Instituto Nacional de Matemática Pura e Aplicada

IMPA

On the Risk Premium for Option Pricing in a Stochastic-Volatility Financial Model: Malliavin and PDE Techniques.

Cesar Augusto Gomez Velez

Tese apresentada para a obtenção do Título de
Doutor em Ciências

Rio de Janeiro

Março, 2007

Agradecimentos

Primeiramente agradeço a Deus e a vida por me permitir realizar este trabalho. Agradeço também de maneira muito especial ao professor Jorge Zubelli, tanto por introduzir-me neste belo e apaixionante campo quanto pela la sua paciente orientação e interesse no desenvolvimento do trabalho.

Agradeço a os professores Carlos Isnard, Andre Nachbin e o mesmo Jorge Zubelli pelo conhecimento que tem dividido comigo.

Agradeço a Leidjane Zacariotto e a Carlos Andre de Roure pelo seu apoio e conselho, a meus companheiros Mahendra Panthee, Javier Solano, Juan Carlos Galvis, Damian Fernandez, Freddy Hernandez, Jimmy Santamaria, Didier Pilot ao professor e amigo José Santiago Fajardo.

Aos professores Matheus Grasselli e Max Oliveira de Souza pelos valiosos comentários sobre o presente trabalho.

Também agradeso a muitas outras pessoas que não cito nestes agradecimentos, mais com quem meu coração esta igualmente grato, a amizade que tem me ofrecido com certeza tem contribuido também a convertir este trabalho numa realidade.

O tempo dividido com cada um deles eles, entre discussoẽs sobre matemática e outros aspectos da vida tem feito da minha passagem pelo IMPA uma das melhores experiências de minha vida. Todos eles fizeram os tempos difíceis mais amenos e os bons momemtos tornaran-se inesqueçíveis. Agradeço a vida ter me colocado perto destas pessoas com quem além de ter aprendido matemática, tenho aprendido sobre as alegrias e tristezas da vida, rindo e chorando mais no final cresçendo como pessoa.

Rio de Janeiro, Brasil
27 de Março de 2007.

Contents

Introduction

A crucial problem in financial mathematics is that of pricing derivative contracts under realistic assumptions. Among these contracts are the so-called *options*. An option gives its holder the right, *but not the obligation*, to buy or sell a certain amount of a financial asset, by a certain date, for a certain strike price. The writer of the option must specify:

- The type of the option: If the option is to buy it is named a *call* whereas if the option is to sell it is called a *put*. These two are called *vanilla* options.
- The underlying asset: Typically, it is a stock, a bond, a currency, or an index;
- The amount of the underlying to be purchased or sold;
- The expiration date of the contract; if the option can be exercised at any date before maturity, it is called an *American* option but, if it can only be exercise at maturity, it is called a *European* option;
- The strike price at which the transaction is performed if the option is exercised.

Let us examine the case of a European call option on a given stock. We will assume that the price of the stock at time t is given by X_t where $X = \{X_t\}_{t\geq 0}$ is a stochastic process in some probability space. Let us assume also that we fixed an initial time t_0 when the value of X_t is x, i.e., $X_{t_0} = x$. Let us call T the expiration date and K the exercise price. Obviously, if K is greater than X_T, the holder of the option has no interest whatsoever in exercising the option. But, if $X_T > K$, then the holder makes a profit of $X_T - K$ by exercising the option, i.e., buying the stock for K and selling it back on the market at X_T. Therefore, the value of the call at maturity is given by

$$(X_T - K)^+ = \max(X_T - K, 0).$$

Actually, we are interested in the price of the option at an initial time t_0. Obviously, at this moment in time we do not know the value $(X_T - K)^+$, since it is also a random variable.

The ground-breaking work of Black-Scholes [BS73] and the analysis of Merton [Mer73], by means of a general approach based on reasonable simplifying assumptions, led to pricing formulas for the vanilla options. In their model, they realized that, at time t_0, the price of the European call of our discussion is completely determined despite the random character of the *payoff* $(X_T - K)^+$ at the maturity time T. In order to introduce the forthcoming ideas of this work and to prepare the notation let us develop informally some ideas concerning the fair pricing of a call option.

At a first glance, and without no prior information, the conditional expected value of the quantity $(X_T - K)^+$

$$\widetilde{U} = \mathbb{E}[(X_T - K)^+ | X_{t_0} = x] \ , \tag{1}$$

would be a good guess for the value of the payoff $(X_T - K)^+$ at time T. Under the usual assumption of riskless borrowing at an interest rate r, such value should be brought to *present value* and thus, at time t_0, the price $\widehat{U}$ of the option would be

$$\widehat{U} = e^{-r(T-t_0)}\mathbb{E}[(X_T - K)^+ | X_{t_0} = x] \ . \tag{2}$$

However, a very important point is the fact that actually the expectation in Equation (2) must be computed not with respect to the probability measure $\mathbb{P}$, associated to the random dynamics that models the stock dynamics of X_t (*physical measure*), but with respect to an *equivalent measure*. This measure is characterized by the property that for the *discounted stock price process* $S_t = e^{-r(t-t_0)}X_t$ we have that the conditional expectations $\mathbb{E}[S_{t_2}|S_{t_1}]$ for $t_0 \leq t_1 \leq t_2 \leq T$ satisfy the property

$$\mathbb{E}[S_{t_2}|S_{t_1}] = S_{t_1} \ . \tag{3}$$

We then say that the discounted stock price process S_t satisfies the *martingale* property with respect to this new probability measure $\mathbb{Q}$ called the *risk neutral measure (r.n.m)* or the *martingale measure*. The main point being that any of the two situations where

