UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING

Electronics and Computer Science

Pathwise Derivatives Valuation as a Cauchy Problem on the Space of Monte Carlo Paths

by

Laurie M. Carver

Thesis for the degree of Doctor of Philosophy

September 2019

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING
Electronics and Computer Science

Doctor of Philosophy

PATHWISE DERIVATIVES VALUATION AS A CAUCHY PROBLEM ON THE SPACE OF MONTE CARLO PATHS

by Laurie M. Carver

It is well-known that, under classical assumptions, the arbitrage-free value of European options contracts in complete continuous Markovian models is given by the solution to a *Cauchy problem* on $\mathbb{R}^d$. Recent research has shown that similar results hold in a *path-dependent* context, whereby values solve an analogous Cauchy problem on the *non-separable* path space $\mathbb{D}^d$ of càdlàg paths, in an *almost sure* sense, that is – on an *implicit* subset with model probability one. This presents difficulties for numerical solution, as practitioners must work with real data time series and at most countably many operations. This thesis resolves this in a wide class of continuous path-dependent market models, by showing that in this context derivatives' valuation is equivalent to solving a Cauchy problem for *any* path in an *explicit* subset of a new Banach space $\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)$ of *Monte Carlo paths*, that is naturally adapted to practitioner intuition and numerical methods.

First, we develop a new framework for the pathwise analysis of market risk models. Based on a new notion of *pathwise variance* that generalises existing notions of *quadratic variation*, we construct the new Banach space $\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)$ and show how its geometry captures practitioner intuition about risk models' volatility. The paths of a wide class of risk models can then be constructed explicitly as a family of integral expressions parameterised by *noise* paths. Second, we show how the Dupire (2009) differential operators for path functionals fit within the general theory of differentiation, and have a rich calculus. We characterise *vertical* differentials as the generators of strongly continuous groups on certain spaces of path functionals, and show how this allows for chain and product rules, and analogues of Taylor's Theorem and smooth approximation. Similarly, the *horizontal* differential is characterised as the generator of a strongly continuous *semi-group* on spaces of *functional processes*. We prove a Fundamental Theorem of Calculus type result for this operator, and show how this framework can be applied to derivatives' risk sensitivities. Third, we develop a theory of pathwise no arbitrage valuation on general Monte Carlo paths. We prove a new *Itô formula* for functional processes generalising that of Föllmer (1981) and Cont and Fournié (2010) to continuous paths with possibly discontinuous pathwise variance, and models which are not necessarily semimartingales. We then derive a corresponding valuation equation and robustness property for hedging error.

Contents

Nomenclature

$\mathbb{X}$	a generic Banach space
$c(\mathbb{X})$	the Banach space of convergent sequences in $\mathbb{X}$
$cs(\mathbb{X})$	the Banach space of convergent series in $\mathbb{X}$
$\ell^p(\mathbb{X})$	the Banach space of p-summable sequences in $\mathbb{X}$
$\mathbb{S}(\mathbb{X})$	the vector space of $\mathbb{X}$-valued simple paths
$\mathbb{W}(\mathbb{X})$	the Banach space of $\mathbb{X}$-valued continuous paths
$\mathbb{D}(\mathbb{X})$	the Banach space of $\mathbb{X}$-valued càdlàg paths
$\mathbb{F}(\mathbb{X})$	the Banach space of $\mathbb{X}$-valued càglàd paths
$\mathbb{G}(\mathbb{X})$	the Banach space of $\mathbb{X}$-valued regulated paths
$\mathbb{G}_0(\mathbb{X})$	the Banach space of $\mathbb{X}$-valued null limit paths
$BV(\mathbb{R}^d)$	the Banach space of $\mathbb{R}^d$-valued paths of bounded variation
x_+, x_-	right-/left-limit paths of a path x
$\Delta_+ x, \Delta_- x, \Delta x$	right-/left-/total jump paths of a path x
$\alpha[u], \beta[u], \gamma[u]$	singular/right-/left-bump paths at u
$\mathfrak{p}$	a generic partition of $[0, T]$
$\mathcal{E}^{\mathfrak{p}}(x)$, $\mathcal{E}_r^{\mathfrak{p}}(x)$, $\mathcal{E}_l^{\mathfrak{p}}(x)$	Euler projection of x on $\mathfrak{p}$; right- and left-continuous versions
$\mathscr{P}$	a generic clock
$\mathrm{var}_{\mathscr{P},p}(x)$	pathwise p-variance of $\mathbb{R}^d$-valued path x along clock $\mathscr{P}$
$[x, y]_{\mathscr{P}}$	pathwise co-variance matrix of $\mathbb{R}^d$-valued paths x, y along $\mathscr{P}$
$\mathbb{M}^p_{\mathscr{P}}(\mathbb{R}^d)$	the Banach space of $\mathbb{R}^d$-valued p-Monte Carlo paths along $\mathscr{P}$
$\mathbb{M}^p(\mathbb{R}^d)$	the Banach space of $\mathbb{R}^d$-valued p-Monte Carlo paths
$\tau(x)$	trader's clock evaluated on path x
$\mathrm{Lip}(\mathbb{X})$, $\mathrm{Lip}_b(\mathbb{X})$	the Banach spaces of Lipschitz/bounded Lipschitz functionals on $\mathbb{X}$
$C_b(\mathbb{X})$, $\mathscr{U}_b(\mathbb{X})$	the Banach spaces of continuous/uniformly continuous, bounded functionals on $\mathbb{X}$
$B^i[t]$	the ith bump operator at t
$\partial^i_{x_t}$	the ith vertical differential at t operator on path functionals
$H(u)$	the horizontal translation by u operator
$\mathscr{D}$	the horizontal differential operator for functional processes
$\tilde{\pi}_t$	the stopping operator at t on paths
π, π_-	the non-anticipative, predictable projections on functional processes
$x \underset{t}{\oplus} y$	the concatenation of paths x and y at t

Acknowledgements

It is standard in the conventional format of theses to thank one's supervisors for their patience and advice. I have tended to read this kind of thing and think: *yada yada yada, whatever*. It is only when one actually tries to *do* this kind of thing that one realises the level of importance this kind of support has. So: *seriously*. The sheer scale of the patience and support shown me by Associate Professor Enrico Gerding, Professor Frank McGroarty, and Dr Ramin Okhrati is of such magnitude that I am afraid I am unable adequately to communicate it, or to thank them enough. The ideas of this thesis and their shortcomings are all mine, but the fact that there is one at all is in large part due to them.

The only people to whom I owe more are my ever-supportive and loving family. My children are simultaneously the greatest of my achievements and the most continually surprising of my adventures; whatever happens, I live for them. Above all, I owe everything to my wife Lorna – only she truly knows the darkness from which she rescued me. All of me is hers.

Thanks also to my parents and brothers, and the extended Carver and Golding clans. Shout outs: David Rieser, Owen Rump, Tim Ham, Jon Pendelton; all staff at *Risk* magazine, past and present, most of all Nick Dunbar, Nick Sawyer, Richard Lee, Mauro Cesa and Duncan Wood; Vladimir Piterbarg, for encouraging me to seek a PhD; the cities of Bristol, London, Los Angeles, Poole, and Bournemouth. RIP Robin Goodchild, Ernie Esser.

For Lorna, and our babies

Chapter 1

Introduction

"He believed in an infinite series of times, in a growing, dizzying net of divergent, convergent and parallel times which approached one another, forked, broke off, or ignored each other through the centuries, and embraced every possibility."

- Jorge Luis Borges, *The Garden of Forking Paths*

"The past has gotta stop! The future's gotta rock!"

- Leftfield feat. Afrika Bambaataa, *Afrika shox*

A *derivative security* is a contract committing its counterparties to make and receive specified payments *contingent on the future outcomes of other variables*, such as stock or commodity prices, interest rates, and so on. A large part of the financial sector is devoted to the buying, selling, design, and management of such contracts – as of end June 2018, the Bank for International Settlements estimates the gross market value of outstanding 'over-the-counter' (OTC) derivatives to be $ 11 trillion, on a notional value of $595 trillion (Bank for International Settlements (2018)). The dependence on *a priori* unknown future variables, and the consequent need to robustly model their potential values, makes derivatives contracts complicated.

The fundamental principle of derivatives valuation is the exclusion of *arbitrage* – that is, riskless profit opportunities. Under suitable technical assumptions this is equivalent to finding a probability measure on the relevant future market states under which the profit and loss of any implementable trading strategy is expected to be zero.

If we restrict attention to contracts depending only on values at expiry, and the market is assumed to evolve as a *Markov* diffusion – that is, with dynamics depending only on its instantaneous values, independent of its history – then stochastic calculus shows that there is a deep connection between the no arbitrage value under classical market assumptions and the solution of a *Cauchy problem* for a corresponding partial differential equation (PDE).

The more general path-dependent setting has until recently relied exclusively on purely *probabilistic* methods. The most important is known as *Monte Carlo simulation*, which generates piecewise constant scenario paths that uniformly approximate risk model paths under the *risk-neutral* measure, and whose variance converges in probability to the corresponding quadratic variation. The Strong Law of Large Numbers then says the corresponding average of the derivative payoff over these scenarios converges to the model's option price.

Recent work by Dupire (2009), Cont and Fournié (2010) and others, has derived *path-dependent* or *functional* partial differential equations (FPDEs) for valuing derivatives in general models, that directly parallel the classical Markov case. Solutions can be given using methods such as forward-backward stochastic differential equations (Cont et al. (2016), El Karoui et al.). An issue with both from the practitioner point of view is that they remain inherently probabilistic, and strictly speaking only apply to an implicit subset of the given path space of probability one. The underlying problem is that the space of paths used in these approaches includes paths with infinite variance as a dense subset – it is 'too big'.

This thesis formulates a new framework for derivatives valuation on an explicit concrete, practically relevant space of paths. We introduce new Banach spaces, $\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)$, of *Monte Carlo paths* for the pathwise analysis of market risk models, whose structure is defined in such a way as to reflect numerical approximation in both the uniform and quadratic variation sense. No arbitrage valuation of derivatives in sufficiently regular continuous market models can then be characterised as a Cauchy problem on $\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)$, rather than on $\mathbb{R}^d$.

The rest of this introductory chapter is structured as follows. In Section 1.1 we will expand on the above comments to further motivate the problem under consideration, and detail the standing assumptions we will be making. In Section 1.2 we will state the aims of this research and the contributions it makes to the existing literature. Finally, in Section 1.3 we give an overview of the chapters to come.

1.1 Motivation: arbitrage and path-dependence

To give the reader the flavour of the problem, and introduce some necessary nomenclature, it is perhaps best to begin with a simple example of a derivative contract.

A (European-style) *vanilla put option* on an *underlying* stock with *spot price* $S(t)$ at time t, entitles one counterparty (the option's buyer or *holder*, who is said to be *long* the option) to sell the stock to the other (its seller or *writer*, who is *short*) at a time in the future (its *maturity*) at a price (its *strike*) fixed at inception. If, at the maturity time T, the stock's price $S(T)$ is less than the strike k, the holder can receive $k - S(T)$ by buying it on the market for $S(T)$ and *exercising* the option to sell it to the writer, who must pay k. If not, the option *expires*, and neither party must do anything.

If we think of the evolution of the spot price over the course of the contract as a *path* or function $S : [0, T] \to \mathbb{R}_+$, then the option *payoff functional* is

$$\Phi_k^{\mathrm{Put}}(S) = (k - S(T))^+ \tag{1.1}$$

The option in effect functions as a kind of insurance for the holder against the stock's price dropping below the strike, provided by the writer, who therefore requires compensation – a price – to enter the contract. In this case the payoff functional depends only on the spot price at maturity, and so can be thought of as a function on $\mathbb{R}$, but this need not be the case.

For instance, an *up-and-out put* is a type of *barrier* option, that behaves like a put option at maturity *provided the spot price has remained below a particular level over the course of the contract*:

$$\Phi_{k,B}^{\mathrm{UOP}}(S) = \Phi_k^{\mathrm{Put}}(S)\, 1_{\{x \in \mathcal{X} \,:\, x < B 1_{[0,T]}\}}(S) \tag{1.2}$$

$$= \begin{cases} (k - S(T))^+ & \text{if } S(u) < B \text{ for all } u \in [0, T] \\ 0 & \text{otherwise} \end{cases} \tag{1.3}$$

where B is the barrier level, and $1_{\{x \in \mathcal{X} \,:\, x < B 1_{[0,T]}\}}$ is the indicator function of the set of paths in some specified class $\mathcal{X}$ whose values all lie below B.

In general, derivatives essentially just involve more complicated versions of this idea – an agreement to exchange different cashflows in various outcomes, possibly depending on the history of the underlying. There are many other path-dependent derivatives, such as *American* options, which can be exercised at any point prior to maturity, *Asian* options, which depend on spot price averages, and *variance swaps* which depend on realised variance. In principle, any path functional could serve as the payoff for a derivative contract.

Where derivatives fundamentally differ from other financial contracts – fire insurance, say – is that there is a market for trading in the underlying, and so the contract's cash flows can be *hedged* by an appropriate dynamic trading strategy. This means that rather than the actuarial approach of a premium calculated to offset the statistical average liability incurred, the contract's price is determined by the idea of *arbitrage*.

Suppose there was a hedging strategy whose value equalled the payoff of the contract in any 'reasonable' scenario, that is, that *replicated* the derivative. A trader that offered the contract at a different price to the cost of the replicating portfolio would leave themselves open to *arbitrage*, and quickly go out of business – another trader could secure riskless cash from them, by shorting the more expensive position and going long the cheaper, pocketing the difference and leaving a zero net position. The basic principle of a derivatives trader is to ensure that instruments are valued so as to exclude this possibility, known in the literature as *no-arbitrage* pricing.

The set of results known as the *Fundamental Theorem of Asset Pricing* essentially say that this *financial* property of a market model is captured in a *probabilistic* property of admitting a *risk neutral* probability measure, with the same null sets as the statistical measure, under which (discounted) asset prices behave as *martingales*, equal to a conditional expectation of their future values. It has been proved in many variants and contexts, from single period finite state models, through discrete time to continuous time semimartingale models (Delbaen and Schachermayer (2006)). Once this measure has been identified, the no arbitrage value of a payoff is simply its (discounted) expectation under this measure.

A popular way of calculating this expectation is known as *Monte Carlo simulation*, in which the expectation is approximated by a sample average of random variables with distribution converging to the model's, using the strong law of large numbers. A typical (continuous) model is an *Itô process* modelling the distribution of the risk factors under the risk-neutral measure. Practitioners generally think of these in the infinitesimal form

$$dS(t) = r(t, \omega)S(t)dt + \sigma(t, \omega)S(t)dW \tag{1.4}$$

where the variable S and the coefficients r, σ are random processes, and the factors $S(t)$ are to ensure positivity, as in an asset price process. The amount of variability of the model is encoded in the *quadratic variation*,

$$d[S] = (\sigma(t, \omega)S(t))^2 dt \tag{1.5}$$

which is arrived at via the 'multiplication table'

$$(dt)^2 = 0 \tag{1.6}$$

$$dtdW = 0 \tag{1.7}$$

$$(dW)^2 = dt \tag{1.8}$$

These random variables will typically be approximated using *piecewise constant* paths with specified increments, of the form

$$S(t_i) - S(t_{i-1}) = r(t_{i-1})S(t_{i-1})\delta t_i + \sigma(t_i)S(t_{i-1})Z(t_i) \tag{1.9}$$

with Z appropriately chosen (pseudo)random variables, that converge uniformly in time, and in variance, as the partitions $\{t_i\}$ get increasingly finer.

In the case of complete Markovian models, another approach originally due to Merton (1973) relates the no arbitrage valuation to the solution of a *Cauchy problem* on $\mathbb{R}_+$. For the Black-Scholes model

$$dS(t) = rS(t)dt + \sigma S(t)dW \tag{1.10}$$

where r, σ are constants, this takes the form

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} + r\big(s\frac{\partial V}{\partial s} - V\big) = 0 \tag{1.11}$$

$$V(T, s) = \varphi(s) \tag{1.12}$$

$$(t, s) \in [0, T] \times \mathbb{R}_+ \tag{1.13}$$

A similar result holds in the more general setting of a contract on d risky assets, each following Markovian models with spatial and time dependent parameters:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{d} \sigma^{ij}(t, s) s^i s^j \frac{\partial^2 V}{\partial s^i s^j} + r(t, s)\big(\sum_{i=1}^{d} s^i \frac{\partial V}{\partial s^i} - V\big) = 0 \tag{1.14}$$

$$V(T, s) = \varphi(s) \tag{1.15}$$

$$(t, s) \in [0, T] \times \mathbb{R}_+^d \tag{1.16}$$

This PDE approach is particularly popular because it uses sensitivities to variables to construct explicit hedging portfolios, via *delta hedging*, in which a stock position is given by the first-order differential with respect to the spot price (Taleb (1997)). Perhaps most importantly, the Black-Scholes hedging portfolio satisfies a *robustness property*. If a trader hedges with the Black-Scholes model, but the stock in fact follows equation 1.4, then the (discounted) profit and loss from the strategy is

$$\text{P\&L} = \frac{1}{2}\int_0^T \frac{\partial^2 V}{\partial s^2} S^2(t)(\sigma^2 - \sigma_r^2)dt \tag{1.17}$$

where σ_r is the *realised* quadratic variation from the process 1.4. In other words, the trader's book evolves according to the error in estimation of the realised volatility by a Black-Scholes *implied* volatility, and relatively easily calculable differentials of Black-Scholes values. The efficacy of this result means it has been called the 'Fundamental Theorem of Derivatives Trading', (Ellersgaard et al. (2017)). It shows that "the option trader's job is really about balancing realised against implied (or pricing) volatility", according to Andreasen (2003). Indeed, "it is hard to imagine that the derivatives industry could exist at all without some result of this kind", according to Davis (2010).

But many derivatives are *path-dependent*, and realistic stock price dynamics are, too (Voit (2001)). Recently, there has been a surge of interest in analogous results that hold in a *pathwise* sense, that is, on each path in some suitable state space of a model. In deep work, Dupire (2009), Cont and Fournié (2010), Cont et al. (2016) and others have developed a *functional Itô calculus*, including notions of *vertical* and *horizontal* differentiation operators for path functionals. In particular, Dupire (2009) derives a Black-Scholes type Cauchy problem for a general derivative,

$$\Delta_t f(t, s_t) + \frac{1}{2}\sigma^2 \Delta_{ss} f(t, s_t) - r(f - s(t)\Delta_s f)(t, s_t) = 0 \tag{1.18}$$

$$f(T, s_T) = \Phi(S_T) \tag{1.19}$$

$$\mathbb{P}[s = S(\omega)] = 1 \tag{1.20}$$

where Δ_x is the vertical and Δ_t is the horizontal differential. Note however, the *almost sure* nature of this claim – it holds only for paths in some (implicit) subset of probability one under the model. The underlying space used in this analysis is $\mathbb{D}^d$, the space of right-continuous paths with left limits, under the uniform norm. Unfortunately, from the point of view of a practitioner this is a somewhat unwieldy object. It is *non-separable*, that is, does not contain a countable dense subset, and so we cannot hope to arbitrarily approximate all such paths with one fixed numerical scheme. Further, it contains many paths without a meaningful *quadratic variation*, and so is not taylored to studying 'volatility'. It is, in a sense, 'too big'. As a result, the aforementioned Cauchy problem type results only hold *almost surely* with respect to some measure $\mathbb{P}$, which *implicitly* defines a set of model states rather than *explicitly*.

Which if any state space would work better? To a practitioner, ultimately all calculations are done in some, high, but finite, dimensional model. What matters is that the *outputs* from this model *converge* in the senses that matter, ideally in a way that allows for robustness to model risk as in 1.17. As discussed above, this means *uniformly* and in *variance*, and the approximations for doing so are the Monte Carlo paths.

This suggests a simple idea:

> *derivatives valuation is the solution of a Cauchy problem on the space of Monte Carlo paths*

This thesis is dedicated to exploring this approach to the valuation problem. The first obstacle is that no such Banach space of Monte Carlo paths has explicitly been studied in the literature to date. The main contribution of this thesis lies in defining and taking seriously such a space as a legitimate mathematical object in its own right, and derivatives valuation as an analytic problem on this space. It is the first work to explicitly introduce these spaces, and to formulate the problem purely in this setting.

1.1.1 A note on assumptions

A short note is in order on the technical asssumptions of this work. Throughout we will consider idealized financial markets obeying the following classical assumptions:

- Infinite divisibility – any asset can be traded in quantities of any real number
- Frictionless trading – there are no transaction costs, taxes, or restrictions on 'short' selling
- Riskless funding – the trader has access to a default-free money market account paying a continuously compounded rate of interest with no credit restrictions

These are modelling assumptions, and are of course not strictly true in the real world. The first is entirely uncontroversial, and essential for any reasonable analysis. For the second, there is a large and growing literature on financial markets with transaction costs (see Kabanov and Safarian (2009) and the references therein), which is an important subculture in mathematical finance. However, since the problem under consideration is unsolved even in the absence of trading restrictions, it is reasonable for the scope of this thesis to neglect them.

The third assumption has been much challenged in the literature in recent years, from papers such as Piterbarg (2010), Burgard and Kjaer (2011) and others. After much controversy (see Carver (2012)) it is now essentially accepted that derivatives valuation must incorporate the various funding costs of trading desks (see Bielecki and Rutkowski (2014) for a survey). Nevertheless, the mathematics of doing so is highly non-trivial, and the classical case remains the first-order benchmark approximation. Since the problem under discussion is unsolved in those circumstances, it is acceptable for the scope of this thesis to neglect funding costs.

1.2 Research aims and contributions

The objectives of this thesis are:

1. To develop a framework for the *pathwise* analysis of market risk models that is taylored to practitioners' intuition and numerical methods
2. To develop a calculus of derivatives' risk sensitivities, in the context of the functional analytic theory of differentiation and operator semigroups
3. To formulate derivatives valuation in continuous models as the solution of a Cauchy problem on an explicit set of paths

The main contributions of this thesis are:

1. In regards to Objective 1, we construct a new notion of *pathwise variance* and new Banach spaces of *Monte Carlo paths* for the pathwise analysis of market risk models, and formulate them as *pathwise* integrals on this space parameterised by noise paths (Chapter 3)
2. In regards to Objective 2, we develop a novel functional analytic approach to the calculus of derivatives' risk sensitivities, based on a characterisation of *vertical* and *horizontal* differentials as *generators* of *(semi)groups* of operators on appropriate Banach spaces of path functionals, and prove useful analogues of Taylor's Theorem, the chain and product rules, and smooth approximation (Chapter 4)
3. In regards to Objective 3, we prove a new *functional Itô formula* for functionals of paths with possibly discontinuous pathwise variance and use it to prove a Cauchy problem for no arbitrage derivatives valuation on an explicit set of model paths, and a corresponding hedging robustness property (Chapter 5)

In more detail, each Objective is tackled in a corresponding Chapter, each of which is based on an original idea of the author.

Chapter 3 describes new Banach spaces of Monte Carlo paths that are natural state spaces for the solutions of risk models, with topologies governed by a *pathwise* Itô isometry (Theorem 3.23). This then allows for a novel approach to analysing the structure of path spaces using a natural group action (Theorems 3.9, 3.20).

Chapter 4 is based on the observation that *the Dupire differentials are the generators of strongly continuous (semi)groups*, consistently embedding them into existing abstract theories of differentiation. This then allows for a powerful calculus including new product and chain rules (Propositions 4.10 and 4.11), a new analogue of Taylor's Theorem (Theorem 4.12), and new density theorems for functionals which are 'smooth' in this new sense (Lemma 4.13, Theorem 4.14).

Chapter 5 uses the framework of Chapters 3 and 4 to prove a new *Itô formula* for Monte Carlo paths with possibly discontinuous pathwise variance (Theorem 5.3). Using a Merton (1973) style argument we then derive the valuation formula and a robustness property for Black-Scholes hedging in the presence of singular volatility shifts.

1.3 Overview

The rest of this thesis is organised as follows. In Chapter 2 we will review the relevant literature. First, in Section 2.1 we review the two main approaches to no arbitrage pricing, via Monte Carlo simulation and the PDE approach in the Markovian context. In Section 2.2 we look at the more recent path-dependent theory, largely based on the *functional Itô calculus*. In Section 2.3 we summarise how this literature leads naturally to the idea of derivatives valuation as solving a Cauchy problem on Monte Carlo paths.

The next three chapters detail the principal contributions. In Chapter 3 we develop a methodology for the pathwise analysis of market risk models. In Section 3.1 we discuss the space of *regulated paths* as the 'natural' space for the study of quantities that can be uniformly approximated by financial time series, and introduce a novel approach to studying its structure using a natural symmetry group action. In Section 3.2 we introduce a notion of *pathwise variance* and use it to define a Banach space of *Monte Carlo paths*, and show how its structure reflects practitioner's intuition about volatility. In Section 3.3 we use integration along a sequence of partitions to show this space helps is connected with the larger path spaces, the probabilists' Skorokhod topology, and the theory of Riemann integration, and construct explicit sets of paths of full probability for an Itô process market model.

In Chapter 4 we introduce a new functional analytic approach to the greeks, based on differential operators on path-dependent functionals. In Section 4.1 we introduce some important spaces of functionals of Monte Carlo paths, and give financial examples. In Section 4.2, we introduce the family of *bumping* operators, a family of strongly continuous groups of operators on path functionals parameterised by time, and prove results on their calculus and approximation by *vertically smooth* functionals. In Section 4.3 the space of *non-anticipative functionals* is introduced as a quotient space of *functional processes*, in contrast to the approach of Cont and Fournié (2010) as functionals on a quotient space of paths. It is shown that in this context, the bumping generators coincide with the Dupire vertical differential, and this observation is applied to the specific cases of derivatives' *delta* and *gamma*. Finally, in Section 4.4 we introduce the *horizontal translation semigroup* on nonanticipative functionals, and show that its infinitesimal generator coincides with the horizontal differential of Dupire (2009).

In Chapter 5 we develop a purely pathwise derivatives trading methodology on Monte Carlo paths. In Section 5.1 we introduce trading strategies and their P&L as functional processes, using *trader's clocks* as pathwise substitutes for stopping times. In Section 5.2 we use the results of Chapter 4 to derive a more general functional Itô formula, that allows for a singular volatility change without a jump in the underlying. Finally, in Section 5.3 we prove a valuation equation for continuous model paths and a corresponding hedging robustness property.

Finally, in Chapter 6 I conclude by discussing the prospects for further research.

Chapter 2

Literature Review

In this Chapter we review the literature on derivatives valuation relevant to our problem. In Section 2.1 I give an overview of general no arbitrage valuation by so-called *risk-neutral expectations*, via *Monte Carlo simulation* and the PDE approach in *Markovian models*. In Section 2.2 we review pathwise valuation methodologies, the *functional Itô calculus* and path-dependent PDEs. Finally, in 2.3, we summarise how these three strands point to a gap in the literature for a pathwise methodology that is more amenable to practitioner intuition and numerical methods.

2.1 No arbitrage valuation and Cauchy problems on $\mathbb{R}^d$

No arbitrage valuation can be a complicated subject, but the main ideas can be illustrated in a simple one-period, finite state model (Bjork (2009), Cox et al. (1979)). Suppose a market opens at $t = 0$, and at $t = 1$ the world will be revealed to be in one of d states from a set $\Omega := \{\omega_1, \ldots, \omega_d\}$ with corresponding *objective* (also known as *physical*, *statistical*, or *real-world*) probabilities $\mathbb{P}_k := \mathbb{P}(\omega_k)$, $k = 1, \ldots, d$.

There are $a+1$ assets which can be traded in the market, whose future values can be modelled as random variables, i.e. functions of the state ω. A riskless *money market account* $\tilde{B}$ grows with interest rate r in every state of the world, i.e. $\tilde{B}_0 = 1$, $\tilde{B}_1 = (1+r)1_\Omega(\omega)$. There are also a risky assets with respective prices $\tilde{S}_0^i$ at $t = 0$, and which at $t = 1$ pay functions

$$\tilde{S}_1^i : \Omega \to (0, \infty) \tag{2.1}$$

for $i = 1, \ldots, a$. In any given state, the return on the stocks $\frac{\tilde{S}_1^i(\omega)}{\tilde{S}_0^i}$ may be greater or lower than $1 + r$, but we require the average return

$$\frac{\mathbb{E}[\tilde{S}_1^i]}{\tilde{S}_0^i} > 1 + r \tag{2.2}$$

for $i = 1, \ldots, a$, since otherwise the risky asset is of no worth to a risk averse investor. For convenience we normalize the assets by accounting for them in terms of the money market, and write these values without the tildes: $S_t := \frac{\tilde{B}_0}{\tilde{B}_t}\tilde{S}_t$; $B_t := 1_\Omega$, for $t = 0, 1$.

A portfolio of Δ^i units of the ith asset costs

$$c[\Delta] = \sum_{i=1}^{a} \Delta^i S_0^i \tag{2.3}$$

and at $t = 1$, has discounted payoff given by the function

$$\psi_\Delta(\omega) := \sum_{i=1}^{m} \Delta^i (S_1^i(\omega) - S_0^i) \tag{2.4}$$

An *arbitrage* in this model is such a portfolio with the property that

$$c[\Delta] < 0$$

$$\psi_\Delta(\omega_k) \geq 0, \;\; k = 1, \ldots, d \tag{2.5}$$

$$\psi_\Delta(\omega_{k^*}) > 0, \;\; \text{some } k^* \tag{2.6}$$

In economic terms, this constitutes a *money pump*: arbitrarily large profits can be extracted at zero risk. It is a basic modelling assumption that competitive markets exclude such possibilities; certainly a practicing trader must attempt to do so. In mathematical terms, this is a statement of linear algebra, which by Farkas's Lemma (Lang (2005)) is equivalent to the existence of a non-empty convex set $\mathcal{Q}$ of vectors $\mathbb{Q} \in \mathbb{R}^n$ such that

$$c[\Delta] = \sum_{k=1}^{d} \mathbb{Q}_k \psi_\Delta(\omega_k) \tag{2.7}$$

for any portfolio position. Geometrically, if there is no arbitrage then there exists a separating hyperplane H between the spot price vector S_0 and the set $\mathcal{K}$ of portfolio payoffs. The normal to H in the direction $S_0 - \mathcal{K}$ fufills 2.7, which is clearly seen to be a convex condition.

Since the money market allows us to replicate the constant payoff function 1_Ω at (discounted) cost $B(0) = 1$, we see that

$$1 = \sum_{k=1}^{d} \mathbb{Q}_k B(1)(\omega_k) = \sum_{k=1}^{d} \mathbb{Q}_k \tag{2.8}$$

for any $\mathbb{Q} \in \mathcal{Q}$, and hence the functions $\mathbb{Q} : \Omega \to \mathbb{R}_+$ are *probability measures*, known as *risk-neutral measures*. Since the cost of 'replicating' a risky asset is simply its current spot price, we have

$$\mathbb{E}^{\mathbb{Q}}[S_1 - S_0] = 0 \tag{2.9}$$

where $\mathbb{E}^{\mathbb{Q}}$ denotes expectation under the measure $\mathbb{Q}$. That is, the (discounted) asset price processes are *martingales* under the risk-neutral measures – "the mathematical expectation of the speculator is zero", as Bachelier (1900) put it in his thesis which founded the subject. Note that the only role played by the objective statistical measure $\mathbb{P}$ was to say which states 'counted'. If we had included another state ω' with probability $\mathbb{P}(\omega') = 0$ the same result would have been obtained; the risk neutral measures are *equivalent* to the statistical measure, in the sense of having the same *null sets*.

Suppose a trader then writes a derivative with (discounted) payoff $\varphi : \Omega \to \mathbb{R}$. What price should she charge her counterparty? She forms a hedging portfolio Δ at cost $c[\Delta]$ with payoff ψ_Δ. Setting the hedging error $\mathrm{Err}(\omega_k) := \psi(\omega_k) - \varphi(\omega_k)$ to be zero in each state ω_k gives d equations in the a unknowns Δ^i. If $a > d$, then the equation is overspecified, and without loss of generality we may express $a - d$ *redundant* assets as linear combinations of the others and rewrite any portfolio in terms of these assets, without affecting the payoffs. If $a = d$, then we may solve for a *unique* d-dimensional vector Δ^*_φ that renders the error zero; any payoff can thus be *replicated*, and the market is said to be *complete*. The set $\mathcal{Q} = \{\mathbb{Q}\}$ is therefore a singleton, with $\mathbb{Q}_k$ being the cost of the replicating strategy for the *Arrow security* 1_{ω_k} (Arrow (1964), Becherer and Davis (2010)).

The case $a < d$ is more complicated. In general the hedging error equation will not have a solution, and so some payoffs will not be replicable; the market is said to be *incomplete*. In general, there will be an *interval* of arbitrage-free prices for a derivative, and traders must choose between them based on their preferences. Methods include minimal cost superhedging, expected utility optimization, and the minimization of some risk measure of the hedging error (Karatzas et al. (1991), Davis (1997), Stoikov (2006))

These results are the simplest versions of the results known as the *Fundamental Theorems of Asset Pricing*. Roughly, the first says that a market is free of arbitrage if and only if there exists an equivalent martingale measure, and the second that if the market is complete this measure is unique (Delbaen and Schachermayer (2006), Harrison and Pliska (1981)). The above argument can be generalised to discrete time steps and a continuum of outcomes without much difficulty. The complications only really come in once we attempt to move to a continuous time model.

A standard model of continuous time finance, as for instance set out in Oksendal (1992), starts with a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is an *abstract* set, $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is a filtration of sigma-algebras modelling the flow of information over time, and $\mathbb{P}$ the (real-world) probability measure on this space under which $W(t)$ is a Brownian motion on $\mathbb{R}^d$. A *market* is then given as an $\mathcal{F}_t$-adapted Itô process of the form

$$dS^0(t) = r(t,\omega)S^0(t)dt \tag{2.10}$$

$$dS^i(t) = \mu^i(t,\omega)dt + \sum_{j=1}^{d} \sigma^{ij}(t,\omega)dW^j(t) \tag{2.11}$$

for $i = 1, \dots, m$, with $S^0(0) = 1, S^i(0) = s^i$, and W a d-dimensional Brownian motion. Positivity for asset prices can be ensured by setting the coefficients to be proportional to $S^i(t)$ and relabelling, i.e.: $\mu^i(t,\omega) \to \mu^i(t,\omega)S^i(t)$, $\sigma^i(t,\omega) \to \sigma^i(t,\omega)S^i(t)$. More generally, we can take the market to be a *semimartingale*, that is a process of the form $S(t) = A(t) + M(t)$, where A is an adapted process of finite variation and M a (local) martingale.

