UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

An application of the Malliavin calculus in finance

by

Gordon Fordred

Submitted in partial fulfillment of the requirements for the degree

Magister Scientiae

in the Department of Mathematics and Applied Mathematics
in the Faculty of Natural and Agricultural Sciences

University of Pretoria
Pretoria

July 2009

Contents

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Abstract

This dissertation provides a brief theoretical introduction to the Malliavin calculus leading to a particular application in finance. The Malliavin calculus concepts are used to aid in the simulation of the Greeks for financial contingent claims. Particular focus is placed on creating efficiency in the more exotic type option simulations, where no closed solution pricing formulae exist.

Acknowledgements

I would like to thank Elvis and my mamma.

DECLARATION

I, the undersigned, hereby declare that the dissertation submitted herewith for the degree Magister Scientiae to the University of Pretoria contains my own, independent work and has not been submitted for any degree at any other university.

Gordon Fordred

2009-09-08

Chapter 1

Introduction

1.1 Overview

This dissertation's intention is to provide an overview of the latest methods in optimising the convergence rates of Monte Carlo and Quasi Monte Carlo methods, particularly the use of the Malliavin calculus approach, in estimating sensitivities of derivatives on financial contingent claims. These sensitivities are more commonly known as the "Greeks".

To this end, a basic overview of the Malliavin calculus will be elaborated upon, exposing its required useful properties.

The modern finance industry has induced an emphasis on risk management, hence, as more advanced and sophisticated financial instruments are developed, so the need arises to develop efficient methods to quantify their sensitivities. Not only are these "Greeks" useful in risk management, they also aid in hedging strategies and help provide price quotes for the derivative product, since bid-ask spreads are usually a percentage of some "Greek".

In many cases, due to the complexity of the product (particularly those with a discontinuous pay-off function and payoff parameters with unknown densities), these measures of risk cannot be expressed as a closed form and numerical methods such as Monte Carlo are required.

The most recent major work in the field was the concept of Malliavin weighting functions by Fournié et al. [6], however Broadie and Glasserman [8] have also provided an excellent method namely the Likelihood Ratio method, for more restrictive situations.

To overcome poor Monte Carlo convergence, many methods have been suggested, such as:

- Broadie and Glasserman [8] differentiate the density function, resulting in the Likelihood Ratio method

- Fournié et al. [6] suggested smoothing the function to be estimated, using integration-by-parts. The results are similar to those of Elworthy [5], who shifted the differential operator from the payoff function to the diffusion kernel, introducing the weighting function. This work implies that the "Greeks" can be expressed as the expected value of a discounted payoff multiplied by a weighting function:

$$\boxed{\text{Greek} = \mathbb{E}_{\mathbb{P}}\left[e^{-\int_0^T r_t dt}\Phi(\alpha)\cdot weight\right]}$$

1.2 Layout

The layout of this document will include an overview of the Malliavin calculus, its setting and a discussion of the associated properties of the method. Finally, an application of the Malliavin calculus to the "Greeks" will be explained in conjunction with the Monte Carlo methods used to optimise the sensitivity estimation.

1.3 What is the Malliavin calculus?

The Malliavin calculus is an infinite dimensional differential calculus on the Wiener Space. It is also known as the Stochastic Calculus of Variations.

1.3.1 History

The original Stochastic Calculus of Variations, now called the Malliavin calculus, was developed by Paul Malliavin in 1976.

The theory was further developed by Bismut [3], Stroock [13], Ustünel [14] and Watanabe [15] amongst others. It was originally designed to study the smoothness of the densities of solutions of stochastic differential equations. One of its striking features is that it provides a probabilistic proof of the celebrated Hörmander theorem, which gives a condition for a partial differential operator to be hypoelliptic. This illustrates the power of this calculus. In the following years many probabilists worked on this topic and the theory was developed further, either as analysis on the Wiener space or in a white noise setting.

Since then, the Malliavin calculus has raised increasing interest and subsequently many of its applications to finance have been found, such as minimal variance hedging and Monte Carlo methods for option pricing. More recently, the Malliavin calculus has also become a useful tool for studying insider trading models and some extended market models driven by Lévy processes or fractional Brownian motion. The "enlargement of filtration" technique plays an important role in the modelling of such problems and the Malliavin calculus can be used to obtain general results about when and how such filtration enlargement is possible. Moreover, when the additional information of the insider is generated by adding the information about the value of one extra random variable, the Malliavin calculus can be used to explicitly find the optimal portfolio of an insider for a utility optimisation problem with logarithmic utility.

1.3.2 Key results

One major result is that the adjoint operator of the Malliavin derivative operator, called the Skorohod integral, has the property of being the extension of the Itô integral for non-adapted processes. One of the most important points in the theory is the integration-by-parts formula, which relates the derivative operator on the Wiener Space and the Skorohod extended stochastic integral.

Malliavin calculus was thus seen as a starting point to developing stochastic calculus for non-adapted processes. This means that stochastic differential equations can be formulated and described where the solution is not adapted to the Brownian filtration.

1.3.3 General setting

The aim is to develop a probabilistic differential stochastic calculus over an infinite dimensional space. This infinite dimensional space will typically be the classical Wiener Space, $(C_0[0,T], \mathcal{F}, \mu)$.

When looking at the Malliavin calculus one has to consider the theory of differential operators defined on suitable Sobolev spaces of Wiener functionals.

Heuristically, Malliavin calculus aims to address quantities such as $\frac{dF}{d\omega}$, where $F \in L^2$ and $\omega \in \Omega$. To define such a term over finite dimensional spaces is relatively straight forward, however, we would like to extend this essentially classical functional calculus to infinite dimensional spaces like $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Malliavin calculus effectively defines the derivatives of functions on the Wiener Space and is also considered a theory of integration-by-parts on the Wiener Space. It further allows us to compute the derivatives of a large set of random variables and processes, whether or not they are adapted to the filtration defined on the Wiener Space.

There are two ways of introducing the Malliavin calculus:

- Using analysis on the Wiener Space describing the Wiener Itô chaos expansion, the n-fold iterated integrals, symmetric functions and the Hermite relations,
- The basic simple process-simple function method with the extension of the operator domain by density.

The latter method will be exposed, since the former is more applicable to the classical notions of the Malliavin calculus and results in areas not needed to support the application of the Malliavin calculus in finance, which the latter illustrates effectively and efficiently for this dissertation.

1.3.4 Financial applications

Some known financial applications exist and include: risk management (sensitivity analysis, stochastic volatility models, insider trading and hedging methods), optimal portfolio theory, conditioning and econometrics. This dissertation will primarily highlight sensitivity analysis.

Chapter 2

Malliavin calculus concepts

2.1 The Wiener Space and the Wiener Measure

For simplicity, we will be working in one dimension only - extensions to higher dimensions follow naturally. In the definition of the Brownian motion we have assumed that the usual probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is arbitrary. Furthermore, the probability space is very abstract, making differentiation concepts on it hard to define. We therefore need a sufficient structure and it is the Wiener Space which has this desired "structure". To commence defining this structure, recall the Wiener process, which will form the basis of all further arguments:

Definition 2.1.1 (Wiener Process). *A continuous-time stochastic process $W(t)$ for $t \geq 0$ with $W(0) = 0$ and such that the increment $W(t) - W(s)$ is Gaussian with mean 0 and variance $t - s$ for any $0 \leq s < t$, and increments for non- overlapping time intervals are independent. Brownian motion (i.e. random walk with random step sizes) is the most common example of a Wiener process.*

Definition 2.1.2 (Wiener Space). *Let $\Omega := C_0([0,T])$ denote the space of real continuous functions on $[0,T]$ with the value 0 at time $t = 0$, i.e.*

$$\Omega := C_0([0,T]) = \{\omega : [0,T] \to \mathbb{R} | \omega \text{ continuous}, \ \omega(0) = 0\}$$

Definition 2.1.3 (Wiener Measure). *The Wiener Measure is the probability law on the space of continuous functions, $C_0([0,T])$, induced by the*

Wiener process.

Proposition 2.1.1 (Norm on the Wiener Space). *For a given $\omega \in C_0([0,T])$ we can define the uniform norm on $C_0([0,T])$*

$$\|\omega\|_\infty := \sup_{t\in[0,T]} |\omega(t)|$$

This uniform norm makes $C_0([0,T])$ a Banach space and the dual, $C_0([0,T])^*$ associated with $C_0([0,T])$ can be identified with $\mathcal{M}([0,T])$, the space of signed measures v on $[0,T]$.

Remark 2.1.1. *By the Weierstrass approximation theorem, the polynomials are dense in $C([0,T])$ (the space of real-valued continuous functions). Thus, polynomials with rational coefficients are also dense in $C([0,T])$.*

By considering components, $C([0,T])$ is a separable Banach space. Since every subset of a separable metric space is separable, this implies $C_0([0,T])$ is separable too. The paths followed by the Wiener process lie in $C_0([0,T])$.

2.1.1 The Wiener Measure

A more formal way of viewing the Wiener process is as a stochastic process taking values over the set of all possible trajectories. Let $C_\alpha[a,b]$ be the set of continuous functions f defined on $[a,b]$ with $f(a) = \alpha$. Let $\Omega = C_0[0,T]$ and $\mathcal{F} = \mathcal{B}(C_0[0,b])$, where $\mathcal{B}(C_0[0,b])$ is the σ-algebra generated by open sets (with respect to the sup metric) of $C_0[0,T]$. The filtration $\mathcal{F}_t$ in this case would be a sequence of sub-σ- algebras.

In 1923, Wiener showed that there is a well-defined measure μ on this measure space, known as the Wiener measure. Elements of $C_0[0,T]$ under the Wiener measure correspond to the sample paths of the Brownian motion, and the probability space $(C_0[0,T], \mathcal{F}, \mu)$ is called the classical Wiener space.

Construction

The construction process is based on a variation of Einstein's probabilistic formula. Let the location of a particle at time t be represented by $x = x(t)$.

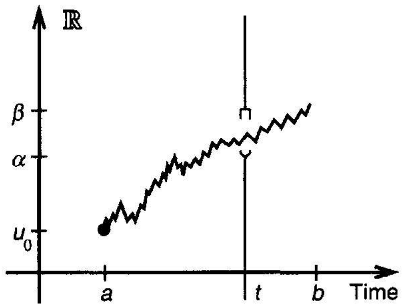

Figure 2.1: Paths starting at u_0 at time a and passing within the interval $(\alpha, \beta]$ at time t

We know that

$$\mathbb{P}[\{\alpha < x(t) \leq \beta\}] = \int_{\alpha}^{\beta} \underbrace{\frac{1}{\sqrt{2\pi t}} e^{\frac{-u^2}{2t}}}_{\text{N(0,t) density}} du$$

We temporarily fix t such that $a < t \leq b$. Suppose $-\infty \leq \alpha < \beta \leq +\infty$. Considering a particle originating at u_0 at time a, we then have, as depicted in Figure 2.1:

$$\mathbb{P}\left[\{\alpha < x(t) \leq \beta \,|\, x(a) = u_o\}\right] = \int_{\alpha}^{\beta} p(u, u_0, t-a) du$$

where

$$p(u, u_0, t) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(u-u_0)^2}{2t}}$$

Clearly, $p(u, u_0, t-a)$ is the $N(u_0, t-a)$ density function.

Wiener wanted to demonstrate the existence of a countably additive probability measure $\mu = \mu_{a,b}$ on $C_0([a, b])$ such that if

$$a = t_0 < t_1 ... < t_n \leq b$$

and if α_j and β_j are extended real numbers so that

$$-\infty \leq \alpha_j < \beta_j \leq +\infty \text{ for } j = 1, 2, ..., n$$

Then,

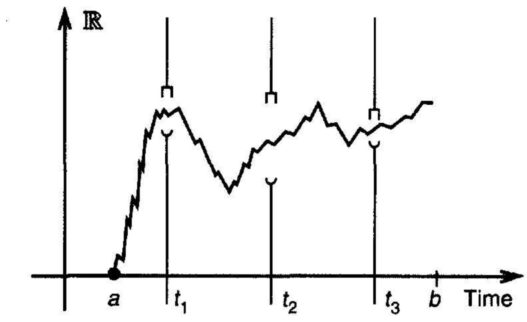

Figure 2.2: Wiener paths starting at 0 at time a and passing through the interval $(\alpha_j, \beta_j]$ at times t_j, for $j = 1, 2, 3$

$$\mu\left(\{x \in C_0([a,b]) : \alpha_j \leq x(t_j) \leq \beta_j, j = 1, 2, ..., n\}\right)$$
$$= \int_{\alpha_n}^{\beta_n} ... \int_{\alpha_2}^{\beta_2} \int_{\alpha_1}^{\beta_1} p(u_1, 0, t_1 - a) p(u_2, u_1, t_2 - t_1) ... p(u_n, u_{n-1}, t_n - t_{n-1}) du_1 ... du_n$$

This can also be written as

$$= \int_{\alpha_n}^{\beta_n} ... \int_{\alpha_1}^{\beta_1} [(2\pi)^n (t_1 - a) ... (t_n - t_{n-1})]^{-\frac{1}{2}} \times e^{-\sum_{j=1}^{n} \frac{(u_j - u_{j-1})^2}{2(t_j - t_{j-1})}} du_1 ... du_n$$

A case of the above scenario is depicted in Figure 2.2 for $n = 3$.

Definition 2.1.4. *For any topological space $\mathcal{T}$, $\mathcal{B} = \mathcal{B}(\mathcal{T})$, the Borel class of $\mathcal{T}$ is the σ-algebra generated by the open subsets of $\mathcal{T}$. We are interested in $\mathcal{B}(C_0([a,b]))$.*

Remark 2.1.2. *The Wiener measure will be a measure defined on a σ-algebra containing the "Borel class" of $C_0([a,b])$.*

The collection of subsets of $C_0([a,b])$ will be denoted by I and will be called *intervals.*

$$I = I(t_1, ..., t_n : (\alpha_1, \beta_1] \times (\alpha_2, \beta_2] \times ... \times (\alpha_n, \beta_n])$$
$$= \{x \in C_0([a,b]) : \alpha_j \leq x(t_j) \leq \beta_j, j = 1, 2, ..., n\}$$

The collection of all such intervals (also called cylinder sets) will be denoted as $\mathcal{I}$. Wiener wanted to develop a Lebesgue type integral over the infinite dimensional space $C_0([a,b])$, hence being able to use Lebesgue theory. It can be shown that $\mathcal{I}$ is a semi-algebra and μ is well defined and countably-additive on $\mathcal{I}$ and further that $\sigma(\mathcal{I}) = \mathcal{B}(C_0([a,b]))$.

The measure space $(C_0([a,b]), \mathcal{B}(C_0([a,b])), \mu)$ can be completed, producing $\tilde{\mathcal{S}}$, the σ-algebra of Wiener measurable sets and the complete measure space $(C_0([a,b]), \tilde{\mathcal{S}}, \mu)$. The above process is known as the Carathéodory extension process.

When considering the interval $[0,T]$, we can clearly understand the reasoning for defining $\Omega = C_0([0,T])$, as the *Wiener space.*

We can visualise each path

$$t \to W(t, \omega)$$

of the Wiener process starting at 0 as an element ω of $C_0([0,T])$, the subspace of real continuous functions. We can thus identify $W(t,\omega)$ with the value $\omega(t)$ at time t of an element $\omega \in C_0([0,T])$, by way of the coordinate map:

$$W(t, \omega) = \omega(t)$$

With this identification, the Wiener process merely becomes the space $\Omega = C_0([0,T])$ and the probability measure $\mathbb{P}$ of the Wiener process becomes the measure μ defined on the cylinder sets of Ω by

$$\mu\left(\{\omega : \omega(t_1) \in F_1, ..., \omega(t_k) \in F_k\}\right) = \mathbb{P}\left(\{W(t_1) \in F_1, ..., W(t_k) \in F_k\}\right)$$

$$= \int_{F_1 \times ... \times F_k} \rho(t_1, x, x_1)\rho(t_2 - t_1, x_1, x_2)...\rho(t_k - t_{k-1}, x_{k-1}x_k)dx_1...dx_k$$

where $F_i \subset \mathbb{R}$ for $i = 1, 2, .., k$; $0 \leq t_1 < t_2 < ... < t_k$ and

$$\rho(t, x, y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2}|x-y|^2}$$

The measure μ is called the *Wiener measure* on Ω. We will write $L^2(\Omega)$ for $L^2(\mu)$ and $L^2([0,T] \times \Omega)$ for $L^2(\lambda \times \mu)$ where λ is the Lebesgue measure on $[0,T]$.

A verbal construction of the Wiener measure

Let $(\Xi, \mathcal{G}, \nu)$ be a probability space and let $\beta = (\beta_t)_{t\in[0,T]}$ be a Brownian motion with respect to the probability measure ν. $\mathcal{G}$ is the natural filtration generated by the Brownian motion β such that $\mathcal{G} = \mathcal{G}_T$ where $\mathcal{G}_t := \sigma(\beta_s | 0 \leq s \leq t)$.

