{Im}(\lambda)t)$ for all $t \in [-\tau,\infty)$, but as before a simple calculation shows

$$\frac{1}{t}\int_0^t s\cos(\mathrm{Im}(\lambda)s)ds = \frac{\sin(\mathrm{Im}(\lambda)t)}{\mathrm{Im}(\lambda)} + \frac{\cos(\mathrm{Im}(\lambda)t) - 1}{t\mathrm{Im}(\lambda)^2},$$

which oscillates indefinitely and so $x_0(\cdot;\psi) \notin \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$, a contradiction. Thus it must be when $v_0(\mu) = 0$, all elements of Λ' have multiplicity one. In this case,

$$r_\tau(t) = \sum_{j=1}^{n} \{c_j\cos(\beta_j t) + k_j\sin(\beta_j t)\} + g(t),$$

where $c_j, k_j \in \mathbb{R}\setminus\{0\}$. Now we fix the initial condition $\psi = 0$ which yields, $x(t,0) = (r_\tau * f)(t)$. Now assume first that $n > 1$ which means at least one zero is not at the origin. Without loss of generality we may assume $\beta_1 \neq 0$. Choose $f(t) = k_1\sin(\beta_1 t) - c_1\cos(\beta_1 t)$

and define $F(t) := \int_0^t f(s)ds$, we see $\int_0^t x(u,0)du = \int_0^t r_\tau(u)F(t-u)du$ with,

$$F(t) = \frac{c_1}{\beta_1}\cos(\beta_1 t) + \frac{k_1}{\beta_1}\sin(\beta_1 t) - \frac{c_1}{\beta_1}.$$

Now we can calculate $\int_0^t x(u,0)du$ semi-explicitly,

$$\begin{aligned}
&\frac{1}{t}\int_0^t r_\tau(u)F(t-u)du \\
&= \sin(\beta_1 t)\left(\frac{c_1 k_1}{\beta_1} + \frac{(c_1^s - k_1^2)}{2t\beta_1^2}\right) + \cos(\beta_1 t)\left(\frac{c_1 k_1}{t\beta_1^2} + \frac{(c_1^2 - k_1^2)}{2\beta_1}\right) - \frac{c_1 k_1}{t\beta_1^2} \\
&\quad + \frac{1}{t}\sum_{j=1}^{n}\{A_{1,j}\cos(\beta_j t) + A_{2,j}\sin(\beta_j t) + A_{3,j}\cos(\beta_1 t) + A_{4,j}\sin(\beta_1 t) + A_{5,j}\} \\
&\quad + \frac{1}{t}\int_0^t g(u)F(t-u)du,
\end{aligned}$$

where $A_{i,j} \in \mathbb{R}$. The summation term will vanish as $t \to \infty$ and the exponential estimate on g ensures the integral term will also vanish. However we see the first term will oscillate indefinitely as the constants c_1, k_1 and β_1 are non-zero. This means $x(\cdot;0) \notin \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ which is a contradiction. Now if n=1 there is only a single root on the imaginary axis. If this root is not at the origin then the above argument still holds, so we assume the sole root is at the origin. In this case $r_\tau(t) = c + g(t)$ where $c \in \mathbb{R}\backslash\{0\}$. Now set $\psi = 0$ and $f(t) = 1/c$, then $x(t,0) = t + (g * f)(t)$. As $g \in L^1(\mathbb{R}_+;\mathbb{R})$ and $f \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$, Lemma 2.4.3 implies the convolution has a Cesáro limit but the linear function $f(t) = t$ will not, which yields the final contradiction. Thus we must have that $v_0(\mu) < 0$ and the claim is proven. □

Proposition 2.4.1, Theorem 2.4.1 and remark 2.2.1, provide the proof of our next theorem.

Theorem 2.4.2. *Let $x(\cdot,\psi)$ be the solution to (2.2.4).*

*(**A**) Suppose $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R})$, then for all $\psi \in C([-\tau,0];\mathbb{R})$ the following hold true,*

(i) $\int_\cdot^{\cdot+\theta} f(s)ds \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ for all $\theta \in (0,1]$ $\iff$ $x(\cdot,\psi) \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$.

(ii) $f \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ $\iff$ $x(\cdot,\psi) \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ and $\dot{x}(\cdot,\psi) \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$.

Moreover in each case the limits are given explicitly, that is as $t \to \infty$,

(i)

$$\frac{1}{t}\int_0^t \int_s^{s+\theta} f(u)duds \to C\theta; \qquad \frac{1}{t}\int_0^t x(s,\psi)ds \to \frac{-C}{\mu([-\tau,0])}.$$

(ii)

$$\frac{1}{t}\int_0^t f(s)ds \to C; \quad \frac{1}{t}\int_0^t x(s,\psi)ds \to \frac{-C}{\mu([-\tau,0])}; \quad \frac{1}{t}\int_0^t \dot{x}(s,\psi)ds \to 0.$$

where $C \in \mathbb{R}$.

(B) *The following are equivalent,*

(i) $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R})$.

(ii) *For all* $\psi \in C([-\tau,0];\mathbb{R})$ *and* $f \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ *we have* $x(\cdot,\psi) \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$.

(iii) *For all* $\psi \in C([-\tau,0];\mathbb{R})$ *and* f *s.t.* $\int_\cdot^{\cdot+\theta} f(s)ds \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ *for all* $\theta \in (0,1]$ *we have* $x(\cdot,\psi) \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$.

2.4.3 Example

In this section we show the statements $f \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ and $t \mapsto \int_t^{t+\theta} f(s)ds \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ for each $\theta \in (0,1]$ are in fact not equivalent. In general the first implies the second. To see this recall that if $f \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ then by Theorem 2.4.1 part (ii), we necessarily have solutions of (2.2.1) obeying $x \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$. But then applying part (i) of Theorem 2.4.1 yields $t \mapsto \int_t^{t+\theta} f(s)ds \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ for each $\theta \in (0,1]$. This argument, although mathematically sound, is unsatisfactory due to its indirect nature. Indeed this can be proven in a direct fashion with an apt application of Lemma 2.4.3. Let $f \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$ and introduce the function $g(t) = \theta^{-1}\chi_{\{t\in[0,\theta]\}}(t)$. Then $g \in L^1(\mathbb{R}_+;\mathbb{R})$, hence by Lemma 2.4.3 $(f * g) \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$, but for $t > \theta$,

$$(f * g)(t) = \int_0^\theta g(s)f(t-s)ds = \frac{1}{\theta}\int_{t-\theta}^t f(s)ds.$$

Thus we have proven as claimed that,

$$f \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R}) \implies t \mapsto \int_t^{t+\theta} f(s)ds \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R}) \text{ for all } \theta \in (0,1]. \tag{2.4.5}$$

The next proposition illustrates that the reverse implication is not true.

Proposition 2.4.2. *Let* $\beta : \mathbb{R}_+ \to \mathbb{R}_+ \in C^2(\mathbb{R}_+;\mathbb{R})$ *and strictly increasing. Assume* $\beta(t), \beta'(t) \to \infty$ *as* $t \to \infty$. *Then* $f : \mathbb{R}_+ \to \mathbb{R}$, $t \mapsto \beta'(t)\sin(\beta(t))t$ *obeys,*

$$t \mapsto \int_t^{t+\theta} f(s)ds \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R}) \textit{ for all } \theta \in (0,1],$$

while $f \notin Ces(\mathbb{R}_+;\mathbb{R})$.

An explicit example of a function (up to a multiplicative constant) which satisfies the assumptions of proposition 2.4.2 is,

$$f : \mathbb{R}_+ \to \mathbb{R},\, t \mapsto t^\alpha \sin(t^{\alpha+1})t \quad \alpha > 0.$$

Others examples of this type can easily be constructed and their existence illustrates once again that the smoothing effect induced by the mapping $t \mapsto \int_t^{t+\theta} f(s)ds$ means the stability of solutions of (2.2.1) is very robust under *"poorly-behaved"* perturbations. Although proposition 2.4.2 tells us that the reverse implication of (2.4.5) is not true in general we provide a side condition which yields the converse in the case that the function is positive.

Proposition 2.4.3. *Let $f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R}_+)$. The following are equivalent,*

$$(i)\ \lim_{t\to\infty} \frac{1}{t}\int_0^t \int_s^{s+\theta} f(u)duds = C\theta \text{ for all } \theta \in (0,1]; \quad \lim_{t\to\infty}\frac{1}{t}\int_t^{t+1} f(s)ds = 0.$$

$$(ii)\ \lim_{t\to\infty}\frac{1}{t}\int_0^t f(s)ds = C.$$

Proof of proposition 2.4.3. The fact that (ii) implies (i) is clear from the discussion at the beginning of this section and from the identity,

$$\frac{1}{t}\int_t^{t+1} f(s)ds = \frac{1}{t}\int_0^{t+1} f(s)ds - \frac{1}{t}\int_0^t f(s)ds.$$

To see the other implication we write for $t > \theta$,

$$\begin{aligned} &\frac{1}{t}\int_0^t \int_s^{s+\theta} f(u)duds \qquad (2.4.6)\\ &\qquad = \frac{1}{t}\int_0^\theta f(u)(u-\theta)du + \frac{\theta}{t}\int_0^t f(u)du + \frac{1}{t}\int_t^{t+\theta} f(u)(t+\theta-u)du. \end{aligned}$$

The first and third term on the right hand side will vanish upon sending $t \to \infty$ and thus we obtain the desired limit. □

2.4.4 Comparison with integral equations

To this point we have considered integro-differential and functional equations and have shown solutions to perturbed equations may admit a Cesàro limit even when the perturbation function does not. This phenomenon is impossible for integral equations, and in this section we provide a proof. Consider the integral equation,

$$x(t) = \int_0^t k(t-s)x(s)ds + f(t), \quad t \geq 0, \qquad (2.4.7)$$

where $f \in L^1_{loc}(\mathbb{R}_+;\mathbb{R})$ and $k \in L^1(\mathbb{R}_+;\mathbb{R})$. In this regime the integral resolvent of k denoted by r_k is the solution to the convolution equation $r_k = k + r_k * k$, on the halfline $\mathbb{R}_+$. The standard variation of constants formula yields, $x(t) = f(t) + (r_k * f)(t), t \geq 0$. Now let $r_k \in L^1(\mathbb{R}_+;\mathbb{R})$. Assume $f \in \text{Ces}(\mathbb{R}_+;\mathbb{R})$; then Lemma 2.4.3 implies $x \in \text{Ces}(\mathbb{R}_+;\mathbb{R})$. Conversely assume $x \in \text{Ces}(\mathbb{R}_+;\mathbb{R})$, then with $k \in L^1(\mathbb{R}_+;\mathbb{R})$ we can rearrange (2.4.7) and apply Lemma 2.4.3 to give $f \in \text{Ces}(\mathbb{R}_+;\mathbb{R})$.

2.5 Conclusion

In this chapter we have characterised when the solution of a perturbed linear functional differential equation lies in either the space $\text{Ces}(\mathbb{R}_+;\mathbb{R})$ or $L^p(\mathbb{R}_+;\mathbb{R})$ and also provided necessary and sufficient conditions to ensure the solution of a perturbed linear integro-differential Volterra equation lies in $\text{Ces}(\mathbb{R}_+;\mathbb{R})$ or $L^p(\mathbb{R}_+;\mathbb{R})$ [2]. It is interesting to reflect on the improvements made with respect to the classical theory. For classical admissibility theory, three things are needed to ensure solutions of (2.2.1) reside in a particular space V while operating under the assumptions that $r \in L^1(\mathbb{R}_+;\mathbb{R})$ and $\nu \in M(\mathbb{R}_+;\mathbb{R})$, namely:

1. The resolvent obeys, $r \in V$.

2. If $g \in L^1(\mathbb{R}_+;\mathbb{R})$ then the operator $h \mapsto g * h$ must map V into itself.

3. The perturbation function obeys $f \in V$.

The results in this chapter have shown that the third point can be significantly weakened. We can have $f \notin V$ and yet solutions still reside in V. This is contingent on the perturbation function admitting a decomposition of the form $f = f_1 + f_2$ such that $f_1 \in V$ and $t \mapsto \int_0^t f_2(s)ds \in V$. Even after scrutiny of the proofs of Lemmas 2.3.1, 2.3.2, 2.4.1 and 2.4.2 it seems unclear how one may generally describe such spaces. However even with this said, it is the authors opinion that given any "*reasonable*" space that is usually considered for admissibility theory that satisfies point 2 above, a decomposition result like the one described can be proven. The results from this chapter suggest that in order to obtain sharp results for perturbed dynamical systems the correct condition to study is (2.1.3). In the following chapters we shall show this is also the case for stochastic equations. In regards to the deterministic theory, an interesting extension would be to consider emulating such methodology for infinite dimensional equations like those found in Prüss [101].

[2] under the assumption of an integrable resolvent.

2.6 Proofs

Proof of identities (2.3.2) *and* (2.4.2). Note that (2.3.2) is clearly a special case of (2.4.2), obtained by setting $\theta = 1$, hence we prove (2.4.2). First let $t > \theta$, integrate f_2 over $[0, t]$ and apply Fubini's theorem to the second integral to obtain,

$$\begin{aligned}
\int_0^t f_2(s)ds &= \int_0^t f(s)ds - \int_0^t \frac{1}{\theta}\int_{s-\theta}^s f(u)duds \\
&= \int_0^t f(s)ds - \frac{1}{\theta}\int_{-\theta}^0 \int_0^{u+\theta} f(u)dsdu - \frac{1}{\theta}\int_0^{t-\theta}\int_u^{u+\theta} f(u)dsdu \\
&\qquad - \frac{1}{\theta}\int_{t-\theta}^t \int_u^t f(u)dsdu.
\end{aligned}$$

Recalling that we extended $f(t)$ to be zero for $t < 0$ shows the second term above is zero. Then evaluating the inner integrals of the last two terms gives,

$$\int_0^t f_2(s)ds = \int_0^t f(s)ds - \int_0^{t-\theta} f(s)ds - \frac{1}{\theta}\int_{t-\theta}^t f(s)(t-u)ds.$$

Now combining all integrals on the right and then applying Fubini's theorem for a final time yields,

$$\int_0^t f_2(s)ds = \frac{1}{\theta}\int_{t-\theta}^t \int_{t-u}^{\theta} f(s)duds = \frac{1}{\theta}\int_0^{\theta}\int_{t-u}^t f(s)dsdu.$$

If $t \in (0, \theta]$ we once again integrate f_2 over $[0, t]$,

$$\begin{aligned}
\int_0^t f_2(s)ds &= \int_0^t f(s)ds - \int_0^t \frac{1}{\theta}\int_{s-\theta}^s f(u)duds \\
&= \int_0^t f(s)ds - \int_0^{\theta} \frac{1}{\theta}\int_0^s f(u)duds,
\end{aligned}$$

where the second equality follows again from the fact $f(t) = 0$ for $t < 0$. Now we apply Fubini's theorem to the second integral on the right and evaluate the resulting inner integral,

$$\int_0^t f_2(s)ds = \int_0^t f(s)ds - \int_0^\theta \frac{1}{\theta}\int_0^s f(u)duds.$$
$$= \int_0^t f(s)ds - \frac{1}{\theta}\int_0^t f(u)(t-u)du.$$

Combining both integrals and applying Fubini's theorem for the final time gives,

$$\int_0^t f_2(s)ds = \frac{1}{\theta}\int_0^t \int_{t-s}^\theta f(s)duds = \frac{1}{\theta}\int_0^\theta \int_{t-u}^t f(s)dsdu.$$

Thus (2.4.2) is proven. □

Proof of Lemma 2.4.3. We only prove the statement for convolutions with finite measures, as the result for L^1 functions is a special case wherein the measure is absolutely continuous with respect to the Lebesgue measure. We show the difference $\left|\frac{1}{t}\int_0^t (f*\nu)(s)ds - C\nu(\mathbb{R}_+)\right|$, tends to zero as $t\to\infty$. Let $F(t) := \int_0^t f(s)ds$, then this difference (after an application of Fubini's theorem) becomes,

$$\left|\frac{1}{t}\int_{[0,t]} F(t-u)\nu(du) - C\nu(\mathbb{R}_+)\right|.$$

We first consider only the case when $|C|>0$ as the case when $C=0$ is simpler and follows from an analogous argument. Let $\varepsilon\in(0,1)$ be arbitrary and fix the positive constants K,T and α such that,

$$\left|\frac{1}{K}\int_0^K f(s)ds - C\right| < \frac{\varepsilon}{5|\nu|(\mathbb{R}_+)}; \qquad |\nu|([T,\infty)) < \frac{\varepsilon}{5|C|};$$

$$\alpha < \frac{\varepsilon}{5}\min\left\{\frac{1}{|C||\nu|(\mathbb{R}_+)}, \left(1-\frac{K}{K+1}\right)\right\}.$$

Then fix,

$$t > \max\left\{(K+1), (T+K), \frac{T}{\alpha}, \left(\frac{5|\nu|(\mathbb{R}_+)}{\varepsilon}\sup_{s\in[0,K]}\left|\int_0^s f(s)ds\right|\right)\right\}.$$

Now consider,

$$\begin{aligned}
&\Big|\frac{1}{t}\int_{[0,t]} F(t-u)\nu(du) - C\nu(\mathbb{R}_+)\Big| \\
&\qquad \leq \left|\frac{1}{t}\int_{[0,t-K]} F(t-u)\nu(du) - C\nu(\mathbb{R}_+)\right| + \frac{|\nu|(\mathbb{R}_+)}{t}\sup_{s\in[0,K]}\left|\int_0^s f(s)ds\right| \\
&\qquad < \left|\frac{1}{t}\int_{[0,t-K]} F(t-u)\nu(du) - C\nu(\mathbb{R}_+)\right| + \frac{\varepsilon}{5}.
\end{aligned}$$

Next we make the substitution $s = t-u$,

$$\begin{aligned}
&\Big|\frac{1}{t}\int_{[0,t-K]} F(t-u)\nu(du) - C\nu(\mathbb{R}_+)\Big| + \frac{\varepsilon}{5} \\
&\qquad = \Big|\frac{1}{t}\int_{[K,t]} s\left[\frac{F(s)}{s} - C\right]\nu(ds) + \frac{C}{t}\int_{[K,t]} s\nu(ds) - C\nu(\mathbb{R}_+)\Big| + \frac{\varepsilon}{5} \\
&\qquad \leq \Big|\frac{C}{t}\int_{[K,t]} s\nu(ds) - C\nu(\mathbb{R}_+)\Big| + \int_{[K,t]}\left|\frac{F(s)}{s} - C\right||\nu|(ds) + \frac{\varepsilon}{5} \\
&\qquad < \Big|\frac{C}{t}\int_{[K,t]} s\nu(ds) - C\nu(\mathbb{R}_+)\Big| + \frac{\varepsilon|\nu|([K,t])}{5|\nu|(\mathbb{R}_+)} + \frac{\varepsilon}{5} \\
&\qquad \leq \Big|\frac{C}{t}\int_{[K,t]} s\nu(ds) - C\nu(\mathbb{R}_+)\Big| + \frac{2\varepsilon}{5}.
\end{aligned}$$

Now we undo the previous substitution and set $u = t - s$,

$$\begin{aligned}
&\Big|\frac{C}{t}\int_{[K,t]} s\nu(ds) - C\nu(\mathbb{R}_+)\Big| + \frac{2\varepsilon}{5} \\
&\qquad = \Big|\frac{C}{t}\int_{[0,t-K]} (t-u)\nu(du) - C\nu(\mathbb{R}_+)\Big| + \frac{2\varepsilon}{5} \\
&\qquad \leq \Big|C\nu([0,t-K]) - C\nu(\mathbb{R}_+)\Big| + \Big|\frac{C}{t}\int_{[0,t-K]} u\,\nu(du)\Big| + \frac{2\varepsilon}{5} \\
&\qquad < \frac{\varepsilon}{5} + \Big|\frac{C}{t}\int_{[0,t-K]} u\,\nu(du)\Big| + \frac{2\varepsilon}{5},
\end{aligned}$$

where the last inequality follows from the fact that $t > T + K$ and

$$|C\nu([0,t-K]) - C\nu(\mathbb{R}_+)| = |C||\nu|([t-K,\infty)).$$

Now our definition of α ensures $\alpha t < t - K$ when $t > K + 1$, thus we can write

$$\begin{aligned}
&\Big|\frac{C}{t}\int_{[0,t-K]} u\,\nu(du)\Big| + \frac{3\varepsilon}{5} \\
&\qquad = \Big|\frac{C}{t}\int_{[0,\alpha t]} u\,\nu(du) + \frac{C}{t}\int_{[\alpha t,t-K]} u\,\nu(du)\Big| + \frac{3\varepsilon}{5} \\
&\qquad \leq \alpha\,|C||\nu|(\mathbb{R}_+) + \frac{|C|}{t}\int_{[\alpha t,t-K]} |u|\,|\nu|(du) + \frac{3\varepsilon}{5} \\
&\qquad < \frac{\varepsilon}{5} + \frac{|C|}{t}\int_{[\alpha t,t]} |u|\,|\nu|(du) + \frac{3\varepsilon}{5} \\
&\qquad < \frac{\varepsilon}{5} + |C||\nu|([\alpha t,\infty)) + \frac{3\varepsilon}{5}.
\end{aligned}$$

Now finally as $\alpha t > T$ we have,

$$|C||\nu|([\alpha t,\infty)) + \frac{4\varepsilon}{5} < \frac{\varepsilon}{5} + \frac{4\varepsilon}{5} = \varepsilon,$$

as required. □

Proof of proposition 2.4.2. Throughout this proof the notation β^{-1} denotes the functional inverse of β and not the reciprocal.

$$\frac{1}{t}\int_1^t f(s)ds = -\cos(\beta(t)) + \frac{\cos(\beta(1))}{t} + \frac{1}{t}\int_{\beta(1)}^{\beta(t)} \frac{\cos(u)}{\dot\beta(\beta^{-1}(u))}du.$$

Now suppose the integral on the right is convergent, then we will have,

$$\frac{1}{t}\int_1^t f(s)ds = -\cos(\beta(t)) + o(1),$$

and thus $f \notin \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$. Now we show this integral is indeed convergent.

$$\int_{\beta(1)}^{\beta(t)} \frac{\cos(u)}{\dot\beta(\beta^{-1}(u))}du = \frac{\sin(\beta(t))}{\dot\beta(t)} - \frac{\sin(\beta(1))}{\dot\beta(1)} + \int_1^t \frac{\ddot\beta(u)}{(\dot\beta(u))^2}\sin(\beta(u))du.$$

As $\dot\beta(t) \to \infty$ as $t \to \infty$ we need only estimate the final integral.

$$\left|\int_1^t \frac{\ddot\beta(u)}{(\dot\beta(u))^2}\sin(\beta(u))du\right| \leq \int_{\dot\beta(1)}^{\dot\beta(t)} \frac{1}{s^2}ds,$$

which is a convergent integral as required. Next we show $t \mapsto \int_t^{t+\theta} f(s)ds \in \mathrm{Ces}(\mathbb{R}_+;\mathbb{R})$. Let,

$$I(t) := \cos(1) + \int_{\beta(1)}^{\beta(t)} \frac{\cos(u)}{\dot\beta(\beta^{-1}(u))}du,$$

then from the above analysis we have $I(t) \to I^* \in \mathbb{R}$. It follows that,

$$\int_t^{t+\theta} f(s)ds = -\theta\cos(\beta(t+\theta)) - t\cos(\beta(t+\theta)) + t\cos(\beta(t)) + I(t+\theta) - I(t).$$

It is clear that,

$$\frac{1}{t}\int_0^t \cos(\beta(u+\theta))du \to 0 \text{ as } t \to \infty,$$

and also,

$$\frac{1}{t}\int_0^t I(u+\theta) - I(u)du \to 0 \text{ as } t \to \infty.$$

Finally we show $\frac{1}{t}\int_0^t u\cos(\beta(u))du \to 0$ and the claim will be proven. Making the substitution $\omega = \beta(u)$, we see,

$$\int_1^t u\cos(\beta(u))du = \int_{\beta(1)}^{\beta(t)} \frac{\beta^{-1}(\omega)}{\dot{\beta}(\beta^{-1}(\omega))}\cos(\omega)d\omega.$$

Introducing $\gamma(t) := \frac{\beta^{-1}(t)}{\dot{\beta}(\beta^{-1}(t))}$, we note,

$$\gamma'(t) = \frac{1}{(\dot{\beta}(\beta^{-1}(t)))^2} - \frac{\ddot{\beta}(\beta^{-1}(t))}{(\dot{\beta}(\beta^{-1}(t)))^3}\beta^{-1}(t),$$

then integrating by parts once again yields,

$$\begin{aligned}\int_{\beta(1)}^{\beta(t)} \frac{\beta^{-1}(u)}{\dot{\beta}(\beta^{-1}(u))}\cos(u)du =& \frac{t\sin(\beta(t))}{\dot{\beta}(t)} - \frac{\sin(\beta(1))}{\dot{\beta}(1)} - \int_{\beta(1)}^{\beta(t)} \frac{\sin(u)}{(\dot{\beta}(\beta^{-1}(u)))^2}du \\ &+ \int_{\beta(1)}^{\beta(t)} \frac{\ddot{\beta}(\beta^{-1}(u))}{(\dot{\beta}(\beta^{-1}(u)))^3}\beta^{-1}(u)\sin(u)du.\end{aligned}$$

Upon dividing by t it is easy to show the first three terms on the right hand side will vanish as $t \to \infty$ so we focus only on the last, majorising and making the substitution $\omega = \beta^{-1}(u)$ gives,

$$\left|\int_{\beta(1)}^{\beta(t)} \frac{\ddot{\beta}(\beta^{-1}(u))}{(\dot{\beta}(\beta^{-1}(u)))^3}\beta^{-1}(u)\sin(u)du\right| \leq \int_1^t \frac{\omega\ddot{\beta}(\omega)}{(\dot{\beta}(\omega))^2}dw.$$

Now for $\varepsilon \in (0,1)$ we can write,

$$\begin{aligned}\int_1^t \frac{\omega\ddot{\beta}(\omega)}{(\dot{\beta}(\omega))^2}dw &= \int_1^{\varepsilon t} \frac{\omega\ddot{\beta}(\omega)}{(\dot{\beta}(\omega))^2}dw + \int_{\varepsilon t}^t \frac{\omega\ddot{\beta}(\omega)}{(\dot{\beta}(\omega))^2}dw \\ &\leq \varepsilon t\int_1^\infty \frac{\ddot{\beta}(\omega)}{(\dot{\beta}(\omega))^2}dw + t\int_{\varepsilon t}^\infty \frac{\ddot{\beta}(\omega)}{(\dot{\beta}(\omega))^2}dw.\end{aligned}$$

Then dividing across by t, sending $t \to \infty$ and then $\varepsilon \to 0$ yields a zero limit as required.

□

Chapter 3

Stochastic Volterra and Functional Equations with Additive Noise

The material in this chapter is based on the following article:

1. J. A. D. Appleby and E. Lawless. Solution space characterisation of perturbed linear discrete and continuous stochastic Volterra convolution equations: the ℓ^p and L^p cases. *submitted* (2024) [14].

3.1 Introduction

Over the last four decades, a substantial literature has been developed concerning the qualitative behaviour of both deterministic and stochastic, dynamical systems with memory. For systems with finite memory the deterministic literature is summarised in classical texts such as [44, 72] while for the stochastic literature recent monographs by Mao and Shaiket [85, 112] have appeared. When considering deterministic Volterra equations one can consult [38, 53], however to the best of our knowledge such a monograph for stochastic Volterra equations is yet to appear. In this chapter we endeavour to extend some of the results in Chapter 2 to the stochastic case which leads us to the following $\mathbb{R}^d$-valued stochastic Volterra integrodifferential equation,

$$dX(t) = \left(f(t) + \int_{[0,t]} \nu(ds)X(t-s) \right) dt + \sigma(t)dB(t), \quad t \geq 0. \tag{3.1.1}$$

Here ν is understood to be a matrix-valued finite signed Borel measure and B an m-dimensional Brownian motion. The general theory of stochastic Volterra equations has a long history, we mention the following influential articles [24, 97, 99] for general existence and uniqueness results. Such general equations have also seen prominent use in applications, asymptotic stability of stochastic hereditary systems is considered in [94] while [25] provides numerous applications in physics such as studying the effects of random heat sources on electrical currents. The asymptotic behaviour of solutions to equations of type (3.1.1) was first considered by Chan and Williams [32]. Following on from this work, an abundance of literature has appeared considering various types of behaviour both qualitative and quantitative. Asymptotic behaviour is considered in [3, 6, 7, 8, 19] while subexponential and exponential asymptotic behaviour is explored in [4, 9]. Attention has also been given to equations with non semimartingale drivers in both finite [46, 48] and infinite dimensions [47] while equations with multiplicative noise (both linear and non-linear) have also received due study [5, 83, 87, 113]. Our primary goal is to provide necessary and sufficient conditions on the forcing terms f and σ in order to guarantee the trajectories of solutions to (3.1.1) are almost surely p-integrable functions in time. Indeed sufficient conditions are easily obtained, but to the best of our knowledge a complete characterisation has up to this has point remained unknown. The first barrier at hand is to characterise when solutions to the underlying deterministic equation (i.e when $\sigma = 0$) are elements of $L^p(\mathbb{R}_+;\mathbb{R}^d)$. This was carried out in the scalar case in Chapter 2 and will be developed here. The condition generated does not impose any restrictions directly on f but rather on a continuous linear functional applied to f, namely we must have for each component,

$$t \mapsto \int_t^{t+\theta} f_i(s)ds \in L^p(\mathbb{R}_+;\mathbb{R}) \quad \text{for all } \theta > 0, \tag{3.1.2}$$

to ensure solutions to the underlying deterministic equation are p-integrable for $p \geq 1$. This extension to finite dimensions follows as a simple corollary to Theorem 3.3.1 below. Turning our attention to σ we see a dichotomy of cases; for $p \geq 2$ one needs each component of the matrix σ to satisfy,

$$t \mapsto \int_t^{t+\theta} \sigma_{ij}^2(s)ds \in L^{\frac{p}{2}}(\mathbb{R}_+;\mathbb{R}) \quad \text{for all } \theta > 0, \tag{3.1.3}$$

while for $p \in [1, 2)$ we require,

$$n \mapsto \int_n^{n+1} \sigma_{ij}^2(s)ds \in \ell^{\frac{p}{2}}(\mathbb{Z}_+;\mathbb{R}). \tag{3.1.4}$$

The fact $p = 2$ is the border case is not surprising and arises due to Itô's isometry. Thus if we make the standard assumption that the underlying resolvent of the measure ν is integrable[1], then our main result shows that for $p \geq 1$ (along with any norm on $\mathbb{R}^d$), the solution to (3.1.1) satisfying,

$$\|X(\cdot)\| \in L^p(\mathbb{R}_+;\mathbb{R}) \textit{ almost surely}, \tag{3.1.5}$$

is equivalent to f satisfying (3.1.2) (component wise) and each component of σ obeying

[1]Recall this is equivalent to $\det[zI_{d\times d} - \hat{\nu}(z)] \neq 0$ for $\text{Re}(z) \geq 0$ where $z \in \mathbb{C}$ and $\hat{\nu}$ denotes the Laplace transform of the measure ν.

either (3.1.3) or (3.1.4) depending on the value of p. The utility of this result lies upon the fact that the perturbation functions may be very ill-behaved and yet conditions (3.1.2), (3.1.3) and (3.1.4) may still be satisfied. An explicit example was provided in section 2.3.1 while a new example is provided in section 3.5 below. In the context of equation (3.1.1), characterising when a process has almost surely p-integrable sample paths is of interest due to the fact it is equivalent to the p^{th}-mean of the process being integrable, i.e (3.1.5) is equivalent to $\int_0^\infty \mathbb{E}[\|X(s)\|^p]ds < +\infty$ (see Theorem 3.3.1 below). Such a condition is often a first stop on the way to proving $\mathbb{E}[\|X(t)\|^p] \to 0$ as $t \to \infty$, which is of great interest in all domains where such models are employed in practise. Pairing this integrability condition with a pathwise regularity result on the mapping $t \mapsto \mathbb{E}[\|X(t)\|^p]$ such as Lipschitz (or uniform) continuity yields the desired convergence to zero. For stochastic functional differential equations the classical approach is via construction of an appropriate Lyapunov functional which in turn gives the integrability of the p^{th}-mean, this is outlined in [73, Chapter 4]. The drawbacks of this method lie in the difficulty of constructing such a Lyapunov functional and the reward for such hard work is a sufficient condition; necessary conditions are not easily obtained. The approach in this chapter is simpler and provides characterisations of p-integrable paths rather than generating sufficient conditions, our analysis is direct and focuses on the conditions (3.1.2), (3.1.3) and (3.1.4). Additionally we obtain precise information about the almost sure asymptotic behaviour of the paths (Theorem 3.3.4) without needing to analyse the path regularity of the p^{th}-mean.

We make the observation that all results proven for equation (3.1.1) are also true if one considers instead stochastic functional differential equations of the form,

$$dX(t) = \left(f(t) + \int_{[-\tau,0]} \mu(ds)X(t+s) \right) dt + \sigma(t)dB(t), \quad t \geq 0; \quad X(t) = \psi(t) \quad t \leq 0. \tag{3.1.6}$$

Here the measure-valued kernel μ is concentrated on the compact set $[-\tau, 0]$ where $\tau > 0$ is the fixed delay parameter and the initial data is a $C([-\tau, 0]; \mathbb{R}^d)$-valued random variable. We shall often write $X(t, \psi)$ when referring to solutions of (3.1.6) to emphasize the dependence on the random initial function and to distinguish between the solution of the Volterra equation. In fact in this case we do not need any assumption regarding the integrability of the resolvent: this comes out as a necessary condition in the proof.

It transpires that in order to prove our main result, we require an auxiliary lemma regarding the summability of sequences of certain random variables which is lifted directly from the analysis of the discrete time analogue of (3.1.1). By this we mean the $\mathbb{R}^d$-valued stochastic Volterra summation equation[2],

[2]We shall use the notation X to represent solutions to both (3.1.1) and (3.1.7), it should be clear by the context to which equation we are referring.

$$X(n+1) = X(n) + \sum_{j=0}^{n} K(n-j)X(j) + f(n) + \sigma(n)\xi(n+1), \quad n \in \mathbb{Z}^{,} \qquad (3.1.7)$$

where ξ is some $i.i.d$ sequence of random vectors, f, σ are deterministic and the kernel K is a matrix valued sequence with entries in $\ell^1(\mathbb{Z}_+; \mathbb{R})$. Equations of this type have also drawn much attention from the literature, for essential background on the deterministic theory we refer the reader to [104]. Like in the stochastic case there are many variations of (3.1.7) one can study. For work focused on a fixed number of time lags often mean square behaviour for systems with both additive and multiplicative noise is a priority [11, 110]. The monograph by Shaiket [111] provides an excellent overview for such equations. In the Volterra case asymptotic behaviour of the trajectories is considered for both linear deterministic equations [18, 43, 57] and stochastic equations [20, 26]. Our results are of a qualitative flavour, i.e in order for solutions of (3.1.7) to obey $\|X\| \in \ell^p(\mathbb{Z}_+; \mathbb{R})$ almost surely, it is necessary and sufficient to have $\|f\|, \|\sigma\| \in \ell^p(\mathbb{Z}_+; \mathbb{R})$. There is a trade-off between what assumptions one makes about the noise ξ and the structure you impose on the matrix σ. For instance you can allow the components of ξ to follow essentially any arbitrary distribution (this is made precise by definition 3.2.1) and depend on one another in any way you please but this forces σ to be diagonal. If one wishes to have a general matrix σ then for the above result to hold we need to restrict ourselves to Gaussian sequences with independent components.

The layout of this chapter is as follows. In section 3.2.1 we discuss the notation and definitions necessary to study the discrete equation (3.1.7). Section 3.2.2 focuses on a general noise sequence without componentwise independence while section 3.2.3 handles the case of Gaussian noise. We then move on to study the continuous time equation, introducing notation in section 3.3.1, the main results in section 3.3.2 and additional results regarding the asymptotic behaviour of sample paths in section 3.3.3. In section 3.4 we discuss stochastic functional equations before providing examples in section 3.5 to emphasize the utility of theoretical results. Finally we conclude with a discussion of results in section 3.6.

3.2 Discrete results

3.2.1 Notation and mathematical preliminaries

We denote the standard basis vectors in $\mathbb{R}^d$ by $\mathbf{e}_1, \dots, \mathbf{e}_d$. The sum of all basis vectors in $\mathbb{R}^d$ will be denoted $\mathbf{e}_{[d]}$. The standard innerproduct on $\mathbb{R}^d$ is denoted by $\langle \cdot, \cdot \rangle$. We say the $d \times m$ matrix-valued sequence $a : \mathbb{Z}_+ \to \mathbb{R}^{d\times m}$, $n \mapsto a(n)$ obeys $a \in \ell^p(\mathbb{Z}_+; \mathbb{R}^{d\times m})$ if each component sequence, $(a(n))_{i,j} \in \ell^p(\mathbb{Z}_+; \mathbb{R})$ for $i = 1, \dots, d$ and $j = 1, \dots, m$, where $(a(n))_{i,j} := \langle \mathbf{e}_i, a(n)\mathbf{e}_j \rangle$. For scalar valued objects we shall use the notation $|\cdot|$ to represent the standard euclidean norm on $\mathbb{R}$. We will use the common short hand $X \overset{d}{=} Y$, to denote when two random variables X and Y are equal in distribution and the notation χ to denote indicator random variables/functions.

We introduce the following class of real valued random variables, which will be used in the proofs of a large proportion of converse results throughout this section.

Definition 3.2.1. *Let ξ be a real valued random variable. We say $\xi \in \mathbb{D}$ if there exists distinct, measurable and bounded, Borel sets B_1, B_2 such that,*

(i) $\mathbb{P}[\xi \in B_i] > 0$ *for* $i = 1, 2$ *with* $p_1 := \mathbb{P}[\xi \in B_1]$ *and* $p_2 := \mathbb{P}[\xi \in B_2]$.

(ii) $\mathbb{E}[\xi\chi_{\{\xi \in B_i\}}] \neq 0$ *for* $i = 1, 2$ *with* $e_1 := \mathbb{E}[\xi\chi_{\{\xi \in B_1\}}]$ *and* $e_2 := \mathbb{E}[\xi\chi_{\{\xi \in B_2\}}]$.

(iii) $p_2 e_1 - p_1 e_2 \neq 0$.

Remark 3.2.1. *The class of random variables introduced in definition 3.2.1 is certainly nonstandard. The definition is made in this manner, mainly for technical reasons that appear in subsequent proofs. However upon some reflection one realises that this class of random variables is very large indeed. It is the author's opinion that after excluding random variables with singular laws, the only random variables that are not elements of the space $\mathbb{D}$ are the almost surely constant random variables. However no proof is provided for this conjecture. Instead we prove that if a random variable ξ has a distribution with either an absolutely continuous part or at least two isolated atoms then $\xi \in \mathbb{D}$(see proposition 3.2.1 below, the proof of which is relegated to the final section of this chapter).*

Proposition 3.2.1. *Let ξ be a real valued random variable whose distribution has either an absolutely continuous part or at least two isolated atoms (with non-zero probability), then $\xi \in \mathbb{D}$.*

3.2.2 General noise

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space. We study the following, $\mathbb{R}^d$-valued Stochastic Volterra summation equation,

$$X(n+1) = X(n) + \sum_{j=0}^{n} K(n-j)X(j) + f(n) + \sigma(n)\xi(n+1), \quad n \in \mathbb{Z}_+; \quad X(0) = \xi. \tag{3.2.1}$$

Here K and σ are real, deterministic $d \times d$ matrix-valued sequences and f is a real, deterministic $\mathbb{R}^d$ valued sequence. Each $\xi(n)$ is a random vector in $\mathbb{R}^d$. We make the following standing assumptions throughout this subsection,

$$\sigma(n) \text{ is a sequence of diagonal matrices.} \tag{3.2.2}$$

$$\xi(n) \text{ is an i.i.d sequence of } \mathbb{R}^d\text{-valued random vectors,} \tag{3.2.3}$$

$$\langle \xi(n), \mathbf{e}_j \rangle \overset{d}{=} \xi_j \text{ where } \xi_j \in \mathbb{D}, \text{ for } j = 1, \dots, d. \tag{3.2.4}$$

We assume our probability space is rich enough to support the corresponding sequence of random vectors $(\xi(n))_{n\geq 1}$ and equip it with the natural filtration generated by $\xi(n)$, i.e $\mathcal{F}_n = \sigma(\xi(k) : 1 \leq k \leq n)$ with the convention that $\mathcal{F}_0 = \{\emptyset, \Omega\}$. It is worth stressing some common assumptions that we are **not** making. We do not require the random variables $\langle \xi(n), \mathbf{e}_j \rangle$ and $\langle \xi(n), \mathbf{e}_i \rangle$ for $i \neq j$, to be independent nor to we require that they have the same distribution. We now introduce an auxiliary object which allows efficient study of equation (3.2.1). Let the so called resolvent, be the $\mathbb{R}^{d\times d}$-valued solution to the following matrix equation,

$$R(n+1) = R(n) + \sum_{j=0}^{n} K(n-j)R(j), \quad n \in \mathbb{Z}_+; \quad R(0) = I_{d\times d}. \tag{3.2.5}$$

Thus the well-known variation of constants formula then gives,

$$X(n) = R(n)\xi + \sum_{j=1}^{n} R(n-j)\left[f(j-1) + \sigma(j-1)\xi(j)\right], \quad n \in \mathbb{Z}_+. \tag{3.2.6}$$

In the sequel we will always assume,

$$K \in \ell^1(\mathbb{Z}_+; \mathbb{R}^{d\times d}), \quad R \in \ell^1(\mathbb{Z}_+; \mathbb{R}^{d\times d}). \tag{3.2.7}$$

This is a standard assumption one needs to make in order to infer any reasonable convergence results for solutions to (3.2.1), see [20].

Theorem 3.2.1. *Let $p \in [1, \infty)$. Suppose $\xi(n)$ obeys (3.2.3)-(3.2.4) and additionally that $\langle \xi(n), \mathbf{e}_j \rangle \in L^p(\Omega)$ for $j = 1, \dots, d$. Let K and R obey (3.2.7), σ obey (3.2.2) and let X be the solution of (3.2.1). Then the following are equivalent,*

(i) $f \in \ell^p(\mathbb{Z}_+, \mathbb{R}^d)$ and $\sigma \in \ell^p(\mathbb{Z}_+; \mathbb{R}^{d\times d})$,

(ii) $X \in \ell^p(\mathbb{Z}_+, \mathbb{R}^d)$ almost surely.

An interesting question to ask is whether p-summability is ever possible for the solution of (3.2.1) without having the deterministic perturbations f and σ being p-summable. One might conjecture that this is possible but upon inspection of the proof of the converse result of Theorem 3.2.1 we see that this cannot be the case. We state this explicitly as a corollary, whose proof follows directly from the proof of Theorem 3.2.1.

Corollary 3.2.1. *Suppose $\xi(n)$ obeys (3.2.3)-(3.2.4). Let K and R obey (3.2.7), σ obey (3.2.2) and let X be the solution of (3.2.1). If $X \in \ell^p(\mathbb{Z}_+, \mathbb{R}^d)$ a.s. for $1 \le p < \infty$, then $f \in \ell^p(\mathbb{Z}_+, \mathbb{R}^d)$ and $\sigma \in \ell^p(\mathbb{Z}_+; \mathbb{R}^{d\times d})$.*

Theorem 3.2.1 essentially tells us that a "*stabilisation by noise*" effect cannot occur. It seems the only way such a phenomena can occur is if one relaxes the identically distributed assumption on the noise sequence to allow the possibility of $\xi(n) \to 0$ in an appropriate sense. Before proving these results we prepare a lemma.

Lemma 3.2.1. *Let $(f(n))_{n\ge 0}$ and $(\sigma(n))_{n\ge 0}$ be deterministic (scalar) sequences and $(\xi(n))_{n\ge 1}$ be an i.i.d sequence of (scalar) random variables with $\xi(n) \stackrel{d}{=} \xi$ where $\xi \in \mathbb{D}$. If for any $p \in [1, \infty)$,*

$$\sum_{n=0}^{\infty} |f(n) + \sigma(n)\xi(n+1)|^p < +\infty \quad a.s., \tag{3.2.8}$$

then $f, \sigma \in \ell^p(\mathbb{Z}_+)$.

Proof of Lemma 3.2.1. First we show that (3.2.8) implies both f and σ are null sequences and hence bounded. We note that (3.2.8) implies that $f(n) + \sigma(n)\xi \stackrel{n\to\infty}{\longrightarrow} 0$ almost surely, to see this consider the following argument.

By (3.2.8) we have for all $\varepsilon > 0, \mathbb{P}\left[|f(n) + \sigma(n)\xi(n+1)| > \varepsilon, \text{ i.o.}\right] = 0$. Now fix an $\varepsilon > 0$ and suppose $\mathbb{P}\left[|f(n) + \sigma(n)\xi| > \varepsilon, \text{ i.o.}\right] > 0$. By the first Borel-Cantelli lemma this implies

$$\sum_{n=0}^{\infty} \mathbb{P}\left[|f(n) + \sigma(n)\xi| > \varepsilon\right] = +\infty,$$

which by the equality in distribution of $\xi(n)$ and ξ, forces,

$$\sum_{n=0}^{\infty} \mathbb{P}\left[|f(n) + \sigma(n)\xi(n+1)| > \varepsilon\right] = +\infty.$$

Then leveraging the independence of the $\xi(n)$'s, we apply the second Borel-Cantelli lemma to give,

$$\mathbb{P}\left[|f(n) + \sigma(n)\xi(n+1)| > \varepsilon, \text{ i.o.}\right] = 1,$$

which is a contradiction and so we must have $\mathbb{P}\left[|f(n) + \sigma(n)\xi| > \varepsilon, \text{ i.o.}\right] = 0$. As ε was arbitrary, this gives $f(n) + \sigma(n)\xi \stackrel{n\to\infty}{\longrightarrow} 0$ almost surely, as required. Next we fix ω_1 and ω_2 such that $\xi(\omega_1) \neq \xi(\omega_2)$, which can always be done as $\xi \in \mathbb{D}$. Thus,

$$(f(n) + \sigma(n)\xi(\omega_1)) - (f(n) + \sigma(n)\xi(\omega_2)) \to 0,$$

which yields,

$$\sigma(n)(\xi(\omega_1) - \xi(\omega_2)) \to 0,$$

as $\xi(\omega_1) - \xi(\omega_2) \neq 0$, thus σ is a null sequence which forces f to be a null sequence and so they are both bounded. Now let B_1 and B_2 be the two Borel sets which come from definition 3.2.1 and introduce the sequences,

$$X_n := |f(n) + \sigma(n)\xi(n+1)|^p \chi_{\{\xi(n+1)\in B_1\}},$$
$$Y_n := |f(n) + \sigma(n)\xi(n+1)|^p \chi_{\{\xi(n+1)\in B_2\}}.$$

The boundedness f and σ along with the truncation of the noise sequence $\xi(n)$ means X_n and Y_n are in fact almost surely uniformly bounded random variables. Clearly from (3.2.8) we have that $X_n, Y_n \in \ell^p(\mathbb{Z}_+)$ almost surely, this along with uniform boundedness allows us to apply Kolmogorov's two-series test [115, Chapter IV] which yields,

$$\sum_{n=0}^{\infty} \mathbb{E}[X_n] < +\infty, \quad \sum_{n=0}^{\infty} \mathbb{E}[Y_n] < +\infty.$$

Now focusing on the first series we see,

$$\begin{aligned}
+\infty &> \sum_{n=0}^{\infty} \mathbb{E}\left[|f(n) + \sigma(n)\xi(n+1)|^p \chi_{\{\xi(n+1)\in B_1\}}\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[|f(n) + \sigma(n)\xi|^p \chi_{\{\xi\in B_1\}}\right] \\
&\geq \sum_{n=0}^{\infty} \left|\mathbb{E}\left[f(n)\chi_{\{\xi\in B_1\}} + \sigma(n)\xi\chi_{\{\xi\in B_1\}}\right]\right|^p \\
&= \sum_{n=0}^{\infty} |p_1 f(n) + e_1\sigma(n)|^p,
\end{aligned}$$

where the inequality follows from Jensen's inequality as $p \geq 1$ and p_1, e_1 are constants from definition 3.2.1. An identical argument using $\sum_{n=0}^{\infty} \mathbb{E}[Y_n]$ also yields,

$$\sum_{n=0}^{\infty} |p_2 f(n) + e_2\sigma(n)|^p < +\infty.$$

Next we leverage the fact that $p_2 e_1 - p_1 e_2 \neq 0$. Consider,

$$\begin{aligned}
\sigma(n)(p_2e_1 - p_1e_2) &= \sigma(n)(p_2e_1 - p_1e_2) + p_1p_2f(n) - p_1p_2f(n) \\
&= p_2(p_1f(n) + e_1\sigma(n)) - p_1(p_2f(n) + e_2\sigma(n)).
\end{aligned}$$

Thus we have,

$$|p_2e_1 - p_1e_2|^p \sum_{n=0}^{\infty} |\sigma(n)|^p \leq |p_2|^p \sum_{n=0}^{\infty} |p_1 f(n) + e_1\sigma(n)|^p + |p_1|^p \sum_{n=0}^{\infty} |p_2 f(n) + e_2\sigma(n)|^p.$$

As both series on the right have been shown to be finite we have that $\sigma \in \ell^p(\mathbb{Z}_+)$. A similar argument also shows $f \in \ell^p(\mathbb{Z}_+)$ and the proposition in proven. $\square$

Proof of Theorem 3.2.1. To prove the forward implication it is enough to use (3.2.6) and show $f(n)+\sigma(n)\xi(n+1)$ is an element of $\ell^p(\mathbb{Z}_+;\mathbb{R}^d)$ almost surely, as $\ell^p(\mathbb{Z}_+;\mathbb{R}^d)$ is closed under convolution with $\ell^1(\mathbb{Z}_+;\mathbb{R}^{d\times d})$ sequences (and $R \in \ell^1(\mathbb{Z}_+;\mathbb{R}^{d\times d})$ by assumption (3.2.7)). This amounts to showing the component sequence $\langle f(n) + \sigma(n)\xi(n+1), \mathbf{e}_i\rangle \in \ell^p(\mathbb{Z}_+;\mathbb{R})$ almost surely for an arbitrary $i \in \{1,\dots,d\}$. The triangle inequality and $\langle \xi(n), \mathbf{e}_i\rangle \in L^p(\Omega)$ yield,

$$\mathbb{E}\left[|\langle f(n) + \sigma(n)\xi(n+1), \mathbf{e}_i\rangle|^p\right] = \mathbb{E}[|f_i(n) + \sigma_{ii}(n)\xi_i(n+1)|^p] \le |f_i(n)|^p + C_p\,|\sigma_{ii}(n)|^p\,,$$

where $C_p = \mathbb{E}\left[|\xi_i(n+1)|^p\right] = \mathbb{E}\left[|\xi_i|^p\right]$ is independent of n by (3.2.4). Invoking (i) yields,

$$\sum_{n=0}^{\infty} \mathbb{E}\left[|\langle f(n) + \sigma(n)\xi(n+1), \mathbf{e}_i\rangle|^p\right] < \infty,$$

which implies,

$$\sum_{n=0}^{\infty} |\langle f(n) + \sigma(n)\xi(n+1), \mathbf{e}_i\rangle|^p < \infty \quad a.s.$$

For the reverse implication we use the identity,

$$X(n+1) - X(n) - \sum_{j=0}^{n} K(n-j)X(j) = f(n) + \sigma(n)\xi(n+1).$$

As $K \in \ell^1(\mathbb{Z}_+;\mathbb{R}^{d\times d})$ and $X \in \ell^p(\mathbb{Z}_+;\mathbb{R}^d)$ almost surely, the right hand side must also be an element of $\ell^p(\mathbb{Z}_+;\mathbb{R}^d)$ almost surely. Thus we have for each $i \in \{1,\dots,d\}$,

$$\sum_{n=0}^{\infty} |f_i(n) + \sigma_{ii}(n)\xi_i(n+1)|^p < \infty \quad a.s.$$

Then by Lemma 3.2.1 we have $f_i, \sigma_{ii} \in \ell^p(\mathbb{Z}_+;\mathbb{R})$ and so $f \in \ell^p(\mathbb{Z}_+,\mathbb{R}^d)$ and $\sigma \in \ell^p(\mathbb{Z}_+;\mathbb{R}^{d\times d})$.