Key to the use of these models in practice is the concept of *quadratic variation*. Two $\mathbb{R}$-valued stochastic process X, Y have *quadratic co-variation process* $[X, Y]$ if for each $\varepsilon > 0$ there exists $\delta > 0$ such that if a partition $\mathfrak{p}$ has *mesh* $\sup |\mathfrak{p}_{k+1} - \mathfrak{p}_k| < \delta$, then

$$\mathbb{P}[|\sum_{k=0}^{m(\mathfrak{p})} \big(X(\omega,\mathfrak{p}_{k+1}) - X(\omega,\mathfrak{p}_k)\big)\big(Y(\omega,\mathfrak{p}_{k+1}) - Y(\omega,\mathfrak{p}_k)\big) - [X,Y](\omega,t)| \geq \varepsilon] < \varepsilon \tag{2.12}$$

The idea is that the discrete-time covariance computed with respect partitions converges in probability to the quadratic co-variation, in the limit as the partitions become 'infinitely finely grained' (Revuz and Yor (2004), Jacod and Shiryaev (2002)).

For a general $\mathbb{R}^d$-valued process we write $[X]^{ij} := [X^i, X^j]$ for the *co-variation matrix*. In stochastic calculus there is a simple rule for calculating $[S]$ for continuous semimartingales, and especially Itô processes S. Namely, the quadratic variation equals that of the (local) martingale part, and in the case of an Itô process, it is determined by the 'multiplication table'

$$(dt)^2 = 0 \tag{2.13}$$

$$dt dW^i = 0 \tag{2.14}$$

$$dW^i dW^j = I_{i,j} dt \tag{2.15}$$

where I is the $d \times d$ identity matrix. For a market given by a general Itô process S as above this gives

$$d[S,S](t) = \sum_{i,j=1}^{d} \sigma^{ij}(t,\omega)\sigma^{ji}(t,\omega)dt \tag{2.16}$$

The key step in continuous time derivatives valuation is in itroducing a *market price of risk*. If there is no arbitrage in the model 3.72, then by a similar, if slightly more technical, argument as in the simplest case there exists a process θ such that, in the notation of 2.10

$$\sigma\theta = \mu - rS, \tag{2.17}$$

see Theorem 12.2.8 of Oksendal (1992). If this so-called *market price of risk* process is such that $M(t) := \exp(-\int_0^t \theta(s)dW(s) - \frac{1}{2}\int_0^t |\theta(s)|^2 ds)$ is a $\mathbb{P}$-martingale (this holds, in particular, if it satisfies a mild integrability condition known as the *Novikov condition*, see Oksendal (1992) page 162), then *Girsanov's Theorem* ensures the existence of an equivalent probability measure $\mathbb{Q}$ for which the process $W_{\mathbb{Q}}(t) := W(t) + \int_0^t \theta du$ is a $\mathbb{Q}$-Brownian motion. Hence the modelling equations can be rewritten

$$dS^0(t) = r(t,\omega)S^0(t)dt \tag{2.18}$$

$$\begin{aligned} dS^i(t) &= r(t,\omega)S^i(t)dt + \sum_{j=1}^{m} \sigma^{ij}(t,\omega)(dW^j(t) + \theta^j dt) \\ &= r(t,\omega)S^i(t)dt + \sum_{j=1}^{d} \sigma^{ij}(t,\omega)dW^j_{\mathbb{Q}}(t) \end{aligned} \tag{2.19}$$

Discounting, and recognising the remainder as a stochastic integral and martingale, shows it is a risk-neutral measure. Note that in an incomplete market, $d < m$, the market price of risk will not be unique; in fact, it turns out that the market is complete, and the market price is unique if and only if σ has a left-inverse, i.e. $m = d$ (Karatzas et al. (1991)).

The problem then becomes: how can one calculate this risk-neutral expectation? A popular method is *Monte Carlo simulation* (Glasserman (2004)). In this set-up, the model 2.18 is approximated using *piecewise constant* paths with specified increments, of the form

$$S(t_i) - S(t_{i-1}) = r(t_{i-1})S(t_{i-1})\delta t_i + \sigma(t_i,\omega_i)Z(t_i,\omega_i) \tag{2.20}$$

with Z appropriately chosen (pseudo)random variables to approximate the (risk-neutral) Brownian motion, that converge uniformly in time, and in variance, as the partitions $\{t_i\}$ get increasingly finer. The strong law of large numbers then shows that the sample average

$$\frac{1}{n}\sum_{k=1}^{n} \Phi(S(\omega_k)) \tag{2.21}$$

converges to the expectation, and hence no arbitrage price. The use of Monte Carlo methods in option pricing is extensive (see for example, Broadie and Glasserman (1996), Boyle (1977), Longstaff and Schwartz (2001)).

In the case of a complete Markovian model, a fundamental relationship between expectations and PDEs means things are considerably simplifed. The classic argument is due to Merton (1973). Consider a market with a riskless cash account paying a continuously compounding rate r of interest, and a single risky stock with spot price obeying the stochastic differential equation (SDE)

$$S(t) = S(0) + \int_0^t \mu S(u)du + \int_0^t \sigma S(u)dW(u) \tag{2.22}$$

where the *yield* $\mu > r$, and *volatility* $\sigma > 0$ are constants, and W is a Brownian motion on $\mathbb{R}$.

Suppose a trader takes a short position in a European option paying the function $\varphi(S(T))$ at maturity T. She decides to value the option position using a (deterministic, suitably smooth) function $V(t,s)$ of time and spot only, so that she would pay $V(t,S(t))$ in cash at time t in exchange for being relieved of the position. She then hedges by taking a position in the stock, given by another such (deterministic, suitably smooth) function $\Delta(t,s)$, financed by borrowing from the cash account. Writing the value for the portfolio on this basis as $\Pi(t, s) := s\Delta(t, s) - V(t, s)$, the profit and loss over the interval $[t, t + \delta]$ is

$$\begin{aligned} \text{P\&L} &= \Pi(t+\delta, S(t+\delta)) - \Pi(t, S(t)) \\ &= \int_t^{t+\delta} \Delta(u, S(u)) dS(u) - (V(t+\delta, S(t+\delta)) - V(t, S(t)) \end{aligned} \tag{2.23}$$

where the first summand is the cumulative change in the hedge position, and the second the change in the (short) option position according to her valuation. Using equation 2.22 to expand dS and *Itô's lemma* to expand V then gives

$$\begin{aligned} \text{P\&L} = &\int_t^{t+\delta} \mu S(u)\big(\Delta(u, S(u)) - \frac{\partial V}{\partial s}(u, S(u))\big)\, du \\ &+ \int_t^{t+\delta} \sigma S(u)\big(\Delta(u, S(u)) - \frac{\partial V}{\partial s}(u, S(u))\big) dW(u) \\ &- \int_t^{t+\delta} \frac{\partial V}{\partial t}(u, S(u)) + \frac{1}{2}\sigma^2 S(u)^2 \frac{\partial^2 V}{\partial s^2}(u, S(u))\, du \end{aligned} \tag{2.24}$$

Choosing the function $\Delta(t, s) := \frac{\partial V}{\partial s}(t, s)$ renders the first two summands zero; the remaining summand is absolutely continuous, and so represents an infinitesimally riskless position, that grows proportionately to dt rather than dS. Alternatively, liquidating her holdings and putting the value $\Pi(u, S(u))$ in the cash account is a similarly infinitesimally riskless position yielding

$$r \int_t^{t+\delta} \big(\Delta(u, S(u))S(u) - V(u, S(u))\big) du \tag{2.25}$$

Choosing V such that their difference is anything other than zero would result in the position being vulnerable to arbitrage: a strategy of buying the cheaper of the two and shorting the dearer in sufficient scale yields profits proportional to the time elapsed. Since this must be true for any $\delta > 0$, and the spot price path has positive probability of visiting any region of $\mathbb{R}_+$, the function V must satisfy the following *Cauchy problem* for the *Black-Scholes* PDE:

$$\begin{aligned} \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 V}{\partial s^2} + rs\frac{\partial V}{\partial s} - rV &= 0 \\ V(T, s) &= \varphi(s) \end{aligned} \tag{2.26}$$

for $s \in \mathbb{R}_+$. If this problem is solvable, the solution must be the arbitrage-free price for the model 2.22. The only free parameter of the equation is the stock's volatility σ; it is independent of the stock's yield μ, while the interest rate r is of course exogenously observable in the market. PDE methods are widely used in a vast literature on derivatives pricing (Wilmott et al. (1998), Gatheral (2006))

However, the truly remarkable property of the Black-Scholes methodology as developed by Merton is its *robustness* to model risk (Davis (2010), Ellersgaard et al. (2017)). Suppose that a trader hedges using the Black-Scholes model with parameter σ, but rather than the SDE 2.22 the stock instead follows the general Itô process model of 2.10. The hedging P&L can be expanded again using Itô's formula as before, but because of equation 2.26 we may conclude that

$$\text{P\&L} = \frac{1}{2}\int_0^T \frac{\partial^2 V}{\partial s^2} S^2(t)(\sigma^2 - \sigma_r^2)dt \tag{2.27}$$

where σ_r is the *realised* quadratic variation from the process 2.10. For many options, in particular vanilla calls and puts, the second differential term is positive, and hence the trader makes a profit provided her *implied volatility* σ overestimates the realised.

The use of constant parameters in the above discussion was primarily for convenience; in fact the entire above analysis goes through if $\sigma(s)$ is a uniformly invertible $\mathbb{R}^{d\times d}$-valued, and $r = r(s)$, $\mu = \mu(s)$ are $\mathbb{R}^d$-valued, functions of time and spot (a *Markov* model) to yield the Cauchy problem

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sum_{i,j=1}^{d} \sigma^{ij}(t,s)\sigma^{ji}(t,s)s^i s^j \frac{\partial^2 V}{\partial s^i s^j} + \sum_{i=1}^{d} r(t,s)s^i \frac{\partial V}{\partial s^i} - r(t,s)V = 0$$
$$V(T,s) = \varphi(s) \tag{2.28}$$

for $s \in \mathbb{R}^d_+$, as long as the resulting SDE 2.22 has a well-defined solution that allows for the application of Itô's Lemma.

A modern treatment of the Black-Scholes PDE as the generator of a semigroup is given in Gonzalez-Gaxiola and Santiago (2012). While the Black-Scholes model allows for closed-form type solutions, a general Markov model will not. However operator semigroup methods are the next best thing, see Garman (1985), Barucci et al. (1999), Altomare and Attalienti (2002a), Altomare and Attalienti (2002b), Zhou and Wang (2009), Linetsky (2004), Linetsky (2007). More recently, authors have found applications of semigroups in concrete derivatives problems. For instance, Linetsky (2002) looks at their role in pricing exotics, and a perturbation approach was used for the particular case of a so-called Constant Elasticity of Variance model by Colombo et al. (2003). Other applications include volatility derivatives Albanese et al. (2009), barrier options Kato and Yamada (2014), and life insurance reserves Fahrenwaldt (2015).

2.2 Pathwise derivatives valuation

The literature on *pathwise* derivatives valuation really starts with the remarkable paper of Bick and Willinger (1994), which deserves to be more widely known. In it they showed that many of the advantages of the Black-Scholes PDE approach remained in the *non-probabilistic* setting of functions of particular *paths*, for which a particular notion of pathwise quadratic variation, due to Föllmer (1981), exists.

To understand Föllmer's quadratic variation, suppose $\mathcal{T}_n$ is a sequence of partitions of $[0, T]$ with *mesh* converging to zero. A path $x : [0, T] \to \mathbb{R}$ is said to have *finite quadratic variation in the Follmer sense* if the atomic measures

$$\xi_n := \sum_{t_i \in \mathcal{T}_n} (x(t_{i+1}) - x(t_i))^2 \delta_{t_i} \tag{2.29}$$

converge *weakly* to a Radon measure on $[0, T]$ with distribution function $[x]$, whose atomic part coincides with the square of the jumps of x. That is, if

$$\int_{[0,T]} f(u)\, \xi_n(du) \to \int_{[0,T]} f(u)\, [x](du) \tag{2.30}$$

for all continuous $f : [0, T] \to \mathbb{R}$, and

$$[x](t) = [x]_c(t) + \sum_{u \in [0,t]} (\Delta x(u))^2 \tag{2.31}$$

where $[x]_c$ is continuous. By explicitly considering the convergence terms in the expression, Föllmer (1981) proved the formula, analogous to the classical Itô formula,

$$\begin{aligned} F(x(t)) = F(x(0)) + \int_0^t F'(x(s-))dx(s) + \frac{1}{2} \int_0^t F''(x(s-)) \\ + \sum_{s \le t} \left(F(x(s)) - F(x(s-)) - F'(x(s-))\Delta x(s) - \frac{1}{2} F''(x(s-))(\Delta x(s))^2 \right. \end{aligned} \tag{2.32}$$

holds for such paths if F is a C^2 function.

Bick and Willinger (1994) applied this to paths with constant quadratic variation in this sense given by σ, and showed that for a call option, the Black-Scholes formula and corresponding hedging strategy replicated the option along this path, provided the stock did not finish at the strike. Remarkably, as noted by Mishura and Schied (2016), the point is that the key assumption of the Black-Scholes model that makes it work as it does is not the assumption of a lognormal distribution for the asset, but the *constant quadratic variation*.

The next major breakthrough was the seminal paper of Dupire (2009), which developed the first truly path-dependent Itô calculus. Motivated by the idea of derivatives prices as functions taking paths as arguments, but needing to account for the flow of information, he considered as a domain a *union*

$$\Lambda := \bigcup_{t \in [0,T]} \mathbb{D}_t \tag{2.33}$$

where $\mathbb{D}_t$ is the space of cádlág paths on the interval $[0, t]$. He then defined a *functional* as a function on this union, with a particular topology, and used these as the basic building blocks for valuation functionals.

He then introduced what he called *space* and *time* differential operators for these functionals, by intuitively perturbing the graphs. If $x_t \in \mathbb{D}_t$, then its vertical perturbation x_t^ε is

$$x_t^\varepsilon(u) := \begin{cases} x(u) & \text{if } u < t \\ x(t) + \varepsilon & \text{if } u = t \end{cases} \tag{2.34}$$

The vertical differential is then

$$\lim_{\varepsilon \to 0} \frac{f(x_t^\varepsilon) - f(x_t)}{\varepsilon} \tag{2.35}$$

when this exists. The horizontal perturbation $x_{t,h} \in \mathbb{D}_{t+h}$ is

$$x_{t,h}(u) := \begin{cases} x(u) & \text{if } u < t \\ x(t) & \text{if } t \le u \le t + h \end{cases} \tag{2.36}$$

The horizontal differential is similarly then

$$\lim_{h \downarrow 0} \frac{f(x_{t,h}) - f(x_t)}{h} \tag{2.37}$$

when this exists. Note that the time differential is *one-sided*, based on right-translations. He went on to prove a *functional Itô formula* for functionals evaluated on the paths of a continuous semimartingale X, of the form

$$f(x_T) = f(x_0) + \int_0^T \Delta_x f(x_t) dX(t) + \int_0^T \Delta_t f(x_t) dt + \frac{1}{2} \int_0^T \Delta_{xx} f(x_t) d\langle x_t \rangle \tag{2.38}$$

where $\Delta_x, \Delta_{xx}, \Delta_t$ denote the first- and second space derivative and the time derivative respectively, and $\langle . \rangle$ is the (ordinary, probabilist's) quadratic variation. He then used this to derive a *Feynman-Kac* formula for functionals and a *martingale representation formula* for writing conditional expectations as stochastic integrals that extended the *Clark-Ocone formula* of Malliavin calculus. Finally, he used Merton's argument to derive a *functional* valuation equation

$$\Delta_t V(t, s_t) + \frac{1}{2}\sigma^2 \Delta_{ss} V(t, s_t) - r(V - s(t)\Delta_s V)(t, s_t) = 0 \tag{2.39}$$

$$V(T) = \Phi \tag{2.40}$$

$$\mathbb{P}[s = S(\omega)] = 1 \tag{2.41}$$

on the paths s of a continuous semimartingale. This is of a similar form as the Black-Scholes equation 2.26, except everything is path-dependent, and Dupire applied it to the case of a (path-dependent) *Asian* option. Note however, the *almost sure* nature of this claim – it holds only for paths in some (implicit) subset of paths of probability one under the model.

Various technical aspects of Dupire's formalism were considered somewhat unwieldy, for instance the use of a union in 2.33. In a sequence of papers and the book Cont et al. (2016), Cont and co-authors have done the most work on formalising and expanding Dupire's ideas, as well as incorporating Föllmer's quadratic variation. Cont and Fournié (2010) and Cont and Fournié

(2013), brought in a formalism based on what they called *non-anticipative functionals*, functions $f([x_t]) = f(t, x)$ defined on *equivalence classes of paths* that agree up to some t, rather than a union of path spaces. Consequently their *vertical* and *horizontal* differentials have a slightly different character, for instance, the vertical differential is then

$$\lim_{\varepsilon \to 0} \frac{f(t, x + \varepsilon 1_{[t,T]}) - f(t, x)}{\varepsilon} \tag{2.42}$$

while the horizontal differential is

$$\lim_{h \downarrow 0} \frac{f(t + h, x_t) - f(t, x)}{h} \tag{2.43}$$

They derived extended Itô formulas for paths with Föllmer's quadratic variation, while incorporating jumps and functions of quadratic variation. In the book Cont et al. (2016), they extend this to *positive principles* for the solutions of *functional Kolmogorov equations* like the valuation equation 2.39, as well as a theory of path-dependent stochastic control. Further applications have been to the the solution via the Euler scheme Cont and Lu (2016), and the calculation of greeks by integration by parts type formulae Jazaerli and Saporito (2013).

2.3 Summary

We saw in Section 2.1 that the absence of arbitrage in a market model is about how self-financing trading strategies can be used to construct a *probability measure* on its *state space*. The state space can only be approximated, uniformly and in variance, by *Monte Carlo simulation*. However, in a Markovian model, Merton's classic argument shows that *if* a no arbitrage value exists as a $C^{1,2}$ function of time and spot, then it must satisfy a particular PDE. The Markov property means that the state space can be simplifed to a product $[0, T] \times \mathbb{R}_+$, and so can be solved by a semigroup of operators acting on functions on $\mathbb{R}_+$, with a convenient product formula. We saw in Section 2.2 that *functional Itô calculus* furnishes analogous results for the general case of a path-dependent model almost surely satisfying an FPDE, on a state space of the *càdlàg paths*, using *Föllmer's quadratic variation* a notion of pathwise variance that requires the weak convergence of discrete time variance computations to a measure of a particular Lebesgue decomposition.

There is a qualitative difference between the simple one-period, finite state model and the continuous time counterpart: in the former we can legitimately list the state space as a set of scenarios. The complexity of the construction of the continuous time model means we lose some concreteness in the possible outcomes, and must speak of statements holding 'almost surely'. To which set of paths does this refer?

From a practitioner's point of view it means paths which can be approximated using discrete time, numerical methods: limits, uniformly and in variance, of Monte Carlo simulations. However, the pathwise methodology takes $\mathbb{D}^d$ as the setting, for which many of the paths do not have finite volatility, in the sense of Föllmer's quadratic variation necessary for the use of his Itô formula. It is 'too big'.

There has been some discussion of this in the literature. Schied and Voloshchenko (2015) notes that the set of paths which *do* have quadratic variation in this sense is in particular not closed under linear combinations, i.e. *is not a vector space*, which means analysis is difficult and in particular intuitive use of the quadratic variation as an inner product will not work. Davis et al. (2018) uses partitions that depend on the path to construct 'badly behaved' quadratic variation. Chiu and Cont (2018) suggest a connection with the Skorokhod topology. This begs the questions: is $\mathbb{D}^d$ the 'right space' to work with? Is Föllmer's quadratic variation the 'right' definition of volatility?

It is the position of this thesis is that the answer to these questions are *no*. A simpler, more general definition of variance that in contrast to Föllmer's quadratic variation doesn't put constraints on the Lebsgue decomposition of the limit measure, yields a vector space for its domain and a genuine inner product type structure. This is covered in Chapter 3.

Both the definitions of the vertical differential, due to Dupire (2009) and Cont and Fournié (2010) respectively, are somewhat odd from the perspective of the analyst. Differential operators are usually defined on (dense subsets of) functions *on a vector space*, as the *generator* of a *(semi)group* of some topological or analytical type. In more plain terms, as the infinitesimal action of some kind of translation. The abstract theory of differentiation from this perspective is well developed and powerful, Engel and Nagel (1999). By contrast, vertical and horizontal differentials are defined on functionals of *unions* or *equivalence classes of paths*, stopped at any of a continuum of points in time. Further, they each require the satisfaction of *a continuum of* limit conditions of difference quotients. As well as being unwieldy and difficult to check in principle, this rules out the possibility of generating any kind of semigroup, and hence the well-developed theory mentioned above. This begs the questions: is this the right way to define vertical/horizontal differentiation? Is this the right way to treat the flow of information?

It is the position of this thesis that the answers to these questions are also *no*. Rather than defining a vertical differential on functionals of equivalence classes of paths using a continuum of limiting difference quotients, the easier approach is to define it *at each separate time* on *functionals of paths*, parameterised families of which can *then* be quotiented to account for the flow of information. This way, it can be seen to be the generator of a strongly continuous group on various (Banach) spaces. This seemingly pedantic point is actually very efficacious as it yields calculus results much more easily. Similarly, the horizontal differential is the generator of a strongly continuous *semigroup*. This is covered in Chapter 4.

The core idea of no arbitrage pricing is contained in that simplest model at the beginning of Section 2.1: *states should be priced consistently with the spot market*; implicit in that is that we can articulate which states 'matter' as a set of concrete scenarios. The results of Chapters 3 and 4 enable us to distinguish which paths matter and which do not, and furnish them with an inner product structure that reflects intuition about (co-)variaton, and hence to construct trading strategies and the notion of arbitrage. This is covered in Chapter 5.

Chapter 3

Monte Carlo Spaces and Pathwise Analysis of Market Models

In this Chapter we address Objective 1 of the thesis Aims from Section 1.2, by introducing Banach spaces of *Monte Carlo paths*, and using them to develop a framework for the pathwise analysis of market risk models.

A fundamental requirement for the management of market risk is the modelling of the evolution of risk factors. Practitioners inform their decision making using concrete sets of future scenarios, and working backwards to the present. As such they tend to think of models as determining plausible market dynamics, using a reasonable function of underlying random 'noise'. This is epitomised by *Monte Carlo simulation* methods, which use a fixed discretisation of a time interval to project piecewise constant trajectories.

However, there is of course no natural finest timescale for discretisation, and so some limiting type notion is required. The main mathematical objects used for this are *Itô processes*, random variable-valued functions of time that *with probability one* obey equations based on *stochastic integrals*. An extremely useful property of such integrals, known as the *Itô isometry*, gives an explicit form for their variance which greatly enables their manipulation.

However, the underlying probability spaces for such models are typically 'too big', in the sense that the probability measure is concentrated on a set that is 'exceptional' in some sense. Being random variables, and therefore only *defined* up to sets of probability zero, strictly speaking Itô processes on these spaces do not give concrete scenarios, but only equivalence classes containing many unsuitable, irrelevant trajectories. For instance, the probability space of the benchmark Black-Scholes model is usually taken to be the space of continuous functions on an interval. Smooth functions are of course dense in this space, but with probability one any model path fails to be differentiable at all times.

This Chapter develops a *pathwise* methodology for analysing risk models, which deterministically project concrete trajectories from candidate 'noise' paths, in line with practitioner intuition.

The solutions live in newly defined Banach spaces whose norms are directly related to piecewise constant approximation, both uniformly and in variance, which we accordingly call *Monte Carlo spaces*. These are purely functional analytic, non-probabilistic, objects whose structure is closely tailored to the operations of financial modelling. The upshot is a reduction of models' probability spaces and a realignment of their topologies, so that models can be thought of as unambiguously determining concrete scenarios for pathwise risk management, as functions of random noise. This will allow us to develop pathwise sensitivity analysis, replication, and valuation methods for derivatives in Chapters 4 and 5.

The rest of this Chapter is organised as follows. In Section 3.1, motivated by the need to consider uniform limits of discrete *time series* in finance, we introduce our basic modelling framework of *paths*, and show how the space of such limits is governed by dense subsets of the interval which we call *clocks*. The space permits a natural group action which permutes the underlying clocks, and whose orbits are copies of familiar *sequence spaces*. In Section 3.2 we introduce Monte Carlo spaces as natural subspaces, and show that their structure embeds practitioners' intuition about (co)-volatility. Finally, In Section 3.3 we introduce an integral with respect to a clock, and show how it can be used to furnish sets of paths with full probability in a generic Ito process model.

3.1 From time series to regulated paths

The basic objects of finance theory are *time series*: sets of data parameterised by time, such as stock prices, interest rates, trading strategies and so on. The modelling problem in finance is essentially how to plausibly extend a given time series from the present to a future date T. We represent this mathematically using *paths*, by which we mean a function $x : [0,T] \to \mathbb{X}$ for some *Banach space* $\mathbb{X}$ (see, for instance, Megginson (1998), Folland (2013), Aliprantis et al. (1999)).

For a simple example, if we choose $\mathbb{X} = \mathbb{R}^d$, the path $t \to x(t) = \sum_{i=1}^{d} x^i(t)e^i$ could represent the evolution over time of d spot market prices, or other risk factors, where e^i, $i = 1, \dots, d$ denotes the canonical basis in $\mathbb{R}^d$. Alternatively, if $\mathbb{X}$ is some space of *path functionals* – e.g. functions of the $\mathbb{R}^d$-valued paths defined above – then we will call an $\mathbb{X}$-valued path $t \to \varphi(t)$ a *functional process*, which might for instance represent a rule to hold $\varphi(t)(x)$ units of a stock at time t when the market follows the path x, ie a *trading strategy*.

The use of continuous time is an idealising assumption. Any financial time series in the real world assigns numbers to time *intervals*, and so defines a piecewise constant path; for instance, the price of a stock on an exchange is updated each time it is traded. To formalise this, recall that a *partition* $\mathfrak{p}$ of $[0,T]$ is a finite set $\{\mathfrak{p}_k : 0 = \mathfrak{p}_0 < \mathfrak{p}_1 < \dots < \mathfrak{p}_m = T\}$, for some $m = m(\mathfrak{p})$; the set of all such partitions is written $\mathcal{P}$. A path which is constant on each of the open intervals $\mathcal{I}_k := (\mathfrak{p}_{k-1}, \mathfrak{p}_k)$, $k = 1, \dots, m(\mathfrak{p})$, of a partition $\mathfrak{p}$ is said to be a $\mathfrak{p}$-*step path*, and we will write the space of such $\mathfrak{p}$-step paths as $\mathbb{S}^{\mathfrak{p}}(\mathbb{X})$. However, there is *a priori* no limit for how fine a partition one should use – so-called 'high frequency traders' are in a perpetual arms race on exactly this issue. Taking the union over all possible partitions, the space of *step paths*

$$\mathbb{S}(\mathbb{X}) := \bigcup_{\mathfrak{p} \in \mathcal{P}} \mathbb{S}^{\mathfrak{p}}(\mathbb{X}) \tag{3.1}$$

houses all possible financial time series, and so is in some sense the fundamental vector space we work with. We will spend the rest of this thesis studying its closures in various senses, and the corresponding operators and functionals.

We will need several norms for paths, but probably the most important is the *supremum* or *uniform* norm

$$||x||_\infty := \sup_{t \in [0,T]} ||x(t)||_{\mathbb{X}} \tag{3.2}$$

and it is easy to see that the *bounded paths* $\mathbb{B}(\mathbb{X}) := \{x : [0,T] \to \mathbb{X} : ||x||_\infty < \infty\}$ is a Banach space containing $\mathbb{S}(\mathbb{X})$. However, it is 'too big' – for instance, it is easy to see the indicator function of the rationals $1_{\mathbb{Q}} \in \mathbb{B}(\mathbb{R})$ is not the uniform limit of step paths. So we will require a little more regularity. The left-limit x_- of a path x is defined for $t \in (0,T]$ as

$$x_-(t) := x(t-) = \lim_{\varepsilon \downarrow 0} x(t-\varepsilon) \tag{3.3}$$

when this limit exists in $\mathbb{X}$, and $x_-(0) := x(0)$. Similarly, the right-limit path x_+ is defined for $t \in [0,T)$ as

$$x_+(t) := x(t+) = \lim_{\varepsilon \downarrow 0} x(t+\varepsilon) \tag{3.4}$$

when this limit exists in $\mathbb{X}$, and $x_+(T) := x(T)$. We can think of these as operators defined on subspaces of $\mathbb{B}(\mathbb{X})$, and note that each of them operators is an *idempotent*: $(x_+)_+ = x_+$, $(x_-)_- = x_-$. The following subspaces defined via these operators are of fundamentally importance.

Definition 3.1. (Regulated, càdlàg, càglàd, null limit paths)

A path $x : [0,T] \to \mathbb{X}$ is said to be *regulated* if both the left x_- and right limit x_+ paths exist; the space of such paths will be denoted $\mathbb{G}(\mathbb{X}) := \mathrm{dom}(.)_+ \cap \mathrm{dom}(.)_-$. We define the subspaces:

- the *càdlàg* or *RCLL path space* $\mathbb{D}(\mathbb{X}) := \{x \in \mathbb{G}(\mathbb{X}) : x = x_+\}$ of *right-continuous paths with left limits*
- the *càglàd* or *LCRL path space* $\mathbb{F}(\mathbb{X}) := \{x \in \mathbb{G}(\mathbb{X}) : x = x_-\}$ of *left-continuous paths with right limits*
- the *continuous* path space $\mathbb{W}(\mathbb{X}) := \{x \in \mathbb{G}(\mathbb{X}) : x = x_+ = x_-\}$,
- the *left null-*, *right null-*, and *null limit* path spaces, $\mathbb{G}_{0-}(\mathbb{X}) := \{x \in \mathbb{G}(\mathbb{X}) : x_- = 0\}$, $\mathbb{G}_{0+}(\mathbb{X}) := \{x \in \mathbb{G}(\mathbb{X}) : x_+ = 0\}$, $\mathbb{G}_0(\mathbb{X}) := \mathbb{G}_{0-}(\mathbb{X}) \cap \mathbb{G}_{0+}(\mathbb{X})$

see Figure 3.1.

PATHS AND DISCONTINUITIES

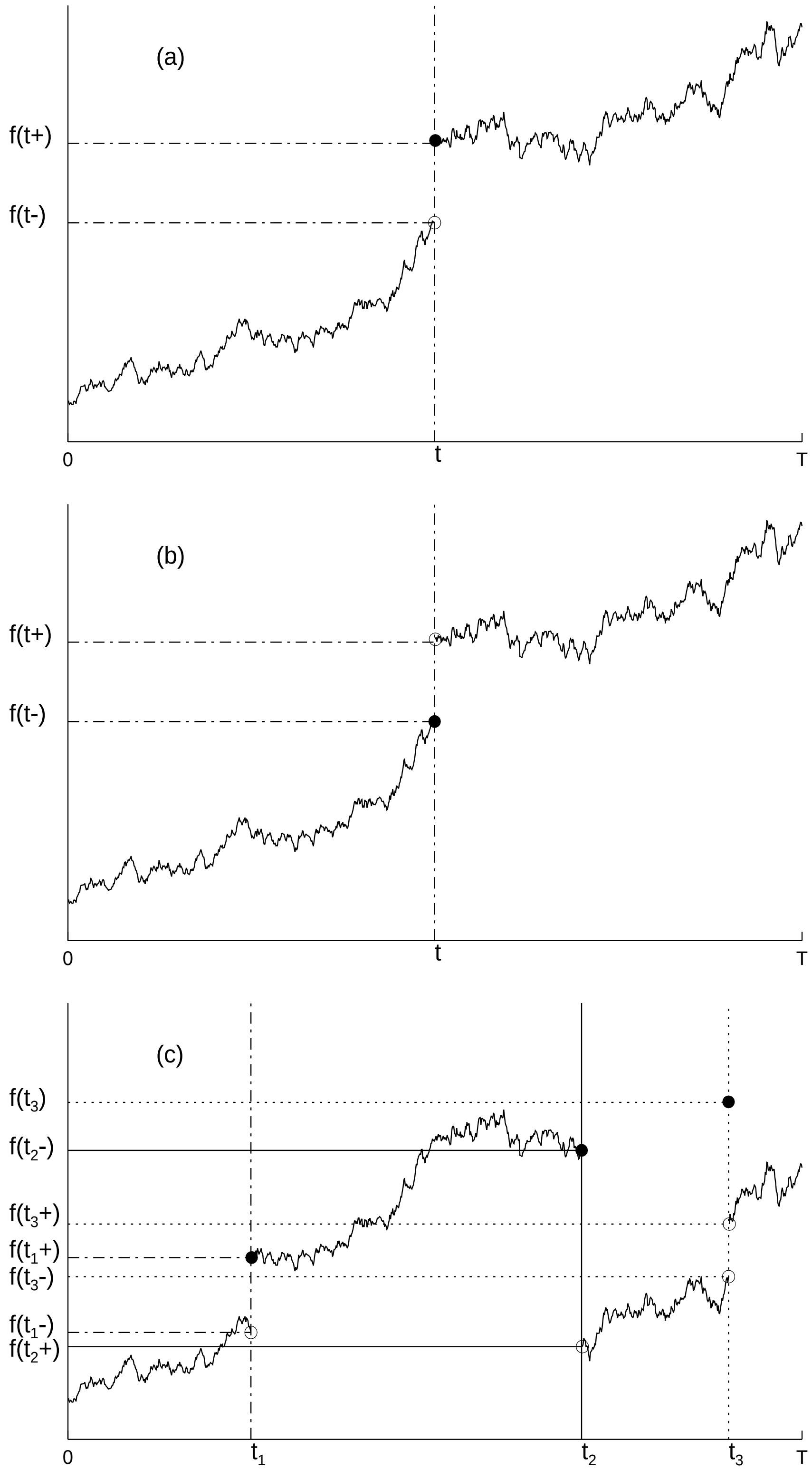

Figure 3.1: Generic (a) cádlág, (b) cáglád, and (c) regulated paths.

These spaces are known to varying degrees in the analysis literature, though only the *càdlàg* and *càglàd* spaces are well known in finance. The terminology comes from French; e.g. 'càdlàg' stands for the acronym *continu à droite, limite à gauche*, and similarly for càglàd. The choice of $\mathbb{W}$ for the continuous paths is to stand for 'Wiener', as it is of course the canonical state space for the Wiener process. Regulated paths were championed by the Bournbaki movement as an alternative setting for integration to the Riemann theory (Berberian (1979)), but are now mainly considered only in specialist analysis texts. The 'left/right null limit' spaces are even more obscure, and the terminology is the author's. When $\mathbb{X} = \mathbb{R}^d$, we write simply $\mathbb{G}^d := \mathbb{G}(\mathbb{R}^d)$, $\mathbb{D}^d := \mathbb{D}(\mathbb{R}^d)$, $\mathbb{F}^d := \mathbb{F}(\mathbb{R}^d)$, $\mathbb{W}^d := \mathbb{W}(\mathbb{R}^d)$.