Since Brownian motion is continuous, it can then be regarded as a mapping from Ξ into Ω (the Wiener Space as defined before) via the mapping of ξ from Ξ to the continuous function

$$\xi \times t \to \beta_t(\xi).$$

We now equip Ω with the σ-algebra $\mathcal{F}$ generated by the finite-dimensional cylinder set

$$\{\omega | \omega(t_1) \in A_1, ..., \omega(t_n) \in A_n\}\ , 0 \leq t_1 < ... < t_n \leq T \text{ and } A_1, A_2, ..., A_n \in \mathcal{B}$$

where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$.

Now the Brownian motion, $\xi \to \beta(\xi)$ can be regarded as a measurable mapping from

$$(\Xi, \mathcal{G}, \nu) \text{ to } (\Omega, \mathcal{F})$$

and at the same time inducing a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ given by

$$\mathbb{P}\left(\{\omega | \omega(t_1) \in A_1, ..., \omega(t_n) \in A_n\}\right) = \nu\left(\beta_{t_1} \in A_1, ..., \beta_{t_n} \in A_n\right).$$

This measure is called the Wiener Measure.

Defining the coordinate mapping process

$$B_t : \Omega \to \mathbb{R}$$

on the Wiener Space by

$$B_t(\omega) := \omega(t),$$

we now note that the process

$$\boldsymbol{B} = (B_t)_{t\in[0,T]}$$

has the same distribution under $\mathbb{P}$ as

$$\beta = (\beta_t)_{t\in[0,T]}$$

has under ν.

Hence, $\boldsymbol{B} = (B_t)_{t\in[0,T]}$ is a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. Further to this, $\mathcal{F} = \mathcal{F}_T$, where $\mathcal{F} := \sigma(B_s | 0 \leq s \leq t)$ is the σ-algebra generated by the Brownian motion, $\boldsymbol{B} = (B_t)_{t\in[0,T]}$.

- In essence the coordinate mapping on the Wiener Space becomes a Brownian motion under the Wiener measure.
- Hence, when needing to work with Brownian motion, one can use the coordinate mapping on the Wiener Space instead of some arbitrary abstract probability space.

2.2 Malliavin calculus setting

For simplicity, only the one dimensional case is considered as extensions are clear for higher dimensions.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space and $L^2(\Omega, \mathcal{F}, \mathbb{P})$ denote the set of square integrable random variables on this space. Furthermore the space is equipped with the filtration $(\mathcal{F}_t)$ generated by one-dimensional Brownian motion, W_t.

2.2.1 Functionals of Brownian motion and their spaces

The objective is to differentiate functionals of Brownian motion of the form:

$$F : \Omega \to \mathbb{R}$$

or at least those of a certain "nice" subclass of functions.

Let $L^2([0,T])$ denote the Hilbert space of deterministic square integrable functions $h : [0,T] \to \mathbb{R}$. For $h \in L^2([0,T])$ we define the following random variable

$$W(h) = \int_0^T h(t) dW_t$$

the usual Itô integral with respect to Brownian motion. Note that $W(h)$ is a Gaussian random variable with $\mathbb{E}[W(h)] = 0$ and by the Itô isometry,

$$\mathbb{E}[W(h)W(g)] = \int_0^T h(t)g(t)dt = (h, g)_{L^2([0,T])}$$

Definition 2.2.1. *A stochastic process, $W = \{W(h), h \in L^2([0,T])\}$, defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called an isonormal Gaussian process (on $L^2([0,T])$), if W is a centred Gaussian family of random variables, such that*

$$(W(h), W(g))_\Omega = (h, g)_{L^2([0,T])}$$

for all $h, g \in L^2([0,T])$

The closed subspace $\mathcal{H}_1 \subset L^2(\Omega)$ of such random variables is then isometric to $L^2([0,T])$ and is called the space of zero-mean Gaussian random variables.

Definition 2.2.2. *Let* $C_p^\infty(\mathbb{R}^n)$ *denote the set of infinitely differentiable functions (smooth functions)* $f : \mathbb{R}^n \to \mathbb{R}$*, such that* f *and all its partial derivatives of all orders have polynomial growth.*

We denote by $\mathcal{S}$ *the class of random variables (functionals), of the form*

$$F = f(W(h_1), ..., W(h_n))$$

where $f \in C_p^\infty(\mathbb{R}^n)$ *and* $h_1, ...h_n \in L^2([0,T])$.

Note that $\mathcal{S}$ is then a dense subspace of $L^2(\Omega)$.

The general context then becomes one consisting of a probability space, $(\Omega, \mathcal{F}, \mathbb{P})$ and a Gaussian subspace $\mathcal{H}_1$, of $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

2.2.2 Differentiation on the Wiener Space

As stated before we wish to develop an infinite dimensional differential calculus on a measure space like $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

It is known that Brownian motion is nowhere differentiable with respect to time. One can, however, define the concept of differentiation of random variables with respect to perturbations in the underlying Brownian motion.

We will first consider the directions of the above mentioned perturbations.

The Cameron Martin space

We should recall that the sample space Ω can be identified with the space of continuous functions (the classical Wiener Space). We can then consider the subspace

$$\mathbb{H}^1 = \{\gamma \in \Omega : \gamma = \int_0^t h(s)ds, h \in L^2([0,T])\},$$

or

$$\mathbb{H}^1 = \{h \in C_0[0,T] : h' \in L^2([0,T])\}$$

namely, the space of continuous functions with square integrable derivatives. This space is isomorphic to $L^2([0,T])$ and is called the Cameron-Martin space.

We equip $\mathbb{H}^1$ with an inner product defined by

$$(g,h)_{\mathbb{H}^1} = \int_0^T h'(t)g'(t)dt$$

The Stone Weierstrass theorem, provides a reason for $\mathbb{H}^1$ being dense in $C_0[0,T]$.

Definition 2.2.3 (Cameron Martin subspace directions). *Let $h \in L^2([0,T])$ be a deterministic, square integrable function with respect to the Lebesgue measure $\lambda(dt) = dt$ on $[0,T]$. We will consider directions of the form*

$$\gamma(t) = \int_o^t h(s)ds$$

We see that $t \to \gamma(t)$ is continuous on $[0,T]$ and $\gamma(0) = 0$. Hence $\gamma \in \Omega$ (the Wiener Space) and will therefore be a valid direction.

It turns out that obtaining a theory for derivatives in all directions is still an open problem, hence we will define directional derivatives of random variables in the directions of elements of the Cameron-Martin subspace. This will generalise allowing derivatives in the directions of isonormal Gaussian processes which will be sufficient for our needs. We will now apply the idea of the Fréchet derivative to the classical Wiener Space $(C_0[0,T], \mathcal{F}, \mu)$, where μ is the Wiener measure.

Definition 2.2.4 (Fréchet derivative in the strong sense). *Let $F : \Omega \to \mathbb{R}$ be a random variable. The directional derivative (in the strong sense) of F in the all the directions of $\gamma(t) = \int_0^t h(s)ds$ with $h \in L^2([0,T])$ (elements making up the Cameron Martin subspace) at the point ω is defined by:*

$$\boldsymbol{D}_\gamma F := \lim_{\varepsilon \to 0} \frac{F(\omega + \varepsilon\gamma) - F(\omega)}{\varepsilon} = \frac{d}{d\varepsilon}\left[F(\omega + \varepsilon\gamma)\right]_{\varepsilon=0}$$

if the limit exists in $L^2(\Omega)$.

If there exists a $\psi(t,\omega) \in L^2([0,T]\times\Omega)$ such that

$$\mathbf{D}_\gamma F(\omega) = \int_0^T \psi(t,\omega)h(t)dt$$

then F is said to be *differentiable* and the *derivative* of F is defined to be

$$\boxed{\mathbf{D}_t F := \psi(t,\omega) \in L^2([0,T]\times\Omega)}$$

The set of all differentiable random variables is denoted by $\mathcal{D}_{1,2}$

Another view of differentiating F comes via the following definition:

Definition 2.2.5. *Following from Definition 2.2.2. For $F \in \mathcal{S}$ we define the stochastic process*

$$D_t F := \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(W(h_1),...,W(h_n))h_i(t)$$

It can be shown that $DF \in L^2([0,T]\times\Omega)$

Since D operates on functions in the form of partial derivatives, general properties of chain rule, product rule and linearity are shared.

We now have two "D" operators:

- $$\mathbf{D}_t F := \psi(t,\omega) \in L^2([0,T]\times\Omega)$$
- $$D_t F := \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(W(h_1),...,W(h_n))h_i(t) \in L^2([0,T]\times\Omega)$$

For both definitions we have obtained a linear operator

$$D : \mathcal{S} \subset L^2(\Omega) \to L^2([0,T]\times\Omega)$$

$$\boldsymbol{D} : \mathcal{D}_{1,2} \subset L^2(\Omega) \to L^2([0,T]\times\Omega)$$

To extend the domains of linear operators D and $\boldsymbol{D}$ ($\mathcal{S}$ and $\mathcal{D}_{1,2}$ respectively), we now introduce the following norm, $\|\cdot\|_{1,2}$

$$\|F\|_{1,2} = \|F\|_{L^2(\Omega)} + \|D_t F\|_{L^2([0,T]\times\Omega)}, \qquad F \in \mathcal{S}$$

$$\|F\|_{1,2} = \|F\|_{L^2(\Omega)} + \|\mathbf{D}_t F\|_{L^2([0,T]\times\Omega)}, \qquad F \in \mathcal{D}_{1,2}$$

We define $\mathbb{D}_{1,2}$ as the closure of $\mathcal{S}$ in the norm $\|\cdot\|_{1,2}$. Then

$$D : \mathbb{D}_{1,2} \subset L^2(\Omega) \to L^2([0,T] \times \Omega)$$

is a closed unbounded operator with dense domain $\mathbb{D}_{1,2}$. This will be elaborated on further later.

Remark 2.2.1. *At this point we desire two things:*

- *A general concept of a derivative in more general measure spaces*
- *Hope that $\mathcal{D}_{1,2}$ is a Sobolev space under the $\|\cdot\|_{1,2}$ norm.*

However, derivatives in the sense of Fréchet derivatives provide neither. The reason being that we will be interested in random variables, F that are defined $\mathbb{P}$-a.s.

The Fréchet derivative is implicitly dependent on the continuity of F. We therefore need to adapt our notion of a derivative, to one that is independent of the topological structure of Ω, hence, we need a derivative that acts in the weak sense. When working with the classical Wiener Space, it is clear that the existence of the Fréchet derivative of a random variable F, depends on the existence of a continuous version of F. There are random variables F that do not posses a continuous version.

This shows that $\mathcal{D}_{1,2}$ is not complete and therefore cannot be a Sobolev Space.

To finalise the point, the Fréchet derivative is not sufficient to enable the extension of the theory to a more general setting. The solution involves the introduction of the Malliavin derivative, which is merely the generalisation of the Fréchet derivative defined in the weak sense. The Malliavin derivative provides the solution to both problems previously encountered above.

Remark 2.2.2. *A good analogy of the situation between Fréchet and Malliavin derivatives for a random variable F, is the comparison between Riemann and Lebesgue integration for some function f. Riemann and Fréchet provide both a theoretical foundation and intuitive understanding, but both methods suffer from the same problem of domains being incomplete spaces. Lebesgue and Malliavin's work served to solve this problem.*

At this stage it is not clear whether $\mathcal{D}_{1,2}$ is closed under the defined norm. In other words, whether a $\|\cdot\|_{1,2}$-Cauchy sequence in $\mathcal{D}_{1,2}$ converges to an element of $\mathcal{D}_{1,2}$. To overcome this problem, we analyse the family of random variables, $\mathcal{S}$ (Wiener polynomials), further:

Remark 2.2.3 ($\mathcal{S}$ is dense in $L^2(\Omega)$). *From the martingale convergence theorem and the monotone class theorem, it follows that the set of random variables,*

$$\{f(W_{t_1}, ..., W_{t_n}); t_i \in [0,T], f \in C_p^\infty(\mathbb{R}^n)\}$$

is dense in $L^2(\Omega)$.

$C_p^\infty(\mathbb{R}^n)$, denotes the space of infinitely differentiable rapidly decreasing functions on $\mathbb{R}^n$.

From the above and the Fourier transform, the linear span of the set

$$\left\{ \exp\left(\int_0^T h_s dW_s - \frac{1}{2}\int_0^T h_s^2 ds \right); h \in L^2([0,T]) \right\}$$

is dense in $L^2(\Omega)$. Due to the analyticity of the characteristic function of the Wiener measure, the elements of the above set can be approximated by polynomials, hence the polynomials are dense in $L^2(\Omega)$.

By virtue of the chain rule, the family of Wiener polynomials is differentiable, i.e. $\mathcal{S} \subset \mathcal{D}_{1,2}$.

Lemma 2.2.1. *Let* $F(\omega) = p\left(\int_0^T h_1(t)dW_t, \int_0^T h_2(t)dW_t, ..., \int_0^T h_n(t)dW_t\right) \in \mathcal{S}$*. Then* $F \in \mathcal{D}_{1,2}$ *and*

$$\mathbf{D}_t F(\omega) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} p\left(\int_0^T h_1(t)W_t, \int_0^T h_2(t)dW_t, ..., \int_0^T h_n(t)dW_t\right) \cdot h_i(t)$$

By letting $\theta_i = \int_0^T h_i(t)dW_t$ *this can be rewritten as:*

$$\mathbf{D}_t F(\omega) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} p\left(\theta_1, \theta_2, ..., \theta_n\right) \cdot h_i(t)$$

Proof. Let

$$\psi(t, \omega) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} p\left(\theta_1, \theta_2, ..., \theta_n\right) \cdot h_i(t)$$

Since

$$\sup_{s \in [0,T]} \mathbb{E}\left[|W(s)|^N\right] < \infty \qquad N \in \mathbb{N}$$

We can now look at:

$$\frac{1}{\varepsilon}\left[F(\omega + \varepsilon\gamma) - F(\omega)\right]$$

Recall that $\gamma(t) = \int_0^t g(s)ds$ for some $g \in L^2([0,T])$. Let us first consider $F(\omega + \varepsilon\gamma)$.

$$\begin{aligned} F(\omega + \varepsilon\gamma) &= p\left(\int_0^T h_1(t)d(\omega + \varepsilon\gamma), ..., \int_0^T h_n(t)d(\omega + \varepsilon\gamma)\right) \\ &= p\left(\theta_1 + \int_0^T h_1(t)d(\varepsilon\gamma), ..., \theta_n + \int_0^T h_n(t)d(\varepsilon\gamma)\right) \\ &= p\left(\theta_1 + \varepsilon\int_0^T h_1(s)g(s)ds, ..., \theta_n + \varepsilon\int_0^T h_n(s)g(s)ds\right) \\ &= p\left(\theta_1 + \varepsilon(h_1, g), ..., \theta_n + \varepsilon(h_n, g)\right) \end{aligned}$$

Now let us look at $\mathbf{D}_\gamma(\theta_i)$ for $i = 1, ..., n$

$$\begin{aligned}\mathbf{D}_\gamma(\theta_i) &= \mathbf{D}_\gamma\left(\int_0^T h_i dW_t\right)\\ &= \frac{1}{\varepsilon}\left[\int_0^T h_i d\omega(s) + \varepsilon(h_i, g) - \int_0^T h_i d\omega(s)\right]\\ &= (h_i, g) \text{ for } i = 1, ..., n\end{aligned}$$

So

$$\mathbf{D}_\gamma F(\omega) \longrightarrow \frac{\partial p}{\partial x_1}(\theta_1, ..., \theta_n)\cdot \mathbf{D}_\gamma(\theta_1) + ... + \frac{\partial p}{\partial x_n}(\theta_1, ..., \theta_n)\cdot \mathbf{D}_\gamma(\theta_n) \text{ in } L^2(\Omega)$$

as $\varepsilon \to 0$

Hence,

$$\mathbf{D}_\gamma F(\omega) = \int_0^T \psi(t, \omega) g(t) dt$$

□

Definition 2.2.6. *$\mathbb{D}_{1,2}$ is defined to be the closure of the family $\mathcal{S}$ with respect to the norm $\|\cdot\|_{1,2}$.*

Then $\mathbb{D}_{1,2}$ consists of all $F \in L^2(\Omega)$ for which there exists $F_n \in \mathcal{S}$ such that

$$F_n \to F \text{ in } L^2(\Omega) \text{ as } n \to \infty$$

and

$$\{\mathbf{D}_t F_n\}_{n=1}^{\infty} \text{ is convergent in } L^2([0, T] \times \Omega).$$

In this case it is tempting to define

$$D_t F := \lim_{n\to\infty} \mathbf{D}_t F_n$$

However, for this to work, we need to ensure that this representation defines $D_t F$ uniquely. In other words, if there exists another sequence $G_n \in \mathcal{S}$ such that

$$G_n \to F \text{ in } L^2(\Omega) \text{ as } n \to \infty$$

and

$$\{\mathbf{D}_t G_n\}_{n=1}^{\infty} \text{ is convergent in } L^2([0, T] \times \Omega),$$

does it follow that

$$\lim_{n\to\infty} \mathbf{D}_t F_n = \lim_{n\to\infty} \mathbf{D}_t G_n \ ?$$

This can be answered by considering the difference $H_n = F_n - G_n$ and proving the closability of the operator $\mathbf{D}_t$.