$\square$

3.2.3 Gaussian noise

If we specialise to the case where $\xi(n)$ is an i.i.d sequence of Gaussian random vectors with zero mean and whose components are independent, then we can relax the assumption of σ being a diagonal matrix. From now on σ is an $\mathbb{R}^{d\times m}$ valued sequence and the noise vector $\xi(n)$ will take values in $\mathbb{R}^m$ for some $m \geq d$. Introduce the following assumptions,

$$\xi(n) \text{ is an i.i.d sequence of } \mathbb{R}^m\text{- valued, zero mean Gaussian random vectors.} \quad (3.2.9)$$

$$\langle \xi(n), \mathbf{e}_i \rangle \text{ is independent of } \langle \xi(n), \mathbf{e}_j \rangle \text{ for all } i \neq j \text{ and for all } n \in \mathbb{N}. \quad (3.2.10)$$

Theorem 3.2.2. *Suppose $\xi(n)$ obeys (3.2.9)-(3.2.10) and let K and R obey (3.2.7). Let X be the solution of (3.2.1). Then for $p \in [1, \infty)$ the following are equivalent,*

(i) $f \in \ell^p(\mathbb{Z}_+, \mathbb{R}^d)$ and $\sigma \in \ell^p(\mathbb{Z}_+; \mathbb{R}^{d\times m})$,

(ii) $X \in \ell^p(\mathbb{Z}_+, \mathbb{R}^d)$ almost surely.

For a proof of Theorem 3.2.2 follow the proof of Theorem 3.2.1 verbatim, the only difference being in the converse, invoke Lemma 3.2.2 instead of Lemma 3.2.1.

Lemma 3.2.2. *Let $(f(n))_{n\geq 0}$ and $(\sigma(n))_{n\geq 0}$ be deterministic, $\mathbb{R}^d$ and $\mathbb{R}^{d\times m}$-valued sequences respectively. Let $\xi(n)$ obey (3.2.9)-(3.2.10). Then for $p \in [1, \infty)$, if*

$$f + \sigma\xi \in \ell^p(\mathbb{Z}_+; \mathbb{R}^d) \quad a.s.,$$

then $f \in \ell^p(\mathbb{Z}_+; \mathbb{R}^d)$ and $\sigma \in \ell^p(\mathbb{Z}_+; \mathbb{R}^{d\times m})$.

Proof of Lemma 3.2.2. We have for each $i \in \{1, \dots, d\}$,

$$\sum_{n=0}^{\infty} \left| f_i(n) + \sum_{j=1}^{m} \sigma_{ij}(n)\xi_j(n+1) \right|^p < \infty \quad a.s.$$

Define,

$$\zeta_i(n) := \sum_{j=1}^{m} \sigma_{ij}(n)\xi_j(n+1).$$

Then $\zeta_i(n)$ is sequence of independent, normally distributed random variables with zero mean. Now consider,

$$\zeta_i(n) = \sqrt{\mathbb{E}[\zeta_i(n)^2]}\frac{\zeta_i(n)}{\sqrt{\mathbb{E}[\zeta_i(n)^2]}} = \sqrt{\mathbb{E}[\zeta_i(n)^2]}\tilde{\zeta}_i(n),$$

where $\tilde{\zeta}_i(n)$ is an i.i.d sequence of standard normal random variables[3]. Our series then becomes,

$$\sum_{n=0}^{\infty} \left| f_i(n) + \sqrt{\mathbb{E}[\zeta_i(n)^2]}\tilde{\zeta}_i(n) \right|^p < \infty \quad a.s.$$

[3]We need not worry that $\sqrt{\mathbb{E}[\zeta_i(n)^2]} = 0$, as if this is the case then $\zeta_i(n) = 0$ almost surely and so this term will have no contribution to the series. Thus in this instance we define $\tilde{\zeta}_i(n) := \zeta_i(n)$.

Invoking Lemma 3.2.1 yields,

$$\sum_{n=0}^{\infty}|f_i(n)|^p+\sum_{n=0}^{\infty}|\sqrt{\mathbb{E}[\zeta_i(n)^2]}|^p<\infty.$$

Recall that,

$$\mathbb{E}[\zeta_i(n)^2]=\sum_{j=1}^{m}\sigma_{ij}^2(n)\mathbb{E}[\xi_j(n+1)^2]=\sum_{j=1}^{m}C_j\sigma_{ij}^2(n).$$

Thus,

$$+\infty>\sum_{n=0}^{\infty}|\sqrt{\mathbb{E}[\zeta_i(n)^2]}|^p=\sum_{n=0}^{\infty}\left(\sum_{j=1}^{m}C_j\sigma_{ij}^2(n)\right)^{p/2}\geq C\sum_{n=0}^{\infty}\sum_{j=1}^{m}|\sigma_{ij}(n)|^p,$$

for some $C>0$. The the last inequality follows from the fact that norms on finite dimensional vector spaces are equivalent and that $p\geq 1$. Thus $f_i,\sigma_{ij}\in\ell^p(\mathbb{Z}_+;)$ for $i\in\{1,\dots,d\}, j\in\{1,\dots,m\}$. The claim is proven.

□

3.3 Continuous Results

3.3.1 Notation

We say a function $f:\mathbb{R}_+\to\mathbb{R}^{d\times m}$, $t\mapsto f(t)$ obeys $f\in L^p(\mathbb{R}_+;\mathbb{R}^{d\times m})$ if each component function $(f(t))_{i,j}\in L^p(\mathbb{R}_+;\mathbb{R})$ for $i=1,\dots,d$ and $j=1,\dots,m$, where $(f(t))_{i,j}:=\langle\mathbf{e}_i,f(t)\mathbf{e}_j\rangle$. We make this slightly non-standard choice of notation to emphasize the fact all proofs are carried out component-wise. However this definition is equivalent to say $\|f\|\in L^p(\mathbb{R}_+;\mathbb{R})$ for any chosen norm on $\mathbb{R}^{d\times m}$. Let $M(E;\mathbb{R}^{d\times d})$ denote the space of $d\times d$ matrix valued, finite, signed Borel measures on $E\subset\mathbb{R}$. Given $\nu\in M(E;\mathbb{R}^{d\times d})$ the total variation measure of each component is denoted $|\nu_{ij}|$. The space of continuous functions $f:\mathbb{R}_+\to\mathbb{R}^{d\times m}$ is denoted by $C(\mathbb{R}_+;\mathbb{R}^{d\times m})$. If X is an $\mathbb{R}^d$-valued stochastic process then once again we note the statement $\mathbb{E}[|X_i(\cdot)|^p]\in L^p(\mathbb{R}_+;\mathbb{R})$ for each $i\in\{1,\dots,d\}$ is equivalent to saying, $\mathbb{E}[\|X(\cdot)\|^p]\in L^p(\mathbb{R}_+;\mathbb{R})$ for any chosen norm on $\mathbb{R}^d$. When there is no scope for ambiguity to arise, $\|\cdot\|$ will denote an arbitrary norm on the given finite dimensional vector space under consideration. For scalar valued objects we shall use the notation $|\cdot|$ to represent the standard euclidean norm on $\mathbb{R}$. If $a,b\in\mathbb{R}$, $a\wedge b$ and $a\vee b$ will denote *min* and *max* of the two numbers respectively.

3.3.2 Main Results

Consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ supporting an m-dimensional standard Brownian motion, $(B_t)_{t\geq 0}$. We study the following $\mathbb{R}^d$-valued stochastic integrodifferential Volterra equation given by,

$$dX(t) = \left(f(t) + \int_{[0,t]} \nu(ds)X(t-s)\right)dt + \sigma(t)dB(t), \quad t \geq 0; \quad X(0) = \xi. \tag{3.3.1}$$

Here $f \in C(\mathbb{R}_+; \mathbb{R}^d)$, $\nu \in M(\mathbb{R}_+; \mathbb{R}^{d\times d})$ and $\sigma \in C(\mathbb{R}_+; \mathbb{R}^{d\times m})$. Note as B is a standard m-dimensional Brownian Motion we have $\langle B(t), \mathbf{e}_i\rangle$ and $\langle B(t), \mathbf{e}_j\rangle$ are independent for all $i \neq j$. Theorem 1.2.1 ensures the existence of a unique solution. To see this consider for two arbitrary functions $\phi, \varphi \in C(\mathbb{R}_+; \mathbb{R}^d)$,

$$\left| f(t) + \int_{[0,t]} \nu(ds)\phi(t-s) - f(t) + \int_{[0,t]} \nu(ds)\varphi(t-s)\right| \leq |\nu|(\mathbb{R}_+) \sup_{s\leq t} |\phi(s) - \varphi(s)|.$$

As ν is a finite measure the drift satisfies the global Lipschitz condition in Theorem 1.2.1. Assuming the perturbation functions are continuous is strictly stronger than needed for purely existence and uniqueness results. This can be weakened to $f \in L^1_{loc}(\mathbb{R}_+; \mathbb{R}^d)$ and $\sigma \in L^2_{loc}(\mathbb{R}_+; \mathbb{R}^{d\times m})$ and equation 3.3.1 will still have a unique solution. However to simplify the proofs in this chapter we impose the stronger condition of continuity for the perturbation functions.

We write precisely the conditions on the forcing functions f and σ which we shall be primarily concerned with throughout the rest of this chapter.

$$\text{For } p \in [1,\infty), \quad \int_{\cdot}^{\cdot+\theta} f_i(s)ds \in L^p(\mathbb{R}_+; \mathbb{R}) \quad \forall\, \theta > 0, \quad i \in \{1, \ldots, d\}, \tag{3.3.2}$$

$$\text{For } p \in [2,\infty), \int_{\cdot}^{\cdot+\theta} \sigma_{ij}^2(s)ds \in L^{\frac{p}{2}}(\mathbb{R}_+; \mathbb{R}) \quad \forall\, \theta > 0,\ i \in \{1, \ldots, d\},\ j \in \{1, \ldots, m\}, \tag{3.3.3}$$

$$\text{For } p \in [1,2), \int_{\cdot}^{\cdot+1} \sigma_{ij}^2(s)ds \in \ell^{\frac{p}{2}}(\mathbb{Z}_+; \mathbb{R}), \quad i \in \{1, \ldots, d\},\ j \in \{1, \ldots, m\}. \tag{3.3.4}$$

Next we introduce the continuous time counterpart of (3.2.5). The differential resolvent for (3.3.1) is the $d \times d$ matrix-valued solution to the following equation,

$$\dot{r}(t) = \int_{[0,t]} \nu(ds)r(t-s), \quad t > 0; \quad r(0) = I_{d\times d}. \tag{3.3.5}$$

In the sequel we shall employ the following standing assumption

$$r \in L^1(\mathbb{R}_+; \mathbb{R}^{d\times d}). \tag{3.3.6}$$

For the readers convenience we once again recall a classical characterisation of condition (3.3.6). If $\nu \in M(\mathbb{R}_+;\mathbb{R}^{d\times d})$, then the differential resolvent r of ν satisfies

$$r \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d}) \iff \det[zI_{d\times d} + \int_{\mathbb{R}_+} e^{-zt}\nu(dt)] \neq 0, \quad \mathrm{Re}(z) \geq 0.$$

For a thorough study of equation (3.3.5) we refer the reader to [53, Chapter 3]. With preliminaries dispensed with we can now present our main result.

Theorem 3.3.1. *Let r obey (3.3.6) and X be the solution of (3.3.1). Then we have the following dichotomy,*

*(**A**) For all $p \in [2,\infty)$, following are equivalent.*

- *(i)* $\mathbb{E}[\|X(\cdot)\|^p] \in L^1(\mathbb{R}_+;\mathbb{R})$,
- *(ii)* $X(\cdot) \in L^p(\mathbb{R}_+;\mathbb{R}^d)$ *almost surely,*
- *(iii)* f *obeys* (3.3.2) *and* σ *obeys* (3.3.3).

*(**B**) For all $p \in [1,2)$, following are equivalent.*

- *(i)* $\mathbb{E}[\|X(\cdot)\|^p] \in L^1(\mathbb{R}_+;\mathbb{R})$,
- *(ii)* $X(\cdot) \in L^p(\mathbb{R}_+;\mathbb{R}^d)$ *almost surely,*
- *(iii)* f *obeys* (3.3.2) *and* σ *obeys* (3.3.4).

Proof of Theorem 3.3.1. Theorem 3.3.3 along with Lemma 3.3.1 yield the claim.

□

Before we discuss the proof of Theorem 3.3.1 we mention the assumption that $r \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$ is in a certain sense necessary. By this we mean a converse result of the type given by Theorem 3.9 in [53] also holds in the stochastic case even though its proof is essentially a trivial consequence of the deterministic theorem. Nonetheless we state it here for completeness.

Theorem 3.3.2. *Let X be the solution of (3.3.1). Then the following statements are true,*

*(**A**) For all $p \in [2,\infty)$, following are equivalent.*

- *(i)* $r \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$;
- *(ii) For all $\xi \in \mathbb{R}$, f and σ satisfying (3.3.2) and (3.3.3) respectively, we have* $\mathbb{E}[\|X(\cdot)\|^p] \in L^1(\mathbb{R}_+;\mathbb{R})$;
- *(iii) For all $\xi \in \mathbb{R}$, f and σ satisfying (3.3.2) and (3.3.3) respectively, we have* $X(\cdot) \in L^p(\mathbb{R}_+;\mathbb{R}^d)$ *almost surely.*

(*B*) *For all $p \in [1,2)$, following are equivalent.*

(i) $r \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$;

(ii) *For all $\xi \in \mathbb{R}$, f and σ satisfying* (3.3.2) *and* (3.3.4) *respectively, we have* $\mathbb{E}[\|X(\cdot)\|^p] \in L^1(\mathbb{R}_+;\mathbb{R})$;

(iii) *For all $\xi \in \mathbb{R}$, f and σ satisfying* (3.3.2) *and* (3.3.4) *respectively, we have* $X(\cdot) \in L^p(\mathbb{R}_+;\mathbb{R}^d)$ *almost surely.*

Proof of Theorem 3.3.2. For both statements (**A**) and (**B**), the fact that (i) $\implies$ (ii) & (iii) follows directly from Theorem 3.3.1; hence we need only show (iii) $\implies$ (i). We prove this for (**A**) and (**B**) simultaneously by setting $\sigma = 0$, hence conditions (3.3.3) and (3.3.4) are automatically satisfied. Now fix $f \in L^p(\mathbb{R}_+;\mathbb{R}^d)$, each component of f must also satisfy (3.3.2). This follows from Theorem 3.3.1 with $\nu(dt) = -\delta_0(dt)I_{d\times d}$ where δ_0 is a point mass measure at zero. The claim then follows from [53, Theorem 3.9]. □

Although Theorem 3.3.2 is not entirely satisfactory it seems one cannot do better. One would ideally like to prove that $r \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$ starting from the hypothesis that for all initial conditions $\xi \in \mathbb{R}$ the solution obeyed $X(\cdot) \in L^p(\mathbb{R}_+;\mathbb{R}^d)$ almost surely. However even in the deterministic case (i.e with $\sigma = 0$) such a result is unavailable. One case in which a much stronger result can be proven is that of finite memory; this is achieved by Theorem 3.4.1 in section 3.4. In this situation one relies heavily on an asymptotic expansion for the resolvent which is far more explicit than what one would get in the Volterra setting.

Next we turn to the task of proving Theorem 3.3.1. In order to efficiently study equation (3.3.1) we introduce the simpler equation which is embedded in (3.3.1), namely if one sets $\nu(dt) = -\delta_0(dt)I_{d\times d}$ then we obtain,

$$dY(t) = (f(t) - Y(t))\,dt + \sigma(t)dB(t), \quad t \geq 0; \quad Y(0) = 0. \tag{3.3.7}$$

Next we introduce a lemma which shows the p-integrability of solutions to (3.3.1) is equivalent to the p-integrability of solutions to (3.3.7). This is a very convenient result as (3.3.7) admits an explicit solution and is much more tractable from an analytic point of view.

Lemma 3.3.1. *Let r obey* (3.3.6)*, X be the solution of* (3.3.1) *and Y be the solution of* (3.3.7)*. Then for all $p \in [1,\infty)$,*

(i) $\mathbb{E}[\|X(\cdot)\|^p] \in L^1(\mathbb{R}_+;\mathbb{R}) \iff \mathbb{E}[\|Y(\cdot)\|^p] \in L^1(\mathbb{R}_+;\mathbb{R})$.

(ii) $X(\cdot) \in L^p(\mathbb{R}_+;\mathbb{R}^d)$ *almost surely* $\iff Y(\cdot) \in L^p(\mathbb{R}_+;\mathbb{R}^d)$ *almost surely.*

Remark 3.3.1. *This lemma has deep consequences. Essentially this lemma tells us that the non-Markovianity of the Volterra equation has no bearing whatsoever on whether the* *paths are integrable functions in time; this is purely dictated by the perturbation functions. A priori there seems no reason to suspect this is the case, this result suggests in order for the path dependence to have significant impact on the qualitative behaviour of the solution one needs stronger memory, i.e one could consider instead σ-finite measures for the kernel.*

Proof of Lemma 3.3.1. Suppose $\mathbb{E}[\|X(\cdot)\|^p] \in L^1(\mathbb{R}_+;\mathbb{R})$. Rewrite (3.3.1) as,

$$dX(t) = (f(t) + Q(t) - X(t))\, dt + \sigma(t)dB(t),$$

where $Q(t) = X(t) + \int_{[0,t]} \nu(ds)X(t-s)dt$. For the i^{th} component of Q we have using standard estimates,

$$\mathbb{E}[|Q_i(t)|^p] \leq C\mathbb{E}[|X_i(t)|^p] + C\sum_{j=1}^{d} \left(|\nu_{ij}| * \mathbb{E}[|X_j(\cdot)|^p]\right)(t),$$

where $C > 0$. Then by our supposition and the fact that $|\nu_{ij}|$ is a finite measure, the right hand side must be integrable and so $\mathbb{E}[|Q_i(t)|^p]$ is integrable. Next we write down X in terms of Y and Q. The i^{th} component of $X(t)$ is given by,

$$\begin{aligned}X_i(t) &= \xi_i e^{-t} + \int_0^t e^{-(t-s)}(f_i(s) + Q_i(s))ds + \sum_{j=1}^{m}\int_0^t e^{-(t-s)}\sigma_{ij}(s)dB_j(s)\\ &= \xi_i e^{-t} + \int_0^t e^{-(t-s)}Q_i(s)ds + Y_i(t).\end{aligned}$$

Rearranging yields,

$$Y_i(t) = X_i(t) - \xi_i e^{-t} - \int_0^t e^{-(t-s)}Q_i(s)ds.$$

This representation for $Y_i(t)$ gives the following estimate,

$$\mathbb{E}[|Y_i(t)|^p] \leq C\mathbb{E}[|X_i(t)|^p] + C\mathbb{E}[|\xi_i|^p]e^{-pt} + C\int_0^t e^{-(t-s)}\mathbb{E}[|Q_i(s)|^p]ds,$$

where $C > 0$. The first term on the right hand side is integrable by supposition and the third is integrable as $\mathbb{E}[|Q_i(t)|^p]$ has been shown to be integrable by the estimate above. Hence $\mathbb{E}[|Y_i(\cdot)|^p] \in L^1(\mathbb{R}_+;\mathbb{R})$ and so $\mathbb{E}[\|Y(\cdot)\|^p] \in L^1(\mathbb{R}_+;\mathbb{R})$.

Next we show the reverse implication. Hence suppose $\mathbb{E}[||Y(\cdot)||^p] \in L^1(\mathbb{R}_+;\mathbb{R})$. Introduce the process $Z(t) = X(t) - Y(t)$, then we have $Z \in C^1(\mathbb{R}_+;\mathbb{R}^d)$ and so,

$$\dot{Z}(t) = \int_{[0,t]} \nu(ds)Z(t-s) + g(t),$$

where $g(t) = Y(t) + \int_{[0,t]} \nu(ds)Y(t-s)$. The i^{th} component of Z then satisfies,

$$Z_i(t) = \sum_{j=1}^{d} r_{ij}(t)\xi_j + \sum_{j=1}^{d} \int_0^t r_{ij}(t-s)g_j(s)ds.$$

This yields the estimate,

$$\mathbb{E}[|Z_i(t)|^p] \le C \sum_{j=1}^{d} |r_{ij}(t)|^p \mathbb{E}[|\xi_j|^p] + C \sum_{j=1}^{d} \int_0^t |r_{ij}(t-s)| \mathbb{E}[|g_j(s)|^p] ds,$$

where $C > 0$. Assumption (3.3.6) ensures $r \in L^1(\mathbb{R}_+; \mathbb{R}^{d\times d})$, thus we need only show $\mathbb{E}[|g_j(t)|^p]$ is integrable. Now,

$$\mathbb{E}[|g_i(t)|^p] \le C\mathbb{E}[|Y_i(\cdot)|^p] + C \sum_{j=1}^{d} \left(|\nu_{ij}| * \mathbb{E}[|Y_j(\cdot)|^p]\right)(t),$$

where $C > 0$. Once again as $|\nu_{ij}|$ is a finite measure and $\mathbb{E}[\|Y(\cdot)\|^p]$ is integrable (by supposition), it follows that $\mathbb{E}[|g_i(t)|^p]$ is integrable and so $\mathbb{E}[|Z_i(t)|^p]$ is integrable, thus $\mathbb{E}[\|Z(\cdot)\|^p] \in L^p(\mathbb{R}_+; \mathbb{R})$. Then by the definition of Z, $X(t) = Z(t) + Y(t)$, thus each component satisfies,

$$\int_0^\infty \mathbb{E}[|X_i(t)|^p] dt \le C \int_0^\infty \mathbb{E}[|Z_i(t)|^p] + \mathbb{E}[|Y_i(t)|^p] dt < +\infty,$$

where $C > 0$. Thus $\mathbb{E}[\|X(\cdot)\|^p] \in L^1(\mathbb{R}_+; \mathbb{R})$ as required. The second assertion follows by using the exact same estimates modulo taking expectations. □

With Lemma 3.3.1 at hand we can now devote our attention to the study of the OU type process (3.3.7). The proof of the following theorem is the main piece of analysis throughout this chapter.

Theorem 3.3.3. *Let r obey (3.3.6) and Y be the solution of (3.3.7). Then we have the following dichotomy,*

*(**A**) For all $p \in [2, \infty)$, the following are equivalent.*

(i) $\mathbb{E}[\|Y(\cdot)\|^p] \in L^1(\mathbb{R}_+; \mathbb{R})$,

(ii) $Y(\cdot) \in L^p(\mathbb{R}_+; \mathbb{R}^d)$ almost surely,

(iii) f obeys (3.3.2) and σ obeys (3.3.3).

*(**B**) For all $p \in [1, 2)$, following are equivalent.*

(i) $\mathbb{E}[\|Y(\cdot)\|^p] \in L^1(\mathbb{R}_+; \mathbb{R})$,

(ii) $Y(\cdot) \in L^p(\mathbb{R}_+; \mathbb{R}^d)$ almost surely,

(iii) f obeys (3.3.2) and σ obeys (3.3.4).

Proof of Theorem 3.3.3. **(A)**: $(iii) \implies (i)$. The i^{th} component of Y has representation,

$$Y_i(t) = \int_0^t e^{-(t-s)} f_i(s)ds + \sum_{j=1}^m \int_0^t e^{-(t-s)}\sigma_{ij}(s)dB_j(s).$$

This yields the estimate,

$$\mathbb{E}[|Y_i(t)|^p] \leq C\left|\int_0^t e^{-(t-s)} f_i(s)ds\right|^p + C\sum_{j=1}^m \mathbb{E}\left[\left|\int_0^t e^{-(t-s)}\sigma_{ij}(s)dB_j(s)\right|^p\right],$$

where $C > 0$. As shown in Chapter 2, $\int_0^t e^{-(t-s)} f_i(s)ds \in L^p(\mathbb{R}_+;\mathbb{R})$ if and only if $\int_\cdot^{\cdot+\theta} f_i(s)ds \in L^p(\mathbb{R}_+;\mathbb{R})$ for all $\theta > 0$. Thus we need only focus on the second term. For each $j \in \{1,\dots,m\}$ we have,

$$\frac{\int_0^t e^{-(t-s)}\sigma_{ij}(s)dB_j(s)}{\left(\int_0^t e^{-2(t-s)}\sigma_{ij}^2(s)ds\right)^{1/2}} \sim \mathcal{N}(0,1).$$

Hence we must have

$$\mathbb{E}\left[\left|\int_0^t e^{-(t-s)}\sigma_{ij}(s)dB_j(s)\right|^p\right] = C_p\left(\int_0^t e^{-2(t-s)}\sigma_{ij}^2(s)ds\right)^{p/2},$$

Where C_p is a constant representing the p^{th} moment of a standard normal random variable. Then once again by [12], the left hand side is integrable if and only if $\int_\cdot^{\cdot+\theta} \sigma_{ij}^2(s)ds \in L^{\frac{p}{2}}(\mathbb{R}_+;\mathbb{R})$ for all $\theta > 0$. Thus $\mathbb{E}[|Y_i(t)|^p]$ is integrable and the first claim is proven.

(A): $(i) \implies (ii)$. An application of Fubini's theorem yields the claim.

(A): $(ii) \implies (iii)$. Recall the i^{th} component of Y obeys,

$$dY_i(t) = (f_i(t) - Y_i(t))dt + \sum_{j=1}^m \sigma_{ij}(t)dB_j(t).$$

Integrating this equation over the interval $[t, t+1]$ gives,

$$Y_i(t+1) - Y_i(t) + \int_t^{t+1} Y_i(s)ds = \int_t^{t+1} f_i(s)ds + \sum_{j=1}^m \int_t^{t+1} \sigma_{ij}(s)dB_j(s).$$

Condition (ii) ensures the three terms on the left hand side are elements of $L^p(\mathbb{R}_+;\mathbb{R})$

almost surely. Thus we have,

$$\int_0^\infty \left| \int_t^{t+1} f_i(s)ds + \sum_{j=1}^m \int_t^{t+1} \sigma_{ij}(s)dB_j(s) \right|^p dt < \infty \quad a.s.$$

We break this integral into two series, namely,

$$\sum_{n=0}^\infty \int_{2n}^{2n+1} \left| \int_t^{t+1} f_i(s)ds + \sum_{j=1}^m \int_t^{t+1} \sigma_{ij}(s)dB_j(s) \right|^p dt + \\ \sum_{n=0}^\infty \int_{2n+1}^{2n+2} \left| \int_t^{t+1} f_i(s)ds + \sum_{j=1}^m \int_t^{t+1} \sigma_{ij}(s)dB_j(s) \right|^p dt < \infty \quad a.s.$$

We estimate these series below by,

$$\sum_{n=0}^\infty \left| \int_{2n}^{2n+1} \int_t^{t+1} f_i(s)dsdt + \sum_{j=1}^m \int_{2n}^{2n+1} \int_t^{t+1} \sigma_{ij}(s)dB_j(s)dt \right|^p + \\ \sum_{n=0}^\infty \left| \int_{2n+1}^{2n+2} \int_t^{t+1} f_i(s)dsdt + \sum_{j=1}^m \int_{2n+1}^{2n+2} \int_t^{t+1} \sigma_{ij}(s)dB_j(s)dt \right|^p .$$

Now introduce the following notation,

$$f_{i,1}(n) := \int_{2n}^{2n+1} \int_t^{t+1} f_i(s)dsdt; \quad f_{i,2}(n) := \int_{2n+1}^{2n+2} \int_t^{t+1} f_i(s)dsdt$$
$$u_{i,j}(n) := \int_{2n}^{2n+1} \int_t^{t+1} \sigma_{ij}(s)dB_j(s)dt; \quad v_{i,j}(n) := \int_{2n+1}^{2n+2} \int_t^{t+1} \sigma_{ij}(s)dB_j(s)dt.$$

So our two series become,

$$\sum_{n=0}^\infty \left| f_{i,1}(n) + \sum_{j=1}^m u_{i,j}(n) \right|^p + \sum_{n=0}^\infty \left| f_{i,2}(n) + \sum_{j=1}^m v_{i,j}(n) \right|^p < +\infty \quad a.s.$$

An application of the stochastic Fubini theorem yields the following representations for $u_{i,j}(n)$ and $v_{i,j}(n)$,

$$u_{i,j}(n) = \int_{2n}^{2n+2} \left[s \wedge (2n+1) - 2n \vee (s-1) \right] \sigma_{ij}(s)dB_j(s),$$
$$v_{i,j}(n) = \int_{2n+1}^{2n+3} \left[s \wedge (2n+2) - (2n+1) \vee (s-1) \right] \sigma_{ij}(s)dB_j(s).$$

Recall $\wedge, \vee$ denote the min and max respectively. We note that $u_{i,j}(n)$ and $v_{i,j}(n)$ are sequences of independent Gaussian random variables (with mean zero), this is because each member of the sequences are Wiener integrals where the domains of integration are non-overlapping. It is a general fact that such integrals are independent. Furthermore we can define,

$$U_i(n) := \sum_{j=1}^{m} u_{i,j}(n); \quad V_i(n) := \sum_{j=1}^{m} v_{i,j}(n).$$

Now $U_i(n)$ and $V_i(n)$ are Gaussian random variables which follows from the independence of the components of the m-dimensional Brownian Motion. These can be rewritten as follows[4]

$$\begin{aligned} U_i(n) &= \sqrt{\mathbb{E}[U_i(n)^2]}\frac{U_i(n)}{\sqrt{\mathbb{E}[U_i(n)^2]}} = \sqrt{\mathbb{E}[U_i(n)^2]}\tilde{U}_i(n), \\ V_i(n) &= \sqrt{\mathbb{E}[V_i(n)^2]}\frac{V_i(n)}{\sqrt{\mathbb{E}[V_i(n)^2]}} = \sqrt{\mathbb{E}[V_i(n)^2]}\tilde{V}_i(n). \end{aligned}$$

Now $\tilde{U}_i(n)$ and $\tilde{V}_i(n)$ are i.i.d sequences of standard normal random variables. This yields the final form of our two series,

$$\sum_{n=0}^{\infty}\left|f_{i,1}(n) + \sqrt{\mathbb{E}[U_i(n)^2]}\tilde{U}_i(n)\right|^p + \sum_{n=0}^{\infty}\left|f_{i,2}(n) + \sqrt{\mathbb{E}[V_i(n)^2]}\tilde{V}_i(n)\right|^p < +\infty \quad a.s. \quad (3.3.8)$$

Now we can invoke Lemma 3.2.1 to claim,

$$\sum_{n=0}^{\infty}\left|\sqrt{\mathbb{E}[U_i(n)^2]}\right|^p + \sum_{n=0}^{\infty}\left|\sqrt{\mathbb{E}[V_i(n)^2]}\right|^p < +\infty.$$

Next we take a closer look at both summands, as U_i and V_i are linear combinations of independent normal random variables (with zero mean) it follows that,

$$\mathbb{E}[U_i(n)^2] = \sum_{j=1}^{m}\mathbb{E}[u_{i,j}(n)^2]; \quad \mathbb{E}[V_i(n)^2] = \sum_{j=1}^{m}\mathbb{E}[v_{i,j}(n)^2]$$

[4] As in the discrete case we need not worry that $\mathbb{E}[U_i(n)^2] = 0$ or $\mathbb{E}[U_i(n)^2] = 0$ for some n, in this case we define $\tilde{U}_i(n) := U_i(n)$ and $\tilde{V}_i(n) = V_i(n)$ for such values of n. If however both $\mathbb{E}[U_i(n)^2] = 0$ and $\mathbb{E}[V_i(n)^2] = 0$ for all n then care in needed. Using the identity (3.3.8), Itô's isometry will force $\sigma_{ij}(t) = 0$ for $t \in [2n, 2n+2]$. But this holds for all n and so it forces $\sigma_{ij}(t) = 0$ for all $t \geq 0$ and so the desired integrability condition will be trivially satisfied.

Now as norms on finite dimensional vector spaces are equivalent, we have,

$$\sum_{n=0}^{\infty}\left|\sqrt{\mathbb{E}[U_i(n)^2]}\right|^p = \sum_{n=0}^{\infty}\left(\sum_{j=1}^{m}\mathbb{E}[u_{i,j}(n)^2]\right)^{\frac{p}{2}} \geq C_1\sum_{n=0}^{\infty}\left(\sum_{j=1}^{m}\left|\mathbb{E}[u_{i,j}(n)^2]^{\frac{1}{2}}\right|\right)^p$$
$$\sum_{n=0}^{\infty}\left|\sqrt{\mathbb{E}[V_i(n)^2]}\right|^p = \sum_{n=0}^{\infty}\left(\sum_{j=1}^{m}\mathbb{E}[v_{i,j}(n)^2]\right)^{\frac{p}{2}} \geq C_2\sum_{n=0}^{\infty}\left(\sum_{j=1}^{m}\left|\mathbb{E}[v_{i,j}(n)^2]^{\frac{1}{2}}\right|\right)^p,$$

with $C_1, C_2 > 0$. Furthermore as $p \geq 1$ and both $\mathbb{E}[u_{i,j}(n)^2]$ and $\mathbb{E}[v_{i,j}(n)^2]$ are non-negative, we have,

$$+\infty > \sum_{n=0}^{\infty}\left|\sqrt{\mathbb{E}[U_i(n)^2]}\right|^p \geq C_1\sum_{n=0}^{\infty}\left(\sum_{j=1}^{m}\mathbb{E}[u_{i,j}(n)^2]^{\frac{1}{2}}\right)^p \geq C_1\sum_{n=0}^{\infty}\sum_{j=1}^{m}\left(\mathbb{E}[u_{i,j}(n)^2]\right)^{\frac{p}{2}}$$
$$+\infty > \sum_{n=0}^{\infty}\left|\sqrt{\mathbb{E}[V_i(n)^2]}\right|^p \geq C_2\sum_{n=0}^{\infty}\left(\sum_{j=1}^{m}\mathbb{E}[v_{i,j}(n)^2]^{\frac{1}{2}}\right)^p \geq C_2\sum_{n=0}^{\infty}\sum_{j=1}^{m}\left(\mathbb{E}[v_{i,j}(n)^2]\right)^{\frac{p}{2}}.$$

Thus it must be the case that,

$$\sum_{n=0}^{\infty}\left(\mathbb{E}[u_{i,j}(n)^2]\right)^{\frac{p}{2}} + \sum_{n=0}^{\infty}\left(\mathbb{E}[v_{i,j}(n)^2]\right)^{\frac{p}{2}} < +\infty,$$

for each $j \in \{1, \ldots, m\}$. Now making use of Itô's isometry, we have for each j,

$$\sum_{n=0}^{\infty}\left(\mathbb{E}[u_{i,j}(n)^2]\right)^{\frac{p}{2}} = \sum_{n=0}^{\infty}\left(\int_{2n}^{2n+2}\left[s\wedge(2n+1) - 2n\vee(s-1)\right]^2\sigma_{ij}^2(s)ds\right)^{\frac{p}{2}} < +\infty$$
$$\sum_{n=0}^{\infty}\left(\mathbb{E}[v_{i,j}(n)^2]\right)^{\frac{p}{2}} = \sum_{n=0}^{\infty}\left(\int_{2n+1}^{2n+3}\left[s\wedge(2n+2) - (2n+1)\vee(s-1)\right]^2\sigma_{ij}^2(s)ds\right)^{\frac{p}{2}} < +\infty$$

Next we obtain a lower estimate for the summands for both series.

$$\begin{aligned}\mathbb{E}[u_{i,j}(n)^2] &= \int_{2n}^{2n+2}\left[s\wedge(2n+1) - 2n\vee(s-1)\right]^2\sigma_{ij}^2(s)ds \\ &= \int_{2n}^{2n+1}(s-2n)^2\sigma_{ij}^2(s)ds + \int_{2n+1}^{2n+2}(2n+2-s)^2\sigma_{ij}^2(s)ds \\ &\geq \int_{2n+\frac{1}{2}}^{2n+1}\frac{\sigma_{ij}^2(s)}{4}ds + \int_{2n+1}^{2n+\frac{3}{2}}\frac{\sigma_{ij}^2(s)}{4}ds.\end{aligned}$$

Similarly we have,

$$\begin{aligned}\mathbb{E}[v_{i,j}(n)^2] &= \int_{2n+1}^{2n+3} \left[s \wedge (2n+2) - (2n+1) \vee (s-1)\right]^2 \sigma_{ij}^2(s)ds \\ &= \int_{2n+1}^{2n+2} (s-2n-1)^2 \sigma_{ij}^2(s)ds + \int_{2n+2}^{2n+3} (2n+3-s)^2 \sigma_{ij}^2(s)ds \\ &\geq \int_{2n+\frac{3}{2}}^{2n+2} \frac{\sigma_{ij}^2(s)}{4} ds + \int_{2n+2}^{2n+\frac{5}{2}} \frac{\sigma_{ij}^2(s)}{4} ds.\end{aligned}$$

Now consider,

$$\frac{1}{C}\left|\int_{2n}^{2n+2} \sigma_{ij}^2(s)ds\right|^{p/2} \leq \left|\int_{2n}^{2n+\frac{1}{2}} \sigma_{ij}^2(s)ds\right|^{p/2} + \left|\int_{2n+\frac{1}{2}}^{2n+\frac{3}{2}} \sigma_{ij}^2(s)ds\right|^{p/2} + \left|\int_{2n+\frac{3}{2}}^{2n+2} \sigma_{ij}^2(s)ds\right|^{p/2},$$

where $C > 0$. The last two terms on the right hand side are summable by the above estimates on $\mathbb{E}[u_{i,j}(n)^2]$ and $\mathbb{E}[v_{i,j}(n)^2]$. To see that the first term is summable consider,

$$\sum_{n=0}^{\infty}\left|\int_{2n+2}^{2n+\frac{5}{2}} \sigma_{ij}^2(s)ds\right|^{p/2}.$$

We know this sum is finite by virtue of the above estimates on $\mathbb{E}[v_{i,j}(n)^2]$. Thus making the change of indexing variable $n = l - 1$, yields,

$$\sum_{l=1}^{\infty}\left|\int_{2l}^{2l+\frac{1}{2}} \sigma_{ij}^2(s)ds\right|^{p/2}.$$

Hence we have,

$$\sum_{n=0}^{\infty}\left|\int_{2n}^{2n+2} \sigma_{ij}^2(s)ds\right|^{p/2} < +\infty. \tag{†}$$

Finally we consider,

$$\begin{aligned}\int_0^{\infty}\left|\int_t^{t+1} \sigma_{ij}^2(s)ds\right|^{p/2} dt &= \sum_{n=0}^{\infty}\int_{2n}^{2n+2}\left|\int_t^{t+1} \sigma_{ij}^2(s)ds\right|^{p/2} dt \\ &\leq \sum_{n=0}^{\infty}\int_{2n}^{2n+2}\left|\int_{2n}^{2n+3} \sigma_{ij}^2(s)ds\right|^{p/2} dt \\ &= 2\sum_{n=0}^{\infty}\left|\int_{2n}^{2n+3} \sigma_{ij}^2(s)ds\right|^{p/2} \\ &\leq C\sum_{n=0}^{\infty}\left|\int_{2n}^{2n+2} \sigma_{ij}^2(s)ds\right|^{p/2} + C\sum_{n=0}^{\infty}\left|\int_{2n+2}^{2n+3} \sigma_{ij}^2(s)ds\right|^{p/2},\end{aligned}$$

where $C > 0$. We have just proven the first term is finite which also implies the second is finite after a change of indexing variable. Thus we have shown,

$$\int_{\cdot}^{\cdot+1} \sigma_{ij}^2(s)ds \in L^{\frac{p}{2}}(\mathbb{R}_+;\mathbb{R}).$$

We note that this can easily be shown to be equivalent to $\int_{\cdot}^{\cdot+\theta} \sigma_{ij}^2(s)ds \in L^{\frac{p}{2}}(\mathbb{R}_+;\mathbb{R})$ for all $\theta > 0$. As i and j were arbitrary throughout this argument, this holds for all components of the matrix σ. Recall that,

$$\int_{\cdot}^{\cdot+\theta} \sigma_{ij}^2(s)ds \in L^{\frac{p}{2}}(\mathbb{R}_+;\mathbb{R}) \text{ for all } \theta > 0 \iff \mathbb{E}\left[\left|\int_0^{\cdot} e^{-(\cdot-s)}\sigma_{ij}(s)dB_j(s)\right|^p\right] \in L^1(\mathbb{R}_+;\mathbb{R}).$$

Thus it must be the case that,

$$\int_0^\infty \left|\int_0^t e^{-(t-s)}\sigma_{ij}(s)dB_j(s)\right|^p dt < +\infty \quad a.s. \tag{$\dagger'$}$$

We now recall that,

$$\int_0^t e^{-(t-s)} f_i(s)ds = Y_i(t) - \sum_{j=1}^m \int_0^t e^{-(t-s)}\sigma_{ij}(s)dB_j(s).$$

As all terms on the right are elements of $L^p(\mathbb{R}_+;\mathbb{R})$ almost surely it follows the left hand side is also an element of $L^p(\mathbb{R}_+;\mathbb{R})$. But once again by the results from Chapter 2 this is equivalent to,

$$\int_{\cdot}^{\cdot+\theta} f_i(s)ds \in L^p(\mathbb{R}_+;\mathbb{R}) \text{ for all } \theta > 0.$$

As this hold for all $i \in \{1,\dots,d\}$, the proof of **(A)** is complete.

(B): $(iii) \implies (i)$. Follow the proof from **(A)** and replace the L^p condition on $\int_t^{t+\theta} \sigma_{ij}^2(s)ds$ with the summability condition $\int_n^{n+1} \sigma_{ij}^2(s)ds \in \ell^{p/2}$ and invoke Lemma 3.3.2 (stated directly after this proof).

(B): $(i) \implies (ii)$. Follow the proof from **(A)**.

(B): $(ii) \implies (iii)$. Follow the proof from **(A)** up until $(\dagger)$. Thus we have that

$$\sum_{n=0}^\infty \left|\int_n^{n+1} \sigma_{ij}^2(s)ds\right|^{\frac{p}{2}} < +\infty.$$

Recall by Lemma 3.3.2, we have that,

$$\int_{\cdot}^{\cdot+\theta} \sigma_{ij}^2(s)ds \in \ell^{\frac{p}{2}}(\mathbb{Z}_+;\mathbb{R}) \text{ for all } \theta > 0 \iff \mathbb{E}\left[\left|\int_0^{\cdot} e^{-(\cdot-s)}\sigma_{ij}(s)dB_j(s)\right|^p\right] \in L^1(\mathbb{R}_+;\mathbb{R}).$$

Then follow the proof of **(A)** from ($\dagger'$) onwards.

$\square$

Lemma 3.3.2. *Let f be a non-negative continuous function and $\beta > 0$. Then for $p \in (0,1)$, the following are equivalent,*

(i) $t \mapsto \int_0^t e^{-\beta(t-s)} f(s)ds \in L^p(\mathbb{R}_+;\mathbb{R})$.

(ii) $\int_n^{n+1} f(s)ds \in \ell^p(\mathbb{Z}_+;\mathbb{R})$.

Proof of Lemma 3.3.2. $(ii) \implies (i)$: Define $v(t) := \int_0^t e^{-\beta(t-s)} f(s)ds$. From the definition of v, we have for all $t \in [n, n+1]$,

$$v(t) = v(n)e^{-\beta(t-n)} + \int_n^t e^{-\beta(t-s)} f(s)ds. \tag{3.3.9}$$

Thus we have the estimate (noting that v is non-negative),

$$v(t) \le v(n) + \int_n^{n+1} f(s)ds.$$

Raising both sides to the power of p and integrating, we obtain,

$$\int_n^{n+1} v(t)^p dt \le \int_n^{n+1} \left(v(n) + \int_n^{n+1} f(s)ds \right)^p dt = \left(v(n) + \int_n^{n+1} f(s)ds \right)^p.$$

Summing over n,

$$\int_0^\infty v(t)^p dt \le \sum_{n=0}^\infty v(n)^p + \sum_{n=0}^\infty \left(\int_n^{n+1} f(s)ds \right)^p.$$

The second series is finite by assumption. Thus we need only show $v \in \ell^p(\mathbb{Z}_+;\mathbb{R})$. Setting $t = n+1$ in (3.3.9) and estimating yields,

$$v(n+1) \le v(n)e^{-\beta} + \int_n^{n+1} f(s)ds.$$

Thus we can estimate the solution to this difference inequality by

$$v(n) \le \sum_{j=1}^n e^{-\beta(n-j)} \left(\int_{j-1}^j f(s)ds \right).$$

Thus, as $p \in (0,1)$,

$$v(n)^p \le \sum_{j=1}^n e^{-\beta p(n-j)} \left(\int_{j-1}^j f(s)ds \right)^p.$$

The right hand side is a convolution of two summable sequences and is thus summable.

Hence $v(n)^p \in \ell^1(\mathbb{Z}_+;\mathbb{R})$ and we are done.

$(i) \implies (ii)$: Returning once again to (3.3.9), we get the estimate,

$$v(t) \geq v(n)e^{-\beta(t-n)} \geq v(n)e^{-\beta}.$$

Thus,

$$\int_n^{n+1} v(t)^p dt \geq v(n)^p e^{-\beta p}.$$

Summing we obtain,

$$+\infty > \int_0^\infty v(t)^p dt \geq \sum_{n=0}^\infty v(n)^p e^{-\beta p}.$$

Hence $v(n) \in \ell^p(\mathbb{Z}_+;\mathbb{R})$. But once again with $t = n+1$ in (3.3.9) we obtain,

$$\begin{aligned} v(n+1) &= v(n)e^{-\beta} + e^{-\beta(n+1)} \int_n^{n+1} e^{\beta s} f(s)ds \\ &\geq e^{-\beta} \int_n^{n+1} f(s)ds. \end{aligned}$$

Thus we must have $\int_n^{n+1} f(s)ds \in \ell^p(\mathbb{Z}_+;\mathbb{R})$. The claim is proven.

□

3.3.3 Almost sure asymptotic behaviour

When within the regime of $p-$integrable sample paths of solutions to (3.3.1), one can easily show that the asymptotic behaviour of the sample paths are almost surely determined by the asymptotic behaviour of the underlying deterministic equation.