Obviously $\mathbb{G}(\mathbb{X})$ is quite a general class of functions, including polynomials, smooth functions, continuous functions, piecewise continuous functions, etc. As an example of a function which is *not* regulated, again take $1_{\mathbb{Q}}$, the indicator of the set of rationals. The càdlàg and càglàd paths are naturally isomorphic via time reversal: if x is càdlàg, then $\overleftarrow{x}(t) := x(T-t)$ is cáglád, and vice versa. The choice of whether to use càdlàg and càglàd paths to model a phenomenon is about how to incorporate a jump. It is natural to model a risk factor as immediately including the jump value, and a trading strategy as a function of these risk factors that adjusts *immediately after*. The spaces $\mathbb{D}^d$ and $\mathbb{F}(\mathbb{D}^d)$ will house these two objects, and accordingly be most important to us.

The *jump* at t of a regulated path x is

$$\Delta x(t) := x_+(t) - x_-(t) \in \mathbb{X} \tag{3.5}$$

It can be decomposed as the sum of the jumps *from the right* $\Delta_+ x(t) := x_+(t) - x(t)$ and *from the left*, $\Delta_- x(t) := x(t) - x_-(t)$, note that $\Delta_- x(0) = \Delta_+ x(T) = 0$. If x is càdlàg, $\Delta_+ x \equiv 0$ so $\Delta x(t) = \Delta_- x(t)$; similarly if x is càglàd, $\Delta_- x \equiv 0$ so $\Delta x(t) = \Delta_+ x(t)$; and of course if x is continuous, $\Delta_- x = \Delta_+ x = \Delta x = 0$.

Since for any regulated path $x_+(t) - \Delta_+ x(t) = x(t) = x_-(t) + \Delta_- x(t)$, we have the decompositions

$$x = x_+ - \Delta_+ x \tag{3.6}$$
$$x = x_- + \Delta_- x \tag{3.7}$$
$$x = \frac{1}{2}\left(x_+ + x_-\right) + \frac{1}{2}\left(\Delta_- x - \Delta_+ x\right) \tag{3.8}$$

The set of *ε-jump times* of $x \in \mathbb{G}(\mathbb{X})$ is

$$\mathcal{J}_\varepsilon(x) := \{t : ||\Delta_+ x(t)|| + ||\Delta_- x(t)|| \geq \varepsilon\} \tag{3.9}$$

and we write $\mathcal{J}(x) := \bigcup_{\varepsilon>0} \mathcal{J}_\varepsilon(x)$ for the set of *jump times*.

The proof of the following classical result is instructive, so we include it here for ease of reference.

Proposition 3.2. *For any $x \in \mathbb{G}(\mathbb{X})$ and $\varepsilon > 0$, $\mathcal{J}_\varepsilon(x)$ is finite, and so $\mathcal{J}(x)$ is countable.*

Proof. Given $x \in \mathbb{G}(\mathbb{X})$, the existence of left and right limits means that for each $t \in [0, T]$ and $\varepsilon > 0$ there exists $\delta(t, \varepsilon) > 0$ such that if $u, v \in B(t, \delta(t, \varepsilon))$, the open ball centred at t of radius $\delta(t, \varepsilon)$, then

$$(t-u)(t-v) > 0 \implies ||x(u) - x(v)|| < \varepsilon \tag{3.10}$$

where the product condition concisely ensures we only consider when u, v approach t from the same side. In particular, if $t' \in B(t, \delta(t, \varepsilon)) \setminus t$, then for sufficiently small η, $t' \pm \eta$ are on the same side of t as t', and so $||\Delta x(t')|| = ||x(t'+) - x(t'-)|| < \varepsilon$. Since $[0, T]$ is compact, finitely many such balls, say M, cover it and so $||x||_\infty \leq M\varepsilon$. If we call the finite set consisting of the centres of these balls $C(\varepsilon)$, then $\mathcal{J}_\varepsilon(x) \subset C(\varepsilon)$. □

In particular Proposition 3.2 shows that regulated paths are bounded and almost everywhere continuous, and so are Bochner integrable (or Lebesgue integrable, in the $\mathbb{R}^d$-valued case). The condition 3.10 suggests that the regulated property can be interpreted as a form of continuity; this is indeed the case, as shown in the recent paper of Cichon et al. (2018). Although this is in some sense a deep result, the construction of the relevant topological space is somewhat technical, and of limited use in finance. However, the following partial result is fairly straightforward, and describes a fundamental decomposition. It is essentially a new interpretation of Corollary 1 of Berberian (1978), in terms of a standard Banach space of functions, which we now briefly recall. For a locally compact Hausdorff space Y, $C_0(Y, \mathbb{X})$ denotes the space of continuous $\mathbb{X}$-valued functions *vanishing at infinity*, that is, continuous $f : Y \to \mathbb{X}$ with the property that for each $\varepsilon > 0$ there exists a compact set $K \subset Y$ such that if $y \in K^c$ then $||f(y)||_{\mathbb{X}} < \varepsilon$. For notational convenience we set $\mathcal{T}_d$ to be $[0, T]$ endowed with the (locally compact, Hausdorff) discrete topology.

Corollary 3.3. *If $x \in \mathbb{G}(\mathbb{X})$, then $\Delta_- x, \Delta_+ x \in \mathbb{G}_0(\mathbb{X})$. Further, the null limit paths $\mathbb{G}_0(\mathbb{X}) \cong C_0(\mathcal{T}_d, \mathbb{X})$, and we have the vector space decompositions $\mathbb{G}(\mathbb{X}) \cong \mathbb{D}(\mathbb{X}) \oplus C_0(\mathcal{T}_d, \mathbb{X}) \cong \mathbb{F}(\mathbb{X}) \oplus C_0(\mathcal{T}_d, \mathbb{X})$.*

Proof. Since the right- and left-limit operators are idempotents, it is immediate from the definition that $(\Delta_- x)_+ = (\Delta_- x)_- = 0$, and similarly for the right-jumps. Since compactness in $\mathcal{T}_d$ coincides with having finitely many elements, $C_c(\mathcal{T}_d, \mathbb{X})$ is simply the paths with finite support, and their closure is exactly the null limit paths. The rest follows from the decomposition formulae 3.6, and the obvious fact that $\mathbb{G}_0(\mathbb{X}) \cap \mathbb{D}(\mathbb{X}) = \mathbb{G}_0(\mathbb{X}) \cap \mathbb{F}(\mathbb{X}) = \{0\}$. □

For $u \in [0, T]$, define three families of *simple* scalar paths $\alpha[u] \in \mathbb{G}^1, \beta[u] \in \mathbb{D}^1, \gamma[u] \in \mathbb{F}^1$ by

$$\alpha[u](t) := \begin{cases} 1 & \text{if } t = u \\ 0 & \text{otherwise} \end{cases} \tag{3.11}$$

$$\beta[u](t) := 1_{[u,T]}(t) = \begin{cases} 1 & \text{if } u \leq t \leq T \\ 0 & \text{otherwise} \end{cases} \tag{3.12}$$

and

$$\gamma[u](t) := 1_{(u,T]}(t) = \begin{cases} 1 & \text{if } u < t \le T \\ 0 & \text{otherwise} \end{cases} \tag{3.13}$$

These are the prototypes of discontinuous 'shock' paths: the $\alpha[u]$ are isolated shocks; the $\beta[u]$ are right-continuous, reflecting an instantaneously incorporated jump; while the $\gamma[u]$ are left-continuous, reflecting a jump incorporated infitesimally afterward. We will use these elementary paths to build up more general paths using appropriate projection operators, a bit like a basis. Note that $\alpha[u] = \beta[u] - \gamma[u]$, and $\beta[0] = 1_{[0,T]}$ is the unit path, while $\gamma[T] = 0$. If $u < v$, we can write the indicators of half-open intervals as $1_{[u,v)} = \beta[u] - \beta[v]$, $1_{(u,v]} = \gamma[u] - \gamma[v]$, and of an open interval as $1_{(u,v)} = \gamma[u] - \beta[v]$. Finally, note that if $t \neq t'$, $||\alpha[t] - \alpha[t']|| = ||\beta[t] - \beta[t']|| = ||\gamma[t] - \gamma[t']|| = 1$; in particular this shows that each of $\mathbb{G}^1, \mathbb{D}^1$ and $\mathbb{F}^1$ is *non-separable* – none has a countable dense subset.

A general $\mathfrak{p}$-step path $x \in \mathbb{S}^{\mathfrak{p}}(\mathbb{X})$ may be written

$$x = x_0 + \sum_{k=1}^{m(\mathfrak{p})} x_{2k-1}\gamma[\mathfrak{p}_{k-1}] + x_{2k}\beta[\mathfrak{p}_k] \tag{3.14}$$

for some $x_k \in \mathbb{X}$, $k = 0, 1, \ldots, 2m(\mathfrak{p}) + 1$. The corresponding map $i : \mathbb{S}^{\mathfrak{p}}(\mathbb{X}) \to \mathbb{X}^{2m(\mathfrak{p})+1}$ given by $i(x)_k := x_k$ gives an isomorphism $\mathbb{S}^{\mathfrak{p}}(\mathbb{X}) \cong \mathbb{X}^{2m(\mathfrak{p})+1}$, and $i(x)_0 = x(0)$, $i(x)_{2k-1} = \Delta_+ x(\mathfrak{p}_{k-1})$, $i(x)_{2k} = \Delta_- x(\mathfrak{p}_k)$, for $k = 1, \ldots, m(\mathfrak{p})$. If we write the space of right-continuous $\mathfrak{p}$-step paths $\mathbb{S}^{\mathfrak{p}}_r(\mathbb{X}) = \mathbb{S}^{\mathfrak{p}}(\mathbb{X}) \cap \mathbb{D}(\mathbb{X})$, then $i(x)_{2k-1} = 0$, $\mathbb{S}^{\mathfrak{p}}_r(\mathbb{X}) \cong \mathbb{X}^{m(\mathfrak{p})+1}$ and $\{\beta[t]\}$ spans $\mathbb{S}_r(\mathbb{X}) := \bigcup_{\mathfrak{p}} \mathbb{S}^{\mathfrak{p}}_r(\mathbb{X})$. Similarly, if $x \in \mathbb{S}^{\mathfrak{p}}_l(\mathbb{X}) := \mathbb{S}^{\mathfrak{p}}(\mathbb{X}) \cap \mathbb{F}(\mathbb{X})$, the left-continuous $\mathfrak{p}$-step paths, then $i(x)_{2k-1} = 0$, and so $\mathbb{S}^{\mathfrak{p}}_l(\mathbb{X}) \cong \mathbb{X}^{m(\mathfrak{p})+1}$ and $\{\gamma[t]\}$ spans $\mathbb{S}_l(\mathbb{X}) := \bigcup_{\mathfrak{p}} \mathbb{S}^{\mathfrak{p}}_l(\mathbb{X})$.

We now introduce some new operators, which will be very helpful to the analysis of path spaces. Given a partition $\mathfrak{p}$ and $x \in \mathbb{G}(\mathbb{X})$, let $\mathcal{R}_k(x)$ be the closure of $x(\mathcal{I}_k)$. An operator $\mathcal{E}^{\mathfrak{p}} : \mathbb{G}(\mathbb{X}) \to \mathbb{S}^{\mathfrak{p}}(\mathbb{X})$ such that $\mathcal{E}^{\mathfrak{p}}(x)(t) = \lambda_k(x) \in \mathcal{R}_k(x)$ if $t \in (\mathfrak{p}_{k-1}, \mathfrak{p}_k)$, for $k = 1, \ldots, m(\mathfrak{p})$, and $\mathcal{E}^{\mathfrak{p}}(x)(\mathfrak{p}_k) = x(\mathfrak{p}_k)$, for $k = 0, 1, \ldots, m(\mathfrak{p})$, will be called an *Euler projection* on $\mathbb{G}(\mathbb{X})$ (the name is after the common numerical scheme for piecewise constant approximation of SDE solutions). We may write it as

$$\mathcal{E}^{\mathfrak{p}}(x) = \sum_{k=0}^{m(\mathfrak{p})} x(\mathfrak{p}_k)\alpha[\mathfrak{p}_k] + \sum_{k=1}^{m(\mathfrak{p})} \lambda_k(x)\big(\gamma[\mathfrak{p}_{k-1}] - \beta[\mathfrak{p}_k]\big) \tag{3.15}$$

Taking the right-limit we get an operator $\mathcal{E}^{\mathfrak{p}}_r : \mathbb{D}(\mathbb{X}) \to \mathbb{S}_r(\mathbb{X})$. It is natural to choose $\lambda_k(x) = x(\mathfrak{p}_k+) = x(\mathfrak{p}_k)$, in which case we get

$$\mathcal{E}^{\mathfrak{p}}_r(x) := x(0) + \sum_{k=1}^{m(\mathfrak{p})} \big(x(\mathfrak{p}_k) - x(\mathfrak{p}_{k-1})\big)\beta[\mathfrak{p}_k] \tag{3.16}$$

We will mean this choice of λ whenever we refer to 'the Euler projection on $\mathbb{D}(\mathbb{X})$'. It is useful to define the *differences along* $\mathfrak{p}$ of a path x to be the sequence $\delta^{\mathfrak{p}} x \in c_{00}(\mathbb{X})$, with components

$$\delta_k^{\mathfrak{p}} x := x(\mathfrak{p}_k) - x(\mathfrak{p}_{k-1}) \tag{3.17}$$

for $k = 1, \ldots, m(\mathfrak{p})$, and $\delta_k^{\mathfrak{p}} x = 0$ for $k = 0$ and $k > m(\mathfrak{p})$. In that case we have

$$\mathcal{E}_r^{\mathfrak{p}}(x) := x(0) + \sum_{k=1}^{m(\mathfrak{p})} \delta_k^{\mathfrak{p}} x \beta[\mathfrak{p}_k] \tag{3.18}$$

Similarly, taking the left-limit and choosing $\lambda_k(x) = x(\mathfrak{p}_{k+1}-) = x(\mathfrak{p}_{k+1})$, we get $\mathcal{E}_l^{\mathfrak{p}} : \mathbb{F}(\mathbb{X}) \to \mathbb{S}_l(\mathbb{X})$

$$\mathcal{E}_l^{\mathfrak{p}}(x) := x(0) + \sum_{k=0}^{m(\mathfrak{p})-1} \big(x(\mathfrak{p}_{k+1}) - x(\mathfrak{p}_k)\big)\gamma[\mathfrak{p}_k] \tag{3.19}$$

and we will always mean this by 'the Euler projection on $\mathbb{F}(\mathbb{X})$'.

The following classical theorem is fundamental, as it says that the regulated/cádlág/cáglád paths are the uniform closures of step/left-continuous step/right-continuous step paths respectively, and so house the limiting objects approximated by real-world time series. We include a proof based on Dieudonné (1993), but emphasising the role of the Euler projections.

Theorem 3.4. *If $\mathbb{X}$ is a Banach space (resp. algebra) then $\mathbb{G}(\mathbb{X})$, $\mathbb{D}(\mathbb{X})$ and $\mathbb{F}(\mathbb{X})$ are Banach spaces (resp. algebras) in which the step paths $\mathbb{S}(\mathbb{X})$, $\mathbb{S}_r(\mathbb{X})$ and $\mathbb{S}_l(\mathbb{X})$ respectively are dense.*

Proof. If $\mathbb{X}$ is a Banach algebra it is easy to check that the path spaces are algebras under the pointwise product $x.y(t) := x(t).y(t)$, and that this is continuous with respect to the uniform norm; we need only prove they coincide with the uniform closures of the relevant step paths when $\mathbb{X}$ is merely a Banach space.

Suppose $x \in \overline{\mathbb{S}(\mathbb{X})}$. Then for each $\varepsilon > 0$ there exists a partition $\mathfrak{p}$ and $x_\varepsilon \in \mathbb{S}^{\mathfrak{p}}(\mathbb{X})$ such that $||x - x_\varepsilon||_\infty < \frac{\varepsilon}{3}$. Since $\mathbb{S}(\mathbb{X}) \subset \mathbb{G}(\mathbb{X})$, for any $t \in [0, T]$ there exists $\delta > 0$ such that $||x_\varepsilon(u) - x_\varepsilon(v)|| < \frac{\varepsilon}{3}$ whenever $0 < (t-u)(t-v) < \delta$. Hence for such u, v, $||x(u) - x(v)|| \leq ||x(u) - x_\varepsilon(u)|| + ||x_\varepsilon(u) - x_\varepsilon(v)|| + ||x(v) - x_\varepsilon(v)|| < \varepsilon$, so $\overline{\mathbb{S}(\mathbb{X})} \subset \mathbb{G}(\mathbb{X})$. A similar argument shows $\overline{\mathbb{S}_r(\mathbb{X})} \subset \mathbb{D}(\mathbb{X})$and $\overline{\mathbb{S}_l(\mathbb{X})} \subset \mathbb{F}(\mathbb{X})$.

Now suppose $x \in \mathbb{G}(\mathbb{X})$ and $\varepsilon > 0$. Consider the partition $\mathfrak{p}^\varepsilon = \mathfrak{p}^\varepsilon(x) := \{0, T\} \cup C(N)$ where $\frac{1}{N} < \varepsilon$ and $C(N) := \{c_k\}$ is the set whose points are the centres of the balls constructed in the proof of proposition 3.2; the closure $\mathcal{R}_{\mathfrak{p}_k^\varepsilon}(x)$ of the image of $(\mathfrak{p}_k^\varepsilon, \mathfrak{p}_{k+1}^\varepsilon)$ under x has diameter less than ε. The result then follows using the Euler projection on this partition: on $\mathbb{G}(\mathbb{X})$, set $x_\varepsilon := \mathcal{E}^{\mathfrak{p}_\varepsilon}(x)$ for some $\lambda_k(x) \in \mathcal{R}_{\mathfrak{p}_k^\varepsilon}(x)$; then $||x - x_\varepsilon|| < \varepsilon$. On $\mathbb{D}(\mathbb{X})$ and $\mathbb{F}(\mathbb{X})$ use instead $x_\varepsilon := \mathcal{E}_r^{\mathfrak{p}_\varepsilon}(x)$, and $x_\varepsilon := \mathcal{E}_l^{\mathfrak{p}_\varepsilon}(x)$ respectively. □

Theorem 3.4 constructs approximations of general regulated paths as limits of step paths defined on sequences of partitions. The remainder of this section is dedicated to understanding how the choice of this sequence of partitions is reflected in the structure of $\mathbb{G}(\mathbb{X})$. We start with the following definition of some concepts which are convenient for parameterising numerical approximations.

Definition 3.5. (Timescales, refining sequences of partitions, and clocks)

We will call a countable dense subset of $[0, T]$ a *timescale*. A sequence $\mathscr{P} := (\mathfrak{p}^n)$ of partitions will be said to be *dense* if its *union* $\mathscr{P}_\infty := \cup\mathscr{P} = \{\mathfrak{p}^n_k : k = 0, 1, \dots, m(\mathfrak{p}^n); n \in \mathbb{N}\}$ is a timescale, or equivalently if the *mesh* $|\mathfrak{p}^n| := \max_k |\mathfrak{p}^n_k - \mathfrak{p}^n_{k-1}| \to 0$. Such a dense sequence of partitions will be said to be *refining* if it is in addition (strictly) increasing, i.e. $\mathfrak{p}^n \subsetneq \mathfrak{p}^{n+1}$. A *clock* with timescale $\mathscr{T}$ is a refining sequence of partitions with union $\mathscr{T}$.

'Mesh' and 'refining' are common terms in the literature; the terminology of 'timescales' and 'clocks' is due to the author. The prototypical example of a clock is given by the *dyadic rationals* $\mathscr{Q} = (\mathfrak{q}^n)$, with $\mathfrak{q}^n_k := \frac{kT}{2^n}$, $k = 0, 1, \dots, 2^n$, and timescale $\mathscr{Q}_\infty := \{t \in [0, T] : \frac{2^n t}{T} \in \mathbb{N}\}$.

For a given clock $\mathscr{P}$, we set the $\mathscr{P}$-step paths $\mathbb{S}^{\mathscr{P}}(\mathbb{X}) := \cup_n \mathbb{S}^{\mathfrak{p}^n}(\mathbb{X})$, and the right- and left-continuous versions $\mathbb{S}^{\mathscr{P}}_r(\mathbb{X}) := \mathbb{S}^{\mathscr{P}}(\mathbb{X}) \cap \mathbb{D}(\mathbb{X})$, $\mathbb{S}^{\mathscr{P}}_l(\mathbb{X}) := \mathbb{S}^{\mathscr{P}}(\mathbb{X}) \cap \mathbb{F}(\mathbb{X})$. We write their closures in the uniform norm as $\mathbb{G}_{\mathscr{P}}(\mathbb{X}) := \overline{\mathbb{S}^{\mathscr{P}}(\mathbb{X})}$, $\mathbb{D}_{\mathscr{P}}(\mathbb{X}) := \overline{\mathbb{S}^{\mathscr{P}}_r(\mathbb{X})}$, $\mathbb{F}_{\mathscr{P}}(\mathbb{X}) := \overline{\mathbb{S}^{\mathscr{P}}_l(\mathbb{X})}$, and call them the *$\mathscr{P}$-regulated*, *$\mathscr{P}$-càdlàg*, *$\mathscr{P}$-càglàd* paths, respectively. Note that by construction these spaces are separable whenever $\mathbb{X}$ is, and that because of the unit distance between distinct $\beta[t]$, $\gamma[t]$, $\mathbb{G}_{\mathscr{P}}(\mathbb{X}) = \{x \in \mathbb{G}(\mathbb{X}) : \mathcal{J}(x) \subset \mathscr{P}\}$, $\mathbb{D}_{\mathscr{P}}(\mathbb{X}) = \{x \in \mathbb{D}(\mathbb{X}) : \mathcal{J}(x) \subset \mathscr{P}\}$, $\mathbb{F}_{\mathscr{P}}(\mathbb{X}) = \{x \in \mathbb{F}(\mathbb{X}) : \mathcal{J}(x) \subset \mathscr{P}\}$. In particular, $\mathbb{W}(\mathbb{X}) = \mathbb{D}_{\mathscr{P}}(\mathbb{X}) \cap \mathbb{F}_{\mathscr{P}}(\mathbb{X})$.

We now characterise these spaces, and hence $\mathbb{G}(\mathbb{X})$, in terms of classical *sequence spaces*, which we briefly recall here for ease of reference (see Megginson (1998), Diestel (1984), Leonard (1976) for more). The space of *sequences in* $\mathbb{X}$ is written $\mathbb{X}^{\mathbb{N}}$. The *partial sum* operator $s : \mathbb{X}^{\mathbb{N}} \to \mathbb{X}^{\mathbb{N}}$ defined as

$$s(x)_n := \sum_{k=1}^{n} x_k \tag{3.20}$$

is a linear isomorphism, with inverse $s^{-1}(x)_1 = x_1$, $s^{-1}(x)_n = x_n - x_{n-1}$, $n > 1$.

Definition 3.6. (Sequence spaces)

- The space of *eventually zero sequences in* $\mathbb{X}$,

$$c_{00}(\mathbb{X}) := \{x = (x_n) \in \mathbb{X}^{\mathbb{N}} : \text{there exists } N \text{ s.t. } x_n = 0 \text{ if } n \geq N\} \tag{3.21}$$

  is invariant under s, and in some sense the 'minimal' space containing $\mathbb{X}^n$ for each n

- The space of *sequences converging to zero in* $\mathbb{X}$,

$$c_0(\mathbb{X}) := \{x = (x_n) \in \mathbb{X}^{\mathbb{N}} : x_n \to 0 \text{ as } n \to \infty\} \tag{3.22}$$

and of *convergent sequences in* $\mathbb{X}$,

$$c(\mathbb{X}) := \{x = (x_n) \in \mathbb{X}^{\mathbb{N}} : \text{ there exists } x^* \in \mathbb{X} \text{ s.t. } x_n \to x^* \text{ as } n \to \infty\} \tag{3.23}$$

are Banach spaces under the norm $||x||_\infty := \sup_n ||x_n||$. We have $c_0(\mathbb{X}) = \overline{c_{00}(\mathbb{X})}$, and $c(\mathbb{X}) \cong c_0(\mathbb{X}) \oplus \mathbb{X}$

- The space of *convergent series in* $\mathbb{X}$,

$$cs(\mathbb{X}) := \{x = (x_n) \in \mathbb{X}^{\mathbb{N}} : s(x) \in c(\mathbb{X})\} \tag{3.24}$$

is a Banach space under the norm $||x||_{cs} := ||s(x)||_\infty = \sup_n || \sum_{k=1}^n x_k ||$, and is isometrically isomorphic to $c(\mathbb{X})$ by construction

- For $1 \le p < \infty$, the space of *p-summable sequences in* $\mathbb{X}$,

$$\ell^p(\mathbb{X}) := \{x = (x_n) \in \mathbb{X}^{\mathbb{N}} : \sum_{k=1}^{n} ||x_k||^p \in c(\mathbb{R})\} \tag{3.25}$$

is a Banach space under the norm $||x||_{\ell^p(\mathbb{X})} := (\lim_{n\to\infty} \sum_{k=1}^n ||x_k||^p)^{\frac{1}{p}}$, in which $c_{00}(\mathbb{X})$ is dense

- The space of *bounded sequences in* $\mathbb{X}$,

$$\ell^\infty(\mathbb{X}) := \{x = (x_n) \in \mathbb{X}^{\mathbb{N}} : ||x||_\infty < \infty\} \tag{3.26}$$

is a Banach space under $||.||_\infty$

Clearly $\ell^1(\mathbb{X}) \subset cs(\mathbb{X})$, $\ell^1(\mathbb{X}) \subset \ell^p(\mathbb{X}) \subset \ell^\infty(\mathbb{X})$, and it is well known that $c(\mathbb{X})^* \cong c_0(\mathbb{X})^* \cong \ell^1(\mathbb{X}^*)$; if $1 < p, q < \infty$ then $\ell^p(\mathbb{X})^* \cong \ell^q(\mathbb{X}^*)$ where $\frac{1}{p} + \frac{1}{q} = 1$; and $\ell^1(\mathbb{X})^* \cong \ell^\infty(\mathbb{X}^*)$ (Leonard (1976)).

The following lemma is a new result identifying copies of a sequence space within path spaces, indexed by clocks.

Lemma 3.7. *For any clock* $\mathscr{P}$ *we have the isometric isomorphisms* $\mathbb{G}_{\mathscr{P}}(\mathbb{X}) \cong \mathbb{D}_{\mathscr{P}}(\mathbb{X}) \cong \mathbb{F}_{\mathscr{P}}(\mathbb{X}) \cong c(\mathbb{X})$.

Proof. We concentrate initially on the claim for $\mathscr{P}$*-càdlàg* paths. For any such path x and $\varepsilon > 0$, there exists $n \in \mathbb{N}$ and a finite sequence x_k^n, $k = 0, 1, \dots, m(\mathfrak{p}^n)$, such that the $\mathfrak{p}^n$-step path

$$x_\varepsilon := \sum_{k=0}^{m(\mathfrak{p}^n)} x_k^n \beta[\mathfrak{p}_k^n] \tag{3.27}$$

satisfies $||x - x_\varepsilon||_\infty < \varepsilon$.

There are two natural orders on $\mathscr{P}_\infty$: the order inherited from $[0,T]$ and the lexicographic order whereby $\mathfrak{p}^n_k \prec \mathfrak{p}^{n'}_{k'}$ if $n < n'$ or if $n = n'$ and $k < k'$. We may list $\mathscr{P}_\infty$ in this latter order as a sequence, which we write $\mathfrak{r} := (\mathfrak{r}_k)$. We may then form the sequence $y(x) \in \mathbb{X}^{\mathbb{N}}$, with entry $y(x)_k = y_k = x^n_l$, where $\mathfrak{r}_k = \mathfrak{p}^n_l$. Clearly the correspondence $x \leftrightarrow y$ is linear and bijective onto its image, and by the above argument we have

$$x(t) = \sum_{\mathfrak{r}_k \le t} y_k = \lim_{n\to\infty} \sum_{\mathfrak{r}_k \prec \mathfrak{p}^n_{m(\mathfrak{p}^n)}} y_k 1_{\mathfrak{r}_k \le t} \tag{3.28}$$

for each t. Define the norms $||.||_t$ on $\mathbb{X}^{\mathbb{N}}$ by

$$||y||_t = ||\sum_{\mathfrak{r}_k \le t} y_k||_{\mathbb{X}} \tag{3.29}$$

and

$$||y||^* = \max_{t\in[0,T]} ||y||_t \tag{3.30}$$

which exists as for each y, $t \to ||y||_t$ is continuous. This makes the isomorphism isometric, and by construction the image is $cs(\mathbb{X}) \cong c(\mathbb{X})$. The proof for $\mathbb{F}_{\mathscr{P}}(\mathbb{X})$ and $\mathbb{G}_{\mathscr{P}}(\mathbb{X})$ are similar. □

So far we have worked tended to work with a fixed clock; but the proof of proposition 3.4 constructs *for a given path* x a clock $\mathfrak{p}^\epsilon(x)$, which in particular satisfies $\mathcal{J}_\epsilon(x) \subset \mathfrak{p}^\epsilon(x)$, and approximations along it. We will formalise this construction with a convenient concept due to the author, which we call a *trader's clock*; the idea is to add to a fixed clock the first times the risk factors move a sufficiently large amount, as a trader monitoring her portfolio at an increasingly fine grid of times might.

Definition 3.8. (Trader's clock)

A *trader's clock* is a sequence of maps $\tau^n : \mathbb{G}(\mathbb{X}) \to \mathcal{P}$ such that there exists a *resolution* sequence $r_n > 0$ such that $r_n \to 0$ as $n \to \infty$, and

$$\tau^n_k(x) = \inf\{t \in [\tau^n_{k-1}(x), p^n_k] \mid |x(t) - x(\tau^n_{k-1}(x))| \ge r_n > 0\} \tag{3.31}$$

where $p^n_k = \inf \mathfrak{p}^n \cap [\tau^n_{k-1}(x), T]$, and $\mathscr{P} := (\mathfrak{p}^n)$ is a fixed clock, independent of x.

In the probabilistic setting this is recognisable as a particular case of a *stopping time*; it is formalised in the deterministic setting for the first time here. While its particular values may depend on the path, a trading clock is completely specified by its (countable) path-independent data: a 'background' clock, usually the dyadic rationals $\mathcal{Q}$ for convenience, and a resolution $r \in c_0(\mathbb{R})_+$. The definition is recursive: the successor of a given time in τ^n is either the next time the path is further than r_n from its value at that time, or the next member of $\mathfrak{p}^n$ – whichever comes first. In particular, for each $x \in \mathbb{G}(\mathbb{X})$, $\tau^n(x)$ is at least as fine as $\mathfrak{p}^n$, and so $\tau_\infty(x) := \cup_n \tau^n(x)$ is dense. Moreover, $\mathcal{J}_{r_n}(x) \subset \tau^n(x)$ so in particular, $\sup_{t\in[0,T]\setminus\tau^n} |\Delta x(t)| \to 0$ as $n \to \infty$.

Theorem 3.4 can now be restated as: *a path is regulated if and only if it is the uniform limit of a sequence Euler projections* $\mathcal{E}^{\tau^n(.)}(.)$ *along a trading clock* $\tau^n(.)$. Note that since the trading clock depends on the path, the operators $\mathcal{E}^{\tau^n(.)}$ are in general *non-linear*, mapping a regulated path x to a sequence $\mathcal{E}^{\tau^n(x)}(x) \in \mathbb{G}_{\tau(x)}(\mathbb{X})$.

If we think of the interval $[0,T]$ as an ordered topological space, then the appropriate notion of an *order-preserving automorphism* of it is a continuous, strictly increasing bijection $\lambda : [0,T] \to [0,T]$, and we write $\Lambda := \mathrm{Aut}[0,T]$ for the group consisting of such functions (Lang (2005)). Intuitively, these correspond to reparameterisations of the time variable. If $\mathscr{T} \subset [0,T]$ is a timescale, then for any $\lambda \in \Lambda$, $\lambda(\mathscr{T})$ is also a timescale; we can think of Λ as a subgroup of the permutation group of the set of timescales, specifically those which preserve the order of $[0,T]$. The idea behind the main result of this section, due to the author, is that this permuting action is 'well behaved' and 'fills up' $\mathbb{G}(\mathbb{X})$.

Theorem 3.9. *(Structure Theorem for* $\mathbb{G}(\mathbb{X})$*)*

There exists a faithful action $a : \Lambda \to Aut(\mathbb{G}(\mathbb{X}))$ *of* Λ *on* $\mathbb{G}(\mathbb{X})$*, given on the simple paths by composition*

$$a_\lambda(\alpha[u])(t) := \alpha[u](\lambda(t)) \tag{3.32}$$

$$a_\lambda(\beta[u])(t) := \beta[u](\lambda(t)) \tag{3.33}$$

$$a_\lambda(\gamma[u])(t) := \gamma[u](\lambda(t)) \tag{3.34}$$

and extended to $\mathbb{G}(\mathbb{X})$ *by linearity and continuity. The corresponding set of orbits is parameterised by the clocks* $\mathscr{P}$*, and*

$$\mathbb{G}(\mathbb{X})\big/_{\Lambda} \cong \mathbb{G}_{\mathscr{P}}(\mathbb{X}) \cong c(\mathbb{X}) \tag{3.35}$$

Proof. It is easy to check 3.32 gives a well-defined action of Λ on $\mathbb{S}(\mathbb{X})$ and hence $\mathbb{G}(\mathbb{X})$. If some $\lambda^* \in \Lambda$ was such that $x = x \circ \lambda^*$ for all x, then in particular it would fix all continuous paths, and so $\lambda(t) = t$ and $\alpha_\lambda = I$; hence the action is faithful. For a given $\lambda \in \Lambda$, the effect on a sequence $x_n \in \mathbb{S}(\mathbb{X})$ of the action is to permute its underlying clock's timescale $\mathscr{T} \to \lambda^{-1}(\mathscr{T})$. By Theorem 3.4, α_λ is an automorphism of $\mathbb{G}(\mathbb{X})$, and in particular defines a family of isometric isomorphisms $\mathbb{G}_{\mathscr{P}}(\mathbb{X}) \to \mathbb{G}_{\mathscr{P}'}(\mathbb{X})$. Define the equivalence relation $\sim_\Lambda$ on $\mathbb{G}(\mathbb{X})$ by $x \sim_\Lambda y \iff$ there exists $\lambda \in \Lambda$ such that $y = \alpha_\lambda(x)$, and the orbits $\mathscr{O}(x) = \{y \in \mathbb{G}(\mathbb{X}) : x \sim_\Lambda y\}$ for $x \in \mathbb{G}(\mathbb{X})$. By Theorem 3.4 there is some $\mathscr{P}$ such that $x \in \mathbb{G}_{\mathscr{P}}(\mathbb{X})$, so there is a sequence $\xi^n \in c_{00}(\mathbb{X})$ such that $\sum_k \xi^n_k \in c(\mathbb{X})$, and $x = \lim (i^{\mathscr{P}}_n)^{-1}(\xi^n)$. Since $\alpha_\lambda((i^{\mathscr{P}}_n)^{-1}(\xi^n)) = (i^{\lambda^{-1}(\mathscr{P})}_n)^{-1}(\xi^n)$, a regulated path y is in $\mathscr{O}(x)$ if and only if $y = \lim (i^{\lambda^{-1}(\mathscr{P})}_n)^{-1}(\xi^n)$. So $\mathscr{O}(x)$ is the set of those paths which are uniform limits of the step paths based on the same sequence ξ^n, under different choices of clocks. Hence $\mathbb{G}(\mathbb{X})\big/_{\Lambda} \cong cs(\mathbb{X}) \cong c(\mathbb{X})$. □

Theorem 3.9 is important for numerical purposes because it effectively decomposes the regulated paths into a (separable) sequence space of value increments, and a (separable) group governing when they occur. In fact it is behind a familiar tool from the theory of stochastic process, namely the *Skorokhod metric*, which is exactly the distance along the graph of this group action. Although $\mathbb{G}(\mathbb{X})$ is not complete under this metric (for example, the image of a path under the sequence $\lambda_n(u) := (\frac{u}{T})^n$ yields a non-convergent Cauchy sequence), it can be shown that there is an equivalent metric for which it is (see Billingsley (1968), pp 111 – 115). Since any λ is specified by its values on the dyadic rationals, in the resulting topology $\mathbb{G}(\mathbb{X})$ is separable.