Theorem 2.2.2. *(Closability of the operator $\boldsymbol{D}_t$)*

Suppose $\{H_n\}_{n=1}^{\infty} \subset \mathcal{S}$ has the properties:

$$H_n \to 0 \ \textit{in} \ L^2(\Omega) \ \textit{as} \ n \to \infty$$

and

$$\{\mathbf{D}_t H_n\}_{n=1}^{\infty} \ \textit{is convergent in} \ L^2([0,T] \times \Omega).$$

Then

$$\lim_{n\to\infty} \mathbf{D}_t H_n = 0$$

The proof uses the integration-by-parts result.

Lemma 2.2.3. *(Integration-by-parts)*

Suppose $F \in \mathcal{D}_{1,2}, \varphi \in \mathcal{D}_{1,2}$ and $\gamma(t) = \int_0^t g(s)ds$ with $g \in L^2([0,T])$. Then

$$\mathbb{E}[\mathbf{D}_\gamma F \cdot \varphi] = \mathbb{E}\left[F \cdot \varphi \cdot \int_0^T g dW\right] - \mathbb{E}[F \cdot \mathbf{D}_\gamma \varphi]$$

Proof: See [11] Section 4.8 and later in the chapter.

By the Lemma

$$\mathbb{E}[\mathbf{D}_\gamma H_n \cdot \varphi] = \mathbb{E}\left[H_n \varphi \cdot \int_0^T g dW\right] - \mathbb{E}[H_n \cdot \mathbf{D}_\gamma \varphi]$$

$$\to 0 \ \textit{as} \ n \to \infty \ \textit{for all} \ \varphi \in \mathcal{S}$$

Since $\{\mathbf{D}_\gamma H_n\}_{n=1}^{\infty}$ converges in $L^2(\Omega)$ and $\mathcal{S}$ is dense in $L^2(\Omega)$, we then conclude that $\mathbf{D}_\gamma H_n \to 0$ in $L^2(\Omega)$ as $n \to \infty$.

Since this holds for all $\gamma(t) = \int_0^t g ds$, we find that

$$\mathbf{D}_t H_n \to 0 \text{ in } L^2([0,T] \times \Omega)$$

□

Definition 2.2.7 (Malliavin Derivative). *Let $F \in \mathbb{D}_{1,2}$ so that there exists a $\{F_n\} \subset \mathcal{S}$ such that*

$$F_n \to F \text{ in } L^2(\Omega) \text{ as } n \to \infty$$

and

$$\{\mathbf{D}_t F_n\}_{n=1}^{\infty} \text{ is convergent in } L^2([0,T] \times \Omega)$$

Then we define

$$D_t F = \lim_{n\to\infty} \mathbf{D}_t F_n$$

and

$$D_\gamma F = \int_0^T D_t F \cdot g(t) dt$$

$$\text{for all } \gamma(t) = \int_0^t g(s) ds \text{ with } g \in L^2([0,T])$$

We will call $D_t F$ the ***Malliavin Derivative*** *of F.*

Remark 2.2.4. *At this stage we have two apparently different definitions for the derivative of F:*

1. *The derivative $\mathbf{D}_t F$ of $F \in \mathcal{D}_{1,2}$*

2. *The Malliavin derivative $D_t F$ of $F \in \mathbb{D}_{1,2}$*

The next result will show that the two derivatives coincide for $F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$

Lemma 2.2.4. *Let $F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$ and suppose that $\{F_n\} \subset \mathcal{S}$ has the properties*

$F_n \to F$ in $L^2(\Omega)$ as $n \to \infty$ and $\{\mathbf{D}_t F_n\}_{n=1}^{\infty}$ is convergent in $L^2([0,T] \times \Omega)$

Then

$$\mathbf{D}_t F = \lim_{n\to\infty} \mathbf{D}_t F_n$$

Hence

$$D_t F = \mathbf{D}_t F \textit{ for } F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$$

Proof. We are effectively given that $\mathbf{D}_\gamma F_n$ converges in $L^2(\Omega)$ for each $\gamma(t) = \int_0^t g(s)ds$ where $g \in L^2([0,T])$. Using the integration-by-parts lemma we obtain

$$\mathbb{E}[(\mathbf{D}_\gamma F_n - \mathbf{D}_\gamma F) \cdot \varphi] = \mathbb{E}\left[(F_n - F) \cdot \varphi \cdot \int_0^T gdW\right] - \mathbb{E}[(F_n - F) \cdot \mathbf{D}_\gamma \varphi]$$
$$\to 0 \text{ for all } \varphi \in \mathcal{S}$$

Hence $\mathbf{D}_\gamma F_n \to \mathbf{D}_\gamma F$ in $L^2(\Omega)$ and we then have

$$\mathbf{D}_t F = \lim_{n\to\infty} \mathbf{D}_t F_n$$

□

By definition, $\mathcal{S} \subseteq \mathbb{D}_{1,2} \subseteq L^2(\Omega)$. Since the closure of $\mathcal{S}$ with respect to the norm $\|\cdot\|_{1,2}$ is equal to $L^2(\Omega)$, it is tempting to conclude that $\mathbb{D}_{1,2} = L^2(\Omega)$ as well. However, it can be shown that $\mathbb{D}_{1,2} \subsetneq L^2(\Omega)$ such that the two norms are not equivalent.

For elements in $\mathbb{D}_{1,2}$ we can now define a derivative at the limit of $D_t F_n$. This is the so called Malliavin derivative. Since the derivatives coincide for $F \in \mathcal{D}_{1,2} \cap \mathbb{D}_{1,2}$ we will use the notation $D_\gamma F$ for the directional derivative and $D_t F$ for the derivative of such random variables.

2.3 Differentiation Rules

Proposition 2.3.1 (The Chain rule). *Let $F \in \mathcal{D}_{1,2}$ and let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function. Then $f(F(\omega)) \in \mathcal{D}_{1,2}$ and*

$$D_t f(F(\omega)) = f'(F(\omega))D_t F(\omega)$$

Proof. Using the key definition of the directional derivative we get

$$\begin{aligned} D_\gamma f(F(\omega)) &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[f(F(\omega + \varepsilon\gamma)) - f(F(\omega))\right] \\ &= \lim_{\varepsilon \to 0}\left[\frac{f(F(\omega + \varepsilon\gamma)) - f(F(\omega))}{F(\omega + \varepsilon\gamma) - F(\omega)} \cdot \frac{F(\omega + \varepsilon\gamma) - F(\omega)}{\varepsilon}\right] \\ &= f'(F(\omega))D_\gamma F(\omega) \end{aligned}$$

Since $F \in \mathcal{D}_{1,2}$ we know that $D_t F(\omega)$ will exist in $L^2([0,T] \times \Omega)$, so we may write

$$\begin{aligned} D_\gamma f(F(\omega)) &= f'(F(\omega))D_\gamma F(\omega) \\ &= f'(F(\omega)) \int_0^T D_t F(\omega) h(t) dt \\ &= \int_0^T f'(F(\omega)) D_t F(\omega) h(t) dt \end{aligned}$$

Hence $f(F) \in \mathcal{D}_{1,2}$ and $D_t f(F(\omega)) = f'(F(\omega))D_t F(\omega)$

□

Proposition 2.3.2 (The Product rule). *If $F, G \in \mathcal{D}_{1,2}$ then $FG \in \mathcal{D}_{1,2}$ and*

$$D_t(F(\omega)G(\omega)) = (D_t F(\omega))G(\omega) + F(\omega)(D_t G(\omega))$$

Proof. Using the key definition of the directional derivative again, we get

$$\begin{aligned}
&D_\gamma(F(\omega)G(\omega))\\
&=\lim_{\varepsilon\to 0}\frac{F(\omega+\varepsilon\gamma)G(\omega+\varepsilon\gamma)-F(\omega)G(\omega))}{\varepsilon}\\
&=\lim_{\varepsilon\to 0}\frac{F(\omega+\varepsilon\gamma)G(\omega+\varepsilon\gamma)-F(\omega)G(\omega+\varepsilon\gamma)+F(\omega)G(\omega+\varepsilon\gamma)-F(\omega)G(\omega))}{\varepsilon}\\
&=\lim_{\varepsilon\to 0}\frac{[F(\omega+\varepsilon\gamma)-F(\omega)]\,G(\omega+\varepsilon\gamma)+F(\omega)\,[G(\omega+\varepsilon\gamma)-G(\omega)]}{\varepsilon}\\
&=\lim_{\varepsilon\to 0}\left(\frac{F(\omega+\varepsilon\gamma)-F(\omega)}{\varepsilon}G(\omega+\varepsilon\gamma)+F(\omega)\frac{G(\omega+\varepsilon\gamma)-G(\omega)}{\varepsilon}\right)\\
&=(D_\gamma F(\omega))G(\omega)+F(\omega)(D_\gamma G(\omega))
\end{aligned}$$

Since we were given that $F, G \in \mathcal{D}_{1,2}$ we get

$$\begin{aligned}
D_\gamma(F(\omega)G(\omega)) &= (D_\gamma F(\omega))G(\omega)+F(\omega)(D_\gamma G(\omega))\\
&=\int_0^T (D_tF(\omega))h(t)dt\cdot G(\omega)+F(\omega)\int_0^T (D_tG(\omega))h(t)dt\\
&=\int_0^T \left[(D_tF(\omega))G(\omega)+F(\omega)(D_tG(\omega))\right]h(t)dt
\end{aligned}$$

Hence,

$$D_t(F(\omega)G(\omega)) = (D_tF(\omega))G(\omega)+F(\omega)(D_tG(\omega))$$

□

Proposition 2.3.3. *Let $F \in \mathbb{D}_{1,2}$ be $\mathcal{F}_s$-adapted. Then D_tF will be $\mathcal{F}_s$-adapted such that for $t > s$ we will have:*

$$D_tF = 0$$

Proof. The result will be proved for a special case. Consider a random variable of the form

$$F(\omega) = \exp\left(\int_0^T h(u)dW_u - \frac{1}{2}\int_0^T h^2(u)du\right) \tag{2.3.1}$$

where $h \in L^2([0,T])$ is deterministic. Notice that the Novikov condition is satisfied, implying that F is an exponential martingale.

We also have from the chain rule and the below example, that

$$D_t F = F h(t)$$

Now,

$$\begin{aligned}
D_t \mathbb{E}[F|\mathcal{F}_s] &= D_t \exp\left(\int_0^s h(u)dW_u - \frac{1}{2}\int_0^s h^2(u)du\right) \\
&= D_t \exp\left(\int_0^T h(u)\mathbf{1}_{[0,s]}(u)dW_u - \frac{1}{2}\int_0^s h^2(u)du\right) \\
&= \exp\left(\int_0^s h(u)dW_u - \frac{1}{2}\int_0^s h^2(u)du\right) h(t)\mathbf{1}_{[0,s]}(t) \\
&= \mathbb{E}[F|\mathcal{F}_s]\, h(t)\mathbf{1}_{[0,s]}(t) \\
&= \mathbb{E}[Fh(t)|\mathcal{F}_s]\, \mathbf{1}_{[0,s]}(t) \\
&= \mathbb{E}[D_t F|\mathcal{F}_s]\, \mathbf{1}_{[0,s]}(t)
\end{aligned}$$

In the above computation we have used the fact that F is a martingale, the chain rule and example (2.3.6). The above result can be extended to the linear span of random variables of the form (2.3.1). Since this linear span is dense in $L^2(\Omega)$ it would seem reasonable that the result would hold true for more general random variables. Clearly the result would not hold for all $F \in L^2(\Omega)$ since it involves the Malliavin derivative of F, which does not exist for all $F \in L^2(\Omega)$. In Nualart [10] and Øksendal [11] the result is proved for the more general $F \in \mathbb{D}_{1,2}$.

In particular if $F \in \mathbb{D}_{1,2}$ is $\mathcal{F}_s$-adapted we get

$$\begin{aligned} D_t F &= D_t \mathbb{E}[F|\mathcal{F}_s] \\ &= \mathbb{E}[D_t F|\mathcal{F}_s]\mathbf{1}_{[0,s]}(t) \end{aligned}$$

So $D_t F$ is $\mathcal{F}_s$-adapted and $D_t F = 0$ if $t > s$.

□

The final result in this differentiation section is a representation of the integrand from Itô's representation theorem. See Øksendal [11] Theorem. 4.3.3. This theorem is the cornerstone of the martingale approach to optimal portfolio choice, where the integrand represents the optimal investment strategy. Since the representation theorem only gives the existence of such an investment strategy, the following proposition is an important financial application of the Malliavin calculus. The result has been used to solve optimal portfolio problems in complete markets with Monte-Carlo simulation.

Proposition 2.3.4 (The Clark-Ocone formula). *Let $F \in \mathbb{D}_{1,2}$ be $\mathcal{F}_T$-adapted. Then*

$$F(\omega) = \mathbb{E}[F] + \int_0^T \mathbb{E}[D_t F|\mathcal{F}_T](\omega)d\omega(t) \tag{2.3.2}$$

Proof. The proof, again, will be for the special exponential martingale case as represented by (2.3.1). This proof is very similar to the proof of Itô's integral representation. Let us define the stochastic process F_t by:

$$F_t(\omega) := \exp\left(\int_0^t h(u)dW_u - \frac{1}{2}\int_0^t h^2(u)du\right) \tag{2.3.3}$$

and let $F := F_T$. We now introduce the auxiliary process

$$Z_t := \int_0^t h(u)dW_u - \frac{1}{2}\int_0^t h^2(u)du$$

equivalently

$$dZ_t := h(t)dW_t - \frac{1}{2}h^2(t)dt$$

This allows us to represent the dynamics of F_t, given by Itô's lemma when letting $f(x) = e^x$, as

$$\begin{aligned} dF_t &= F_t dZ_t + \frac{1}{2} F_t (dZ_t)^2 \\ &= F_t \left(h(t) dW_t - \frac{1}{2} h^2(t) dt \right) + \frac{1}{2} F_t h^2(t) dt \\ &= F_t h(t) dW_t \end{aligned}$$

We also know

$$\begin{aligned} \mathbb{E}\left[D_t F | \mathcal{F}_t\right] &= \mathbb{E}\left[F h(t) | \mathcal{F}_t\right] \\ &= \mathbb{E}\left[F | \mathcal{F}_t\right] h(t) \\ &= F_t h(t) \end{aligned}$$

The dynamics of F in integral form is equivalent to

$$\begin{aligned} F &= F_0 + \int_0^T F_t h(t) dW_t \\ &= 1 + \int_0^T F_t h(t) dW_t \end{aligned}$$

It is now clear that $\mathbb{E}[F] = 1$ and we have shown $\mathbb{E}\left[D_t F | \mathcal{F}_t\right] = F_t h(t)$, hence

$$F = \mathbb{E}[F] + \int_0^T \mathbb{E}\left[D_t F | \mathcal{F}_t\right] dW_t$$

The result has been shown for a special case. By linearity, the representation holds for linear combinations of exponentials of the form (2.3.3). In the general case, any random variable F in $L^2(\Omega, F_T, \mathbb{P})$ can be approximated in mean square by a sequence F_n of linear combinations of exponentials of the form (2.3.3) and hence $F \in \mathbb{D}_{1,2}$. See Nualart [10] for a full proof.

□

2.3.1 Examples

Example 2.3.5. *Let* $F(\omega) = W_t(\omega) = \omega(t)$. *Then*

$$F(\omega + \varepsilon\gamma) = \omega(t) + \varepsilon\gamma(t)$$

so that

$$\begin{aligned} D_\gamma F(\omega) &= \lim_{\varepsilon \to 0} \frac{F(\omega + \varepsilon\gamma) - F(\omega)}{\varepsilon} \\ &= \gamma(t) \\ &= \int_0^t h(s)ds \\ &= \int_0^T \mathbf{1}_{[0,t]}(s)h(s)ds \end{aligned}$$

So from the Malliavin derivative definition

$$D_s W_t = \mathbf{1}_{[0,t]}(s)$$

□

The Malliavin derivative should be considered as a perturbation of the underlying Brownian motion. For Brownian path changes at time $s \leq t$, the entire future path will also change. For times $s > t$ no change has occurred at time t, hence the use of the indicator function.