Theorem 3.3.4. *Let X be the solution of (3.3.1) and $r \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$ be the solution of (3.3.5). Assume $X \in L^p(\mathbb{R}_+;\mathbb{R}^d)$ a.s. for some $p \geq 1$. Then,*

$$\|X(t) - (r * f)(t)\| \longrightarrow 0 \text{ as } t \to \infty \quad a.s. \tag{3.3.10}$$

This result picks up precisely the almost sure asymptotic behaviour of the paths but in terms of the resolvent r, which could be seen as unsatisfactory. In general in this regime the asymptotic behaviour of the mapping, $t \mapsto (r * f)(t)$ cannot easily be determined. If one wishes for solutions to vanish almost surely, we of course need $(r * f)(t) \to 0$ as $t \to \infty$ which as outlined in Chapter 2, is equivalent to $t \mapsto \int_t^{t+\theta} f(s)ds \to 0$ as $t \to \infty$ for all $\theta > 0$, however it is not evident that this will be the case. In the special case when $X \in L^1(\mathbb{R}_+;\mathbb{R}^d)$ almost surely, one can remove the dependence on the resolvent in (3.3.10) and obtain the almost sure asymptotic behaviour of the trajectories purely in terms of the perturbation function f.

Corollary 3.3.1. *Let X be the solution of (3.3.1) and $r \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$ be the solution of (3.3.5). Assume $X \in L^1(\mathbb{R}_+;\mathbb{R}^d)$ a.s., then,*

$$\left\| X(t) - \int_0^1 \int_{t-u}^t f(s)dsdu \right\| \longrightarrow 0 \text{ as } t \to \infty \quad a.s., \tag{3.3.11}$$

where $\int_0^1 \int_{t-u}^t f(s)dsdu$ is a d-dimensional vector whose i^{th} entry is given by $\int_0^1 \int_{t-u}^t f_i(s)dsdu$.

Proof of corollary 3.3.1. By Theorem 3.3.4 we have that,

$$\|X(t) - (r * f)(t)\| \longrightarrow 0 \text{ as } t \to \infty \quad a.s. \tag{3.3.12}$$

so we need only focus on the asymptotic behaviour of $(r * f)(t)$. Now by Theorem 3.3.1 we know $\int_{\cdot}^{\cdot+\theta} f_i(s)ds \in L^1(\mathbb{R}_+;\mathbb{R})$ for all $\theta > 0$, $i \in \{1,\dots,d\}$. Hence we may apply Lemma 2.3.2 in Chapter 2 to claim $f_i = f_{i,1} + f_{i,2}$ where $f_{i,1} \in L^1(\mathbb{R}_+;\mathbb{R})$ and $\int_0^t f_{i,2}(s)ds \in L^1(\mathbb{R}_+;\mathbb{R})$. Thus we may write the vector valued function $f = f_1 + f_2$ where f_1 and f_2 are defined in the obvious way. Now with $f_3(t) := \int_0^t f_2(s)ds$ we follow the proof of Theorem 2.3.1 from Chapter 2 to obtain,

$$\begin{aligned}(r * f)(t) &= (r * f_1)(t) + f_3(t) + (r' * f_3)(t) \\ &= (r * f_1)(t) + f_3(t) + (\nu * r * f_3)(t),\end{aligned} \tag{3.3.13}$$

where the last equality follows from equation (3.3.5). Now as $r \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$ we know also that $r \in BC_0(\mathbb{R}_+;\mathbb{R}^{d\times d})$ and so $r * f_1 \in BC_0(\mathbb{R}_+;\mathbb{R}^d)$. Additionally as $\nu \in M(\mathbb{R}_+;\mathbb{R}^{d\times d})$ and $\nu * r \in BC_0(\mathbb{R}_+;\mathbb{R}^{d\times d})$ we have $\nu * r * f_3 \in BC_0(\mathbb{R}_+;\mathbb{R}^d)$. Thus, we have,

$$\|X(t) - f_3(t)\| \le \|X(t) - (r * f)(t)\| + \|(r * f_1)(t)\| + \|(\nu * r * f_3)(t)\|,$$

Thus we have $\|X(t) - f_3(t)\| \to 0$ as $t \to \infty$. Finally the explicit representation of f_3 is given in the proof of Lemma 2.3.2 in Chapter 2, namely,

$$f_3(t) = \int_0^1 \int_{t-u}^t f(s)dsdu.$$

The claim is proven. □

In the introduction we discussed that often times a strategy to infer information about asymptotic behaviour of solutions is to first prove some kind of integrability result and then infer convergence to zero. We will use Theorem 3.3.4 and corollary 3.3.1 to highlight how this line of attack may yield severely sub optimal results for even the simplest equations. We shall momentarily kill the deterministic perturbation f and consider only the scalar equation,

$$dY(t) = -Y(t)dt + \sigma(t)dB(t); \quad Y(0) = 0. \tag{3.3.14}$$

Now $\sigma \in C(\mathbb{R}_+;\mathbb{R})$ and B is a scalar Brownian motion. Theorem 3.3.4 ensures the condition $n \mapsto \int_n^{n+1} \sigma^2(s)ds \in \ell^{1/2}(\mathbb{Z}_+;\mathbb{R})$ is enough to force $|Y(t)| \to 0$ almost surely as $t \to \infty$. However it was shown in [6] that Y converging to zero almost surely is equivalent to,

$$\sum_{n=0}^{\infty} \sqrt{\int_n^{n+1} \sigma^2(s)ds} \exp\left(\frac{-\varepsilon}{\int_n^{n+1} \sigma^2(s)ds}\right) < +\infty, \tag{3.3.15}$$

for all $\varepsilon > 0$. But,

$$\sum_{n=0}^{\infty} \sqrt{\int_n^{n+1} \sigma^2(s)ds} \exp\left(\frac{-\varepsilon}{\int_n^{n+1} \sigma^2(s)ds}\right) \le \sum_{n=0}^{\infty} \sqrt{\int_n^{n+1} \sigma^2(s)ds},$$

uniformly in epsilon, so if one thinks about convergence to zero in terms of condition (3.3.15), Y being almost surely integrable is most definitely sufficient but nowhere near necessary. If we add back in the deterministic perturbation f, the situation is even worse, Theorem 3.3.4 and corollary 3.3.1 demonstrate integrability is no longer even sufficient and in fact one needs to impose side conditions on f to ensure convergence. Next we provide such a convergence characterisation in the special case when σ is a diagonal matrix.

Theorem 3.3.5. *Let r obey (3.3.6), X be the solution of (3.3.1), $m = d$ and σ be a diagonal matrix. Then the following are equivalent,*

(i) Each component of f and σ satisfy,

$$\int_t^{t+\theta} f_i(s)ds \xrightarrow{t\to\infty} 0 \text{ for all } \theta > 0; \ \sum_{n=0}^{\infty} \sqrt{\int_n^{n+1} \sigma_{ii}^2(s)ds} \exp\left(\frac{-\varepsilon}{\int_n^{n+1} \sigma_{ii}^2(s)ds}\right) < +\infty \tag{3.3.16}$$

for all $\varepsilon > 0$.

(ii) $\|X(t)\| \longrightarrow 0$ as $t \to \infty$, almost surely.

Proof of Theorem 3.3.5. The process satisfying $dY(t) = (f(t) - Y(t))dt + \sigma(t)dB(t)$ is now just a concatenation of scalar processes in which almost sure convergence to zero for each component is equivalent to (i) [13, Theorem 11]. Now as in the proof of Lemma 3.3.1 consider the process $Z(t) = X(t) - Y(t)$ and use the fact $Y \in BC_0(\mathbb{R}_+;\mathbb{R}^d)\, a.s.$ and $\nu \in M(\mathbb{R}_+;\mathbb{R}^{d\times d})$ to conclude $Z \in BC_0(\mathbb{R}_+;\mathbb{R}^d)$. As $X(t) = Z(t) + Y(t)$ we have $X \in BC_0(\mathbb{R}_+;\mathbb{R}^d)\, a.s.$ and the forward implication is proven. Now we need only show $X \in BC_0(\mathbb{R}_+;\mathbb{R}^d) \quad a.s. \implies Y \in BC_0(\mathbb{R}_+;\mathbb{R}^d) \quad a.s.$, after which we can conclude by applying Theorem 11 in [13]. We once again continue as in the proof of Lemma 3.3.1 and consider

$$dX(t) = (f(t) + Q(t) - X(t))\, dt + \sigma(t)dB(t),$$

where $Q(t) = X(t) + \int_{[0,t]} \nu(ds)X(t-s)dt$. Now as $\nu \in M(\mathbb{R}_+; \mathbb{R}^{d\times d})$ and $X \in BC_0(\mathbb{R}_+; \mathbb{R}^d)$ $a.s.$ we must have $Q \in BC_0(\mathbb{R}_+; \mathbb{R}^d)$ $\quad a.s.$. Next we obtain for each component,

$$Y_i(t) = X_i(t) - \xi_i e^{-t} - \int_0^t e^{-(t-s)} Q_i(s)ds$$

from which it is clear $Y_i \in BC_0(\mathbb{R}_+; \mathbb{R})$ $\quad a.s.$ for each component, and thus the theorem is proven.

□

We close off this section with a proof of Theorem 3.3.4 in which we shall make use of the following lemma.

Lemma 3.3.3. *Let $f \in L^1_{loc}(\mathbb{R}_+; \mathbb{R}_+)$ and assume for $p \geq 1$ that,*

$$\int_0^\infty \left(\int_t^{t+\theta} f(s)ds \right)^p dt < +\infty \quad \text{for all } \theta > 0.$$

Then for each sequence $(a_n)_{n\geq 0}$ with $a_0 = 0$ and $0 < \alpha \leq a_{n+1} - a_n \leq \beta$ where $\alpha, \beta > 0$, we have,

$$\sum_{n=0}^\infty \left(\int_{a_n}^{a_{n+1}} f(s)ds \right)^p < +\infty.$$

Proof of Theorem 3.3.4. Introduce the $\mathbb{R}^d$-valued process with dynamics $dM(t) = -M(t)dt + \sigma(t)dB(t)$ with zero initial condition. By Theorem 3.3.1 and Lemma 3.3.1, $X \in L^p(\mathbb{R}_+; \mathbb{R}^d)$ forces $M \in L^p(\mathbb{R}_+; \mathbb{R}^d)$ and in particular $\mathbb{E}\left[|M(t)|^p\right] \in L^1(\mathbb{R}_+; \mathbb{R}^d)$. As $t \mapsto \mathbb{E}\left[|M_i(t)|^p\right]$ is continuous and integrable, for each $i \in \{1, \cdots, d\}$ there exists a sequence $a_n \nearrow \infty$ with $0 \leq a_{n+1} - a_n \leq 1$ such that $\sum_{n=0}^\infty \mathbb{E}\left[|M_i(a_n)|^p\right] < +\infty$. To see this note that we have,

$$\begin{aligned} +\infty > & \int_0^\infty \mathbb{E}\left[|M_j(s)|^p\right] ds \\ & = \sum_{n=0}^\infty \int_n^{n+1} \mathbb{E}\left[|M_j(s)|^p\right] ds \geq \sum_{n=0}^\infty \inf_{s\in[n,n+1]} \mathbb{E}\left[|M_j(s)|^p\right] = \sum_{n=0}^\infty \mathbb{E}\left[|M_j(a_n)|^p\right], \end{aligned}$$

where for each $n \in \mathbb{Z}_+$,

$$a_n := \min \left\{ x \in [n, n+1] : \mathbb{E}\left[|M_j(x)|^p\right] = \inf_{s\in[n,n+1]} \mathbb{E}\left[|M_j(s)|^p\right] \right\}.$$

The continuity of $t \mapsto \mathbb{E}\left[|M_j(t)|^p\right]$ ensures the infimum over compact intervals is always attained and by construction we have $0 \leq a_{n+1} - a_n \leq 1$ and $a_n \nearrow \infty$.

Now for $t \in [a_n, a_{n+1}]$ each component satisfies,

$$M_i(t)e^t = e^t M_i(a_n) + \sum_{j=1}^{m} \int_{a_n}^{t} e^s \sigma_{ij}(s) dB_j(s).$$

Using standard estimates we obtain the following for some $C > 0$,

$$\mathbb{E}\left[\sup_{a_n \le t \le a_{n+1}} \left|e^t M_i(t)\right|^p\right] \le C e^{p a_{n+1}} \mathbb{E}[|M_i(a_n)|^p] + C \sum_{j=1}^{m} \mathbb{E}\left[\sup_{a_n \le t \le a_{n+1}} \left|\int_{a_n}^{t} e^s \sigma_{ij}(s) dB_j(s)\right|^p\right]. \tag{$*$}$$

Now we can apply the Burkholder-Davis-Gundy Inequality to each term in the finite sum on the right hand side, this yields for some $C > 0$,

$$\begin{aligned}\mathbb{E}\left[\sup_{a_n \le t \le a_{n+1}} \left|\int_{a_n}^{t} e^s \sigma_{ij}(s) dB_j(s)\right|^p\right] &\le C \left(\int_{a_n}^{a_{n+1}} e^{2s} \sigma_{ij}^2(s) ds\right)^{p/2} \\ &\le C e^{p(a_{n+1})} \left(\int_{a_n}^{a_{n+1}} \sigma_{ij}^2(s) ds\right)^{p/2}.\end{aligned}$$

Plugging this estimate back into ($*$) we obtain for some new constant $C' > 0$,

$$\mathbb{E}\left[\sup_{a_n \le t \le a_{n+1}} \left|e^t M_i(t)\right|^p\right] \le C' e^{p(a_{n+1})} \mathbb{E}[|M_i(a_n)|^p] + C' e^{p(a_{n+1})} \left(\int_{a_n}^{a_{n+1}} \sigma_{ij}^2(s) ds\right)^{p/2}.$$

Dividing across by $e^{p(a_{n+1})}$ on both sides we see,

$$e^{-p(a_{n+1}-a_n)} \mathbb{E}\left[\sup_{a_n \le t \le a_{n+1}} |M_i(t)|^p\right] \le C' \mathbb{E}[|M_i(a_n)|^p] + C' \left(\int_{a_n}^{a_{n+1}} \sigma_{ij}^2(s) ds\right)^{p/2}.$$

Now the first term on the right hand side is summable by construction of the sequence a_n. To show the second term is summable we need to distinguish two cases; the first case is when $p \in [1, 2)$. Note by the construction of the sequence a_n we have,

$$\left(\int_{a_n}^{a_{n+1}} \sigma_{ij}^2(s) ds\right)^{p/2} \le \left(\int_{n}^{n+1} \sigma_{ij}^2(s) ds + \int_{n+1}^{n+2} \sigma_{ij}^2(s) ds\right)^{p/2},$$

and once again we use the fact that norms on finite dimensional vector spaces are equivalent to obtain the bound for some $C > 0$,

$$\left(\int_{n}^{n+1} \sigma_{ij}^2(s) ds + \int_{n+1}^{n+2} \sigma_{ij}^2(s) ds\right)^{p/2} \le C^{1/p} \left(\left(\int_{n}^{n+1} \sigma_{ij}^2(s) ds\right)^{1/2} + \left(\int_{n+1}^{n+2} \sigma_{ij}^2(s) ds\right)^{1/2}\right)^{p}.$$

Now as $p \geq 1$ we use a standard estimate to obtain for some new constant $C' > 0$,

$$\left(\int_{a_n}^{a_{n+1}} \sigma_{ij}^2(s)ds\right)^{p/2} \leq C'\left(\int_{n}^{n+1} \sigma_{ij}^2(s)ds\right)^{p/2} + C'\left(\int_{n+1}^{n+2} \sigma_{ij}^2(s)ds\right)^{p/2}.$$

Both terms on the right are summable by virtue of Theorem 3.3.1. For the case when $p \geq 2$ we invoke Theorem 3.3.1 and apply Lemma 3.3.3. Thus we have that

$$+\infty > \sum_{n=0}^{\infty} e^{-p(a_{n+1}-a_n)}\mathbb{E}\left[\sup_{a_n \leq t \leq a_{n+1}} |M_i(t)|^p\right] \geq e^{-p}\sum_{n=0}^{\infty}\mathbb{E}\left[\sup_{a_n \leq t \leq a_{n+1}} |M_i(t)|^p\right].$$

As the right hand side is finite we can apply the Kolmogorov two series test,

$$\sum_{n=0}^{\infty} \sup_{a_n \leq t \leq a_{n+1}} |M_i(t)|^p < \infty.$$

Thus we necessarily must have,

$$\sup_{a_n \leq t \leq a_{n+1}} |M_i(t)|^p \longrightarrow 0 \text{ as } n \to \infty \quad a.s.$$

This clearly yields $M_i(t) \to 0$ almost surely for each i. As in the proof of Lemma 3.3.1 we introduce the process $Z(t) = X(t) - M(t)$, which yields,

$$\dot{Z}(t) = (\nu * Z)(t) + f(t) + (\nu * M)(t) + M(t).$$

By variation of constants we obtain,

$$Z(t) - (r * f)(t) = r(t)\xi + (r * \nu * M)(t) + (r * M)(t).$$

As $r \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$, $\nu \in M(\mathbb{R}_+;\mathbb{R}^{d\times d})$ and $M \in BC_0(\mathbb{R}_+;\mathbb{R}^d)$, the right hand side will converge to the zero vector upon sending $t \to \infty$. Thus we have (for any norm on $\mathbb{R}^d$),

$$\lim_{t\to\infty} \|X(t) - (r * f)(t)\| \leq \lim_{t\to\infty} (\|Z(t) - (r * f)(t)\| + \|M(t)\|) = 0 \quad a.s.$$

$\square$

3.4 Stochastic Functional Differential Equations

Once again, consider a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$ supporting an m-dimensional standard Brownian motion, $(B_t)_{t\geq 0}$. We now study the following $\mathbb{R}^d$-valued stochastic functional differential equation (for some fixed delay parameter $\tau > 0$) given by,

$$dX(t,\psi) = \left(f(t) + \int_{[-\tau,0]} \mu(ds)X(t+s,\psi) \right) dt + \sigma(t)dB(t), \quad t \geq 0; \quad X(t,\psi) = \psi(t) \quad t \leq 0. \tag{3.4.1}$$

Here ψ is a $C([-\tau,0];\mathbb{R}^d)$-valued $\mathcal{F}_0$-measurable random variable, $f \in C(\mathbb{R}_+;\mathbb{R}^d)$, $\mu \in M([-\tau,0];\mathbb{R}^{d\times d})$ and $\sigma \in C(\mathbb{R}_+;\mathbb{R}^{d\times m})$. We make the standing assumption that,

$$\mathbb{E}\left[\sup_{t\in[-\tau,0]} \|\psi(t)\|^2 \right] < +\infty. \tag{3.4.2}$$

This is needed to apply Theorem 1.2.1, guaranteeing existence and uniqueness. As in the previous section, B is a standard m-dimensional Brownian Motion where $\langle B(t), \mathbf{e}_i \rangle$ and $\langle B(t), \mathbf{e}_j \rangle$ are independent for all $i \neq j$. With only trivial modifications, essentially all results from section 3.3 carry over to the finite memory problem, but can be significantly improved. When considering functional equations one can remove any side assumptions on the resolvent leading to much stronger results. We denote by r_τ the differential resolvent of the measure ν which solves the following matrix equation,

$$r_\tau'(t) = \int_{[-\tau,0]} \mu(ds) r_\tau(t+s), \quad t > 0; \quad r_\tau(0) = I_{d\times d}. \tag{3.4.3}$$

Unlike for Volterra equations we do not need to make a standing assumption that $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$, although we do note that this is equivalent to r_τ obeying the estimate $\|r_\tau(t)\| \leq Ce^{-\alpha t}$, for some $C, \alpha > 0$ and an arbitrary norm on $\mathbb{R}^{d\times d}$. We state the following lemma whose proof is identical to that of Lemma 3.3.1,

Lemma 3.4.1. *Let $p \in [1,\infty)$. Assume $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$, $\psi \in L^p(\Omega; C([-\tau,0];\mathbb{R}^d))$, X be the solution of (3.4.1), and Y be the solution of (3.3.7). Then the following are true,*

(i) $\mathbb{E}[\|X(\cdot,\psi)\|^p] \in L^1([-\tau,\infty);\mathbb{R}) \iff \mathbb{E}[\|Y(\cdot)\|^p] \in L^1(\mathbb{R}_+;\mathbb{R})$.

(ii) $X(\cdot,\psi) \in L^p([-\tau,\infty);\mathbb{R}^d)$ *almost surely* $\iff Y(\cdot) \in L^p(\mathbb{R}_+;\mathbb{R}^d)$ *almost surely.*

Remark 3.4.1. *Lemma 3.4.1 provides further evidence to the conjecture that memory has no effect on the integrability of the trajectories. Thus even equations with possibly wildly different delay structures (such as equations (3.3.1) and (3.4.1)) are completely equivalent on the level of p-integrable trajectories.*

Theorem 3.4.1. *Let $p \in [1,\infty)$. Let X be the solution of (3.4.1) and suppose $\psi \in L^p(\Omega; C([-\tau,0];\mathbb{R}^d))$. Then the following statements are true,*

(A) *If $p \in [2,\infty)$, the following are equivalent.*

(i) For each initial function ψ, $\mathbb{E}[\|X(\cdot,\psi)\|^p] \in L^1([-\tau,\infty);\mathbb{R})$,

(ii) *For each initial function* ψ, $X(\cdot,\psi) \in L^p([-\tau,\infty);\mathbb{R}^d)$ *almost surely,*

(iii) $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$, f *obeys* (3.3.2) *and* σ *obeys* (3.3.3).

(B) *If* $p \in [1,2)$, *the following are equivalent.*

(i) *For each initial function* ψ, $\mathbb{E}[\|X(\cdot,\psi)\|^p] \in L^1([-\tau,\infty);\mathbb{R})$,

(ii) *For each initial function* ψ, $X(\cdot,\psi) \in L^p([-\tau,\infty);\mathbb{R}^d)$ *almost surely,*

(iii) $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$, f *obeys* (3.3.2) *and* σ *obeys* (3.3.4).

Remark 3.4.2. *The reader should note how much stronger Theorem 3.4.1 is compared to Theorems 3.3.1 and 3.3.2 for the Volterra equation. For the SFDE we obtain a much cleaner characterisation and a far stronger converse when claiming the necessity of* $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$. *A similar situation occurs when considering the asymptotic behaviour of the paths, see Theorem 3.4.2 below.*

Before proving Theorem 3.4.1 we recall some common notation in the field of functional differential equations. If we set $\sigma = 0$ in (3.4.1) we obtain the equation,

$$\dot{x}(t) = \int_{[-\tau,0]} \mu(du)x(t+u) + f(t), \quad t \geq 0; \quad x(t) = \psi(t), \quad t \leq 0, \tag{3.4.4}$$

whose solution is given by,

$$x(t,\psi) = r_\tau(t)\psi(0) + \int_{[-\tau,0]} \mu(ds)\left(\int_s^0 r_\tau(t+s-u)\psi(u)du\right) + \int_0^t r_\tau(t-s)f(s)ds. \tag{3.4.5}$$

Let,

$$x_0(t,\psi) := r_\tau(t)\psi(0) + \int_{[-\tau,0]} \mu(ds)\left(\int_s^0 r_\tau(t+s-u)\psi(u)du\right), \tag{3.4.6}$$

so we can rewrite (3.4.5) as $x(t,\psi) = x_0(t,\psi) + (r_\tau * f)(t)$. Thus x_0 solves equation (3.4.4) with the perturbation term switched off. The so called characteristic matrix is given by $\Delta(\lambda) = \lambda I_{d\times d} - \int_{[-\tau,0]} \mu(ds)e^{\lambda s}$ wherein the asymptotic behaviour of r_τ is governed by the roots of the characteristic equation $\det(\Delta(\lambda)) = 0$. Let $\Lambda = \{\lambda \in \mathbb{C} : \det(\Delta(\lambda)) = 0\}$ and define $v_0(\mu) := \sup\{\mathrm{Re}(\lambda) : \lambda \in \Lambda\}$ with $\Lambda' = \{\lambda \in \Lambda : \mathrm{Re}(\lambda) = v_0(\mu)\}$.

Proof of Theorem 3.4.1. Essentially an identical argument as in the proof of Theorem 3.3.1 gives $(iii) \implies (i)$ for both (**A**) and (**B**). In order to show $(i) \implies (ii)$ we need to show that $(i) \implies r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$ to which we could then follow the line of proof once again from Theorem 3.3.1. Thus we assume (i) and aim to show $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$. First we note that for an arbitrary deterministic ψ we have $X(t,\psi) - X(t,0) = x_0(t,\psi)$ and thus we have the estimate,

$$\int_0^\infty \mathbb{E}[\|x_0(t,\psi)\|]^p dt \leq C\int_0^\infty \mathbb{E}[\|X(t,\psi)\|]^p dt + C\int_0^\infty \mathbb{E}[\|X(t,\psi)\|]^p dt < +\infty,$$

for some constant $C > 0$. But ψ is deterministic thus $x_0(\cdot,\psi) \in L^p([-\tau,\infty);\mathbb{R}^d)$. Let $\lambda \in \Lambda'$; then by exercise I.3.8 in [44] we have for some $v \in \mathbb{R}^d\backslash\{0\}$, $x_0(t,\psi) = \exp(\text{Re}(\lambda)t)v$. If $v_0(\mu) \geq 0$ this contradicts $x_0(,\psi) \in L^p([-\tau,\infty);\mathbb{R}^d)$ and so we must have $v_0(\mu) < 0$, but this yields the exponential estimate $\|r_\tau(t)\| \leq Ce^{-\alpha t}$, for some $C, \alpha > 0$ and so $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$ as required. Thus we follow the proof of Theorem 3.3.1 verbatim to claim $(i) \implies (ii)$ for both $(\mathbf{A})$ and $(\mathbf{B})$. In the same manner one can show $(ii) \implies r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$ thus we once again follow the proof of Theorem 3.3.1 to claim $(ii) \implies (iii)$ for both $(\mathbf{A})$ and $(\mathbf{B})$. $\square$

Our final theorem regarding SFDEs is concerned with almost sure convergence to zero in the regime where σ is diagonal and once again we improve significantly on the Volterra case by removing all assumptions on the resolvent. We note also that a version of Theorem 3.3.4 also holds for the functional equation but we do not state this as a theorem.

Theorem 3.4.2. *Let ψ obey (3.4.2), X be the solution of (3.3.1), $m = d$ and σ be a diagonal matrix. Then the following are equivalent,*

(i) $r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$, each component of f and σ satisfy (3.3.16),

(ii) For each initial function ψ, $\|X(t,\psi)\| \longrightarrow 0$ as $t \to \infty$, almost surely.

Proof of Theorem 3.4.2. The proof of $(i) \implies (ii)$ follows directly from Theorem 3.3.5 with only trivial modifications. For the reverse implication first one needs to show $(ii) \implies r_\tau \in L^1(\mathbb{R}_+;\mathbb{R}^{d\times d})$, which follows from identical arguments as in the proof of Theorem 3.4.1. Thus we can once again (with only trivial modifications) follow the proof of Theorem 3.3.5 to yield the claim. $\square$

3.5 Examples

This section is devoted to providing explicit examples of perturbation functions which highlight the utility of the theory presented and also dispel any thought that the conditions (3.3.2) and (3.3.3) we have generated are superfluous and in fact just equivalent to imposing that $f, \sigma^2 \in L^p$ in the appropriate sense. This is indeed not the case. All functions presented in this section will be scalar valued. In Chapter 2 we provided the example,

$$f(t) = e^{\alpha t}\sin(e^{\beta t}),$$

where $0 < \alpha < \beta$ and show $\int_\cdot^{\cdot+\theta} f(s)ds \in L^p(\mathbb{R}_+;\mathbb{R})$ despite the fact that $\int_0^\infty |f(s)|^p ds = +\infty$ for any $p \geq 1$. This takes care of the deterministic perturbation. However for the stochastic perturbation we are concerned with σ^2 rather than σ itself, thus the positivity means such oscillatory functions as above cannot be considered. Instead one must think about functions that have very large deviations over very small intervals; we construct such a function below.

For each $n \in \mathbb{N}$, let $a_n < 1/2$ and h_n be positive sequences and suppose that $g(t) = 0$ for $[n, n+a_n]$ and $[n+1-a_n, n+1]$ and on $[n+a_n, n+1/2]$, g is linear with $g(n+a_n) = 0$ and $g(n+1/2) = h_n$, while on $[n+1/2, n+1-a_n]$, g is linear with $g(n+1-a_n) = 0$. Then g is continuous, has a spike of width $1-2a_n$ and maximal height h_n. Below we give an explicit example,

$$g(t) \coloneqq \begin{cases} 0, & t \in [0,2], \\ 0, & t \in [n, n+a_n], \\ \frac{h_n}{(\frac{1}{2}-a_n)}t - \frac{h_n(n+a_n)}{(\frac{1}{2}-a_n)}, & t \in [n+a_n, n+\frac{1}{2}], \\ \frac{-h_n}{(\frac{1}{2}-a_n)}t + \frac{h_n(n+1-a_n)}{(\frac{1}{2}-a_n)}, & t \in [n+\frac{1}{2}, n+1-a_n], \\ 0, & t \in [n+1-a_n, n+1], \end{cases}$$

$n \geq 2$, with $h_n \coloneqq n^\beta$ and $a_n \coloneqq \frac{1}{2} - \frac{1}{n^{\beta+1}}$, where $\beta > 0$. Figure 3.1 provides a graph of the first five spikes of g.

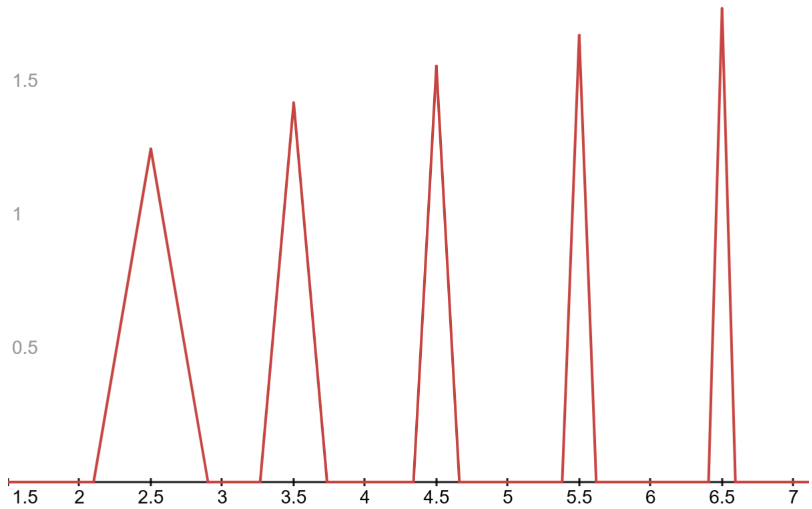


Figure 3.1: Graph of $g(t)$ with $\beta = 0.32$

The following proposition shows this function has the desired behaviour.

Proposition 3.5.1. *Let $g : \mathbb{R}_+ \to \mathbb{R}_+$ be defined as above, then for each $p \in (1, \infty)$ we have,*

$$\int_0^\infty \left| \int_t^{t+1} g(s)ds \right|^p dt < +\infty; \qquad \int_0^\infty |g(s)|^p ds = +\infty.$$

Proof of proposition 3.5.1. One can easily verify that,

$$\int_n^{n+1} g(s)ds = \frac{1}{n}.$$

Next consider,

$$\begin{aligned}
\int_0^\infty \left| \int_t^{t+1} g(s)ds \right|^p dt &= \sum_{n=2}^\infty \int_n^{n+1} \left(\int_t^{t+1} g(s)ds \right)^p dt \\
&\leq \sum_{n=2}^\infty \int_n^{n+1} \left(\int_n^{n+2} g(s)ds \right)^p dt \\
&= \sum_{n=2}^\infty \left(\int_n^{n+1} g(s)ds + \int_{n+1}^{n+2} g(s)ds \right)^p \\
&\leq C_p \left(\sum_{n=2}^\infty \frac{1}{n^p} + \frac{1}{(n+1)^p} \right) < +\infty,
\end{aligned}$$

where $C_p > 0$ is a constant depending on p, hence the first assertion is proven. For the second assertion consider the estimate,

$$\int_0^\infty g(s)^p ds \geq \sum_{n=1}^\infty \int_n^{n+1} g(s)^p \chi_{\{g(s)\geq 1\}}(s)ds,$$

thus we need only show the series on the right is divergent and then we are done. Now as $p > 1$ we have,

$$\int_n^{n+1} g(s)^p \chi_{\{g(s)\geq 1\}}(s)ds \geq \int_n^{n+1} g(s)\chi_{\{g(s)\geq 1\}}(s)ds.$$

Now as $\beta > 0$ we have $\limsup_{t\to\infty} g(t) = +\infty$, moreover the running maxima is non-decreasing hence there exists an $m \in \mathbb{N}$ such that $\sup_{t\in[n,n+1]} g(t) > 1$ for all $n \geq m$. Fix such an m (recall as g is continuous this sup is always attained). One can verify that for $n \geq m$,

$$\int_n^{n+1} g(s)\chi_{\{g(s)\geq 1\}}(s)ds = \frac{1}{n} - \frac{2}{n^{\beta+1}} + \frac{1}{n^{2\beta+1}},$$

which, as $\beta > 0$, is not a summable sequence. Thus we have,

$$\int_0^\infty g(s)^p ds \geq \sum_{n=m}^\infty \int_n^{n+1} g(s)^p \chi_{\{g(s)\geq 1\}}(s)ds \geq \sum_{n=m}^\infty \left(\frac{1}{n} - \frac{2}{n^{\beta+1}} + \frac{1}{n^{2\beta+1}}\right) = +\infty,$$

as required. □

Thus consider the following scalar equation,

$$dX(t) = \left(e^{\alpha t}\sin(e^{\beta t}) + \int_{[0,t]} \nu(ds)X(t-s)\right)dt + \sigma(t)dB(t),$$

where $\sigma(t) := \sqrt{g(t)}$ where $g(t)$ is defined as in proposition 3.5.1. Then Theorem 3.3.1 tells us that, $X \in L^p(\mathbb{R}_+;\mathbb{R})$ $a.s.$ despite the fact that $\int_0^\infty |f(s)|^p ds = +\infty$ and $\int_0^\infty |\sigma^2(s)|^{p/2} ds = +\infty$ for any $p \geq 2$.

3.6 Conclusion

In this chapter we have provided characterisations of when perturbed integrodifferential Volterra equations with additive noise have almost surely p-integrable trajectories, when the associated p^{th}-mean process resides in L^1 and the almost sure p-summability of solutions to analogous equations in discrete time. The main driver of these results was the connection between the Volterra equation (3.3.1) and an associated SDE (3.3.7). The results in this chapter suggest the qualitative (and possibly quantitative) behaviour of such Volterra equations is equivalent to that of a much simpler SDE and the authors conjecture this equivalence may extend to the identification of top Lyapunov exponents.

In addition to studying the qualitative behaviour we have also engaged in the discussion of almost sure convergence to zero in various regimes and supplied a characterisation of such convergence in a special case. As outlined in section 3.4, the relation between such Volterra equations and SFDEs is very strong. When passing to the finite memory problem one can obtain even stronger results as compared to the Volterra case.

3.7 Proofs

Proof of proposition 3.2.1. Suppose the distribution of ξ has two isolated atoms at α and β, then we can find epsilon Balls such that $\mathbb{E}[\xi\chi_{\{\xi\in B_\varepsilon(\alpha)\}}] = \alpha\mathbb{P}(\xi=\alpha)$ and $\mathbb{E}[\xi\chi_{\{\xi\in B_\varepsilon(\beta)\}}] = \beta\mathbb{P}(\xi=\beta)$. Now assume,

$$\mathbb{P}(\xi=\alpha)\mathbb{E}[\xi\chi_{\{\xi\in B_\varepsilon(\beta)\}}] = \mathbb{P}(\xi=\beta)\mathbb{E}[\xi\chi_{\{\xi\in B_\varepsilon(\alpha)\}}].$$

Clearly this forces $\beta=\alpha$ which is a contradiction thus we take the two bounded Borel sets to be the epsilon balls defined above and we have $\xi\in\mathbb{D}$. Now assume the distribution of ξ has an absolutely continuous part, and denote its density by f. Fix an arbitrary non zero t in the interior of the support of f and choose $a<b$ such that $\mathbb{P}(\xi\in(a,t))=\mathbb{P}(\xi\in(t,b))$[5]. Now assume,

$$\mathbb{P}(\xi\in(a,t))\mathbb{E}[\xi\chi_{\{\xi\in(t,b)\}}] = \mathbb{P}(\xi\in(t,b))\mathbb{E}[\xi\chi_{\{\xi\in(a,t)\}}],$$

which yields,

$$\int_t^b xf(x)dx = \int_a^t xf(x)dx.$$

A simple estimate yields,

$$t\int_t^b f(x)dx \le \int_a^t xf(x)dx.$$

But by construction we then must have,

$$t\int_a^t f(x)dx \le \int_a^t xf(x)dx.$$

This yields the following inequality,

$$0 \le \int_a^t (t-x)f(x)dx \le 0.$$

Thus the integrand must be zero *a.e.* but this impossible. Thus we have the desired contradiction. Thus we choose our Borel sets to be (a,t) and (t,b) which gives us $\xi\in\mathbb{D}$ as required.

[5]This can always be done due to the continuity of the distribution function restricted to the support of f.

□

Proof of Lemma 3.3.3 . Let $\theta = 2\beta$. We note that for all $t \in [a_n, a_{n+1}]$ we have $[a_{n+1}, a_n + 2\beta] \subset [t, t+2\beta]$. Thus,

$$
\begin{aligned}
+\infty > \int_0^{\infty} \left(\int_t^{t+2\beta} f(s)ds \right)^p dt &= \sum_{n=0}^{\infty} \int_{a_n}^{a_{n+1}} \left(\int_t^{t+2\beta} f(s)ds \right)^p dt \\
&\geq \sum_{n=0}^{\infty} \int_{a_n}^{a_{n+1}} \left(\int_{a_{n+1}}^{a_n+2\beta} f(s)ds \right)^p dt \\
&= \sum_{n=0}^{\infty} (a_{n+1} - a_n) \left(\int_{a_{n+1}}^{a_n+2\beta} f(s)ds \right)^p \\
&\geq \alpha \sum_{n=0}^{\infty} \left(\int_{a_{n+1}}^{a_n+2\beta} f(s)ds \right)^p \\
&\geq \alpha \sum_{n=0}^{\infty} \left(\int_{a_{n+1}}^{a_{n+2}} f(s)ds \right)^p,
\end{aligned}
$$

where the last inequality comes from the fact $a_{n+2} \leq \beta + a_{n+1} \leq 2\beta + a_n$.

□

Chapter 4

Stochastic Functional Differential Equations with Multiplicative Noise

The material in this chapter is based on the following article:

1. J. A. D. Appleby and E. Lawless. Mean square asymptotic stability characterisation of perturbed linear stochastic functional differential equations. *Applied Numerical Mathematics*, 200;80-109 (2024) [10].

4.1 Introduction

In this Chapter we turn our attention to perturbed scalar SFDEs with multiplicative noise. A variety of stability types can be considered, but from the outset of stability studies in stochastic functional differential (or evolution) equations the asymptotic behaviour, and specifically the convergence in the mean square, has attracted a great deal of attention (see for instance the Haussmann [61], Ichikawa [64], Mao [81], Mizel and Trutzer [94], Mohammed [95]). Substantial bibliographies can be found in the important monographs by Mao [85] and Shaikhet [112]. Moreover, since linear equations with multiplicative noise are so fundamental to mathematical analysis of hereditary systems, it is appropriate that they receive due study. Concretely, the equation of interest in this Chapter is the following:

$$dX(t) = \left(f(t) + \int_{[-\tau,0]} X(t+s)\nu(ds) \right) dt + \left(\sigma(t) + \int_{[-\tau,0]} X(t+s)\mu(ds) \right) dB(t), \tag{4.1.1}$$

where ν and μ are finite signed Borel measures on $[-\tau, 0]$. It is worth noting that the analysis in this Chapter is strictly confined to scalar equations with no clear generalisation of results to the multi-dimensional setting. The reason for this will be clear upon inspection of the proofs, one would expect such a concession should come with substantial rewards in the form of very strong and precise Theorems; we believe this is indeed the case.

The results in this Chapter are able to characterise exactly the conditions which give rise to the various types of mean square convergence. In particular, we are able to show that $\mathbb{E}[X^2(t,\psi)] \to 0$ as $t \to \infty$ for every continuous initial condition ψ with finite second moment is *equivalent to* the mean square convergence to zero of the solution to the unperturbed equation (i.e with $f = \sigma = 0$), alongside the conditions

$$\int_t^{t+\theta} f(s)\,ds \to 0, \quad t \to \infty \text{ for each } \theta \in (0,1], \quad \int_t^{t+1} \sigma^2(s)\,ds \to 0 \quad t \to \infty.$$

If exponential convergence of $\mathbb{E}[X^2(t)]$ to zero is desired as $t \to \infty$, this is equivalent the mean square convergence to zero of the solution to the unperturbed equation, alongside the following exponential decay conditions on f and σ:

$$\text{There is } C > 0,\, \beta > 0 \text{ such that } \left|\int_0^t e^{\beta s} f(s)\,ds\right| \le C, \quad t \ge 0, \quad \int_0^\infty e^{2\beta s}\sigma^2(s)\,ds < +\infty.$$

Finally, if we want mean square integrability, as characterised above, this is equivalent the mean square convergence to zero of the solution to the unperturbed equation, alongside the following square integrability conditions fulfilled by f and σ:

$$t \mapsto \int_t^{t+\theta} f(s)\,ds \in L^2(\mathbb{R}_+) \text{ for each } \theta > 0; \quad \sigma \in L^2(\mathbb{R}_+).$$

These results constitute a solid advance in the theory, since *coincident necessary and sufficient conditions on f and σ are imposed which guarantee the appropriate type of convergence.* Until now, sufficient condition results abound, but such an exact characterisation has not been achieved. Other results achieve something of this goal, but here we are able to exactly characterise the underlying stability condition, as well as the precise conditions on the forcing term which enable the results to hold.

It is to be noted that these conditions do not place pointwise bounds on f and σ, but are rather conditions on certain types of averages of f and σ. In particular for convergence to zero and mean square integrability we see exactly the same conditions imposed on the forcing functions as in the previous two Chapters. The exponential integral properties noted above have occurred already in the literature for *affine* stochastic Volterra integrodifferential equations (see e.g. [9, 83]), in which there is no state dependence in the diffusion term.

Clearly the results stated above rely on whether or not one has a strong understanding of the behaviour of the solution to the unperturbed equation. A full characterisation of the mean–square behaviour of general scalar linear SFDEs was produced in Appleby, Riedle and Mao [16]. They studied the following equation:

$$dU(t) = \left(\int_{[-\tau,0]} U(t+s)\nu(ds)\right) dt + \left(\int_{[-\tau,0]} U(t+s)\mu(ds)\right) dB(t) \tag{4.1.2}$$

with continuous and deterministic initial function ψ, it has been shown that the global mean square asymptotic stability of the zero solution of this equation is equivalent to

$$r(t) \to 0 \text{ as } t \to \infty, \quad \int_0^\infty \left(\int_{[-\tau,0]} r(t+s)\mu(ds) \right)^2 dt < 1,$$

where r is the differential resolvent[1] of the underlying deterministic differential equation i.e.,

$$r'(t) = \int_{[-\tau,0]} r(t+s)\nu(ds), \quad t > 0; \quad r(0) = 1, \quad r(t) = 0 \quad t \in [-\tau, 0).$$

Indeed, if the mean square convergence to zero occurs, it must do so exponentially fast. There is a very extensive literature on exponential mean square stability for SFDEs with finite delay, with great advances being made by Mao and co–workers: some representative and foundational works from this school include [82, 84, 86, 88] and the themes of this chapter and more recent advances is reflected in the monograph [85]. Other papers which seek to give a characterisation of the mean square asymptotic stability of solutions of unforced SFDEs include [31], [80] and [5]. A corresponding characterisation for linear stochastic difference equations is presented by Appleby and Lawless in [11].

The chapter is organised as follows: Section 4.2 gives a precise formulation of the problem, together with some background theory. Section 4.3 deduces linear Volterra integral equations for the mean square of X, as well as some auxiliary functions (which are mean–squares of other processes). Section 4.4 states the main results and discusses hypotheses. Section 4.5 gives a brief conclusion and Section 4.6 contains the proofs of the main results.

4.2 Mathematical Preliminaries

For the following general results on SFDEs and stochastic analysis, the reader may refer to the monographs [85] and [72]. Let us fix a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration $(\mathcal{F}(t))_{t\geq 0}$ satisfying the usual conditions and let $B = \{B(t) : t \geq 0\}$ be a one–dimensional Brownian motion on this space. Let $\tau > 0$ and ψ be a $C([-\tau, 0]; \mathbb{R})$–valued $\mathcal{F}(0)$–measurable random variable with

$$\|\psi\|^2 := \mathbb{E}\left[\sup_{t\in[-\tau,0]} \psi^2(t) \right] < +\infty, \tag{4.2.1}$$

recalling that $C([-\tau, 0]; \mathbb{R})$ is the space of continuous functions $\varphi : [-\tau, 0] \to \mathbb{R}$ equipped with the norm $|\varphi| = \sup_{t\in[-\tau,0]} |\varphi(t)|$. We assume also ψ is independent of the Brownian

[1] Here we do **NOT** employ the notation of Chapter 3 and use r instead of r_τ to denote the functional resolvent. This is due to the fact in this chapter only equations with finite memory are considered, hence this notation change should not cause confusion.

motion B. Note that the finiteness of $\|\psi\|$ and the Dominated Convergence Theorem ensure that $t \mapsto \mathbb{E}[\psi^2(t)]$ is continuous on $[-\tau, 0]$. Denote

$$\phi(t) := \sqrt{\mathbb{E}[\psi^2(t)]}, \quad t \in [-\tau, 0], \tag{4.2.2}$$

so that $\phi^2(t) = \mathbb{E}[\psi^2(t)]$ for $t \in [-\tau, 0]$. Both ϕ and ϕ^2 are continuous.

We first state the unperturbed equation whose asymptotic behaviour is of paramount importance. If we let ν and μ be in $M([-\tau, 0]; \mathbb{R})$, the space of finite Borel measures on $[-\tau, 0]$, the unperturbed equation has the following form:

$$\begin{aligned} dU(t) &= \left(\int_{[-\tau,0]} U(t+s)\nu(ds) \right) dt + \left(\int_{[-\tau,0]} U(t+s)\mu(ds) \right) dB(t), \quad t \geq 0 \\ U(t) &= \psi(t), \quad t \leq 0, \end{aligned} \tag{4.2.3}$$

where ψ has the properties indicated above. (4.2.3) is so–called differential shorthand for

$$\begin{aligned} U(t) &= U(0) + \int_0^t \left(\int_{[-\tau,0]} U(s+u)\nu(du) \right) ds + \int_0^t \left(\int_{[-\tau,0]} U(s+u)\mu(du) \right) dB(s), \quad t \geq 0 \\ U(t) &= \psi(t), \quad t \leq 0. \end{aligned}$$

For every ψ as specified above, there exists a unique, continuous, adapted process $U = \{U(t, \psi) : t \geq -\tau\}$ which satisfies (4.2.3). This process U is a so–called strong solution of (4.2.3), and U has finite second moments (see Theorem 1.2.1). This means that

$$\mathbb{E}[U^2(t)] < +\infty, \quad \text{for all } t \geq -\tau,$$

and, a fortiori,

$$\mathbb{E}\left[\sup_{-\tau \leq s \leq t} U^2(s) \right] < +\infty, \quad \text{for all } t \geq -\tau.$$

The equation (4.2.3) was studied extensively by Appleby et al. [16] in which they gave a full characterisation of the mean square behaviour of (4.2.3). This includes a set of necessary and sufficient conditions which ensures $\mathbb{E}[U^2(t, \psi)] \to 0$ as $t \to \infty$ for all initial functions ψ. It should be noted in [16] the authors only considered deterministic initial functions: however with the additional assumption (4.2.1), *a condition which we will impose throughout this chapter without further reference*, all of their results carry over to the case of random initial functions. In [16] it was found that the stochastic stability is heavily dependent on the behaviour of the underlying deterministic equation and moreover the fundamental resolvent. In this chapter we demonstrate that this still prevails and so we introduce both of these objects in detail. The deterministic[2], unperturbed delay equation

[2]To use the term *deterministic* to describe equation (4.2.4) is technically incorrect due to the presence of the random initial function. However its dynamics are indeed deterministic so it is in this spirit that

is given by,

$$\dot{x}_0(t) = \int_{[-\tau,0]} x_0(t+u)\nu(du), \quad t \geq 0, \tag{4.2.4}$$
$$x_0(t) = \psi(t), \quad t \in [-\tau, 0],$$

where both ν and ψ are defined as above. There are many texts that deal with deterministic delay equations; for further analysis of (4.2.4) we refer the reader to [44, 58]. The underlying integral resolvent is the unique locally absolutely continuous function $r : [0, \infty) \to \mathbb{R}$ which satisfies

$$r(t) = 1 + \int_0^t \int_{[\max\{-\tau,-s\},0]} r(s+u)\nu(du)ds, \quad t \geq 0. \tag{4.2.5}$$

The above equation can be written in differential form (as was done in the previous chapters) by specifying $r(0) = 1$ and $r(t) = 0$ for all $t < 0$. In all results throughout this chapter we need to make assumptions on the asymptotic behaviour of the resolvent r and its connection with the measure ν. The following description of the asymptotic behaviour is standard and may be found in [44, 58]. As pointed out in Appleby et al. [16], the following conditions on solutions to (4.2.5) are all equivalent:

(a) $r(t) \to 0$, as $t \to \infty$;

(b) $r \in L^1(\mathbb{R}_+)$;

(c) $r \in L^2(\mathbb{R}_+)$.