3.2 Pathwise variance and Monte Carlo spaces

In the context of risk models, we are interested primarily in the case $\mathbb{X} = \mathbb{R}^d$, where d is the number of independent sources of risk, so we now mainly work in that setting. The spaces in Section 3.1 are of paths which are *uniform* limits of step paths; they give no explicit constraint for how much member paths can vary over time. In this Section we introduce Banach spaces of *Monte Carlo paths* which provide a natural setting to do so.

For a fixed partition $\mathfrak{p}$, any $x \in \mathbb{S}^{\mathfrak{p}}(\mathbb{R}^d)$ may be written

$$x = \sum_{k=0}^{m(\mathfrak{p})} x_k^{\mathfrak{p}} \beta[\mathfrak{p}_k] \tag{3.36}$$

where $x_{\cdot}^{\mathfrak{p}} \in c_{00}(\mathbb{R}^d)$ is the vector of *co-ordinates* of x. Consider the operator mapping x to $\mathrm{var}_{\mathfrak{p},p}(x) \in \mathbb{S}^{\mathfrak{p}}(\mathbb{R})$, where

$$\mathrm{var}_{\mathfrak{p},p}(x) := \sum_{k=0}^{m(\mathfrak{p})} |x_k^{\mathfrak{p}}|^p \beta[\mathfrak{p}_k] \tag{3.37}$$

Note that this is an increasing path with values in $\mathbb{R}$. We will call this operator the is the *$p-$variation along* $\mathfrak{p}$, and in the case $p = 2$ the *pathwise variance along* $\mathfrak{p}$.

Of particular interest is the case $p = 2$, $d = 1$; in that case we may define the *pathwise p-co-variance along* $\mathfrak{p}$ as

$$[x, y]_{\mathfrak{p}} := \sum_{k=0}^{m(\mathfrak{p})} x_k^{\mathfrak{p}}.y_k^{\mathfrak{p}} \, \beta[\mathfrak{p}_k] \tag{3.38}$$

and note that it satisfies a *polarization identity*: $[x, y]_{\mathfrak{p}} = \frac{1}{4}\big(\mathrm{var}_{\mathfrak{p},2}(x + y) + \mathrm{var}_{\mathfrak{p},2}(x - y)\big)$. We may extend this definition to the general $\mathbb{R}^d$ case via the matrix $[x, y]_{\mathfrak{p}}$ with i, j component given by

$$[x, y]_{\mathfrak{p}}^{ij} := [x^i, y^j]_{\mathfrak{p}} \tag{3.39}$$

For a given clock $\mathscr{P} = (\mathfrak{p}^n)$, any $x \in \mathbb{S}_r^{\mathscr{P}}(\mathbb{R}^d)$ is in $\mathbb{S}^{\mathfrak{p}^n}(\mathbb{R}^d)$ for sufficiently high n; we may then define a corresponding operator via

$$\mathrm{var}_{\mathscr{P},p}(x)(t) = \lim_{n\to\infty} \mathrm{var}_{\mathfrak{p}^n,p}(x)(t) \tag{3.40}$$

where the limit is taken pointwise in t, again defines an increasing path, and always exists for step paths since the sequence is eventually constant. Taking the pth root of the final value gives us a norm

$$||x||_{\mathscr{P},p} := \left(\mathrm{var}_{\mathscr{P},p}(x)(T)\right)^{\frac{1}{p}} \tag{3.41}$$

on $\mathbb{S}_r^{\mathscr{P}}(\mathbb{R}^d)$. It is then natural to consider its completion, and similarly for the càglàd and regulated step paths. Since we wish to retain the uniform convergence as well, this motivates the following fundamental definition.

Definition 3.10. The *right* (resp. *left*, resp. *regulated*) $(\mathscr{P}, p)-$*Monte Carlo path space* $\mathbb{M}_{\mathscr{P}}^p(\mathbb{R}^d)_r$ (resp. $\mathbb{M}_{\mathscr{P}}^p(\mathbb{R}^d)_l$, $\mathbb{M}_{\mathscr{P}}^p(\mathbb{R}^d)_{reg}$) is the completion of $\mathbb{S}_r^{\mathscr{P}}(\mathbb{R}^d)$ (resp. $\mathbb{S}_l^{\mathscr{P}}(\mathbb{R}^d)$, $\mathbb{S}^{\mathscr{P}}(\mathbb{R}^d)$) in the norm

$$||.||_{\mathscr{P},p} := \max(||.||_\infty, ||.||_{\mathscr{P},p}) \tag{3.42}$$

The *continuous* $(\mathscr{P}, p)-$*Monte Carlo path space* is $\mathbb{M}_{\mathscr{P}}^p(\mathbb{R}^d)_c := \mathbb{M}_{\mathscr{P}}^p(\mathbb{R}^d)_r \cap \mathbb{W}(\mathbb{R}^d)$.

These are obviously separable Banach spaces by construction, and since $||.||_p \leq ||.||_{p'}$ for $p' \leq p$, we have the inclusion $\mathbb{M}_{\mathscr{P}}^{p'}(\mathbb{R}^d)_l \subset \mathbb{M}_{\mathscr{P}}^p(\mathbb{R}^d)_l$. To avoid tedious repetition we will mainly work with càdlàg Monte Carlo path spaces from now on; obviously due to the canonical isomorphism parallel results hold also for càglàd. We will only really be interested in the cases $p = 1, 2$. It is obvious that the convergence of the 2-norm in this case implies its convergence for each component, and therefore the convergence of the co-variance matrix $[x, x]^{ij}$.

Perhaps surprisingly, given their ubiquity in numerical approximations of stochastic processes, these spaces are considered for the first time here as legitimate functional analytic objects in their own right. The novelty of this definition, as opposed to for instance the analysts' notion of *finite p-variation* paths, is that firstly we are explicitly fixing the variation to be calculated along the clock $\mathscr{P}$, and secondly we also require the uniform convergence of the step path approximations. Although these may seem to be seriously restrictive, it turns out not to be, at least for continuous models.

Since the Euler projections are $\mathscr{P}$-step path approximations of given càdlàg/càglàd paths, we have the following characterisation of the $(\mathscr{P}, p)-$Monte Carlo space.

Proposition 3.11. *A path x is in $\mathbb{M}_{\mathscr{P}}^p(\mathbb{R}^d)_r$ if and only if $x \in \mathbb{D}_{\mathscr{P}}(\mathbb{R}^d)$ and*

$$\lim_{n\to\infty} \sum_{k=1}^{m(\mathfrak{p}^n)} |\delta_k^{\mathfrak{p}^n} x|^p \tag{3.43}$$

exists.

Proof. Any $x \in \mathbb{M}_{\mathscr{P}}^p(\mathbb{R}^d)_r$ is the uniform limit of $\mathscr{P}$-step paths, which we may take to be the corresponding Euler projections. The expression 3.43 is simply the p-norm of the nth Euler projection, and so its convergence is the other condition for membership. □

The main reason we are interested in Monte Carlo path spaces is its relationship to the *quadratic variation* of stochastic processes, as shown in the following definition and Theorem.

Definition 3.12. (Probabilist's quadratic variation) A stochastic process X has *quadratic variation process* $[X]$ if for each $\varepsilon > 0$ there exists $\delta > 0$ such that if a partition $\mathfrak{q}$ has mesh less than δ, then

$$\mathbb{P}[|\sum_{k=0}^{m(\mathfrak{q})} |X(\omega, \mathfrak{q}_{k+1}) - X(\omega, \mathfrak{q}_k)|^2 - [X](\omega, t)| \geq \varepsilon] < \varepsilon \tag{3.44}$$

Theorem 3.13. *Let $\mathbb{P}$ be a probability measure on $\mathbb{D}^d$ such that the process $X(\omega, t)$ is a square integrable continuous semimartingale. Then for any fixed clock $\mathscr{P}$ with summable mesh $\sum |\mathfrak{p}^n| < \infty$,*

$$\mathbb{P}[\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_c] = 1 \tag{3.45}$$

and

$$var_{\mathscr{P},2}(X(\omega, .)) = [X](\omega, .) \tag{3.46}$$

$\mathbb{P}$*-almost surely.*

Proof. Since X has almost surely continuous paths on the compact interval $[0, T]$, they are uniformly continuous; hence the Euler projections $\mathcal{E}^{\mathfrak{p}^n}(X) := X(\omega, 0) + \sum_{k=1}^{m(\mathfrak{p}^n)} \delta_k^{\mathfrak{p}^n} X(\omega, .)\beta[\mathfrak{p}_k]$ almost surely converge uniformly. Fix t. By definition, $|X(\omega)_t|_{\mathfrak{p}^n,2} := |\sum_{\mathfrak{p}_k \leq t} |X(\omega, \mathfrak{p}_{k+1}) - X(\omega, \mathfrak{p}_k)|^2$, so for any $\varepsilon > 0$ we may choose sufficiently large n so that $\mathbb{P}[|X(\omega)_t|_{\mathfrak{p}^n,2} - [X](\omega, t)| \geq \varepsilon] < \varepsilon$. Passing to a subsequence if necessary and taking the limit completes the proof. □

It should be emphasised that the clock $\mathscr{P}$ must be fixed independently of the paths; as shown by Davis et al. (2018), by allowing it to vary we may construct a version of the quadratic variation that is almost surely maximally badly behaved, in a paerticular sense.

The point of Theorem 3.13 is that we may use $\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)$ as a *state space* for continuous market risk models. But why *should* we use these as state spaces, rather than $\mathbb{W}^d$ or $\mathbb{D}^d$? First, Monte Carlo spaces exclude the many paths in these classical state spaces that violate an essential property of the models, namely having finite quadratic variation. Second, that the topological structure of Monte Carlo spaces by construction reflects this. This is illustrated by the following result, a pathwise version of a result from stochastic processes, with essentially the same proof.

Lemma 3.14. *For any clock $\mathscr{P} = (\mathfrak{p}^n)$, if $x \in \mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_c$ then*

$$||x||_{\mathscr{P},2} = 0 \tag{3.47}$$

Proof. For any n and path x we have $\sum_{k=0}^{m(\mathfrak{p}^n)} |\delta_k^n x|^2 \leq \max_k |\delta_k^n x| \text{ var}_{\mathfrak{p}^n,1}(x)$. If $x \in \mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_{reg}$, then the second term converges to $||x||_{\mathscr{P},1}$, while if x is continuous, it is uniformly continuous on $[0, T]$, and hence the first term converges to zero. □

This gives us the ammunition to ensure that the pathwise co-variation really is an inner product.

Theorem 3.15. *For any clock $\mathscr{P}$,*

$$\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_r \big/ \mathcal{Z} \tag{3.48}$$

is a Hilbert space, with inner product

$$[x, y]_{\mathscr{P}} := \lim_{n\to\infty} [x, y]_{\mathfrak{p}^n} \tag{3.49}$$

where $\mathcal{Z} := \{x \in \mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_r \ : \ ||x||_{\mathscr{P},2} = 0\}$.

Proof. Since the semi-norm $x \to ||x||_{\mathscr{P},2}$ is continuous on $\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_r$, $\mathcal{Z}$ is a closed subspace and we may consider the quotient Banach space $\mathcal{Q} = \mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_r + \mathcal{Z}$, consisting of cosets $x + \mathcal{Z}$ and endowed with the norm

$$||x + \mathcal{Z}||_{\mathcal{Q}} = \inf_{z\in\mathcal{Z}} ||x - z||_{\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_r} \tag{3.50}$$

It remains to show this norm is Hilbertian. To that end, note that for any $z \in \mathcal{Z}$, $||x - z||_{\mathscr{P},2} \leq ||x||_{\mathscr{P},2} + ||z||_{\mathscr{P},2} = ||x||_{\mathscr{P},2}$. By Lemma 3.14, $\mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_c \subset \mathcal{Z}$. Any step path $s \in \mathbb{S}^{\mathscr{P}}(\mathbb{R}^d)$ can be uniformly approximated by a sequence in $\mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_c$; for instance, a linear interpoliation works. Hence $\mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_c$ is dense in $\mathbb{D}_{\mathscr{P}}(\mathbb{R}^d)$, and for any $x \in \mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_r$ there exists $z_n \in z \in \mathcal{Z}$ such that $||x-z_n||_\infty \to 0$. Hence $||x+\mathcal{Z}||_{\mathcal{Q}} = ||x||_{\mathscr{P},2}$. Since this latter is the limit of a sequence of Hilbertian seminorms, by considering the polarization identity on each and taking the limit, we see it is Hilbertian. □

Of particular interest is the set of paths which have finite pathwise variance along a trading clock:

$$\mathbb{V}^d_\tau := \{x \in \mathbb{D}^d \ : \ ||x||_{\tau(x),2} < \infty\} \tag{3.51}$$

These form an effective state space for jump processes, under a Skorokhod style topology. However, we will focus on continuous models in this thesis.

3.3 Integration along a clock and pathwise versions of market models

We now begin the process of exploring the structure of these spaces. First, it turns out that *finite absolute variation* Monte Carlo paths, i.e. $\mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)$, serve as the natural dual objects to the uniform limit paths, $\mathbb{D}^d_{\mathscr{P}}, \mathbb{F}^d_{\mathscr{P}}, \mathbb{G}^d_{\mathscr{P}}$. The pairing is of course via integration, though unlike in the Riemann-Stieltjes theory we will allow our 'integrals' to depend on the sequence of partitions, and explicitly construct them as paths.

Definition 3.16. (Pathwise integral along a clock)

The *integral along* a partition $\mathfrak{p}$ of a path $f \in \mathbb{F}^d$ with respect to $x \in \mathbb{D}^d$ is the $\mathbb{R}$-valued step path given by

$$\int_0^t f\, d^{\mathfrak{p}} x := \sum_{k=0}^{m(\mathfrak{p})} (f(\mathfrak{p}_k), \delta_k^{\mathfrak{p}} x)_{\mathbb{R}^d} \beta[\mathfrak{p}_k](t) \tag{3.52}$$

We say f is *integrable along* $\mathscr{P} = (\mathfrak{p}^n)$ *in* $\mathbb{M}^p_{\mathscr{P}}(\mathbb{R}^d)_{\text{reg}}$ if

$$\int_0^t f\, d^{\mathscr{P}} x := \lim_{n\to\infty} \int_0^t f\, d^{\mathfrak{p}^n} x \tag{3.53}$$

exists in $\mathbb{M}^p_{\mathscr{P}}(\mathbb{R}^d)_{\text{reg}}$.

Note that by construction, $f \in \mathbb{F}^1$ is *Riemann-Stieltjes integrable* with respect to x if and only if it is integrable with respect to x along any clock $\mathscr{P}$, and this integral is independent of the choice of $\mathscr{P}$, i.e. is invariant under the action of Λ. So this integral represents a generalisation of the ordinary Riemann-Stijetles theory. Furthermore, if f is integrable in $\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_{\text{reg}}$, then by construction

$$\begin{aligned}
\text{var}_{\mathscr{P},p}\Big(\int_0^{\cdot} f\, d^{\mathscr{P}} x\Big)(t) &= \lim_{n\to\infty} \sum_{k=0}^{m(\mathfrak{p}^n)} |(f(\mathfrak{p}_k^n), \delta_k^{\mathfrak{p}^n} x)_{\mathbb{R}^d}|^2 \beta[\mathfrak{p}_k^n](t) \\
&= \lim_{n\to\infty} \sum_{k=0}^{m(\mathfrak{p}^n)} \sum_{i,j=1}^{d} f^i(\mathfrak{p}_k^n) f^j(\mathfrak{p}_k^n) \delta_k^{\mathfrak{p}^n} x^i \delta_k^{\mathfrak{p}^n} x^j \beta[\mathfrak{p}_k^n](t)
\end{aligned} \tag{3.54}$$

for each t.

This notion of integral gives us the following duality results.

Lemma 3.17. *For a fixed clock $\mathscr{P} = (\mathfrak{p}^n)$, and any $F \in \mathbb{F}_{\mathscr{P}}(\mathbb{R}^d)^*$, there exists a unique $f \in \mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_{reg}$ such that*

$$F(q) = q(0)f(0) + \int_0^T q\, d^{\mathscr{P}} f \tag{3.55}$$

for all $q \in \mathbb{F}^1_{\mathscr{P}}$, and $||F|| = var_{\mathscr{P},1}(f)$.

Proof. We prove the case $d = 1$ for simplicity of notation; the general case is similar.

Let $F \in \mathbb{F}^{1*}_{\mathscr{P}}$, $g_k^n = 1_{[0,\mathfrak{p}_k^n]}$, and $f_k^n = F(g_k^n)$. Define the step paths

$$f_n(t) := f_0^n \alpha(0) + \sum_{k=1}^{m(\mathfrak{p}^n)} f_k^n (g_k^n(t) - g_{k-1}^n(t)) \tag{3.56}$$

clearly $||f_n||_\infty \leq ||F||$. Moreover, by the continuity and linearity of F,

$$\begin{aligned}
\mathrm{var}_{\mathfrak{p}^n,1}(f_n) &= |f_0^n| + \sum_{k=1}^{m(\mathfrak{p}^n)} |f_k^n - f_{k-1}^n| \\
&= F\big((\mathrm{sgn} f_0^n) g_0^n + \sum_{k=1}^{m(\mathfrak{p}^n)} \mathrm{sgn}(f_k^n - f_{k-1}^n)(g_k^n - g_{k-1}^n)\big) \\
&\leq ||F||_{\mathbb{F}^{1*}_{\mathscr{P}}} \qquad (3.57)
\end{aligned}$$

so $f \in \mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_{\mathrm{reg}}$, and $||f|| \leq 2||F||$. If

$$s := s(0) g_0^n + \sum s(\mathfrak{p}_k^n)(g_k^n - g_{k-1}^n) \qquad (3.58)$$

is a $\mathfrak{p}^n$ step path, then by continuity and linearity,

$$\begin{aligned}
F(s) &= s(0) F(g_0^n) + \sum s(\mathfrak{p}_k^n)(F(g_k^n) - F(g_{k-1}^n)) \\
&= s(0) F(g_0^n) + \int_0^t s d^{\mathscr{P}} f \qquad (3.59)
\end{aligned}$$

and $|F(s)| \leq ||s|| \mathrm{var}_{\mathscr{P},1}(f)$. Since n was arbitrary, and such step paths are dense in $\mathbb{F}^1_{\mathscr{P}}$, by continuity of F we are done. □

Corollary 3.18. *We have* $\mathbb{F}_{\mathscr{P}}(\mathbb{R}^d)^* \cong \mathbb{D}_{\mathscr{P}}(\mathbb{R}^d)^* \cong \mathbb{G}_{\mathscr{P}}(\mathbb{R}^d)^* \cong \mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_{reg}$.

Proof. Since any regulated $g \in \mathbb{G}_{\mathscr{P}}(\mathbb{R}^d)$ can be written as a sum $g = g_- + \Delta_- g$, we need only to extend a linear functional $F \in \mathbb{F}_{\mathscr{P}}(\mathbb{R}^d)^*$ to $\mathbb{G}_0(\mathbb{R}^d)$. The necessary pairing is given by

$$F(g) = g(0) f(0) + \int_0^T g_- d^{\mathscr{P}} f + \sum_{u \in [0,T]} \Delta_- g(u) \Delta_- f(u) \qquad (3.60)$$

where $f \in \mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_{\mathrm{reg}}$ is the path obtained from Lemma 3.17. □

The Monte Carlo spaces are of course by construction built out of skeletons with respect to a fixed clock $\mathscr{P}$. We may now relate the above new concepts and results to the classical concept of *total variation* and *bounded variation paths*, which are independent of a clock.

Definition 3.19. (Total variation)

The *total variation* of a path x is the path

$$\mathcal{V}(x)(t) := \sup_{\mathfrak{p} \in \mathcal{P}} \mathrm{var}_{\mathfrak{p},1}(x) \qquad (3.61)$$

where the supremum is over all partitions $\mathfrak{p}$ of $[0, T]$.

If $\mathcal{V}(x)(T) < \infty$, then x is said to have *bounded q-variation*; the space of bounded variation paths is written $BV(\mathbb{X})$, and endowed with the *variation norm*:

$$||x||_{BV} := ||x(0)|| + \mathcal{V}(x)(T) \tag{3.62}$$

It is a classical fact that $BV(\mathbb{X})$ is a Banach space under the variation norm, in which the space of step paths is dense. Another important fact is that a real-valued path has bounded variation if and only if it is the difference of two increasing functions. In particular, since for $x \in \mathbb{M}^p_{\mathscr{P}}(\mathbb{R}^d)_{reg}$ the p-variation $\mathrm{var}_{\mathscr{P},p}(x)(.)$ is bounded and increasing, it has bounded variation.

The following Theorem is new in its specific form (see below for a more detailed discussion).

Theorem 3.20. *The dual spaces $\mathbb{G}(\mathbb{R}^d)^* \cong \mathbb{D}(\mathbb{R}^d)^* \cong \mathbb{F}(\mathbb{R}^d)^* \cong BV(\mathbb{R}^d)$. Furthermore, $BV(\mathbb{R}^d)$ is invariant under the action of Λ given in Theorem 3.9, and any $x \in BV(\mathbb{R}^d)$ has the decomposition*

$$x = C(x) + \sum_{u\in[0,T]} \Delta_{-}x(u)\beta[u] + \Delta_{+}x(u)\gamma[u] \tag{3.63}$$

with $C(x) \in \mathbb{W}(\mathbb{R}^d)$, where the sum has countably many non-zero terms and converges absolutely.

Proof. By Theorem 3.9 the spaces $\mathbb{G}(\mathbb{R}^d), \mathbb{D}(\mathbb{R}^d), \mathbb{F}(\mathbb{R}^d)$ are the unions of the respective clock-specific spaces $\mathbb{G}_{\mathscr{P}}(\mathbb{R}^d), \mathbb{D}_{\mathscr{P}}(\mathbb{R}^d), \mathbb{F}_{\mathscr{P}}(\mathbb{R}^d)$. It is clear from the definition that

$$BV(\mathbb{R}^d) = \bigcap_{\mathscr{P}} \mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_{reg} \tag{3.64}$$

where the intersection is over all clocks, and the norm $||x||_{BV}$ is equivalent to $\sup_{\mathscr{P}} ||.||_{\mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_{reg}}$. The first and second statements then follow from Lemma 3.18. The formula 3.63 is then simply due to the absolute convergence of the jump part. □

This result is of independent interest in functional analysis, and in this form strictly speaking new. Kaltenborn (1934) was the first to consider that problem, and found $\mathbb{G}^1 \cong BV^1 \oplus \ell^1$; later Webb (1973) proved the result for $\mathbb{F}^1$, while Tvrdy (1996) proved a corresponding result for *regular* (also known as *equiregulated*) paths, satsifying $x = \frac{1}{2}(x_+ + x_-)$. In the more general case of $\mathbb{X}$ being a Banach algebra, Berberian (1978) used the Gelf'and Naimark theorem to prove $\mathbb{G}(\mathbb{X})$ is isomorphic to the continuous functions on a (very badly behaved) topological space, and consequently characterised $\mathbb{G}(\mathbb{X})^*$ as a space of Radon measures by the Riesz representation Theorem. However, none has prior to this characterised the dual as *solely* the bounded variation paths. In effect, our proof simply shows that Kaltenborg could have brought his additional ℓ^1 summand term into the definition of the BV function corresponding to a general continuous linear functional, by defining the pairing as we have.

The point about this characterisation of the dual of the càglàd paths is that it allows us to relate the pathwise variance to Föllmer's quadratic variation. Recall that a path $x : [0,T] \to \mathbb{R}$ is said to have *finite quadratic variation in the Föllmer sense* if the atomic measures

$$\xi_n := \sum_{t_i \in \mathcal{T}_n} (x(t_{i+1}) - x(t_i))^2 \delta_{t_i} \tag{3.65}$$

converge *weakly* to a Radon measure on $[0,T]$ with distribution function $[x]$, whose atomic part coincides with the square of the jumps of x. That is, if

$$\int_{[0,T]} f(u)\, \xi_n(du) \to \int_{[0,T]} f(u)\, [x](du) \tag{3.66}$$

for all continuous $f : [0,T] \to \mathbb{R}$ (strictly speaking this is the *weak star* topology that the space of Radon measures has due to its dual pairing with the continuous paths), and

$$[x](t) = [x]_c(t) + \sum_{u \in [0,t]} (\Delta x(u))^2 \tag{3.67}$$

where $[x]_c$ is continuous. Comparing 3.65 to 3.37, we see that the pathwise variance along a partition is the distribution function of Follmer's atomic measure. The point of the next result is that, rather than the weak-star convergence as a measure based on the dual pairing with continuous functions, the pathwise variance is the weak-star convergence of the bounded variation distribution function based on the dual pairing with $\mathbb{F}^d$.

Corollary 3.21. *A càdlàg path $x \in \mathbb{D}_{\mathscr{P}}$ is in $\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_r$ if and only if the pathwise co-variance matrix satisfies*

$$\int_0^T f(u)[x,x]^{ij}_{\mathfrak{p}^n}(du) \to \int_0^T f(u)[x,x]^{ij}_{\mathscr{P}}(du) \tag{3.68}$$

for any $f \in \mathbb{F}^d_{\mathscr{P}}$

Proof. This is immediate once we recognise that the weak-star topology on $\mathbb{M}^1_{\mathscr{P}}(\mathbb{R}^d)_r$ is simply pointwise convergence at the points of the clock:

$$\int_0^T 1_{[0,t]}(u)[x,x]^{ij}_{\mathfrak{p}^n}(du) = [x,x]^{ij}_{\mathfrak{p}^n}(t) - [x,x]^{ij}_{\mathfrak{p}^n}(0) = [x,x]^{ij}_{\mathfrak{p}^n}(t) \tag{3.69}$$

so pointwise convergence coincides with integration against the paths $1_{[0,t]}$, $t \in \mathscr{P}$. Since the latter are dense in $\mathbb{F}^d_{\mathscr{P}}$ we are done. □

The role of a market risk model can be illustrated by the following schematic:

$$\textit{market data} \;+\; \textit{“noise”} \;\underset{MODEL}{\longrightarrow}\; \textit{future market scenarios}$$

The model's job is to take some data on the historical evolution of the market and some suitably random 'noise', to project a plausible *future* evolution of the market. We will now introduce a quite general class of continuous market risk models, and then construct for any such model a set of paths of full probability. First we make the following fundamental definition.

Definition 3.22. The set of *homogenous noise* paths in $\mathbb{R}^d$ is

$$\Omega_v := \{\omega \in \mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_c \ : \ [\omega, \omega](t) = t\} \tag{3.70}$$

Theorem 3.23. *(Pathwise Itô isometry)*

For $f \in \mathbb{F}$ and any homogenous noise path ω_v, $\int_0^{\cdot} f d^{\mathscr{P}}\omega \in \mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_c$, and

$$var_{\mathscr{P},p}(\int_0^{\cdot} f d^{\mathscr{P}}\omega) = \int_0^{\cdot} |f(u)|^2 du \tag{3.71}$$

Proof. This is immediate from the definition of homogenous noise paths, the Lebesgue integrability of the regulated path $|f(u)|^2$, and equation 3.54. □

We will now consider a class of market models known as *Itô processes*.

Definition 3.24. (Itô process)

An *Itô process* is a stochastic process $X : \Omega \times [0,T] \to \mathbb{R}^d$ on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$$X(t) = X(0) + \int_0^t u(s,\omega) ds + \int_0^t v(s,\omega) dW(s) \tag{3.72}$$

where W is a $\mathbb{P}$-Brownian motion, u, v are stochastic processes adapted to $\mathcal{F}_t$ and

$$\mathbb{P}[\int_0^t |u(s,\omega)| ds + \int_0^t |v(s,\omega)|^2 ds] = 1 \tag{3.73}$$

Strictly speaking, we can weaken the assumptions to allow u, v to be adapted to a (larger) filtration, as long as W remains a martingale under the enlargement (see Oksendal (1992)) but this assumption will be sufficient for our purposes.

Theorem 3.25. *Let X be an Itô process model, and the function $\tilde{X} : [0,T] \times \Omega_v \to \mathbb{R}^d$ be given by*

$$\tilde{X}(t,\omega) := X(0,\omega) + \int_0^t u(s,\omega) ds + \int_0^t v(s,\omega) d^{\mathscr{P}}\omega(s) \tag{3.74}$$

Then $\mathbb{P}[\omega \ : \ \exists t, X(t) \neq \tilde{X}(t)] = 0$

Proof. Since W is a Brownian motion, $\mathbb{P}[\Omega_v] = 1$; on this set the expression 3.72 reduces to 3.74. □

Chapter 4

Calculus of Monte Carlo Path Functionals and the 'Greeks'

In this Chapter we address Objective 2 of the thesis Aims from Section 1.2, by embedding the Dupire (2009) and Cont and Fournié (2010) treatment of derivatives' risk sensitivities into the standard functional analytic context for differentiation, of *semigroup theory*.

In practice, derivatives' greeks are often calculated via 'bumping' procedures (Glasserman (2004)). For instance, suppose a derivative on a stock is valued by a functional V of time t and the path S_t of the spot price up to t. The sensitivity of the value to an instantaneous change in spot – the *delta* – will often be approximated by the difference quotient

$$\Delta_\varepsilon V(t, S_t) := \frac{V(t, S_t + \varepsilon_t) - V(t, S_t)}{\varepsilon_t} \tag{4.1}$$

where S_t denotes the path of the spot price up until t, and ε_t is a 'small perturbation' at t, in some sense.

The difference quotient is suggestive of a differential operation, and in the case of a Markov model the value functional reduces to an ordinary function of time and spot price at that time, and the limit $\lim_{\varepsilon \to 0} \Delta_\varepsilon V$ reduces to an ordinary partial differential, $\frac{\partial V}{\partial S}$.

The path-dependent case was first considered by Dupire (2009). As we saw in Chapter 2, he defined *non-anticipative functionals* as parameterised families of functions on a *union* of paths defined on increasing time intervals – i.e. with $F(t, .)$ defined on paths defined on $[0, t]$ – and developed his *functional Itô calculus* based on *vertical* and *horizontal differentials* as limits of similar expressions to 4.1 at each time. Cont and his co-authors then reformulated them as functions of *equivalence classes of paths* under the stopping operation, enabling the development of a weak theory of vertical differentiation and *Functional Kolmogorov Equations* (Cont et al. (2016)).

However, both approaches present considerable technical difficulties, and in particular neither is well-suited to the semigroup formalism of ordinary PDEs. For one, semigroups are families of operators on a single space that houses the boundary conditions. But the 'boundary conditions' of our problem – payoff functionals – are *not* non-anticipative functionals, and so are not in the natural domain spaces of operators built out of vertical differentials, as we would expect our semigroups to be. For another, non-anticipative functionals are defined on a complicated metric space of equivalence classes of paths, and *a priori* little is known about manipulating operators on such a space.

In this Chapter we will reformulate vertical and horizontal differentiation of non-anticipative functionals in such a way as lends itself to the functional analytic methods of semigroup theory. In 4.1 introduce some classes of functionals on Monte Carlo spaces, and give some key examples from finance. In Section 4.2 we will consider a natural idea of differentiation for functionals as the generator of a strongly continuous group of isometries, and show how it leads to several analogues of ordinary differential calculus results including a Taylor Theorem and a density result for functionals which are smooth in this sense. In Section 4.3 we show how when this differential an appropriate application of this differential to *functional processes* recovers the Dupire/Cont vertical differential, and derivatives risk sensitivities such as *delta* and *gamma*. Finally, in Section 4.4 we show that the Dupire horizontal differential is also the generator of a strongly continuous semigroup, and a representation for it as a limit of right time differentials on a sequence of finite dimensional subspaces.

4.1 Path functionals

The basic idea of functional analysis is *duality* – to capture the properties of a space through corresponding spaces of *functionals*, that is maps with range in the (complex) scalars. In this section we introduce some of the most important spaces of functionals on the path spaces defined in Chapter 3 above.

In this section Z denotes a generic Banach space. The space of continuous, bounded functionals on Z will be denoted $C_b(Z)$, and of *uniformly* continuous, bounded functionals $\mathscr{U}_b(Z)$; it is well known that each is a Banach space when endowed with the supremum norm $||.||_\infty$, and that these encode the (norm) topological properties of Z. The vector spaces of continiuous and uniformly continuous functionals on Z will similarly be written $C(Z), \mathscr{U}(Z)$, though in general these do not have a sensible norm structure.