Example 2.3.6. *Let* $F(\omega) = \int_0^T f(s)dW_s(\omega) = \int_0^T f(s)d\omega(s)$ *where* $f \in L^2([0,T])$ *is a deterministic function. We then have*

$$F(\omega + \varepsilon\gamma) = \int_0^T f(s)d(\omega(s) + \varepsilon\gamma(s))$$

Now we can consider the directional derivative

$$\begin{aligned} D_\gamma F(\omega) &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[F(\omega + \varepsilon\gamma) - F(\omega) \right] \\ &= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\int_0^T f(s) d(\omega(s) + \varepsilon\gamma(s)) - \int_0^T f(s) d\omega(s) \right] \\ &= \int_0^T f(s) d\gamma(s) \\ &= \int_0^T f(s) h(s) ds \end{aligned}$$

since we know that the directions are of the form $\gamma(t) = \int_0^t h(s)ds$

Now using the definition link between directional and Malliavin derivatives

$$D_s F(\omega) = f(s)$$

Notice that if we chose $f(s) = \mathbf{1}_{[0,t]}(s)$ *we would be dealing with* $F(\omega) = W_t(\omega) = \omega(t)$ *leading to the result*

$$D_s W_t = \mathbf{1}_{[0,t]}(s)$$

□

Example 2.3.7. *Let* $F(\omega) = f(W_t(\omega)) = f(\omega(t))$ *where* f *is differentiable. By the chain rule and example 2.3.5 we then have*

$$D_s F(\omega) = f'(W_t(\omega)) \mathbf{1}_{[0,s]}(t)$$

□

Example 2.3.8. *Let* $F(\omega) = \int_0^T f(W_t(\omega)) dW_t = \int_0^T f(\omega(t)) d\omega(t)$ *then we have*

$$\begin{aligned} F(\omega + \varepsilon\gamma) &= \int_0^T f(\omega(t) + \varepsilon\gamma(t)) d(\omega(t) + \varepsilon\gamma(t)) \\ &= \int_0^T f(\omega(t) + \varepsilon\gamma(t)) d\omega(t) + \varepsilon \int_0^T f(\omega(t) + \varepsilon\gamma(t)) d\gamma(t) \end{aligned}$$

Hence the directional derivative is as follows:

$$
\begin{aligned}
& D_\gamma F(\omega) \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[F(\omega + \varepsilon\gamma) - F(\omega) \right] \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\int_0^T f(\omega(t) + \varepsilon\gamma(t)) d\omega(t) + \varepsilon \int_0^T f(\omega(t) + \varepsilon\gamma(t)) d\gamma(t) - \int_0^T f(\omega(t)) d\omega(t) \right] \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\int_0^T \left[f(\omega(t) + \varepsilon\gamma(t)) - f(\omega(t)) \right] d\omega(t) + \varepsilon \int_0^T f(\omega(t) + \varepsilon\gamma(t)) d\gamma(t) \right] \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left[\int_0^T \left[f(\omega(t) + \varepsilon\gamma(t)) - f(\omega(t)) \right] d\omega(t) \right] + \int_0^T f(\omega(t)) d\gamma(t) \\
&= \int_0^T f'(\omega(t))\gamma(t) d\omega(t) + \int_0^T f(\omega(t)) d\gamma(t) \\
&= \int_0^T f'(\omega(t)) \left(\int_0^t h(s) ds \right) d\omega(t) + \int_0^T f(\omega(t)) h(t) dt \\
&= \int_0^T \left(\int_s^T f'(\omega(t)) d\omega(t) \right) h(s) ds + \int_0^T f(\omega(s)) h(s) ds \\
&= \int_0^T \left(\int_s^T f'(\omega(t)) d\omega(t) + f(\omega(s)) \right) h(s) ds \\
&= \int_0^T \left(\int_s^T f'(W_t(\omega)) dW_t(\omega) + f(W_s(\omega)) \right) h(s) ds
\end{aligned}
$$

Now using the standard definition of the Malliavin derivative we have that

$$
D_s \int_0^T f(W_t(\omega)) dW_t = \int_s^T f'(W_t(\omega)) dW_t + f(W_s(\omega))
$$

□

2.4 Integration on the Wiener Space

2.4.1 Integration-by-parts on the Wiener Space

Proposition 2.4.1 (Integration-by-parts). *Let $F, G \in \mathcal{D}_{1,2}$ and define $\gamma(t) = \int_0^t h(s)ds$ for $h \in L^2([0,T])$ then*

$$
\mathbb{E}\left[\int_0^T (D_t F) h_t dt \right] = \mathbb{E}\left[F \int_0^T h_t dW_t \right] \tag{2.4.1}
$$

Proof. From the definition of the directional derivative we get

$$\begin{aligned}
\mathbb{E}\left[\int_0^T (D_tF)h_t dt\right] &= \int_\Omega D_\gamma F(\omega)d\mathbb{P}(\omega) \\
&= \int_\Omega \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[F(\omega+\varepsilon\gamma)-F(\omega)\right]d\mathbb{P}(\omega) \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_\Omega \left[F(\omega+\varepsilon\gamma)-F(\omega)\right]d\mathbb{P}(\omega) \\
&= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left[\underbrace{\int_\Omega F(\omega+\varepsilon\gamma)d\mathbb{P}(\omega)}_{\kappa} - \int_\Omega F(\omega)d\mathbb{P}(\omega)\right]
\end{aligned}$$

Since $h \in L^2([0,T])$, εh will satisfy Novikov's condition which ensures that

$$M_t := \exp\left(-\varepsilon\int_0^t h_s dW_s - \frac{1}{2}\varepsilon^2\int_0^t h_s^2 ds\right)$$

is a $\mathbb{P}$-Martingale. By the Girsanov theorem

$$\widetilde{W}_t := W_t + \varepsilon\int_0^t h_s ds$$

is a Brownian motion under the measure $\widetilde{\mathbb{P}}$ defined by

$$\begin{aligned}
\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} &= M_T \\
&= \exp\left(-\varepsilon\int_0^T h_s dW_s - \frac{1}{2}\varepsilon^2\int_0^T h_s^2 ds\right) \\
&= \exp\left(-\varepsilon\int_0^T h_s d\widetilde{W}_s + \frac{1}{2}\varepsilon^2\int_0^T h_s^2 ds\right)
\end{aligned}$$

Now reconsider κ in the above calculations

$$\begin{aligned}
\int_\Omega F(\omega+\varepsilon\gamma)d\mathbb{P}(\omega) &= \int_\Omega F(\widetilde{\omega})\exp\left(\varepsilon\int_0^T h_s d\widetilde{W}_s(\omega) - \frac{1}{2}\varepsilon^2\int_0^T h_s^2 ds\right)d\widetilde{\mathbb{P}}(\omega) \\
&= \int_\Omega F(\omega)\exp\left(\varepsilon\int_0^T h_s dW_s(\omega) - \frac{1}{2}\varepsilon^2\int_0^T h_s^2 ds\right)d\mathbb{P}(\omega)
\end{aligned}$$

Hence,

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T (D_t F) h_t dt\right] \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[\int_\Omega F(\omega + \varepsilon\gamma) d\mathbb{P}(\omega) - \int_\Omega F(\omega) d\mathbb{P}(\omega)\right] \\
&= \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \int_\Omega F(\omega)\left[\exp\left(\varepsilon \int_0^T h_s dW_s(\omega) - \frac{1}{2}\varepsilon^2 \int_0^T h_s^2 ds\right) - 1\right] d\mathbb{P}(\omega) \\
&= \int_\Omega F(\omega) \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[\exp\left(\varepsilon \int_0^T h_s dW_s(\omega) - \frac{1}{2}\varepsilon^2 \int_0^T h_s^2 ds\right) - 1\right] d\mathbb{P}(\omega) \\
&= \int_\Omega F(\omega) \frac{d}{d\varepsilon}\left[\exp\left(\varepsilon \int_0^T h_s dW_s(\omega) - \frac{1}{2}\varepsilon^2 \int_0^T h_s^2 ds\right) - 1\right]_{\varepsilon=0} d\mathbb{P}(\omega) \\
&= \int_\Omega \left[F(\omega) \int_0^T h_s dW_s(\omega)\right] d\mathbb{P}(\omega) \\
&= \mathbb{E}\left[F \int_0^T h_t dW_t\right]
\end{aligned}$$

□

Corollary 2.4.2. *Let* $F, G \in \mathcal{D}_{1,2}$ *and define* $\gamma(t) = \int_0^t h_s ds$ *for* $h \in L^2([0,T])$*. Then*

$$\mathbb{E}\left[G \int_0^T (D_t F) h_t dt\right] = \mathbb{E}\left[FG \int_0^T h_t dW_t\right] - \mathbb{E}\left[F \int_0^T (D_t G) h_t dt\right]$$

Proof. From the product rule we have $FG \in \mathcal{D}_{1,2}$ and

$$D_t(FG) = F D_t G + G D_t F$$

Using the integration-by-parts proposition with "FG" replacing F we get

$$\begin{aligned}
\mathbb{E}\left[FG \int_0^T h_t dW_t\right] &= \mathbb{E}\left[\int_0^T (D_t(FG)) h_t dt\right] \\
&= \mathbb{E}\left[F \int_0^T (D_t G) h_t dt\right] + \mathbb{E}\left[G \int_0^T (D_t F) h_t dt\right]
\end{aligned}$$

□

An alternative proof of the integration-by-parts result from Friz [7] is in the form of a lemma:

Lemma 2.4.3 (Integration-by-parts). *Let* $h \in H = L^2([0,T])$*,* $F \in \mathcal{S}$ *and we also have* $W(h) = \int_0^T hdW$

Then

$$\mathbb{E}[\langle DF, h\rangle_H] = \mathbb{E}\left[F\int_0^T hdW\right] \tag{2.4.2}$$

Proof. By homogeneity and without loss of generality, we can set $\|h\| = 1$. Also, we can find an f such that $F = f(W(h_1), ..., W(h_n))$, with (h_i) orthonormal in H and having $h = h_1$. Then using the standard integration-by-parts:

$$\begin{aligned}\mathbb{E}[\langle DF, h\rangle_H] &= \mathbb{E}\left[\sum_i \partial_i f(W(h_1), ..., W(h_n))h_i \cdot h\right] \\ &= \mathbb{E}\left[\sum_i \partial_i f(W(h_1), ..., W(h_n))\langle h_i, h\rangle\right]\end{aligned}$$

However, all h_i's are orthonormal by assumption, such that, $\langle h_i, h_j\rangle = 0$, $\forall i \neq j$. Hence we are left with:

$$\begin{aligned}\mathbb{E}[\langle DF, h\rangle_H] &= \mathbb{E}[\partial_1 f \cdot \langle h_1, h\rangle] \\ &= \mathbb{E}[\partial_1 f \cdot \|h\|] \text{ since } h_1 = h \\ &= \int_{\mathbb{R}^n} \partial_1 f(x)(2\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{2}} dx\end{aligned}$$

Consider $(\mathbb{R}, \lambda)$. Let f be smooth with compact support, then by the translation invariance of the Lebesgue measure:

$$\int f(x+h)d\lambda = \int f(x)d\lambda$$

Dividing by h and letting $h \to 0$ we have:

$$\int f' d\lambda = 0$$

Let $f = fg$ then:

$$\int f'gd\lambda = -\int fg'd\lambda$$

This implies that integration-by-parts is the infinitesimal expression of a measure invariance.

Continuing with the Lemma proof:

$$\begin{aligned}\mathbb{E}[\langle DF, h\rangle_H] &= \int_{\mathbb{R}^n} \partial_1 f(x)(2\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{2}} dx \\ &= -\int_{\mathbb{R}^n} f(x)(2\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{2}}(-x_1)dx\end{aligned}$$

Remember on the space $(\mathbb{R}^n, \nu^n)$ we have a n-dimensional Gaussian measure,

$$d\nu^n(x) = (2\pi)^{-\frac{n}{2}} e^{-\frac{|x|^2}{2}} dx$$

So,

$$\mathbb{E}[\langle DF, h\rangle_H] = \int_{\mathbb{R}^n} f(x)x_1 d\nu^n = \mathbb{E}[F \cdot W(h_1)] = \mathbb{E}[F \cdot W(h)]$$

Now, letting $F, G \in \mathcal{S}$ and applying equation (2.4.2) to smooth functionals FG, we then have:

$$\mathbb{E}[\langle D(FG), h\rangle_H] = \mathbb{E}[F\langle DG, h\rangle_H] + \mathbb{E}[G\langle DF), h\rangle_H] = \mathbb{E}[FGW(h)] \tag{2.4.3}$$

□

2.4.2 The Skorohod Integral and its properties

The Skorohod integral of a stochastic process can be constructed from the Wiener Itô chaos expansion. It is apparent that the Skorohod integral coincides with the adjoint operator of the Malliavin differential operator. For a Skorohod integrable process h_t and a Malliavin differentiable random variable F, the Skorohod integral, $S(h)$, is defined as

$$\langle F, S(h)\rangle_{L^2(\Omega)} = \langle D_t F, h\rangle_{L^2([0,T]\times\Omega)}$$

where $\langle\cdot,\cdot\rangle$ defines the inner product. The following definition omits the technical conditions for a process to be Skorohod integrable.

Definition 2.4.1 (Nualart Definition 1.3.1 [10]). *If h is Skorohod integrable, we define the Skorohod integral of h, as the element $S(h) := \int_0^T h_t \delta W_t \in L^2(\Omega)$ that satisfies*

$$\mathbb{E}[FS(h)] = \mathbb{E}\left[\int_0^T (D_t F) h_t dt\right]$$

for all $F \in \mathbb{D}_{1,2}$

Proposition 2.4.4. *If h_t is $\mathcal{F}_t$-adapted, the Skorohod integral coincides with the Itô integral when it is defined. In other words*

$$\int_0^T h_t \delta W_t = \int_0^T h_t dW_t$$

Proof. The result will only be proved for deterministic functions $h_t \in L^2([0,T])$. We let $F, G \in \mathbb{D}_{1,2}$ be two Malliavin differentiable random variables. Now, using the definition of the Skorohod integral and the proved integration-by-parts property in Corollary 2.4.2, we have

$$\begin{aligned}
\mathbb{E}[GS(Fh)] &= \mathbb{E}\left[\int_0^T (D_t G) F h_t dt\right] \\
&= \mathbb{E}\left[F \int_0^T (D_t G) h_t dt\right] \\
&= \mathbb{E}\left[GF \int_0^T h_t dW_t\right] - \mathbb{E}\left[G \int_0^T (D_t F) h_t dt\right] \\
&= \mathbb{E}\left[G\left(F \int_0^T h_t dW_t - \int_0^T (D_t F) h_t dt\right)\right]
\end{aligned}$$

Since this must hold for all $G \in \mathbb{D}_{1,2}$, an inner product argument will give

$$S(Fh) = F \int_0^T h_t dW_t - \int_0^T (D_t F) h_t dt$$

If we let $F = 1$ we can see that the Itô and Skorohod integrals coincide for deterministic $L^2([0,T])$ functions h_t.

Clearly, $S(1 \cdot h) = 1 \cdot \int_0^T h_t dW_t - \int (D_t 1) h_t dt$ so that $S(h) = \int_0^T h_t dW_t$.

If we take F to be an $\mathcal{F}_s$-measurable random variable and $h_t = \mathbf{1}_{(s,u](t)}$ we then obtain

$$\begin{aligned} S(Fh) &= F \int_0^T h_t dW_t - \int_0^T (D_t F) h_t dt \\ &= F \int_s^u dW_t - \int_s^u (D_t F) dt \\ &= F(W_u - W_s) - 0 \end{aligned}$$

The reason for $D_t F = 0$ for $t > s$ is due to F being $\mathcal{F}_s$-measurable.

We can further write

$$Fh_t = 0 \cdot \mathbf{1}_{[0,s]}(t) + F \cdot \mathbf{1}_{(s,u]}(t) + 0 \cdot \mathbf{1}_{(u,T]}(t)$$

We can see that F is an elementary process, hence

$$\begin{aligned} \int_0^T F h_t dW_t &= 0 \cdot (W_s - W_0) + F \cdot (W_u - W_s) + 0 \cdot (W_T - W_u) \\ &= F \cdot (W_u - W_s) \end{aligned}$$

So, in this case, the Itô and Skorohod integral coincide as well

$$S(Fh) = \int_0^T F h_t dW_t$$

To show the result for any $\mathcal{F}_t$-adapted process $h_t \in L^2([0,T] \times \Omega)$, one can use an approximation argument.

□

Proposition 2.4.5 (Alternative approach to already proved result). *Let F be a Malliavin differentiable random variable, then*

$$\int_0^T F h_t \delta W_t = F \int_0^T h_t \delta W_t - \int_0^T (D_t F) h_t dt$$

where h_t is Skorohod integrable. Furthermore, if h_t is $\mathcal{F}_t$ adapted we have

$$\int_0^T F h_t \delta W_t = F \int_0^T h_t dW_t - \int_0^T (D_t F) h_t dt$$

Proof. Let G be a Malliavin differentiable random variable. Now, using the product rule and the integration-by-parts property, we get:

$$\begin{aligned}\mathbb{E}\left[\int_0^T (D_t G) F h_t dt\right] &= \mathbb{E}\left[\int_0^T \underbrace{[(D_t(GF)) - G(D_t F)]}_{\text{Product Rule}} h_t dt\right] \\ &= \mathbb{E}\left[\int_0^T D_t(GF) h_t dt - G \int_0^T (D_t F) h_t dt\right]\end{aligned}$$

Now, using the definition of the Skorohod integral

$$\mathbb{E}[F S(h)] = \mathbb{E}\left[\int_0^T (D_t F) h_t dt\right] = \mathbb{E}\left[F \int_0^T h_t \delta W_t\right]$$

We have

$$\begin{aligned}\mathbb{E}\left[\int_0^T (D_t G) F h_t dt\right] &= \mathbb{E}\left[GF \int_0^T h_t \delta W_t - G \int_0^T (D_t F) h_t dt\right] \\ &= \mathbb{E}\left[G\left[F \int_0^T h_t \delta W_t - \int_0^T (D_t F) h_t dt\right]\right]\end{aligned}$$

Using the definition of the Skorohod integral again, we can rewrite the left hand side of the above equation as:

$$\mathbb{E}\left[G \int_0^T F h_t \delta W_t\right] = \mathbb{E}\left[G\left[F \int_0^T h_t \delta W_t - \int_0^T (D_t F) h_t dt\right]\right]$$

However, we were given that h_t is Skorohod integrable and that h_t was $\mathcal{F}_t$-adapted such that $\int_0^T h_t \delta W_t = \int_0^T h_t dW_t$ so that

$$\mathbb{E}\left[G\int_0^T Fh_t dW_t\right] = \mathbb{E}\left[G\left[F\int_0^T h_t\delta W_t - \int_0^T (D_tF)h_t dt\right]\right]$$

and because this should be true for all G, again the result follows from an inner product argument.