Henceforth the above relations will be used interchangeably without reference. An important detail regarding the stability of the resolvent is that whenever any of the above conditions are fulfilled r tends to zero *exponentially* fast. This fact was used in Section 3.4 of Chapter 3, for completeness we provide a more detailed discussion in the present chapter. One looks for solutions to (4.2.5) of exponential type which leads to a transcendental characteristic equation. For all $\lambda \in \mathbb{C}$ we may define

$$h(\lambda) = \lambda - \int_{[-\tau,0]} e^{\lambda s}\nu(ds).$$

We pause to note that the second term on the right hand side has the character of a Laplace transform, and accordingly we will use the notation

$$\hat{\nu}(\lambda) := \int_{[-\tau,0]} e^{\lambda s}\nu(ds), \quad \lambda \in \mathbb{C}.$$

we will continue to refer to equation (4.2.4) as *deterministic*.

For measurable functions f defined on $[0,\infty)$, the usual Laplace transform is defined by

$$\hat{f}(\lambda) := \int_0^\infty e^{-\lambda s} f(s)\, ds$$

for λ in appropriate regions of $\mathbb{C}$.

Returning to a discussion of the solutions of the characteristic equation, it is standard ([44, Chapter I]) that the set $\Lambda := \{\lambda \in \mathbb{C} : h(\lambda) = 0\}$ is non–empty and that there is a finite $v_0(\nu) \in \mathbb{R}$ such that

$$v_0(\nu) = \sup\{\mathrm{Re}(\lambda) : \lambda \in \Lambda\}.$$

Finally, it is the case that $r(t) \to 0$ as $t \to \infty$ is equivalent to $v_0(\nu) < 0$. The significance of the number $v_0(\nu)$ is that it enables us to obtain a definite exponential bound on r. Specifically, for all $\alpha > v_0(\nu)$ we have $r(t) = o(\exp(\alpha t))$ for $t \to \infty$. Indeed, we have a global exponential bound: for each $\alpha > v_0(\nu)$ there is a $K = K_\alpha > 0$ such that $|r(t)| \le K_\alpha e^{\alpha t}$ for all $t \ge 0$. This global exponential bound is inherited by $x_0(\cdot,\psi)$ which can be seen via a variation of constants formula

$$x_0(t,\psi) = r(t)\psi(0) + \int_{[-\tau,0]} \left(\int_s^0 r(t+s-u)\psi(u)du \right) \nu(ds), \quad t \ge 0.$$

Taking the triangle inequality, one obtains the bound

$$|x_0(t,\psi)| \le C_\alpha e^{\alpha t} \sup_{s\in[-\tau,0]} |\psi(s)|, \quad t \ge 0,$$

where $C_\alpha > 0$ is a constant depends on α and ν, but not on ψ. Since we assume that ψ is random with $\mathbb{E}[\sup_{s\in[-\tau,0]} |\psi(s)|^2] < +\infty$, we get the estimate

$$\mathbb{E}[x_0^2(t;\psi)] \le C_\alpha^2 e^{2\alpha t} \mathbb{E}\left[\sup_{s\in[-\tau,0]} |\psi(s)|^2 \right], \quad t \ge 0,$$

for all $\alpha > v_0(\nu)$. We therefore observe that if $v_0(\nu) < 0$, then all solutions tend to zero exponentially fast for deterministic initial conditions, and also exponentially fast in mean square if the initial function has a finite mean square in the sense given above.

The converse of this result is also true (namely that if all solutions of (4.2.4) tend to zero, then $v_0(\nu) < 0$).

To see this, we make a general observation. If $\lambda \in \Lambda$, and $\psi(t) = e^{\lambda t}$ for $t \in [-\tau, 0]$, then $x_0(t,\psi) = e^{\lambda t}$ for all $t \ge 0$. If λ is real, then this furnishes a real–valued solution; in the case that $\lambda \in \mathbb{C}$, we can use the observation that the conjugate of λ, $\bar{\lambda}$ is also in Λ, to get real valued solutions. Note first that $x(t, a\psi_1 + b\psi_2) = ax(t,\psi_1) + bx(t,\psi_2)$ for $t \ge 0$ and any $a, b \in \mathbb{C}$ and any continuous complex–valued initial functions ψ_1 and ψ_2. Taking $\psi_1(t) = e^{\lambda t}$ and $\psi_2(t) = e^{\bar{\lambda} t}$ and $a = b = 1/2$, we see that the real–valued initial function $\psi(t) = \mathrm{Re}(e^{\lambda t})$ for $t \geq 0$ gives rise to the real–valued solution $x(t,\psi) = \mathrm{Re}(e^{\lambda t})$ for $t \geq 0$; likewise, taking $a = 1/(2i)$ and $b = -1/(2i)$, the real–valued initial function $\psi(t) = \mathrm{Im}(e^{\lambda t})$ for $t \geq 0$ gives rise to the real–valued solution $x(t,\psi) = \mathrm{Im}(e^{\lambda t})$ for $t \geq 0$.

Now, let $x_0(t,\psi) \to 0$ as $t \to \infty$ for all $\psi \in C([-\tau,0],\mathbb{R})$. Suppose, by way of contradiction, that $v_0(\nu) \geq 0$. Since in fact $\sup\{\mathrm{Re}(\lambda) : \lambda \in \Lambda\} = \max\{\mathrm{Re}(\lambda) : \lambda \in \Lambda\}$ we have that there is $\lambda \in \Lambda$ such that $\mathrm{Re}(\lambda) = v_0(\nu)$. Now take $\psi(t) = \mathrm{Re}(e^{\lambda t})$ for $t \in [-\tau, 0]$. Then $x(t,\psi) = \mathrm{Re}(e^{\lambda t})$ for $t \geq 0$. But since $\mathrm{Re}(\lambda) = v_0(\nu) \geq 0$, $\limsup_{t\to\infty} |x(t,\psi)| > 0$. But this contradicts the supposition that $x(t,\psi) \to 0$ as $t \to \infty$, so we have that $v_0(\nu) < 0$, as needed.

Here, and in what follows, we place great weight on understanding the asymptotic behaviour of components in the solution of the perturbed stochastic equation, because it is a common feature in our proofs that we decouple the behaviour of the perturbed stochastic equation into parts which depend on either the functional appearing in the diffusion coefficient or on the underlying deterministic equation, i.e., the resolvent. Before discussing the mean square asymptotic stability of (4.2.3) we need to introduce the notation

$$G(f_t) := \int_{[-\tau,0]} f(t+u)\mu(du), \text{ for all } f \in C[0,\infty),$$

for a measure $\mu \in M$. We recall the main result[3] from Appleby et al. [16] which states

$$\lim_{t\to\infty} \mathbb{E}[U^2(t,\psi)] = 0 \text{ for all } \psi \text{ obeying (4.2.1)} \quad \Longleftrightarrow \quad \begin{cases} r \in L^2(\mathbb{R}_+), \\ \|G(r_\cdot)\|_{L^2(\mathbb{R}_+)} < 1, \end{cases} \tag{4.2.6}$$

where for an L^2 function f on $[0,\infty)$ we use the conventional notation

$$\|f\|_{L^2(\mathbb{R}_+)} := \left(\int_0^\infty f^2(s)\,ds\right)^{1/2}.$$

Although this stability theorem is stated in terms of objects that are not part of the problem data[4], it still offers some support to our intuition that provided the underlying deterministic system is stable (the first condition in (4.2.6)) and the perturbation term is "*small*" (the second condition in (4.2.6)), then the unperturbed stochastic equation remains asymptotically stable in a mean square sense.

We now introduce the perturbed equation which is defined on the same probability

[3]This result excludes the pathological case in which deterministic solutions solve the stochastic equation i.e., $G([x_0]_t) \equiv 0$ for all initial functions, meaning the solution is no longer stochastic. We also exclude such cases throughout this chapter.

[4]It should be noted that upon further analysis, these conditions can be expressed as explicit conditions on the measures μ and ν. This however is not the aim of this Thesis and will be addressed in a future work.

space introduced above

$$dX(t) = \left(f(t) + \int_{[-\tau,0]} X(t+s)\nu(ds)\right) dt + \left(\sigma(t) + \int_{[-\tau,0]} X(t+s)\mu(ds)\right) dB(t),\ t \geq 0, \tag{4.2.7}$$
$$X(t) = \psi(t), \quad t \leq 0,$$

where $f, \sigma \in C(\mathbb{R}_+; \mathbb{R})$ are deterministic functions, and ψ has the same properties as in the solution of (4.2.3). As with condition (4.2.1), *we will assume this continuity and determinism of f and σ throughout the chapter without further reference.* For existence and uniqueness of solutions with finite second moments of (4.2.7), in the sense described above for (4.2.3), we refer the reader to the monograph by Mao [85]. For completeness we recall some notation for the perturbed deterministic equation

$$\dot{x}(t) = f(t) + \int_{[-\tau,0]} x(t+u)\nu(du), \quad t \geq 0, \tag{4.2.8}$$
$$x(t) = \psi(t), \quad t \leq 0.$$

Using Laplace transforms one can readily obtain a variation of constants formula for solutions of (4.2.8) namely

$$x(t,\psi) = r(t)\psi(0) + \int_{[-\tau,0]} \left(\int_s^0 r(t+s-u)\psi(u)du\right) \nu(ds) + \int_0^t r(t-s)f(s)ds, \tag{4.2.9}$$

for $t \geq 0$. Recall from Chapter 2 we can extend a finite measure on $[-\tau, 0]$ to $\mathbb{R}_+$ (2.2.5). We do this to make use of standard results for convolutions with finite measures. Finally with our notation for equation (4.2.4) we may rewrite equation (4.2.9) as,

$$x(t,\psi) = x_0(t,\psi) + x_1(t), \quad t \geq 0, \tag{4.2.10}$$

where $x_1(t) \coloneqq (r * f)(t)$. We do this to exploit the fact that x_1 is independent of ψ and non–random, while x_0 is independent of f and random: it also allows us to make use of the results we stated and deduced for x_0 earlier.

To keep notation clean, from this point on we will frequently omit the dependence on the initial condition, will we write $x_0(t) = x_0(t,\psi)$, $x(t) = x(t,\psi)$ and $X(t) = X(t,\psi)$ for solutions of (4.2.4), (4.2.8) and (4.2.7) respectively.

4.3 Volterra Equations for the Mean Square

Following the spirit of Appleby at al. [16] we define a new process

$$Y(t) = \sigma(t) + \int_{[-\tau,0]} X(t+s)\mu(ds), \quad t \geq 0. \tag{4.3.1}$$

This allows us to readily write down a variation of constants formula for solutions of (4.2.7). By Lemma 6.1 from Reiß et al. [105], we have

$$X(t) = \begin{cases} x(t) + \int_0^t r(t-s)Y(s)dB(s), & t \geq 0, \\ \psi(t), & t \in [-\tau, 0], \end{cases} \tag{4.3.2}$$

where r is the resolvent given by equation (4.2.5). Although (4.3.2) does not give an explicit solution for X, it does allow us to readily write down a deterministic Volterra equation for the mean square of X, and also an expression relating the mean square of Y to the mean square of X. In so doing, the question of studying the mean square of the stochastic equation is converted into one of studying the solution of certain deterministic convolution integral equations, to which the extensive— and much more widely understood— theory of deterministic equations can be applied. As such, the following result can be considered *the most important one in this chapter*. This is because it not only forms the basis for the particular asymptotic results derived here, but acts as a springboard in the future to a very complete understanding of the mean square of solutions of perturbed SFDEs, where the perturbations f and σ may have other interesting properties.

Theorem 4.3.1. *Let X be the solution of (4.2.7). Then we have for all $t \geq 0$,*

$$\mathbb{E}[X^2(t)] = \mathbb{E}[x^2(t)] + \int_0^t r^2(t-s)\mathbb{E}[Y^2(s)]ds, \tag{4.3.3}$$

where Y, defined by (4.3.1), obeys for all $t \geq 0$,

$$\mathbb{E}[Y^2(t)] = \mathbb{E}\left[(\sigma(t) + G(x_t))^2\right] + \int_0^t G^2(r_{t-s})\mathbb{E}[Y^2(s)]ds. \tag{4.3.4}$$

Proof of Theorem 4.3.1. Squaring (4.3.2) gives

$$X^2(t) = x^2(t) + 2x(t)\int_0^t r(t-s)Y(s)dB(s) + \left(\int_0^t r(t-s)Y(s)dB(s)\right)^2. \tag{4.3.5}$$

We first consider the second cross term on the right hand side; by letting $t_0 \in [0,t]$ be arbitrary, we define

$$M(t) := 2x(t_0)\int_0^t r(t_0-s)Y(s)dB(s), \quad t\geq 0.$$

Taking expectations and using Itô's isometry one can show

$$\mathbb{E}[|M(t)|] \leq \mathbb{E}[x^2(t_0)] + \int_0^t r^2(t_0-s)\mathbb{E}[Y^2(s)]ds,$$

so $\mathbb{E}[|M(t)|] < \infty$ for all $t \geq 0$. On the other hand, because $x(t_0)$ has finite expectation, is independent of B and is $\mathcal{F}(0)$–measurable, we have that for $t \geq s \geq 0$

$$\mathbb{E}[M(t)|\mathcal{F}(s)] = 2x(t_0)\mathbb{E}\left[\int_0^t r(t_0-u)Y(u)dB(u)\middle|\mathcal{F}(s)\right] = M(s),$$

using the fact that the second factor in M is a standard Itô integral and thus is a martingale. Therefore, M is a martingale and so $\mathbb{E}[M(t)] = 0$ for all $t \geq t_0 \geq 0$. In particular $\mathbb{E}[M(t_0)] = 0$. As t_0 was chosen arbitrarily, this means that

$$\mathbb{E}\left[2x(t)\int_0^t r(t-s)Y(s)dB(s)\right] = 0, \quad t\geq 0.$$

To deal with the squared term in (4.3.5), we proceed similarly. Fix $t_0 \geq 0$ and for $t \geq 0$ define

$$N(t) = \int_0^t r(t_0-s)Y(s)dB(s), \quad t\geq 0.$$

Since X has finite second moments (this follows from Theorem 1.2.1), so does Y, and therefore N is a martingale with finite second moments. Therefore, by Itô's isometry, we have

$$\mathbb{E}[N^2(t)] = \int_0^t r^2(t_0-s)\mathbb{E}[Y^2(s)]ds.$$

Now take $t = t_0$, so that

$$\mathbb{E}\left[\left(\int_0^{t_0} r(t_0-s)Y(s)dB(s)\right)^2\right] = \int_0^{t_0} r^2(t_0-s)\mathbb{E}[Y^2(s)]ds.$$

Since t_0 is arbitrary, we may replace it by t, and therefore taking expectations on both sides of (4.3.5), we get

$$\mathbb{E}[X^2(t)] = \mathbb{E}[x^2(t)] + \int_0^t r^2(t-s)\mathbb{E}[Y^2(s)]ds,$$

as required. Next we prove that $Y(t)$ obeys (4.3.4). Letting $t \geq \tau$, and using Fubini's theorem for stochastic integrals we can show

$$\begin{aligned}
Y(t) &= \sigma(t) + \int_{[-\tau,0]} X(t+s)\mu(ds) \\
&= \sigma(t) + \int_{[-\tau,0]} \left(x(t+s) + \int_0^{t+s} r(t+s-u)Y(u)dB(u) \right) \mu(ds) \\
&= \sigma(t) + G(x_t) + \int_{[-\tau,0]} \int_0^{t+s} r(t+s-u)Y(u)dB(u)\mu(ds) \\
&= \sigma(t) + G(x_t) + \int_0^t \left(\int_{[\max\{-\tau,u-t\},0]} r(t+s-u)\mu(ds) \right) Y(u)dB(u) \\
&= \sigma(t) + G(x_t) + \int_0^t G(r_{t-u})Y(u)dB(u),
\end{aligned}$$

where in the last line we used the fact that $r(t) = 0$ for all $t < 0$. Notice that the integral is of the same form as that in (4.3.2), and that the first term on the right hand side has the same properties as x, namely, independence from B, finite moments and $\mathcal{F}(0)$–measurability. Therefore, we can compute $\mathbb{E}[Y^2(t)]$ by following exactly the same steps as used to compute $\mathbb{E}[X^2(t)]$ above. Doing this, we get

$$\mathbb{E}[Y^2(t)] = \mathbb{E}\left[(\sigma(t) + G(x_t))^2 \right] + \int_0^t G^2(r_{t-s})\mathbb{E}[Y^2(s)]ds,$$

for $t > \tau$. It remains to obtain the corresponding integral equation for $\mathbb{E}[Y^2(t)]$ for $t \in [0,\tau]$. Proceeding as before, we have that

$$\begin{aligned}
Y(t) &= \sigma(t) + \int_{[-\tau,-t]} X(t+s)\mu(ds) + \int_{[-t,0]} X(t+s)\mu(ds) \\
&= \sigma(t) + \int_{[-\tau,-t]} \psi(t+s)\mu(ds) + \int_{[-t,0]} x(t+s)\mu(ds) \\
&\quad + \int_{[-t,0]} \int_0^{t+s} r(t+s-u)Y(u)dB(u)\mu(ds) \\
&= \sigma(t) + G(x_t) + \int_{[-t,0]} \int_0^{t+s} r(t+s-u)Y(u)dB(u)\mu(ds).
\end{aligned}$$

Then by invoking the stochastic Fubini theorem, and once again using the fact that $r(t) = 0$ for all $t \le 0$, we arrive once more at

$$Y(t) = \sigma(t) + G(x_t) + \int_0^t G(r_{t-u})Y(u)dB(u), \quad t \in [0,\tau].$$

Squaring and taking expectations, we get the same expression for $\mathbb{E}[Y^2(t)]$ deduced above on $[\tau,\infty)$, and this completes the proof. $\square$

4.4 Main Results for Asymptotic Stability

With the integral equations for $\mathbb{E}[Y^2]$ and $\mathbb{E}[X^2]$ in hand, we are now ready to present asymptotic results: in this section we present three results characterising certain types of stability for solutions of (4.3.3). Although the equations for the mean square are deterministic, there are two special challenges to meet. Firstly, these integral equations are written in terms of objects such as x, which are not part of the problem data, and our goal is to determine conditions for asymptotic behaviour which can be stated more directly in terms of, and with minimal dependence on, the problem data. Secondly, we wish to present necessary and sufficient conditions for certain types of stability, and we will try to do this by imposing conditions on the perturbing functions f and σ *which do not depend on the resolvent* r*, or the measures* ν *and* μ.

4.4.1 Reformulation and preliminaries

In trying to keep notation as clean as possible we find the following definitions to be useful when proving all results in this section. If we define $Z(t) := (r^2 * \mathbb{E}[Y^2])(t)$, then the equations for the mean square become

$$\begin{aligned}\mathbb{E}[X^2(t)] &= \mathbb{E}[x^2(t)] + Z(t),\\ Z(t) &= \left(r^2 * \mathbb{E}\left[(\sigma + G(x.))^2\right]\right)(t) + (G^2(r.) * Z)(t),\end{aligned} \tag{4.4.1}$$

for $t \geq 0$; the second equation was obtained by taking the convolution with r^2 across equation (4.3.4). Further defining

$$\gamma(t) := \left(r^2 * \mathbb{E}\left[(\sigma + G(x.))^2\right]\right)(t), \quad t \geq 0, \tag{4.4.2}$$

finally yields

$$Z(t) = \gamma(t) + (G^2(r.) * Z)(t), \quad t \geq 0. \tag{4.4.3}$$

Since the behaviour of x depends on that of x_0 and $x_1 = r * f$, which are known directly, the asymptotic behaviour of the mean square of X is clinched by getting the asymptotic behaviour of Z. In this direction, it makes sense to introduce an integral resolvent ρ which is independent of γ, but in terms of which Z can be expressed. Let ρ obey the equation

$$\rho(t) = G^2(r_t) + (G^2(r.) * \rho)(t), \quad t \geq 0, \tag{4.4.4}$$

(see [53, Ch. 2]). Then

$$Z(t) = \gamma(t) + (\rho * \gamma)(t), \quad t \geq 0.$$

We begin this section with a lemma that provides an integrability result on ρ.

To do so, we need first to deal with a special case, in which $\|G(r)\|_{L^2(\mathbb{R}^+)} = 0$. If this is the case, then $G(r_t) = 0$ a.e. $t \geq 0$. Taking Laplace transforms across this equation gives $\hat{\mu}(z)\hat{r}(z) = 0$ for values of $z \in \mathbb{C}$ for which $\mathrm{Re}(z) > v_0(\nu)$ (for these values of z we are guaranteed that $\hat{r}(z)$ is well–defined; since μ is finite and supported on $[-\tau, 0]$, $\hat{\mu}(z)$ is defined for all $z \in \mathbb{C}$). But since for $\mathrm{Re}(z) > v_0(\nu)$, we have $z - \hat{\nu}(z) \neq 0$ and $\hat{r}(z)(z - \hat{\nu}(z)) = 1$, it follows that $\hat{\mu}(z) = 0$ for all $\mathrm{Re}(z) > v_0(\nu)$. This implies that $\mu(E) = 0$ for all Borel sets $E \subseteq [-\tau, 0]$. As a consequence, in the case when $\|G(r)\|_{L^2(\mathbb{R}^+)} = 0$, we have that X obeys the SFDE

$$dX(t) = \left(f(t) + \int_{[-\tau,0]} X(t+s)\nu(ds) \right) dt + \sigma(t)\, dB(t), \quad t \geq 0.$$

Thus X has the representation

$$X(t,\psi) = x_0(t,\psi) + \int_0^t r(t-s)f(s)\, ds + \int_0^t r(t-s)\sigma(s)\, dB(s), \quad t \geq 0,$$

or

$$X(t,\psi) = x(t,\psi) + \int_0^t r(t-s)\sigma(s)\, dB(s), \quad t \geq 0, \quad \text{a.s.}$$

The mean square is given explicitly by

$$\mathbb{E}[X^2(t)] = \mathbb{E}[x^2(t)] + \int_0^t r^2(t-s)\sigma^2(s) ds, \tag{4.4.5}$$

and can be studied by direct deterministic methods. To summarise, we have shown that $\|G(r.)\|_{L^2(\mathbb{R}^+)} = 0$ if and only if μ is almost everywhere zero (the reverse implication is trivial) and in these cases, the mean square is given directly by (4.4.5). We will sometimes need to treat the situation when μ is zero idiosyncratically in our proofs, and in many cases appealing to (4.4.5) directly suffices.

The following lemma is needed when $\|G(r.)\|_{L^2(\mathbb{R}^+)} > 0$; in the case when $\|G(r.)\|_{L^2(\mathbb{R}^+)} = 0$, the lemma is not needed, and a direct appeal to (4.4.5) can be made instead.

Lemma 4.4.1. *Let ρ be the integral resolvent of (4.4.3), $r \in L^2(\mathbb{R}_+)$ and $0 < \|G(r.)\|^2_{L^2(\mathbb{R}_+)} < 1$. Then there is an $\alpha > 0, \alpha' > 0$ such that the function*

$$\Gamma(\lambda) := \int_0^\infty e^{2\lambda s} G^2(r_s) ds, \tag{4.4.6}$$

is well defined for $\lambda \in [0, \alpha)$, and furthermore,

$$\int_0^\infty e^{2\epsilon s}\rho(s) ds < \infty,$$

for all $\epsilon \in [0, \alpha')$ where α' is the unique number such that $\Gamma(\alpha') = 1$.

Proof of Lemma 4.4.1. To prove the first assertion, note the assumption on $r \in L^2(\mathbb{R}_+)$

gives us the estimate

$$|r(t)| \leq Ke^{-\alpha t}, \quad t \geq 0,$$

and for some $K > 0$ and $\alpha > 0$ (where $-\alpha > v_0(\nu)$). One can extend this estimate to $G(r_.)$:

$$|G(r_t)| = \left| \int_{[-\tau,0]} r(t-s)\nu(ds) \right| \leq e^{-\alpha t} \int_{[-\tau,0]} e^{\alpha s} |\nu|(ds) = K' e^{-\alpha t},$$

for some $K' > 0$, thus we have $G^2(r_t) \leq Ce^{-2\alpha t}$. Here we are using the conventional notation $|\nu|$ for the total variation measure of ν, which is a positive and finite measure in $M[-\tau, 0]$ (see e.g., [109, Thm. 6.2]). We also exploit here the fundamental estimate

$$\left| \int_{[-\tau,0]} f(t+s)\nu(ds) \right| \leq \int_{[-\tau,0]} |f(t+s)||\nu|(ds),$$

for measurable functions f, which can be deduced from the case for finite measures on $[0, \infty)$ (see [53, Thm. 3.4.5]).

Using the estimate for $G(r_.)$ in (4.4.6) gives, for all $0 \leq \lambda < \alpha$,

$$\Gamma(\lambda) \leq \int_0^\infty e^{-2s(\alpha-\lambda)} ds < +\infty.$$

On the other hand, Γ is clearly non–decreasing on its maximal interval of existence. Moreover, since $t \mapsto G(r_t)$ is non–trivial, it follows that either there is a finite $\beta > 0$ such that Γ is well defined on $[0, \beta)$ and $\Gamma(\lambda) \to \infty$ as $\lambda \to \beta^-$ or that Γ is defined on $[0, \infty)$ and $\Gamma(\lambda) \to \infty$ as $\lambda \to \infty$. Therefore, Γ is continuous and increasing, $\Gamma(0) < 1$ and $\Gamma(\lambda) \to \infty$ as $\lambda \to \beta^-$ (where $\beta = \infty$ is possible). Thus by the intermediate value theorem, there is a unique $\alpha' < \beta$ such that $\Gamma(\alpha') = 1$. Since Γ is well defined on $[0, \alpha) \subseteq [0, \beta)$ we have $\alpha \leq \beta$. Thus, we may choose any $\epsilon \in (0, \alpha')$ such that $\Gamma(\epsilon) < 1$.

For the second statement, choose such an $\epsilon \in (0, \alpha')$, and scale equation (4.4.4) by $e^{2\epsilon t}$ to get

$$\rho(t)e^{2\epsilon t} = G^2(r_t)e^{2\epsilon t} + \int_0^t e^{2\epsilon(t-s)} G^2(r_{t-s}) \cdot e^{2\epsilon s} \rho(s) ds, \quad t \geq 0.$$

Using the notation $\rho_\epsilon(t) := \rho(t)e^{2\epsilon t}$ and $G^2_\epsilon(r_t) := e^{2\epsilon t} G^2(r_t)$ we can rewrite the above equation as

$$\rho_\epsilon(t) = G^2_\epsilon(r_t) + \int_0^t G^2_\epsilon(r_{t-s}) \rho_\epsilon(s) ds, \quad t \geq 0. \tag{4.4.7}$$

Now integrating equation (4.4.7), applying Fubini's theorem and using the fact that $r(t) = 0$ for $t < 0$, we see that

$$\int_0^\infty \rho_\epsilon(s) ds = \frac{\int_0^\infty G^2_\epsilon(r_s) ds}{1 - \int_0^\infty G^2_\epsilon(r_s) ds},$$

where we have exploited the fact that $\Gamma(\epsilon) = \int_0^\infty G_\epsilon^2(r_s)ds < 1$. Thus

$$\int_0^\infty \rho_\epsilon(s)ds = \int_0^\infty e^{2\epsilon s}\rho(s)ds,$$

is well defined for all $\epsilon \in [0, \alpha')$. $\square$

4.4.2 Statement of main results

With the above preliminaries dispensed with, we are in a position to state our main results on asymptotic behaviour. We consider three types of convergence of solutions, and describe necessary and sufficient conditions on the perturbations f and σ, as well as the underlying unperturbed equation U, for which each type of convergence result holds.

The common theme of the results is two–fold: first of all, in order that solutions X of the perturbed equation have the appropriate behaviour in mean square, it is necessary that the solutions U of the unperturbed equation tend to zero in mean square, and indeed this forms part of the sufficient conditions for convergence of the perturbed equations too. The second common feature is that the behaviour of the perturbing terms f and σ can be quite irregular or "out of control" on a pointwise basis, but nevertheless the mean square of the solution will be well–behaved. Roughly speaking, if σ is such that $\int_t^{t+1} \sigma^2(s)\,ds$ has the appropriate decay property, and $\int_t^{t+\theta} f(s)\,ds$ has the appropriate decay property for all $\theta \in (0, 1]$, then $\mathbb{E}[X^2]$ will have the decay property. In fact, these average conditions on f and σ turn out to be necessary for the appropriate decay property in the mean square of X: such decay in the mean square implies these "sectional averages" of f and σ must have the stipulated decay too.

Theorem 4.4.1. *Let X be the solution to equation (4.2.7). Suppose that ψ obeys (4.2.1). Then the following conditions **(A)** and **(B)** are equivalent:*

(A) *(i)* $r \in L^1(\mathbb{R}_+)$,

(ii) $\|G(r_\cdot)\|_{L^2(\mathbb{R}_+)} < 1$,

(iii) For all $\theta \in (0, 1]$, $\int_t^{t+\theta} f(s)ds \to 0$ *as* $t \to \infty$,

(iv) $\int_t^{t+1} \sigma^2(s)ds \to 0$ *as* $t \to \infty$.

(B) $\lim_{t\to\infty} \mathbb{E}[X^2(t, \psi)] = 0$ *for all* $\psi \in C([-\tau, 0]; \mathbb{R})$.

Note that conditions (i) and (ii) in **(A)** are equivalent to $\mathbb{E}[U^2(t, \psi)] \to 0$ as $t \to \infty$ for all ψ. Therefore, the convergence of the solution in mean square to zero is equivalent to the global asymptotic stability of the unperturbed equation, coupled with the decay properties of f and σ in (iii) and (iv).

We note that the conditions (iii) and (iv) are fulfilled for functions f and σ for which $f(t) \to 0$ as $t \to \infty$ and $\sigma(t) \to 0$ as $t \to \infty$. However, f and σ can be substantially less well–behaved, and still the conditions (iii) and (iv) can be fulfilled.

Recall the example from Section 3.5 in Chapter 3. There we gave an example of a positive spike function g, with the running maximum having an arbitrarily fast rate of growth. Thus the peaks in g can be arbitrarily high, provided it has a sufficiently short duration. Taking $\sigma = \sqrt{g}$ supplies an example with $\limsup_{t\to\infty} |\sigma(t)| = +\infty$ but for which condition (iv) applies.

An example with no sign restrictions is $f(t) = e^{\alpha t}\sin(e^{\beta t})$ for $t \geq 0$ and $0 < \alpha < \beta$. It was shown in section 2.3.1 of Chapter 2 that $\int_t^{t+\theta} f(s)ds = O(e^{-(\beta-\alpha)t})$ as $t \to \infty$ for each $\theta > 0$. Thus condition (iii) applies. We notice that the condition with an absolute value inside the integral is too restrictive. For instance, if f obeys

$$\int_t^{t+\theta} |f(s)|\, ds \to 0 \quad \text{as } t \to \infty, \text{ for any } \theta > 0,$$

then this condition implies $\mathbf{(A)}$(iii), and is equivalent to $\mathbf{(A)}$(iii) when f does not change sign on $[0, \infty)$. However, it can be shown for f obeying (2.3.5) with $\alpha \in (0, \beta)$ the integral $\int_t^{t+\theta} |f(s)|\, ds$ diverges as $t \to \infty$. To see this, write $T = e^{\beta t}$, $A = e^{\beta\theta}$ and note that $\epsilon := 1 - \alpha/\beta \in (0, 1)$; integration by substitution gives

$$\int_t^{t+\theta} |f(s)|\, ds = \frac{1}{\beta}\int_T^{AT} u^{-\epsilon}|\sin(u)|\, du.$$

As $\epsilon \in (0, 1)$, consider intervals on which $|\sin(u)| \geq 1/2$. Bounding the integral below by considering only these intervals, we see that the lower bound grows at a rate

$$C\int_T^{AT} u^{-\epsilon}\, du \geq C'T^{1-\epsilon},$$

where C and C' are T–independent and strictly positive. Since $\epsilon \in (0, 1)$, the lower bound diverges, and hence

$$\int_t^{t+\theta} |f(s)|\, ds \geq C'e^{(\beta-\alpha)t}, \quad t \to \infty.$$

In applications, understanding when convergence to limiting values is exponentially fast is often important. We turn to this next. First, it is not hard to show that the mean square asymptotic stability of the unperturbed equation implies the exponential decay in the mean square of U. It is therefore natural to ask what conditions on f and σ preserve this exponential convergence in the mean square of solutions of (4.2.7). Once again, this exponential convergence occurs if and only if the unperturbed equation is mean square asymptotically stable, and f and σ obey an exponential decay bound.

Theorem 4.4.2. *Let X be the solution to equation* (4.2.7)*. Suppose that ψ obeys* (4.2.1)*. Then the following conditions (**A**) and (**B**) are equivalent:*

*(**A**)* *(i)* $r \in L^1(\mathbb{R}_+)$,

(ii) $\|G(r.)\|_{L^2(\mathbb{R}_+)} < 1$,

(iii) There is a $\beta_1 > 0$ *such that* $\int_0^\infty e^{2\beta_1 s}\sigma^2(s)ds < \infty$.

(iv) There is a $\beta_2 > 0$ *such that* $t \mapsto \left|\int_0^t e^{\beta_2 s} f(s)ds\right|$ *is uniformly bounded.*

(B) $\mathbb{E}[X^2(t;\psi)] \le C^2(\psi, f, \sigma)e^{-2\alpha(f,\sigma)t}$, *for all* $\psi \in C([-\tau, 0];\mathbb{R})$ *with* $\alpha(f,\sigma) > 0$.

We notice once again from conditions (iii) and (iv) that neither f nor σ need to obey pointwise exponential bounds, but that rather they exhibit exponential decay “on average”.

The conditions **(A)**(iii) and **(A)**(iv) are equivalent to conditions which appear stronger, and give more freedom to choose the exponents β_1 and β_2. In fact, **(A)**(iii) and **(A)**(iv) give exponential integrability for all β sufficiently small. This is clear in the condition for σ: if

$$\int_0^\infty e^{2\beta_1 s}\sigma^2(s)ds < \infty,$$

then obviously for all $\beta < \beta_1$,

$$\int_0^\infty e^{2\beta s}\sigma^2(s)ds \le \int_0^\infty e^{2\beta_1 s}\sigma^2(s)ds < \infty.$$

However, this is perhaps less obvious in the case of the condition **(A)**(iii) on f. Assume, as in **(A)**(iii), that there is a $\beta_2 > 0$ and $B > 0$ such that

$$\left|\int_0^t e^{\beta_2 s} f(s)ds\right| \le B, \quad t \ge 0.$$

We will show that this implies

$$\left|\int_0^t e^{\beta s} f(s)ds\right| \le 2B, \quad t \ge 0.$$

for all $\beta \in (0, \beta_2]$.

To prove this claim, let $\beta \in (0, \beta_2]$ and define $u_\beta'(t) = -\beta u_\beta(t) + f(t)$ for $t \ge 0$, with $u_\beta(0) = 0$. Notice that $|u_{\beta_2}(t)| \le Be^{-\beta_2 t}$ for all $t \ge 0$ by hypothesis. Let $\delta_\beta = u_\beta - u_{\beta_2}$. Then

$$\delta_\beta' = -\beta u_\beta + \beta_2 u_{\beta_2} = -\beta(\delta_\beta + u_{\beta_2}) + \beta_2 u_{\beta_2}.$$

Therefore

$$\delta_\beta(t) = \int_0^t e^{-\beta(t-s)}(\beta_2 - \beta)u_{\beta_2}(s)\, ds, \quad t \ge 0.$$

Thus

$$|\delta_\beta(t)| \le B(\beta_2 - \beta)e^{-\beta t}\int_0^t e^{(\beta-\beta_2)s}\, ds \le B(\beta_2 - \beta)e^{-\beta t}\int_0^\infty e^{-(\beta_2-\beta)s}\, ds = Be^{-\beta t}.$$

Therefore $u_\beta(t) = \delta_\beta(t) + u_{\beta_2}(t)$ obeys

$$|u_\beta(t)| \le Be^{-\beta t} + Be^{-\beta_2 t} \le 2Be^{-\beta t}.$$

But since u_β is the convolution of $e^{-\beta t}$ and f, this gives

$$\left| \int_0^t e^{-\beta(t-s)} f(s)\, ds \right| \le 2Be^{-\beta t}, \quad t \ge 0,$$

which gives the desired β–uniform estimate claimed above, namely

$$\left| \int_0^t e^{\beta s} f(s)\, ds \right| \le 2B, \quad t \ge 0, \quad \beta \in (0, \beta_2].$$

Notice the character of the exponential bound in **(B)**: the estimate of the rate of decay α can depend on f and σ (and of course, on r), but it does not depend on ψ: the ψ–dependence is instead confined to the multiplier of the decaying exponential. Of course, we also expect f–, σ– and r–dependence in this multiplier. In the proof, we do not attempt to make a very fine estimate of α: however, scrutiny of the proof suggests that the faster the decay in r, σ and f, larger is the estimate on α, and the faster the rate of mean square convergence to zero.

As pointed out earlier, the function $f(t) = e^{\alpha t}\sin(e^{\beta t})$ for $t \ge 0$, where $0 < \alpha < \beta$, obey an exponentially decaying estimate of the form

$$\int_t^{t+\theta} f(s)\, ds = \mathrm{O}(e^{-(\beta-\alpha)t}), \quad t \to \infty$$

for any choice of $\theta > 0$, despite the fact that f itself is exponentially unbounded. We show now that this exponential decay arises in exactly the form necessary for Theorem 4.4.2. To see this, note for any $\eta > 0$ that

$$\int_0^t e^{\eta s} f(s)\, ds = \int_0^t e^{(\eta+\alpha)s} \sin(e^{\beta s})\, ds = \frac{1}{\beta} \int_1^T u^{(\eta+\alpha)/\beta - 1} \sin(u)\, du$$

where $T = e^{\beta t}$. Integration by parts yields

$$\begin{aligned}\int_1^T u^{(\eta+\alpha)/\beta-1} \sin(u)\, du = -T^{(\eta+\alpha)/\beta-1}\cos(T) + \cos(1)& \\ + \left(\frac{\eta+\alpha}{\beta} - 1\right) \int_1^T u^{(\eta+\alpha)/\beta-2} \cos(u)\, du&.\end{aligned}$$

Therefore, as $T \to \infty$, the right hand side is bounded provided $0 < \eta < \beta - \alpha$, and an

η–independent upper bound can be obtained. Hence, for each $\eta \in (0, \beta - \alpha)$ we have that

$$\left| \int_0^t e^{\eta s} f(s)\, ds \right| \le B, \qquad \text{for all } t \ge 0,$$

so if $\int_0^\infty e^{2\beta_1 s}\sigma^2(s)\, ds < +\infty$ for some $\beta_1 > 0$, then we will have exponential decay in the mean–square (contingent on the unperturbed equation being globally asymptotically stable in the mean–square).

Theorem 4.4.2 shows that if the decay in f and σ is not exponential, we do see exponential convergence in the mean square. However, in applications it is often of interest to know if solutions are integrable in the mean square. Thus, we ask what conditions are necessary and sufficient for

$$\int_0^\infty \mathbb{E}[X^2(t, \psi)]\, dt < +\infty.$$

As in previous theorems, the asymptotic mean square stability of the unperturbed equation is essential. But since this implies also the exponential decay to zero of the mean square of U, this means that

$$\int_0^\infty \mathbb{E}[U^2(t, \psi)]\, dt < +\infty,$$

is necessary for the mean square integrability of X. Moreover, this mean square integrability will be preserved provided f and σ satisfy the appropriate square integrability conditions.

Theorem 4.4.3. *Let X be the solution to equation (4.2.7). Suppose that ψ obeys (4.2.1). Then the following conditions **(A)** and **(B)** are equivalent:*

(A) *(i)* $r \in L^1(\mathbb{R}_+)$,

(ii) $\|G(r.)\|_{L^2(\mathbb{R}_+)} < 1$,

(iii) $t \mapsto \int_t^{t+\theta} f(s)ds \in L^2(\mathbb{R}_+)$ *for* $\theta > 0$,

(iv) $\sigma \in L^2(\mathbb{R}_+)$.

(B) $\mathbb{E}[X^2(\cdot;\psi)] \in L^1(\mathbb{R}_+)$, *for all* $\psi \in C([-\tau, 0]; \mathbb{R})$.

Notice that $\sigma \in L^2(\mathbb{R}_+)$ is equivalent to

$$t \mapsto \int_t^{t+1} \sigma^2(s)\, ds \in L^1(\mathbb{R}_+)$$

so the condition on σ can still be framed in terms of the average over unit intervals, as in earlier theorems.

4.5 Conclusion

In this chapter we gave characterisations of convergence to zero, integrability and exponential convergence to zero, of the mean square of scalar linear perturbed stochastic functional differential equations with multiplicative noise. Using the machinery developed in both Chapter 2 and Chapter 3 we identified the exact criteria on the deterministic forcing functions required to obtain the desired asymptotic behaviour from solutions. As previously discussed, pointwise estimates on the perturbation functions are not required. Instead we impose conditions on certain averages over compact sets which allows for highly irregular behaviour on a pointwise basis.

Generalising these results for multi- dimensional equations may require alternative methods as a direct representation for the mean square in terms of the coupled system of Volterra integral equations is no longer available. The first step would be to study the the multi-dimensional version of the unperturbed equation from the paper of Appleby, Mao and Riedle [16], for which a complete characterisation of the mean square is still an open problem.

4.6 Proofs

Proof of Theorem 4.4.1. We begin with $(\mathbf{A}) \implies (\mathbf{B})$.

Let $Z(t) = (r^2 * \mathbb{E}[Y^2])(t)$. Then, as pointed out before we have

$$\mathbb{E}[X^2(t)] = \mathbb{E}[x^2(t)] + Z(t), \quad t \geq 0,$$
$$Z(t) = \left(r^2 * \mathbb{E}\left[(\sigma + G(x_.))^2\right]\right)(t) + (G^2(r_.) * Z)(t), \quad t \geq 0.$$

Define $x_1(t) = 0$ for $t \leq 0$ and $x_1(t) = (r * f)(t)$ for $t \geq 0$. Then

$$x_1'(t) = \int_{[-\tau,0]} x_1(t+u)\nu(du) + f(t), \quad t > 0. \tag{4.6.1}$$

As shown in Chapter 2, $(\mathbf{A})$(iii) is equivalent to $x_1(t) \to 0$ as $t \to \infty$. This gives $\mathbb{E}[x^2(t)] \to 0$ as $t \to \infty$. To see this, consider the bound

$$\mathbb{E}[x^2(t)] \leq 2\mathbb{E}[x_0^2(t)] + 2x_1^2(t).$$

We have shown $x_1(t) \to 0$ as $t \to \infty$ and we know condition $(\mathbf{A})$(i) ensures $\mathbb{E}[x_0^2(t)] \leq Ce^{-2\alpha t}$ for some $\alpha < -v_0(\nu)$. Hence $\mathbb{E}[x^2(t)] \to 0$ as $t \to \infty$, as claimed. Next we want to show $Z(t) \to 0$ as $t \to \infty$. Recall

$$Z(t) = \gamma(t) + (G^2(r_.) * Z)(t), \quad t \geq 0,$$

with $\gamma(t) = \left(r^2 * \mathbb{E}\left[(\sigma + G(x.))^2\right]\right)(t)$. We have the immediate inequality

$$\gamma(t) \leq \left(r^2 * \left(2\sigma^2 + 2\mathbb{E}[G^2(x.)]\right)\right)(t).$$

Recall $|r(t)| \leq Ke^{-\alpha t}$ for all $t \geq 0$ and some constant $K, \alpha > 0$ (condition **(A)**(i) ensures we can always do this).

Thus we have

$$(r^2 * \sigma^2)(t) = \int_0^t r^2(t-s)\sigma^2(s)ds \leq K^2 \int_0^t e^{-2\alpha(t-s)}\sigma^2(s)ds.$$

Then once again the results from Chapter 2 ensure the right-hand side converging to zero is equivalent to **(A)**(iv). Hence $(r^2 * \sigma^2)(t) \to 0$ as $t \to \infty$. Now we need only show $\mathbb{E}[G^2(x_t)] \to 0$ as $t \to \infty$ which will ensure $\gamma(t) \to 0$ as $t \to \infty$. Notice in the case when μ is zero that this is automatically true, and that we can conclude that $\gamma(t) \to 0$ as $t \to \infty$ directly. Dealing with the case of non–trivial μ, observe that $G(x_t) = G([x_0]_t) + G([x_1]_t)$ which yields the inequality

$$G^2(x_t) \leq 2G^2([x_0]_t) + 2G^2([x_1]_t).$$

Now

$$|G([x_0]_t)| \leq \int_{[-\tau,0]} |x_0(t+u)||\mu|(du),$$

and so

$$\begin{aligned}
G^2([x_0]_t) &\leq \left(\int_{[-\tau,0]} |x_0(t+u)||\mu|(du)\right)^2 \\
&= \int_{[-\tau,0]}\int_{[-\tau,0]} |x_0(t+s)| \cdot |x_0(t+u)||\mu|(du)|\mu|(ds) \\
&\leq \int_{[-\tau,0]}\int_{[-\tau,0]} \left(\frac{1}{2}x_0^2(t+s) + \frac{1}{2}x_0^2(t+u)\right)|\mu|(du)|\mu|(ds) \\
&= |\mu|\left([-\tau,0]\right) \cdot \int_{[-\tau,0]} x_0^2(t+s)|\mu|(ds).
\end{aligned}$$

Thus, by taking expectations and using the exponential estimate $\mathbb{E}[x_0^2(t)] \leq C(\psi)e^{-2\alpha t}$ implied by condition **(A)**(i), we see that

$$\begin{aligned}
\mathbb{E}[G^2([x_0]_t)] &\leq C(\psi)e^{-2\alpha t}|\mu|\left([-\tau,0]\right) \cdot \int_{[-\tau,0]} e^{-2\alpha s}|\mu|(ds) \\
&\leq K(\psi)e^{-2\alpha t},
\end{aligned}$$

for some constant K. Thus we have $\mathbb{E}[G^2([x_0]_t)] \to 0$ as $t \to \infty$. Next, write $\tilde{\mu}$ in terms of μ as in (2.2.5). Then

$$G([x_1]_t) = (\tilde{\mu} * x_1)(t).$$

As shown above, $x_1(t) \to 0$ as $t \to \infty$, and as $\mu \in M([0,\infty))$ is finite, we have once again that the above convolution tends to zero as $t \to \infty$. Hence $G^2([x_1]_t) \to 0$ as $t \to \infty$. This now gives us that $\gamma(t) \to 0$ as $t \to \infty$.

Next we make use of a variation of constants formula obeyed by Z (see [53, Thm. 2.3.5]):

$$Z(t) = \gamma(t) + (\gamma * \rho)(t),$$

where ρ is defined by (4.4.4). Conditions **(A)** (i) and (ii) satisfy the assumptions of Lemma 4.4.1 and thus $\rho \in L^1(\mathbb{R}_+)$. This along with $\gamma(t) \to 0$ as $t \to \infty$ ensures that $Z(t) \to 0$ as $t \to \infty$ which completes the proof of the forward implication (**A**) $\implies$ (**B**).

In the case when μ is trivial, we note that

$$\mathbb{E}[X^2(t)] = \mathbb{E}[x^2(t)] + (r^2 * \sigma^2)(t), \quad t \geq 0.$$

We have shown above that both terms on the right hand side tend to zero as $t \to \infty$, so the implication (**A**) $\implies$ (**B**) proven in this case also. We now show (**B**) $\implies$ (**A**).

Step 1: (**B**) $\implies$ **(A)**$(i)\&(iii)$:

Equation (4.3.3) immediately tells us that $\mathbb{E}[X^2(t,\psi)] \geq \mathbb{E}[x^2(t,\psi)]$ for all ψ. Thus for deterministic ψ we have $x^2(t,\psi) \to 0$ as $t \to \infty$. Note x solves equation (3.4.1) with the noise term switched off, hence we can invoke Theorem 3.4.2 and claim both **(A)**(i) and (iii) hold.