A functional that satisfies

$$|\Phi(x) - \Phi(y)| \le l||x - y|| \tag{4.2}$$

for some $l > 0$ is said to be *Lipschitz continuous* or just simply *Lipschitz*, and the vector space of such functionals is written $\mathrm{Lip}(Z)$. The smallest such l is known as its *Lipschitz constant*, which we will write $l(\Phi)$. This is a semi-norm on $\mathrm{Lip}(Z)$ for which $l(\Phi) = 0 \iff \Phi \equiv c$, and so

$||\Phi||_{\text{Lip}} = |\Phi(0)| + l(\Phi)$ defines a norm on $\text{Lip}(Z)$, under which it is in fact a Banach space. The space of bounded Lipschitz functionals will be written $\text{Lip}_b(Z)$, and is a Banach algebra under the norm $||\Phi||_{\text{Lip}_b} := ||\Phi||_\infty + l(\Phi)$ and the obvious pointwise multiplication $\Phi.\Psi(x) := \Phi(x)\Psi(x)$ (see Weaver (1999) for more on Lipschitz spaces). Clearly, $\text{Lip}_b(Z) \subset \text{Lip}(Z)$ and $\text{Lip}_b(Z) \subset \mathcal{U}_b(Z) \subset C_b(Z)$.

In general it is natural to think of time series as right-continuous step paths, as we think of information being incorporated as times passes. Therefore the evolution of risk factors will be modelled by càdlàg paths $x \in \mathbb{D}^d$. In order to carry out a satisfactory pathwise analysis of the corresponding derivatives contracts written on these paths we will need to understand this space. First, it has a familiar measurability structure. For $t \in [0, T]$, the evaluation maps $\epsilon_t^i : \mathbb{D}^d \to \mathbb{R}^d$ defined by

$$\epsilon_t^i(x) = x^i(t) \tag{4.3}$$

are Lipschitz continuous, $i = 1, \dots, d$. These govern the relevant notion of *measurability* on $\mathbb{D}^d$; intuitively, we are interested in information that can be constructed from observations of the risk factors. We set $\mathcal{F}_t$ to be the sigma algebra generated by the set of evaluation functionals $\{\epsilon_t^i \mid 0 \le u \le t, i = 1, \dots, d\}$, and note that these sigma algebras form a *filtration*: $\mathcal{F}_t \subset \mathcal{F}_u$ if $t \le u$.

We would like to construct a sufficiently rich class of functionals on $\mathbb{D}^d$. The right-continuous Euler projections $\mathcal{E}_r^{\mathfrak{p}} : \mathbb{D}^d \to \mathbb{S}_r^{\mathfrak{p}}(\mathbb{R}^d)$ map paths to step paths, which are essentially finite dimensional objects, determined by the $m(\mathfrak{p})$-dimensional vector of their jumps at the partition points. Using these, we can construct a large class of functionals, in particular furnishing embedded copies of $C_b(\mathbb{R}^n)$ in $C_b(\mathbb{D}^d)$ for all n.

Definition 4.1. (Euler functionals)

Given a partition $\mathfrak{p}$, we define the *Euler operator* $\mathcal{S}^{\mathfrak{p}}$ mapping $\mathbb{R}^{m(\mathfrak{p})}$-dimensional functions to functionals on $\mathbb{D}^d$ by

$$\mathcal{S}^{\mathfrak{p}}(\varphi)(x) := \varphi\big(\mathcal{E}_r^{\mathfrak{p}}(x)(\mathfrak{p}_0), \Delta\mathcal{E}_r^{\mathfrak{p}}(x)(\mathfrak{p}_1), \dots, \Delta\mathcal{E}_r^{\mathfrak{p}}(x)(\mathfrak{p}_{m(\mathfrak{p})})\big) \tag{4.4}$$

for $x \in \mathbb{D}^d$. Conversely, for a functional $\Phi : \mathbb{D}^d \to \mathbb{R}$ we define the *Euler functional*

$$\Phi_*^{\mathfrak{p}} := \Phi \circ \mathcal{E}_r^{\mathfrak{p}} \tag{4.5}$$

Proposition 4.2.

For any partition $\mathfrak{p}$, the Euler operator $\mathcal{S}^{\mathfrak{p}}$ preserves continuity and uniformly continuity, and gives isometric injections $C_b(\mathbb{R}^{d\times(m(\mathfrak{p})+1)}) \hookrightarrow C_b(\mathbb{D}^d)$, $\mathcal{U}_b(\mathbb{R}^{d\times(m(\mathfrak{p})+1)}) \hookrightarrow \mathcal{U}_b(\mathbb{D}^d)$. Further, given a trading clock τ, if Φ is continuous we have $\Phi_*^{\tau^n} \to \Phi$ pathwise, and if $\Phi \in \mathcal{U}_b(\mathbb{D}^d)$, in norm.

Proof. We define $\mathcal{S}^{\mathfrak{p}}[\varphi](x) := \varphi\big(\mathcal{E}_r^{\mathfrak{p}}(x)(\mathfrak{p}_0), \Delta\mathcal{E}_r^{\mathfrak{p}}(x)(\mathfrak{p}_1), \ldots, \Delta\mathcal{E}_r^{\mathfrak{p}}(x)(\mathfrak{p}_{m(\mathfrak{p})})\big)$; clearly it is an algebra homomorphism. Since the Euler projection is surjective onto the right continuous $\mathfrak{p}$-step paths, which in turn is naturally isomorphic to $\mathbb{R}^{d\times(m(\mathfrak{p})+1)}$, the isometry property is similarly easy: $||\mathcal{S}^{\mathfrak{p}}[\varphi]|| = \sup_{x\in\mathbb{R}^{d\times(m(\mathfrak{p})+1)}} |\varphi(x)| = ||\varphi||$. The second claim follows from the fact that the Euler projections converge to the identity in the strong operator topology, and continuity/uniform continuity. □

Some concrete examples of *payoff functionals*, representing unit notional positions in derivatives contracts, are in order. For simplicity of exposition we take $d = 1$, and think of the single risk factor x as the log-price of a stock with spot price $S(t) = e^{x(t)}$.

Example 4.1. *(Payoff functionals)*

- *The 'vanilla'* European call *and* put option *payoffs at strike k with maturity T,*

$$\Phi_{k,T}^{Call}(S) = (S(T) - k)^+ \tag{4.6}$$

$$\Phi_{k,T}^{Put}(S) = (k - S(T))^+ \tag{4.7}$$

are Lipschitz continuous (see Figure 4.1) with $l(\Phi_{k,T}^{Call}) = l(\Phi_{k,T}^{Put}) = 1$.

VANILLA OPTION PAYOFFS

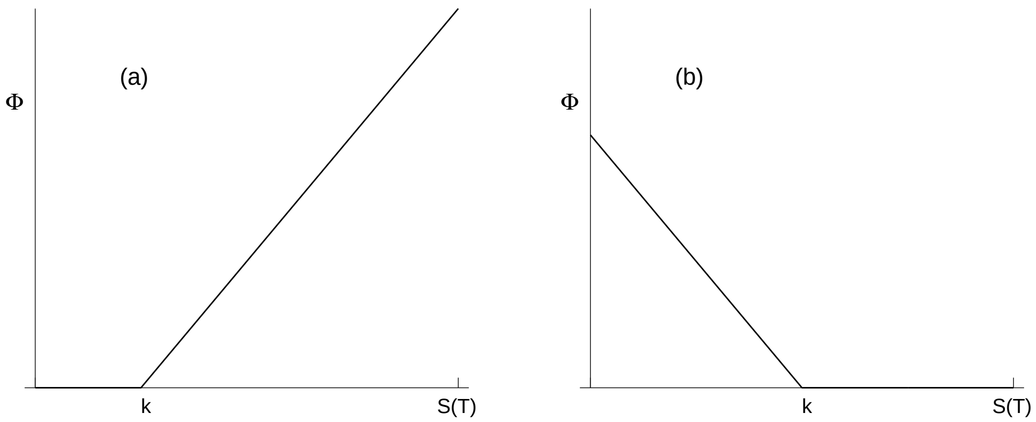

Figure 4.1: 'Vanilla' option payoffs: (a) call, (b) put.

- *Given $\varphi : \mathbb{R} \to \mathbb{R}$, the European payoff*

$$\Phi^{\varphi}(S) := \varphi(S(T)) \tag{4.8}$$

is continuous/resp. uniformly continuous/resp. Lipschitz/resp. measurable if and only if φ is continuous/resp. uniformly continuous/resp. Lipschitz continuous/resp. measurable

■ *The* digital *or* binary option *payoff*

$$\Phi_E^{Dig}(S) := 1_E(S) \tag{4.9}$$

where $E \in \mathcal{F}_T$, is $\mathcal{F}_T$ measurable, but not continuous. Usually we will only be interested in the case $E := \{x \mid x(t_i) \in [a_i, b_i]\}$, for some $t_i \in [0, T]$, $a_i < b_i \in \mathbb{R}$, $i = 1, \ldots, n$ (see Figure 4.2).

DIGITAL OPTION PAYOFF

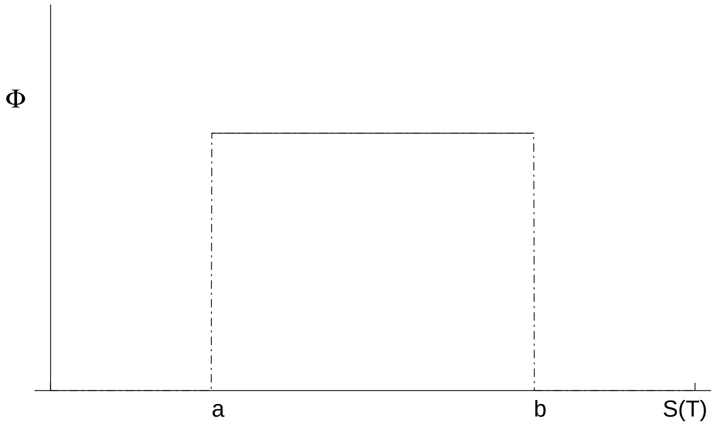

Figure 4.2: 'Digital' option payoff.

■ *The* up-and-out put option *payoff with barrier $B > 0$*

$$\Phi_{k,T,B}^{UOP}(S) = (k - S(T))^{+} 1_{\{S<B\}}(S) \tag{4.10}$$

is $\mathcal{F}_T$ measurable, but not continuous

■ *A single leg maturity T* variance swap *payoff struck at k is:*

$$\Phi_k^{VS}(S) := v^2(S) - k^2 \tag{4.11}$$

where

$$v^2(S) := \frac{1}{n} \sum_{i=1}^{n} \big(\log \frac{S(t_i)}{S(t_{i-1})} \big)^2 = \frac{1}{n} \sum_{i=1}^{n} \big(x(t_i) - x(t_{i-1}) \big)^2 \tag{4.12}$$

is the realised variance *functional, over the dates t_i, $i = 0, 1, \ldots, n$, $t_n = T$, is continuous, and uniformly continuous away from zero*

■ *An* Asian *option payoff is a functional depending on the (left limit of) the (continuously sampled, geometric)* average log-price*, given for $t \in [0, T]$ by*

$$A(t)(x) := \exp(\frac{1}{t} \int_0^t x(u)du) \tag{4.13}$$

For instance, the average strike call option *payoff is*

$$\Phi_T^A(S) = (S(T) - A(T-))^+ \tag{4.14}$$

Note that, since x is càdlàgit is bounded and Lebesgue almost everywhere continuous, and so the integral converges.

The payoffs of portfolios of derivatives can be represented as linear combinations; this is the basic idea behind *structuring*, building payoffs with specific characteristics for clients. For instance, a simple example is given by a *call spread* $\Phi^{CSpd}(S) := \Phi^{\text{Call}}_{k_1,T}(S) - \Phi^{\text{Call}}_{k_2,T}(S)$ where $k_1 \le k_2$ (see figure...). *Condors* are another class of option strategies, with the Lipschitz continuous payoff

$$\Phi^{Con}_{k_1,k_2,\varepsilon,T}(S) = \frac{1}{\delta}\big(\Phi^{\text{Call}}_{k_1-\delta,T}(S) - \Phi^{\text{Call}}_{k_1,T}(S) - \Phi^{\text{Call}}_{k_2,T}(S) - \Phi^{\text{Call}}_{k_2+\delta,T}(S)\big) \tag{4.15}$$

for $k_1, k_2 \in \mathbb{R}$, $\varepsilon > 0$ (see Figure 4.3).

CONDOR OPTION STRATEGY

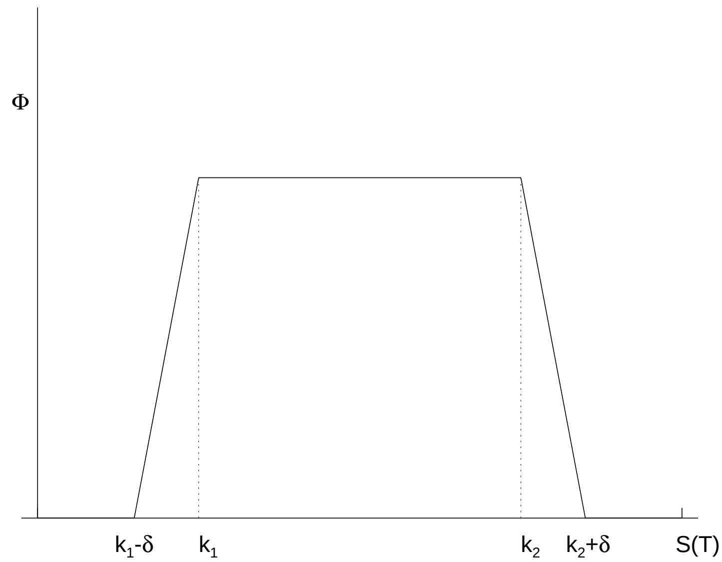

Figure 4.3: 'Condor' option strategy payoff.

These will be particularly useful for our purposes, as they can be thought of as basic building blocks for structuring, as shown in the following lemma. The idea is that by sending $\delta \to 0$ we may uniformly approximate any digital payoff, and hence any measurable functional under the L^p-norm (see Figure 4.4).

Lemma 4.3. *(Density of Lipschitz functionals in L^p)*

$Lip_b(\mathbb{D}^d)$ is dense in $L^p(\mathbb{P})$, $1 \le p < \infty$, for any probability measure $\mathbb{P}$ on the measurable space $(\mathbb{D}^d, \mathscr{F}_T)$.

APPROXIMATION OF DIGITAL BY CONDORS

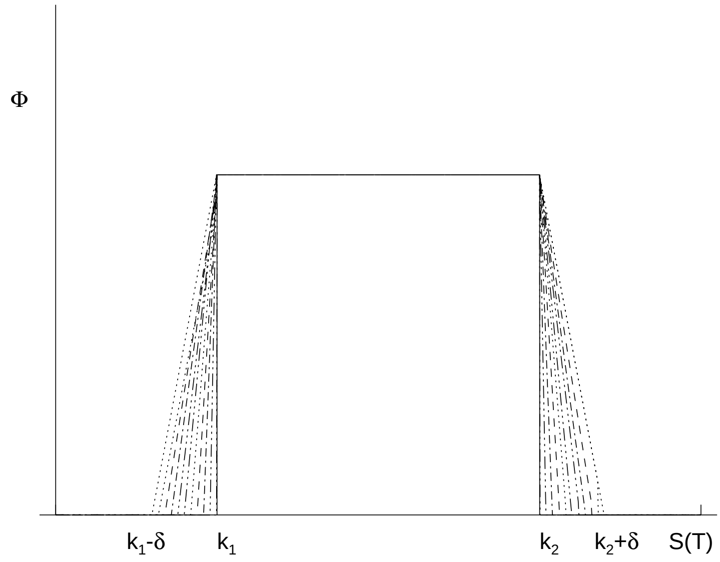

Figure 4.4: Uniform approximation of digital option payoff by condor strategies as $\delta \to 0$.

Proof. The sigma-algebra $\mathscr{F}_T$ is generated by the family of sets of the form

$$B(t_1, \ldots, t_n; a_1, b_1, \ldots a_n, b_n) := \{x \mid x(t_i) \in [a_i, b_i], i = 1, \ldots, n\} \tag{4.16}$$

and so linear combinations of the corresponding indicator functions are dense in L^p. These are given by

$$1_B(x) = \prod_{i=1}^{n} 1_{[a_i,b_i]}(x(t_i)) \tag{4.17}$$

For each $\epsilon > 0$, the bounded Lipschitz functional $\Phi^{Con}_{a,b,\epsilon,t}(x)$ converges monotonically to $1_{[a,b]}(x(t))$. By the Lebesgue monotone convergence theorem, we have

$$||1_{[a,b]} - \Phi^{Con}_{a,b,\varepsilon,t}||^p_p = \mathbb{E}[|1_{[a,b]} - \Phi^{Con}_{a,b,\varepsilon,t}|^p] \to 0 \tag{4.18}$$

as $\varepsilon \to 0$. By the algebra property of Lip_b, any product as in equation 4.17 can be similarly approximated by a suitable product of such Lipschitz approximating payoffs. □

Since $||.||_\infty \leq ||.||_{\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)}$, we have $\text{Lip}_b(\mathbb{D}^d) \subset \text{Lip}_b(\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d))$, so the previous result is true of $\text{Lip}_b(\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d))$ also.

Corollary 4.4. *(Density of bounded Lipschitz Euler functionals in L^p)*

If $\mathbb{P}$ is a probability measure on $(\mathbb{D}^d, \mathscr{F}_T)$ and $X \in L^p(\mathbb{P})$, then for each $\varepsilon > 0$ and any trading clock τ, there exists a bounded Lipschitz functional $\Phi \in Lip_b(\mathbb{D}^d)$ and $N \in \mathbb{N}$ such that $||X - \Phi^{\tau^n}_*||_p < \varepsilon$ for $n \geq N$. If $X = X(\omega, T)$ is the final value of a continuous semimartingale then *for each* $\varepsilon > 0$ *and any refining sequence of partitions* $\mathscr{P}$ *there exists* $\Phi \in Lip_b(\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d))$ *and* $N \in \mathbb{N}$*, such that* $||X - \Phi_*^{\mathfrak{p}^n}||_p < \varepsilon$ *for* $n \geq \mathbb{N}$.

Proof. This follows from Lemma 4.3, and Propositions 3.13, 4.2. □

4.2 Bumping groups and vertical differentiation

The real line has a group action on itself by *translation*, $t \to t + u$, which naturally lifts to a corresponding family of operators $tr(.)$ on functions, defined by $tr(u)f(t) = f(t+u)$. This furnishes a representation of the additive group $\mathbb{R}$ as *isometries* on the space $\mathscr{U}_b(\mathbb{R})$ that is *strongly continuous*, in the sense that $tr(u_n)f \to f$ in these spaces if $u_n \to 0$ in $\mathbb{R}$. The *differentiation* operator $\frac{d}{dx}$ is then the *infinitesimal action* of this group, in the sense that

$$\frac{d}{dx}f(x) := \lim_{\varepsilon \to 0}\frac{f(x+\varepsilon) - f(x)}{\varepsilon} = \lim_{\varepsilon \to 0}\frac{tr(\varepsilon)f(x) - tr(0)f(x)}{\varepsilon} \tag{4.19}$$

when this is well-defined. The classical *Taylor series* then allows us to recover the translation group as an 'exponential function' of differentiation,

$$f(x+u) = tr(u)f(x) = \sum_{k=0}^{\infty}\frac{u^k}{k!}(\frac{d}{dx})^k f\,(x) =: e^{u\frac{d}{dx}}f\,(x), \tag{4.20}$$

for sufficiently regular functions.

This then naturally generalises to $\mathbb{R}^d$, via the notion of *partial derivatives* $\frac{\partial}{\partial x_i}$ that generate translations in the direction of each of the Cartesian co-ordinates, out of which any general translation can be 'built'. In this Section we will see how this is paralleled in the path functional setting, by the notions of *bumping* groups that build translations out of vertical perturbations, and the corresponding *vertical differentiation* operators analogous to partial differentiation.

Again we will write Z for a generic Banach space, and $\mathcal{B}(Z)$ for the space of bounded linear operators on Z. Recall that the topology of pointwise convergence on $\mathcal{B}(Z)$, $A_n \to A \iff A_n z \to Az$ for all $z \in Z$, is known as the *strong operator topology* (SOT). The following definition is fundamental to the general theory of differentiation, and the rest of this thesis.

Definition 4.5. (Semigroups, groups of operators and orbit functions)

A *semigroup* of operators on a Banach space Z is a map $G : \mathbb{R}_+ \to \mathcal{B}(Z)$ satisfying

$$G(0) = I \tag{4.21}$$

$$G(u+v) = G(u)G(v) \text{ for all } u, v \in \mathbb{R}_+ \tag{4.22}$$

The *orbit function* of a semigroup G at $z \in Z$ is the map $\zeta_G^z : \mathbb{R}_+ \to Z$ defined by $\zeta_G^z(u) = G(u)z$. A *group* of operators on Z is a map $G : \mathbb{R} \to \mathcal{B}(Z)$, such that $G(u)$ is invertible for each u, satisfying

$$G(0) = I \tag{4.23}$$

$$G(u+v) = G(u)G(v) \text{ for all } u, v \in \mathbb{R} \tag{4.24}$$

$$G(u)^{-1} = G(-u) \tag{4.25}$$

The *orbit function* of a group G at $z \in Z$ is defined similarly $\zeta_G^z : \mathbb{R} \to Z$ by $\zeta_G^z(u) = G(u)z$.

A semigroup G of operators is said to be *strongly continuous* if it is right-continuous at zero in the SOT:

$$||G(u)z - z|| \to 0 \text{ as } u \downarrow 0 \tag{4.26}$$

for each $z \in Z$. Similarly, a group is strongly continuous if it is continuous at zero: $||G(u)z - z|| \to 0$ as $u \to 0$. Note that if G is a strongly continuous group of operators, then for $u \in \mathbb{R}_+$, $G^+(u) := G(u)$ and $G^-(u) := G(-u)$ are strongly continuous semigroups.

Much as Euclidean space has the canonical basis e^i, there is a natural set with which we can build generic paths in $\mathbb{G}^d$: the vertical bumps $\beta^i[t](u) := e^i 1_{[t,T]}(u)$, $\gamma^i[t](u) := e^i 1_{(t,T]}(u)$. It turns out that translation by the $\beta^i[t]$ underlies the idea of risk sensitivities, so we will focus on this. The corresponding *bumping* operator evaluates a functional on the 'bumped' path $x + \varepsilon\beta^i[t]$ (see figure 4.5).

Definition 4.6. (Vertical bumping operators)

For $t \in [0, T]$, the *bump operators* for coordinate i at time t are defined for functionals Φ on $\mathbb{G}^d$ by

$$B_t^i(u)\Phi\,(x) := \Phi(x + u\beta^i[t]) \tag{4.27}$$

Clearly, for each $t \in [0, T]$, $i = 1, \ldots, d$, the bump operator $u \to B_t^i(u)$ defines a group:

$$B_t^i(0)\Phi(x) = \Phi(x) \tag{4.28}$$

$$B_t^i(u+v)\Phi(x) = B_t^i(v)\Phi(x + u\beta^i[t]) = B_t^i(v)B_t^i(u)\Phi(x), \tag{4.29}$$

$$B_t^i(-u)B_t^i(u)\Phi(x) = \Phi(x) \tag{4.30}$$

Note also that for $t \neq t'$ the operators *commute*: $B_{t'}^i(u')B_t^i(u)\Phi(x) := \Phi(x + u\beta^i[t] + u'\beta^i[t']) = B_t^i(u)B_{t'}^i(u')\Phi(x)$.

The question then becomes: for what class of functionals is this group strongly continuous? To answer this, note that $||u\beta^i[t]||_\infty = |u|$. Meanwhile, $\delta_k^{\mathfrak{p}}(u\beta^i[t]) = 0$ unless $k = k(t)$, in which case $\delta_{k(t)}^{\mathfrak{p}}(u\beta^i[t]) = u$, so $||u\beta^i[t]||_{BV} = |u|$, and $||u\beta^i[t]||_{2,\mathscr{P}} = u^2$. This immediately gives us the following examples.

‘BUMPED’ PATH

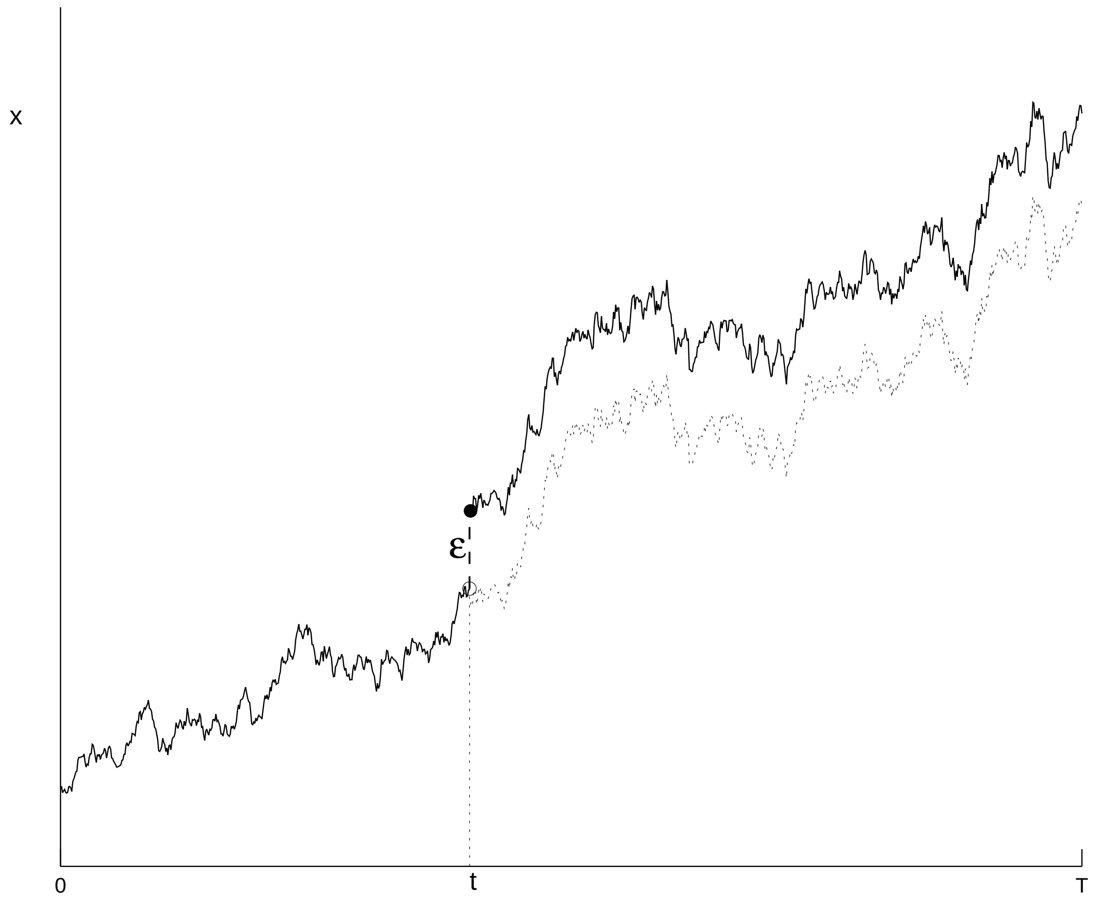

Figure 4.5: The action of a vertical bump function on a path.

Example 4.2. *For $t \in [0, T]$, $i = 1, \ldots, d$, the bumping operator B_t^i defines a strongly continuous group on*

- *each of $\mathcal{U}_b(\mathbb{G}^d)$, $\mathcal{U}_b(\mathbb{D}^d)$, $\mathcal{U}_b(BV^d)$ and $\mathcal{U}_b(\mathbb{M}^2_{\mathcal{P}}(\mathbb{R}^d))$, since by uniform continuity $||B_t^i(u)\Phi - \Phi||_\infty = \sup_x |\Phi(x + u\beta^i[t]) - \Phi(x)| \to 0$ as $|u| \to 0$.*
- *each of $Lip(\mathbb{G}^d), Lip(\mathbb{D}^d), Lip(BV^d)$, since for such functionals $|\Phi(x + u\beta^i[t]) - \Phi(x)| \leq l(\Phi)|u|$, so $l(B_t^i(u)\Phi - \Phi) \leq |u|l(\Phi)$, and hence $||B_t^i(u)\Phi - \Phi||_{Lip} = |\Phi(u\beta^i[t]) - \Phi(0)| + |u|l(\Phi) \leq 2|u|l(\Phi) \to 0$ as $|u| \to 0$.*
- *each of $Lip_b(BV^d), Lip_b(\mathbb{G}^d), Lip_b(\mathbb{D}^d)$, by combining the above two examples.*

We will call a space of path functionals on which B_t^i is a strongly continuous group for each $i = 1, \ldots, d$, a $B_t - space$, and denote a generic such space $\mathbb{Y}$ in the sequel.

It will often be useful to consider the functions $\zeta_t^{\Phi,x} : \mathbb{R} \to \mathbb{R}^d$, given by evaluating the orbit functions of the vertical bumping groups on a given path: $\zeta_t^{\Phi,x}(u) := \left(\zeta^{\Phi}_{B_t^i}(u)(x)\right) = \left(\Phi(x + u\beta^i[t])\right)$, which we will call the *vertical orbit function at* Φ, *x and t*. Note that the orbit functions inherit the continuity/uniform continuity/Lipschitz contintuity of the corresponding functional.

The analogue of differentiation for a semigroup of operators is given by the *generator*.

Definition 4.7. The generator A of a strongly continuous semigroup $G(.)$ of operators on a Banach space Z is the operator

$$Az := \lim_{\varepsilon \downarrow 0} \frac{G(\varepsilon)z - z}{\varepsilon} \tag{4.31}$$

with domain $dom(A) := \{z \in Z \mid Az \text{ exists in } Z\}$.

Note that this definition is specific to the space Z, so strictly speaking we must distinguish between the generators of the same (semi)group defined on different spaces. If G is a strongly continuous group of operators, then the generators A^+, A^- of G^+, G^- are like right- and left-differentials. Generators are so-called because of the following fundamental theorem (see Engel and Nagel (1999), Chapter II, Lemma 1.3 and Theorem 1.4), analogous to the fundamental theorem of calculus. We will need to make use of the theory of *Banach space-valued* or *Bochner* integration (Pisier (2016)).

Theorem 4.8. *The generator A of a strongly continuous semigroup G on a Banach space Z is a closed and densely defined linear operator on Z which determines G uniquely. In particular, for each $t \geq 0$, $z \in Z$, the integral operator $\int_0^t G(s)zds \in dom(A)$ and we have*

$$G(t)z - z = A\int_0^t G(s)zds = \int_0^t AG(s)zds = \int_0^t G(s)Azds \tag{4.32}$$

Equation 4.32 can be thought of as an analogue of the *Fundamental Theorem of Calculus* in the semigroup setting. We will not reproduce the whole proof here, but note that the first equation follows from the fact that

$$\begin{aligned}
\frac{1}{\varepsilon}\Big(G(\varepsilon)\int_0^t G(s)zds - \int_0^t G(s)zds\Big) &= \frac{1}{\varepsilon}\Big(\int_0^t G(s+\varepsilon)zds - \int_0^t G(s)zds\Big) && (4.33)\\
&= \frac{1}{\varepsilon}\int_\varepsilon^{t+\varepsilon} G(s)zds - \frac{1}{\varepsilon}\int_0^t G(s)zds && (4.34)\\
&= \frac{1}{\varepsilon}\int_t^{t+\varepsilon} G(s)zds - \frac{1}{\varepsilon}\int_0^\varepsilon G(s)zds && (4.35)\\
&\to G(t)z - z && (4.36)
\end{aligned}$$

as required. In particular $AG(u)z = G(u)Az$ for each $u \geq 0$; iterating this relation we see that $G(u)$ maps $dom(A^n)$ to itself – that is, $dom(A^n)$ is *invariant* – which we can think of as G 'preserving smoothness'.

In our setting, the fundamental differential operators for risk sensitivities are the generators of the vertical bumping groups B_t^i on a B-space $\mathbb{Y}$ – a directional differential *in the direction of vertical bumps*.

Definition 4.9. (Vertical differentiation)

The generator $\partial^i_{x_t}$ of the vertical bumping group B^i_t at t on a B-space $\mathbb{Y}$, defined as

$$\partial^i_{x_t}\Phi(x) := \lim_{\varepsilon\to 0}\frac{B^i_t(\varepsilon)\Phi(x)-\Phi(x)}{\varepsilon} = \lim_{\varepsilon\to 0}\frac{\Phi(x+\varepsilon\beta^i[t])-\Phi(x)}{\varepsilon} = \zeta^{\Phi\ i\ \prime}_{x,t}(0) \tag{4.37}$$

is called the ith *vertical differential at t*, defined on the domain

$$\mathbb{Y}^{1,i}_t := \{\Phi\in\mathbb{Y} \mid \partial^i_{x_t}\Phi(x)\in\mathbb{Y} \text{ for } i=1,\dots,d\} = \cap^d_{i=1}\mathbb{Y}^{1,i}_t \tag{4.38}$$

The *vertical gradient at t* puts these components together and maps functionals into the product space $\mathbb{Y}^d$. It will be written $\nabla_t := (\partial^i_{x_t})$, where $i = 1,\dots,d$, and has domain $\mathbb{Y}^1_t = \cap^d_{i=1}\mathbb{Y}^{1,i}_t$.

This naturally generalises to higher orders $k \geq 2$ of differentiability, via $\mathbb{Y}^k_t := \{\Phi\in\mathbb{Y} \mid \nabla_t\Phi(x)\in\mathbb{Y}^{k-1}_t\}$. A functional Φ is said to be *vertically smooth at t* if $\Phi\in\mathbb{Y}^{\infty,i}_t := \bigcap_{k\in\mathbb{N}}\mathbb{Y}^k_t$, and *vertically smooth* if it is vertically smooth at every $t\in[0,T]$. These are Banach spaces under the obvious graph norms. We denote the vector space of vertically smooth functionals by $\mathbb{Y}^\infty_{[0,T]}(\mathbb{D}^d) := \bigcap_{t\in[0,T]}\mathbb{Y}^\infty_t(\mathbb{D}^d)$.

Because the bumping operators form strongly continuous groups, we may avail ourselves of the well-developed theory of operator semigroups, and find that vertical differentials behave in many ways analogous to ordinary differentials on $\mathbb{R}^d$. Many of the following results are consequences of this theory; however, this theory is not well-known to many working in the finance literature, nor is our example known in the semigroup literature. Rather than a string of obscure-looking references to that literature, we will endeavour to provide direct proofs for the most important results we need.

It is easy to see that orbit functions inherit the degree of differentiability of the functionals. If $\Phi\in C^1_t$, then $\zeta^\Phi_{x,t}\in C^1(\mathbb{R},\mathbb{R}^d)$ and $\zeta^{\Phi\ i\ \prime}_{x,t}(u) = \lim_{\varepsilon\to 0}\frac{B^i[t](\varepsilon)B^i[t](u)-B^i[t](u)}{\varepsilon}\Phi(x) = B^i[t](u)\zeta^{\Phi\ i\ \prime}_{x,t}(0) = {}^i\nabla_t\Phi(x+u\beta^i[t])$. This seemingly trivial observation turns out to render calculus rules of vertical differentiation in our setting much easier to prove than in the Cont et al setting, where we lack a counterpart because the set of equivalence classes of stopped paths is not a vector space. A Leibniz product rule and a chain rule follow easily.