□

2.4.3 Integration examples

Example 2.4.6. *If we let $h_t = W_t$ we have*

$$\begin{aligned}\int_0^T W_t\delta W_t &= \int_0^T W_t dW_t \\ &= \frac{1}{2}W_T^2 - \frac{1}{2}T\end{aligned}$$

since W_t is $\mathcal{F}_t$-adapted and in this case the Skorohod integral coincides the Itô integral.

Example 2.4.7. *It we let $h_t = 1$ and $F = W_T$. Using the result of the integration-by-parts proposition, where F is a Malliavin differentiable random variable:*

$$\int_0^T Fh_t\delta W_t = F\int_0^T h_t\delta W_t - \int_0^T (D_tF)h_t dt,$$

then

$$\begin{aligned}\int_0^T W_T\delta W_t &= W_T\int_0^T \delta W_t - \int_0^T (D_tW_T)dt \\ &= W_T(W_T - W_0) - \int_0^T \mathbf{1}_{[0,T]}(t)dt \\ &= W_T^2 - T\end{aligned}$$

Note that for specific applications to finance, to follow, the following adjunction shorthand notation will be used:

$$\int_0^T h_t \delta W_t := D^*(u_t)$$

for u_t $\mathcal{F}_t$ adapted or not.

2.5 Malliavin calculus application to Itô processes

This section focusses on finding the Malliavin derivatives of classical processes. These processes are diffusions, i.e. solutions of the following stochastic differential equation:

$$\begin{aligned} dX_t &= b(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 &= x \end{aligned}$$

where W_t is a standard Brownian motion, b and σ are, for existence and uniqueness, continuously differentiable, globally Lipschitz with bounded first derivatives and linear growth. In integral form, the stochastic differential equation can be written as:

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s$$

Now, with the above process we associate its first variation process, $(Y_t)_{t \geq 0} = \frac{\partial}{\partial x} X_t$:

$$\begin{aligned} dY_t &= b'(t, X_t)Y_t dt + \sigma'(t, X_t)Y_t dW_t \\ Y_0 &= 1 \end{aligned}$$

Theorem 2.5.1 (The Malliavin Derivative of an Itô Process). *For all $t \geq 0$ and $X_t \in \mathbb{D}_{1,2}$ the Malliavin derivative of X_t is given by:*

$$D_s X_t = Y_t Y_s^{-1} \sigma(s, X_s) \mathbf{1}_{\{s \leq t\}}$$

Proof. $X_t \in \mathbb{D}_{1,2}$ according to Nualart (Theorem 2.2.1) [10], by a limit of simple processes argument.

Recall the properties of commutation between Malliavin derivatives and integrals:

For any u_t adapted:

- $D_s\left(\int_0^T u_t dW_t\right) = u_s + \int_s^T D_s u_t dW_t$
- $D_s\left(\int_0^T u_t dt\right) = \int_s^T D_s u_t dt$

Then,

$$\begin{aligned}
D_s X_t &= D_s\left(\int_0^t \sigma(u, X_u) dW_u\right) + D_s\left(\int_0^t b(u, X_u) du\right) \\
&= \sigma(s, X_s) + \int_s^t D_s \sigma(u, X_u) dW_u + \int_s^t D_s b(u, X_u) du \\
&= \sigma(s, X_s) + \int_s^t \sigma'(u, X_u) D_s X_u dW_u + \int_s^t b'(u, X_u) D_s X_u du
\end{aligned}$$

That is,

$$D_s X_t = \sigma(s, X_s) exp\left\{\int_s^t \left(b' - \frac{\sigma'^2}{2}\right) du + \int_s^t \sigma' dW_u\right\}$$

So $(D_s X_t)_{t\geq 0}$ is a solution of the same stochastic differential operator as Y_t with a different initial condition., $D_s X_s = \sigma(s, X_s)$. There is also the added constraint that $D_s X_t = 0$ for $s > t$. This means that $D_s X_t$ and $Y_t \mathbf{1}_{\{s\leq t\}}$ are proportional with a constant of proportionality given by the initial conditions quotient $\frac{\sigma(s,X_s)}{Y_s}$.

Hence

$$D_s X_t = Y_t Y_s^{-1} \sigma(s, X_s) \mathbf{1}_{\{s\leq t\}}$$

□

Differentiation of the Black Scholes process

In the Black Scholes case, the diffusion process is characterised with $b(t, X_t) = bX_t$ and $\sigma(t, X_t) = \sigma X_t$. Also due to the solution of the future asset price:

$$Y_t = \frac{\partial}{\partial x} X_t = \frac{\partial}{\partial x} x e^{\mu t + \sigma W_t} = \frac{X_t}{x}$$

$\mu = b - \frac{\sigma^2}{2}$

We thus have:

$$D_s X_t = \sigma X_t \mathbf{1}_{\{s \leq t\}}$$

□

The above results will be heavily used in the Malliavin calculus application to finance chapter.

Chapter 3

Malliavin calculus applied to finance

To demonstrate an application of the Malliavin calculus in finance, this dissertation will now expose its power in helping to speed up calculations of risk management sensitivity measures (the Greeks) for options with non-closed analytical pricing functions, as is the case with most exotic options, where prices may depend on the entire price path.

3.1 Vanilla and Exotic Options

An option is a financial instrument that gives the holder the right to receive certain cash payoffs under certain conditions. For this privilege, the holder pays a premium to the writer of the option. Traditional options have been traded for hundreds of years and are fairly well understood. Whether European or American expiry, they are sometimes called "vanilla" options. In contrast, "exotic" options such as Asian, Barrier, Lookback and Binary are much more recent innovations. While a vanilla option pays off depending on the price of the underlying asset, the payoff of an exotic option typically depends on some function of the price of the underlying asset, or a relationship between several underlying assets.

3.2 Option sensitivities

3.2.1 Delta hedging

Derivative positions are "Delta hedged", by combining the option position with a position in the underlying asset to form a portfolio, where portfolio value does not change in reaction to changes in the price of the underlying over a short period of time. If the portfolio is continuously dynamically hedged, it will earn the risk free rate of interest.

Since changes in the underlying are the primary source of risk in a derivatives portfolio, the first order relationship between underlying value change and derivative value change as well as the associated convexity (second order relationship) adjustment are crucial to maintain the delta hedge. Further consideration of changes in the derivative value to changes in volatility, time to maturity and interest rate changes should also be factored in to produce an efficient total hedge. When an investor has a short position in derivatives, significant exposure to changes in volatility exists. Invariably, a Delta-Gamma-Vega hedge would require the volatility offset to be achieved with offsetting options. For the above reasons, the sensitivities of derivative values to changes in a parameter are of paramount risk importance.

3.2.2 The Greeks

Setting and option pricing

The economy is modeled on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with continuous trading over the time horizon $t \in [0, T]$. Information evolves according to the adapted filtration $\{\mathcal{F}_t, t \in [0, T]\}$ generated by the standard one dimensional Wiener process $(W_t)_{t\in[0,T]}$

The price of a contingent claim with expiry at time T, is traditionally calculated as the expected value of the discounted payoff value in a risk neutral probability measure-uniquely defined in complete markets with no arbitrage.

Hence, letting $X = X(\alpha)$ be a random variable depending on a parameter α, the price of the contingent claim at $t = 0$, is represented as:

$$\mathcal{P}(\alpha) = \mathbb{E}_{\mathbb{P}}\left[\widetilde{\Phi}(X(\alpha), \alpha)\Big|\mathcal{F}_0\right] \tag{3.2.1}$$

where $\widetilde{\Phi}$ (the discounted payoff function at expiry) is generally non-smooth. This expectation will be conditional to the information available today, described by the σ-algebra, $\mathcal{F}_0$ and is with respect to the risk neutral probability measure $\mathbb{P}$. Note that all future references to expectation $\mathbb{E}[\cdot]$ when calculating option prices and sensitivities are with respect to the risk neutral probability measure $\mathbb{P}$.

As an example, the payoff function for a European call option will be $\Phi(S_T) = (S_T - K)^+$ with K being the strike price. Hence, $X(\alpha) = S_T$. The price of the call is given as:

$$\mathcal{P} = \mathbb{E}\left[\Phi(S_T)e^{-\int_0^T r_s ds}\Big|\mathcal{F}_0\right] \tag{3.2.2}$$

where S_t is the underlying asset price process and r_s represents the risk free rate process. We also assume the underlying asset price process follows a geometric Brownian motion characterised by the diffusion equation:

$$\frac{dS_t}{S_t} = rdt + \sigma dW_t \tag{3.2.3}$$

equivalently

$$S_t = S_0 + r\int_0^t S_s ds + \sigma \int_0^t S_s dW_s \tag{3.2.4}$$

where r and σ are the constant risk free interest rate and volatility respectively. This is the most typical model used to describe asset prices. In finally solving the differential equation, we have the unique continuous strong solution of (3.2.3) with initial condition $S_0 = x$:

$$S_T = xe^{\{\mu T + \sigma W_T\}}$$

where $\mu = r - \frac{\sigma^2}{2}$ and $\{W_t\}_{t\in[0,T]}$ is the Wiener process.

Then by using the probability density function of S_T, we obtain the explicit integral

$$\mathcal{P}(x) = \int_{-\infty}^{\infty} e^{-rT}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy \tag{3.2.5}$$

Greek calculation methods and their types

A “Greek”, is a measure of the sensitivity of the option price with respect to a parameter, so:

$$\frac{\partial \mathcal{P}(\alpha)}{\partial \alpha} = \frac{\partial \mathbb{E}[\Phi(X(\alpha),\alpha)]}{\partial \alpha} = \mathbb{E}\left[\Phi'(X(\alpha),\alpha)\frac{\partial X(\alpha)}{\partial \alpha}\right]$$

The Greek letters are summarised as follows:

Greek	Sensitivity
Δ (Delta)	∂_S
Γ (Gamma)	∂_S^2
ρ (Rho)	∂_r
$\mathcal{V}$ (Vega)	∂_σ
θ (Theta)	∂_t

3.2.3 The numerical approach to the Greeks

Finite difference method

The main numerical approach to calculating an option price sensitivity is the finite difference approach, where prices are shifted by small amounts to approximate the sensitivity metric. As an example, we consider the sensitivity of the contingent claim with respect to the underlying asset price. The key requirement here is that the option prices, or at least a method to estimate them accurately, is necessary prior to estimation.

Using a centred difference scheme we have:

$$\begin{aligned}\Delta &\simeq \frac{\mathcal{P}(x+\frac{\varepsilon}{2}) - \mathcal{P}(x-\frac{\varepsilon}{2})}{\varepsilon} \\ &\simeq \mathbb{E}_{\mathbb{P}}\left[e^{-\int_0^T r_s ds} \frac{\Phi\left(S_T^{x+\frac{\varepsilon}{2}}\right) - \Phi\left(S_T^{x-\frac{\varepsilon}{2}}\right)}{\varepsilon} \middle| \mathcal{F}_0 \right]\end{aligned}$$

It is clear that

$$\lim_{\varepsilon \to 0} \left[\frac{\Phi\left(S_T^{x+\frac{\varepsilon}{2}}\right) - \Phi\left(S_T^{x-\frac{\varepsilon}{2}}\right)}{\varepsilon} \right] = \frac{\partial}{\partial x}\Phi(S_T)$$

Likelihood ratio method

The idea of the likelihood ratio method is to report the derivative of our payoff function on the density of the parameters of this function using an integration-by-parts formula. The process begins with switching the expectation and differentiation terms and is well explored in Broadie and Glasserman's [8] work on the subject.

In other words, the aim is to avoid taking the derivative of the payoff function, by using an integration-by-parts procedure. If Φ is a.s. differentiable with derivatives of polynomial growth, we are allowed to interchange the integral and differential operators; by virtue of the dominated convergence theorem.

In its most general case, let X be the parameter of the payoff function Φ and let g be the density function of X. The price of an option is then given by:

$$\mathcal{P} = \mathbb{E}\left[e^{-rT}\Phi(X)\right] = \int_0^\infty e^{-rT}\Phi(x)g(x)dx$$

Supposing that the payoff function does not depend on the parameter λ, the derivative of the price of the option with respect to λ is given by:

$$\begin{aligned}\frac{d\mathcal{P}}{d\lambda} &= \int_0^\infty e^{-rT}\Phi(x)\frac{\partial g(x)}{\partial \lambda}dx \\ &= \int_0^\infty e^{-rT}\Phi(x)\frac{\partial \ln(g(x))}{\partial \lambda}g(x)dx \\ &= \mathbb{E}\left[e^{-rT}\Phi(X)\frac{\partial \ln(g(X))}{\partial \lambda}\right]\end{aligned}$$

So when we know the distribution of the parameter of our function payoff, then we can express the Greeks as the expectation of the payoff multiplied by a weight, which does not depend on the payoff function.

Recalling the case for the European call option, we obtained the explicit integral for the price of the option from the probability density function of S_T

$$\mathcal{P}(x) = \int_{-\infty}^{\infty} e^{-rT}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy$$

We can show a proportional relationship between $\frac{\partial\Phi}{\partial x}$ and $\frac{\partial\Phi}{\partial y}$.
We have

$$\frac{\partial}{\partial x}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right) = e^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}$$

and

$$\frac{\partial}{\partial y}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right) = x\sigma\sqrt{T}e^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}$$

So

$$\frac{\partial}{\partial x}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right) = \frac{1}{x\sigma\sqrt{T}}\frac{\partial}{\partial y}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)$$

We can now establish an integration-by-parts result:

$$
\begin{aligned}
\frac{\partial}{\partial x}\mathcal{P} &= \frac{\partial}{\partial x}\left(\int_{-\infty}^{\infty} e^{-rT}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy\right)\\
&= e^{-rT}\int_{-\infty}^{\infty}\frac{\partial}{\partial x}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy\\
&= \frac{e^{-rT}}{x\sigma\sqrt{T}}\int_{-\infty}^{\infty}\frac{\partial}{\partial y}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy\\
&= \frac{e^{-rT}}{x\sigma\sqrt{T}}\left[\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}\right]_{-\infty}^{\infty}\\
&\quad+\frac{e^{-rT}}{x\sigma\sqrt{T}}\left[\int_{-\infty}^{\infty} y\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy\right]\\
&= \frac{e^{-rT}}{x\sigma\sqrt{T}}\left[\int_{-\infty}^{\infty}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)\frac{y}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy\right]\\
&= e^{-rT}\int_{-\infty}^{\infty}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)\underbrace{\frac{y}{x\sqrt{2\pi\sigma^2 T}}e^{-\frac{y^2}{2}}dy}_{\text{Lognormal Density}}
\end{aligned}
$$

$$
\text{Lognormal for } y \mapsto \frac{\log(\frac{x}{S_0}) - \left(r-\frac{\sigma^2}{2}\right)T}{\sigma T} = \frac{W_T}{\sigma\sqrt{T}} \text{ for } x = S_T
$$

$$
\begin{aligned}
&e^{-rT}\int_{-\infty}^{\infty}\Phi\left(xe^{rT+\sigma\sqrt{T}y-\frac{1}{2}\sigma^2 T}\right)\frac{W_T}{xT\sigma\sqrt{2\pi}}e^{-\frac{y^2}{2}}dy\\
&= \mathbb{E}_{\mathbb{P}}\left[\frac{e^{-rT}}{x\sigma T}W_T\Phi(S_T)\right]
\end{aligned}
$$

□

$$
\frac{\partial}{\partial x}\mathcal{P} = \mathbb{E}_{\mathbb{P}}\left[\underbrace{\frac{e^{-rT}}{x\sigma T}W_T}_{\text{Weight function}}\cdot\Phi(S_T)\right] \qquad (3.2.6)
$$

That which we set out to achieve has materialised by removing the differential operator. We have introduced a weight function $\frac{e^{-rT}}{x\sigma T}W_T$, which is independent of the payoff function, Φ and is straight forward to simulate using Monte Carlo methods. Now even if the payoff function is not continuous, the integration-by-parts method smoothes the payoff function with a weight independent of the payoff function and makes the method more efficient in general.

The main difference between the Greek simulation method and that of finite differences, is that the finite difference method requires price simulation whereas this method merely requires a *weight* to facilitate the simulation.

3.3 The Malliavin derivative approach to the Greeks

Assume we wish to calculate $\mathbb{E}[f'(X)Y]$. X and Y are two random variables. If the density functions of X and Y are unknown, we would need to remove the derivative, such that:

$$\mathbb{E}[f'(X)Y] = \mathbb{E}[f(X)H]$$

for some new random variable H.