Step 2: (**B**) $\implies$ **(A)**(iv) :

Let Z be defined as above. Since we have $\mathbb{E}[X^2(t)] \geq Z(t)$, we automatically have $Z(t) \to 0$ as $t \to \infty$, and in the same way we also have $\gamma(t) \to 0$ as $t \to \infty$ for all ψ. Now fix ψ to be deterministic: then $\gamma(t) \to 0$ as $t \to \infty$ implies

$$\int_0^t r^2(t-s)\,(\sigma(s) + G(x_s))^2\,ds \to 0, \quad t \to \infty.$$

The continuity of r (and the fact $r(0) = 1$) ensures that for all $k \in [0,1)$, there exists an η_k such that $r^2(t) \geq k$ for all $t \in [0,\eta_k]$. Thus for $t \geq \eta_k$ we have

$$\int_{t-\eta_k}^t r^2(t-s)\,(\sigma(s) + G(x_s))^2\,ds \geq \int_{t-\eta_k}^t k\,(\sigma(s) + G(x_s))^2\,ds \geq 0.$$

But we also have that

$$\int_{t-\eta_k}^{t} r^2(t-s)\left(\sigma(s)+G(x_s)\right)^2 ds \le \int_0^t r^2(t-s)\left(\sigma(s)+G(x_s)\right)^2 ds,$$

so we must have

$$\int_{t-\eta_k}^{t} \left(\sigma(s)+G(x_s)\right)^2 ds \to 0, \text{ as } t\to\infty.$$

If we can select a $k\in[0,1)$ such that $\eta_k=1$, we arrive at

$$\int_{t-1}^{t} \left(\sigma(s)+G(x_s)\right)^2 ds \to 0, \quad t\to\infty.$$

If no such k can be selected, we proceed as follows. We have for some $k\in[0,1)$ that $\eta_k<1$ (othwerwise there is nothing to prove) and that

$$\int_{t-\eta_k}^{t} \left(\sigma(s)+G(x_s)\right)^2 ds \to 0, \quad t\to\infty.$$

Replacing t by $t-\eta_k$ yields

$$\int_{t-2\eta_k}^{t-\eta_k} \left(\sigma(s)+G(x_s)\right)^2 ds \to 0, \quad t\to\infty,$$

and combining these limits gives

$$\int_{t-2\eta_k}^{t} \left(\sigma(s)+G(x_s)\right)^2 ds \to 0 \quad t\to\infty.$$

We continue in this manner until we find an $n\in\mathbb{N}$ such that $n\eta_k>1$: this then implies

$$\begin{aligned}\int_{t-1}^{t} \left(\sigma(s)+G(x_s)\right)^2 ds &\le \int_{t-n\eta_k}^{t} \left(\sigma(s)+G(x_s)\right)^2 ds\\ &= \sum_{j=1}^{n}\int_{t-j\eta_k}^{t-(j-1)\eta_k} \left(\sigma(s)+G(x_s)\right)^2 ds.\end{aligned}$$

Passing to the limit we see

$$\int_{t-1}^{t} \left(\sigma(s)+G(x_s)\right)^2 ds \to 0, \text{ as } t\to\infty,$$

which holds irrespective of the value of η_k. Notice in the case when μ is trivial that $G(x_t)=0$ for all $t\ge 0$, and so **(A)** (iv) holds automatically.

We continue now in the case when μ is non–trivial, in which case we cannot expect $G(x_t)$ to automatically be zero. We have already shown that **(B)** implies $x(t)\to 0$ as $t\to\infty$, which in turn ensures $G(x_t)\to 0$ as $t\to\infty$. Thus $\int_{t-1}^{t} G^2(x_s)ds\to 0$ as $t\to\infty$, so we get

$$\int_{t-1}^{t} \sigma^2(s)ds + \int_{t-1}^{t} 2\sigma(s)G(x_s)ds \to 0, \quad t \to \infty. \tag{4.6.2}$$

Note by the Cauchy–Schwarz inequality that we have

$$\left|\int_{t-1}^{t} 2\sigma(s)G(x_s)ds\right|^2 \leq \int_{t-1}^{t} \sigma^2(s)ds \int_{t-1}^{t} 4G^2(x_s)ds, \tag{4.6.3}$$

and that the second integral on the right hand side tends to zero as $t \to \infty$. Thus, using (4.6.2) and (4.6.3), for all $\epsilon > 0$, there exists $T_1(\epsilon)$ and $T_2(\epsilon)$ such that for all $t > T(\epsilon) := \max\{T_1(\epsilon), T_2(\epsilon)\}$ we have

$$\left|\int_{t-1}^{t} \sigma^2(s)ds + \int_{t-1}^{t} 2\sigma(s)G(x_s)ds\right| < \epsilon \text{ and } \left|\int_{t-1}^{t} 2\sigma(s)G(x_s)ds\right| < \epsilon \cdot \sqrt{\int_{t-1}^{t} \sigma^2(s)ds.}$$

Hence for $t \geq T(\epsilon)$, we have

$$\begin{aligned}\int_{t-1}^{t} \sigma^2(s)ds &= \left|\int_{t-1}^{t} \sigma^2(s)ds + \int_{t-1}^{t} 2\sigma(s)G(x_s)ds - \int_{t-1}^{t} 2\sigma(s)G(x_s)ds\right| \\ &< \epsilon + \epsilon \cdot \sqrt{\int_{t-1}^{t} \sigma^2(s)ds.}\end{aligned}$$

Let $p(x) := x^2 - \epsilon x - \epsilon$. With $x := \sqrt{\int_{t-1}^{t} \sigma^2(s)ds} \geq 0$, we have $p(x) < 0$. This implies

$$|x| < \frac{\epsilon + \sqrt{\epsilon^2 + 4\epsilon}}{2}.$$

Hence

$$\int_{t-1}^{t} \sigma^2(s)ds < \frac{\epsilon + \sqrt{\epsilon^2 + 4\epsilon}}{2}, \quad t \geq T(\epsilon).$$

Since ϵ is arbitrary, we have $\int_{t-1}^{t} \sigma^2(s)ds \to 0$ as $t \to \infty$, which is condition **(A)**(iv).

Step 3: **(B)** $\implies$ **(A)**(ii) :

The first case we must consider is when the measure μ is the zero measure. This gives $G(r.) \equiv 0$ so that $\|G(r.)\|_{L^2(\mathbb{R}_+)} = 0$ is automatically less than one, and **(A)**(ii) automatically holds.

From here we exclude the case where μ is the zero measure. Since **(B)** holds, we have that $r(t) \to 0$ as $t \to \infty$, and indeed that $r \in L^1(\mathbb{R}_+)$. Since μ is a finite measure, we therefore have that $G(r.)$ is in $L^1(\mathbb{R}_+)$. Also, the fact that $r(t) \to 0$ as $t \to \infty$ implies that $G(r_t) \to 0$ as $t \to \infty$. Therefore, we have that $G^2(r.) \in L^2(\mathbb{R}_+)$. Hence $\|G(r.)\|_{L^2(\mathbb{R}_+)} < +\infty$. On the other hand, since we are now considering non–trivial μ, our earlier arguments show that $\|G(r.)\|_{L^2(\mathbb{R}_+)} > 0$.

In the proof of Step 3, we deduced that $\gamma(t) \to 0$ as $t \to \infty$ so our first task will be to show that $\gamma(t)$ is in fact strictly positive on some non–trivial interval. There are two cases we must consider.

Case 1: $\sigma(t) + G([x_1]_t) \not\equiv 0$ for all $t \geq 0$.

Take $\psi \equiv 0$ so that

$$\sigma(t) + G(x_t) = \sigma(t) + G([x_0]_t) + G([x_1]_t) = \sigma(t) + G([x_1]_t) =: \gamma_1(t),$$

and hence $\gamma(t) = (r^2 * \gamma_1^2)(t)$ for $t \geq 0$. Note as γ_1 is continuous (and not identically equal to zero) there exists an interval $(t_1, t_2) \subset [0, \infty)$ such that $\gamma_1^2(t) \geq \eta$ for all $t \in (t_1, t_2)$ for some $\eta > 0$. Let $t \geq \theta$, where $\theta > 0$ will be chosen later: then

$$\gamma(t) = \int_0^t r^2(t-s)\gamma_1^2(s)ds \geq \int_{t-\theta}^t r^2(t-s)\gamma_1^2(s)ds \geq \inf_{u\in[0,\theta]} r^2(u) \int_{t-\theta}^t \gamma_1^2(s)ds.$$

Now we choose θ small enough such that $t_1 + \theta < t_2$ and that $\inf_{u\in[0,\theta]} r^2(u) \geq \frac{1}{2}$. Now choose t such that $t \in (t_1 + \theta, t_2)$, which means we have,

$$\gamma(t) \geq \inf_{u\in[0,\theta]} r^2(u) \cdot \eta\theta \geq \frac{\eta\theta}{2}.$$

Case 2: $\sigma(t) + G([x_1]_t) \equiv 0$ for all $t \geq 0$.

Since μ is non–zero, we may choose ψ so that $\int_{[-\tau,0]} \mu(dt)\psi(t) \neq 0$. This ensures $G([x_0]_0) \neq 0$. Then we get

$$\sigma(t) + G(x_t) = \sigma(t) + G([x_0]_t) + G([x_1]_t) = G([x_0]_t) =: \gamma_2(t).$$

The continuity of γ_2 and the fact $\gamma_2(0) \neq 0$ ensures $\gamma_2^2(t) > \eta'$ for all $t \in [0, t_2')$ for some $t_2' > 0$. Thus by a similar argument as in **Case 1** we must have γ strictly positive on some non–trivial interval.

Thus we have concluded when $\mu \not\equiv 0$, we can find a deterministic ψ such that there exists an $\tilde{\eta} > 0$ and an interval (t_1', t_2') so that $\gamma(t, \psi) \geq \tilde{\eta} > 0$ for all $t \in (t_1', t_2')$. Recall $Z(t) = \gamma(t) + (\gamma * \rho)(t)$, where ρ is defined as in (4.4.4). Now suppose $\|G(r.)\|_{L^2(\mathbb{R}_+)} \geq 1$: by the Renewal Theorems 3.1.4 and 3.1.5 in Alsmeyer [2], there exists a $\lambda \geq 0$ such that

$\rho(t)/e^{\lambda t} \to c > 0$ as $t \to \infty$. Let $T > 0$ be arbitrary and choose $t \geq T$. Then

$$\frac{Z(t)}{e^{\lambda t}} \geq \frac{1}{e^{\lambda t}} \int_0^T \rho(t-s)\gamma(s)ds = \int_0^T \left[\frac{\rho(t-s)}{e^{\lambda(t-s)}} - c\right] e^{-\lambda s}\gamma(s)ds + c\int_0^T e^{-\lambda s}\gamma(s)ds.$$

Hence

$$\liminf_{t\to\infty} e^{-\lambda t} Z(t) \geq c\int_0^T e^{-\lambda s}\gamma(s)ds.$$

By hypothesis, $Z(t) \to 0$ as $t \to \infty$. Using this and the fact that $c > 0$, we must have

$$\int_0^T e^{-\lambda s}\gamma(s)ds = 0, \text{ for all } T > 0.$$

But since γ is strictly positive on a non–trivial interval, the above integral cannot be zero and hence we reach our desired contradiction. Thus $\|G(r.)\|_{L^2(\mathbb{R}_+)} < 1$ which proves condition **(A)**(ii) and hence the reverse implication **(B)** implies **(A)**. Since we already proved that **(A)** implies **(B)**, the proof is complete. □

Proof of Theorem 4.4.2. We first show $(\mathbf{A}) \implies (\mathbf{B})$.

With $Z(t,\psi)$ defined as in the proof of Theorem 4.4.1 we have $\mathbb{E}[X^2(t)] = \mathbb{E}[x^2(t)] + Z(t)$. The first object we study is $\mathbb{E}[x^2(t)]$, which can be estimated by $\mathbb{E}[x^2(t)] \leq 2\mathbb{E}[x_0^2(t)] + 2x_1^2(t)$. Condition **(A)**(i) ensures $\mathbb{E}[x_0^2(t)] \leq C(\psi)e^{-2\alpha t}$ where $\alpha > 0$ is such that $|r(t)| \leq Ke^{-\alpha t}$ for $t \geq 0$ and some constant $K > 0$. Next we let u be the solution of

$$u'(t) = -\beta_2 u'(t) + f(t), \quad t \geq 0, \tag{4.6.4}$$

with $u(0) = 0$, where β_2 is chosen such that condition **(A)**(iv) holds. We can now estimate $u(t)$ for $t \geq 0$:

$$|u(t)| \leq e^{-\beta_2 t}\left|\int_0^t e^{\beta_2 s} f(s)ds\right| \leq C(f)e^{-\beta_2 t}.$$

Define $\delta = x_1 - u$, where u is understood now to solve (4.6.4). Then $\delta(t) = 0$ for $t \leq 0$. Write $\tilde{\nu}$ in terms of ν as in (2.2.5) so that

$$\delta'(t) = (x_1 * \nu)(t) + \beta_2 u(t) = (\delta * \nu)(t) + (\tilde{\nu} * u) + \beta_2 u(t).$$

With $v(t) := (\tilde{\nu} * u) + \beta_2 u(t)$, δ obeys the variation of constants formula

$$\delta(t) = \int_0^t v(t-s)r(s)ds, \quad t \geq 0.$$

Observe that v depends only on u and ν and so it too obeys an exponential estimate of the form $|v(t)| \le C(f)e^{-\beta_2 t}$ for $t \ge 0$. Combining this with the exponential estimate on r and the equation above for δ we obtain

$$|\delta(t)| \le C(f)Ke^{-\min(\alpha-\epsilon,\beta_2)t}, \quad t \ge 0,$$

for arbitrarily small $\epsilon < \alpha$. This estimate follows from

$$|\delta(t)| \le C(f)K\int_0^t e^{-\beta_2(t-s)}e^{-\alpha s}\,ds,$$

and by estimating the integral in the cases $\alpha \ge \beta_2$ and $\alpha < \beta_2$. Now we can combine our estimates on δ and u to get an estimate on x_1. Thus we have proved the estimate

$$|x_1(t)| \le C'(f)e^{-\min(\alpha-\epsilon,\beta_2)t}, \quad t \ge 0.$$

Notice that this also gives the estimate

$$|G([x_1]_t)| \le C_2(f)e^{-\min(\alpha-\epsilon,\beta_2)t}, \quad t \ge 0.$$

Putting together the estimates for x_1 and $\mathbb{E}[x_0^2]$ yields

$$\mathbb{E}[x^2(t,\psi)] \le C(f,\psi)e^{-2\min(\alpha-\epsilon,\beta_2)t}, \quad t \ge 0.$$

Next we focus on $Z(t)$; let ρ and γ be defined as in (4.4.4) and (4.4.2) respectively. We first estimate γ:

$$\gamma(t) \le 2\int_0^t r^2(t-s)\sigma^2(s)ds + 2\int_0^t r^2(t-s)\mathbb{E}[G^2(x_s)]ds. \tag{4.6.5}$$

Considering the first term on the right hand side of (4.6.5), condition **(A)**(iii) ensures there exists a β_1 such that

$$\int_0^\infty e^{2\beta s}\sigma^2(s)ds < \infty, \text{ for all } \beta \le \beta_1.$$

Thus

$$\int_0^t r^2(t-s)\sigma^2(s)ds \le K^2e^{-2\alpha t}\int_0^t e^{2\alpha s}\sigma^2(s)\,ds. \tag{4.6.6}$$

When $\alpha \le \beta_1$, the integral is uniformly bounded, and the first term grows no faster than $e^{-2\alpha t}$. When $\alpha > \beta_1$, we have

$$\begin{aligned}\int_0^t r^2(t-s)\sigma^2(s)ds &\le K^2e^{-2\alpha t}\int_0^t e^{2\alpha s}\sigma^2(s)\,ds = K^2e^{-2\alpha t}\int_0^t e^{2(\alpha-\beta_1)s}e^{2\beta_1 s}\sigma^2(s)\,ds \\ &\le K^2e^{-2\alpha t}e^{2(\alpha-\beta_1)t}\int_0^t e^{2\beta_1 s}\sigma^2(s)\,ds \le K^2e^{-2\alpha t}e^{-2\beta_1 t}\int_0^\infty e^{2\beta_1 s}\sigma^2(s)\,ds,\end{aligned}$$

so that the integral is $O(e^{-2\beta_1 t})$ as $t \to \infty$. Thus, we have that $(r^2 * \sigma^2)(t) = O(e^{-2\min(\alpha,\beta_1)t})$ as $t \to \infty$.

At this moment, we have enough information to conclude the proof in the case when μ is zero, so we halt the general argument to dispense with this trivial case. We have already obtained the estimate

$$\mathbb{E}[x^2(t,\psi)] \le C(f,\psi)e^{-2\min(\alpha-\epsilon,\beta_2)t}, \quad t \ge 0,$$

and we have just shown that $(r^2 * \sigma^2)(t) \le C(\sigma)e^{-2\min(\alpha,\beta_1)t}$. Therefore

$$\mathbb{E}[X^2(t)] = \mathbb{E}[x^2(t)] + (r^2 * \sigma^2)(t) \le C(f,\sigma,\psi)e^{-2\min(\alpha-\epsilon,\beta_2,\beta_1)t}, \quad t \ge 0.$$

The exponent on the right hand side depends on f and σ through β_2 and β_1, but there is no dependence in the exponent in ψ.

In the rest of the proof, we concentrate on the case where μ is non–trivial. To control the second term on the right hand side of (4.6.5) we need an estimate on $\mathbb{E}[G^2(x_t)]$; one can get this by following an identical argument in the proof of Theorem 4.4.1. We have

$$\mathbb{E}[G^2([x_0]_t)] \le K(\psi)e^{-2\alpha t}, \quad t \ge 0.$$

Using the estimate earlier obtained for $G([x_1]_t)$, we get

$$G^2([x_1]_t) \le C_2^2(f)e^{-2\min(\alpha-\epsilon,\beta_2)t}, \quad t \ge 0.$$

This implies that

$$\mathbb{E}[G^2(x_t)] \le C(f,\psi)e^{-2\min(\alpha-\epsilon,\beta_2)t}, \quad t \ge 0.$$

Therefore

$$\begin{aligned}\int_0^t r^2(t-s)\mathbb{E}[G^2(x_s)]\,ds &\le C(f,\psi)K^2\int_0^t e^{-2\alpha(t-s)}e^{-2\min(\alpha-\epsilon,\beta_2)s}\,ds \\ &\le C'(f,\psi)e^{-2\min(\alpha-\epsilon,\beta_2)t}.\end{aligned}$$

Hence we have

$$|\gamma(t)| \le C(f,\sigma,\psi)e^{-2\alpha(f,\sigma)t}, \quad t \ge 0,$$

where $\alpha(f,\sigma) := \min(\alpha-\epsilon,\beta_1,\beta_2)$. The f and σ dependence here comes from the f and σ dependence on β_2 and β_1 respectively.

Since we are now tackling the case when μ is non–zero, we note that conditions **(A)**(i) and **(A)**(ii) can be used to apply Lemma 4.4.1, so we have (recall $G_\lambda^2(r_s) = e^{2\lambda s}G^2(r_s)$)

$$\int_0^\infty e^{2\lambda s}\rho(s)ds = \frac{\int_0^\infty G_\lambda^2(r_s)ds}{1-\int_0^\infty G_\lambda^2(r_s)ds} =: K',$$

which is finite for any $\lambda \in [0, \alpha')$ where $\Gamma(\alpha') = 1$. Hence set $\lambda := \min\{\alpha(f,\sigma), \alpha' - \epsilon\}$ for an arbitrarily small ϵ. Here the choice of λ clearly depends on f and σ, while α' depends on ν and μ, but not on ψ. Since $\lambda \leq \alpha$, we have

$$\begin{aligned} Z(t) &= \int_0^t \gamma(t-s)\rho(s)ds + \gamma(t) \\ &\leq C(f,\sigma,\psi)\int_0^t e^{-2\lambda(t-s)}\rho(s)ds + C(f,\sigma,\psi)e^{-2\lambda t} \\ &= C(f,\sigma,\psi)e^{-2\lambda t}\left(\int_0^\infty e^{2\lambda s}\rho(s)ds + 1\right) \\ &\leq C'(f,\sigma,\psi)e^{-2\lambda t}, \end{aligned}$$

where the finiteness of the integral at the penultimate step follows from $\lambda < \alpha'$. Lastly, we have

$$\mathbb{E}[X^2(t)] = \mathbb{E}[x^2(t)] + Z(t) \leq C(f,\psi)e^{-2\min(\alpha-\epsilon,\beta_2)t} + C'(f,\sigma,\psi)e^{-2\lambda t} \leq C''(f,\sigma,\psi)e^{-2\lambda t}.$$

Since $\lambda > 0$ depends on f and σ, but not on ψ, we have completed the proof that **(A)** implies **(B)**.

We now show (**B**) $\implies$ (**A**) :

Step 1: (**B**) $\implies$ **(A)**(i) and **(A)**(ii):

By hypothesis we have $\mathbb{E}[X^2(t,\psi)] \to 0$ as $t \to \infty$ for all ψ and so Theorem 4.4.1 implies conditions **(A)**(i) and **(A)**(ii).

Step 2: (**B**) $\implies$ **(A)**(iv):

Using the fact that $\mathbb{E}[X^2(t,\psi)] \geq \mathbb{E}[x^2(t,\psi)]$, if we choose ψ to be deterministic we obtain $x^2(t,\psi) \leq C(f,\sigma,\psi)e^{-2\alpha t}$ for all $t \geq 0$ (here $\alpha = \alpha(f,\sigma)$). Using equation (4.2.10) and setting $\psi = 0$ gives $x_1^2(t) = x^2(t,0) \leq C(f,\sigma,0)e^{-2\alpha t}$ for $t \geq 0$, which implies $|x_1(t)| \leq Ce^{-\alpha t}$ for $t \geq 0$. Let u be the solution to

$$u'(t) = -\beta_2 u(t) + f(t), \quad t \geq 0,$$

with $u(t) = 0$ for all $t \leq 0$ and $\beta_2 \in (0, \alpha)$. Extend x_1 to be zero for $t < 0$ and define $\delta = x_1 - u$. Using the fact that x_1 obeys (4.6.1), we see that for $t > 0$

$$\delta'(t) = \beta_2 u(t) + (x_1 * \nu)(t) = -\beta_2\delta(t) + \beta_2 x_1(t) + (x_1 * \nu)(t),$$

Define $v(t) := \beta_2 x_1(t) + (x_1 * \nu)(t)$ for $t \geq 0$. We note $|v(t)| \leq Ce^{-\alpha t}$ for all $t \geq 0$ by virtue of the estimate on x_1 above and ν being finite. Solving for the above equation for δ gives

$$|\delta(t)| = \left|\int_0^t e^{-\beta_2(t-s)} v(s)ds\right| \leq Ce^{-\beta_2 t}\int_0^\infty e^{-(\alpha-\beta_2)s}ds.$$

The integral on the right hand side is finite as $\beta_2 \in (0,\alpha)$. Thus we have $|\delta(t)| \leq C'e^{-\beta_2 t}$ for all $t \geq 0$. Once again using the definition of δ and the estimates obtained for x_1 and δ we see that $|u(t)| \leq (C + C')e^{-\beta_2 t}$ for $t \geq 0$. But

$$u(t) = \int_0^t e^{-\beta_2(t-s)s} f(s)ds,$$

and so

$$\left|\int_0^t e^{-\beta_2(t-s)} f(s)ds\right| \leq (C + C')e^{-\beta_2 t}, \quad t \geq 0$$

which implies

$$\left|\int_0^t e^{\beta_2 s} f(s)ds\right| \leq C + C' =: B, \quad t \geq 0.$$

Thus condition **(A)**(iv) is proven.

Step 3: **(B)** $\implies$ **(A)**(iii):

Take ψ to be deterministic. From **(B)**, and using definitions (4.4.1) and (4.4.3), we have straight away that $Ce^{-2\alpha t} \geq \mathbb{E}[X^2(t,\psi)] \geq Z(t,\psi) \geq \gamma(t,\psi)$ for $t \geq 0$, where $\alpha = \alpha(f,\sigma) > 0$. Since ψ is deterministic, this gives

$$\int_0^t r^2(t-s)\left(\sigma(s) + G(x_s)\right)^2 ds \leq Ce^{-2\alpha t}, \quad t \geq 0.$$

The continuity of r and the fact $r(0) = 1$ ensure that there exists an $\eta \in (0,1)$ such that $r^2(t) \geq \frac{1}{2}$ for all $t \in [0,\eta)$. Thus for $t \geq \eta$, we have

$$Ce^{-2\alpha t} \geq \int_0^t r^2(t-s)\left(\sigma(s) + G(x_s)\right)^2 ds \geq \frac{1}{2}\int_{t-\eta}^t \left(\sigma(s) + G(x_s)\right)^2 ds.$$

Now let $m \in \mathbb{N}$ be the minimal integer such that $m\eta \geq 1$. Thus for all $t \geq m\eta := T'$ we have

$$\int_{t-1}^{t} (\sigma(s)+G(x_s))^2\,ds \le \int_{t-m\eta}^{t} (\sigma(s)+G(x_s))^2\,ds = \sum_{j=0}^{m-1}\int_{t-(j+1)\eta}^{t-j\eta} (\sigma(s)+G(x_s))^2\,ds$$
$$\le \sum_{j=0}^{m-1} 2Ce^{-2\alpha(t-j\eta)} = C'e^{-2\alpha t}.$$

We have $T' \in [1,2)$. To see this, note that $(m-1)\eta < 1$ so $T' = m\eta < 1+\eta < 2$. Thus the above estimate holds for all $t \ge T'$ and therefore for all $t \ge 1$ (modulo an alternative constant C''). In other words, we have obtained the estimate

$$\int_{t-1}^{t} (\sigma(s)+G(x_s))^2\,ds \le C''e^{-2\alpha(f,\sigma)t}, \quad t \ge 1.$$

Next, since $Ce^{-2\alpha t} \ge \mathbb{E}[X^2(t,\psi)] \ge x^2(t,\psi)$ for $t \ge 0$, we have $|x(t,\psi)| \le \sqrt{C}e^{-\alpha t}$ for $t \ge 0$. Thus, as μ is a finite measure on $[-\tau, 0]$, $G(x_t)$ inherits an exponential estimate from $x(t)$, so that $|G(x_t)| \le K'e^{-\alpha t}$ for some $K' > 0$. Next for $t \ge 1$, we get

$$\int_{t-1}^{t} \sigma^2(s)ds + \int_{t-1}^{t} 2\sigma(s)G(x_s)ds \le \int_{t-1}^{t} \sigma^2(s)ds + \int_{t-1}^{t} 2\sigma(s)G(x_s)ds + \int_{t-1}^{t} G^2(x_s)ds$$
$$\le C''e^{-2\alpha t}.$$

Clearly this yields

$$\int_{t-1}^{t} \sigma^2(s)ds - \left|\int_{t-1}^{t} 2\sigma(s)G(x_s)ds\right| \le \int_{t-1}^{t} \sigma^2(s)ds + \int_{t-1}^{t} 2\sigma(s)G(x_s)ds \le C''e^{-2\alpha t}.$$

On the other hand, by the Cauchy–Schwarz inequality, and using the exponential estimate for $|G(x_t)|$, we get

$$\left|\int_{t-1}^{t} 2\sigma(s)G(x_s)ds\right|^2 \le 4\int_{t-1}^{t} \sigma^2(s)ds \cdot \int_{t-1}^{t} G^2(x_s)ds \le K_2e^{-2\alpha t}\int_{t-1}^{t} \sigma^2(s)ds,$$

for some constant $K_2 > 0$. Taking the last two estimates together, this implies

$$\int_{t-1}^{t} \sigma^2(s)ds \le C''e^{-2\alpha t} + \sqrt{K_2e^{-2\alpha t}\int_{t-1}^{t} \sigma^2(s)ds}, \quad t \ge 1.$$

Write $A := C''e^{-2\alpha t}$, $B := \sqrt{K_2e^{-2\alpha t}}$, and consider $p(x) := x^2 - Bx - A$ for $x \ge 0$. Putting $x := \sqrt{\int_{t-1}^{t} \sigma^2(s)ds} \ge 0$, we see that $p(x) \le 0$. In general, the constraints $x \ge 0$ and $p(x) \le 0$ imply

$$0 \le x \le \frac{B+\sqrt{B^2+4A}}{2}.$$

Hence there is a $C_3 > 0$ such that

$$\int_{t-1}^{t} \sigma^2(s)ds \leq C_3 e^{-2\alpha t}, \quad t \geq 1.$$

Now let $\beta_1 < \alpha$. Then for $t \geq 1$

$$\int_{t-1}^{t} e^{2\beta_1 s}\sigma^2(s)ds \leq e^{2\beta_1 t}\int_{t-1}^{t} \sigma^2(s)ds \leq C_3 e^{-2t(\alpha-\beta_1)}.$$

In particular for any $n \in \mathbb{N}$,

$$\int_{n-1}^{n} e^{2\beta_1 s}\sigma^2(s)ds \leq C'' e^{-2n(\alpha-\beta_1)}.$$

Therefore

$$\int_0^\infty e^{2\beta_1 s}\sigma^2(s)ds \leq \sum_{n=1}^{\infty} C_3 e^{-2n(\alpha-\beta_1)} = C_3 \frac{e^{-2(\alpha-\beta_1)}}{1-e^{-2(\alpha-\beta_1)}}.$$

Hence $\int_0^\infty e^{2\beta_1 s}\sigma^2(s)ds < \infty$ for all $\beta_1 \in (0, \alpha)$. This completes the proof that **(B)** implies **(A)**, and hence both implications are proven. $\square$

Proof of Theorem 4.4.3. We first show $(\mathbf{A}) \implies (\mathbf{B})$.
Using equation (4.4.1), we need only show $\mathbb{E}[x^2], Z \in L^1(\mathbb{R}_+)$. By Theorem 3.4.1 we have $\mathbb{E}[x^2] \in L^1(\mathbb{R}_+)$.

For Z we concentrate first on the case when μ is non–trivial. Consider Z, recalling that if we let γ and ρ be defined as in (4.4.2) and (4.4.4) respectively, then $Z(t) = \gamma(t)+(\gamma*\rho)(t)$ for $t \geq 0$. Conditions **(A)**(i) and **(A)**(ii) ensure Lemma 4.4.1 is applicable and thus $\rho \in L^1(\mathbb{R}_+)$. We need only show $\gamma \in L^1(\mathbb{R}_+)$ in order to show $Z \in L^1(\mathbb{R}_+)$. We have that

$$\gamma(t) \leq 2(r^2 * \sigma^2)(t) + 2(r^2 * \mathbb{E}[G^2(x)])(t).$$

Conditions **(A)**(i) and **(A)**(iv) imply the first term on the right hand side is in $L^1(\mathbb{R}_+)$: hence if we show $\mathbb{E}[G^2(x.)] \in L^1(\mathbb{R}_+)$, we are done. We estimate $\mathbb{E}[G^2(x_t)]$ as before, according to

$$\begin{aligned} G^2(x_t) &\leq \int_{[-\tau,0]}\int_{[-\tau,0]} |x(t+s)||x(t+u)||\mu|(du)|\mu|(ds) \\ &\leq \int_{[-\tau,0]}\int_{[-\tau,0]} \left(\frac{1}{2}x^2(t+s)+\frac{1}{2}x^2(t+u)\right)|\mu|(du)|\mu|(ds) \\ &= |\mu|([-\tau,0]) \cdot \int_{[-\tau,0]} x^2(t+s)|\mu|(ds). \end{aligned}$$

Taking expectations gives

$$\mathbb{E}[G^2(x_t)] \leq |\mu|([-\tau,0]) \cdot \int_{[-\tau,0]} \mathbb{E}[x^2(t+s)]|\mu|(ds).$$

Since $\mathbb{E}[x^2] \in L^1(\mathbb{R}_+)$ and $|\mu|$ is a finite measure, we have $\mathbb{E}[G^2(x_\cdot)] \in L^1(\mathbb{R}_+)$. This completes the proof of the forward implication when μ is non–trivial.

In the case when $\mu = 0$, we have

$$\mathbb{E}[X^2(t)] = \mathbb{E}[x^2(t)] + (r^2 * \sigma^2)(t), \quad t \geq 0.$$

As before Theorem 3.4.1 ensures $\mathbb{E}[x^2] \in L^1(\mathbb{R}_+)$; on the other hand, conditions **(A)**(i) and **(A)**(iv) imply the second term on the right hand side is in $L^1(\mathbb{R}_+)$, as above. Thus we have shown that (**B**) implies (**A**) in the case of trivial μ also.

We now show (**B**) $\implies$ (**A**) :

Step 1: (**B**) $\implies$ (**A**)(i)&(iii):

Recall that $\mathbb{E}[X^2(t,\psi)] \geq \mathbb{E}[x^2(t,\psi)]$ for all $t \geq 0$ for all ψ. Thus $\mathbb{E}[x^2(\cdot,\psi)] \in L^1([-\tau,\infty))$ for all ψ. Which by Theorem 3.4.1 implies (**A**) (i) & (iii).

Step 2: (**B**) $\implies$ **(A)**(iv):

Once again let ψ be deterministic. We have immediately that $\mathbb{E}[X^2(t,\psi)] \geq Z(t,\psi)$ for $t \geq 0$ and so $Z \in L^1(\mathbb{R}_+)$. Additionally, by the definition of Z along with equation (4.4.3) this forces $\gamma \in L^1(\mathbb{R}_+)$. Define

$$A(t) = \int_0^t (\sigma(s) + G(x_s))^2 \, ds, \quad t \geq 0.$$

Integrating equation (4.4.2), and using Fubini's theorem, we see

$$\int_0^T \gamma(t)\, dt = \int_0^T \int_0^t r^2(s) \left(\sigma(t-s) + G(x_{t-s})\right)^2 \, ds \, dt = \int_0^T A(T-s) r^2(s)\, ds.$$

Therefore we have that there is a $B > 0$ such that

$$\int_0^t A(t-s) r^2(s)\, ds \leq B, \quad t \geq 0.$$

We need to show that A, which is non–negative and non–decreasing, tends to a finite limit; we already know from Step 1 that $r \in L^2(\mathbb{R}_+)$. Suppose to the contrary that $A(t) \to \infty$ as

$t \to \infty$. Then for every $M > 0$, there is a $T(M) > 0$ such that for $t \geq T(M)$, $A(t) \geq M$. Now, for $t \geq T(M)$, we have

$$\int_0^t r^2(t-s)A(s)\,ds \geq \int_T^t r^2(t-s)\,ds \cdot M = \int_0^{t-T} r^2(u)\,du \cdot M.$$

Since $r(0) = 1$ and r is continuous, $\int_0^\infty r^2(s)\,ds > 0$, and we have

$$B \geq \liminf_{t\to\infty} \int_0^t r^2(t-s)A(s)\,ds \geq \int_0^\infty r^2(u)\,du \cdot M.$$

Since $M > 0$ is arbitrary, we may let $M \to \infty$, which gives $+\infty > B = \infty$, a contradiction. Therefore, we must have that A tends to a finite limit, which is nothing other than

$$\int_0^\infty (\sigma(s) + G(x_s))^2\,ds < \infty.$$

We notice now in the case when μ is zero that $G(x.) = 0$, so we have $\sigma \in L^2(\mathbb{R}_+)$ as required.

The rest of the proof is devoted to the case when μ is non-trivial. By the above integrability, there exists a $C > 0$, independent of $T > 0$, such that

$$\int_0^T \sigma^2(s)ds + 2\int_0^T \sigma(s)G(x_s)ds + \int_0^T G^2(x_s)ds \leq C, \quad T > 0.$$

This inequality then implies

$$\int_0^T \sigma^2(s)ds - \left|\int_0^T 2\sigma(s)G(x_s)ds\right| \leq \int_0^T \sigma^2(s)ds + \int_0^T 2\sigma(s)G(x_s)ds \leq C,$$

for all $T > 0$. By the Cauchy–Schwarz inequality, we have

$$\left|\int_0^T 2\sigma(s)G(x_s)ds\right|^2 \leq 4\int_0^T \sigma^2(s)ds \cdot \int_0^T G^2(x_s)ds$$

Next, recall that we have shown that $x \in L^2(\mathbb{R}_+)$ in Step 1 above. Since μ is a finite measure, we have that $G(x.) \in L^2(\mathbb{R}_+)$ also. Therefore, there exists $C' > 0$ such that

$$\left|\int_0^T 2\sigma(s)G(x_s)ds\right|^2 \leq C' \int_0^T \sigma^2(s)ds, \quad T > 0,$$

which leads to

$$\int_0^T \sigma^2(s)ds \leq C + \sqrt{C' \int_0^T \sigma^2(s)ds}, \quad T \geq 0.$$

Define $p(x) := x^2 - \sqrt{C'}x - C$ for $x \in \mathbb{R}$. The inequality $p(x) \leq 0$ is satisfied for $x \geq 0$

provided
$$0 \le x \le \frac{\sqrt{C'} + \sqrt{C' + 4C}}{2}.$$
Therefore, as $p\left(\sqrt{\int_0^T \sigma^2(s)ds}\right) \le 0$, we have
$$\int_0^T \sigma^2(s)ds \le \left[\frac{\sqrt{C'} + \sqrt{C' + 4C}}{2}\right]^2, \quad T > 0,$$
which implies that $\sigma \in L^2(\mathbb{R}_+)$, since the bound on the right hand side is independent of T. Hence we have shown condition **(A)**(iv), as required.

Step 4: (**B**) $\implies$ **(A)**(ii):

Suppose finally, by way of contradiction, that $\|G(r.)\|^2_{L^2(\mathbb{R}_+)} \ge 1$. Take ψ deterministic and arbitrary. We know $\mathbb{E}[X^2] \in L^1(\mathbb{R}_+)$ forces both Z and γ to be in $L^1(\mathbb{R}_+)$. Thus integrating equation (4.4.3), using Fubini's theorem and the fact $r(t) = 0$ for all $t < 0$ we see that
$$\int_0^\infty Z(s)ds = \int_0^\infty \gamma(s)ds + \int_0^\infty G^2(r_u)du \cdot \int_0^\infty Z(s)ds. \tag{4.6.7}$$
If $\|G(r.)\|^2_{L^2(\mathbb{R}_+)} = 1$, this implies $\int_0^\infty \gamma(s)ds = 0$. Recall the argument from the proof of Theorem 4.4.1 which showed one can always choose a deterministic ψ such that there exists an non-trivial interval where $\gamma(t)$ is non zero, provided μ is not the zero measure (which is impossible in this case, since we have assumed $\|G(r.)\|^2_{L^2(\mathbb{R}_+)} > 0$). Choosing such a ψ means that $\int_0^\infty \gamma(s)ds \neq 0$, forcing a contradiction.
If $b^2 := \|G(r.)\|^2_{L^2(\mathbb{R}_+)} > 1$ once again choose a ψ such that γ is strictly positive on a non–trivial interval. Then $\int_0^\infty \gamma(s)\, ds > 0$. Moreover, since $Z(t) \ge \gamma(t)$ we have that $\int_0^\infty Z(s)\, ds > 0$. Therefore from (4.6.7) we get
$$\int_0^\infty Z(s)ds > b \int_0^\infty Z(s)\, ds,$$
and since $b > 1$, we arrive again at a contradiction. Thus we must have $\|G(r.)\|^2_{L^2(\mathbb{R}_+)} < 1$, and the theorem is proven. $\square$

Part II

Optimal Consumption

Chapter 5

A Variational Approach to Portfolio Choice

The material in this chapter is based on the following article:

1. P. Guasoni, E. Lawless and H. M. Tai. A variational approach to portfolio choice. *submitted*, (2025) .

5.1 Introduction and Motivation

The financial planning problem is ubiquitous in modern, developed economies. Economic agents have the goal of being financially secure so as to adequately finance current and future consumption needs. This goal is primarily achieved as follows: an agent receives a (usually fixed) sum of money in exchange for labour. The agent uses this sum to finance her current consumption needs (e.g., rent, food, etc.) while saving or investing the excess to plan for a future time when exchanging labour for money is no longer feasible. The question then is: what are the optimal investment and consumption policies?

When attempting to answer this question, there are many factors one must consider. The first one is how to mathematically formulate the goal of the agent: What does it mean to "*choose*" a consumption and investment policy, and according to what criterion is a chosen policy "*optimal*"? One of the major advances in this field came from Harry Markowitz in 1952, wherein he introduced the idea of mean–variance optimality in his paper *Portfolio Selection* [89]. The asset returns were modelled as random variables, and the investor would specify a pre-determined level of desired risk (measured by the variance of the assets) and then seek to obtain the highest possible expected return given the risk level (or vice versa). Once a solid mathematical foundation had been provided, an abundance of subsequent research followed. In particular, Markowitz's theory ultimately led to the development of the so-called Capital Asset Pricing Model (CAPM) by William F. Sharpe in 1964 [114].

A fundamental drawback of these approaches lies in their static nature. Markets are not fixed: they evolve (essentially) continuously over time, and so a dynamic model was needed. The first major breakthrough in dynamic portfolio choice came from Robert C. Merton in his seminal papers in 1969, *Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case* [90], and in 1971, *Optimum Consumption and Portfolio Rules in a Continuous-Time Model* [91]. Merton moved away from mean–variance optimisation to a more general optimality criterion of expected utility maximisation[1], while simultaneously modelling risky assets as stochastic processes instead of static random variables. Utility functions are real-valued maps that describe agent preferences. They are increasing, i.e., we prefer more to less; and concave, i.e., a fixed absolute increase in wealth means more when we are poor than when we are rich. Merton modelled the risky assets as geometric Brownian motion, meaning the return at time t was described by a deterministic drift and a random diffusion component:

$$R_t = \underbrace{\mu\, t}_{\text{drift}} + \underbrace{\sigma\, B_t}_{\text{diffusion}}$$

where $B_t \sim \mathcal{N}(0,t)$, with $\mu > 0$ being the instantaneous expected excess return and $\sigma > 0$ the instantaneous volatility. The agent invests a proportion π of her wealth X in the risky asset and chooses a consumption policy c, all of which are continuous in time. Thus, the return the agent receives on her wealth at time t evolves according to the dynamics:

$$\underbrace{r\,(1-\pi_t)}_{\text{risk free return}} + \underbrace{\pi_t\,(\mu\, t + \sigma\, B_t)}_{\text{return from risky asset}} - \underbrace{c_t,}_{\text{consumption}}$$

where r denotes the rate of return from the risk-free asset. In Merton's original model, he included a term for exogenous labour income and showed that if labour income is constant, then it could essentially be ignored without loss of generality. However, if labour income is stochastic, then this poses a significant increase in mathematical complexity and is better studied as an extension of the baseline model, in which consumption is financed solely from investment income.

Let U denote a so-called utility function: the agent then wishes to choose the policies π and c that maximise her expected utility from consumption over a time horizon $[0,T]$,

$$\max_{\pi,c} \mathbb{E}\left[\int_0^T e^{-\beta(T-s)} U(c_s) ds + \Phi(X_T)\right].$$

Here, $\beta > 0$ denotes a time preference parameter that indicates the agent's level of impatience, and Φ is a function that measures utility from the agent's unconsumed wealth

[1] One can prove, with a specific choice of utility function, that the two optimisation criteria are equivalent.

at the terminal time T. One can also consider the problem over an infinite horizon (i.e., $T = \infty$ and $\Phi = 0$), which models young agents who have a longer investment horizon.

A major change in this model is the focus on consumption rather than wealth. This stems from the intuition that agents gain utility from consuming goods and services purchased using wealth, rather than from solely obtaining wealth itself. Although this was a serious advance in the portfolio optimisation literature, Merton's model suffered from some highly unrealistic assumptions.

Merton assumed the excess return (μ), volatility (σ) and interest rate (r) were constant, in the sequel we refer to these quantities as the *market coefficients*. The existence of indices tracking both volatility and excess return along with central banks varying interest rates are in direct contradiction with this assumption. These quantities have direct influence on individual agents, central banks use interest rates to influence consumption in the wider economy, markets fluctuate in an unpredictable manner, meaning investment opportunities for financial planners are inherently random. It is imperative that these factors are included in models of optimal consumption and portfolio choice, and in doing so one sees stark differences in optimal polices.

Why optimal policies change in the presence of stochastic investment opportunities is a phenomenon known as intertemporal hedging. When stochastic investment opportunities are added to the market model, the risk averse agent will endeavour to limit her exposure to the sources of noise which cannot be directly traded[2]. It is assumed this extra noise is correlated with the noise coming from the risky assets and hence the agent can alter the proportion of her wealth invested in the risky assets in order to hedge against the randomness now inherent in the market coefficients.

The level of correlation between the noise from the risky assets and the noise from the market coefficients determines how much hedging can be achieved. If this correlation is zero then no hedging is possible, optimal investment policies remain unchanged and the market is said to be fully incomplete. If the correlation is non zero but also imperfect, then the market is said to be incomplete and the agent engages in intertemporal hedging to an extent dependent on the level of correlation. If there exists perfect correlation then the market is said to be complete and all external risk from the stochastic market coefficients can be hedged perfectly. A rigorous formulation of market completeness was provided by Bensoussan [22].

5.2 Overview

Incorporating stochasticity into the market coefficients has been a major focus of the portfolio choice literature since Merton's seminal works and its subsequent generalisation

[2] This refers to the real world situation in which interest rates, volatility and excess returns are not financial assets which can be bought and sold directly in an open market.

to multi-dimensions [67]. Merton extended his own model and formulated the optimal consumption problem with stochastic investment opportunities in [92] but did not provide a general solution. Rather he focused on the connection to mean-variance optimisation and the CAPM.

In the following years, Bellman's dynamic programming approach gained popularity and was given a rigorous treatment in the finite horizon setting by Karatzas et.al. [68], wherein the authors characterise optimal polices in terms of the solution to certain Cauchy problems, but under the assumption of bounded model coefficients.

An alternative approach using martingale methods was provided by Cox and Huang [40] wherein asset prices followed a general diffusion and market coefficients satisfied Lipschitz and linear growth type conditions.

These works provided a framework to theoretically solve the optimal consumption problem in a very general setting. However, implementing these methods in order to infer relevant economic insights from optimal polices is difficult. To overcome this challenge, many researchers began to analyse particular parametric models.

A portfolio choice model with Gaussian and mean reverting interest rates was studied by Korn and Kraft [74] while a similar model with consumption was solved by Munk and Sørensen [96]. Benth, Karlsen and Reikvam [23] considered models with non-Gaussian volatility with consumption, while Kallsen and Muhle-Karbe [66] addressed stochastic volatility in a general setting of affine processes but without consumption. Models with stochastic excess returns have also been considered by Liu and Watcher [78, 121] respectively. For brevity we mention that the optimal consumption/portfolio choice problem when markets are incomplete has also received much attention but is far less understood and will not be the focus of this thesis. (See the following works and the references therein for details [40, 45, 50, 55, 56, 62, 69, 71, 106, 122, 123].)

The best possible result is of course to obtain an explicit, closed form solution. This is very rare and has only been achieved in a handful of specific models [33, 34, 55, 56, 75, 78, 121]. The recent monograph by Rogers [107] and the review paper by Liu and Muhle-Karbe [79] offer a thorough overview of the literature. Without explicit solutions, alternative methods must be used to study the behaviour of optimal policies and the effects of intertemporal hedging.

Many of the above cited works make two simplifying assumptions: (i) they consider only a finite investment horizon; and (ii) they ignore consumption, instead maximising expected utility from terminal wealth. For finite horizon problems one can use the results mentioned above which, characterise optimal policies in terms of solutions to certain PDEs and employ numerical methods. This idea works because, in a finite horizon problem, boundary conditions for the associated Hamilton-Jacobi-Bellman (HJB) PDE are implied by terminal utility. This is not the case with an infinite horizon problem, wherein the HJB equation has no obvious boundary conditions that can be used to characterise the value function.

The contribution of this thesis to the complete market infinite horizon optimal consumption literature is twofold. First, under very mild conditions on the market coefficients, we provide a novel characterisation of optimal polices in terms of an associated (and tractable) variational problem. Second, we provide a simple numerical method which does not require boundary conditions and allows for accurate numerical approximation of optimal policies. We compare the accuracy of our scheme with models possessing known explicit solutions and find that it has an error of less than 0.002% (see Appendix A.1).

Thus under certain assumptions we circumvent the need to study the associated HJB equation. We highlight examples of models with highly non-linear coefficients for which our approach yields both verification and an accurate numerical approximation, while HJB equations have no explicit solutions.

We show that for typical parameter values, the relative difference between consumption plans implied by Merton's model and actually optimal polices can range from -85% to 200% when investment opportunities are stochastic. In each model, when the state variable takes on extreme values, we see optimal consumption policies are far more conservative than their myopic counterparts. The hedging demand depends heavily on the sign of the correlation between the state variable and the risky asset, and in certain models, optimal policies can exhibit non-standard behaviour. In a model with constant interest rates and stochastic risk premia, we show that for highly risk averse agents, it can be optimal to short sell the risky asset despite a strictly positive risk premium. See Section 5.4.1.

Section 5.3 introduces the market model, the agent's objective, the HJB equation and the variational problem. We study a complete market with n risky assets wherein the interest rate, excess return, and volatility all depend on an exogenous state variable which follows an autonomous scalar diffusion[3]. The agent's goal is to maximise expected utility from consumption over an infinite horizon subject to isoelastic preferences. In this setting the HJB equation is a second order linear inhomogeneous (potentially) singular ODE with non constant coefficients, a thorough discussion regarding its analysis thus far in the literature is conducted. We state the associated variational problem, which is derived by reverse engineering, i.e. the functional to be minimised is constructed so that the Euler-Lagrange equation is exactly the HJB equation.