Proposition 4.10. *(Leibniz product rule for vertical differentials)*

Suppose $\Phi,\Psi\in C^1_t$, *and set* $\Xi(x)=\Phi(x)\Psi(x)$. *Then* $\Xi\in C^1_t$, *and* $\partial^i_{x_t}\Xi(x) = \Phi(x).\partial^i_{x_t}\Psi(x) + \Psi(x).\partial^i_{x_t}\Phi(x)$.

Proof. Fix a component for vertical differentiation, and for ease of notation embed this in a one dimensional model to avoid a proliferation of co-ordinate indices. Since $\Phi,\Psi\in C^1_t$, their orbit functions $\zeta^{\Phi}, \zeta^{\Psi} \in C^1(\mathbb{R})$, and so we have

$$\frac{B[t](\varepsilon)[\Phi.\Psi] - \Phi.\Psi}{\varepsilon} \equiv \frac{1}{\varepsilon}\big(B[t](\varepsilon)\Phi.(B[t](\varepsilon)\Psi - \Psi) + \Psi.(B[t](\varepsilon)\Phi - \Phi)\big) \tag{4.39}$$

$$= \zeta^{\Phi}_{x,t}(\varepsilon)\big(\frac{\zeta^{\Psi}_{x,t}(\varepsilon) - \zeta^{\Psi}_{x,t}(0)}{\varepsilon}\big) + \zeta^{\Psi}_{x,t}(0)\big(\frac{\zeta^{\Phi}_{x,t}(\varepsilon) - \zeta^{\Phi}_{x,t}(0)}{\varepsilon}\big) \tag{4.40}$$

$$\to \zeta^{\Phi}_{x,t}(0).\nabla\zeta^{\Psi}_{x,t}(0) + \zeta^{\Psi}_{x,t}(0).\nabla\zeta^{\Phi}_{x,t}(0) \tag{4.41}$$

$$= \Psi(x).\nabla_t\Phi(x) + \Phi(x).\nabla_t\Psi(x) \tag{4.42}$$

□

Proposition 4.11. *Suppose $F, G \in C^1_t$, $G(x) \neq 0$. Then $H := F \circ G \in C^1_t$, and*

$$\partial^i_{x_t} H(x) = \partial^i_{x_t} F(G(x))\partial^i_{x_t} G(x) \tag{4.43}$$

Proof. Again, we treat a one dimensional case for ease of notation. Since $F, G \in C^1_t$, their orbit functions $\zeta^F, \zeta^G \in C^1(\mathbb{R})$, and so we have for $|\varepsilon| > 0$,

$$F(x + \varepsilon\beta[t]) := \zeta^F_{x,t}(\varepsilon) = \zeta^F_{x,t}(0) + \varepsilon\zeta^{F\,\prime}_{x,t}(0) + \varepsilon f(\varepsilon) \tag{4.44}$$

$$= F(x) + \varepsilon\nabla_t F(x) + \varepsilon f(\varepsilon) \tag{4.45}$$

$$G(x + \varepsilon\beta[t]) := \zeta^G_{x,t}(\varepsilon) = \zeta^G_{x,t}(0) + \varepsilon\zeta^{G\,\prime}_{x,t}(0) + \varepsilon g(\varepsilon) \tag{4.46}$$

$$= G(x) + \varepsilon\nabla_t G(x) + \varepsilon g(\varepsilon) \tag{4.47}$$

where f, g are functions on $\mathbb{R}$ such that $\frac{f(\varepsilon)}{\varepsilon} \to 0$ as $\varepsilon \to 0$, and similarly for g. Without loss of generality, we may assume $f(0) = g(0) = 0$.

So

$$B[t](\varepsilon)H(x) - H(x) = F(G(x + \varepsilon\beta[t])) - F(G(x)) \tag{4.48}$$

$$= \zeta^F_{G(x),t}(\varepsilon(\nabla_t G(x) + g(\varepsilon))) - \zeta^F_{G(x),t}(0) \tag{4.49}$$

$$\begin{aligned} &= \zeta^F_{G(x),t}(0) + \zeta^{F\,\prime}_{G(x),t}(0)\varepsilon(\nabla_t G(x) + g(\varepsilon)) \\ &+ \varepsilon(\nabla_t G(x) + g(\varepsilon))f(\varepsilon(\nabla_t G(x) + g(\varepsilon))) - \zeta^F_{G(x),t}(0) \end{aligned} \tag{4.50}$$

and hence

$$\frac{B[t](\varepsilon)H(x) - H(x)}{\varepsilon} = \nabla_t F(G(x))\nabla_t G(x) + g(\varepsilon) + (\nabla_t G(x) + g(\varepsilon))f(\varepsilon(\nabla_t G(x) + g(\varepsilon)),$$

and all but the first term tends to zero as $\varepsilon \to 0$. □

In particular, the chain rule means we may freely change variables between spot prices S and log-spot prices x, transforming the vertical differential via $\partial^i_{x_t}\Phi(S) := S^i(t)\partial^i_{x_t}\Psi(x)$, where $\Psi(x) = \Phi(S(x))$, much as in the finite dimensional context.

Another very useful result is the following analogue of *Taylor's theorem with remainder*. This will be pivotal in our later proof of the path-dependent version of *Itô's lemma*.

Theorem 4.12. *(Vertical Taylor theorem)*

If $\Phi \in C^n_t$ *and* $u > 0$,

$$B[t](u)\Phi(x) = \sum_{k=0}^{n-1} \frac{u^k}{k!} \nabla^k_t \Phi(x) + \frac{1}{(n-1)!} \int_0^u (u-v)^{n-1} B[t](v) \nabla^n_t \Phi(x) dv \tag{4.51}$$

Proof. Since

$$\frac{d}{du} B[t](u)\Phi = \lim_{\varepsilon \to 0} B[t](u)\big(\frac{B[t](\varepsilon) - I}{\varepsilon}\big)\Phi = B[t](u)\nabla_t \Phi \tag{4.52}$$

we have $\left(\frac{d}{du}\right)^k (B[t](u)\Phi) = B[t](u)\nabla^k_t \Phi$. We may now repeatedly integrate equation 4.32 by parts:

$$B[t](u)\Phi(x) = \Phi(x) + \int_0^u B[t](s)\Phi(x) ds \tag{4.53}$$

$$= \Phi(x) + \big[(u-s)B[t](s)\nabla\Phi(x)\big]_{s=0,u} + \int_0^u (u-s)B[t](u)\nabla^2\Phi(x) ds \tag{4.54}$$

$$= \Phi(x) + u\nabla_t \Phi(x) + \big[\frac{1}{2}(u-s)^2 B[t](u)\nabla^2\Phi(x)\big]_{s=0,u} + \int_0^u \frac{1}{2}(u-s)^2 B[t](u)\nabla^2\Phi(x) ds \tag{4.55}$$

and so on. The nth iteration gives equation 4.51. □

Let us have some examples.

Example 4.3. *(Vertical differentiation of path functionals)*

- *The coordinate functionals* $\epsilon[t']$ *have* $\nabla_t \epsilon[t'](x) = 1_{[t',T]}(t) = \beta[t'](t)$, $\nabla^k_t \epsilon[t'] = 0$ *for* $k > 1$.

- *For a refining sequence of partitions* τ^n, *a* τ^n*-cylindrical functional* $\Phi^\tau_n(x) = \varphi(x(\tau^n_0), \ldots, x(\tau^n_{m(\tau^n)}))$ *is vertically differentiable at t if and only if* φ *is a differentiable function of the components* $x(\tau^n_k)$ *with* $\tau^n_k \geq t$, *in which case*

$$\nabla_t \Phi^\tau_n(x) = \sum_{k=0}^{m(\tau^n)} \partial_k \varphi(x(\tau^n))\, \beta[t](\tau^n_k(x)) \tag{4.56}$$

- *Suppose* $d = 1$, S *is the price of a stock, and* $S(T) \neq k$. *Then* Φ^{Call}, Φ^{Put} *are vertically differentiable at* S *for each* $t \in [0, T]$, *and*

$$\nabla_t \Phi^{Call}(S) = 1_{(k,\infty)}(S(T)) \tag{4.57}$$

and $\nabla_t \Phi^{Put}(S) = -1_{[0,k)}(S(T))$. *At paths with* $S(T) = k$, *we have,* $\nabla_t^+ \Phi^{Call}(S) = 1$, $\nabla_t^- \Phi^{Call}(S) = 0$, *and similarly for* Φ^{Put}.

- *The digital option payoff functional,* Φ_E^{Dig}, *for* $E := \{S \mid S(t_i) \in [a_i, b_i]\ i = 1, \dots, n\}$ *has* $\nabla_t \Phi_E^{Dig}(S) = 0$ *for all* t *at* S *such that* $S(t_i) \notin \{a_i, b_i\}$.

- *The* up-and-out put option *payoff with barrier* $B > 0$, $\Phi_B^{UOP}(S) = (k - S(T))^+ 1_{\{S<B\}}(S)$ *has* $\nabla_t \Phi_B^{UOP}(S) = -1_{[0,k)}(S(T))$ *for* S *such that* $S(u) \neq B$ *for all* u. *It has one-sided vertical differentials on paths that 'touch' the barrier. For instance, suppose a path* $S^* < B$ *except at time* u^* *where* $S^*(u^*) = B$. *Then* $\nabla_t^- \Phi_B^{UOP}(S^*) = \nabla_t^- \Phi^{Put}(S^*)$, *but* $\nabla_t^+ \Phi_B^{UOP}(S^*) = 0$.

- *Consider a variance swap single leg payoff functional* $\Phi_k^{VS}(S)$ *as in Example 4.1, and suppose* $t \in [t_{i^*}, t_{i^*+1})$. *The orbit function is* $\zeta_{x,t}^{VS}(\varepsilon) = \frac{1}{n}\sum_{i=1}^{i^*-1}\big(\log \frac{S(t_i)}{S(t_{i-1})}\big)^2 + \frac{1}{n}\big(\log \frac{S(t_i^*)+\varepsilon}{S(t_{i^*-1})}\big)^2 + \frac{1}{n}\sum_{i=i^*+1}^{n}\big(\log \frac{S(t_i)+\varepsilon}{S(t_{i-1})+\varepsilon}\big)^2 + k^2$, *so differentiating and setting* $\varepsilon = 0$, *we have*

$$\nabla_t \Phi_k^{VS}(S) = \frac{2}{n}\frac{1}{S(t_{i^*-1})}\log\frac{S(t_{i^*})}{S(t_{i^*-1})} + \frac{2}{n}\sum_{i=i^*+1}^{n}\Big(\frac{1}{S(t_i)} - \frac{1}{S(t_{i-1})}\Big)\log\frac{S(t_i)}{S(t_{i-1})} \tag{4.58}$$

$$= \frac{2}{n}\sum_{i=i^*}^{n}\frac{1}{S(t_{i^*-1})}(x(t_{i^*}) - x(t_{i^*-1})) - \frac{2}{n}\sum_{i=i^*}^{n}\frac{1}{S(t_{i^*})}(x(t_{i^*+1}) - x(t_{i^*})) \tag{4.59}$$

- *The average log-price to maturity functional* $A(T)(x) = \exp(\frac{1}{T}\int_0^T x(u)du)$ *has orbit function* $\zeta_{x,t}^{A(T)}(\varepsilon) = \exp(\frac{1}{T}\int_0^T (x(u) + \varepsilon 1_{[t,T]})du) = \exp(\frac{(T-t)\varepsilon}{T})A(T)(x)$, *so*

$$\nabla_t A(T)(x) = \frac{(T-t)}{T}A(T)(x) \tag{4.60}$$

We will say that $\Phi \in C_t^\infty$ is *vertically entire at* t, if the *vertical Taylor series* $\sum_{k=0}^{\infty}\frac{u^k}{k!}\nabla_t^k\Phi(x)$ converges for each $u > 0$. Much as in the finite dimensional case, it turns out such smooth objects are *dense* in C_b. The following lemma is based on a result of Gelfand for the general abstract case (Gel'fand (1939)).

Lemma 4.13. *For any* $\Phi \in C_b(\mathbb{D}^d)$, $\varepsilon > 0$ *and* $t \in [0, T]$, *there exists an entire functional* $\Phi_\varepsilon \in C_t^\infty(\mathbb{D}^d)$ *such that* $||\Phi - \Phi_\varepsilon|| < \varepsilon$.

Proof. The idea is to create a 'smoothed' version of Φ, by averaging the effect of a *Gaussian perturbation*, and using integration by parts. Define the operators $\mathcal{G}_t^\sigma : C_b \to C_b$ by

$$\mathcal{G}_t^\sigma(\Phi)(x) := \mathbb{E}[\Phi(x + \sigma Y\beta[t])] = \int_{-\infty}^{\infty} B[t](\sigma s)\Phi(x)\mathfrak{g}(s)ds \tag{4.61}$$

where Y is a standard normally distributed random variable, $\mathfrak{g}(s) := \frac{1}{2\pi}e^{-\frac{s^2}{2}}$ the corresponding probability density function, and $\sigma > 0$. By the group property and strong continuity of $B[t]$, for each $\varepsilon > 0$ there exists $\delta > 0$ such that if $|u| < \delta$, $||B[t](u)\Phi - \Phi|| < \varepsilon$. Hence

$$||\mathcal{G}_t^\sigma(\Phi) - \Phi|| = ||\int_{-\infty}^{\infty} \big(B[t](\sigma s)\Phi(x) - \Phi(x)\big)\mathfrak{g}(s)ds|| \leq \varepsilon + 2||\Phi||(1 - \mathcal{N}(\frac{\delta}{\sigma})) \tag{4.62}$$

where $\mathcal{N}(s) := \mathbb{P}[Y \leq s]$ is the standard normal cumulative probability distribution function. So $\mathcal{G}_t^\sigma(\Phi) \to \Phi$ in C_b as $\sigma \to 0$.

We now show $\mathcal{G}_t^\sigma(\Phi)$ is vertically entire at t. We have

$$\frac{1}{\varepsilon}\big(B[t](\varepsilon) - I\big)\mathcal{G}_t^\sigma(\Phi)(x) = \frac{1}{\varepsilon}\int_{-\infty}^{\infty}(B[t](\sigma s + \varepsilon) - B[t](\sigma s))\Phi(x)\mathfrak{g}(s)ds \tag{4.63}$$

$$= \int_{-\infty}^{\infty} B[t](s)\Phi(x)\frac{\mathfrak{g}(\frac{s-\varepsilon}{\sigma}) - \mathfrak{g}(\frac{s}{\sigma})}{\varepsilon}\frac{ds}{\sigma} \tag{4.64}$$

Hence, by the dominated convergence theorem and the smoothness of the Gaussian, $\mathcal{G}_t^\sigma(\Phi) \in C_t^\infty$ and

$$\nabla_t^k \mathcal{G}_t^\sigma(\Phi)(x) = \int_{-\infty}^{\infty} B[t](s)\Phi(x)\Big(\frac{d}{ds}\Big)^k \mathfrak{g}(\frac{s}{\sigma})\frac{ds}{\sigma} \tag{4.65}$$

It is well known (see, for instance, Huang and Yan (2000), Appendix A) that $\big(\frac{d}{ds}\big)^k \mathfrak{g}(\frac{s}{\sigma}) = (\frac{-1}{\sigma^2})^k \mathrm{He}_k(\frac{s}{\sigma})\mathfrak{g}(\frac{s}{\sigma})$, where He_k is the kth *Hermite polynomial*, and that these form an orthonormal basis of $L^2(d\mathfrak{g})$. So $||\nabla_t^k \mathcal{G}_t^\sigma(\Phi)|| \leq \sigma^k ||\Phi|| |\mathbb{E}[\mathrm{He}_k(\sigma Y)]| \leq \sigma^k ||\Phi||$, and the Taylor series converges. □

Theorem 4.14. *$C^\infty_{[0,T]}(\mathbb{D}^d)$ is dense in $\mathcal{U}_b(\mathbb{D}^d)$.*

Proof. We use a similar trick to Lemma 4.13, but replace the single Gaussian perturbation with one based on a *Brownian motion*. For $\Phi \in C_{ub}(\mathbb{D}^d)$, let

$$\mathcal{W}^\sigma(\Phi)(x) := \mathbb{E}[\Phi(x + \sigma W)] \tag{4.66}$$

where W is a standard Brownian motion, and again $\sigma > 0$. For fixed $x \in \mathbb{D}^d$ the integrand $\omega \to \Phi(x + \sigma W(\omega))$ is a continuous, bounded functional on $\mathbb{W}^d$. By the definition of the Wiener measure, for any refining sequence of partitions τ^n,

$$\mathcal{W}_n^\sigma(\Phi)(x) := \mathbb{E}[\Phi(x + \sigma \mathcal{E}_n^\tau(W))] = \prod_{k=1}^{m(\tau^n)} \mathcal{G}_{\tau_k^n}^{\sigma(\tau_k^n - \tau_{k-1}^n)}(\Phi)(x) \to \mathcal{W}^\sigma(\Phi)(x) \tag{4.67}$$

as $n \to \infty$. Since Brownian paths are continuous, the Euler projections converge uniformly, so by the *uniform* continuity of Φ, $||\mathcal{W}^\sigma(\Phi) - \mathcal{W}_n^\sigma(\Phi)|| \to 0$ as $\sigma \to 0$, for any fixed $n \in \mathbb{N}$. By repeated application of equation 4.62, for any such n and $\varepsilon > 0$ we may pick $\delta > 0$ such that

$$||\mathcal{W}_n^\sigma(\Phi) - \Phi|| < m(\tau^n)\Big(\frac{\varepsilon}{m(\tau^n)} + 2.||\Phi||(1 - \mathcal{N}(\frac{\delta}{\sigma}))\Big) \tag{4.68}$$

Hence $||\mathcal{W}^\sigma(\Phi) - \Phi|| \to 0$ as $\sigma \to 0$ as required.

We must now show the vertical smoothness of $\mathcal{W}^\sigma(\Phi)$. Pick $t \in [0, T]$, and a refining sequence of partitions τ^n with $t \in \tau^n$ for each n. The action of the bumping group effectively shifts the mean of the Gaussian distribution of the subsequent increments. Using the formula for changing the mean of the Gaussian density, the orbit function of $\mathcal{W}_n^\sigma(\Phi)(x)$ at t is

$$\varphi_n(\varepsilon) := B[t](\varepsilon)\mathcal{W}_n^\sigma(\Phi) = \mathbb{E}[\Phi(x + \sigma\mathcal{E}_n^\tau(W)) \prod_{\tau_k^n \geq t} e^{-\frac{\varepsilon}{\sigma}\Delta W(\tau_k^n) + \frac{1}{2}(\frac{\varepsilon}{\sigma})^2(\tau_k^n - \tau_{k-1}^n)}] \tag{4.69}$$

We then have

$$\nabla_t \mathcal{W}_n^\sigma(\Phi) = \varphi_n'(0) = \sum_{\tau_k^n \geq t} \mathbb{E}[\Phi(x + \sigma\mathcal{E}_n^\tau(W))\frac{-\Delta W(\tau_k^n)}{\sigma}] = \mathbb{E}[\Phi(x + \sigma\mathcal{E}_n^\tau(W))(\frac{W(t) - W(T)}{\sigma})] \tag{4.70}$$

The term in the product in 4.69 almost surely converges to the random variable

$$\mathcal{E}(\varepsilon) := e^{\frac{\varepsilon}{\sigma}(W(T) - W(t)) + \frac{1}{2}(\frac{\varepsilon}{\sigma})^2(T-t)} \tag{4.71}$$

and note that $\mathcal{E}(\varepsilon) > 0$ $\mathbb{P}$-a.s., $\mathbb{E}[\mathcal{E}(\varepsilon)] = 1$, and $\mathcal{E}(\varepsilon)$ is the Radon-Nikodym derivative of a probability measure. Using the boundedness of Φ, the dominated convergence theorem then implies that 4.69 converges, say to $\psi(\varepsilon) := \mathbb{E}[\Phi(x + \sigma W)\mathcal{E}(\varepsilon)]$, and that this is therefore equal to the orbit function $\zeta_{x,t}^{\mathcal{W}^\sigma(\Phi)}(\varepsilon)$. This is differentiable at zero and

$$\nabla_t \mathcal{W}^\sigma(\Phi)(x) = \psi'(0) = \mathbb{E}[\Phi(x + \sigma W)(\frac{W(t) - W(T)}{\sigma})] \tag{4.72}$$

Noting that $\frac{W(t) - W(T)}{\sigma}$ is invariant under the transformation $W \to W + u\beta[t]$, we have

$$\nabla_t^k \mathcal{W}^\sigma(\Phi) = \mathbb{E}[\Phi(x + \sigma W)\Big(\frac{W(t) - W(T)}{\sigma}\Big)^k] \tag{4.73}$$

as required. □

Theorem 4.14 is a key result, and should be interpreted as the analogue of the density of smooth functions in the main spaces for analysis of PDEs in finite dimensions. The vector space $C_{[0,T]}^\infty(\mathbb{D}^d)$ will form a *core* for general vertical differential operators in the same way $C^\infty(\mathbb{R}^d)$ does for operators on $\mathbb{R}^d$ – that is, a dense subset where it is 'well behaved'. Note that the assumption of uniform continuity is necessary, but that this includes many payoff functionals, such as vanilla options. This implies the following important corollary.

Corollary 4.15. *$C^{\infty}_{[0,T]}(\mathbb{D}^d)$ is dense in $L^p(\mathbb{P})$ for any probability measure $\mathbb{P}$ on $(\mathbb{D}^d, \mathcal{F})$.*

Proof. This follows immediately from Theorems 4.3 and 4.14 . □

4.3 Functional processes, risk sensitivities and ‘delta’

The price of a derivative contract at any fixed time should be determined by the paths of the relevant risk factors, that is, should be a path functional. So we are naturally led to consider paths *with values in spaces of path functionals*. That is, *functional processes*, maps

$$F : [0,T] \to Z \tag{4.74}$$

where Z is some Banach space of path functionals $z : \mathbb{D}^d \to \mathbb{R}$.

We may consider left, right limits and jump functional process $z(t+), z(t-), \Delta z(t)$ of functional processes, in the usual way, remembering that they are with respect to the norm of Z. We then define the corresponding Banach spaces $\mathbb{G}(Z), \mathbb{F}(Z), \mathbb{D}(Z)$ as discussed previously. For instance, if $Z = C_b(\mathbb{D}^d)$, the norm on $\mathbb{G}(C_b(\mathbb{D}^d))$ is

$$||F||_{\mathbb{G}(C_b(\mathbb{D}^d))} := \sup_{t\in[0,T]} ||F(t)||_{C_{ub}} = \sup_{t\in[0,T]} \sup_{x\in\mathbb{D}} |F(t)(x)| \tag{4.75}$$

Generally we will write $F(t,x)$ when evaluating functional processes, rather than the somewhat clunky $F(t)(x)$, when there is no confusion.

An important point for coherent modelling is that the price at time t must rely only on information available up to t. That is, the evolution of prices should be given by *non-anticipative functionals*, whose values on a given path at a given time do not depend on the ‘future’ of that path. As noted above, the existing approaches to defining this notion – as functions on a vector bundle, or a metric space of equivalence classes of paths – are not amenable to the methods of operator semigroups, so we develop a slightly different one here.

Definition 4.16. (Stopping operators)

The *stopping operator at $t \in [0,T]$*, $\pi_t : \mathbb{D}^d_T \to \mathbb{D}^d_T$, is defined by

$$\pi_t x(u) := x(u \wedge t) = \begin{cases} x(u) & u \in [0,t) \\ x(t) & u \in [t,T] \end{cases} \tag{4.76}$$

We will also use the notation $x_t := \pi_t x$ at times, for brevity.

Note that π_t is a projection, $\pi_t^2 = \pi_t$. We may also define the *predictable projection*,

$$\pi_{t-} x := \begin{cases} x(u) & u \in [0,t) \\ x(t-) & u \in [t,T] \end{cases} \tag{4.77}$$

STOPPED PATH

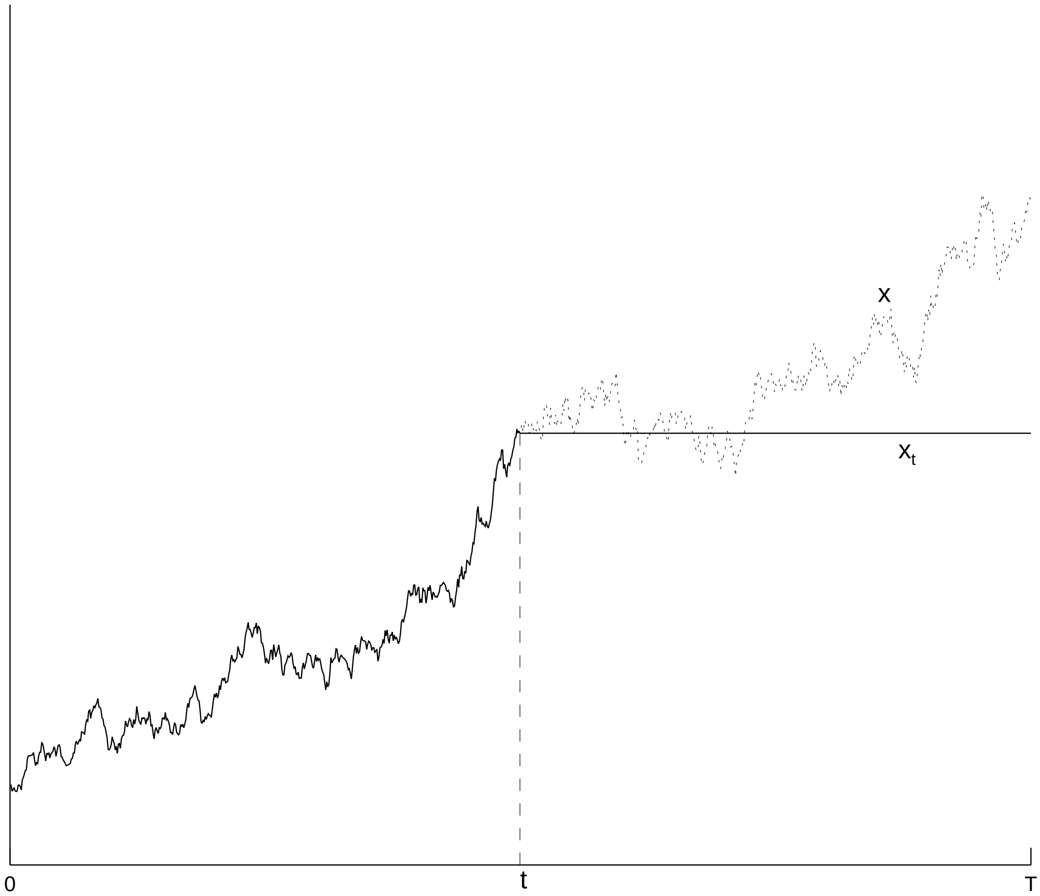

Figure 4.6: The action of the stopping operator on a path.

which clearly also satisfies $\pi_{t-}^2 = \pi_{t-}$, and we will similarly sometimes write $x_{t-} := \pi_{t-}x$. For any $\varepsilon > 0$, we have

$$\pi_{t-}x - \pi_{t-\varepsilon}x := \begin{cases} 0 & u \in [0, t-\varepsilon) \\ x(u) - x(t-\varepsilon) & u \in [t-\varepsilon, t) \\ x(t-) - x(t-\varepsilon) & u \in [t, T] \end{cases} \tag{4.78}$$

The right-continuity of x ensures the middle expression can be made arbitrarily small as $\varepsilon \to 0$, while the existence of a left limit ensures the same of $x(t-) - x(t-\varepsilon)$, so $x_{t-\varepsilon} \to x_{t-}$ in $\mathbb{D}^d$. Obviously, $x_t = x_{t-} + \Delta x(t)\beta[t]$, and in particular continuous paths are predictable.

Proposition 4.17. *If $F \in \mathbb{D}(\mathcal{U}_b(\mathbb{D}^d))$ then for each $x \in \mathbb{D}^d$, the path $t \to F(t, x_t) \in \mathbb{D}^1$. Similarly if $F \in \mathbb{F}(\mathcal{U}_b(\mathbb{D}^d))$, then $t \to F(t, x_t) \in \mathbb{F}^1$.*

Proof. First, since for each $x \in \mathbb{D}^d$

$$|F(t \pm \varepsilon, x) - F(t, x)| \leq ||F(t \pm \varepsilon) - F(t)|| \tag{4.79}$$

the path $t \to F(t,x)$ is càdlàg. Similarly,

$$|F(t \pm \varepsilon, x_{t\pm\varepsilon}) - F(t, x_t)| \leq |F(t \pm \varepsilon, x_{t\pm\varepsilon}) - F(t \pm \varepsilon, x_t)| + |F(t \pm \varepsilon, x_t) - F(t, x_t)| \quad (4.80)$$

The second term is bounded by $||F(t \pm \varepsilon) - F(t)|| \to 0$ as $\varepsilon \downarrow 0$. For the first, since $F(t,.) \in \mathcal{U}_b(\mathbb{D}^d)$, for each $\eta > 0$ there exists $\delta > 0$ such that $||F(t \pm \varepsilon, x) - F(t \pm \varepsilon, y)|| < \eta$ whenever $||x - y|| < \delta$. Since $x \in \mathbb{D}^d$, for sufficiently small ε, $||x_{t\pm\varepsilon} - x_t|| < \delta$ and the first claim is proved. The proof for the second claim is similar. □

If $x, y \in \mathbb{D}^d_T$, then we may form their *concatenation at t* path

$$x \oplus_t y := x_t + \big(y - x(t)\big)\beta[t] \quad (4.81)$$

Given two sets of paths $A, B \subset \mathbb{D}^d$ we define the *concatenation product* $A \oplus B := \{a \oplus_t b \mid a \in A, b \in B, t \in [0,T]\} \subset \mathbb{D}^d$. If $A = \{a\}$ consists of a single path we will abuse this notation slightly and write $a \oplus B$. Note that if A, B are both closed in $\mathbb{D}^d$ then so is $A \oplus B$, and that $\pi_t^{-1}(x_t) = \{x \oplus_t y \mid y \in \mathbb{D}^d_T\} = x \oplus \mathbb{D}^d$ which we can also think of as $\{x_t + z(. - t)\beta[t] \mid z \in \mathbb{D}^d_{T-t},\ z(0) = 0\}$.

The parameterised family of stopping operators on paths then lifts to a projection on functionals, which by slight abuse of notation we will denote by the same letter: $\pi_t\Phi(x) := \Phi(x_t) = \pi_t^2\Phi(x)$. The stopping operator on functionals is a contraction for the supremum norm:

$$||\pi_t\Phi||_\infty = \sup_{x\in\mathbb{D}} |\Phi(x_t)| \leq \sup_{x\in\mathbb{D}} |\Phi(x)| = ||\Phi|| \quad (4.82)$$

and so is continuous on C_b.

For a functional process F, define

- the *predictable part* π_F of F by $\pi\_F(t,x) := \lim_{\varepsilon\downarrow 0} F(t, x_{t-\varepsilon})$
- the *non-anticipative part* πF of F by $\pi F\,(t,x) := \pi_t F(t)(x) = F(t, x_t)$, and
- the *anticipative part* αF of F by $\alpha F(t,x) := \big(I - \pi\big)F\,(t,x) = F(t,x) - F(t,x_t)$

Obviously these are all linear mappings. A functional process F is said to be *predictable* if $F = \pi\_F$, and *non-anticipative* if $F = \pi F$, and *anticipative* if $F \neq 0$ and $\alpha F \neq 0$, with zero of course here being the identically zero functional process. Clearly predictable functional processes are non-anticipative; each has anticipative part equal to the zero functional process, but the former is independent of the instantaneous path value, and the latter only of the strictly prior past.

Example 4.4. *Let $x \in \mathbb{D}^1$, and $F(t,x) := \int_0^t x(s)ds$, and note that since $\mathcal{J}(x)$ is countable the integral exists. Since $F(t,x) - F(t-\varepsilon,x) = \int_{t-\varepsilon}^t x(s)ds \to 0$ as $\varepsilon \to 0$, F is predictable.*

If $Z = C_b$, then α is a continuous linear projection on $\mathbb{D}(C_b)$. Consequently we may form the quotient space under the relation $F \sim G \iff F - G \in A$.

Definition 4.18. (Regulated, càdlàg, càglàd, non-anticipative C_b-functionals)

The space of *Regulated non-anticipative C_b-functionals* is the quotient space $\mathbb{G}_\pi(C_b) := \mathbb{G}(C_b)\big/A$ of $\mathbb{G}(C_b)$ by the closed subspace $A := \text{Im}\ \alpha$ of anticipative càdlàg functionals. The quotient norm is given by

$$||F||_\pi = \sup_{t\in[0,T]} ||\pi_t F(t)||_{C_b} = \sup_{t\in[0,T]} \sup_{x\in\mathbb{D}} |F(t, x_t)| \tag{4.83}$$

Similarly, the spaces of *càdlàg* and *càglàd non-anticipative C_b-functionals* are the corresponding quotient spaces, $\mathbb{D}_\pi(C_b) := \mathbb{D}(C_b)\big/A$ and $\mathbb{F}_\pi(C_b) := \mathbb{F}(C_b)\big/A$, with the same norm.

The first thing to note about this definition of non-anticipative functionals is that they are *equivalence classes of functional processes* that at a given time t give the same value to paths that are stopped at t, rather than *continuous functions defined on a metric space of equivalence classes of paths stopped at t* as in Cont et al. (2016). This may seem a pedantic point – we will discuss it in more detail below – but will actually be key to adapting semigroup methods to the path-dependent setting. It will sometimes be useful to refer to them by their unique value on the pairs, $[F](t, x) = F(t, x_t)$, for $F \in [F]$. If we need to make the equivalence class structure explicit we will write $[F]$, $F + A$, or πF interchangeably, with the presumption that $F \in \pi^{-1}([F])$ is a functional process of the relevant class.

We may now give some examples.