Montero and Kohatsu-Higa [9], set out a routine to fulfill the above need. Let $Z = f(X)$ and apply the D operator on Z:

$$D_s Z = f'(X) D_s X$$

Add Y to the expression by multiplying both sides by $Yh(s)$ where h is an arbitrary process. h could depend on X, Y or both or even another random variable.

$$Yh(s) D_s Z = f'(X) Y h(s) D_s X$$

Now integrate over $s \in [0, T]$

$$\int_0^T Yh(s) D_s Z ds = \int_0^T f'(X) Y h(s) D_s X ds = f'(X) Y \int_0^T h(s) D_s X ds$$

So,

$$f'(X)Y = \int_0^T \frac{Yh(s) D_s Z}{\int_0^T h(v) D_v X dv} ds$$

Hence,

$$\mathbb{E}[f'(X)Y] = \mathbb{E}\left[\int_0^T (D_s Z) u_s ds\right]$$

with

$$u_s = \frac{Yh(s)}{\int_0^T h(v) D_v X dv}$$

Now, by applying the duality principle,

$$\mathbb{E}[f'(X)Y] = \mathbb{E}\left[f(X) D^* \left(\frac{Yh(\cdot)}{\int_0^T h(v) D_v X dv}\right)\right] \tag{3.3.1}$$

Letting,

$$H \equiv D^* \left(\frac{Y h(\cdot)}{\int_0^T h(v) D_v X dv} \right)$$

leads us to what we set out to achieve.

3.4 "Vanilla" options

3.4.1 European style options

The Greeks will be calculated for a European option, for the key sensitivities when considering a "Total hedge" i.e. Delta, Gamma and Vega.

Delta

Referring to (3.2.1), let $X(\alpha) = S_T$

$$\Delta = \frac{\partial}{\partial S_0}\mathbb{E}[e^{-rT}\Phi(S_T)] = e^{-rT}\mathbb{E}\left[\frac{\partial S_T}{\partial S_0}\Phi'(S_T)\right] = \frac{e^{-rT}}{S_0}\mathbb{E}\left[\Phi'(S_T)S_T\right]$$

Using the Malliavin integration-by-parts formula, (3.3.1) we get:

$$\Delta = \frac{e^{-rT}}{S_0}\mathbb{E}\left[\Phi(S_T)D^*\left(\frac{S_T}{\int_0^T D_v S_T dv}\right)\right]$$

Consider $\int_0^T D_v S_T dv$,

$$D_u S_T = D_u(S_0 e^{\{\mu T + \sigma W_T\}}) = S_T \cdot \sigma D_u(W_T) = \sigma S_T \mathbf{1}_{[0,T]}(u)$$

Hence

$$\int_0^T D_v S_T dv = \sigma S_T T$$

Now for

$$D^*\left(\frac{S_T}{\int_0^T D_v S_T dv}\right) = D^*\left(\frac{S_T}{\sigma S_T T}\right) = D^*\left(\frac{1}{\sigma T}\right),$$

using

$$\int_0^T F u_t dW_t = F\int_0^T u_t dW_t - \int_0^T (D_t F) u_t dt$$

with $F = \frac{1}{\sigma T}$ and $u_t \equiv 1$ we can see that

$$D^*\left(\frac{1}{\sigma T}\right) = \frac{W_T}{\sigma T}$$

So

$$\Delta = e^{-rT}\mathbb{E}\left[\Phi(S_T)\frac{W_T}{S_0\sigma T}\right]$$

□

Vega

Vega ($\mathcal{V}$), an anglo-hellenic term with no true Greek meaning, represents the sensitivity of the option price with respect to changes in the underlying volatility. In other words,

$$\mathcal{V} = \frac{\partial}{\partial\sigma}\mathbb{E}\left[e^{-rT}\Phi(S_T)\right] = \mathbb{E}\left[e^{-rT}\Phi'(S_T)\frac{\partial S_T}{\partial\sigma}\right]$$

Recall

$$S_T = S_0 e^{\mu T + \sigma W_T}$$

where $\mu = r - \frac{\sigma^2}{2}$. So, $\frac{\partial S_T}{\partial\sigma} = S_T(W_T - \sigma T)$, leaving Vega being represented as:

$$\mathcal{V} = \mathbb{E}\left[e^{-rT}\Phi'(S_T)S_T(W_T - \sigma T)\right]$$

Now using the standard integration-by-parts procedure

$$\begin{aligned}
\mathcal{V} &= \mathbb{E}\left[e^{-rT}\Phi(S_T)D^*\left(\frac{S_T(W_T - \sigma T)}{\int_0^T D_v S_T dv}\right)\right] \\
&= \mathbb{E}\left[e^{-rT}\Phi(S_T)D^*\left(\frac{S_T(W_T - \sigma T)}{\sigma S_T T}\right)\right] \\
&= \mathbb{E}\left[e^{-rT}\Phi(S_T)D^*\left(\frac{W_T}{\sigma T} - 1\right)\right]
\end{aligned}$$

Now focussing on $D^*\left(\frac{W_T}{\sigma T} - 1\right)$, we end up with $\frac{1}{\sigma T}D^*(W_T) - W_T$, where $D^*(W_T) = W_T^2 - T$ as show in Example 2.4.7, then Vega simplifies to:

$$\mathcal{V} = \mathbb{E}\left[e^{-rT}\left[\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma}\right]\Phi(S_T)\right]$$

Gamma

Gamma (Γ), represents the second order dependence of the price of the option with respect to values of the underlying.

$$\Gamma = \frac{\partial^2}{\partial S_0^2}\mathbb{E}\left[e^{-rT}\Phi(S_T)\right] = \frac{e^{-rT}}{S_0^2}\mathbb{E}\left[S_T^2\Phi''(S_T)\right]$$

Now using the integration-by-parts/duality 3.3.1, with $h \equiv 1$,

$$\Gamma = \frac{e^{-rT}}{S_0^2}\mathbb{E}\left[\Phi'(S_T)D^*\left(\frac{S_T^2}{\int_0^T D_v S_T dv}\right)\right] = \frac{e^{-rT}}{S_0^2}\mathbb{E}\left[\Phi'(S_T)D^*\left(\frac{S_T}{\sigma T}\right)\right]$$

We now focus on the stochastic integral $D^*\left(\frac{S_T}{\sigma T}\right)$, where we again use

$$\int_0^T F u_t dW_t = F\int_0^T u_t dW_t - \int_0^T (D_t F)u_t dt$$

letting $F = \frac{S_T}{\sigma T}$ and $u_t \equiv 1$:

$$\begin{aligned} D^*\left(\frac{S_T}{\sigma T}\right) &= \frac{S_T}{\sigma T}\int_0^T 1 dW_t - \frac{1}{\sigma T}\int_0^T D_t S_T dt \\ &= \frac{S_T}{\sigma T}\int_0^T 1 dW_t - \frac{1}{\sigma T}\sigma T S_T \\ &= S_T\left(\frac{W_T}{\sigma T} - 1\right) \end{aligned}$$

We thus have

$$\Gamma = \frac{e^{-rT}}{S_0^2}\mathbb{E}\left[\Phi'(S_T)S_T\left(\frac{W_T}{\sigma T} - 1\right)\right]$$

Applying integration-by-parts again:

$$\Gamma = \frac{e^{-rT}}{S_0^2}\mathbb{E}\left[\Phi(S_T)D^*\left(\frac{S_T\left(\frac{W_T}{\sigma T} - 1\right)}{\int_0^T D_v S_T dv}\right)\right]$$

Again we have to simplify the stochastic integral

$$\begin{aligned} D^*\left(\frac{S_T\left(\frac{W_T}{\sigma T} - 1\right)}{\int_0^T D_v S_T dv}\right) &= D^*\left(\frac{\left(\frac{W_T}{\sigma T} - 1\right)}{\sigma T}\right) \\ &= \frac{1}{\sigma T}\left[D^*\left(\frac{W_T}{\sigma T}\right) - W_T\right] \end{aligned}$$

Furthermore, using the result from Example 2.4.7 again and combining all the components, we have:

$$\Gamma = \mathbb{E}\left[\frac{e^{-rT}}{S_0^2\sigma T}\left(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma}\right)\Phi(S_T)\right]$$

Also, comparing the Gamma result to Vega, we notice the following relationship:

$$\Gamma = \frac{\mathcal{V}}{S_0^2\sigma T}$$

3.4.2 The explicit analytical computations of the European Greeks

The Greeks for the European type options can be written as expressions, since the density function of S_T has a closed formula, viz.

$$p(x) = \frac{1}{x\sqrt{2\pi\sigma^2 T}}\exp\left(\frac{-\left[\log\left(\frac{x}{S_0}\right) - \mu T\right]^2}{2\sigma^2 T}\right)$$

Recall that the price of the option is given as:

$$\mathcal{P} = \mathbb{E}\left[e^{-rT}\Phi(S_T)\right] = \int_0^\infty e^{-rT}\Phi(x)p(x)dx \qquad (3.4.1)$$

Now obtaining the Greeks is a simple case of partial differentiation with respect to the parameter concerned. We have to assume that the payoff function is independent of the parameter we are finding the sensitivity of. In this case it is obvious that the payoff function $\Phi(x) = (x-K)^+$ is not dependent on S_0.

Explicit European Delta

$$\begin{aligned}
\Delta &= \int_0^\infty e^{-rT}\Phi(x)\frac{\partial p(x)}{\partial S_0}dx \\
&= \int_0^\infty e^{-rT}\Phi(x)\frac{\partial \log p(x)}{\partial S_0}p(x)dx \\
&= \mathbb{E}\left[e^{-rT}\Phi(S_T)\left(\frac{\partial \log p(x)}{\partial S_0}\right)_{x=S_T}\right]
\end{aligned}$$

The last equation looks similar to the Malliavin result, where the derivative of the payoff function has been effectively removed leaving a suitable weight, i.e., $\frac{\partial \log p(x)}{\partial S_0}$ at $x = S_T$.

Also, after some calculus:

$$\begin{aligned}\left(\frac{\partial \log p(x)}{\partial S_0}\right)_{x=S_T} &= \frac{1}{S_0\sigma^2 T}\left[\log\left(\frac{x}{S_0}\right) - \mu T\right]_{x=S_T} \\ &= \frac{W_T}{S_0\sigma T}\end{aligned}$$

Hence we have:

$$\Delta = \mathbb{E}\left[e^{-rT}\Phi(S_T)\frac{W_T}{S_0\sigma T}\right],$$

which is the same as the result using the Malliavin calculus approach.

Explicit European Vega

Again we have the payoff function Φ, being independent of the sensitivity parameter σ.

$$\begin{aligned}\mathcal{V} &= \int_0^\infty e^{-rT}\Phi(x)\frac{\partial p(x)}{\partial \sigma}dx \\ &= \int_0^\infty e^{-rT}\Phi(x)\frac{\partial \log p(x)}{\partial \sigma}p(x)dx \\ &= \mathbb{E}\left[e^{-rT}\Phi(S_T)\left(\frac{\partial \log p(x)}{\partial \sigma}\right)_{x=S_T}\right]\end{aligned}$$

Again, after some calculus:

$$\begin{aligned}\left(\frac{\partial \log p(x)}{\partial \sigma}\right)_{x=S_T} &= -\frac{1}{\sigma} + \frac{1}{T\sigma^3}\left[\log\left(\frac{x}{S_0}\right) - \mu T\right]^2_{x=S_T} - \frac{1}{\sigma}\left[\log\left(\frac{x}{S_0}\right) - \mu T\right]_{x=S_T} \\ &= -\frac{1}{\sigma} + \frac{W_T^2}{\sigma T} - W_T\end{aligned}$$

Hence

$$\mathcal{V} = \mathbb{E}\left[e^{-rT}\Phi(S_T)\left(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma}\right)\right]$$

which is identical to the expression obtained via the Malliavin calculus.

Explicit European Gamma

When considering Gamma, a careful use of the chain rule is needed. For Delta we had the weight as $\left(\frac{\partial \log p(x)}{\partial S_0}\right)_{x=S_T}$. Therefore the Gamma weight is given by:

$$\frac{\partial}{\partial S_0}\left(\frac{\partial \log p(x)}{\partial S_0}\right) = \left(\frac{\partial \log p(x)}{\partial S_0}\right)^2 + \frac{\partial^2 \log p(x)}{\partial S_0^2}$$

We then calculate

$$\left(\frac{\partial \log p(x)}{\partial S_0}\right)^2 = \frac{W_T^2}{S_0^2 \sigma^2 T^2}$$

and

$$\begin{aligned}\frac{\partial^2 \log p(x)}{\partial S_0^2} &= \frac{\partial}{\partial S_0}\left[\frac{1}{S_0 \sigma^2 T}\left(\log\left(\frac{x}{S_0}\right) - \mu T\right)\right] \\ &= -\frac{W_T}{S_0^2 \sigma T} - \frac{1}{S_0^2 \sigma^2 T}\end{aligned}$$

Finally, Gamma is as the Malliavin calculus expression before:

$$\Gamma = \mathbb{E}\left[\frac{e^{-rT}}{S_0^2 \sigma T}\left(\frac{W_T^2}{\sigma T} - W_T - \frac{1}{\sigma}\right)\Phi(S_T)\right]$$

Notice that results from the Malliavin calculus related procedure, are equivalent to the respective partial derivatives (of the probability density function) of the analytical price of European style options with the Black Scholes Merton partial differential equation solution.

3.4.3 European numerical results

A plain vanilla European call option will be considered to demonstrate the Malliavin approach applied to numerical approximations of the Greeks. The payoff in this case will be

$$\Phi(X) = (X - K)^+$$

Delta Numerically

From the Black Scholes Merton formula, delta is easily seen as:

$$\Delta = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1(K)} e^{-\frac{x^2}{2}} dx$$

where

$$d_1(x) = \frac{1}{\sigma\sqrt{T}}\left[\log\left(\frac{S_0}{x}\right) + \left(r + \frac{1}{2}\sigma^2\right)T\right]$$

Suppose we have the following parameters:

r	0.1
σ	0.2
T	1
S_0	100
K	100

The following MATLAB code is used to perform the Monte Carlo approximation of the delta for the scenario.

```
% European Call option Delta Approximation

% using Malliavin calculus

% Set Parameters


clear s_0=100; x=100; r=0.1; t=1; sig=0.2; sim_cnt=10000;

%Assumptions

%input. "t" is in years assume no dividends

%Final Wiener path points at t=T

W_T=normrnd(0,sqrt(t),1,sim_cnt);

%Final Stock price and payoff

final_stock_price=s_0.*exp((r-sig*sig*0.5)*t+sig*sqrt(t)*W_T);

payoff_matrix=vertcat(final_stock_price-x,zeros(1,sim_cnt));

payoff=max(payoff_matrix);


delta_mallaivin=(1/sim_cnt).*(dot(payoff,W_T).*(exp(-r*t)/(s_0*sig*t)));

analytic_delta=BLSDELTA(s_0,x,r,t,sig,0); %=0.7257
```

```
%Plot Malliavin Delta Progression

for i=1:sim_cnt
    malliavin_delta_progression(i) =
    1/i).*(dot(payoff(1:i),W_T(1:i)).*(exp(-r*t)/(s_0*sig*t)));

    est_variance(i)=var(payoff(1:i).*W_T(1:i));
end

plot(1:sim_cnt,malliavin_delta_progression,
     1:sim_cnt,BLSDELTA(s_0,x,r,t,sig,0)*ones(sim_cnt,1))
grid on;
axis([0 sim_cnt 0.5 1]);
```

The exact delta of 0.7257 is well approximated by the Malliavin approach.

Vega Numerically

The explicit formula for Vega is given as

$$\mathcal{V} = S_0\sqrt{\frac{T}{2\pi}}e^{-\frac{[d_1(K)]^2}{2}}$$

with $d_1(K)$ given as before. The scenario parameters are the same as those used previously.

Gamma Numerically

The explicit formula for Gamma is given as

$$\Gamma = \frac{1}{S_0\sqrt{2\pi\sigma^2 T}}e^{-\frac{[d_1(K)]^2}{2}}$$

with $d_1(K)$ given as before. We will use the same scenario parameters as before too.