Section 5.4 gives our two main results, the first of which being an existence, uniqueness, and regularity result for the variational problem along with a characterisation of the value function as the unique solution to the variational problem, subject to the well posedness of a certain martingale problem. This result holds under mild C^2-Hölder regularity assumptions on the model coefficients, along with a non-restrictive positivity stipulation

[3]The correlation between the risky assets and the state variable is also allowed to depend on the state variable itself.

on the related myopic consumption policy, and an assumption of the existence of a stationary distribution for the state variable with an adjusted drift. We note this result is essentially model free and all assumptions depend only on the market coefficients.

Our second main result is a verification argument which provides conditions, once again explicitly given in terms of the market coefficients, for the existence and uniqueness of the previously mentioned martingale problem. We close this section by giving three examples of non-trivial models, for which the developed theory is applicable, and provide numerical approximations of the optimal policies.

In section 5.5 we study the asymptotic behaviour of optimal policies and provide a second, alternative, verification result. Under certain pointwise assumptions on the market coefficients, which encapsulate a large proportion of models, we obtain pointwise bounds on the optimal consumption policy and the intertemporal hedging component present in the optimal investment policy.

A discussion and conclusion is provided in section 5.6 while all proofs are presented in section 5.7.

5.3 Model and problem formulation

5.3.1 Market Model

The market includes a safe asset with price $(S_t^0)_{t\geq 0}$ and n risky assets with prices $(S_t^i)_{t\geq 0}$. Investment opportunities are driven by a *state variable* Y which follows an autonomous diffusion. This results in a market with dynamics,

$$\frac{dS_t^0}{S_t^0} = r(Y_t)\,dt, \tag{5.3.1}$$

$$\frac{dS_t^i}{S_t^i} = r(Y_t)\,dt + dR_t^i, \qquad 1 \leq i \leq n. \tag{5.3.2}$$

The cumulative excess returns R^i and the state variable have dynamics

$$dR_t^i = \mu_i(Y_t)\,dt + \sum_{j=1}^{n} \sigma_{ij}(Y_t)\,dZ_t^j, \qquad 1 \leq i \leq n, \tag{5.3.3}$$

$$dY_t = b(Y_t)\,dt + a(Y_t)\,dW_t, \tag{5.3.4}$$

$$d\langle Z^i, W\rangle_t = \rho_i(Y_t)dt, \quad 1 \leq i \leq n; \quad \sum_{i=1}^{n} \rho_i^2(Y_t) = 1. \tag{5.3.5}$$

where $Z = (Z^1, \ldots, Z^n)$ and W are multivariate and scalar Brownian motions respectively. In order to ease notation, the dependence on the state variable Y_t is omitted unless ambiguity arises. All elements of $\mathbb{R}^n$ will be understood as $n \times 1$ matrices with $\top$ denoting matrix transposition. We denote the covariation matrix of the returns processes $(R^i_t)_{1\leq i\leq n}$ as $\Sigma = \sigma\sigma^\top$ and set $\Upsilon = \sigma\rho a$ as the $n\times 1$ covariation matrix between the risky assets and the state variable. The open connected set $E \subseteq \mathbb{R}$ denotes the range of the state variable Y. Our first task is to construct an $\mathbb{R}^n \times E$-valued process, (R, Y) with satisfies (5.3.3), (5.3.4) and (5.3.5). We follow a similar approach as in Guasoni and Robertson [54] and Guasoni and Wang [56]. Let $\Omega = C(\mathbb{R}_+; \mathbb{R}^n \times E)$ (throughout $\mathbb{R}_+ = [0, \infty)$) equipped with the metric,

$$d(\omega_1, \omega_2) = \sum_{n=1}^{\infty} \frac{1}{2^n} \left(\frac{\sup_{0\leq t\leq n} |\omega_1(t) - \omega_2(t)|}{1 + \sup_{0\leq t\leq n} |\omega_1(t) - \omega_2(t)|} \right), \tag{5.3.6}$$

where $|\cdot|$ is the usual Euclidean distance. We shall use $(\mathcal{B}_t)_{t\geq 0}$ to denote the filtration generated by the coordinate process $G_t(\omega) : \mathbb{R}_+ \times \Omega \to \mathbb{R}^n \times E : (t, \omega) \mapsto \omega(t)$. We note under the metric d, $\mathcal{B} := \sigma\left(\bigcup_{t\geq 0} \mathcal{B}_t\right)$ is equal to the Borel sigma algebra on (Ω, d), Stroock and Varadhan [118, 1.3]. The following assumption equips the measurable space $(\Omega, \mathcal{B})$ with a probability measure, from which the market model is constructed. Let $C^{k,\alpha}(E, \mathbb{R}^{n\times m})$ denote the space of $\mathbb{R}^{n\times m}$-valued k times continuously differentiable functions on E whose k^{th} order partial derivatives are α-Hölder continuous for some $\alpha \in (0, 1)$.

Assumption 5.3.1. (Well posedness)

(i) For some $\alpha \in (0, 1)$, $r \in C^{1,\alpha}(E; \mathbb{R})$, $\mu \in C^{1,\alpha}(E; \mathbb{R}^n)$, $b \in C^{1,\alpha}(E; \mathbb{R})$, $a^2 \in C^{2,\alpha}(E; \mathbb{R})$, $\Sigma \in C^{2,\alpha}(E; \mathbb{R}^{n\times n})$ and $\Upsilon \in C^{2,\alpha}(E; \mathbb{R}^n)$. For all $y \in E$, $\Sigma(y)$ is strictly positive definite and $a^2(y) > 0$.

(ii) Let $x = (0, y) \in \mathbb{R}^n \times E$[4]. There exists a unique solution $\mathbb{P}^x$ to the martingale problem[5] on $\mathbb{R}^n \times E$ for,

$$\hat{L} = \frac{1}{2} \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} \hat{A}_{i,j}(\cdot) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n+1} \hat{b}_i(\cdot) \frac{\partial}{\partial x_i}$$
$$\hat{A} = \begin{pmatrix} \Sigma & \Upsilon \\ \Upsilon^\top & a^2 \end{pmatrix}; \quad \hat{b} = \begin{pmatrix} \mu \\ b \end{pmatrix}.$$

Part (i) above ensures $\hat{A}$ and $\hat{b}$ are well defined and sets the minimum amount of regularity for the coefficients which is required in the sequel. Part (ii) ensures the existence of a probability measure $\mathbb{P}^x$ on $(\Omega, \mathcal{B}, \mathcal{B}_t)$ which satisfies $\mathbb{P}^x(G_0 = x) = 1$ and ensures the

[4] We choose the starting point $(0, y)$ to emphasise constructing the process (R, Y) is essentially dependent on the state variable Y only.

[5] In this context a martingale problem for a given generator L (a second order linear elliptic partial differential operator) is a problem in finding a probability measure on the space of continuous functions such that the process $f(x_t) - f(x_0) - \int_0^t Lf(x_s)ds$ is a martingale with respect to the constructed probability measure. The martingale property should hold for all smooth and compactly supported functions f. Here we have used x to denote the coordinate process on the function space of interest.

process $M_t = f(G_t) - f(G_0) - \int_0^t \hat{L} f(G_s) ds$, $t \geq 0$ is a martingale with respect to $(\mathcal{B}_t)_{t\geq 0}$ for any $f \in C^2(\mathbb{R}^n \times E; \mathbb{R})$ with compact support. As the probability measure depends on the starting point $x = (0, y)$, it is henceforth denoted $\mathbb{P}^y$. Note if one defines $\mathcal{B}_{t+} := \bigcap_{s\geq t} \mathcal{B}_s$ then M_t is also an $(\mathcal{B}_{t+}, \mathbb{P}^x)$-martingale. With this observation we claim there exists an $n+1$-dimensional Brownian motion B defined on an extension $(\tilde{\Omega}, \mathcal{F}, \tilde{\mathbb{P}}^y)$ of $(\Omega, \mathcal{B}, \mathbb{P}^y)$ such that (G_t, B_t), $(\tilde{\Omega}, \mathcal{F}, \mathcal{F}_t, \tilde{\mathbb{P}}^y)$ is a weak solution to the following SDE:

$$d\begin{pmatrix} R_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \mu \\ b \end{pmatrix} dt + \begin{pmatrix} \sigma & 0 \\ a\rho^\top & a\sqrt{1-\rho^\top\rho} \end{pmatrix} dB_t; \quad \begin{pmatrix} R_0 \\ Y_0 \end{pmatrix} = \begin{pmatrix} 0 \\ y \end{pmatrix} \in (\mathbb{R}^n \times E) \tag{5.3.7}$$

where (R, Y) is defined as the projection of G onto the first n and final entries respectively, this follows from Karatzas and Shreve [70, Proposition 5.4.11]. We verify this with a direct calculation:

$$\begin{pmatrix} \sigma & 0 \\ a\rho^\top & a\sqrt{1-\rho^\top\rho} \end{pmatrix} \begin{pmatrix} \sigma^\top & a\rho \\ 0 & a\sqrt{1-\rho^\top\rho} \end{pmatrix} = \begin{pmatrix} \sigma\sigma^\top & \sigma\rho a \\ (\sigma\rho a)^\top & a^2 \end{pmatrix} = \begin{pmatrix} \Sigma & \Upsilon \\ \Upsilon^\top & a^2 \end{pmatrix}.$$

Next let $\tilde{B} = (B_1, \ldots, B_n)^\top$ and define $(Z, W) = (\tilde{B}^\top, \tilde{B}^\top\rho + \sqrt{1-\rho^\top\rho} B_{n+1})$. Hence the process (R, Y) constructed above obeys the dynamics (5.3.3), (5.3.4) and (5.3.5). Thus to finish the market construction, we define the risky assets as $S_t^i := S_0^i \exp\left(\int_0^t r(Y_s) ds\right) \mathcal{E}(R^i)_t$ for some $S_0^i > 0$, where $\mathcal{E}(X)$ denotes the stochastic exponential of the semi-martingale X, see Protter [100, Section II.8].

We shall work in a complete market so we will enforce the condition $\rho^\top\rho = 1$. With this condition satisfied we see the Brownian motion W driving the state variable is simply a linear combination of the Brownian motions driving the excess return process. We pair this parametric stipulation with the following standing assumption which rigorously ensures the market is arbitrage free and complete in the sense of Delbaen and Schachermayer [42] and Harrison and Pliska [59, 60] respectively.

Assumption 5.3.2. (Market is arbitrage free and complete)
There exists a unique probability measure $\mathbb{Q}^y$ such that $\mathbb{Q}^y|_{\mathcal{F}_t}$ and $\tilde{\mathbb{P}}^y|_{\mathcal{F}_t}$ are equivalent for every $t \in \mathbb{R}_+$ and S/S^0 is a $\mathbb{Q}^y$-local martingale.

Assumptions 5.3.1 and 5.3.2 apply throughout the rest of this chapter without further reference.

5.3.2 Optimisation Problem

An investor with initial wealth x, who chooses investment and consumption policies represented by proportions of their wealth X. Denoting π_t the proportion of wealth invested in the risky assets and l_t the consumption to wealth ratio (i.e if c_t is the agents consumption process then $c_t = X_t l_t$), the agent's wealth process then satisfies,

$$\frac{dX_t^{\pi,l}}{X_t^{\pi,l}} = r\,dt + \pi_t^\top\,dR_t - l_t\,dt. \tag{5.3.8}$$

The goal is to maximise expected utility from consumption over an infinite horizon wherein future utility is discounted according to the time-preference parameter β.

$$\max_{(\pi,l)\in\mathcal{A}} \mathbb{E}\left[\int_0^\infty e^{-\beta t}\frac{(X_t\,l_t)^{1-\gamma}}{1-\gamma}dt\right]. \tag{5.3.9}$$

Here $\gamma \in \mathbb{R}^+ \setminus \{1\}$ denotes the agents risk aversion and $\mathcal{A}$ the set of admissible polices. Such a set consists of all pairs of $\mathcal{F}_t$-adapted processes (π, l), where π is integrable with respect to R and $l_t, X_t^{\pi,l} \geq 0$ $a.s.$ for all $t \geq 0$. First we discuss the parameter regime in which our analysis takes place, this requires two assumptions on the model coefficients.

Assumption 5.3.3. $\inf_{y\in E}\kappa(y) > 0$, *where*

$$E \ni y \mapsto \kappa(y) = \frac{\beta}{\gamma} + \left(1-\frac{1}{\gamma}\right)\left(r(y) + \frac{\mu(y)^\top\Sigma^{-1}(y)\mu(y)}{2\gamma}\right) \tag{5.3.10}$$

Assumption 5.3.4. *The mapping*

$$\eta\colon E \to \mathbb{R}_+,\, y \mapsto \frac{1}{a^2(y)}\exp\left(2\int_{x_0}^y \frac{\tilde{b}(u)}{a^2(u)}du\right), \tag{5.3.11}$$

is a normalised probability density on E where

$$\tilde{b} : E \to \mathbb{R},\, y \mapsto b(y) - \left(1-\frac{1}{\gamma}\right)\left(\Upsilon(y)^\top\Sigma^{-1}(y)\mu(y)\right). \tag{5.3.12}$$

In the special case of constant investment opportunities (the Merton model), the optimal consumption-wealth ratio is $\kappa = \frac{\beta}{\gamma} + \left(1-\frac{1}{\gamma}\right)\left(r + \frac{\mu^\top\Sigma^{-1}\mu}{2\gamma}\right)$ [90, 91]. In order for the Merton problem to be well posed it is necessary for $\kappa > 0$. A recent paper by Herdegen, Hobson and Jerome [63] provides an in depth discussion of the well posedness of the Merton problem and solution techniques in various parameter regimes. In this thesis the coefficients depend on the state variable and so a priori, the function κ may take on negative values (depending on the particular parametric form of a given model). Despite this observation we impose the restriction that the range of the mapping (5.3.10) be some strict subset of $\mathbb{R}_+$, which is easy to check from the model coefficients and is guaranteed to hold in the empirically relevant case of $\gamma > 1$ along with positive interest rates. Further weakening Assumption 5.3.3 may be possible with specific parametric restrictions on a model by model basis.

The function (5.3.11) takes the form of the stationary density of a scalar diffusion with drift and diffusion coefficients $\tilde{b}$ and a respectively. Thus, Assumption 5.3.4 holds for a large class of models that extend beyond that of common state variables with stationary distributions possessing exponentially decaying tails.

We conjecture the value function,

$$V(x,y) = \sup_{(\pi,l)\in\mathcal{A}} \mathbb{E}\left[\int_0^\infty e^{-\beta t} U(X_t l_t)\, dt \,\middle|\, X_0 = x, Y_0 = y\right], \tag{5.3.13}$$

is homogeneous with respect to wealth, i.e of the form $V(x,y) = \frac{x^{1-\gamma}}{1-\gamma} g(y)^\gamma$, where $(x,y) \in \mathbb{R}_+ \times E$. The HJB equation has the general form

$$\beta V(x,y) - \sup_{\pi,l}\left\{\mathcal{L}^{\pi,l} V(x,y) + \frac{(x\,l)^{1-\gamma}}{1-\gamma}\right\} = 0 \tag{5.3.14}$$

where the differential operator is defined in [98, Section 3.4.1]. Plugging the Ansatz $V(x,y) = \frac{x^{1-\gamma}}{1-\gamma} g(y)^\gamma$ into (5.3.14), the HJB equation is then

$$\begin{aligned} r + \frac{\beta}{1-\gamma} =& \frac{\gamma b\, g'}{(\gamma-1)g} + \frac{\gamma a^2\, g''}{2(\gamma-1)g} + \frac{\gamma a^2 (g')^2}{2g^2} \\ &- \sup_{\pi,l}\left(\pi^\top \mu - \frac{\gamma}{2}\pi^\top \Sigma \pi + \gamma \pi^\top \Upsilon \frac{g'}{g} + \frac{g^{-\gamma} l^{1-\gamma}}{1-\gamma} - l\right) \end{aligned} \tag{5.3.15}$$

and the candidate optimal controls are

$$\hat{\pi}(y) = \frac{1}{\gamma}\Sigma^{-1}(y)\mu(y) + \Sigma^{-1}(y)\Upsilon(y)\frac{g'(y)}{g(y)}, \qquad \hat{l}(y) = g(y)^{-1}, \tag{5.3.16}$$

which, are substituted into (5.3.15), yielding:

$$\frac{a^2}{2} g'' + \left(b + \left(1 - \frac{1}{\gamma}\right)\left(\Upsilon^\top \Sigma^{-1}\mu\right)\right) g' - \left(\frac{\beta}{\gamma} + \left(1 - \frac{1}{\gamma}\right)\left(r + \frac{\mu^\top \Sigma^{-1}\mu}{2\gamma}\right)\right) g = -1.$$

We define the linear differential operator

$$\mathcal{L}g = \frac{a^2}{2} g'' + \left(b + \left(1 - \frac{1}{\gamma}\right)\left(\Upsilon^\top \Sigma^{-1}\mu\right)\right) g' - \left(\frac{\beta}{\gamma} + \left(1 - \frac{1}{\gamma}\right)\left(r + \frac{\mu^\top \Sigma^{-1}\mu}{2\gamma}\right)\right) g \tag{5.3.17}$$

or (recall (5.3.10) and (5.3.12)) simply $\mathcal{L}g = \frac{a^2}{2} g'' + \tilde{b}\, g' - \kappa\, g$. Thus the HJB equation admits the compact representation

$$\mathcal{L}g = -1. \tag{5.3.18}$$

In some cases closed form solutions to (5.3.18) have been found. Wachter [121] gave an explicit solution when the excess return followed an OU process which was then generalised for so called *quadratic* models by Liu [78]. Guasoni and Wang provide explicit solutions for two models, one with a stochastic Sharpe ratio which is proportional to the square root of the (strictly positive) state variable [56] and another with Vasicek interest rates [55]. These however are very limited cases which focus on a relatively tractable class of models. In general, equation (5.3.18) does not admit an explicit solution.

One can prove the existence of a $C^2(E;\mathbb{R})$ solution to (5.3.18) via sub and super solutions, this is done, even in the case of incomplete markets, by Guasoni and Wang [56, Section 3]. In [56] the authors provide a verification argument but this crucially relies on a strong assumption about the the solution to the HJB equation which is difficult to check, namely it requires knowledge about the pointwise behaviour of the quantity g'/g . The sub and super solution approach provides pointwise bounds on g but yields no information about the derivative, hence the verification argument in [56] is not applicable when explicit solutions are not available.

Known numerical methods cannot be readily employed either as (5.3.18) does not imply any particular boundary conditions. Even if initial conditions for g and g' were available and one obtained a numerical solution this does not yield verification. Thus standard numerical methods alone, even if they were applicable, are not suitable.

The variational approach provided in this thesis circumvents these problems by providing an alternative method to analytically study the solution of the HJB equation while simultaneously yielding verification and a simple numerical method which does not require boundary conditions.

5.3.3 Variational Problem

Next we introduce the variational problem by defining the functional,

$$I : X \to \mathbb{R},\, u \mapsto I(u) \coloneqq \int_E f(y, u(y), u'(y))dy, \tag{5.3.19}$$

for $u \in X$ where X is some space of real valued functions (to be specified) and the mapping f is defined as

$$f : E \times \mathbb{R} \times \mathbb{R} \to \mathbb{R},\, (y, u, \xi) \mapsto \left(\frac{a^2(y)\xi^2}{2} + \kappa(y)u^2 - 2u\right)\eta(y). \tag{5.3.20}$$

The minimisation problem is then,

$$\min\{I(u) : u \in X\}. \tag{5.3.21}$$

If there exists a $C^2(E;\mathbb{R})$ solution to (5.3.21) then it will satisfy the classical Euler-Lagrange equations which in this case takes the form

$$\frac{d}{dy}\left[\frac{\partial}{\partial u'}f(y, u, u')\right] = \frac{\partial}{\partial u}f(y, u, u'). \tag{5.3.22}$$

Plugging the definition of f one obtains,

$$\frac{a^2}{2}u'' + \left(aa' + \frac{a^2\eta'}{2\eta}\right)u' - \kappa u = -1. \tag{5.3.23}$$

Upon substituting the definition of η (5.3.11), equation (5.3.23) reduces to

$$\frac{a^2}{2}u'' + \tilde{b}\,u' - \kappa\,u = -1, \tag{5.3.24}$$

which is exactly $\mathcal{L}u = -1$ i.e, the Hamilton-Jacobi-Bellman equation. Hence the following proposition has been proven.

Proposition 5.3.1. *The classical form Euler-Lagrange equation for the variational problem* (5.3.21) *is the Hamilton-Jacobi-Bellman equation* (5.3.18)*.*

To understand the intuition which relates this deterministic variational problem and the optimal consumption, rewrite the functional I as,

$$I(u) = \int_E \left(\frac{a^2(y)(u'(y))^2}{2} + \kappa(y)\left(u(y) - \frac{1}{\kappa(y)}\right)^2 - \frac{1}{\kappa(y)}\right)\eta(y)dy. \tag{5.3.25}$$

Recall that κ is the myopic consumption-wealth ratio and our candidate optimal policy is $\hat{l}(y) = 1/g(y)$ where g solves the HJB equation. Hence the variational formulation shows that the agent wants to stay close to the myopic consumption ratio (in terms of squared error), but penalises changes in the consumption policy by the presence of the squared derivative term. The intuition is now clear, the agent wishes to stay as close to the myopic consumption ratio as possible while avoiding drastic changes in the consumption rate. If κ is constant then (5.3.21) can by solved by inspecting (5.3.25), of course the solution would be $u(y) = 1/\kappa$ which is exactly the Merton consumption. Hence the variational approach encapsulates the classical case. Additionally the need for Assumption 5.3.3 is now clear, if κ were negative this intuition breaks down and it may well be the case that one could construct a sequence of functions u_n such that $I(u_n) \to -\infty$ as $n \to \infty$.

Before presenting the main theorem it is still necessary to specify the function space X in which the variational problem is solved. The natural choice is the weighted Sobolev space

$$W^{1,2}_{\eta,a^2}(E) := \left\{u \in L^1_{loc}(E;\mathbb{R}) : \int_E |u(y)|^2\eta(y)dy + \int_E |\xi(y)|^2 a^2(y)\eta(y)dy < +\infty\right\}. \tag{5.3.26}$$

In (5.3.26) ξ is understood to be the weak derivative of u, i.e $\xi \in L^1_{loc}(E;\mathbb{R})$ and

$$\int_E u(y)\varphi'(y)dy = -\int_E \xi(y)\varphi(y)dy, \tag{5.3.27}$$

for all $\varphi \in C_c^\infty(E, \mathbb{R})$. For ease of notation,

$$H(E) := W^{1,2}_{\eta,a^2}(E).$$

5.4 Main Results

We are now in a position to state our main results.

Theorem 5.4.1. *Let Assumptions 5.3.1-5.3.4 hold. Then,*

(i) There exists a unique solution $g \in C^2(E; \mathbb{R})$ to the minimisation problem (5.3.21).

(ii) The unique solution g to the minimisation problem (5.3.21) solves the HJB equation.

(iii) If there exists a unique solution $\hat{\mathbb{P}}$ to the martingale problem on $\mathbb{R}^n \times E$ for,

$$\hat{L} = \frac{1}{2}\sum_{i,j=1}^{n+1} \hat{A}_{i,j}\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{n+1} \hat{b}_i \frac{\partial}{\partial x_i}$$
$$\hat{A} = \begin{pmatrix} \Sigma & \Upsilon \\ \Upsilon^\top & a^2 \end{pmatrix}, \quad \hat{b} = \begin{pmatrix} \frac{\mu}{\gamma} + \Upsilon \frac{g'}{g} \\ \tilde{b} + a^2 \frac{g'}{g} \end{pmatrix}.$$

Then, $V(x,y) = \frac{x^{1-\gamma}}{1-\gamma} g(y)^\gamma$ and the optimal policies are,

$$\pi^*(y) = \frac{1}{\gamma}\Sigma^{-1}(y)\mu(y) + \Sigma^{-1}(y)\Upsilon(y)\frac{g'(y)}{g(y)}, \qquad l^*(y) = \frac{1}{g(y)}.$$

The utility of Theorem 5.4.1 obviously hinges on the ability to check its assumptions in concrete models. The main issue is that the solution of the variational problem g is present in the drift coefficient of the martingale problem, but this function is not explicitly known. Our next result shows that it is often the case that explicit pointwise estimates on g'/g are not required.

Assumption 5.4.1. *The scale function p for the scalar diffusion*

$$dZ_t = \tilde{b}(Z_t)dt + a(Z_t)dW_t \tag{5.4.1}$$

obeys,

$$\lim_{y \to \partial E_+} p(y) = +\infty; \qquad \lim_{y \to \partial E_-} p(y) = -\infty$$

where $\partial E_+, \partial E_-$ are the right and left boundaries of E respectively.

Recall for an arbitrary $c \in E$,

$$p(y) := \int_c^y \exp\left(-2\int_c^u \frac{\tilde{b}(s)}{a^2(s)}ds\right)du; \quad y \in E. \tag{5.4.2}$$

Assumption 5.4.1 ensures the diffusion defined by (5.4.1) does not exit E in finite time. This will be true in virtually all models as (5.4.1) is the state variable with an adjusted drift (for which non explosion is always assumed).

Theorem 5.4.2. *Let Assumptions 5.3.1-5.3.4 and 5.4.1 hold. Let g be the unique solution to the minimisation problem (5.3.21). Then the value function is $V(x,y) = \frac{x^{1-\gamma}}{1-\gamma}g(y)^\gamma$ and the optimal policies are*

$$\pi^*(y) = \frac{1}{\gamma}\Sigma^{-1}(y)\mu(y) + \Sigma^{-1}(y)\Upsilon(y)\frac{g'(y)}{g(y)}; \qquad l^*(y) = \frac{1}{g(y)}.$$

In a large class of models, Theorem 5.4.2 reduces the entire optimal consumption problem to checking some elementary conditions on the coefficients of the market model. Before we proceed with further model free analysis we provide some examples of models for which Theorem 5.4.2 applies.

5.4.1 Applications of Theorem 5.4.2

In this subsection we provide examples of three models solved using Theorems 5.4.1 and 5.4.2, wherein an explicit solution to the HJB equation is unavailable. We state each model, provide discussion regarding the models characteristics and implications and provide numerical solutions for optimal policies with realistic market coefficients.

At the end of this subsection, we state a proposition that shows each model fits into our newly developed framework. Throughout this section, Y_∞ denotes the weak limit of the state variable and hence will be distributed according to the stationary distribution of Y.

Non-linear mean-reverting excess returns:

Example 5.4.1.

$$\begin{aligned} r(Y_t) = r; \quad dR_t &= \mu Y_t\,dt + \sigma\,dZ_t, \\ dY_t &= b(\theta - Y_t)^3\,dt + a\,dW_t, \end{aligned} \tag{5.4.3}$$

where $r, \mu, \sigma, b, \theta, a > 0$ are constants and $\rho^2 = 1$. In this regime,

$$\tilde{b}(y) = b(\theta - y)^3 - \lambda y \qquad \kappa(y) = \frac{\beta}{\gamma} + \left(1 - \frac{1}{\gamma}\right)\left(r + y^2\frac{\mu^2}{2\gamma\sigma^2}\right)$$

where $\lambda := \left(1 - \frac{1}{\gamma}\right) \frac{a\rho\mu}{\sigma}$ and

$$\eta(y) = \frac{C}{a^2} \exp\left(-\frac{b(\theta - y)^4}{2a^2} - \frac{\lambda y^2}{a^2}\right).$$

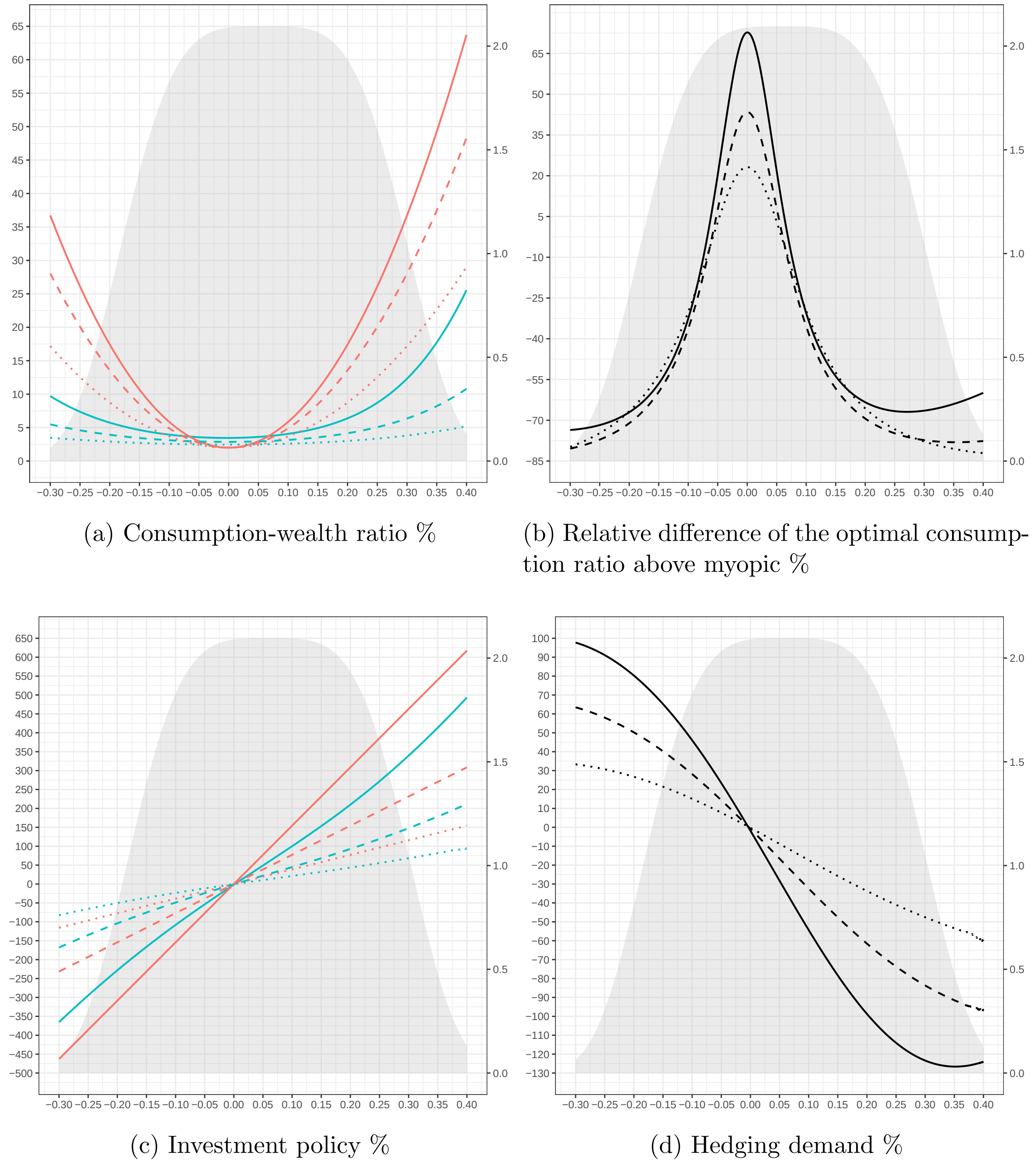
In this model of stochastic excess returns the state variable is a mean reverting OU type process. The cubic non-linearity has two affects: (i) it strongly penalises large deviations from the long term mean; and (ii) significantly reduces the strength of mean reversion when the state variable is close to the long term mean.

The Sharpe ratio is a linear function of the state variable, which means that extreme deviations in both the positive and negative direction result in better investment opportunities for the agent because short selling is permitted. The intuition, confirmed in Figure 5.1, is that consumption is higher when the state variable is far away from its long term mean.

Below we numerically solve the variational problem associated with model (5.4.3) and report both the optimal consumption-wealth ratio and investment policy and compare the differences in optimal polices from the myopic agent. Details of the numerical scheme are provided in Appendix A.1

Parameter	r	μ	σ	b	θ	a	β	ρ
Value	0.02	1	0.18	0.4	0.06	0.031	0.02	1

Table 5.1: Parameter values for model (5.4.3), where μ and σ are chosen so that the expected excess return (with respect to the state variables stationary distribution) is 0.06.



(a) Consumption-wealth ratio %

(b) Relative difference of the optimal consumption ratio above myopic %

(c) Investment policy %

(d) Hedging demand %

Figure 5.1: For every plot the horizontal axis is the state variable Y (with $\mathbb{P}(-0.3 \leq Y_\infty \leq 0.4) \approx 0.995$), the right vertical axis is the stationary density of Y, and the shaded area is the region under the graph of the stationary density of Y. For each plot three types of agents are considered, $\gamma = 2$ (solid line), $\gamma = 4$ (dashed line), and $\gamma = 8$ (dotted line) respectively. Consumption-wealth ratio l (top left panel, left vertical axis, in percent) and investment ratio π (bottom left panel, left vertical axis, in percent) of the Myopic (red) vs Optimal (blue) policies. Relative difference of the optimal consumption ratio above myopic $l^*/\kappa - 1$ (top right panel, left vertical axis, in percent) and the hedging demand $\Sigma^{-1}\Upsilon\frac{g'}{g}$ (bottom right panel, left vertical axis, in percent). Market parameters are given in Table 5.1

As expected, both the consumption and investment policy are far more conservative than their myopic counterparts. The relative difference between the consumption policies ranges between −85% and 70% although, we see the difference begins to decrease when the returns are significantly positive for smaller values of risk aversion. For large positive returns the hedging demand stabilises and then begins to increases for the less risk averse agent similar to the consumption policy.

An interesting feature of the optimal consumption policy is the lack of symmetry which, is more prominent for the myopic policy. The intuition is clear: when the returns process becomes highly negative the price plummets, and the potential to make profits from short selling diminishes unless large amounts of leverage is acquired. The result is more conservative consumption in comparison to similar increases in returns.

In Figure 5.1 we see the amount of leverage the agent takes on is slightly higher in bad times than in good times but they do not significantly differ. For $\gamma = 2$, when the excess returns are 30% the agent invests approximately 3.4 times her net worth in the risky asset. In contrast, when excess returns are −30%, the agent takes a short position of approximately 3.6 times her net worth. This asymmetry indicates a reluctance to significantly increase the amount of leverage in bad times, which would otherwise be necessary to finance higher consumption.

Power law Sharpe Ratio:

Example 5.4.2.

$$\begin{aligned} r(Y_t) = r; \quad dR_t &= \mu Y_t^{\delta}\, dt + \sigma\sqrt{Y_t}\, dZ_t, \\ dY_t &= b(\theta - Y_t)\, dt + a\sqrt{Y_t}\, dW_t, \end{aligned} \tag{5.4.4}$$

where $r, \mu, \sigma, b, \theta, a > 0$ are constants, $\rho^2 = 1$ and $\delta \geq \frac{1}{2}$. In this regime,

$$\tilde{b}(y) = b(\theta - y) - \lambda y^{\delta} \qquad\qquad \kappa(y) = \frac{\beta}{\gamma} + \left(1 - \frac{1}{\gamma}\right)\left(r + y^{2\delta-1}\frac{\mu^2}{2\gamma\sigma^2}\right)$$

where $\lambda := \left(1 - \frac{1}{\gamma}\right)\frac{a\rho\mu}{\sigma}$. We impose the restriction, $\frac{1}{2} < \frac{b\theta}{a^2} < 1$. The lower bound ensures $Y_t > 0$ *a.s.*, the upper bound is a technical condition needed for verification. η takes the form,

$$\eta(y) = \frac{C}{a^2} y^{\frac{2b\theta}{a^2}-1} \exp\left(-\frac{2by}{a^2} - \frac{2\lambda y^{\delta}}{a^2\delta}\right),$$

where $C > 0$ is a normalising constant. To ensure η is a well defined density the following parameter restrictions must hold for different values of δ,

$$\begin{cases} \lambda \in \mathbb{R}, & \delta \in [\frac{1}{2}, 1), \\ b + \lambda > 0, & \delta = 1, \\ \lambda > 0, & \delta \in (1, \infty). \end{cases}$$

In this model for stochastic risk premia, the parameter δ determines the functional form of the Share ratio. For $\delta \in [\frac{1}{2}, \frac{3}{2})$ the Sharpe ratio is a concave function of the state variable, for $\delta > \frac{3}{2}$ it is convex and for $\delta = \frac{3}{2}$ it is linear. For all values of δ a higher Y implies better investment opportunities. The special case of $\delta = 1$ was solved explicitly by Guasoni and Wang [56].

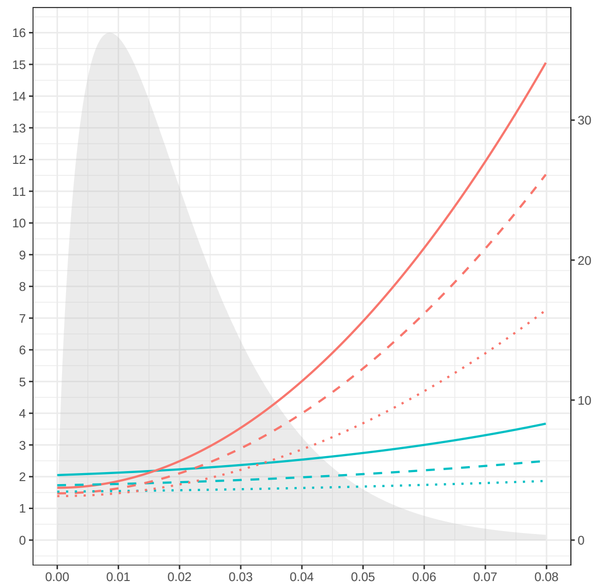
There is a trade-off between the value of δ and the correlation between Y and the asset returns. To ensure that η is a well defined density and that our previous theorems apply, a model with negative correlation must be combined with a concave Sharpe ratio. If the correlation is positive then our framework is compatible with either concavity or convexity. Below we solve this model for the critical case of $\delta = \frac{3}{2}$ which results in a linear Sharpe ratio in the state variable.

Upon inspecting figure 5.2, we see consumption is extremely stable across all states of nature and drastically differs from the myopic policy. For the least risk averse agent ($\gamma = 2$) the consumption policy increases from approximately 2% when the Sharpe ratio is zero to approximately 3.6% when the Sharpe ratio is 1.036. In contrast, the myopic agent consumes at a rate of 15% when the Sharpe ratio reaches 1.036. For higher values of risk aversion ($\gamma = 8$) the optimal consumption policy is essentially linear and increases from its low point of 1.5% to only 2% at its peak. The optimal investment policy is relatively less stable and increases as a concave function of the state variable. The agent consumes more as Y_t increases suggesting the income effect is present, although the actual rate of increase in consumption becomes almost negligible for low values of risk aversion. This finding is similar to the model of Guasoni and Wang [56] wherein $\delta = 1$ and the income effect dominates.

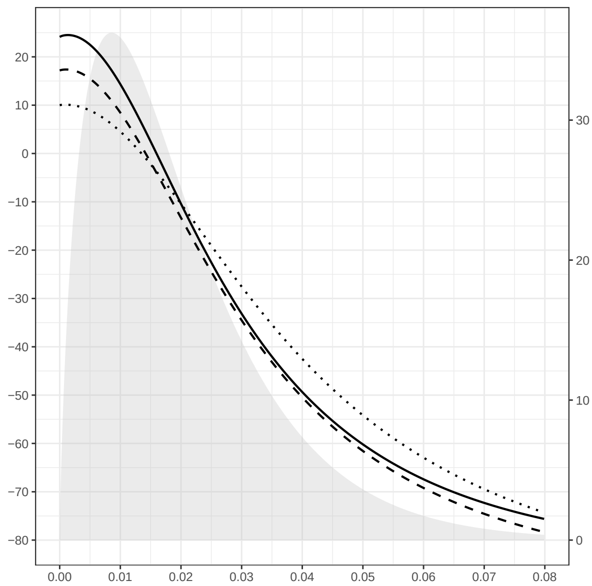
The most interesting aspect of this model is that the agent short sells the risky asset despite a small but strictly positive risk premium. This phenomenon was also observed by Liu [78] in a bond allocation problem and a portfolio choice problem without consumption wherein the state variable is a CIR process. Liu explained this phenomena as a possible feature of power utility with risk aversion less than one, however in our model shorting risky assets with strictly positive risk premium occurs for each value of risk aversion under consideration ($\gamma = 2, 4, 8$). In our model this occurs as the correlation between the state variable and the risky asset is one, meaning the agent takes a short position in the risky asset to hedge against negative movements in the state variable. Hence as the Sharpe ratio approaches zero there is no myopic demand and the agent takes a net short position.

Parameter	r	μ	σ	b	θ	a	β	δ	ρ
Value	0.013	17.705	1.366	0.07	0.02	0.04	0.02	1.5	1

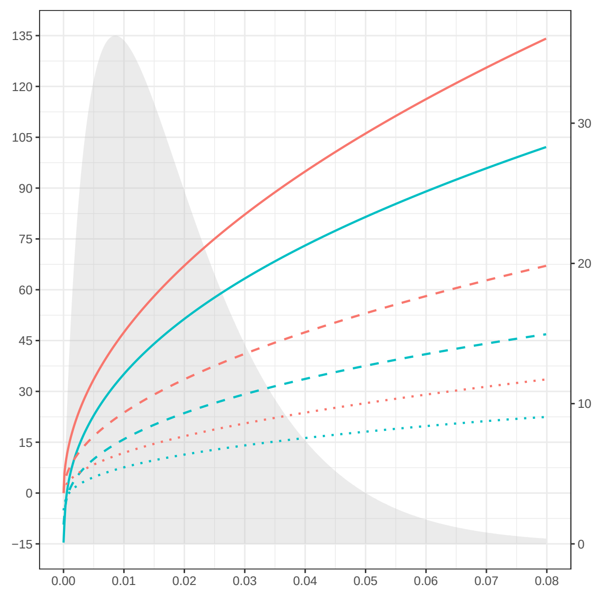
Table 5.2: Parameters for model (5.4.4), μ and σ are chosen so that the expected excess return and volatility (with respect to the state variables stationary distribution) are 0.06 and 0.18.



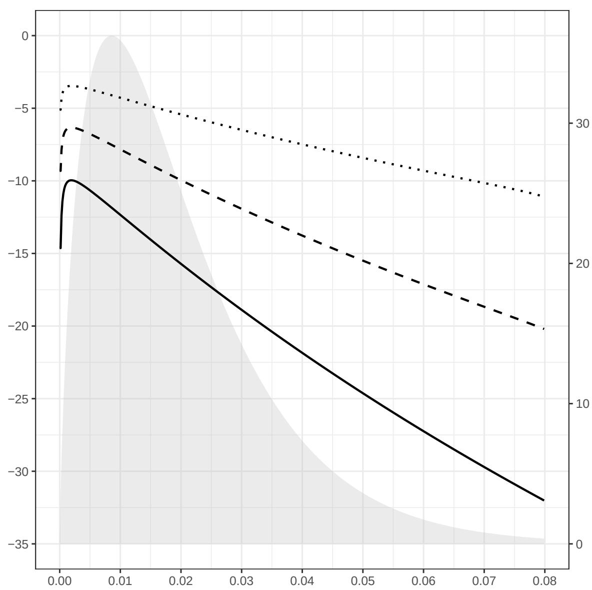
(a) Consumption-wealth ratio %



(b) Relative difference of the optimal consumption ratio above myopic %



(c) Investment policy %



(d) Hedging demand %

Figure 5.2: For every plot the horizontal axis is the state variable Y (with $\mathbb{P}(0 \le Y_\infty \le 0.8) \approx 0.995$), the right vertical axis is the stationary density of Y, and the shaded area is the region under the graph of the stationary density of Y. For each plot three types of agents are considered, $\gamma = 2$ (solid line), $\gamma = 4$ (dashed line), and $\gamma = 8$ (dotted line) respectively. Consumption-wealth ratio l (top left panel, left vertical axis, in percent) and investment ratio π (bottom left panel, left vertical axis, in percent) of the Myopic (red) vs Optimal (blue) policies. Relative difference of the optimal consumption ratio above myopic $l^*/\kappa - 1$ (top right panel, left vertical axis, in percent) and the hedging demand $\Sigma^{-1}\Upsilon\frac{g'}{g}$ (bottom right panel, left vertical axis, in percent). Market parameters are given in Table 5.2

Stochastic interest rates:

Example 5.4.3.

$$\begin{aligned} r(Y_t) = Y_t; \quad dR_t &= \mu\, dt + \sigma\, dZ_t, \\ dY_t &= b(\theta - Y_t)\, dt + a\sqrt{Y_t}\, dW_t, \end{aligned} \tag{5.4.5}$$

where $\mu, \sigma, b, \theta, a > 0$ are constants and $\rho^\top \rho = 1$. In this regime,

$$\tilde{b}(y) = b(\theta - y) - \lambda\sqrt{y} \qquad\qquad \kappa(y) = \frac{\beta}{\gamma} + \left(1 - \frac{1}{\gamma}\right)\left(y + \frac{\mu^2}{2\gamma\sigma^2}\right)$$

where $\lambda := \left(1 - \frac{1}{\gamma}\right)\frac{a\rho\, u}{\sigma}$. We impose the restrictions $\frac{1}{2} < \frac{b\theta}{a^2} < 1$. The lower bound ensures $Y_t > 0$ $a.s.$, η is a well defined density which takes the form,

$$\eta(y) = \frac{C}{a^2} y^{\frac{2b\theta}{a^2} - 1} \exp\left(-\frac{2}{a^2}(by + 2\lambda\sqrt{y})\right),$$

where $C > 0$ is a normalising constant. The bound $\frac{b\theta}{a^2} < 1$ is a technical condition needed for verification.

In this model the interest rate is a almost surely strictly positive CIR process and is negatively correlated with the shocks from the risky asset. The numerical solution is given in figure 5.3. The optimal consumption ratio is linear in the state variable and the relative difference of optimal consumption above myopic seems to approach a constant level as the interest rate becomes large. In times of low interest rates the less risk averse agent consumes more than the highly risk averse agent but this is reversed when interest rates increase.

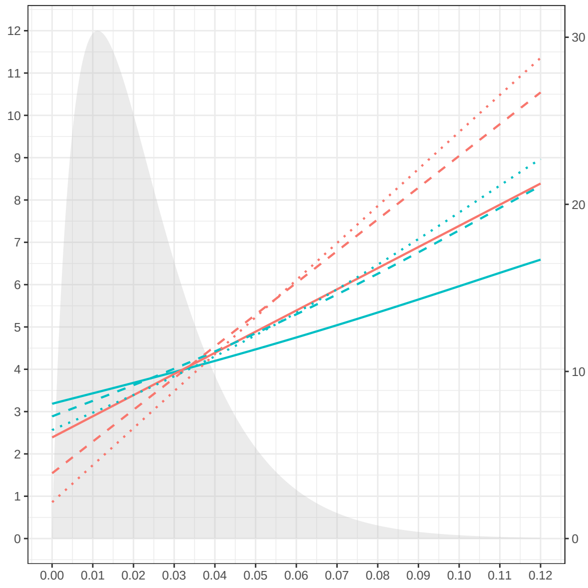
As the state variable is negatively correlated with the risky asset, the hedging demand is strictly positive for interest rates between $0 - 12\%$ offsetting negative shocks to the interest rate. The optimal investment policy peaks at a certain point for all values of risk aversion and subsequently begins to decrease as the interest rate becomes increasingly large. This effect is more pronounced when looking directly at hedging demand, as myopic demand is constant. The hedging component in this model is

$$\frac{a\rho g'(y)\sqrt{y}}{\sigma g(y)}.$$

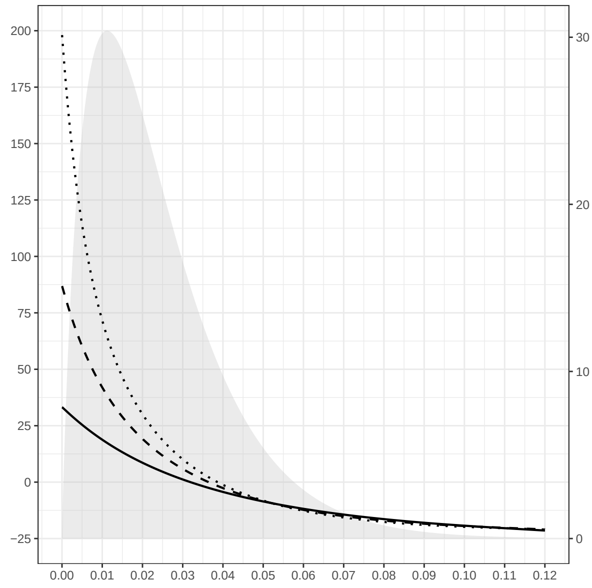
We show in the proof of Proposition 5.4.1 that $\frac{g'(y)}{g(y)} \leq 0$ for $y > \theta$, hence as $\rho = -1$ the hedging component is positive whenever the interest rate is above its long term mean. This indicates the agent will never short the risky asset in order to buy more of the risk free. In fact she always invests a larger proportion of her wealth in the risky asset than the myopic agent.