Example 4.5. *(Non-anticipative functionals)*

- *Let $\Phi \in C_b$. Then the stopped functional process $F_\Phi(t, x) := \Phi(x_t) \in \pi C(T; C_b)$*
- *Let $\mathfrak{p}^n$ be a refining sequence of partitions, and the functions $\varphi_t \in C^1(\mathbb{R}^{m(\mathfrak{p}^n)})$ for each $t \in [0, T]$. The* cylindrical functional process *$F(t, x) = \varphi_t(x(\mathfrak{p}^n))$, is non-anticipative if and only if*

$$\partial_k \varphi_t(x(\mathfrak{p}^n)) = 0 \text{ if } t < \mathfrak{p}^n_k \tag{4.84}$$

- *Let $0 \leq r \in \mathbb{D}(T; C_{ub})$ be a positive non-anticipative functional, representing the short rate on a cash account, as a function of d risk factors x. Then the value $S^0(t, x) := e^{\int_0^t r(s,x)ds}$ at time t of a dollar invested at $t = 0$, is predictable.*
- *Let $\mathbb{P}(x_t)$ be a family of probability measures on $\mathbb{D}^d_T$, indexed by stopped paths x_t, such that $\mathbb{P}(x_t)[\pi_t^{-1}(x)] = 1$, and suppose $\Phi \in \bigcap L^1(\mathbb{P}(x_t))$. Then*

$$V^{\mathbb{P}}[\Phi](t, x) = \mathbb{E}^{\mathbb{P}(x_t)}[\ \Phi(x \oplus_t X)] \tag{4.85}$$

  is a non-anticipative functional, where $X(t, \omega)$ is a stochastic process under $\mathbb{P}(x_t)$. As we shall see, arbitrage-free valuation functionals are of this form, typically with $\mathbb{P}(x_t)$ the regular conditional expectation *of a* local martingale measure.

■ *Let $d = 1$ and consider a market in a single stock with spot price path $S \in \mathbb{D}_T$. The* Black-Scholes *value at time t of a derivative contract with payoff Φ, given the price path to date S_t, is given by*

$$V_\sigma^{BS}[\Phi](t, S_t) := \mathbb{E}_t[e^{-r(T-t)}\Phi(S \oplus_t S(t)X)] \tag{4.86}$$

provided this expectation exists, where X is the path of the stochastic process

$$X(u, \omega) := M(u-t, W(u-t, \omega))1_{[t,T]}(u) \tag{4.87}$$

with $M(s,x) := \exp\left((r - \frac{1}{2}\sigma^2)s + \sigma x\right)$, $\mathbb{E}_t$ is the expectation under a probability measure $\mathbb{P}_t$ on $\mathbb{D}_{T-t}^d$ under which W is Brownian motion, and r and σ the interest rate and volatility parameters respectively. Clearly, this is non-anticipative; the dependence on the path is entirely down to the current value $S(t) = s$, and it may be considered as an ordinary function: $V_\sigma^{BS}[\Phi](t, S_t) = V_\sigma^{BS}[\Phi](t, S(t)\beta[t]) = \mathbb{E}_t[e^{-r(T-t)}\Phi(sX_T)] =: v_\Phi^{BS}(t,s)$. For a simple European option as in 4.8 with payoff $\Phi_\varphi \in L^2(\mathbb{P}_t)$, this gives the one-dimensional integral

$$v_\Phi^{BS}(t,s) = \int_{-\infty}^{\infty} e^{-r(T-t)}\varphi(S(t)e^z)\, \mathfrak{g}\Big(\frac{z - r(T-t)}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t}\Big)dz \tag{4.88}$$

where $\mathfrak{g}(x) := \frac{1}{2\pi}\exp(-\frac{x^2}{2})$ is the density of a standard normal random variable. Clearly, $V_\sigma^{BS}[\Phi_\varphi] \in C(0,T;C_{ub})$.

Note that in these last two 'probabilistic' examples the non-anticipative functionals are 'versions' of conditional expectations, in the sense that they almost surely coincide with the conditional expectation of some stochastic process, but *remain defined pathwise*, not merely almost surely.

It is important to understand the interaction between stopping and bumping operators. We have

$$B[t'](u)\pi_t\Phi(x) = B[t'](u)\Phi(x_t) \tag{4.89}$$

$$= \Phi\big(x_t + u\beta[t']\big), \tag{4.90}$$

while

$$\pi_t B[t'](u)\Phi(x) = \pi_t\Phi(x + u\beta[t']) \tag{4.91}$$

$$= \Phi\big(x_t + u\beta[t'](t)\beta[t']\big) \tag{4.92}$$

$$= \begin{cases} \Phi(x_t) & t \in [0,t') \\ \Phi(x_t + u\beta[t']) & t \in [t',T] \end{cases} \tag{4.93}$$

whence

$$[B[t'], \pi_t]\Phi(x) = \big(\Phi(x_t + u\beta[t']) - \Phi(x_t)\big)1_{[0,t')}(t) \tag{4.94}$$

where the *commutator* $[X, Y] = XY - YX$. Intuitively: *stopping commutes with simultaneous or earlier bumping, but not with later.*

If a non-anticipative functional $F = \pi F$ has $F(t) \in C_t^1$ for each t, then we may define a vector-valued non-anticipative functional ∇F, with ith component given by

$$(\nabla F)^i\,(t, x) = \pi_t \partial^i_{x_t} F(t)(x) = \partial^i_{x_t} F(t, x_t) \tag{4.95}$$

Risk sensitivities are then defined via the application of this differential operator. If a non-anticipative functional $V(t, S)$ represents the value of a derivative at time t on the d-dimensional spot price path S, then $\nabla V(t, S)$ represents its *delta* – the sensitivity of the value with respect to instantaneous shocks in the spot price. The matrix analogous to the Hessian, with i, j entry given by the non-anticipative functional

$$\Gamma_{ij}(t, S) := \partial^i_{x_t} \partial^j_{x_t} V(t, S) \tag{4.96}$$

is known as the (cross-) *gamma* – the sensitivity of the delta with respect to instantaneous shocks in the spot price. If we consider a model in which the prevailing interest rate r is a risk factor we may add it in as an additional variable and define the *rho* as $\rho := \nabla^{d+1} V(t, S, r)$. Similarly for other variables – for a model in which volatilities of the k assets are to be risk factors, we set $d = 2k$, label the first k variables as spot prices S and the next k to be their path-dependent volatilities σ. Then $\nu(t, S) := \nabla^{k+i} V(t, S, \sigma)$, $i = 1, \ldots, k$, is known as the *vega* – the sensitivity of the delta with respect to instantaneous shocks in the volatility.

Example 4.6. *(Vertical differentiation of non-anticipative functionals)*

- *A predictable functional process F has $B_t(\varepsilon)F = F$ for all $\varepsilon > 0$ and $t \in [0, T]$, and so* $\nabla_t F(t, x) \equiv 0$.
- *By equations 4.56 and 4.84, a non-anticipative cylindrical functional process $F(t, x) = \varphi_t(x(\mathfrak{p}^n))$ with $\varphi_t \in C^1(\mathbb{R}^{m(\mathfrak{p}^n)})$ for each $t \in [0, T]$, has $F \in \bigcap_{t\in[0,T]} C_t^1$, and*

$$\nabla_t F(t, x_t) = \partial_{k^*} \phi_t(x(\mathfrak{p}^n)) \tag{4.97}$$

  where $t \in [\mathfrak{p}^n_{k^*}, \mathfrak{p}^n_{k^*+1})$

- *Consider the Black-Scholes valuation functional, 4.86. A standard trick to find a* weak representation *of its delta is to absorb the 'bump' into the Brownian motion in the expectation, and then use Girsanov's theorem to transform the probability measure. Since $(S + \varepsilon\beta[t])(t).X_T = S(t).X^\varepsilon_T$, where*

$$X^\varepsilon_T(u, \omega) := M(u - t, W(u - t, \omega) + \frac{1}{\sigma}\log(\varepsilon))1_{[t,T]}(u) \tag{4.98}$$

$$= M(u - t, W^\varepsilon(u - t, \omega))1_{[t,T]}(u) \tag{4.99}$$

and $W^{\varepsilon}(u-t,\omega) = W(u-t,\omega) + \frac{1}{\sigma}\log(\varepsilon)$*, we have*

$$B[t](\varepsilon)V_{\sigma}^{BS}[\Phi](t,S_t) = \mathbb{E}^{\mathbb{P}_t^{\varepsilon}}[e^{-r(T-t)}\Phi(S \oplus_t S(t)X)] \tag{4.100}$$

where $\mathbb{P}_t^{\varepsilon}$ *is the probability measure on* $\mathbb{D}_{T-t}$ *equivalent to* $\mathbb{P}_t$ *under which* W^{ε} *is a Brownian motion. By* Girsanov's theorem, $\frac{d\mathbb{P}^{\varepsilon}}{d\mathbb{P}}(u,\omega) = \exp(W(u-t,\omega) + \frac{1}{\sigma}\log(\varepsilon) - (u-t) + \frac{1}{2\sigma^2}(\log(\varepsilon))^2)$*. Therefore,*

$$\Delta[\Phi](t,S_t) := \nabla_t V_{\sigma}^{BS}[\Phi](t,S_t) = \lim_{\varepsilon\to 0}\frac{\mathbb{E}^{\mathbb{P}_t^{\varepsilon}} - \mathbb{E}^{\mathbb{P}_t}}{\varepsilon}[e^{-r(T-t)}\Phi(S \oplus_t S(t)X)] \tag{4.101}$$

$$= \frac{d}{d\varepsilon}|_{\varepsilon=0}\mathbb{E}^{\mathbb{P}_t}[e^{-r(T-t)}\Phi(S \oplus_t S(t)X)\frac{d\mathbb{P}_t^{\varepsilon}}{d\mathbb{P}_t}] \tag{4.102}$$

$$= \mathbb{E}^{\mathbb{P}_t}[e^{-r(T-t)}\Phi(S \oplus_t S(t)X)\Psi^{\Delta}[\Phi]] \tag{4.103}$$

where $\Psi^{\Delta}[\Phi] \in L^2(\mathbb{P})$ *is the random variable whose existence follows from the Riesz representation theorem if the limit 4.102 exists, and is known as the* Malliavin weight for delta *(see Fournié et al. (2001)). In the case of the simple European option we have*

$$\Delta[\Phi_{\varphi}](t,S_t) = \int_{-\infty}^{\infty} e^{-r(T-t)}\varphi(S(t)e^z).\frac{z}{\sigma(T-t)S(t)}\,\mathfrak{g}\Big(\frac{z-r(T-t)}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t}\Big)dz \tag{4.104}$$

$$= \mathbb{E}^{\mathbb{P}_t}[e^{-r(T-t)}\Phi(S \oplus_t S(t)X)\Psi_T^{\Delta}] \tag{4.105}$$

where the random variable $\Psi_T^{\Delta}(\omega) := \frac{W(T,\omega)}{\sigma(T-t)S(t)}$*, independent of* φ*.*

The requirements on Φ *to ensure these expressions converge are* probabilistic*; it is not clear if they have a pathwise meaning.*

■ *Similarly, the weak representation for the* Black-Scholes gamma

$$\Gamma^{BS}[\Phi_{\varphi}](t,S_t) = \mathbb{E}^{\mathbb{P}_t}[e^{-r(T-t)}\Phi(S \oplus_t S(t)X)\Psi^{\Gamma}[\Phi]] \tag{4.106}$$

with Malliavin weight $\Psi^{\Gamma}[\Phi]$ *exists when the corresponding limit exists in* $L^2(\mathbb{P}_t)$*, by the Riesz representation theorem, as in the above argument for delta. For a simple European option, a straightforward calculation yields* $\Psi_T^{\Gamma}[\Phi](\omega) = \frac{W_T(W_T-\sigma T)-T}{(\Psi_T^{\Delta})^2}$.

How does our set-up for vertical differentiation relate to the existing approaches of Dupire, Cont et al? For convenience we will consider the Cont framework, where they define non-anticipative functionals as functions defined on a *metric space* Λ of *stopped paths*: equivalence classes of paths $[x]_t := \tilde{\pi}_t^{-1}(x) = \{y \in \mathbb{D}^d \mid \tilde{\pi}_t y = \tilde{\pi}_t x\}$, with metric

$$d([x]_t,[y]_{t'}) := ||x_t - y_{t'}|| + |t-t'| \tag{4.107}$$

(see Figure 4.7). Another way to think about this is as a $\mathbb{D}^d$ vector *bundle* over $[0, T]$ (see Lee (2003)). The natural projection maps $[x]_t$ to t, and in fact there is a global co-ordinate system which we may write as (t, x_t), with x_t in the quotient space $\mathbb{D}^d \big/ \tilde{\pi}_t$.

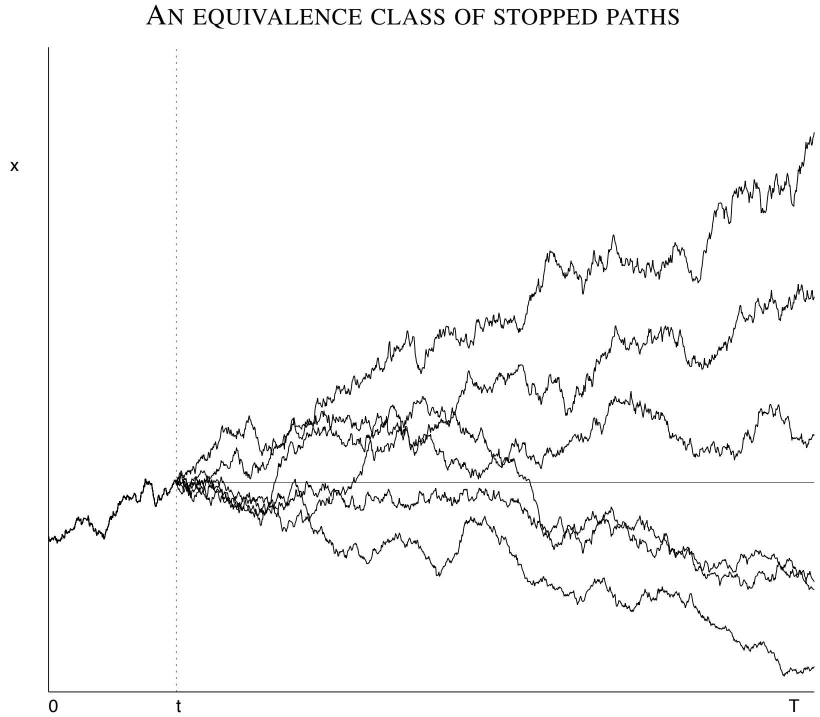

Figure 4.7: Heuristic representation of an equivalence class of paths under the stopping relation.

An important point is that the definition is not as a vector space in itself, merely a metric space. This is partly because adding stopped paths does not seem natural – a naive attempt runs into an 'apples and oranges' problem, since paths stopped at different times are not the same class of object, and so it is not obvious how to add them sensibly. It is possible to define a vector space structure via $[x]_t + \lambda [y]_{t'} := [x + \lambda y]_{t \wedge t'}$, but this is somewhat cumbersome because of the need to take the minimum of the two times, and neither Dupire nor Cont and his co-authors use it. It is more straightforward to simply consider, as we do, a natural vector space of paths like $\mathbb{D}^d$ or $\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)$, and account for the flow of information *in the functionals* by quotienting at that level.

Similarly, when defining their vertical differential, Cont and co-authors consider a difference quotient based on evaluating a non-anticipative functional (in their sense) on an equivalence class of paths stopped at t and the same equivalence class 'bumped' at t, in which all the paths are translated by $\beta^i[t]$ (see figure 4.8). In their formalism, vertical differentials (which we will write as $\tilde{\nabla}$ to distinguish from ours) are defined on these non-anticipative functionals by

$$\tilde{\nabla} \tilde{F}([x]_t) = \lim_{\varepsilon \to 0} \frac{\tilde{F}([x + \varepsilon \beta[t]]_t) - \tilde{F}([x]_t)}{\varepsilon} \tag{4.108}$$

Note that the timing of the bump $\beta[t]$ *depends on when the path has been stopped*; the perturbation 'knows' when the path being considered has been stopped. This makes it like a *continuum*

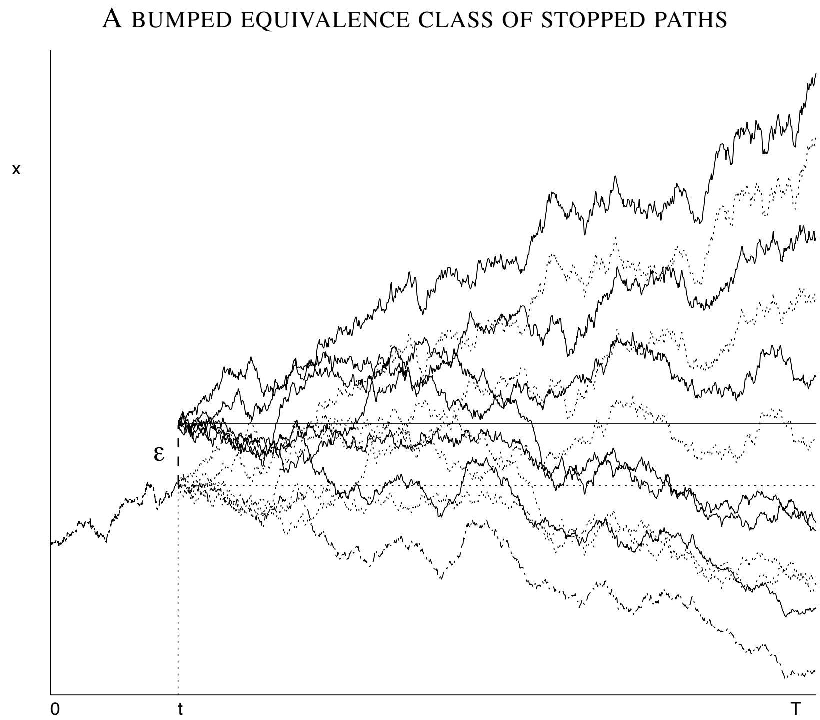

Figure 4.8: Heuristic representation of the action of a vertical bump on an equivalence class of paths under the stopping relation.

of infinitesimal actions of the bumping groups $B[t]$ on a complicated class of appropriate subsets of Λ, rather than a single one, and so the theory of semigroups is not applicable. Essentially, in trying to do everything at once the existing approaches make it more complicated, and less like ordinary differentiation, than it needs to be. In our setting, because we consider each of the continuum of actions *separately* on *a single vector space*, with the stopping operator also considered separately, we can express each as the generator of a semigroup.

For given arguments the definitions correspond. For each t, $x \in \mathbb{D}^d$, the non-anticipative functional (in our sense) $\nabla_t F(t,x)$ coincides with the numerical value of $\tilde{\nabla}\tilde{F}([x]_t)$, and the commutation of the projection with the bumping groups mean the differentials coincide. The advantage of our approach is that, unlike the $\tilde{F}$, which live in the poorly understood space $C(\Lambda)$, our F live in $\mathscr{U}_b(\mathbb{D}^d)$, which is in many ways similar to the analogous spaces for finite dimensions $\mathscr{U}_b(\mathbb{R}^d)$. In particular, since the domain of the functionals is a vector space, we can make use of semigroup analysis, as we saw in Section 4.2.

In some sense the difference in the two formalisms is ‘cosmetic’. However, proving results is considerably easier in our set-up: by quotienting at the level of maps, rather than arguments, we are able to call on semigroup theory, and use the deep parallells to ordinary differentiation, unlike the Dupire-Cont formalism.

4.4 The horizontal translation semigroup and 'theta'

In the Markovian case, a derivative's value is a function of time and spot, and the partial differential with respect to time gives the decay in value due to the passage of time, all else held equal, known as *theta*. In the path-dependent case, Dupire showed how the *horizontal* differential plays the role of a time differential, in which the market's risk factors are 'frozen' at their current value for an infinitesimal time period. As emphasised by Cont and Fournie, this is *not* the same as a partial differential with respect to time in the existing formalism (neither theirs nor Dupire's). However, we shall see that under the new formalism of section 4.3, the horizontal differential is the generator of an appropriate semigroup.

There is a natural translation semigroup on $\mathbb{D}(C_b)$, which we can think of as a (forward) time translation that shifts a functional process ahead, and write as $tr(u)F(t,x) := F((t+u)\wedge T, x)$. The idea of the *horizontal* translation is to correct this for the *flow of information*, using the stopping operator.

Definition 4.19. For $u \in \mathbb{R}_+$, we define the *horizontal translation operator* $H(u)$ on $\mathbb{D}(T;C_b)$ by

$$H(u)F\ (t,x) := F(t+u, x_t) \tag{4.109}$$

if $u \in [0, T-t]$, and zero otherwise.

So horizontal translation is simply the evaluation of the shifted process $t \rightarrow F(t+u,.)$ on the stopped path x_t. From the financial point of view, we can think of this as, for instance, the value of a derivative at $t+u$ using the market data at t, as if it were 'frozen', in Cont and Fournie's terms. By construction, $H(u)F$ is non-anticipative for any functional process F, so in particular $\mathbb{D}_\pi(C_b)$ is *invariant* under its action, and we may use the same notation for the restriction to non-anticipative functionals. A key point is that it inherits the (nilpotent) semigroup structure of translation.

Theorem 4.20. *(Horizontal translation semigroup)*

The horizontal translation $H(u)$ operators form a strongly continuous semigroup of contractions on $\mathbb{D}_\pi(C_b)$.

Proof. Obviously we have $H(0) = I$, and $H(u+v)F(t,x) = F((t+u+v)\wedge T, x_t) = H(v)F((t+u)\wedge T, x_t) = H(u)H(v)F(t,x)$. Strong continuity follows from the strong continuity of $tr(.)$ and the continuity of the non-anticipative projection. The contraction property is obvious from the definition of the norm on $\mathbb{D}_\pi(T;C_b)$.

We must check $H(u)$ is well-defined on $\mathbb{D}_\pi(T;C_b)$, ie independent of the choice of representative from $[F]$. Suppose $F, G \in [F]$, so that for each $t \in [0,T]$, $\pi_t F(t,.) = \pi_t G(t,.)$. Since $\pi_t = \pi_t\pi_{t+u}$, we have for each $x \in \mathbb{D}$, $G(t+u, x_t) = \pi_t\pi_{t+u}G(t+u,x) = \pi_t\pi_{t+u}F(t+u,x) = \pi_t F(t+u,x) = F(t+u,x_t)$ as required. □

We have the corresponding infinitesimal generator

$$\mathcal{D}F(t,x) := \lim_{u\downarrow 0}\frac{H(u)F(t,x)-F(t,x)}{u} \tag{4.110}$$

of the horizontal translation semigroup on its natural domain

$$\mathbb{D}^1(C_b) := \{F \in \mathbb{D}(C_b) \mid \mathcal{D}F \text{ exists}\} \tag{4.111}$$

Note that $\mathcal{D}F \in \mathbb{D}_\pi(C_b)$ – the generator has *non-anticipative* range. It is a *one-sided* differential operator, the infinitesimal action of translation in time from the 'right', i.e. the future, along the path stopped at the present. Accordingly, it makes sense to define it on càdlàg functional processes, which are by definition continuous from the right. Note that this is more general than the Dupire/Cont formalism, since we allow the functional to 'jump', rather than be continuous with respect to the Cont metric 4.107.

Example 4.7. *(Black-Scholes theta)*

Consider again the Black-Scholes valuation functional, 4.86, and set its theta, $\Theta^{BS}[\Phi](t,S_t) := \mathcal{D}V_\sigma^{BS}[\Phi](t,S_t)$. *We have*

$$\frac{H(u)-I}{u}V_\sigma^{BS}[\Phi](t,S_t) = \frac{1}{u}\left(\mathbb{E}^{\mathbb{P}_{t+u}}[e^{-r(T-t-u)}\Phi(S_t \oplus_{t+u} S_t(t+u)X)] - \mathbb{E}^{\mathbb{P}_t}[e^{-r(T-t)}\Phi(S_t \oplus_t S_t(t)X)]\right) \tag{4.112}$$

$$= \frac{1}{u}\left(\mathbb{E}^{\mathbb{P}_{t+u}}[e^{-r(T-t-u)}\Phi(S_t \oplus_{t+u} S(t)X)] - \mathbb{E}^{\mathbb{P}_t}[e^{-r(T-t)}\Phi(S_t \oplus_t S(t)X)]\right) \tag{4.113}$$

$$\to \frac{\partial}{\partial t}v_\Phi^{BS}(t,s) \tag{4.114}$$

as expected.

Following the logic of Theorem 4.8, we may now introduce the corresponding notion of the *horizontal integral.*

Definition 4.21. The horizontal integral of a functional process F is defined for each $v > 0$ as

$$\left(\int_0^v du H(u)F\right)(t,x) = \int_0^{v\wedge(T-t)} du F(t+u, x_t) \tag{4.115}$$

Note that $\pi_t \int_0^v du H(u)F = \int_0^v du H(u)F$, so for each $v > 0$ horizontal integration by construction defines an operator with range in the *non-anticipative* functional processes. The meaning of this should be clear – at any given instant we integrate *future* values of the functional *along the stopped path*, or in financial terms, 'keeping the market fixed'. Naturally we will want to integrate all such future values, up to the horizon T; so as a shorthand we will write $⨍ F\,(t,x) := \int_0^{T-t} du H(u)F\,(t,x) = \int_0^{T-t} du F(t+u,x_t)$.

The following is an analogue of the fundamental theorem of calculus.

Proposition 4.22. *(Horizontal fundamental theorem of calculus)*

The horizontal integral is the right-inverse of $-\mathcal{D}$*:*

$$-\mathcal{D} \fint F\,(t, x_t) = F(t, x_t) \tag{4.116}$$

for all $F \in \mathbb{D}(T; C_b)$.

Proof. If $F \in \mathbb{D}(C_b)$, then $t \to ||F(t)||$ is càdlàg on $[0, T]$, hence bounded, and so in particular $F(., x)$ is integrable on $[0, T]$ for each path x. For fixed t, x, we have

$$\frac{H(u) - I}{u} \fint F\,(t, x) = \frac{1}{u}\Big(\fint F\,(t + u, x_t) - \fint F\,(t, x)\Big) = -\frac{1}{u}\int_0^u dv F(t + v, x_t) \tag{4.117}$$

which converges to $-F(t, x_t)$ by the Lebesgue differentiation theorem. □

Chapter 5

No Arbitrage Valuation on Monte Carlo Model Paths

In this Chapter we address Objective 3 of the thesis Aims, by developing a methodology for pathwise no arbitrage valuation on Monte Carlo paths in a wide class of *path-dependent* models, namely Ito processes.

The value of a portfolio in which a European derivative is hedged using the Black-Scholes delta, while the 'real' stock follows the Itô process 2.10, has (discounted) hedging P&L given by

$$\text{P\&L} = \frac{1}{2}\int_0^T \frac{\partial^2 V}{\partial s^2} S^2(t)(\sigma^2 - \sigma_r^2)dt \tag{5.1}$$

where σ_r is the *realised* quadratic variation. In other words, the trader's book evolves according to the error in estimation of the realised volatility by a Black-Scholes *implied* volatility, and relatively easily calculable differentials of Black-Scholes values. The efficiacy of this result means it has been called the 'Fundamental Theorem of Derivatives Trading', (Ellersgaard et al. (2017)). It shows that "the option trader's job is really about balancing realised against implied (or pricing) volatility", according to Andreasen (2003). Indeed, "it is hard to imagine that the derivatives industry could exist at all without some result of this kind", according to Davis (2010).

But what if the stock does not follow an Itô process? The constraints placed on time series to ensure the convergence of the pathwise variance are more general than this – real world stock prices may not follow a semimartingale, or even a cádlág path. Can we say anything about the robustness of trading strategies in this case?

This Chapter is structured as follows. In Section 5.1 we show how a self-financing condition can be naturally incorporated into a purely pathwise framework. In Section 5.2 we derive a new *Itô formula* for functionals of Monte Carlo paths. Finally, in Section 5.3 we use this result to derive a valuation equation for functionals of Itô processes, and a corresponding hedging robustness formula for a Black-Scholes style hedging strategy.

5.1 Self-financing trading strategies along a partition

Fix a time horizon T, and consider a market of d risky assets with positive spot prices given by the paths $\tilde{S}^i : [0,T] \to \mathbb{R}_+$, $i = 1,\dots,d$, and a strictly positive *numéraire* asset, which for simplicity we may think of as a default-free cash account, and whose spot price follows the path $\tilde{S}^0 : [0,T] \to (0,\infty)$. The corresponding *discounted* spot values are then $S^i = \frac{\tilde{S}^i}{\tilde{S}^0} \geq 0$, $i = 1,\dots,d$, and we set $S^0(t) \equiv 1$. Our fundamental modelling assumption is that the log discounted prices $x(t) = \log S(t)$ are càdlàg, so that $S \in \mathbb{D}^d_+$. Recall from Chapter 4 the *predictable part* of S prior to t is

$$S_{t-}(u) = \begin{cases} S(u) & u < t \\ S(t-) & u \in [t,T] \end{cases} \tag{5.2}$$

Let $\mathfrak{p} := \{\mathfrak{p}_k \mid 0 = \mathfrak{p}_0 < \mathfrak{p}_1 \leq \ldots < \mathfrak{p}_{m(\mathfrak{p})-1} = 1\}$ be a partition. It will be useful to set $\underline{\mathfrak{p}}(t) = \max \mathfrak{p} \cap [0,t) = \mathfrak{p}_{\underline{k}(t)}$ to be the last partition point strictly before t with corresponding index $\underline{k}(t)$, $\mathfrak{p}(t) = \max \mathfrak{p} \cap [0,t] = \mathfrak{p}_{k(t)}$ to be the last partition point up to and including t with corresponding index $k(t)$, and $\overline{\mathfrak{p}}(t) = \min \mathfrak{p} \cap (t,T] = \mathfrak{p}_{\overline{k}(t)}$ the first partition point strictly after t, with corresponding index $\overline{k}(t)$.

Given a map $\varphi : \mathfrak{p} \to C_b(\mathbb{D}^d)^d$, we may define a d-dimensional predictable functional process by setting $\mathfrak{H}(0) \equiv 0$, and $\mathfrak{H}(t) = \pi_{\underline{\mathfrak{p}}(t)-}\varphi(\underline{\mathfrak{p}}(t))$ for $t \in (0,T]$, or more explicitly,

$$\mathfrak{H}(t,S) := \varphi(\underline{\mathfrak{p}}(t), S_{\underline{\mathfrak{p}}(t)-})1_{(0,T]}(t) \tag{5.3}$$

where as usual we write the evaluation of functional processes $F(t,S)$ rather than $F(t)(S)$. We will call such an $\mathfrak{H}$ a *simple risky position* or just a *simple position*. It represents a 'buy-and-hold' portfolio, with $\varphi^i(\mathfrak{p}_k, S_{\mathfrak{p}_k-})$ the number of shares in stock i held immediately after $\mathfrak{p}_k$, up until and inclusive of $\mathfrak{p}_{k+1}$. A negative holding is to be interpreted as a 'short sale', and we assume there are no restrictions on such transactions. Note that $\mathfrak{H}$ is bounded, with sup norm $||\mathfrak{H}|| = \max_{t\in\mathfrak{p}} \sup_{S\in\mathbb{D}^d_+} ||\varphi(t,S_{t-})|| < \infty$.

From the definition of $\underline{\mathfrak{p}}(t)$, we see that for sufficiently small $\varepsilon > 0$ we have $\mathfrak{H}(t-\varepsilon) = \mathfrak{H}(t)$, and

$$\mathfrak{H}(t+\varepsilon,S) = \begin{cases} \mathfrak{H}(t,S) & t \notin \mathfrak{p} \\ \varphi(\mathfrak{p}_k, S_{\mathfrak{p}_k-}) & t = \mathfrak{p}_k \end{cases} \tag{5.4}$$

$$= \varphi(\mathfrak{p}(t), S_{\mathfrak{p}(t)-}) \tag{5.5}$$

So $\mathfrak{H} \in \pi\_\mathbb{F}(C_b(\mathbb{D}^d))$ and the right limit $\mathfrak{H}(t+) = \lim_{\varepsilon\downarrow 0}\mathfrak{H}(t+\varepsilon)$ is also a well-defined predictable functional process. For fixed S, the path $t \to \mathfrak{H}(t,S) \in \mathbb{R}^d$ is piecewise constant and càglàd, and the path $t \to \mathfrak{H}(t+,S)$ is piecewise constant and càdlàg.

A trader may hold such a position by financing it via the cash account. The 'mark-to-market', or liquidation value, of the risky position is the non-anticipative functional process given by taking the dot product with the instantaneous spot price:

$$\mathcal{L}_{\mathfrak{H}}(t,S) := \mathfrak{H}(t,S)S(t) = \varphi(\underline{\mathfrak{p}}(t), S_{\underline{\mathfrak{p}}(t)-}).S(t) \tag{5.6}$$

for $t > 0$, and $\mathcal{L}_{\mathfrak{H}}(0) \equiv 0$. The cost of each change in the position is calculated as the difference in liquidation value, $\mathcal{L}_{\mathfrak{H}}(t+,S) - \mathcal{L}_{\mathfrak{H}}(t,S) = \big(\varphi(\mathfrak{p}(t), S_{\mathfrak{p}(t)-}) - 1_{(0,T]}(t)\varphi(\underline{\mathfrak{p}}(t), S_{\underline{\mathfrak{p}}(t)-})\big).S(t)$. Income from selling excess holdings is deposited, and the funds to make up any shortfall are borrowed, we assume without a credit limit. The cumulative *funding cost* non-anticipative functional process representing the net outflow from the cash account is then

$$\mathfrak{F}_{\mathfrak{H}}(t,S) := \sum_{u\in[0,t]} \mathcal{L}_{\mathfrak{H}}(u+,S) - \mathcal{L}_{\mathfrak{H}}(u,S) = \sum_{u\in[0,t]} \big(\mathfrak{H}(u+,S) - \mathfrak{H}(u,S)\big).S(u) \tag{5.7}$$

$$= \varphi(\mathfrak{p}_0, S_0)S(0) + \sum_{k=0}^{m} \big(\varphi(\mathfrak{p}_{k+1}, S_{\mathfrak{p}_{k+1}-}) - \varphi(\mathfrak{p}_k, S_k)\big).S(\mathfrak{p}_{k+1})\beta[\mathfrak{p}_{k+1}](t) \tag{5.8}$$

where each sum has only finitely many non-zero terms, and we recall $\beta[s] = 1_{[s,T]}$.