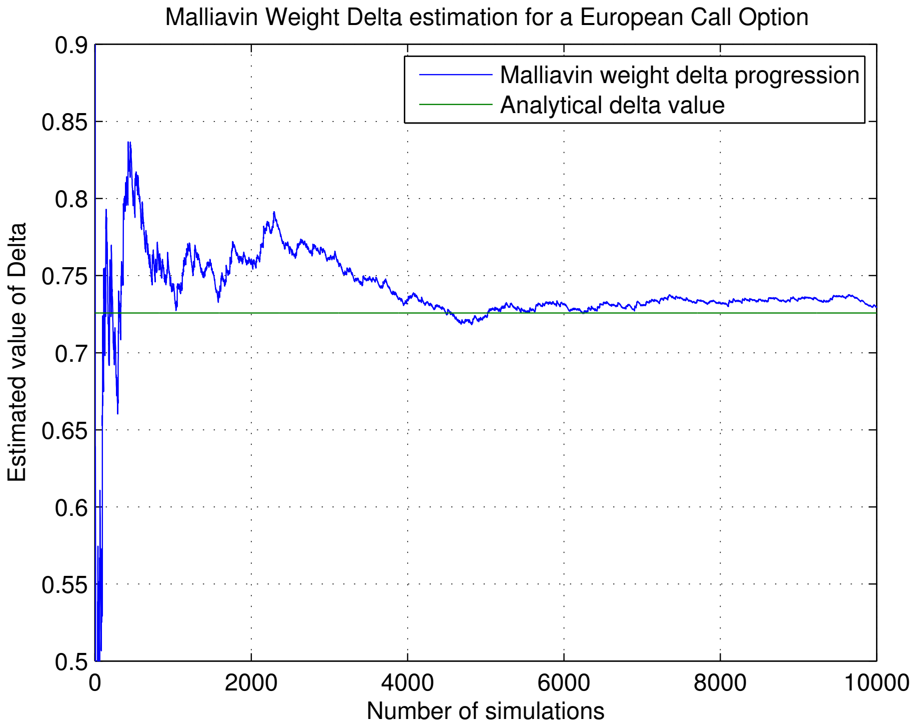

Figure 3.1: The Malliavin progression of an estimated of Delta vs the analytical solution

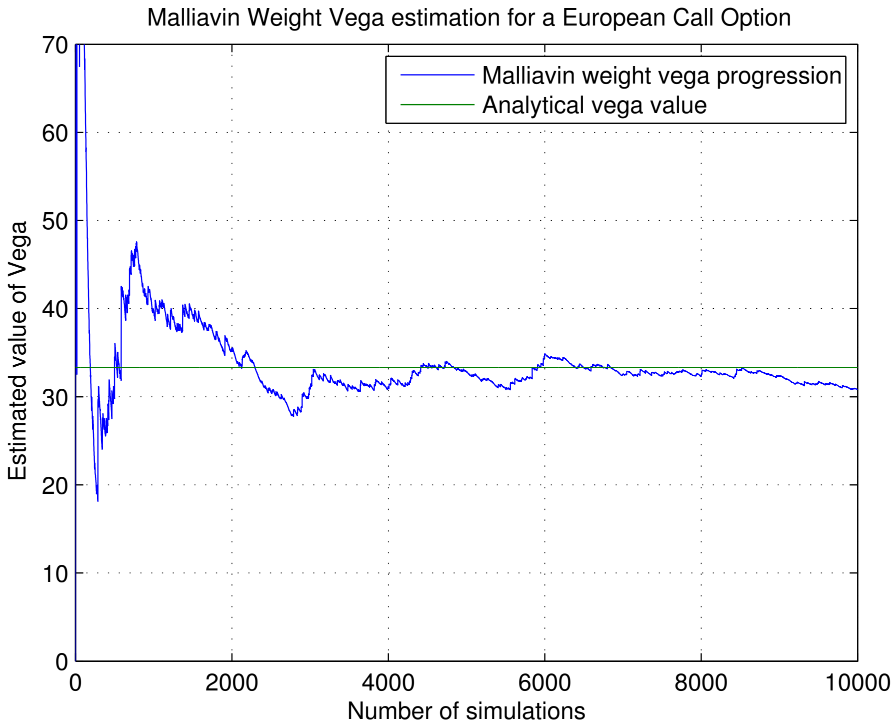

Figure 3.2: The Malliavin progression of an estimated of Vega vs the analytical solution

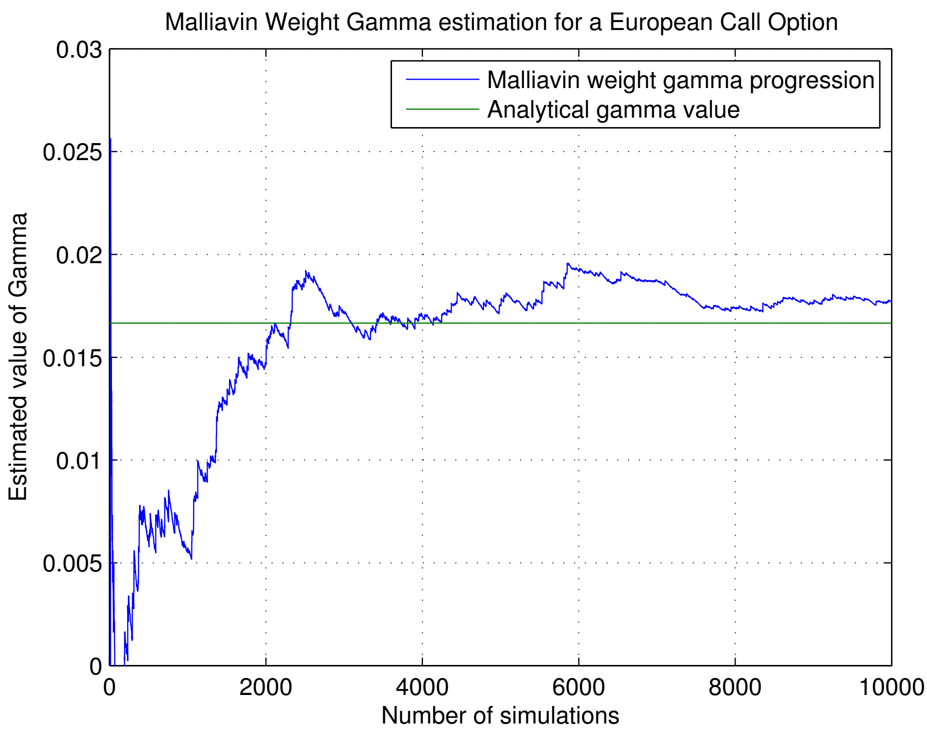

Figure 3.3: The Malliavin progression of an estimated of Gamma vs the analytical solution

3.5 "Exotic" options

3.5.1 Asian style options

Asian options are options in which the underlying variable is the average price over a period of time (path dependent options). Due to this fact, Asian options have a lower volatility, rendering them cheaper relative to their European counterparts.

Asian options are commonly traded on currencies and commodity products which have low trading volumes. They were originally used in 1987 when Banker's Trust Tokyo office used them for pricing average options on crude oil contracts, hence the name "Asian" option.

These options are broadly segregated into three categories; arithmetic average Asians, geometric average Asians and both of these forms can be averaged on a weighted average basis, whereby a given weight is applied to each stock price being averaged. This can be useful for attaining an average on a sample with a highly skewed sample population.

To elaborate on arithmetic averaging, this is seen as being the sum of the sampled asset prices divided by the number of samples:

$$Avg_A = \frac{S_1 + S_2 + ... + S_n}{n}$$

and for geometric averaging, the average value is taken as:

$$Avg_G = \sqrt[n]{S_1 \cdot S_2 \cdot ... \cdot S_n}$$

The payoff functions for Asian options (where the average stock price for either averaging method is denoted as $\overline{S}$) are given as:
For an average price Asian:

$$V = \eta(0, \overline{S} - K)^+ \qquad \text{fixed strike price}$$

and an average strike Asian:

$$V = \eta(0, S_T - \overline{S})^+ \qquad \text{floating strike price}$$

Where η is a binary variable which is set to 1 for a call, and -1 for a put. Asians can be both European style or American style in exercise.

3.5.2 The Asian delta

We now consider the financial sensitivities of options written on the average stock price $\frac{1}{T}\int_0^T S_t dt$ (continuous Asian), instead of only the terminal value S_T as with the European style options, with respect to the underlying.

We have to note that the density function of the random variable $\frac{1}{T}\int_0^T S_t dt$ does not have a known closed formula.

Delta in this case is given by

$$\Delta = \frac{\partial}{\partial S_0}\mathbb{E}\left[e^{-rT}\Phi\left(\frac{1}{T}\int_0^T S_s ds\right)\right] = \frac{e^{-rT}}{S_0}\mathbb{E}\left[\Phi'\left(\frac{1}{T}\int_0^T S_s ds\right)\frac{1}{T}\int_0^T S_u du\right] \tag{3.5.1}$$

We now demonstrate the versatility of the Malliavin integration-by-parts formula in obtaining various expressions for Δ.

Proposition 3.5.1. *Delta as stated by Fournié et al [6] can be shown to be:*

$$\Delta = \frac{2e^{-rT}}{S_0\sigma^2}\mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right)\left(\frac{S_T - S_0}{\int_0^T S_t dt} - \mu\right)\right]$$

with $\mu = r - \frac{\sigma^2}{2}$

Proof. To obtain the result we use 3.5.1 without $\frac{e^{-rT}}{S_0}$ and

$$\mathbb{E}\left[f(X)D^*\left(\frac{Yh(\cdot)}{\int_0^T h(v)D_v X dv}\right)\right] = \mathbb{E}\left[f'(X)Y\right]$$

with $X = \frac{1}{T}\int_0^T S_s ds$, $Y = \frac{1}{T}\int_0^T S_u du$ and $h_t = S_t$ such that

$$\begin{aligned}
&\mathbb{E}\left[\Phi'\left(\frac{1}{T}\int_0^T S_s ds\right)\frac{1}{T}\int_0^T S_u du\right] \\
&= \mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right)D^*\left(\frac{\frac{1}{T}\int_0^T S_u du \cdot S.}{\int_0^T S(v)D_v\left(\frac{1}{T}\int_0^T S_s ds\right)dv}\right)\right]
\end{aligned}$$

Let us first consider $D_v\left(\frac{1}{T}\int_0^T S_s ds\right)$ in the denominator, using the Malliavin derivative of an Itô process.

$$
\begin{aligned}
D_v\left(\frac{1}{T}\int_0^T S_s ds\right) &= \frac{1}{T}\int_0^T D_v S_s ds \\
&= \frac{1}{T}\int_0^T D_v\left(S_0 e^{\mu s+\sigma W_s}\right) ds \\
&= \frac{1}{T}\int_0^T S_s(D_v \sigma W_s) ds \\
&= \frac{\sigma}{T}\int_0^T S_s \mathbf{1}_{\{v\le s\}} ds \\
&= \frac{\sigma}{T}\int_v^T S_s ds
\end{aligned}
$$

Now considering the entire denominator of the fraction on which the adjoint is operating, we have:

$$
\int_0^T S(v) D_v\left(\frac{1}{T}\int_0^T S_s ds\right) dv = \frac{\sigma}{T}\int_0^T S(v)\int_v^T S_s ds dv
$$

Using a version of Fubini's theorem

$$
\int_0^T S(v)\int_v^T S_s ds dv = \frac{1}{2}\left(\int_0^T S(v) dv\right)^2
$$

We have

$$
\begin{aligned}
&\mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right) D^*\left(\frac{\frac{1}{T}\int_0^T S_u du \cdot S.}{\int_0^T S(v) D_v\left(\frac{1}{T}\int_0^T S_s ds\right) dv}\right)\right] \\
&= \mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right) D^*\left(\frac{\int_0^T S_u du \cdot S.}{\frac{\sigma}{2}\left(\int_0^T S(v) dv\right)^2}\right)\right] \\
&= \mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right) \frac{2}{\sigma} D^*\left(\frac{\int_0^T S_u du \cdot S.}{\left(\int_0^T S(v) dv\right)^2}\right)\right] \\
&= \mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right) \frac{2}{\sigma} D^*\left(\frac{S.}{\left(\int_0^T S_s ds\right)}\right)\right] \\
&= \mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right) \frac{2}{\sigma} \int_0^T \left(\int_0^T S_s ds\right)^{-1} S_t dW_t\right]
\end{aligned}
$$

Now, we should consider,

$$D^*\left(\frac{S_\cdot}{\left(\int_0^T S_s ds\right)}\right)=\int_0^T\left(\int_0^T S_s ds\right)^{-1} S_t dW_t$$

within the expectation carefully.

We know that integration-by-parts provides the following relationship:

$$\int_0^T X u_s dW_s = X\int_0^T u_s dW_s - \int_0^T (D_s X) u_s ds$$

Letting the process u_s be S_s and the random variable X be $\frac{1}{\int_0^T S_\kappa d\kappa}$ we then have

$$\int_0^T\left(\int_0^T S_\kappa d\kappa\right)^{-1}\cdot S_t dW_t = \frac{1}{\int_0^T S_\kappa d\kappa}\cdot\int_0^T S_s dW_s - \int_0^T\left(D_s\left(\frac{1}{\int_0^T S_k dk}\right)\cdot S_s\right) ds$$

Looking only at

$$D_s\left(\frac{1}{\int_0^T S_k dk}\right)$$

$$\begin{aligned}
D_s\left(\frac{1}{\int_0^T S_k dk}\right) &= -\left(\int_0^T S_\kappa d\kappa\right)^{-2}\cdot D_s\left(\int_0^T S_\kappa d\kappa\right)\\
&= -\left(\int_0^T S_\kappa d\kappa\right)^{-2}\cdot\int_0^T D_s S_\kappa d\kappa\\
&= -\left(\int_0^T S_\kappa d\kappa\right)^{-2}\cdot\int_0^T S_\kappa \sigma \chi_{\{s\leq\kappa\}}(s) d\kappa\\
&= -\sigma\left(\int_0^T S_\kappa d\kappa\right)^{-2}\cdot\int_s^T S_\kappa d\kappa
\end{aligned}$$

Multiplying by $\frac{\int S_s ds}{\int S_s ds}$ and using the version of Fubini's theorem $\int_0^T S_v \int_v^T S_s ds dv = \frac{1}{2}\left(\int_0^T S_v dv\right)^2$ we get:

$$D_s\left(\frac{1}{\int_0^T S_k dk}\right) = -\sigma\left(\int_0^T S_\kappa d\kappa\right)^{-2} \cdot \frac{\int_0^T S_s \int_s^T S_\kappa d\kappa ds}{\int_0^T S_s ds}$$
$$= -\sigma\left(\int_0^T S_\kappa d\kappa\right)^{-2} \cdot \frac{\frac{1}{2}\left(\int_0^T S_\kappa d\kappa\right)^2}{\int_0^T S_s ds}$$
$$= \frac{-\sigma}{2\int_0^T S_s ds}$$

Returning to:

$$\int_0^T \left(\int_0^T S_\kappa d\kappa\right)^{-1} \cdot S_t dW_t = \frac{1}{\int_0^T S_\kappa d\kappa} \cdot \int_0^T S_s dW_s - \int_0^T \left(D_s\left(\frac{1}{\int_0^T S_k dk}\right) \cdot S_s\right) ds$$
$$= \frac{1}{\int_0^T S_\kappa d\kappa} \cdot \int_0^T S_s dW_s + \frac{\sigma}{2}\int_0^T \left(\frac{1}{\int_0^T S_s ds} \cdot S_s\right) ds$$
$$= \frac{1}{\int_0^T S_\kappa d\kappa} \cdot \int_0^T S_s dW_s + \frac{\sigma}{2}$$

Finally Δ can be written as

$$\Delta = \frac{e^{-rT}}{S_0}\mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right)\left(\frac{2}{\sigma}\frac{\int_0^T S_t dW_t}{\int_0^T S_s ds} + 1\right)\right]$$

Furthermore we know the dynamics of the S_t process.

$$S_T = S_0 + r\int_0^T S_t dt + \sigma\int_0^T S_t dW_t$$

such that

$$\sigma\int_0^T S_t dW_t = \sigma D^*(S_t) = S_T - S_0 - r\int_0^T S_t dt$$

Hence

$$\begin{aligned}
\Delta &= \mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right)\left(\frac{2}{\sigma^2}\left(\frac{S_T - S_0 - r\int_0^T S_t dt}{\int_0^T S_s ds}\right)+1\right)\right] \\
&= \mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right)\left(\frac{2}{\sigma^2}\left(\frac{S_T - S_0}{\int_0^T S_s ds}\right)-\frac{2r-\sigma^2}{\sigma^2}\right)\right] \\
&= \frac{2}{\sigma^2}\mathbb{E}\left[\Phi\left(\frac{1}{T}\int_0^T S_s ds\right)\left(\frac{S_T - S_0}{\int_0^T S_s ds}-\mu\right)\right]
\end{aligned}$$

Including the $\frac{e^{-rT}}{S_0}$ we have the result.

□

3.5.3 The continuous Asian delta numerically

Malliavin weighting method

To perform the numerical computation many thousands of asset price paths are simulated using the given problem parameters. The associated continuous Asian payoffs are calculated

$$\Phi\left(\frac{1}{T}\int_0^T S_t dt\right) = \left(\frac{1}{T}\int_0^T S_t dt - K, 0\right)^+$$

Proposition 3.5.1 will be used to estimate delta. We therefore also need the weight:

$$\frac{S_T - S_0}{\int_0^T S_s ds} - \mu$$

During the numerical procedure, the integral used in both the payoff and the Malliavin weight, $\int_0^T S_s ds$, was approximated as the sum of the simulated paths. A financial year was partitioned into 252 days. Then, for each sample path a version of Proposition 3.5.1 is calculated. Finally, the expectation of these values is taken to approximate delta.

Finite difference method

For the finite difference method, a centered difference scheme was used. Primarily, a price for the Asian option needs to be calculated using the expectation of the discounted payoff of thousands of sample paths. Then, using the same random numbers, the centred difference is taken for the associated price perturbations to estimate a delta.

Note that the finite difference method shows similar efficiency for the Malliavin scheme. The Malliavin scheme is best seen when estimating Gamma for Asian options and when calculating Greeks on extreme discontinuous payoff options e.g. Digital options.