Parameter	r	μ	σ	b	θ	a	β	ρ
Value	1	0.06	0.18	0.04	0.024	0.032	0.02	-1

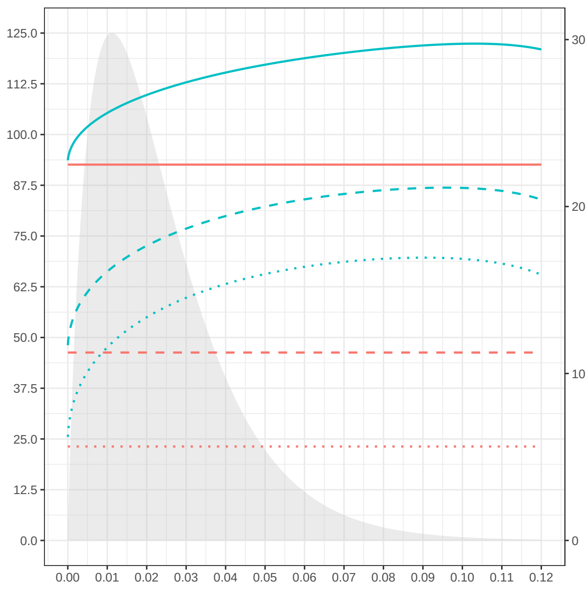
Table 5.3: Parameters for model (5.4.5), r is chosen so that the expected interest rate (with respect to the state variables stationary distribution) is equal to θ.



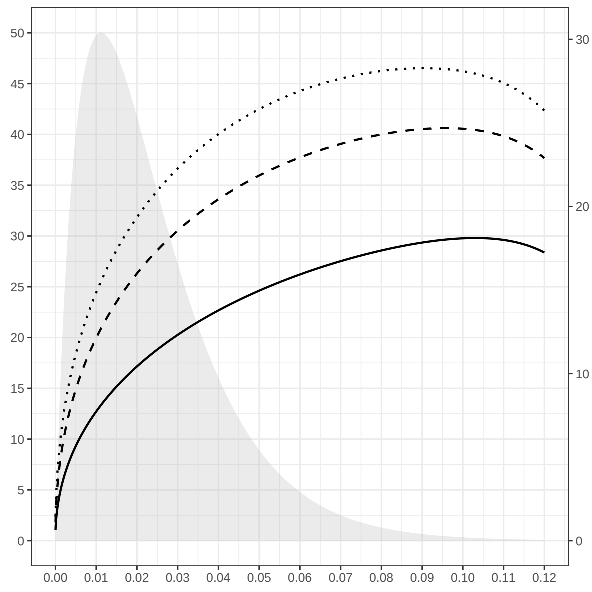
(a) Consumption-wealth ratio %



(b) Relative difference of the optimal consumption ratio above myopic %



(c) Investment policy %



(d) Hedging demand %

Figure 5.3: For every plot the horizontal axis is the state variable Y (with $\mathbb{P}(0 \leq Y_\infty \leq 12) \approx 0.999$), the right vertical axis is the stationary density of Y, and the shaded area is the region under the graph of the stationary density of Y. For each plot three types of agents are considered, $\gamma = 2$ (solid line), $\gamma = 4$ (dashed line), and $\gamma = 8$ (dotted line) respectively. Consumption-wealth ratio l (top left panel, left vertical axis, in percent) and investment ratio π (bottom left panel, left vertical axis, in percent) of the Myopic (red) vs Optimal (blue) policies. Relative difference of the optimal consumption ratio above myopic $l^*/\kappa - 1$ (top right panel, left vertical axis, in percent) and the hedging demand $\Sigma^{-1}\Upsilon\frac{g'}{g}$ (bottom right panel, left vertical axis, in percent). Market parameters are given in Table 5.3

Proposition 5.4.1. *Let $g \in C^2(E;\mathbb{R})$ be the unique solution to the corresponding variational problem for the models (5.4.3)-(5.4.5) respectively. Then the value function is $V(x,y) = \frac{x^{1-\gamma}}{1-\gamma} g(y)^\gamma$ and the optimal controls are,*

$$\pi^*(y) = \frac{1}{\gamma}\Sigma^{-1}(y)\mu(y) + \Sigma^{-1}(y)\Upsilon(y)\frac{g'(y)}{g(y)}; \qquad\qquad l^*(y) = \frac{1}{g(y)}.$$

5.5 Intertemporal Hedging Asymptotics

In the previous section we provided a general verification argument which characterised the value function as the unique solution of a variational problem and gave numerical results for realistic values of the state variable. These results although useful do not provide any information regarding the asymptotic behaviour of the value function (and hence the consumption policy) or its derivative (and hence the intertemporal hedging component). Additionally if the assumptions of Theorem 5.4.2 are not satisfied then one may need to work harder for a verification result.

We address these issues in the following section by obtaining precise pointwise bounds on g'/g, which simultaneously yields asymptotics of the consumption and the intertemporal hedging component, while also providing an alternative verification argument. This is done in three main steps, (i) first the Euler Lagrange equation yields a representation for g' in terms of g. (ii) the maximum principle from the theory of linear elliptic PDEs yields sharp estimates on g. (iii) combining the estimate from step (ii) and the representation of g' from step (i) yields a pointwise bound on g'/g.

The price to be paid in order to obtain precise pointwise estimates on g is of course imposing specific (and more restrictive) pointwise estimates on the model coefficients, something which up until now has not been needed. Throughout this section $g \in C^2(E;\mathbb{R})$ will denote the unique solution to the variational problem (5.3.21). In order to eliminate the proliferation of special cases and to simplify the following exposition, we specialise to the cases where $E = (0,\infty)$ or $E = \mathbb{R}$.

We introduce the notation $U(m) \coloneqq ((-\infty,-m)\cup(m,\infty))\cap E$ where $m > 0$ as we often need to conduct analysis when the state variable takes on extreme values. Additionally $x \lesssim y$ means there exists a constant $C > 0$ (independent of relevant parameters) such that $x \le Cy$.

Assumption 5.5.1. *$\kappa\eta$ is uniformly bounded on $U(m)$ for some $m > 0$ and $a^2\eta \in L^1(U(m);\mathbb{R})$.*

Assumption 5.5.2. *There exists constants $C_1, C_2, C_3 > 0$, $p \ge 0$ and $\psi \in W^{2,1}(E;\mathbb{R}_+)$*

such that,

$$C_1(1+|y|^p) \leq \kappa(y) \leq C_2(1+|y|^p), \quad y \in E, \tag{5.5.1}$$

$$\int_y^{2y} \left(\left|(a^2\psi)''\right| + \left|(\tilde{b}\psi)'\right| \right) du \leq C_3(1+|y|^p), \tag{5.5.2}$$

where $y > 0$ is large enough, ψ is supported on $[y, 2y]$ and $\int_y^{2y} \psi du = \Phi(y)$. Additionally $\Phi(y) \leq y$ for all $y \in E$ and for sufficiently large y, $y \lesssim \Phi(y)$.

Assumption 5.5.3. *Assume there exists $h \in C^2(E;\mathbb{R})$, constants $C_1, C_2, C_3, C_4 > 0$ and $q \geq p \geq 0$ (with p from Assumption 5.5.2) such that for all $|y|$ large enough*

$$-C_1 \leq \mathcal{L}h \leq -C_2 \quad \textit{and} \quad \frac{C_3}{(1+|y|^q)} \leq h(y) \leq \frac{C_4}{(1+|y|^q)}.$$

Assumption 5.5.4. *For large enough $x > 0$, $\inf_{|y|>x} |\tilde{b}(y)| > 0$.*

Assumption 5.5.1 yields the following Lemma:

Lemma 5.5.1. *Let Assumption 5.5.1 hold. Then $a^2(y)\eta(y)g'(y) \to 0$ as $|y| \to \infty$.*

To see why this is useful we integrate the Euler-Lagrange equation (5.3.23) over an interval $[x, y] \subset E$,

$$a^2(y)\eta(y)g'(y) - a^2(x)\eta(x)g'(x) = 2\int_x^y (\kappa(u)g(u) - 1)\,\eta(u)du.$$

As $y \to \infty$, Lemma 5.5.1 yields $a^2(y)\eta(y)g'(y) \to 0$, and rearranging

$$g'(x) = \frac{-2}{a^2(x)\eta(x)} \int_x^\infty (\kappa(u)g(u) - 1)\,\eta(u)du, \tag{5.5.3}$$

which is well defined, as both a^2 and η are strictly greater than zero on the interior of E. This formulation reduces the problem to estimating g, then estimates on the coefficients from the martingale problem in Theorem 5.4.1 are easily obtained. In certain situations it is necessary to be very precise when estimating, for example if the state variable is a square root diffusion then the HJB equation is singular and thus $g'(y)$ may explode as $y \to 0$, in such models the parameter regime will play an important role.

Assumption 5.5.2 yields information about g at the boundary of E. Essentially, it ensures that the candidate optimal consumption policy should, in a certain sense, *agree* with the myopic consumption policy as the state variable approaches the boundary. To be precise, as $y \to \partial E$, if $\kappa(y) \to \infty$ then $g(y)^{-1} \to \infty$, similarly if $\kappa(y) \to C > 0$ then we must have $g(y)^{-1} \to C^* > 0$ where the positive constants C and C^* may differ. Assumption 5.5.3 provides precise pointwise bounds on g via the maximum principle.

This can only be achieved if one has knowledge of the behaviour of g at the boundary which is why Assumption 5.5.2 is crucial.

Assumption 5.5.4 ensures that the quotient $1/\tilde{b}$ is well defined for large values of the state variable and holds in virtually all models. We can now present our second main result.

Theorem 5.5.1. *Let Assumptions 5.3.1-5.3.4 and 5.5.1-5.5.4 hold. Then,*

(i) If $E = \mathbb{R}$, then for all $y \in E$,

$$\frac{1}{(1+|y|^q)} \lesssim g(y) \lesssim \frac{1}{(1+|y|^q)}, \tag{5.5.4}$$

and for sufficiently large $m > 0$ and for all $y \in U(m)$,

$$|g'(y)| \lesssim \frac{1}{\inf_{w \in U(|y|)} |\tilde{b}(w)|} \tag{5.5.5}$$

(ii) If $E = \mathbb{R}_+$ and $\limsup_{y \to 0} \exp(2\int_y^{x_0} \frac{\tilde{b}(u)}{a^2(u)} du) < +\infty$ then (5.5.4) and (5.5.5) holds.

The additional condition for $E = \mathbb{R}_+$ ensures that g' does not explode as the state variable approaches zero. For specific models Theorem 5.5.1 provides exact asymptotics for the value function and has the implication that the leading order asymptotic behaviour of the optimal consumption policy is always greater than or equal to that of the myopic investor. However our economic intuition tells us that the leading order asymptotic behaviour of the optimal consumption policy should be *less than or equal to* that of the myopic investor (for $\gamma > 1$ at least). This leads to the conjecture that the optimal and myopic consumption policies are of the same order of magnitude.

In the next section we provide a general class of examples where this conjecture is true, and verification is carried out by invoking the pointwise estimates from Theorem 5.5.1.

5.5.1 Homoscedastic returns with non-linear one-factor predictability

The excess return is stochastic and follows an additive diffusion with constant noise intensity $a > 0$ and a general drift $b : \mathbb{R} \to \mathbb{R}$, i.e

$$\begin{aligned} r(Y_t) &= r; \quad dR_t^i = \mu_i Y_t \, dt + \sum_{j=1}^n \sigma_{ij} \, dZ_t^j, \qquad 1 \le i \le n, \\ dY_t &= b(Y_t) \, dt + a \, dW_t, \end{aligned} \tag{5.5.6}$$

where $r, \mu_i, \sigma_{ij}, a > 0$ are constants. As in the general market we stipulate that Σ is strictly positive definite and enforce the constraint $\rho^\top \rho = 1$ with the additional assumption that ρ is constant, thus the market is complete. In this regime,

$$\tilde{b}(y) = b(y) - \lambda y \qquad\qquad \kappa(y) = \frac{\beta}{\gamma} + \left(1 - \frac{1}{\gamma}\right)\left(r + y^2 \frac{\mu^\top \Sigma^{-1} \mu}{2\gamma}\right)$$

where $\lambda := \left(1 - \frac{1}{\gamma}\right) a \rho^\top \sigma^{-1} \mu$ and $\mu^\top = (\mu_1, \cdots, \mu_n)$. We make the following assumption on b,

Assumption 5.5.5.

$$(i)\; b \in C^{1,\alpha}(\mathbb{R};\mathbb{R}) \cap Lip(\mathbb{R};\mathbb{R}). \tag{5.5.7}$$

$$(ii)\; \eta \textit{ is a well defined probability density and } \kappa\eta \textit{ is uniformly bounded.} \tag{5.5.8}$$

$$(iii)\;\; \textit{For } |y| \textit{ large enough } \inf_{x > |y|} |\tilde{b}(x)| > 0 \textit{ and } \frac{(1 + y^2)}{\inf_{x>|y|} |\tilde{b}(x)|} \lesssim (1 + |y|). \tag{5.5.9}$$

Proposition 5.5.1. *Let Assumption 5.5.5 hold and $\gamma > 1$. Then the value function is $V(x, y) = \frac{x^{1-\gamma}}{1-\gamma} g(y)^\gamma$ (where $g \in C^2(\mathbb{R};\mathbb{R})$ is the unique solution to the minimisation problem (5.3.21)), the optimal controls are*

$$\pi^*(y) = \frac{1}{\gamma}\Sigma^{-1}\mu\, y + \Sigma^{-1}\Upsilon \frac{g'(y)}{g(y)}, \qquad\qquad l^*(y) = g^{-1}(y),$$

the optimal consumption obeys,

$$(1 + y^2) \lesssim g^{-1}(y) \lesssim (1 + y^2) \quad \textit{for all } y \in E,$$

and the intertemporal hedging component satisfies

$$\left|\frac{g'(y)}{g(y)}\right| \lesssim (1 + |y|) \quad \textit{for all } y \in E.$$

5.6 Conclusion

In this chapter we developed a novel approach to study optimal consumption problems with stochastic investment opportunities in complete markets. We characterised the value function as the unique solution to an associated variational problem, illustrated a numerical scheme to allow finite dimensional approximations of optimal policies and concluded with analytical analysis of the asymptotic behaviour of both consumption streams and intertemporal hedging components.

We emphasise that all assumptions made can be checked directly in terms of model coefficients. Our approach subsumes others in the literature while also providing a simple and accurate numerical method. Three examples of non-trivial models were provided along with a thorough analysis of their numerical solutions. Our findings illustrate the variational approach can be used to model non-standard behaviour in relevant economic quantities such as the risk premium and Sharpe ratios.

5.7 Proofs

5.7.1 Existence, Uniqueness and Regularity for the Variational Problem

Lemma 5.7.1. *Let $E \subseteq \mathbb{R}$, $1 < p < \infty$, $q = \frac{p}{p-1}$ and $g \in L^q(E)$. Then*

$$\ell\colon L^p_\eta(E) \to \mathbb{R};\; f \mapsto \int_E f\, g\, \eta^{1/p}, \qquad \textit{belongs to } L^p_\eta(E)^*.$$

Proof of lemma 5.7.1. Throughout the proof we drop the dependence on the domain E and $\langle x^*, x\rangle$ will denote the dual pairing where x^* is a linear functional and x is a vector. The first step is to construct an isometry between between L^p and L^p_η. Define the mapping,

$$T\colon L^p \to L^p_\eta; f \mapsto f\, \eta^{-1/p}. \tag{5.7.1}$$

This mapping is surjective because given any $g \in L^p_\eta$, set $f = g\, \eta^{1/p} \in L^p$. To show that it is norm preserving, note that for an $f \in L^p$,

$$\|T(f)\|^p_{L^p_\eta} = \int |f\, \eta^{-1/p}|^p \eta = \int |f|^p = \|f\|^p_{L^p}.$$

Thus, under T, L^p and L^p_η are isometric, i.e., $L^p \overset{T}{\cong} L^p_\eta$, hence $L^p_\eta \overset{T^{-1}}{\cong} L^p$. Next, denote the adjoint of T by T^*, which yields the map,

$$T^*\colon (L^p_\eta)^* \to (L^p)^*,$$

whence $(T^*)^{-1} = (T^{-1})^*$. As $L^p \overset{T}{\cong} L^p_\eta$, T^* is also an isometry between $(L^p_\eta)^*$ and $(L^p)^*$. Now let $g \in L^q$ and $\ell \in (L^p)^*$ such that $\langle \ell, h\rangle = \int h\, g$, for all $h \in L^p$. As $\ell \in (L^p)^*$, this implies $(T^*)^{-1}\ell \in (L^p_\eta)^*$, thus for any $f \in L^p_\eta$,

$$\big\langle (T^*)^{-1}\ell, f\big\rangle = \big\langle (T^{-1})^*\ell, f\big\rangle = \big\langle \ell, T^{-1}f\big\rangle = \big\langle \ell, f\,\eta^{1/p}\big\rangle = \int g\, f\, \eta^{1/p}.$$

Hence, for $f \in L^p_\eta$, $g \in L^q$, there exists $\tilde{\ell} \in (L^p_\eta)^*$ such that,

$$\langle \tilde{\ell}, f \rangle = \langle (T^*)^{-1}\ell, f \rangle = \int g\, f\, \eta^{1/p}.$$

$\square$

Lemma 5.7.2. *The functional I defined by (5.3.19) is sequentially weakly lower semi-continuous (l.s.c).*

Proof of Lemma 5.7.2. To show that $I\colon H(E) \to \mathbb{R}$ is sequentially weakly l.s.c, we first show this property for a truncation of I, I_m defined by,

$$I_m(u) := \int_E \left(\frac{a^2\xi^2}{2} + (\kappa \wedge m)u^2 - 2u \right) \eta dx, \tag{5.7.2}$$

where $m \in \mathbb{N}$. Next, fix $u \in H(E)$ and let u_n be a sequence in $H(E)$ such that $u_n \rightharpoonup u \in H(E)$[6]. Now define the mapping,

$$(u, \xi) \mapsto \left(\frac{a^2\xi^2}{2} + (\kappa \wedge m)u^2 - 2u \right) \eta =: f_m(x, u, \xi).$$

This mapping is convex in (u, ξ), therefore

$$f_m(x, u_n, \xi_n) \geq f_m(x, u, \xi) + \partial_u f_m(x, u, \xi) \cdot (u_n - u) + \partial_\xi f_m(x, u, \xi) \cdot (\xi_n - \xi),$$

and, integrating,

$$I_m(u_n) \geq I_m(u) + \int_E \partial_u f_m(x, u, \xi) \cdot (u_n - u) dx + \int_E \partial_\xi f_m(x, u, \xi) \cdot (\xi_n - \xi) dx.$$

It suffices to show that both integrals on the RHS vanish as $n \to \infty$ to conclude that $\liminf_{n\to\infty} I_m(u_n) \geq I_m(u)$. As $u_n \rightharpoonup u$ in $H(E)$, this means $u_n \rightharpoonup u$ in $L^2_\eta(E)$ and $\xi_n \rightharpoonup \xi$ in $L^2_{a^2,\eta}(E)$. Lemma 5.7.1 yields the necessary construction of an element of the dual of the weighted spaces $L^2_\eta(E)$ and $L^2_{a^2,\eta}(E)$ respectively. The first integral satisfies,

$$\int_E \partial_u f_m(x, u, \xi)(u_n - u) dx = \int_E 2((\kappa\wedge m)u - 1)\eta(u_n - u) dx = \int_E 2((\kappa\wedge m)u - 1)\eta^{1/2}(u_n - u)\eta^{1/2} dx.$$

By Lemma 5.7.1, it is enough to show that $2((\kappa \wedge m)u - 1)\eta^{1/2} \in L^2$ to ensure that

[6] Recall that $u_n \rightharpoonup u \in H(E)$ means the sequence u_n converges to u weakly in $H(E)$, i.e for any $x^* \in H(E)^*$ we have $\langle x^*, u_n - u \rangle \to 0$.

$\int_E \partial_u f_m(x,u,\xi)\cdot(u_n-u)dx \to 0$ as $n\to\infty$. Now,

$$\begin{aligned}\|2((\kappa\wedge m)u-1)\eta^{1/2}\|_{L^2}^2 = \int_E 4((\kappa\wedge m)u-1)^2\eta dx &\le 8\int_E(\kappa\wedge m)^2u^2\eta dx + 8\\ &\le 8m^2\int_E u^2\eta dx + 8\\ &\le 8m^2\|u\|_{L^2_\eta} + 8 < \infty,\end{aligned}$$

as required. Likewise the second integral satisfies

$$\int_E \partial_\xi f_m(x,u,\xi)(\xi_n-\xi)dx = \int_E a^2\xi\eta(\xi_n-\xi)dx = \int_E a\xi\eta^{1/2}(\xi_n-\xi)a\eta^{1/2}dx.$$

Once again, by Lemma 5.7.1, $\int_E \partial_\xi f_m(x,u,\xi)\cdot(\xi_n-\xi)dx \to 0$ because $\|a\,\xi\,\eta^{1/2}\|_{L^2} = \|\xi\|_{L^2_{a^2,\eta}} < +\infty$. Thus,

$$\liminf_{n\to\infty} I_m(u_n) \ge I_m(u).$$

Recall $I(u_n)\ge I_m(u_n)$, thus $\liminf_{n\to\infty} I(u_n) \ge \liminf_{n\to\infty} I_m(u_n) \ge I_m(u)$. The monotone convergence theorem yields $I_m(u)\uparrow I(u)$, thus passing to the limit we see

$$\lim_{m\to\infty}\liminf_{n\to\infty} I(u_n) \ge \lim_{m\to\infty} I_m(u) = I(u).$$

But the left hand side is independent of m, therefore

$$\liminf_{n\to\infty} I(u_n) \ge I(u),$$

the proof is complete.

□

Theorem 5.7.1 (Existence and Uniqueness)**.** *Let Assumptions 5.3.1-5.3.4 hold. There exists a unique solution to the minimisation problem* (5.3.21).

Proof of Theorem 5.7.1.

Step 1 (Compactness). First we show that $\inf\{I(u) : u\in H(E)\} = m$, is finite. Fix an arbitrary $u\in H(E)$, then

$$I(u) = \int_E\left(\frac{a^2\xi^2}{2}+\kappa u^2-2u\right)\eta dx = \int_E\left(\frac{a^2\xi^2}{2}+\kappa(u-\kappa^{-1})^2-\kappa^{-1}\right)\eta dx \ge -\int_E \kappa^{-1}\eta dx$$

The last integral is finite because κ is bounded below away from zero and η is a probability density, thus $I(u) > -\infty$. To see that $m < +\infty$, set $u=0$. Hence, m is indeed finite. Now, let u_n be a minimising sequence such that $I(u_n)\to m$ as $n\to\infty$. Such a sequence always exists by definition of an infimum. Next, we show that u_n is a bounded sequence in $H(E)$. Define $C_1 := \inf_{x\in E} \kappa(x) > 0$. Then,

$$\kappa \cdot u_n^2 - 2u_n \geq C_1 \cdot u_n^2 - 2u_n = \frac{C_1}{2}u_n^2 + \frac{C_1}{2}\left(u_n - \frac{2}{C_1}\right)^2 - \frac{2}{C_1} \geq \frac{C_1}{2}u_n^2 - \frac{2}{C_1}.$$

Because u_n is a minimising sequence, I satisfies for sufficiently large n,

$$\begin{aligned} m + 1 \geq I(u_n) &= \int_E \left(\frac{a^2\xi_n^2}{2} + \kappa u_n^2 - 2u_n\right)\eta dx \\ &\geq \int_E \left(\frac{a^2\xi_n^2}{2} + \frac{1}{2}C_1 u_n^2 - 2C_1^{-1}\right)\eta dx \\ &\geq \frac{1}{2}\int_E \left(a^2\xi_n^2 + C_1 u_n^2\right)\eta dx - 2C_1^{-1} \geq C_2\|u_n\|_{H(E)}^2 - 2C_1^{-1}, \end{aligned}$$

where $C_2 = \min\{\frac{1}{2}, C_1\}$. Thus,

$$\|u_n\|_{H(E)}^2 \leq \left(2C_1^{-1} + m + 1\right) C_2^{-1},$$

for sufficiently large n, where the RHS is a strictly positive constant. Thus, passing to a subsequence (and with a slight abuse of indexing notation) u_n is a bounded sequence. Note that Assumptions 5.3.1 and 5.3.4 ensure that $\frac{1}{a^2\eta}, \frac{1}{\eta} \in L^1_{loc}(E;\mathbb{R})$, whence $H(E)$ is a Banach space (Theorem 1.11 in Kufner et. al. [76]). Passing to another subsequence (still denoted u_n), it follows that

$$u_n \rightharpoonup \bar{u} \in H(E).$$

Step 2 (Lower Semicontinuity). By Lemma 5.7.2, I is sequentially weakly l.s.c., hence

$$\liminf_{n\to\infty} I(u_n) \geq I(\bar{u}),$$

and $\bar{u}$ is indeed a minimiser.

Step 3 (Uniqueness). Uniqueness follows from strict convexity of the mapping,

$$(u, \xi) \mapsto \left(\frac{a^2\xi^2}{2} + \kappa u^2 - 2u\right)\eta.$$

$\square$

Theorem 5.7.2 (Regularity of the Solution)**.** *Let $u : E \to \mathbb{R}$, be the unique solution to the above variational problem. Then there exists $v \in C^2(E)$ such that $u = v$ a.e.*

Proof of Theorem 5.7.2. As u solves the variational problem, $u \in H(E)$. Therefore there exists a function $\xi \in L^1_{loc}(E)$ such that

$$\int_E u(x)\varphi'(x)dx = -\int_E \xi(x)\varphi(x)dx \text{ for all } \varphi \in C_c^\infty(E).$$

Defining $v(x) := \int_{x_0}^x \xi(y)dy$ for some $x_0 \in E$, any $\varphi \in C_c^\infty(E)$ satisfies (setting $E = (a,b)$)

$$\begin{aligned}
\int_E v(x)\varphi'(x)dx &= \int_E \left[\int_{x_0}^x \xi(y)dy\right] \varphi'(x)dx = \int_E \int_{x_0}^x \xi(y)\varphi'(x)dydx \\
&= \int_a^b \int_{x_0}^x \xi(y)\varphi'(x)dydx \\
&= \int_a^{x_0} \int_{x_0}^x \xi(y)\varphi'(x)dydx + \int_{x_0}^b \int_{x_0}^x \xi(y)\varphi'(x)dydx \\
&= -\int_a^{x_0} \int_x^{x_0} \xi(y)\varphi'(x)dydx + \int_{x_0}^b \int_{x_0}^x \xi(y)\varphi'(x)dydx \\
&= -\int_a^{x_0} \int_a^y \xi(y)\varphi'(x)dxdy + \int_{x_0}^b \int_y^b \xi(y)\varphi'(x)dxdy \\
&= -\int_a^{x_0} \xi(y)\left[\varphi(y) - \varphi(a)\right]dy + \int_{x_0}^b \xi(y)\left[\varphi(b) - \varphi(y)\right])dy \\
&= -\int_E \xi(y)\varphi(y)dy.
\end{aligned}$$

Thus,

$$\int_E v(y)\varphi'(y)dy = -\int_E \xi(y)\varphi(y)dy = \int_E u(y)\varphi'(y)dy,$$

hence,

$$\int_E (v(y) - u(y))\varphi'(y)dy = 0.$$

As φ is chosen arbitrarily, this equation holds for all $\varphi \in C_c^\infty(E)$, hence $v = u$ a.e.. The solution of the variational problem is unique up to $a.e.$ equivalence, therefore v is also a solution. As v is a minimiser of $I : H(E) \to \mathbb{R}$,

$$\lim_{\varepsilon \to 0} \frac{I(v + \varepsilon\varphi) - I(v)}{\varepsilon},$$

if it exists must be zero, where $\varphi \in C_c^\infty(E)$ is arbitrary. We now show that this limit indeed exists. Dropping function arguments to ease notation,

$$
\begin{aligned}
I(v+\varepsilon\varphi)-I(v) &= \int_E \left(\frac{a^2}{2}\left[(\xi+\varepsilon\varphi')^2-(v')^2)\right]+\kappa\left((v+\varepsilon\varphi)^2-v^2\right)-2(v+\varepsilon\varphi-v))\right)\eta dy \\
&= \int_E \left(\frac{a^2}{2}\left[2\xi\varepsilon\varphi'+\varepsilon^2(\varphi')^2\right]+\kappa(2v\varepsilon\varphi+\varepsilon^2\varphi^2)-2\varepsilon\varphi\right)\eta dy \\
&= \int_E \left(\varepsilon a^2\xi\varphi'+\varepsilon^2\frac{a^2(\varphi')^2}{2}+\varepsilon 2\kappa v\varphi+\varepsilon^2\kappa\varphi^2-2\varepsilon\varphi\right)\eta dy \\
&= \varepsilon\int_E \left(a^2\xi\varphi'+2\kappa v\varphi-2\varphi\right)\eta dy+\varepsilon^2\int_E\left(\frac{a^2(\varphi')^2}{2}+\kappa\varphi^2\right)\eta dy.
\end{aligned}
$$

Thus,

$$
\frac{I(v+\varepsilon\varphi)-I(v)}{\varepsilon}=\int_E \left(a^2\xi\varphi'+2\kappa v\varphi-2\varphi\right)\eta dy+\varepsilon\int_E\left(\frac{a^2(\varphi')^2}{2}+\kappa\varphi^2\right)\eta dy.
$$

The compact support of φ ensures that all integrals are well defined. Letting $\varepsilon\to 0$

$$
\int_E \left(a^2\xi\eta\right)\varphi'+2\left(\kappa v-1\right)\eta\varphi dy=0,
$$

hence,

$$
\int_E \left(a^2\xi\eta\right)\varphi' dy=-\int_E 2\left(\kappa v-1\right)\eta\varphi dy.
$$

As φ is arbitrary, $2\left(\kappa v-1\right)\eta$ is the weak derivative of $a^2\xi\eta$, and integration by parts yields,

$$
-\int_E 2\left(\kappa v-1\right)\eta\varphi dy=\int_E\int_{y_0}^{y} 2\left(\kappa v-1\right)\eta ds\varphi' dy.
$$

Thus,

$$
\int_E \left[a^2\xi\eta-\int_{y_0}^{y} 2\left(\kappa v-1\right)\eta ds\right]\varphi' dy=0.
$$

Once again, as φ is arbitrary,

$$
a^2\xi\eta=\int_{y_0}^{y} 2\left(\kappa v-1\right)\eta ds, \quad a.e.
$$

which, by the regularity of the model's coefficients, implies $\xi\in C(E)$, hence $v\in C^1(E)$. Self regularisation by means of the above equation yields $v\in C^2(E)$ as required. $\square$

5.7.2 Qualitative Verification

The proof of Theorem 5.4.1 requires a lemma that gives an a priori upper and lower bound on the candidate value function g. The proof follows essentially by inspection of the variational problem.

Lemma 5.7.3. *Let g be the solution to the minimisation problem (5.3.21). Then,*

$$0 \leq \frac{1}{\sup_{s\in E}\kappa(s)} \leq g(y) \leq \frac{1}{\inf_{s\in E}\kappa(s)} < +\infty,$$

for all $y \in E$. In addition, $g(y) > 0$ for all $y \in E$.

Proof of Lemma 5.7.3. To prove the claim, rewrite f as

$$f(y,g,g') = \left(\frac{a^2(y)(g'(y))^2}{2} + \kappa(y)\left(g(y) - \kappa^{-1}(y)\right)^2 - \kappa^{-1}(y)\right)\eta(y).$$

Now let $\tilde{g} := g \wedge c$ where $c := \sup_{s\in E}\frac{1}{\kappa(s)}$ which is a finite, strictly positive constant by Assumption 5.3.3. Note $\tilde{g} \in H(E)$ see [124, Corollary 2.1.6]. Now assume there exists a set $E' \subseteq E$ such that $\tilde{g} = c$ on E'. Then,

$$\begin{aligned}
I(g) - I(\tilde{g}) &= \int_E \left(\frac{a^2(g')^2}{2} + \kappa(g-\kappa^{-1})^2 - \kappa^{-1}\right)\eta \\
&\quad - \int_E \left(\frac{a^2\,(\tilde{g}')^2}{2} + \kappa(\tilde{g}-\kappa^{-1})^2 - \kappa^{-1}\right)\eta \\
&= \int_{E'} \left(\frac{a^2}{2}(g')^2 + \kappa\left[(g-\kappa^{-1})^2 - (c-\kappa^{-1})^2\right]\right)\eta \\
&= \int_{E'} \left(\frac{a^2}{2}(g')^2 + \kappa(g-\kappa^{-1}-c+\kappa^{-1})(g-\kappa^{-1}+c-\kappa^{-1})\right)\eta \\
&= \int_{E'} \left(\frac{a^2}{2}(g')^2 + \kappa(g-c)(g+c-2\kappa^{-1})\right)\eta.
\end{aligned}$$

Now as $\tilde{g}|_{E'} = c$ this implies $g \geq c$ on E'. Recall $c = \sup_{s\in E}\frac{1}{\kappa(s)} \implies c \geq \kappa^{-1}$. Thus

$$\kappa(g-c)(g+c-2\kappa^{-1}) \geq \kappa(g-c)(2c-2\kappa^{-1}) \geq 0.$$

Hence the integrand above is non-negative and so $I(g) \geq I(\tilde{g})$, contradicting that g is the minimiser. Thus, either no such set E' can exist or else E' has Lebesgue measure zero which implies $g \leq c = \sup_{s\in E}\frac{1}{\kappa(s)}$ almost everywhere, but g is continuous so this bound must hold on all of E. An identical argument is used to show $g \geq \inf_{s\in E}\frac{1}{\kappa(s)}$. Hence for all $y \in E$ we have proven

$$0 \leq \inf_{s\in E}\kappa^{-1}(s) \leq g(y) \leq \sup_{s\in E}\kappa^{-1}(s) = \frac{1}{\inf_{s\in E}\kappa(s)} < +\infty.$$

To prove $g(y) > 0$ for all $y \in E$ we proceed by contradiction. Assume there exists $x_0 \in E$ such that $g(x_0) = 0$. To analyse the Taylor expansion of g around the point x_0 up to the third order, we require extra regularity in g. Because g satisfies the HJB equation,

$$\frac{a^2(y)}{2}g''(y) + \tilde{b}(y)g'(y) - \kappa(y)g(y) = -1,$$

its second derivative satisfies

$$g''(y) = -\frac{2\tilde{b}(y)}{a^2(y)}g'(y) + \frac{2\kappa(y)}{a^2(y)}g(y) - \frac{2}{a^2(y)}.$$

The terms, $\frac{2\tilde{b}(y)}{a^2(y)}, \frac{2\kappa(y)}{a^2(y)}, \frac{2}{a^2(y)}$ are $C^1(E;\mathbb{R})$ due to the regularity assumptions on the model coefficients (Assumption 5.3.1). We know $g \in C^2(E;\mathbb{R})$ by Theorem 5.7.2, so we may differentiate both sides to obtain,

$$\begin{aligned} g'''(y) = &\left(\frac{-2\tilde{b}'(y)}{a^2(y)} + \frac{4\tilde{b}(y)a'(y)}{a^3(y)}\right)g'(y) - \frac{2\tilde{b}(y)}{a^2(y)}g''(y) \\ &+ \left(\frac{2\kappa'(y)}{a^2(y)} - \frac{4\kappa(y)a'(y)}{a^3(y)}\right)g(y) - \frac{2\kappa(y)}{a^2(y)}g'(y) + \frac{2a'(y)}{a^3(y)}. \end{aligned}$$

which shows $g \in C^3(E;\mathbb{R})$. Next we perform a Taylor expansion on the half open interval $[x_0, x_0 + \varepsilon)$,

$$g(y) = g(x_0) + g'(x_0)(y - x_0) + \frac{g''(x_0)}{2}(y - x_0)^2 + \frac{g'''(\alpha)}{6}(y - x_0)^3,$$

for some $\alpha \in [x_0, x_0 + \varepsilon)$. As $g(y) \geq 0$ for all $y \in E$, $g'(x_0) = 0$. Thus the HJB equation implies that $g''(x_0) = -2/a^2(x_0) < 0$, hence

$$g(y) = \frac{-(y - x_0)^2}{a^2(x_0)} + \frac{g'''(\alpha)}{6}(y - x_0)^3.$$

Now as $y \in (x_0, x_0 + \varepsilon)$,

$$g(y) \leq \frac{-(y - x_0)^2}{a^2(x_0)} + \sup_{s \in [x_0, x_0 + \varepsilon]} |g'''(s)| \frac{(y - x_0)^3}{6}.$$

Now we fix y close enough to x_0 to ensure the right hand side is strictly negative, which gives us our desired contradiction, hence no such point x_0 exists. □

Proof of Theorem 5.4.1. Part (i) is proven by combining Theorems 5.7.1 and 5.7.2. Part (ii) follows directly from part (i) along with Proposition 5.3.1. To prove part (iii) we employ Theorem 3.3 in Guasoni and Wang [56] which has two conditions that need to be satisfied. Our assumption of a unique solution to the martingale problem in part (iii) of Theorem 5.4.1 is exactly the first condition of Theorem 3.3 in [56]. The second condition to check is that $\int_0^\infty g(Y_t)^{-1}dt = \infty \, \hat{\mathbb{P}} - a.s.$, but this is an immediate consequence of the fact g is bounded (and positive), which follows from Lemma 5.7.3. Hence

$$\int_0^\infty g(Y_t)^{-1}dt \geq \int_0^\infty \inf_{x\in E}\kappa(x)dt = +\infty.$$

Thus, Theorem 3.3 in [56] applies and the proof is complete. $\square$

Proof of Theorem 5.4.2. Assumptions 5.3.1-5.3.4 ensure Theorem 5.4.1 applies. Thus we need only check the martingale problem from Theorem 5.4.1 has a unique solution and we are done. Note the coefficients of the martingale problem depend only on the final coordinate, hence it is sufficient to prove the existence of a unique weak solution to the scalar SDE: $dY_t = \left(\tilde{b}(Y_t) + \frac{a^2(Y_t)g'(Y_t)}{g(Y_t)}\right)dt + a(Y_t)dW_t$. Assumption 5.3.1 and Theorem 5.4.1 ensure both the drift and diffusion coefficients are continuous and $a^2(y) > 0$ for all $y \in E$. Thus there exists a unique weak solution up to an exit time τ of the above scalar SDE, see [35, Proposition 2.2]. Next we show the boundary is never reached. Let P denote the scale function for the above SDE, i.e for some $c \in E$,

$$P(y) = \int_c^y \exp\left(-2\int_c^u \frac{\tilde{b}(w)}{a^2(w)} + \frac{g'(w)}{g(w)}dw\right)du.$$

Lemma 5.7.3 implies $g(y) > 0$ for all $y \in E$. Hence

$$\begin{aligned}
P(y) &= \int_c^y \exp\left(-2\int_c^u \frac{\tilde{b}(w)}{a^2(w)} + \frac{d\log(g(w))}{dw}dw\right)du \\
&= \int_c^y \exp\left(-2\int_c^u \frac{\tilde{b}(w)}{a^2(w)}\right)\frac{g^2(c)}{g^2(u)}du
\end{aligned}$$

Now assume $y > c$, then

$$\begin{aligned}
P(y) &= \int_c^y \exp\left(-2\int_c^u \frac{\tilde{b}(w)}{a^2(w)}dw\right)\frac{g^2(c)}{g^2(u)}du \\
&\geq \frac{g^2(c)}{\sup_{s\in E} g^2(s)}\int_c^y \exp\left(-2\int_c^u \frac{\tilde{b}(w)}{a^2(w)}dw\right)du \\
&= \frac{g^2(c)}{\sup_{s\in E} g^2(s)}p(y),
\end{aligned}$$

where p is the scale function for the SDE: $dZ_t = \tilde{b}(Z_t)dt + a(Z_t)dW_t$. By Assumption 5.4.1 $p(y) \to \infty$ as $y \to \partial E_+$, thus $P(y) \to \infty$ as $y \to \partial E_+$. For $y < c$ by an identical argument we have

$$P(y) \leq \frac{g^2(c)}{\sup_{s\in E} g^2(s)}p(y).$$

By Assumption 5.4.1 $p(y) \to -\infty$ as $y \to \partial E_-$, thus $P(y) \to -\infty$ as $y \to \partial E_-$. Hence we have $\mathbb{P}(\tau = \infty) = 1$, see Karatzas and Shreve [70, Chapter 5, Section C] in particular Problem 5.27 and Theorem 5.29. Thus the SDE: $dY_t = \left(\tilde{b}(Y_t) + \frac{a^2(Y_t)g'(Y_t)}{g(Y_t)}\right) dt + a(Y_t)dW_t$ has a unique weak solution on E which implies the martingale problem from Theorem 5.4.1 is well posed. The proof is complete.

□

Proof of Proposition 5.4.1. Assumptions 5.3.1-5.3.4 are trivially true. We now check Assumption 5.4.1 for each model (5.4.3)-(5.4.5) respectively. Throughout this proof $c > 0$ denotes an arbitrary element of E. First we check model (5.4.3), the scale function in this model is given by

$$p(y) = C \int_c^y \exp\left(\frac{b(\theta - u)^4}{2a^2} + \frac{\lambda u^2}{a^2}\right) du$$

for some $C > 0$. By inspection we have $p(y) \to \infty$ as $y \to \infty$ and $p(y) \to -\infty$ as $y \to -\infty$ as required. For model (5.4.4) the scale function is given by

$$p(y) = C \int_c^y u^{-\frac{2b\theta}{a^2}} \exp\left(\frac{2bu}{a^2} + \frac{2\lambda u^\delta}{a^2 \delta}\right) du,$$

for some $C > 0$. The parameter restrictions on λ and δ ensure the leading order exponent in the exponential term is positive and so $p(y) \to \infty$ as $y \to \infty$. Note as $2b\theta > a^2$, the singular term in the integrand will not be integrable at zero. Thus we also have $p(y) \to -\infty$ as $y \to -\infty$ as required. For model (5.4.5) the scale function is given by

$$p(y) = C \int_c^y u^{-\frac{2b\theta}{a^2}} \exp\left(\frac{2}{a^2}(bu - 2\lambda\sqrt{u})\right) du,$$

for some $C > 0$. The analysis for this model is identical to the previous and so the proposition is proven.

□

5.7.3 Asymptotics

Before proving Theorem 5.5.1 we need several preparatory lemmata. Recall throughout this section we let Assumptions 5.3.1-5.3.4 hold. First we prove Lemma 5.5.1 which requires a result from dynamical systems and control theory known as Barbălat's Lemma which we recall for the readers convenience.

Lemma 5.7.4. *Suppose that $f : [0, \infty) \to \mathbb{R}$ is uniformly continuous and that $\lim_{t\to\infty} \int_0^t f(s)ds$ exists. Then $\lim_{t\to\infty} f(t) = 0$.*

Proof. We refer readers to [49] for the proof. □

Using a time reversal argument one can show that Barbălat's Lemma also applies to uniformly continuous functions defined on $(-\infty, 0]$.

Proof of Lemma 5.5.1. Recall g must satisfy the corresponding Euler-Lagrange equation which reads

$$\frac{d}{dy}\left(a^2(y)\eta(y)g'(y)\right) = 2\left(\kappa(y)g(y)-1\right)\eta(y). \tag{5.7.3}$$

Now set $f(y) = a^2(y)\eta(y)g'(y)$ and note that $f \in C^1([m,\infty);\mathbb{R}) \cap L^1([m,\infty);\mathbb{R})$. Continuous differentiability comes from the fact $g \in C^2(E;\mathbb{R})$ and the regularity assumptions on the coefficients of the model. To see why $f \in L^1([m,\infty);\mathbb{R})$ we apply Hölders inequality and use the assumption on the integrability of $a^2\eta$.

$$\int_{[m,\infty)} |g'(y)|a^2(y)\eta(y)dy \le \left(\int_E |g'(u)|^2 a^2(u)\eta(u)du\right)^{\frac{1}{2}} \left(\int_{[m,\infty)} a^2(u)\eta(u)du\right)^{\frac{1}{2}} < +\infty.$$

Now recall that $|g(y)| \le C$ by Lemma 5.7.3, thus,

$$|f'(y)| = |2(\kappa(y)g(y)-1)\eta(y)| \le 2C\kappa(y)\eta(y) + 2\eta(y) \le B,$$

where the last inequality follows from our supposition on the boundedness of $\kappa\eta$ on $U(m)$ (which also implies η is bounded on $U(m)$ as $\inf_{s\in E}\kappa(y) > 0$). Thus f is continuously differentiable with a bounded derivative, which implies f is Lipschitz continuous and thus uniformly continuous. But integrability and uniform continuity imply that f must vanish at infinity by Barbălat's Lemma 5.7.4. An identical argument on $(-\infty,-m]$ yields the claim. □

Lemma 5.7.5. *Let $g \in C^2(E;\mathbb{R})$ be the unique solution to (5.3.21). Let Assumption 5.5.2 hold. Then if $g'(y) \le 0$ ($g'(y) \ge 0$) for all $y \in [m,\infty)$ ($y \in (-\infty,-m]$) where $m > 0$ is large, then,*

$$\lim_{|y|\to\infty} g(y) = 0.$$

Proof of Lemma 5.7.5. Fix $x \in [m,\infty)$. The HJB equation $\left(\frac{a^2}{2}g'' + \tilde{b}g' - \kappa g = -1\right)$ ensures,

$$\int_x^{2x} \left(\frac{a^2}{2}g'' + \tilde{b}g' - \kappa g + 1\right)\psi dy = 0.$$

Performing integration by parts twice yields

$$0 = \int_x^{2x} \left(\frac{a^2}{2}g'' + \tilde{b}g' - \kappa g + 1\right)\psi dy = \int_x^{2x} \left(\frac{1}{2}[a^2\psi]'' - [\tilde{b}\psi]' - \kappa\psi\right) g\, dy + \Phi(x).$$

Rearranging, taking absolute values and using the boundedness of g (which follows from Lemma 5.7.3),

$$\int_x^{2x} \kappa g\psi dy \le \frac{\gamma}{\beta}\int_x^{2x} \left(\left|[a^2\psi]''\right| + \left|[\tilde{b}\psi]'\right|\right) dy + x.$$

Employing (5.5.1) from Assumption 5.5.2 and the fact g is strictly positive and non-increasing to estimate the left hand side below yields,

$$\frac{\alpha x(1+|x|^p)}{C_2} g(2x) \le \int_x^{2x} \kappa g \psi dy.$$

Using (5.5.2) from Assumption 5.5.2 yields,

$$\frac{\gamma}{\beta} \int_x^{2x} \left(\left| [a^2\psi]'' \right| + \left| [\tilde{b}\psi]' \right| \right) dy + x \le \frac{\gamma C_3}{\beta}(1+|x|^p) + x.$$

Combining these estimates yields

$$\frac{\alpha x(1+|x|^p)}{C_2} g(2x) \le \frac{\gamma C_3}{\beta}(1+|x|^p) + x.$$

Rearranging, this estimate implies that $g(2x) \to 0$ as $x \to \infty$. The limit at $-\infty$ is obtained by using the same test function, shifting the support to the interval $[-2x, -x]$. □

Lemma 5.7.6. *Let g be the solution to the minimisation problem (5.3.21) and let Assumption 5.5.2 hold. Then $g(y) \to 0$ as $|y| \to \infty$.*

Proof of Lemma 5.7.6. Fix $\varepsilon \in (0,1)$ and choose x_0 such that $\kappa^{-1}(x_0) \le C_1(1+|x_0|^p)^{-1} = \varepsilon$. Letting $h(x) := C_1(1+|x|^p)^{-1}$, we identify three possible cases,

$$(i)\; g(x_0) > h(x_0) = \varepsilon; \qquad (ii)\; g(x_0) = h(x_0) = \varepsilon; \qquad (iii)\; g(x_0) < h(x_0) = \varepsilon.$$

First consider case (ii), we show $g(y) \le \varepsilon$ for all $y > x_0$. Introduce

$$\tilde{g}(y) = \begin{cases} g(y) \wedge \varepsilon, & y \ge x_0, \\ g(y), & y < x_0. \end{cases}$$

Note that $g(x_0) = h(x_0) = \varepsilon$ ensures $\tilde{g}$ is continuous. Now we assume there exists a set $E' \subseteq [x_0, +\infty)$ such that $\tilde{g}(y) = \varepsilon$ for $y \in E'$. We proceed as in the proof of Lemma 5.7.3. Thus,

$$I(g) - I(\tilde{g}) = \int_{E'} \left(\frac{a^2}{2}(g')^2 + \kappa(g-\varepsilon)(g+\varepsilon-2\kappa^{-1}) \right) \eta.$$

Now for all $y \in E'$,

$$g(y) \ge \varepsilon \implies g(y)+\varepsilon \ge 2\varepsilon \implies g(y)+\varepsilon-2\kappa^{-1}(y) \ge 2\varepsilon - 2\kappa^{-1}(y) \ge 2\varepsilon - 2h(y) \ge 0,$$

where the second last inequality follows by Assumption 5.5.2, and the last because h is monotonically decreasing on $[x_0, \infty)$ with $h(x_0) = \varepsilon$. Thus $I(g) - I(\tilde{g}) \ge 0$, but this contradicts the optimality of g and hence $g(y) \le \varepsilon$ for all $y > x_0$.

Next, consider case (iii) $g(x_0) < h(x_0) = \varepsilon$. If $g(y) < h(y)$ for all $y \in [x_0, +\infty)$ then $g \to 0$ as $y \to \infty$, otherwise there must exist a point $x_1 \in [x_0, +\infty)$ such that $g(x_1) = h(x_1) = \varepsilon' < \varepsilon$. But then the argument in case (ii) yields that $g(y) \leq \varepsilon'$ for all $y > x_1$, whence $g(y) < \varepsilon$ for all $y > x_1$.