A trader with initial cash account balance v at $t = 0$, who then holds a simple risky position $\mathfrak{H}$ together with a cash account balance $v - \mathfrak{F}_{\mathfrak{H}}$ is said to be following a *self-financing simple trading strategy*, or just a *simple trading strategy*, as we will always the self-financing condition unless otherwise stated. The *value* of such a strategy is then simply the sum $\mathcal{V} = \mathcal{V}_{\mathfrak{H},v} = \mathcal{L}_{\mathfrak{H}} + v - \mathfrak{F}_{\mathfrak{H}}$ of the liquidation value of the risky position and the net cash balance. The corresponding *gains* functional process is defined as

$$\mathcal{G}_{\mathfrak{H}}(t,S) := \varphi(\underline{\mathfrak{p}}(t), S_{\underline{\mathfrak{p}}(t)-}).\big(S(t) - S(\underline{\mathfrak{p}}(t))\big) + \sum_{k=1}^{\underline{k}(t)} \varphi(\mathfrak{p}_{k-1}, S_{\mathfrak{p}_{k-1}-})(S(\mathfrak{p}_k) - S(\mathfrak{p}_{k-1})) \tag{5.9}$$

$$= \varphi(\underline{\mathfrak{p}}(t), S_{\underline{\mathfrak{p}}(t)-}).S(t) + \sum_{k=0}^{\underline{k}(t)} \varphi(\mathfrak{p}_{k+1}, S_{\mathfrak{p}_{k+1}-})S(\mathfrak{p}_{k+1}) - \sum_{k=1}^{\underline{k}(t)} \varphi(\mathfrak{p}_k, S_{\mathfrak{p}_k-})S(\mathfrak{p}_{k+1}) - \varphi(\mathfrak{p}_0, S_0)S(0) \tag{5.10}$$

$$= \mathcal{L}_{\mathfrak{H}}(t,S) - \mathfrak{F}_{\mathfrak{H}}(t,S) \tag{5.11}$$

after relabelling indices appropriately, so that $\mathcal{V}_{\mathfrak{H},v} = v + \mathcal{G}_{\mathfrak{H}}$.

It is natural to consider the vertical differentiability of the gains functional process of a simple trading strategy. Since $(S + \varepsilon\beta^i[t])_{t-} = S_{t-}$, we have $\mathfrak{H}(t, S + \varepsilon\beta^i[t]) = \mathfrak{H}(t,S)$, and in fact $\mathfrak{H}(t+, S + \varepsilon\beta^i[t]) = \mathfrak{H}(t,S)$. So

$$(B_t^i(\varepsilon) - I)\mathcal{L}_{\mathfrak{H}}(t,S) = \varepsilon\varphi^i(\underline{\mathfrak{p}}(t), S_{\underline{\mathfrak{p}}(t)-}) = \varepsilon\mathfrak{H}^i(t,S) \tag{5.12}$$

while

$$(B_t^i(\varepsilon) - I)\mathfrak{F}_{\mathfrak{H}}(t,S) = \varepsilon(\mathfrak{H}^i(t+,S) - \mathfrak{H}^i(t,S)) \tag{5.13}$$

and hence

$$\partial^i_{x_t}\mathcal{V}_{\mathfrak{H}}(t,S) = \mathfrak{H}^i(t+,S) = \varphi^i(\mathfrak{p}(t), S_{\mathfrak{p}(t)-}). \tag{5.14}$$

for all v, and similarly $\mathfrak{H}(t,S) = \nabla_{t-}\mathcal{V}_{\mathfrak{H}}(t-,S_{t-})$. Because of the predictability of $\mathfrak{H}$, we have $\partial^i_{x_t}\partial^j_{x_t}\mathcal{V} = 0$, so that in fact $\mathcal{V}_{\mathfrak{H}}(t) \in C^\infty_t(\mathbb{D}^d)$. Alternatively, if V is a functional process such that $V(t) \in C^1_t$, we may set $\varphi(t) = \nabla_t V(t)$ so that $\mathfrak{H}(t,S) = \pi_{\underline{\mathfrak{p}}(t)-}\nabla_{\underline{\mathfrak{p}}(t)}V(\underline{\mathfrak{p}}(t)) = \nabla_{\underline{\mathfrak{p}}(t)}V(\underline{\mathfrak{p}}(t), S_{\underline{\mathfrak{p}}(t)-})$ is a simple trading strategy. For this reason, we will interchangeably refer to simple trading strategies and simple *delta* strategies. They are the basic building blocks for approximating more general functional processes representing the value of derivatives contracts.

5.2 A functional Itô formula for continuous Monte Carlo paths

The key tool for deriving the hedge portfolio of a European contract in the classical Merton set-up is *Itô's lemma*. Here, we prove a new version for Monte Carlo paths, using our pathwise variance rather than the quadratic variation. The proof is very much like Föllmer's, Dupire's and Cont and Fournie, with the only change being the mode of convergence of the variance and less stringent requirements on its Lebesgue decomposition.

The first result of this section is based on the simple idea, originally due to Dupire but also used in the proof of the functional Itô formula of Cont and Fournie, that when evaluated on piecewise constant paths based on a fixed partition, we can break the increment of the value functional process into 'vertical' and 'horizontal' components. In our set-up this appears as the action of our bumping and horizontal translation (semi)groups. We break it out as its own result here for two reasons: first, because it serves as a 'road map' for the proof of the Functional Itô formula 5.3, and second because any strategy or simulation in the real world must of course be based on such a finite partition. For this reason we believe it merits its own statement.

Lemma 5.1. *(The staircase lemma)*

Let $\tilde{V} \in \mathbb{D}^1(\mathbb{Y})$ be a horizontally differentiable, càdlàg functional process with values in a Banach space $\mathbb{Y}$ that is a $\boldsymbol{B}_t$-space for each t in a partition $\mathfrak{p}$, and $\tilde{V}(t) \in \mathbb{Y}^2_t$. Let $V = \pi\tilde{V}$ be its non-anticipative part, and $\mathfrak{H} = \nabla_{\underline{\mathfrak{p}}(t)}V(\underline{\mathfrak{p}}(t), S_{\underline{\mathfrak{p}}(t)-})$ the corresponding self-financing delta trading strategy. Then if $S \in \mathbb{S}^{\mathfrak{p}}_r$,

$$\begin{aligned}\mathcal{G}_{\mathfrak{H}}(T,S) = \sum_{k=1}^{m(\mathfrak{p})} \nabla_{\mathfrak{p}_k}V(\mathfrak{p}_k, S_{\mathfrak{p}_k-}).\delta^{\mathfrak{p}}_k S = &\sum_{k=0}^{m(\mathfrak{p})} V(\mathfrak{p}_{k+1},S) - V(\mathfrak{p}_k,S) \\ &- \sum_{k=0}^{m(\tau^n)} DV(\mathfrak{p}_k,S)(\mathfrak{p}_{k+1}-\mathfrak{p}_k) \\ &- \frac{1}{2}\sum_{k=0}^{m(\tau^n)} \mathrm{Tr}\,\partial^{ij}_{\mathfrak{p}_k}V(\mathfrak{p}_k,S)[S]^{ji}_{\mathfrak{p}} \\ &+ r_n \end{aligned} \tag{5.15}$$

where $r_n \leq \varepsilon_n[S]_{\mathfrak{p}}(T)$ and $\varepsilon_n \to 0$ as $n \to \infty$.

Proof. We construct the sums by going down a 'staircase' of increments, using the vertical and horizontal translation operators. Since $S_{\mathfrak{p}_{k+1}} = S_{\mathfrak{p}_k} + \delta_k^{\mathfrak{p}} S\beta[\mathfrak{p}_{k+1}]$, by the non-anticipativity of V we have

$$\begin{aligned} V(\mathfrak{p}_{k+1}, S) - V(\mathfrak{p}_k, S) &= V(\mathfrak{p}_{k+1}, S_{\mathfrak{p}_{k+1}}) - V(\mathfrak{p}_{k+1}, S_{\mathfrak{p}_k}) \\ &\quad + V(\mathfrak{p}_{k+1}, S_{\mathfrak{p}_k}) - V(\mathfrak{p}_k, S_{\mathfrak{p}_k}) \\ &= \big(B_{\mathfrak{p}_{k+1}}(\delta_k^{\mathfrak{p}} S) - I\big) V(\mathfrak{p}_{k+1}, S_{\mathfrak{p}_k}) \\ &\quad + \big(H(\delta_k \mathfrak{p}) - I\big) V(\mathfrak{p}_k, S) \end{aligned} \tag{5.16}$$

Using the vertical Taylor theorem 4.51 to expand the first increment and the Horizontal Fundamental Theorem of Calculus the second gives

$$\begin{aligned} V(\mathfrak{p}_{k+1}, S) - V(\mathfrak{p}_k, S) = \nabla_{\mathfrak{p}_k} V(\mathfrak{p}_k, S_{\mathfrak{p}_k -})\delta_k^{\mathfrak{p}} S + \frac{1}{2}\delta_k^{\mathfrak{p}} S^{\dagger}.\partial^i_{\mathfrak{p}_k}\partial^j_{\mathfrak{p}_k} V(\mathfrak{p}_k, S)\delta_k^{\mathfrak{p}} S + r_k^n \\ + DV(\mathfrak{p}_k, S)\delta_k \mathfrak{p} \end{aligned} \tag{5.17}$$

with r_k^n the remainder term from the vertical Taylor theorem. Using Theorem 4.51 we see that

$$|r_k^n| \le C|\delta_k^{\mathfrak{p}} S| \sup_{u\in[0,\varepsilon]} |\nabla^2_{S_t} V(\mathfrak{p}_k, S + u\beta[\mathfrak{p}_k]) - \nabla^2_{S_t} V(\mathfrak{p}_k, S)| \tag{5.18}$$

where C bounds both of the uniformly continuous functionals $|\mathscr{D}V(t,S)|$ and $|\nabla^2_{S_t} V(t,S)|$. Summing over k, setting $r_n := \sum_k r_k^n$ and recognising the term in the first-order vertical differential as $\mathcal{G}_{\mathfrak{H}}(T,S)$, gives formula 5.1. □

Note the use of the vertical Taylor theorem is exactly analogous to the use of the ordinary real-variable version in Itô's original proof. Lemma 5.1 gives the gains process at maturity of a delta strategy along a partition as a sum of four terms, each given a line in the statement of the result.

The rest of the section is concerned with taking a continuum limit of this expression; the idea is to approximate a class of spot price paths by the piecewise constant Euler projection path $\mathcal{E}^{\mathfrak{p}^n}(S)$, with jumps at the times of a clock, that will ensure the convergence of each term under an appropriate limiting process. The first is a sum of *jumps* in the value process, while the second is a 'theta' term, involving the horizontal differential DV. The third is a 'gamma' term involving the 'Hessian' matrix $\nabla^{ij}_{\mathfrak{p}_k} V$ of second order vertical differentials and the variance computed along the partition, while the fourth is a remainder term related to this. It turns out that the convergence of is a slightly delicate issue, so we approach it in stages.

We begin with the easiest 'theta' term, which we approach in the full generality of a jumpy path with Euler projection along a trading clock. Note that we do not make any requirements on the convergence of the $\mathfrak{p}$-variance here.

Lemma 5.2. *Let V be as in Lemma 5.1 and τ a trading clock. Then for any $S \in \mathbb{D}^d_+$*

$$\lim_{n\to\infty} \sum_{k=0}^{m(\tau^n)} \mathscr{D}V(\tau^n_k, \mathcal{E}^\tau_n(S))\delta_k\tau^n = \int_0^T \mathscr{D}V(u-, S)du \tag{5.19}$$

Proof. By the strong convergence of A^τ_n to the identity, and since $\mathscr{D}V(t, .)$ is uniformly continuous, for fixed S we have that the (bounded) function $u \to \mathscr{D}V(\tau^n_k(S), A^\tau_n(S))1_{[\tau^n_k(S),\tau^n_{k+1}(S))}(u) \to \mathscr{D}V(u-, S)$ at any continuity point of S. Since $\mathcal{J}(S)$ is countable it has Lebesgue measure zero, so we may conclude by the dominated convergence theorem. □

We may now consider our main result, the full Itô formula.

Theorem 5.3. *(Functional Itô formula for continuous Monte Carlo paths)*

Suppose $S \in \mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_c$ has pathwise variance $[S]_{\mathscr{P}}(u) := Q(u) + J([S]_{\mathscr{P}})(u)$ along $\mathscr{P}$, with Q continuous. Suppose $V = \pi V \in \mathbb{D}^1(\mathbb{Y})$ is a non-anticipative functional process with values in a Banach space $\mathbb{Y}$ which is a B_t-space for each $t \in \mathscr{P}$, such that $V(t) \in \mathbb{Y}^2_t$ for each $t \in \mathscr{P}_\infty$. Then $\int_0^T \nabla_u V(u-, S).d^{\mathscr{P}} S(u)$ exists and

$$V(T, S) = V(0, S_0) + \sum_{u\in[0,T]} \Delta V(u, S) \tag{5.20}$$

$$+ \int_0^T \nabla_u V(u-, S_{u-}).d^{\mathscr{P}} S(u) \tag{5.21}$$

$$+ \int_0^T \mathscr{D}V(u-, S)du + \frac{1}{2}\int_0^T \mathrm{Tr}\left(\partial^i_{x_u}\partial^j_{x_u} V\, Q(S)^{ji}_{\mathscr{P}}(du)\right) \tag{5.22}$$

$$+ \frac{1}{2}\sum_{u\in[0,T]} \mathrm{Tr}\, \partial^i_{x_u}\partial^j_{x_u} V J([S]^{ij}_{\mathscr{P}})(u) \tag{5.23}$$

where the sums have countably many non-zero terms and converges absolutely.

Proof. We consider the expression 5.15 calculated on the Euler projections $\mathcal{E}^{\mathfrak{p}^n}(S)$ along the clock $\mathscr{P} = (\mathfrak{p}^n)$.

By Lemma 5.2, the 'theta' term $\sum_{k=0}^{m(\tau^n)} DV(\mathfrak{p}_k, \mathcal{E}^{\mathfrak{p}}(S))(\mathfrak{p}_{k+1} - \mathfrak{p}_k)$ converges as required. Since $[S]_{\mathscr{P}} \in BV^d$, by Lemma 4.17 the path $t \to \nabla^{ij}_u V(t, .) \in \mathbb{F}^d$. So by 3.18 and the fact that the weak-star topology on BV^d as the dual of $\mathbb{F}^d$ coincides with pointwise convergence, the 'gamma' term $\frac{1}{2}\sum_{k=0}^{m(\mathfrak{p}^n)} \mathrm{Tr}\, \partial^{ij}_{\mathfrak{p}^n_k} V(\mathfrak{p}^n_k, \mathcal{E}^{\mathfrak{p}}(S))[\mathcal{E}^{\mathfrak{p}}(S)]^{ji}_{\mathfrak{p}^n}$ converges as required. The value jump term

$$\sum_{u\in[0,t]} \Delta V(u, S) \tag{5.24}$$

converges, since V is càdlàg. By the Taylor Theorem 4.51, and the continuity of S, the error term $|r^n_k| \le \epsilon^n_k|\delta^n_k S|^2$ where $C \ge \epsilon^n_k \to 0$, and so $\sum_k r^n_k \to 0$ using Lemma 5.3.7 of Cont et al. (2016). The remaining term is the delta term, which must converge by virtue of all the others. □

5.3 No arbitrage valuation and robust hedging on $\mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)$

It turns out that the argument of Merton (1973), and the robustness of the Black-Scholes hedging strategy Davis (2010), is very powerful, and goes through almost unchanged in the path-dependent setting. First, we need to recall the notion of arbitrage, and a classical result on its absence in Itô process models.

Taking inspiration from the simplest definition of absence of arbitrage, that of the single period finite state model, we define

Definition 5.4. (No pathwise arbitrage) A self-financing strategy with value V is said to be an *arbitrage opportunity* for the model paths $\mathcal{X}$ if

1. $V(T,x) \geq 0$ for all $x \in \mathcal{X}$
2. There exists $x^* \in \mathcal{X}$, $T_0 < T$ such that $V(0,x^*) \leq 0$ and $V(T_0,x^*) > 0$

It is said to be *admissible* if there exists $c \geq 0$ such that $V(t,x) \geq -c$ for all $x \in \mathcal{X}$. A model with paths $\mathcal{X}$ has *no pathwise arbitrage* if there is no such admissible arbitrage strategy.

The following is Theorem 12.1.8 a) of Oksendal (1992) (with the minor alteration that we have defined coefficients of the model in multiples of current spot price).

Theorem 5.5. *Suppose a market S follows the Itô process model 3.72, and there exists a process θ such that*

$$\sigma\theta = \mu - r \tag{5.25}$$

and

$$\mathbb{E}[\exp(\frac{1}{2}\int_0^T \theta^2(s)ds)] < \infty \tag{5.26}$$

Then the market almost surely permits no arbitrage.

If S follows 3.72, then we let $\mathscr{X}$ be the set of paths as in 3.74. Since $\mathbb{P}[\mathscr{X}^c] = 0$ this implies there is no arbitrage on the paths.

Theorem 5.6. *(Valuation equation and robustness formula)*

Let $\mathscr{X}$ be the paths of an Itô process model 3.72 satisfying the Novikov condition. Suppose a European derivative has payoff $\Phi \in \mathscr{U}_b(\mathbb{M}^d_{\mathscr{P}})$, and V solves the path-dependent Cauchy problem

$$\mathscr{D}V(t,S) + \frac{1}{2}\sum_{i,j=1}^{d} \sigma^i\sigma^j S(t)^i S(t)^j \partial^i_{S_t}\partial^j_{S_t} V(t,S) + \sum_{i=1}^{d} rS^i(t)\partial^i_{S_t} V(t,S) - rV = 0$$
$$V(T,S) = \Phi(S) \tag{5.27}$$

for all $S \in \mathscr{X}$. *Then* V *is the no arbitrage value of the derivative in the model 3.72, and if a rebalanced delta strategy based on the model is followed but the true stock price path is* $S' \in \mathbb{M}^2_{\mathscr{P}}(\mathbb{R}^d)_c$, *then the discounted profit and loss at maturity is*

$$P\&L = \sum_{u\in[0,t]} \Delta V(u,S) + \frac{1}{2}\sum_{i,j=1}^{d}\int_0^T S^i(t)S^j(t)\Gamma V(t,S')(\sigma^{ij}\sigma^{ji}dt - [S']^{ij}(dt)) \tag{5.28}$$

Further, $\Delta V(u,S) \neq 0$ *if and only if* $\Delta[S] \neq 0$.

Proof. We follow the Merton argument. A trader takes a short position in the derivative and marks it using the non-anticipative functional $V(t,S_t)$ offset by a stock position using the non-anticipative functional $\Delta(t,S_t)$. Using the Itô formula 5.23, we expand the profit and loss

$$\begin{aligned} \text{P\&L} = &\int_t^{t+\delta} \mu(u,S_u)S(u)\big(\Delta(u,S_u) - \partial_{S_t}V(u,S_u)\big)\,du \\ &+ \int_t^{t+\delta} \sigma(u,S_u)S(u)\big(\Delta(u,S_u) - \partial_{S_u}V(u,S_u)\big)dW(u) \\ &- \int_t^{t+\delta} \mathscr{D}V(u,S_u) + \frac{1}{2}\sigma^2(u,S_u)S(u)^2\partial^{ij}_{S_u}V(u,S_u)\,du \end{aligned} \tag{5.29}$$

As in the Markovian case choosing $\Delta(t,S_t) = \nabla_{S_t}V(t,S_t)$ yields a portfolio with predictable growth, which must by the absence of arbitrage equate to the return on the riskless portfolio $V(t,S_t) - \nabla_{S_t}V(t,S_t)$, which is exactly 5.27.

Suppose the trader hedges with the model delta $\nabla_{S_t}V_M$, rebalanced at every jump time j of the pathwise variance, then the self-financing portfolio obeys

$$dV = \nabla_{S_t}V_M dS' + (V - \nabla_{S_t}V_M)dS^0 \tag{5.30}$$

while

$$dV_M = \nabla_{S_t}V_M dS' + \mathscr{D}V_M du + \frac{1}{2}(S'(t))^2\Gamma V_M(t,S')[S'](du) \tag{5.31}$$

and hence, by 5.27, the hedging error $Z = V - V_M$ satisfies

$$\begin{aligned} dZ &= \sum_{u\in[0,T]} \Delta Z(u,S) + \big(r(V - \nabla_{S_t}V_M) - \mathscr{D}V_M\big)dt - \frac{1}{2}(S'(t))^2\Gamma V_M(t,S')[S'](dt) \\ &= rZdt + \frac{1}{2}(S'(t))^2\Gamma V_M(t,S')(\sigma^2 dt - [S'](dt)) \end{aligned} \tag{5.32}$$

as required. Discounting yields 5.28. □

Note here that the pathwise variance is completely general, free of any modelling assumption such as being the path of a semimartingale. Essentially the only assumption is that when the stock price variance is calculated along a clock the variance converges. As such this is the formula that best encapsulates model risk robustness, by putting as few constraints on the real data as can be while keeping the second order term.

Chapter 6

Conclusion and Future Work

In this thesis I have shown how the geometry of the numerical scheme known as *Monte Carlo simulation* can naturally be embedded in a separable Banach space, in which piecewise constant paths are dense. Popular risk models known as Itô processes can then be viewed as random samples from concrete sets of paths in this space parameterised by a suite of homogenous noise paths. The spaces have rich classes of functionals, which admit natural strongly continuous (semi)groups of vertical and horizontal translation, whose generators are differential operators that lie behind risk sensitivities. Derivatives valuation can then be understood as the solution of a *Cauchy problem* on this space involving these operators. Unlike existing approaches, the solutions are for *every* path in a given subset, rather than merely $\mathbb{P}$-almost surely.

Several natural generalisations to this work immediately suggest themselves for future research in mathematical finance.

- *Path-dependent semigroups* – in the Cont and Fournié (2010) framework, the classical solution of the path-dependent PDE 2.39 has been shown to be given by a *forward backward stochastic differential equation*. It is the author's conjecture that these can be reinterpreted as *semigroups* on Monte Carlo functional spaces, as in the case of ordinary PDEs, and the solution written as a product formula

$$V(t) = \lim_{n\to\infty} \prod_{t\leq\tau_k^n} e^{-(\tau_k^n - t)\mathcal{A}_n(\tau_{k+1}^n)} \pi_{\tau_k^n \wedge t}\Phi \tag{6.1}$$

  which can best be interpreted as the limit of a sequence of semigroup solutions of ordinary Cauchy problems, and parallels the idea of *mixed Monte Carlo methods*. This was one of the original motivations for this research, but unfortunately time and technical constraints prevented its proof in time for inclusion in this thesis.

- *Jump models* – only liquid markets can be reasonably modelled by continuous spot prices; in general, illiquidity means spot prices are subject to periodic 'shocks', which are modelled by SDEs with 'jumps'. The proof of the Itô formula 5.3 can be extended to the jump case by following the proof of the corresponding result of Cont and Fournié (2010); by analogy with the Markovian case, one would expect the path-dependent case to involve adding a non-local *jump measure* term to the generator of the semigroup

- *Funding costs* – as mentioned in the introduction, it is now well-known that the classical funding assumptions made in this thesis are inadequate in general, and that true derivatives valuation should incorporate the costs of funding. Again, by analogy with the Markovian case, one would expect the path-dependent case to involve a *semilinear* valuation equation. This requires any semigroup solution to be comprised of *nonlinear* operators

- *American options* – only European-style contracts were considered in this thesis. American options involve the possibility of *early exercise*, which considerably complicates the analysis. Again, by analogy with the Markovian case, one would expect the path-dependent case to involve a *fully nonlinear* valuation equation, again solved by semigroups of *nonlinear* operators

In addition, because Monte Carlo methods are widely used, the use of functional calculus on Monte Carlo spaces as in this thesis potentially has applications in many probabilistic fields that use path-dependent stochastic processes, such as evolutionary biology or plasma physics. From the pure mathematical point of view, the treatment of Monte Carlo spaces here has only scratched the surface, and there would appear a great deal of potential for further research in their functional analysis.

References

C. Albanese, H. Lo, and A. Mijatović. Spectral methods for volatility derivatives. *ArXiv e-prints*, May 2009.

C.D. Aliprantis, K.C. Border, and K.C. Border. *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Studies in economic theory. Springer, 1999.

Francesco Altomare and Antonio Attalienti. Degenerate evolution equations in weighted continuous function spaces, markov processes and the black-scholes equation — part i. *Results in Mathematics*, 42(3):193–211, Nov 2002a.

Francesco Altomare and Antonio Attalienti. Degenerate evolution equations in weighted continuous function spaces, markov processes and the black-scholes equation-part ii. *Results in Mathematics*, 42(3):212–228, Nov 2002b.

J. Andreasen. Derivatives – the view from the trenches. *Presentation at the University of Copenhagen*, 2003.

K. J. Arrow. The Role of Securities in the Optimal Allocation of Risk-bearing1. *The Review of Economic Studies*, 31(2):91–96, 04 1964.

Louis Bachelier. Théorie de la spéculation. *Annales scientifiques de l'École Normale Supérieure*, 3e série, 17:21–86, 1900.

Bank for International Settlements. Quarterly statistical bulletin on derivatives securities. 2018.

Emilio Barucci, Fausto Gozzi, and Vincenzo Vespri. On a semigroup approach to no-arbitrage pricing theory. In Robert C. Dalang, Marco Dozzi, and Francesco Russo, editors, *Seminar on Stochastic Analysis, Random Fields and Applications*, pages 1–14, Basel, 1999. Birkhäuser Basel.

Dirk Becherer and Mark H.A. Davis. *Arrow Debreu Prices*. Wiley, 2010.

S. K. Berberian. The character space of the algebra of regulated functions. *Pacific J. Math.*, 74 (1):15–36, 1978.

S. K. Berberian. Regulated functions: Bourbaki's alternative to the riemann integral. *The American Mathematical Monthly*, 86(3):208–211, 1979.

Avi Bick and Walter Willinger. Dynamic spanning without probabilities. *Stochastic Processes and their Applications*, 50(2):349 – 374, 1994.

Tomasz Bielecki and Marek Rutkowski. Valuation and hedging of contracts with funding costs and collateralization. *SIAM Journal on Financial Mathematics*, 6, 2014.

P. Billingsley. *Convergence of Probability Measures*. Wiley Series in Probability and Mathematical Statistics. Wiley, 1968.

Tomas Bjork. *Arbitrage Theory in Continuous Time*. Oxford University Press, 3 edition, 2009.

Phelim P. Boyle. Options: A monte carlo approach. *Journal of Financial Economics*, 4(3):323 – 338, 1977.

Mark Broadie and Paul Glasserman. Estimating security price derivatives using simulation. *Management Science*, pages 269–285, 1996.

Christoph Burgard and Mats Kjaer. In the balance. *Risk*, pages 72–75, December 2011.

Laurie Carver. Traders v theorists. *Risk*, September 2012.

Henry Chiu and Rama Cont. On pathwise quadratic variation for càdlàg functions. *Electron. Commun. Probab.*, 23:12 pp., 2018.

Kinga Cichon, Mieczyslaw Cichon, and Mohamed M. A. Metwali. On some parameters in the space of regulated functions and their applications. *Carpathian Journal of Mathematics*, 34 (1):17–30, 2018.

Fabrizio Colombo, Massimiliano Giuli, and Vincenzo Vespri. A semigroup approach to no-arbitrage pricing theory: Constant elasticity variance model and term structure models. In Mimmo Iannelli and Günter Lumer, editors, *Evolution Equations: Applications to Physics, Industry, Life Sciences and Economics*, pages 113–126, Basel, 2003. Birkhäuser Basel.

Rama Cont, Vlad Bally, and Lucia Caramellino. *Stochastic integration by parts and Functional Ito calculus (Lectures Notes of the Barcelona Summer School on Stochastic Analysis, Centro de Recerca de Matematica, July 2012)*. Springer (Basel), January 2016.

Rama Cont and David-Antoine Fournié. Change of variable formulas for non-anticipative functionals on path space. *Journal of Functional Analysis*, 259(4):1043 – 1072, 2010.

Rama Cont and David-Antoine Fournié. Functional itô calculus and stochastic integral representation of martingales. *Ann. Probab.*, 41(1):109–133, 01 2013.

Rama Cont and Yi Lu. Weak approximation of martingale representations. *Stochastic Processes and their Applications*, 126(3):857 – 882, 2016.

John C. Cox, Stephen A. Ross, and Mark Rubinstein. Option pricing: A simplified approach. *Journal of Financial Economics*, 7(3):229 – 263, 1979.

Mark Davis. Option pricing in incomplete markets. *Mathematics of Derivative Securities*, 01 1997.

Mark Davis, Jan Obłój, and Pietro Siorpaes. Pathwise stochastic calculus with local times. *Ann. Inst. H. Poincaré Probab. Statist.*, 54(1):1–21, 02 2018.

Mark H.A. Davis. *Black–Scholes Formula*. Wiley, 2010.

F. Delbaen and W. Schachermayer. *The Mathematics of Arbitrage*. Springer Finance. Springer Berlin Heidelberg, 2006.

J. Diestel. *Sequences and series in Banach spaces*. Graduate texts in mathematics. Springer-Verlag, 1984.

J. Dieudonné. *Treatise on analysis*. Number v. 8 in Pure and applied mathematics, a series of monographs and textbooks. Academic Press, 1993.

Bruno Dupire. Functional Itô Calculus. *Bloomberg Portfolio Research Paper No. 2009-04-FRONTIERS*, 2009.

N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71.

Simon Ellersgaard, Martin Jönsson, and Rolf Poulsen. The fundamental theorem of derivative trading – exposition, extensions and experiments. *Quantitative Finance*, 17(4):515–529, 2017.

K.J. Engel and R. Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Graduate Texts in Mathematics. Springer New York, 1999.

Matthias A. Fahrenwaldt. Sensitivity of life insurance reserves via markov semigroups. *Scandinavian Actuarial Journal*, 2015(2):124–140, 2015.

G.B. Folland. *Real Analysis: Modern Techniques and Their Applications*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.

H. Föllmer. Calcul d'itô sans probabilites. In Jacques Azéma and Marc Yor, editors, *Séminaire de Probabilités XV 1979/80*, pages 143–150, Berlin, Heidelberg, 1981. Springer Berlin Heidelberg.

Eric Fournié, Jean-Michel Lasry, Pierre-Louis Lions, and JÃĺrÃťme Lebuchoux. Applications of malliavin calculus to monte-carlo methods in finance. ii. *Finance and Stochastics*, 5:201–236, 05 2001.

Mark B. Garman. Towards a semigroup pricing theory. *The Journal of Finance*, 40(3):847–861, 1985.

J. Gatheral. *The Volatility Surface: A Practitioner's Guide*. Wiley Finance. Wiley, 2006.

I. M. Gel'fand. On one-parameter groups of operators in a normed space. *Dokl. Akad. Nauk SSSR*, 25:713–718, 1939.

P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Applications of mathematics : stochastic modelling and applied probability. Springer, 2004.

Oswaldo Gonzalez-Gaxiola and José Santiago. The black–scholes operator as the generator of a c_0-semigroup and applications. 76, 1 2012.

J.Michael Harrison and Stanley R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*, 11(3):215 – 260, 1981. ISSN 0304-4149.

Z-Y. Huang and J.A. Yan. *Introduction to Infinite Dimensional Stochastic Analysis*. Mathematics and its applications. Kluwer Academic, 2000. ISBN 9787030078186.

J. Jacod and A. Shiryaev. *Limit Theorems for Stochastic Processes*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2002.

S. Jazaerli and Y. F. Saporito. Functional Ito Calculus, Path-dependence and the Computation of Greeks. *ArXiv e-prints*, November 2013.

Y. Kabanov and M. Safarian. *Markets with Transaction Costs: Mathematical Theory*. Springer Finance. Springer Berlin Heidelberg, 2009.

H. S. Kaltenborn. Linear functional operations on functions having discontinuities of the first kind. *Bull. Amer. Math. Soc.*, 40(10):702–708, 10 1934.

Ioannis Karatzas, John P. Lehoczky, and Steven E. Shreve. Martingale and duality methods for utility maximization in incomplete markets. *Siam Journal on Control and Optimization - SIAM*, 29, 05 1991.

Takahashi Akihiko Kato, Takashi and Toshihiro Yamada. A semigroup expansion for pricing barrier options. *International Journal of Stochastic Analysis*, 2014.

S. Lang. *Algebra*. Graduate Texts in Mathematics. Springer New York, 2005.

J.M. Lee. *Introduction to Smooth Manifolds*. Graduate Texts in Mathematics. Springer, 2003.

Isaac Leonard. Banach sequence spaces. *Journal of Mathematical Analysis and Applications*, 54:245âĂŞ265, 04 1976.

V. Linetsky. Exotic spectra. *Risk*, April:85–89, 2002.

Vadim Linetsky. The spectral decomposition of the option value. *International Journal of Theoretical and Applied Finance*, 07(03):337–384, 2004.

Vadim Linetsky. Chapter 6 spectral methods in derivatives pricing. In John R. Birge and Vadim Linetsky, editors, *Financial Engineering*, volume 15 of *Handbooks in Operations Research and Management Science*, pages 223 – 299. Elsevier, 2007.

Francis A. Longstaff and Eduardo S. Schwartz. Valuing american options by simulation: A simple least-squares approach. *Review of Financial Studies*, pages 113–147, 2001.

Robert Megginson. *An Introduction to Banach Space Theory*. Graduate Texts in Mathematics. Springer New York, 1998.

Robert C. Merton. Theory of rational option pricing. *The Bell Journal of Economics and Management Science*, 4(1):141–183, 1973.

Yuliya Mishura and Alexander Schied. Constructing functions with prescribed pathwise quadratic variation. *Journal of Mathematical Analysis and Applications*, 442(1):117 – 137, 2016.

Bernt Oksendal. *Stochastic Differential Equations (3rd Ed.): An Introduction with Applications*. Springer-Verlag, Berlin, Heidelberg, 1992.

Gilles Pisier. *Martingales in Banach Spaces*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2016.

V. Piterbarg. Funding beyond discounting: collateral agreements and derivatives pricing. *Risk*, February:97–102, 2010.

D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2004.

A. Schied and I. Voloshchenko. Pathwise no-arbitrage in a class of Delta hedging strategies. *ArXiv e-prints*, October 2015.

Sasha F. Stoikov. Pricing options from the point of view of a trader. *International Journal of Theoretical and Applied Finance*, 09(08):1245–1266, 2006.

N.N. Taleb. *Dynamic Hedging: Managing Vanilla and Exotic Options*. Series in financial economics and quantitative analysis. Wiley, 1997.

Milan Tvrdy. Linear bounded functionals on the space of regular regulated functions. *Tatra Mountains Mathematical Publications*, 8, 01 1996.

J. Voit. *The Statistical Mechanics of Financial Markets*. Physics and astronomy online library. Springer, 2001.

N. Weaver. *Lipschitz Algebras*. World Scientific, 1999.

G. F. Webb. Dual spaces of spaces of quasi-continuous functions. *Mathematische Nachrichten*, 55:309–323, 1973.

P. Wilmott, J.N. Dewynne, and S. Howison. *Option Pricing: Mathematical Models and Computation*. Oxford Financial, 1998.

Jinke Zhou and Xiaolu Wang. Pricing measures, forward measures and semigroups. 9:411–416, 06 2009.