Numerical experiment

Consider a fixed strike Asian call with the following parameters

r	0.1
σ	0.2
T	1
S_0	100
K	100

and using 10,000 path simulations.

Figure 3.4 shows a graphical view of the estimates. The value of delta is approximately 0.6494.

3.5.4 Digital style options

A digital option (also called a binary option) is a cash settled option that has a discontinuous payoff. Digital options behave similarly to standard options, but the payout is based on whether the option is in the money, not by how much it is in the money. Digital options come in many forms, but the two most basic are cash-or-nothing and asset-or-nothing. Each can be European or American and can be structured as a put or a call.

We will consider a European cash-or-nothing digital call option, with its payoff function given as:

$$\Phi_T = Q \cdot \mathbf{1}_{\{S_T > K\}}$$

Furthermore, a European cash-or-nothing digital pays a fixed amount of money, Q, if it expires in-the-money and nothing otherwise. In other words a European cash-or-nothing digital call option makes a fixed payment if the option expires with the underlier above the strike price. It pays nothing if it expires with the underlier equal to or less than the strike price.

Reiner and Rubinstein's classic 1991 paper [12], introduced these options and a set of closed form analytical formulas, which can be applied to the pricing of these options, giving payoffs within a Black-Scholes framework

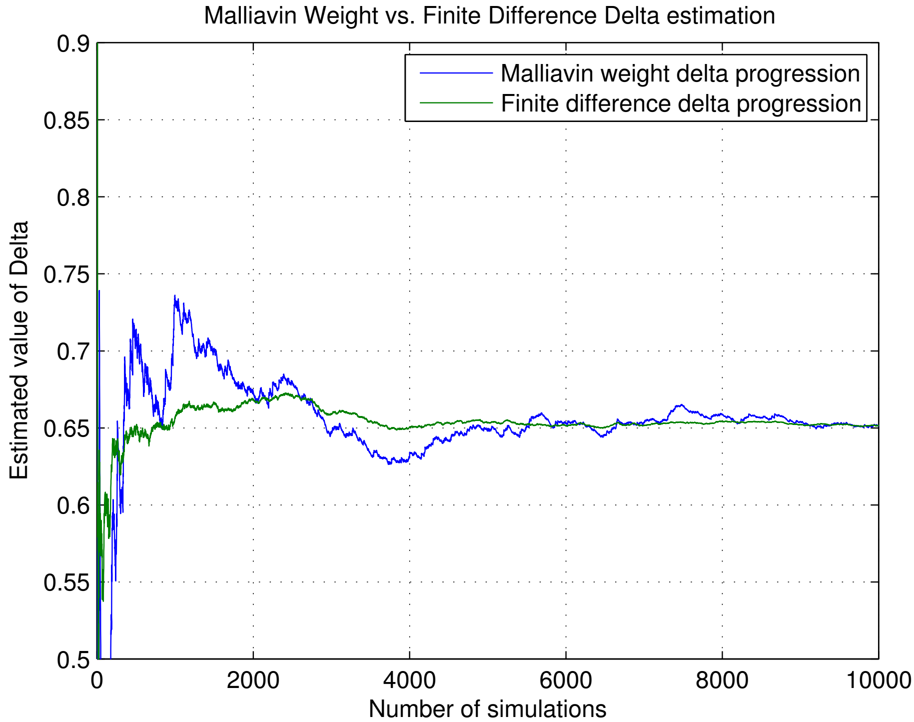

Figure 3.4: The Malliavin progression of an estimate of an Asian delta vs the finite difference progression approximation of the same delta.

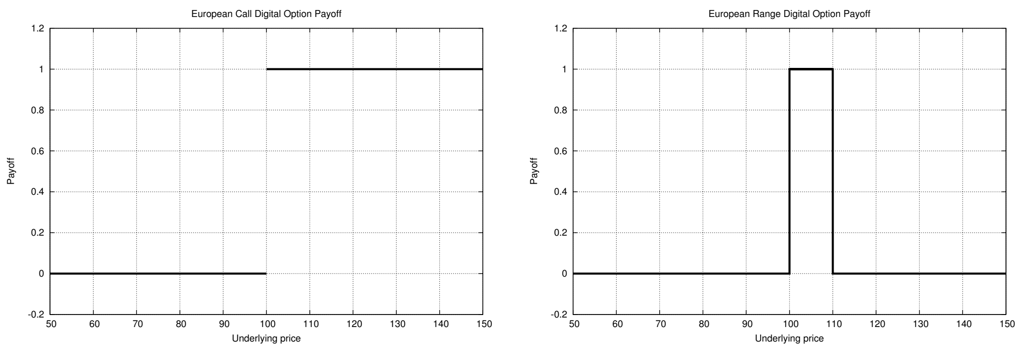

Figure 3.5: Payoff function for European digital call option (left panel), strike price, $X = 100$ and European range option (right panel), payoff being $\mathbf{1}_{[100,110]}$.

shown previously. For standard cash-or-nothing digital options the pricing formula is simple:

$$C_{\text{Digital}} = e^{-rT}N(d_2) \text{ and } P_{\text{Digital}} = e^{-rT}N(-d_2)$$

where, assuming no dividends are present:

$$d_1 = \frac{\ln\frac{S_0}{K} + (r + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}},\ d_2 = d_1 - \sigma\sqrt{T}$$

Notice that the formula for the call premium resembles the amount of cash held in the equivalent hedging portfolio for a vanilla call option, namely $-Ke^{-rT}N(d_2)$.

The analytical solutions for delta and gamma Greeks of European digital options, well illustrated by Avellaneda and Lawrence [1], are the following:

$$\Delta_{\text{Digital}} = \frac{e^{-rT}e^{-\frac{d_2^2}{2}}}{S_0\sqrt{2\pi\sigma^2 T}}$$

$$\Gamma_{\text{Digital}} = -\frac{e^{-rT}e^{-\frac{d_2^2}{2}}}{S_0^2\sigma^2 T\sqrt{2\pi}} \cdot d_1$$

Range Digital

A composite option, consisting of two digital option positions will now be considered. What is known as a range binary is a combination of two digital options. This option has a payoff of $Q \cdot \mathbf{1}_{[a,b]}$, receiving a cash amount Q at expiration should $S_T \in [a, b]$, otherwise the option is worthless at expiration. More specifically the range digital is a combination of longing a digital call with strike price a and shorting a digital call with strike price b. The same payoff can be achieved by longing a digital put with strike price b and shorting a digital put with strike price a.

To demonstrate the differences between Monte Carlo simulations using finite difference and the Malliavin weighting scheme (the identical weighting function as the vanilla call option case is used with the associated payoff function) we shall use the case of a European range digital option with payoff $\Phi_T = \mathbf{1}_{[100,110]}$. See Figure 3.5. The following parameters were used: $\sigma = 0.2$, $T = 1$, $r = 0.1$ and $S_0 = 100$. Delta and gamma are demonstrated in Figures 3.6 and 3.7. Vega is easily inferred from its known relationship with gamma.

Rate of convergence in numerical methods

Benhamou [2] provides the best work on the numerical approximations of the Greeks using Malliavin calculus. I refer to his work to observe the benefit of the Malliavin calculus in terms of the Asian gamma. In the below depictions, of the comparison between the Digital option delta and gamma versus the same estimation via finite difference, it is clear the Malliavin weighted scheme converges to the correct answer quickly with almost no oscillations.

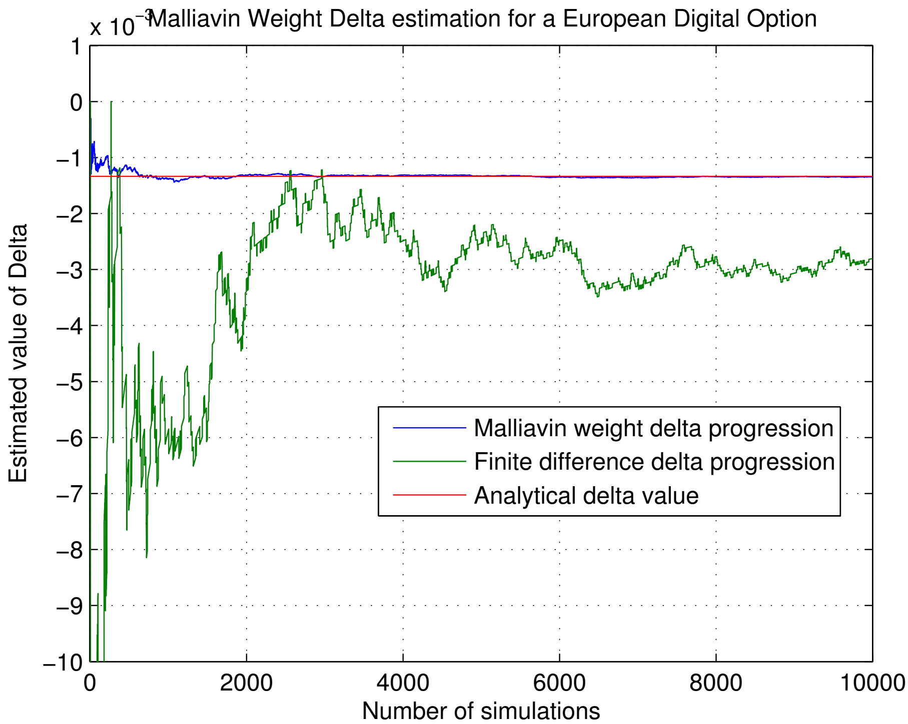

Figure 3.6: The Malliavin progression of an estimate of a European Digital Delta vs the finite difference progression approximation of the same Delta

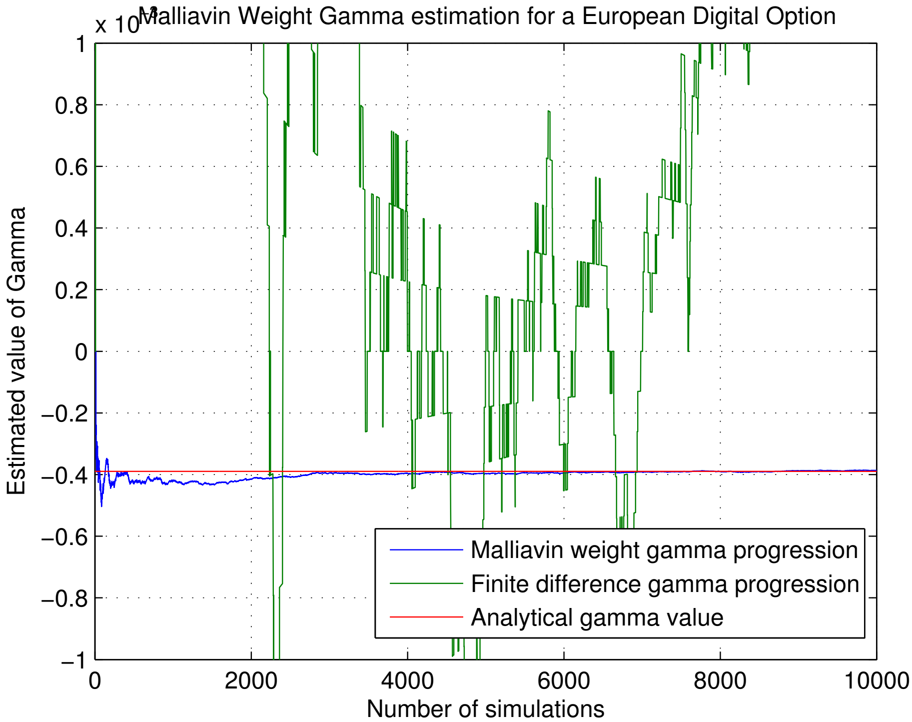

Figure 3.7: The Malliavin progression of an estimate of a European Digital Gamma vs the finite difference progression approximation of the same Gamma

3.6 Monte Carlo simulation and variance reduction

3.6.1 Simulation routine

The basic Monte Carlo method using the Malliavin weight, is as follows:
For interest rates constant

1. Sample a random path for S_t in the risk neutral world
2. Calculate option contingent payoff's
3. Repeat Step 1 and 2 many times
4. Calculate the mean sample payoff, to approximate the expected payoff
5. Discount the expected payoff multiplied by the Malliavin weight

3.6.2 Variance reduction

The choice of h, used in establishing the Malliavin weight in (3.3.1), helps reduce the variance within the Monte Carlo simulation. If Φ is the payoff function and π is the Malliavin weight then we wish to minimise the variance of $\Phi(X_1, ..., X_n) \times \pi$. This would naturally make Monte Carlo simulations converge faster.

One has to ask the question as to how to choose the optimal weight. For each of the Greeks, due to the open choice of h there are infinitely many possible weights. Therefore, a good idea is to look for the weight which minimises the variance of the estimator. Since all the payoff functions are $\mathcal{F}_T$-measurable, the weight with minimal variance will be given by the projection theorem as the conditional expectation of any weight function with respect to the filtration $\mathcal{F}_T$. Elie and Prével [4] provide the following proposition:

Proposition 3.6.1. *The weight function with minimal variance denoted by π_0 is the conditional expectation of any weight function π with respect to the filtration $\mathcal{F}_T$*

$$\pi_0 = \mathbb{E}[\pi|\mathcal{F}_T]$$

This strong result indicates that the best weighting function is always $\mathcal{F}_T$ measurable. Intuitively, we use only the information we have without adding any noise, that is why we get the minimal variance payoff. We see that, in the case of payoff functions which can be expressed in terms of some particular points of the Brownian motion trajectory, the best weight function will be the one expressed in terms of these particular points. However, in term of implementation, we may prefer to choose a generator which is $\mathcal{F}_t$ adapted. In this case, the Skorohod integral can be rewritten as an Itô integral which is much easier to implement. Therefore, it may be more interesting to choose the minimal variance weight beyond the one generated by $\mathcal{F}_t$ adapted generators.

In summary, if π depends only on $(X_1, ..., X_n)$, then π is a variance minimising choice. E.g. $\frac{W_T}{S_0 \sigma T}$ the weight in the European call delta case; S_T is dependent on W_T.

3.7 Conclusion

The dissertation gave a brief history of the Malliavin calculus, particularly the initial rationale for the theory and some of the consequences of the theory.

The major results of the adjoint operator of the Malliavin derivative operator, called the Skorohod integral, being the extension of the Itô integral for non-adapted processes was successfully introduced and explored. Furthermore the integration-by-parts formula, which relates the derivative operator on the Wiener Space and the Skorohod extended stochastic integral was proved in several ways.

As a bridge to the Malliavin calculus' use in finance, the derivative of Itô processes was shown, particularly the derivative of the Black Scholes process.

Calculation of financial contingent claim sensitivities is shown via a direct method, likelihood method, finite difference and the Malliavin weighting scheme. The various methods allowed for strong comparisons to be made, especially via numerical experiments, between the industry used finite difference and Malliavin version.

Despite most practical real world cases having sufficient estimation techniques and the need for daunting stochastic calculus and complicated analytic computation (making the Malliavin method error prone in setting up and potentially overestimating its value) for the Malliavin alternative-the application was still a suitable objective to explore the underlying theory.

In summary, this dissertation has shown that the new Malliavin calculus technique demonstrates extremely powerful alternatives to pricing problems where no explicit formula is available. The key benefit of the Malliavin calculus theory is that it imposes few restrictions on the payoff function and when used correctly, gives very efficient Monte Carlo schemes for the computation of the Greeks.

Bibliography

[1] Avellaneda M., Lawrence P. *Quantitative Modeling of Derivative Securities: From Theory to Practice.* Chapman & Hall / CRC, Boca Raton, 2000.

[2] Benhamou E. An application of Malliavin calculus to continuous time Asian options Greeks. Technical report, Technical Report, London School of Economics, 2000.

[3] Bismut J.M. *Large deviations and the Malliavin calculus.* Boston: Birkhauser., 1984.

[4] Elie R. and Prével J.C. *Application of Malliavin calculus to the computation of the Greeks*, 2003.

[5] Elworthy K.D. Stochastic flows in Riemannian manifolds,. *Diffusion Problems and Related Problems in Analytics*, pages 37–72, 1992.

[6] Fournié, E., Lasry J.M. Lebuchoux J., Lions P.L., Touzi N. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance and Stichastics*, (3):391–412, 1999.

[7] Friz P.K. *An introduction to Malliavin Calculus.* Courant Institute of Mathematical Sciences, New York University, 2005.

[8] Glasserman P Broadie, M. Estimating security price derivatives using simulation. *Management Science*, (42):269–285, 1996.

[9] Montero M., Kohatsu-Higa A. Malliavin calculus applied to finance. *Phisica A*, 320:548–570, 2003.

[10] Nualart D. *The Malliavin Calculus and Related Topics.* Springer, Berlin, 1995.

[11] Øksendal B. *An Introduction to Malliavin Calculus with Applications to Economics*, 1997. Lecture Notes.

[12] Rubinstein M., Reiner E. Unscrambling the binary code. *Risk*, 4:75–83, October 1991.

[13] Stroock D. W. Malliavin calculus, a functional analytic approach. *Journal of Functional Analysis*, 44(2):212–257, 1981.

[14] Ustünel A.S. *An introduction to Analysis on Wiener Space*. Springer, 1992.

[15] Watanabe S. *Stochastic Differential Equations and Malliavin Calculus*. Springer, 1984.