Finally we consider case (i) $g(x_0) > h(x_0) = \varepsilon$. We prove there exists an $x_2 \geq x_0$ such that $g(x_2) = \varepsilon$, then the argument from case (ii) yields $g(y) \leq \varepsilon$ for all $y > x_2$.

We proceed by contradiction and suppose that $g(y) > \varepsilon$ for all $y \in [x_0, \infty)$. We show this supposition implies that $g'(y) \leq 0$ for all $y \in [x_0, \infty)$, whence we can invoke Lemma 5.7.5 to claim $g(y) \to 0$ as $y \to \infty$, a contradiction, and so there must exist a point $x_2 > x_0$ such that $g(x_2) = \varepsilon$, as required.

To show $g'(y) \leq 0$ for all $y \in [x_0, \infty)$ we once again proceed by contradiction and assume $g'(x) > 0$ for some $x \in [x_0, \infty)$. This implies the existence of $\delta > 0$ such that $g(y) > g(x)$ for all $y \in (x, \delta)$. Let $\bar{\delta} > \delta$ be the first point such that $g(\bar{\delta}) = g(x)$. Then introducing the truncation $\tilde{g}(y) := g(y) \wedge g(x)$ will result in $I(g)|_{[x,\bar{\delta})} - I(\tilde{g})|_{[x,\bar{\delta})} \geq 0$ contradicting optimality of g. If no such point $\bar{\delta}$ exists then $I(g)|_{[x,\infty)} - I(\tilde{g})|_{[x,\infty)} \geq 0$ once again contradicting optimality. Hence we must have $g'(y) \leq 0$ for all $y \in [x_0, +\infty)$.

Thus we have proven for any fixed $\varepsilon > 0$, $g(y) < \varepsilon$ for all $y > \max(x_0, x_1, x_2)$. The argument to ensure $g(y) < \varepsilon$ for all $y < -x$ where x is some large positive number is identical.

□

Lemma 5.7.7. *Let $g \in C^2(E; \mathbb{R})$ be the solution to the minimisation problem (5.3.21) and let Assumptions 5.5.2-5.5.3 hold. Then, there exists constants $C_1, C_2, q > 0$ such that, for all $y \in E$,*

$$\frac{C_1}{(1 + |y|^q)} \leq g(y) \leq \frac{C_2}{(1 + |y|^q)}.$$

The proof of Lemma 5.7.7 requires the maximum principle from the theory of elliptic PDEs which we now recall for the readers convenience.

Lemma 5.7.8. *Let $V \subseteq \mathbb{R}^d$ (bounded or unbounded) be a connected open set with a smooth (or empty) boundary. Let $u \in C^2(V), a^{ij}, b^j, c \in C^0(\bar{V})$, a^{ij} strictly elliptic in the interior of E and $c \geq 0$, and assume that*

$$Lu = \left(a^{ij}\partial_i\partial_j + b^j\partial_j - c\right) u \leq 0.$$

Assume the infimum of u is attained in the interior of V, so that $u(x_0) = \inf_{x \in V} u(x)$ for $x_0 \in V$. Then, either $u(x_0) > 0$ or $u(x_0) = u(x)$ for all $x \in V$.

Proof of Lemma 5.7.8. We refer readers to [52, Theorem 3.5] for the proof. □

Proof of Lemma 5.7.7. Assume always $y \in U(m)$ for some large $m > 0$. To prove the upper bound, note that by Assumption 5.5.3, $\mathcal{L}h \leq -C_2$ for some $C_2 > 0$. Thus we also have $\mathcal{L}h \leq -\varepsilon$ for all $\varepsilon \in (0, C_2)$. Recall from the HJB equation that $\mathcal{L}(g) = -1$, therefore

$$\mathcal{L}\left(\frac{h}{\varepsilon} - g\right) = \frac{\mathcal{L}h}{\varepsilon} - \mathcal{L}g \leq -1 + 1 = 0 \text{ for all } \varepsilon \in (0, C_2).$$

By the maximum principle[7] (i.e, Lemma 5.7.8), if the minimum is in the interior, then either $h/\varepsilon - g \geq \inf(h/\varepsilon - g) > 0$ or $h/\varepsilon - g = C$. Note that C must be non-negative as

$$-\kappa C = \mathcal{L}(C) = \mathcal{L}\left(\frac{h}{\varepsilon} - g\right) \leq 0.$$

As $\kappa > 0$ we must also have $C \geq 0$ and so in both cases we see $h/\varepsilon - g \geq 0$. Now if the infimum occurs at the boundary it may be at infinity or $\pm m$, first we check infinity and recall by Assumption 5.5.2, Lemma's 5.7.5 and 5.7.6 hold, which ensure $g(y) \to 0$ as $|y| \to \infty$. Hence

$$\frac{h(y)}{\varepsilon} - g(y) \geq \liminf_{|x| \to \infty} \left(\frac{h(x)}{\varepsilon} - g(x)\right) = 0.$$

Now as $|x| \to m$,

$$\frac{h(y)}{\varepsilon} - g(y) \geq \liminf_{|x| \to m} \left(\frac{h(x)}{\varepsilon} - g(x)\right) = \begin{cases} \dfrac{h(m)}{\varepsilon} - g(m), & \text{as } x \downarrow m, \\ \dfrac{h(-m)}{\varepsilon} - g(-m), & \text{as } x \uparrow -m. \end{cases}$$

Now recall that ε is an arbitrary element of $(0, C_2)$, thus we can always ensure the right hand side is positive for small enough ε. This follows from the boundedness of g. Thus, $g(y) \lesssim h(y)$ as required. Next we prove the lower bound. Now $\mathcal{L}h \geq -C_1$ so in particular $\mathcal{L}h \geq -\delta$ for all $\delta \in (C_1, \infty)$. By Assumption 5.5.3,

$$\mathcal{L}\left(g - \frac{h}{\delta}\right) = \mathcal{L}g - \frac{\mathcal{L}h}{\delta} \leq -1 + 1 = 0.$$

Thus we invoke the maximum principle once again. By similar arguments, in the cases of an interior minimum or a minimum at $\pm\infty$ that the estimate $h(y) \lesssim g(y)$ holds. If the infimum occurs at $\pm m$ then

$$g(y) - \frac{h(y)}{\delta} \geq \liminf_{x \to |m|} \left(g(x) - \frac{h(x)}{\delta}\right) = \begin{cases} g(m) - \dfrac{h(m)}{\delta}, & \text{as } x \to m, \\ g(-m) - \dfrac{h(-m)}{\delta}, & \text{as } x \to -m. \end{cases}$$

[7] Technically one would need to apply the maximum principle on the sets $(-\infty, -m)$ and (m, ∞) separately but the resulting arguments are identical.

Now fix δ large so that the right hand side is positive (recall that g is bounded strictly away from zero (Lemma 5.7.3)). This gives us the estimate $h(y) \lesssim g(y)$. With the assumed estimates on h we have now proven for all $y \in U(m)$,

$$\frac{1}{(1+|y|^q)} \lesssim g(y) \lesssim \frac{1}{(1+|y|^q)}.$$

The continuity and strict positivity of g implies we can trivially extend this estimate to the entirety of E in the case when $E = \mathbb{R}$. When $E = (0, \infty)$ we need to check that $\liminf_{y\to 0} g(y) > 0$ in order to extend our estimate to the whole of E. However this follows by Lemma 5.7.9, which concludes the proof. $\square$

Lemma 5.7.9. *Let g be the solution to the minimisation problem (5.3.21) when $E = \mathbb{R}_+$. Assume $\limsup_{y\to 0^+} \kappa(y) < +\infty$. Then $\liminf_{y\to 0} g(y) > 0$.*

Proof of Lemma 5.7.9. Fix $\varepsilon > 0$ and define $h(x) := \inf_{s\in(0,x]} \frac{1}{\kappa(s)}$. Our assumption that $\limsup_{y\to 0^+} \kappa(y) < +\infty$ ensures $h(x) > 0$ for all $0 < x << 1$. We proceed similarly as in the proof of Lemma 5.7.6 and identify three possible cases,

$$(i)\, g(\varepsilon) > h(\varepsilon); \qquad (ii)\, g(\varepsilon) = h(\varepsilon); \qquad (iii)\, (\varepsilon) < h(\varepsilon).$$

First consider case (ii), we show $g(y) \geq h(\varepsilon) > 0$ for all $y < \varepsilon$. Introduce,

$$\tilde{g}(y) = \begin{cases} g(y) \vee h(\varepsilon), & y < \varepsilon, \\ g(y), & y \geq \varepsilon, \end{cases}$$

where $x \vee y := \max(x, y)$ for $x, y \in \mathbb{R}$. Note that $g(\varepsilon) = h(\varepsilon)$ ensures $\tilde{g}$ is continuous. Assume there exists a set $E' \subseteq (0, \varepsilon]$ such that $\tilde{g}(y) = h(\varepsilon)$ for $y \in E'$. Thus,

$$I(g) - I(\tilde{g}) = \int_{E'} \left(\frac{a^2}{2}(g')^2 + \kappa(g - h(\varepsilon))(g + h(\varepsilon) - 2\kappa^{-1}) \right) \eta.$$

Note that for all $y \in E'$, $g(y) - h(\varepsilon) \leq 0$ and additionally,

$$g(y) \leq h(\varepsilon) \implies g(y) + h(\varepsilon) \leq 2h(\varepsilon) \implies g(y) + h(\varepsilon) - 2\kappa^{-1}(y) \leq 2h(\varepsilon) - 2\kappa^{-1}(y) \leq 0,$$

where the last inequality follows because $h(\varepsilon) = \inf_{s\in(0,\varepsilon]} \frac{1}{\kappa(s)}$. Thus $I(g) - I(\tilde{g}) \geq 0$, but this contradicts the optimality of g and hence $g(y) \geq h(\varepsilon) > 0$ for all $y < \varepsilon$.

Next consider case (i) $g(\varepsilon) > h(\varepsilon) > 0$. If $g(y) > h(y)$ for all $y \in (0, \varepsilon]$ then $\liminf_{y\to 0} g(y) > 0$, otherwise there must exist a point $\varepsilon_1 \in (0, \varepsilon]$ such that $g(\varepsilon_1) = h(\varepsilon_1) > h(\varepsilon) > 0$. Then the argument in case (i) yields that $g(y) > h(\varepsilon_1)$ for all $y < \varepsilon_1$, but this automatically yields $g(y) > h(\varepsilon)$ for all $y \leq \varepsilon$. Thus $\liminf_{y\to 0} g(y) > 0$.

Finally in case (iii) $0 < g(\varepsilon) < h(\varepsilon)$. It follows that $g(y) \geq g(\varepsilon) > 0$ for all $y \in (0, \varepsilon]$ as otherwise truncating g at $g(\varepsilon)$ would further reduce I yielding a contradiction of the optimality of g.

Thus in all cases we see $\liminf_{y \to 0} g(y) > 0$. $\square$

Lemma 5.7.10. *Let $g \in C^2(E; \mathbb{R})$ be the solution to the minimisation problem (5.3.21) and let Assumptions 5.5.1-5.5.4 hold. Then, for some $C > 0$*

$$|g'(y)| \leq \frac{C}{\inf_{w \in U(|y|)} |\tilde{b}(w)|},$$

for all $y \in U(m)$ for some large $m > 0$.

Proof of Lemma 5.7.10. Recall that g must satisfy the corresponding Euler-Lagrange equation, which is

$$\frac{d}{dy}\left(a^2(y)\eta(y)g'(y)\right) = 2\left(\kappa(y)g(y) - 1\right)\eta(y).$$

Integrating this expression over the interval $[x, y] \subset U(m) \cap \mathbb{R}_+$,

$$g'(y)a^2(y)\eta(y) - g'(x)a^2(x)\eta(x) = 2\int_x^y \left(\kappa(u)g(u) - 1\right)\eta(u)du.$$

By Assumptions 5.5.2 and 5.5.3, Lemma 5.7.7 implies

$$|\kappa(u)g(u)| \leq C\frac{(1 + |u|^p)}{(1 + |u|^q)}.$$

As $q \geq p > 0$, this quantity is uniformly bounded, thus this estimate yields

$$|g'(x)a^2(x)\eta(x)| \leq |g'(y)a^2(y)\eta(y)| + C\int_x^y \eta(u)du.$$

By Assumption 5.5.1 we invoke Lemma 5.5.1 yielding $|g'(y)a^2(y)\eta(y)| \to 0$ as $y \to \infty$. Thus taking limits we obtain,

$$|g'(x)a^2(x)\eta(x)| = C\int_x^\infty \eta(u)du.$$

Recalling the definition of η gives,

$$|g'(x)| \leq C\int_x^\infty \frac{1}{a^2(u)} \exp\left(2\int_x^u \frac{\tilde{b}(s)}{a^2(s)}ds\right)du.$$

Assumption 5.5.4 ensures that $|\tilde{b}(u)| \geq \inf_{w\in[x,\infty)} |\tilde{b}(w)| > 0$ for all $u \in [x, \infty)$. Thus,

$$|g'(x)| \leq C \int_x^\infty \frac{\tilde{b}(u)}{\inf_{w\in[x,\infty)} |\tilde{b}(w)| a^2(u)} \exp\left(2\int_x^u \frac{\tilde{b}(s)}{a^2(s)} ds\right) du.$$

The substitution $v = -2\int_x^u \frac{\tilde{b}(s)}{a^2(s)} ds$ gives[8],

$$|g'(x)| \leq \frac{C}{\inf_{w\in[x,\infty)} |\tilde{b}(w)|} \int_0^\infty e^{-v} ds \leq \frac{C}{\inf_{w\in[x,\infty)} |\tilde{b}(w)|}.$$

As x is arbitrary this holds for all points in the set $U(m) \cap \mathbb{R}_+$. The argument to show the estimate on $U(m) \cap \mathbb{R}_-$ is identical except first integrate over $[-y, x]$ at the beginning. □

Proof of Theorem 5.5.1. By Lemma 5.7.7 there exists two constants $C_1, C_2 > 0$ such that

$$\frac{C_1}{(1 + |y|^q)} \leq g(y) \leq \frac{C_2}{(1 + |y|^q)}.$$

By Lemma 5.7.10 there exists two constants $C_3, m > 0$ such that,

$$|g'(y)| \leq \frac{C_3}{\inf_{w\in U(|y|)} |\tilde{b}(w)|} \quad \text{for } y \in U(m).$$

For $E = \mathbb{R}$, the continuity of g' on $[-m, m]$ implies that, combining these estimates one finds some constant $C > 0$ such that

$$\left|\frac{g'(y)}{g(y)}\right| \leq C\left(1 + \frac{(1 + |y|^q)}{\inf_{w\in U(|y|)} |\tilde{b}(w)|} \chi_{\{y\in U(m)\}}\right) \qquad y \in E.$$

If $E = \mathbb{R}_+$, one needs to exclude the possibility that $\limsup_{y\to 0} |g'(y)| = \infty$. The proof of Lemma 5.7.10 to yields the estimate (for some $C > 0$),

$$|g'(x)| \leq \frac{C}{a^2(x)\eta(x)} \int_x^\infty \eta(u) du \leq \frac{C}{a^2(x)\eta(x)}.$$

Plugging in the definition of η (assuming $x < x_0$) it follows that

$$|g'(x)| \leq C \exp\left(2\int_x^{x_0} \frac{\tilde{b}(u)}{a^2(u)} du\right)$$

which has a finite lim sup by assumption. Hence, $\sup_{y\in(0,m]} |g'(y)| < +\infty$ which means that also for $E = \mathbb{R}_+$ the estimates for g and g are combined to obtain the desired result. □

[8]Note this change of variables is valid because $\inf_{w\in[x,\infty)} |\tilde{b}(w)| > 0$, this ensures $\tilde{b}(x)$ does not change sign for large x so the integral is a monotonic function of the upper limit of integration.

Proof of Proposition 5.5.1. The proof follows in two steps, first we check Assumptions 5.3.1-5.3.4 in order to invoke Theorem 5.4.1. Now (5.5.7) ensures Assumption 5.3.1 holds while Assumption 5.3.2 is true via a standard Girsanov argument. As $\gamma > 1$, Assumption 5.3.3 is true by definition and (5.5.8) implies Assumption 5.3.4. Hence Assumptions 5.3.1-5.3.4 are fulfilled and we can apply Theorem 5.4.1.

Step two involves checking there exists a unique solution to the martingale problem in Theorem 5.4.1. Standard existence and uniqueness results (such as Theorem 10.2.2 in Stroock and Varadhan [118]) requires to show that,

$$\|\hat{A}(z)\| \lesssim (1+|z|^2), \quad z \in \mathbb{R}^n \times \mathbb{R}$$
$$\langle z, \hat{b}(z)\rangle \lesssim (1+|z|^2), \quad z \in \mathbb{R}^n \times \mathbb{R},$$

where $\langle \cdot,\cdot\rangle$ denotes the standard inner product on $\mathbb{R}^n \times \mathbb{R}$. Note that in this case $\hat{A}(z)$ is constant, so the first condition is automatically satisfied. For the second condition recall that $\hat{b}$ depends only on the final coordinate, hence we need only check

$$|\hat{b}_{n+1}(y)| \lesssim (1+|y|), \quad y \in \mathbb{R}.$$

Recall

$$\hat{b}(y) = \begin{pmatrix} \frac{\mu(y)}{\gamma} + \Upsilon \frac{g'(y)}{g(y)} \\ \tilde{b}(y) + a^2 \frac{g'(y)}{g(y)} \end{pmatrix}.$$

Because $\tilde{b}$ is globally Lipschitz, it obeys a linear growth bound, while μ is itself linear. Thus if we can show $|g'/g|$ obeys a linear growth bound then the proof will be complete.

We obtain this bound by checking Assumptions 5.5.1-5.5.4 and invoking Theorem 5.5.1. Now (5.5.8) ensures $\kappa\eta$ will be uniformly bounded. As a is constant in this model $a^2\eta$, will certainly be integrable and so Assumption 5.5.1 is satisfied. From the definition of κ it directly follows that we can find two constants $C_1, C_2 > 0$ such that,

$$C_1(1+y^2) \le \kappa(y) \le C_2(1+y^2), \quad y \in \mathbb{R}.$$

Now let ψ be as constructed in the Appendix (A.2.1) (with obvious adjustments made when the support is shifted to the negative half plane). Thus for all $x > 0$ sufficiently large,

$$\begin{aligned}\int_x^{2x} |[a^2\psi(y)]''| + |[\tilde{b}(y)\psi(y)]'|dy &\le \int_x^{2x} a^2|\psi''(y)| + |\tilde{b}'(y)\psi(y)| + |\tilde{b}(y)\psi'(y)|dy \\ &\le C\left(\frac{1}{x} + 1 + \frac{1+x}{x}\right) \\ &\le C(1+x^2),\end{aligned}$$

for some $C > 0$. The second inequality comes from the estimates for ψ and its derivatives along with b being continuously differentiable and Lipschitz continuous which implies b (and therefore $\tilde{b}$) has a bounded derivative and also obeys a global linear growth bound. Hence Assumption 5.5.2 is satisfied with $p = 2$. Now recall $\kappa^{-1}(y) = 1/(\kappa_1 + \kappa_2 y^2)$ where $\kappa_1 := \beta/\gamma + (1 - 1/\gamma)r$ and $\kappa_2 := (1 - 1/\gamma)(\frac{\mu^\top \Sigma^{-1}\mu}{2\gamma})$. Thus we set $h(y) = \kappa^{-1}(y)$ and apply $\mathcal{L}$,

$$\begin{aligned}\mathcal{L}\left(\frac{1}{\kappa}\right) &= \frac{a^2}{2}(\kappa^{-1}(y))'' + \tilde{b}(y)(\kappa^{-1}(y))' - 1 \\ &= \frac{a^2\kappa_2\left(3\kappa_2 - \kappa_1\right)}{\left(\kappa_1 + \kappa_2 y^2\right)^3} - \frac{2\kappa_2\tilde{b}(y)y}{\left(\kappa_1 + \kappa_2 y^2\right)^2} - 1.\end{aligned}$$

Now $\tilde{b}$ inherits the uniform Lipschitz continuity (and therefore a linear growth estimate) from b and thus for large enough $|y|$ the first two terms can be made arbitrarily small and thus one can easily find two constants C_1, C_2 such that,

$$-C_1 \le \mathcal{L}\left(\frac{1}{\kappa}\right) \le -C_2$$

for y outside some large compact set containing the origin. Thus Assumption 5.5.3 is satisfied. Now (5.5.9) is exactly Assumption 5.5.4 thus we can invoke Theorem 5.5.1 to obtain,

$$\left|\frac{g'(y)}{g(y)}\right| \le C\left(1 + \frac{1 + |y|^2}{\inf_{w\in U(|y|)}|\tilde{b}(w)|}\chi(y)_{\{y\in U(x)\}}\right) \quad y \in E,$$

for $x > 0$ sufficiently large, where χ denotes the indicator function. Finally the second part of (5.5.9) implies a linear growth bound for the right hand side for all $|y| > x$ which, by the continuity of g'/g, can extended to all of E.

□

Appendix A

A.1 Numerics

We numerically solve the minimisation problem (5.3.21) via direct discrete approximation and employing finite dimensional optimisation techniques. Note minimising (5.3.25) is equivalent to minimising,

$$\tilde{I}(u) = \int_E \left(\frac{a^2(y)(u'(y))^2}{2} + \kappa(y) \left(u(y) - \frac{1}{\kappa(y)} \right)^2 \right) \eta(y)dy.$$

Let $\bar{E} \subset E$ be compact and Π an equidistant partition of $\bar{E}$. Then

$$\begin{aligned}
\tilde{I}(u) &\approx \int_{\bar{E}} \left(\frac{a^2(y)(u'(y))^2}{2} + \kappa(y) \left(u(y) - \frac{1}{\kappa(y)} \right)^2 \right) \eta(y)dy \\
&\approx \sum_{y_j \in \Pi} \left(\frac{a^2(y_j)(u_{j+1} - u_j)^2}{2h^2} + \kappa(y_j) \left(u_j - \frac{1}{\kappa(y_j)} \right)^2 \right) \int_{y_j}^{y_{j+1}} \eta(x)dx =: F(u_1, \ldots, u_N),
\end{aligned}$$

where we have approximated the derivative by the slope. Employing the L-BFGS optimisation algorithm (see [77]) to minimise $F : \mathbb{R}^N \to \mathbb{R}_+$ yields an approximate numerical solution. To test the effectiveness of this scheme we consider the following model (which is a special case of model (5.4.4) with $\delta = 1$),

$$\begin{aligned}
r(Y_t) = r; \quad dR_t &= \mu Y_t \, dt + \sigma \sqrt{Y_t} \, dZ_t, \\
dY_t &= b(\theta - Y_t) \, dt + a\sqrt{Y_t} \, dW_t,
\end{aligned} \tag{A.1.1}$$

where $r, \mu, \sigma, b, \theta, a > 0$ are constants. In this regime,

$$\tilde{b}(y) = b(\theta - y) - \lambda y \qquad \kappa(y) = \frac{\beta}{\gamma} + \left(1 - \frac{1}{\gamma}\right) \left(r + y\frac{\mu^2}{2\gamma\sigma^2}\right),$$

where $\lambda := \left(1 - \frac{1}{\gamma}\right) \frac{a\rho\mu}{\sigma}$ and

$$\eta(y) = \frac{C}{a^2} y^{\frac{2b\theta}{a^2} - 1} \exp\left(-\frac{2by}{a^2} - \frac{2\lambda y}{a^2}\right),$$

where $C > 0$ is a normalising constant. In this model an explicit solution to the HJB equation was given in [56, Lemma 5.1] and hence we compare the numerical approximation of the optimal consumption with the known explicit solution.

Parameter	r	μ	σ	b	θ	a	β	ρ
Value	0.01	1	1	0.07	0.02	0.04	0.02	1

Table 4: Parameters for model (A.1.1).

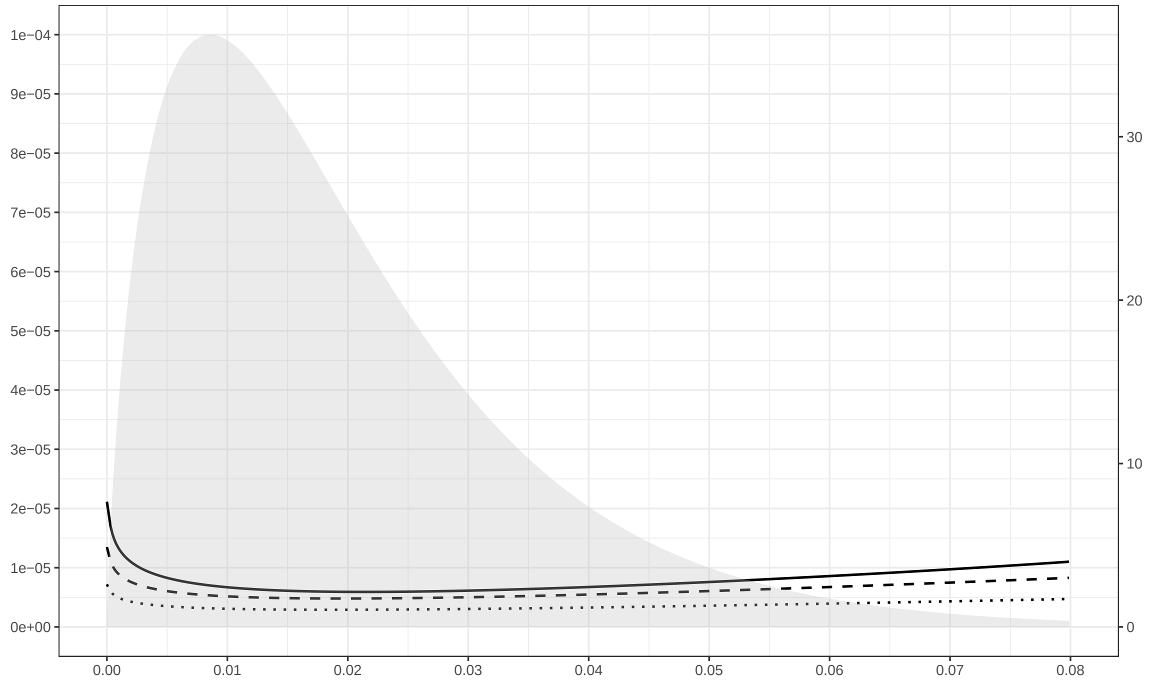


Figure 4: The difference between the exact consumption and the numerical approximation (left vertical axis) against the state variable (horizontal axis, with $\mathbb{P}(0 \leq Y_\infty \leq 0.8) \approx 0.995$) with the stationary density of Y (right vertical axis), for $\gamma = 2$ (solid), $\gamma = 4$ (dashed), and $\gamma = 8$ (dotted).

The naive numerical scheme performs well, as illustrated by the error graph in figure 4, which shows a maximum error of approximately 0.002% within a 99.5% confidence interval for the stationary distribution of the state variable. We note that in some instances a reparametrisation can increase both the speed and accuracy of the numerical scheme.

Rewrite

$$\begin{aligned}\tilde{I}(u) &= \int_E \left(\frac{a^2(y)(u'(y))^2}{2} + \kappa(y) \left(u(y) - \frac{1}{\kappa(y)} \right)^2 \right) \eta(y) dy \\ &= \int_E \left(\frac{a^2(y)(u'(y)\sqrt{\eta(y)})^2}{2} + \kappa(y) \left(u(y)\sqrt{\eta(y)} - \frac{\sqrt{\eta(y)}}{\kappa(y)} \right)^2 \right) dy.\end{aligned}$$

Then set $h = u\sqrt{\eta}$ and minimise the following functional,

$$h \mapsto \int_E \left(\frac{a^2(y)}{2} \left(h'(y) - h(y)\frac{\eta'(y)}{2\eta(y)} \right)^2 + \kappa(y) \left(h(y) - \frac{\sqrt{\eta(y)}}{\kappa(y)} \right)^2 \right) dy.$$

We discretise as above, conduct the minimisation, and then undo the parametrisation to recover the solution to the original optimisation problem. This transformation has the advantage of removing the scaling incurred by the weight function η so that gradient descent based optimisation methods are not affected by near zero gradients induced by the damping from η. This reparametrisation is employed for model (5.4.3).

A.2 Bump function construction

Let $x > 0$ and define following function,

$$\psi : \mathbb{R} \to \mathbb{R} : y \mapsto \begin{cases} \frac{32}{x^2}(y-x)^2, & y \in \left[x, \frac{9x}{8}\right] \\ \frac{-32}{x^2}\left(y - \frac{10x}{8}\right)^2 + 1, & y \in \left[\frac{9x}{8}, \frac{10x}{8}\right] \\ 1, & y \in \left[\frac{10x}{8}, \frac{14x}{8}\right] \\ \frac{-32}{x^2}\left(y - \frac{14x}{8}\right)^2 + 1, & y \in \left[\frac{14x}{8}, \frac{15x}{8}\right] \\ \frac{32}{x^2}(y-2x)^2, & y \in \left[\frac{15x}{8}, 2x\right] \\ 0, & y \notin [x, 2x]. \end{cases} \tag{A.2.1}$$

It is easy to see $\psi \in W^{2,1}(\mathbb{R};\mathbb{R})$. The first derivative is defined classically however the second derivative is understood in the weak sense. It is not hard to obtain the estimates,

$$|\psi(y)| \leq 1; \qquad |\psi'(y)| \leq \frac{C}{|x|}; \qquad |\psi''(y)| \leq \frac{C}{x^2},$$

for some $C > 0$ with $y \in [x, 2x]$. Additionally one can show $\Phi(x) = \int_x^{2x} \psi(y)dy = \frac{3x}{4}$.

Bibliography

[1] E. Abi Jaber. *Volterra processes in finance.* Habilitation à diriger des recherches, Ecole polytechnique, March 2024.

[2] G. Alsmeyer. *Erneuerungstheorie. Analyse stochastischer Regenerationsschemata.* Teubner Skripten zur Mathematischen Stochastik. Stuttgart: B. G. Teubner, 1991.

[3] J. A. D. Appleby. Almost sure stability of linear Itô-Volterra equations with damped stochastic perturbations. *Electron. Comm. Probab.*, 7:223–234, 2002.

[4] J. A. D. Appleby. Subexponential solutions of linear Itô-Volterra equations with a damped perturbation. *Funct. Differ. Equ.*, 11(1-2):5–10, 2004.

[5] J. A. D. Appleby. Mean square characterisation of a stochastic Volterra integrodifferential equation with delay. *Int. J. Dyn. Syst. Differ. Equ.*, 11(3-4):194–226, 2021.

[6] J. A. D. Appleby, J. Cheng, and A. Rodkina. Characterisation of the asymptotic behaviour of scalar linear differential equations with respect to a fading stochastic perturbation. *Discrete Contin. Dyn. Syst.*, pages 79–90, 2011.

[7] J. A. D. Appleby, S. Devin, and D. W. Reynolds. Mean square convergence of solutions of linear stochastic Volterra equations to non-equilibrium limits. *Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal.*, 13B:515–534, 2006.

[8] J. A. D. Appleby, S. Devin, and D. W. Reynolds. Almost sure convergence of solutions of linear stochastic Volterra equations to nonequilibria limits. *J. Integral Equations Appl.*, 19(4):405–437, 2007.

[9] J. A. D. Appleby and A. Freeman. Exponential asymptotic stability of linear Itô-Volterra equations with damped stochastic perturbations. *Electron. J. Probab.*, 8:no. 22, 22, 2003.

[10] J. A. D. Appleby and E. Lawless. Mean square asymptotic stability characterisation of perturbed linear stochastic functional differential equations. *Applied Numerical Mathematics*, 2023.

[11] J. A. D. Appleby and E. Lawless. On the dynamics and asymptotic behaviour of the mean square of scalar linear stochastic difference equations. In *Advances in discrete dynamical systems, difference equations and applications*, volume 416 of *Springer Proc. Math. Stat.*, pages 25–60. Springer, Cham, 2023.

[12] J. A. D. Appleby and E. Lawless. Solution space characterisation of perturbed linear volterra integrodifferential convolution equations: The L^p case. *Applied Mathematics Letters*, 146:108825, 2023.

[13] J. A. D. Appleby and E. Lawless. Characterisation of asymptotic behaviour of perturbed deterministic and stochastic pantograph equations. *arxiv*, 2024.

[14] J. A. D. Appleby and E. Lawless. Solution space characterisation of perturbed linear discrete and continuous stochastic volterra convolution equations: the ℓ^p and L^p cases. *arxiv*, 2024.

[15] J. A. D. Appleby and E. Lawless. Solution space characterisation of perturbed linear functional and integrodifferential volterra convolution equations: Cesàro limits. *arxiv*, 2024.

[16] J. A. D. Appleby, X. Mao, and M. Riedle. Geometric Brownian motion with delay: mean square characterisation. *Proc. Amer. Math. Soc.*, 137(1):339–348, 2009.

[17] J. A. D. Appleby and D. D. Patterson. On the admissibility of unboundedness properties of forced deterministic and stochastic sublinear Volterra summation equations. *Electron. J. Qual. Theory Differ. Equ.*, pages Paper No. 63, 44, 2016.

[18] J. A. D. Appleby and D. D. Patterson. Large fluctuations and growth rates of linear volterra summation equations. *Differential Equations and Applications*, 23(6):1047–1080, 2017.

[19] J. A. D. Appleby and M. Riedle. Almost sure asymptotic stability of stochastic Volterra integro-differential equations with fading perturbations. *Stoch. Anal. Appl.*, 24(4):813–826, 2006.

[20] J. A. D. Appleby, M. Riedle, and A. Rodkina. On asymptotic stability of linear stochastic volterra difference equations with respect to a fading perturbation. *Advanced Studies in Pure Mathematics*, 53:271–282, 2009.

[21] V. E. Beneš. A nonlinear integral equation in the Marcinkiewicz space m_2. *J. Math. and Phys.*, 44:24–35, 1965.

[22] A. Bensoussan. On the theory of option pricing. *Acta Appl. Math.*, 2(2):139–158, 1984.

[23] F. E. Benth, K. H. Karlsen, and K. Reikvam. Merton's portfolio optimization problem in a black and scholes market with non-gaussian stochastic volatility of ornstein-uhlenbeck type. *Mathematical Finance*, 13(2):215–244, 2003.

[24] M. A. Berger and V. J. Mizel. Volterra equations with Itô integrals. I. *J. Integral Equations*, 2(3):187–245, 1980.

[25] M. A. Berger and V. J. Mizel. Volterra equations with Itô integrals. II. *J. Integral Equations*, 2(4):319–337, 1980.

[26] G. Berkolaiko and A. Rodkina. Almost sure convergence of solutions to nonhomogeneous stochastic difference equation. *J. Difference Equ. Appl.*, 12(6):535–553, 2006.

[27] J. P. Bertrandias. Espaces de fonctions bornées et continues en moyenne asymptotique d'ordre p. *Bull. Soc. Math. France Mém.*, 5:106, 1966.

[28] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.

[29] I. V. Boykov, V. A. Roudnev, and A. I. Boykova. Stability of solutions of systems of Volterra integral equations. *Appl. Math. Comput.*, 475:Paper No. 128728, 2024.

[30] H. Brunner. *Volterra integral equations*, volume 30 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2017. An introduction to theory and applications.

[31] E. Buckwar and G. Notarangelo. A note on the analysis of asymptotic mean–square stability properties for systems of linear stochastic delay differential equation. *Discrete and Continuous Dynamical Systems Series B*, 18(6):1521–1531, 2013.

[32] T. Chan and D. Williams. An "excursion" approach to an annealing problem. *Math. Proc. Cambridge Philos. Soc.*, 105(1):169–176, 1989.

[33] H. Chang and K. Chang. Optimal consumption–investment strategy under the vasicek model: Hara utility and legendre transform. *Insurance: Mathematics and Economics*, 72:215–227, 2017.

[34] H. Chang and X. Rong. An investment and consumption problem with cir interest rate and stochastic volatility. *Abstract and Applied Analysis*, 2013(1):219397, 2013.

[35] A. S. Cherny and H. Engelbert. *Singular stochastic differential equations*, volume 1858 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2005.

[36] S. N. Cohen and R. J. Elliott. *Stochastic calculus and applications*. Probability and its Applications. Springer, Cham, second edition, 2015.

[37] C. Corduneanu. *Integral Equations and Stability of Feedback Systems*. Academic Press, New York, 1973.

[38] C. Corduneanu. *Integral equations and applications*. Cambridge Univ. Press, New York, 1990.

[39] C. Corduneanu. *Almost periodic oscillations and waves*. Springer, New York, 2009.

[40] J. C Cox and C. Huang. Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49(1):33–83, 1989.

[41] J. M. Cushing. *Integrodifferential Equations and Delay Models in Population Dynamics*. Lecture Notes in Biomathematics (LNBM, volume 20). Springer, Berlin, 1977.

[42] F. Delbaen and W. Schachermayer. *The mathematics of arbitrage*. Springer Finance. Springer-Verlag, Berlin, 2006.

[43] J. Dibik, M. Ruzickova, E. Schmeidel, and M. Zbaszyniak. Weighted asymptotically periodic solutions of linear Volterra difference equations. *Abstr. Appl. Anal.*, pages Art. ID 370982, 14, 2011.

[44] O. Diekmann, S. A. van Gils, S. M. Verduyn Lunel, and H.-O. Walther. *Delay equations*, volume 110 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1995. Functional, complex, and nonlinear analysis.

[45] D. Duffie, W. Fleming, H. M. Soner, and T. Zariphopoulou. Hedging in incomplete markets with HARA utility. *J. Econom. Dynam. Control*, 21(4-5):753–782, 1997.

[46] N. T. Dung. Linear multifractional stochastic Volterra integro-differential equations. *Taiwanese J. Math.*, 17(1):333–350, 2013.

[47] N. T. Dung. Stochastic Volterra integro-differential equations driven by fractional Brownian motion in a Hilbert space. *Stochastics*, 87(1):142–159, 2015.

[48] N. T. Dung. The density of solutions to multifractional stochastic Volterra integro-differential equations. *Nonlinear Anal.*, 130:176–189, 2016.

[49] B. Farkas and S. Wegner. Variations on barbălat's lemma. *The American Mathematical Monthly*, 123(8):pp. 825–830, 2016.

[50] W. H. Fleming and D. Hernández-Hernández. An optimal consumption model with stochastic volatility. *Finance Stoch.*, 7(2):245–262, 2003.

[51] I.I. Ī. Gikhhman and A. V. Skorohod. *Stochastic differential equations*. Springer-Verlag, New York-Heidelberg, 1972.

[52] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*, volume 224. Springer, 1977.

[53] G. Gripenberg, S.-O. Londen, and O. Staffans. *Volterra Integral and Functional Equations*. Encyclopedia of Mathematics and it's Applications. Cambridge University Press, 1990.

[54] P. Guasoni and S. Robertson. Portfolios and risk premia for the long run. *Ann. Appl. Probab.*, 22(1):239–284, 2012.

[55] P. Guasoni and G. Wang. Consumption and investment with interest rate risk. *J. Math. Anal. Appl.*, 476(1):215–239, 2019.

[56] P. Guasoni and G. Wang. Consumption in incomplete markets. *Finance Stoch.*, 24(2):383–422, 2020.

[57] I. Győri and D. W. Reynolds. On asymptotically periodic solutions of linear discrete Volterra equations. *Fasc. Math.*, PLACEHOLDER(44):53–67, 2010.

[58] J. K. Hale and S. M. V. Lunel. *Introduction to Functional Differential Equations*. Applied Mathematical Sciences. Springer, Berlin, 1993.

[59] J. M. Harrison and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3):215–260, 1981.

[60] J. M. Harrison and S. R. Pliska. A stochastic calculus model of continuous trading: Complete markets. *Stochastic Processes and their Applications*, 15(3):313–316, 1983.

[61] U. G. Haussmann. Asymptotic stability of the linear itô equation in infinite dimensions. *Journal of Mathematical Analysis and Applications*, 65:219–235, 1978.

[62] H. He and N. D. Pearson. Consumption and portfolio policies with incomplete markets and short-sale constraints: the infinite-dimensional case. *J. Econom. Theory*, 54(2):259–304, 1991.

[63] M. Herdegen, D. Hobson, and J. Jerome. An elementary approach to the merton problem. *Mathematical Finance*, 31(4):1218–1239, 2021.

[64] A. Ichikawa. Stability of semilinear stochastic evolution equations. *Journal of Mathematical Analysis and Applications*, 90:12–44, 1982.

[65] G.S. Jordan. Asymptotic stability of a class of integrodifferential systems. *Journal of Differential Equations*, 31(3):359–365, 1979.

[66] J. Kallsen and J. Muhle-Karbe. Utility maximization in affine stochastic volatility models. *International Journal of Theoretical and Applied Finance*, 13(03):459–477, 2010.

[67] I. Karatzas, J. P. Lehoczky, S. P. Sethi, and Steven E. Shreve. Explicit solution of a general consumption/investment problem. *Math. Oper. Res.*, 11(2):261–294, 1986.

[68] I. Karatzas, J. P. Lehoczky, and S. E. Shreve. Optimal portfolio and consumption decisions for a "small investor" on a finite horizon. *SIAM Journal on Control and Optimization*, 25(6):1557–1586, 1987.

[69] I. Karatzas, J. P. Lehoczky, S. E. Shreve, and G. L. Xu. Martingale and duality methods for utility maximization in an incomplete market. *SIAM J. Control Optim.*, 29(3):702–730, 1991.

[70] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[71] T. S. Kim and E. Omberg. Dynamic nonmyopic portfolio behavior. *The Review of Financial Studies*, 9(1):141–161, 06 1996.

[72] V. B. Kolmanovskii and A. Myshkis. *Introduction to the theory and applications of functional-differential equations*, volume 463 of *Mathematics and its Applications*. Kluwer Academic Publishers, Dordrecht, 1999.

[73] V. B. Kolmanovskii and V. R. Nosov. *Stability of functional-differential equations*, volume 180 of *Mathematics in Science and Engineering*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, 1986.

[74] R. Korn and H. Kraft. A stochastic control approach to portfolio problems with stochastic interest rates. *SIAM Journal on Control and Optimization*, 40(4):1250–1269, 2002.

[75] H. Kraft. Optimal portfolios and Heston's stochastic volatility model: an explicit solution for power utility. *Quant. Finance*, 5(3):303–313, 2005.

[76] A. Kufner and B. Opic. How to define reasonably weighted Sobolev spaces. *Comment. Math. Univ. Carolin.*, 25(3):537–554, 1984.

[77] D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Math. Programming*, 45(3):503–528, 1989.

[78] J. Liu. Portfolio selection in stochastic environments. *The Review of Financial Studies*, 20(1):1–39, 2007.

[79] R. Liu and J. Muhle-Karbe. Portfolio choice with stochastic investment opportunities: a user's guide. *https://arxiv.org/abs/1311.1715*, 2013.

[80] M. C. Mackey and I. G. Nechaeva. Solution moment stability in stochastic differential delay equation. *Physical Review E*, 52(4):3366—3376, 1995.

[81] X. Mao. *Exponential stability of stochastic differential equations.* Pure and Applied Mathematics. Marcel Dekker, New York, 1994.

[82] X. Mao. Razumikhin-type theorems on exponential stability of stochastic stochastic functional-differential equations. *Stochastic Processes and Their Applications*, 65(2):233–250, 1996.

[83] X. Mao. Stability of stochastic integro-differential equations. *Stochastic Anal. Appl.*, 18(6):1005–1017, 2000.

[84] X. Mao. Almost sure exponential stability of delay equations with damped stochastic perturbation. *Stochastic Analysis and Applications*, 19(1):67–84, 2001.

[85] X. Mao. *Stochastic differential equations and applications.* Horwood Publishing Limited, Chichester, second edition, 2008.

[86] X. Mao and X. Liao. Almost sure exponential stability of neutral differential difference equations with damped stochastic perturbations. *Electronic Journal of Probability*, 1(8):1–16, 1996.

[87] X. Mao and M. Riedle. Mean square stability of stochastic Volterra integro-differential equations. *Systems Control Lett.*, 55(6):459–465, 2006.

[88] X. Mao and A. Shah. Exponential stability of stochastic delay differential equations. *Stochastics and Stochastics Reports*, 60(1):135–153, 1997.

[89] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[90] R. C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *The Review of Economics and Statistics*, 51(3):247–257, 1969.

[91] R. C. Merton. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3(4):373–413, 1971.

[92] R. C. Merton. An intertemporal capital asset pricing model. *Econometrica*, 41:867–887, 1973.

[93] R. K. Miller. *Nonlinear Volterra Integral Equations.* W.A. Benjamin, Menlo Park, California, 1971.

[94] V J. Mizel and V Trutzer. Stochastic hereditary equations: existence and asymptotic stability. *J. Integral Equations*, 7(1):1–72, 1984.

[95] S.-E. A. Mohammed. *Stochastic Functional Differential Equations.* volume 99 of Research Notes in Mathematics. Pitman, London, 1984.

[96] C. Munk and C. Sørensen. Optimal consumption and investment strategies with stochastic interest rates. *Journal of Banking & Finance*, 28(8):1987–2013, 2004.

[97] É. Pardoux and P. Protter. Stochastic Volterra equations with anticipating coefficients. *Ann. Probab.*, 18(4):1635–1655, 1990.

[98] H. Pham. *Continuous-time stochastic control and optimization with financial applications*, volume 61 of *Stochastic Modelling and Applied Probability.* Springer-Verlag, Berlin, 2009.

[99] P. Protter. Volterra equations driven by semimartingales. *Ann. Probab.*, 13(2):519–530, 1985.

[100] P. Protter. *Stochastic integration and differential equations*, volume 21 of *Applications of Mathematics (New York).* Springer-Verlag, Berlin, 1990. A new approach.

[101] J. Prüss. *Evolutionary integral equations and applications*, volume 87 of *Monographs in Mathematics.* Birkhäuser Verlag, Basel, 1993.

[102] V. F. Pulyaev. Admissibility of certain pairs of spaces with respect to linear Volterra integral equations. *Differentsialcprime nye Uravneniya*, 20(10):1800–1805, 1984.

[103] V. F. Pulyaev and Z. B. Tsalyuk. On the question of admissibility of certain pairs of spaces for linear operators and Volterra equations. *Differentsialcprime nye Uravneniya*, 19(4):684–692, 1983.

[104] Y. N. Raffoul. *Qualitative theory of Volterra difference equations.* Springer, Cham, 2018.

[105] M. Reiß, M. Riedle, and O. van Gaans. On emery's inequality and a variation-of-constants formula. *Stochastic Analysis and Applications*, 25(2):353–379, 2007.

[106] L. Reus and P. Castañeda. A near optimal portfolio rule for the life cycle merton problem. *Computational Economics*, 2024. Publisher Copyright: © The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024.

[107] L. C. G. Rogers. *Optimal investment.* SpringerBriefs in Quantitative Finance. Springer, Heidelberg, 2013.

[108] L. C. G. Rogers and D. Williams. *Diffusions, Markov processes, and martingales. Vol. 1.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Foundations, Reprint of the second (1994) edition.

[109] W. Rudin. *Real and Complex Analysis.* McGraw–Hill, New York, 3rd edition, 1986.

[110] L. Shaikhet. Necessary and sufficient conditions of asymptotic mean square stability for stochastic linear difference equations. *Appl. Math. Lett.*, 10(3):111–115, 1997.

[111] L. Shaikhet. *Lyapunov functionals and stability of stochastic difference equations.* Springer, London, 2011.

[112] L. Shaikhet. *Lyapunov functionals and stability of stochastic functional differential equations.* Springer, Cham, 2013.

[113] L. Shaikhet and J. A. Roberts. Asymptotic stability analysis of a stochastic Volterra integro-differential equation with fading memory. *Dyn. Contin. Discrete Impuls. Syst. Ser. B Appl. Algorithms*, 18(6):749–770, 2011.

[114] William F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442, 1964.

[115] A. N. Shiryaev. *Probability*, volume 95 of *Graduate Texts in Mathematics.* Springer-Verlag, New York, second edition, 1996.

[116] A. Strauss and J. A. Yorke. On asymptotically autonomous differential equations. *Mathematical Systems Theory*, 1:175–182, 1967.

[117] A. Strauss and J. A. Yorke. Perturbation theorems for ordinary differential equations. *Journal of Differential Equations*, 3:15–30, 1967.

[118] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, Berlin-New York, 1979.

[119] Z. B. Tsalyuk. The admissibility of certain pairs of spaces for integral operators and Volterra equations. *Differencialcprime nye Uravnenija*, 13(11):2096–2098, 2111–2112, 1977.

[120] M. Viana and J. M. Espinar. *Differential equations—a dynamical systems approach to theory and practice*, volume 212 of *Graduate Studies in Mathematics.* American Mathematical Society, 2021.

[121] J. Wachter. Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *Journal of Financial and Quantitative Analysis*, 37(1):63–91, 2002.

[122] H. Wang and Z. Hu. Optimal consumption and portfolio decision with stochastic covariance in incomplete markets. *Chaos, Solitons & Fractals*, 138:109775, 2020.

[123] Thaleia Zariphopoulou. A solution approach to valuation with unhedgeable risks. *Finance Stoch.*, 5(1):61–82, 2001.

[124] W. P. Ziemer. *Weakly differentiable functions*, volume 120 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989. Sobolev spaces and functions of bounded variation.