$$\mathbb{E}^{\mathbb{Q}}[S_{t_2}|S_{t_1}] \lessgtr S_{t_1}, \tag{4}$$

is equivalent to the fact of having a positive probability of an *arbitrage opportunity*. Intuitively, if it happens that

$$\mathbb{E}^{\mathbb{Q}}[S_{t_2}|S_{t_1}] < S_{t_1},$$

then with a positive probability the value S_{t_2} would be less than S_t at time $t = t_1$. So it would be better for someone to sell short the stock at t_1 and deposit this amount S_{t_1} in the bank. At time t_2 he will have $e^{r(t_2-t_1)}S_{t_1} > S_{t_1}$ which he can use to return the stock that is now worth S_{t_2}. Thus, with nonzero probability the investor would have made a profit. For a precise definition of non-arbitrage and a proof of the statement above see [Duf01, Sch03, KK01].

One of the main results of Black and Scholes[BS73] is that the price function $\widehat{U}$ in (2) defined as function of the stock price x and a time before maturity $t < T$, i.e. $\widehat{U}(x,t)$[1] satisfies the final value problem:

$$\begin{aligned} &\frac{\partial \widehat{U}}{\partial t} + \frac{\sigma^2 x^2}{2}\frac{\partial^2 \widehat{U}}{\partial x^2} + r(x\frac{\partial \widehat{U}}{\partial x} - \widehat{U}) = 0, \quad 0 < x < \infty, \\ &\widehat{U}(x,T) = (x-K)^+. \end{aligned} \tag{5}$$

In the Black-Scholes (BS) model it is assumed that, under the risk neutral measure, the stock evolves according to the following stochastic differential equation (SDE):

$$\begin{aligned} dX_t &= rX_t dt + \sigma X_t dW_t^Q, \\ X_{t_0} &= x. \end{aligned} \tag{6}$$

Here, W_t^Q is the Brownian motion w.r.t. $\mathbb{Q}$ and the positive constants r and σ are the *interest rate* and the *volatility*, respectively.

Thus, in general the solution U of (5) depends also on the other parameters of the model

$$U(x,t) = U_{BS}(x,t,T,K,r,\sigma^2). \tag{7}$$

It turns out that, if we fix all the other arguments but the volatility, the dependence of U_{BS} on σ defines a $1-1$ map from $\mathbb{R}^+$ onto its range. By *implied volatility* we mean the inverse of such map. Thus, the implied volatility of a traded option is the volatility value that, substituted into the Black-Scholes equation, produces the known price of the option. It is a standard practice in the market to compare option prices by means of their implied volatilities.

Despite its power and impact, the BS model has a number of drawbacks. For example, the implied volatility should be a constant independent of the strike price K and the time to expiration. This, however, is not verified in practice. A number of more realistic models have been suggested in order to mitigate such problems. Among such models, stochastic volatility ones have gained tremendous popularity in the past.

For stochastic volatility models (SVM) it is assumed that the stock price satisfies the following stochastic differential equation

$$dX_t = \mu X_t dt + \sigma_t X_t dW_t \ , \tag{8}$$

where, in contrast to the BS model which assumes that the parameter σ is constant, the volatility evolves according to a stochastic process of the form $\sigma_t = \sigma(Y_t)$, where $\sigma(\cdot)$ is dependent on another suitable stochastic process $Y = \{Y_t\}_{t\geq 0}$. We will concentrate ourselves on the case where Y is an *Ornstein-Uhlenbeck* (OU) type process. In other words, Y satisfies the SDE

$$dY_t = \alpha(m - Y_t)dt + \beta d\widehat{W}_t, \tag{9}$$

[1]These variables (x,t), would correspond to (x,t_0) in formula (2).

where $\widehat{W}_t$ is another Brownian motion possibly correlated to W_t, so we can assume that $d\widehat{W} = \rho dW_t + (1-\rho^2)^{1/2} dZ_t$, were Z_t is a Brownian motion independent of W_t.

Among other features, the process Y given by (9) is normally distributed and has a tendency to revert back to its long-run mean level m, with a velocity that depends on the coefficients α and β.

Despite its elegance, pricing in stochastic volatility models is substantially more complicated. In this context, the market is no longer complete and thus we do not have a unique equivalent risk neutral measure as in the BS model [Duf01]. Such measures are parameterized by functions of the form $\gamma = \gamma(x, y, t)$. For each such risk neutral measure $\mathbb{Q}^\gamma$, the price evolves according to

$$\begin{aligned}
&dX_t = rX_t dt + \sigma(Y_t) X_t dW_t^1, \\
&X_{t_0} = x, \\
&dY_t = \big(\alpha(m - Y_t) - \beta\Lambda(X_t, Y_t, t)\big)dt + \beta(\rho dW_t^1 + (1-\rho^2)^{1/2} dW_t^2), \\
&Y_{t_0} = y \ .
\end{aligned} \tag{10}$$

Here, $\Lambda(x, y, t) = \rho\frac{\mu - r}{\sigma(y)} + (1-\rho^2)^{1/2}\gamma(x, y, t)$ and as usual W_t^1 and W_t^2 are independent Brownian motions.

The option price U in such model is a solution to the partial differential equation [FPS00]

$$\begin{aligned}
&U_t + \frac{\sigma(y)^2 x^2}{2} U_{xx} + \frac{\beta^2}{2} U_{yy} + \rho\beta x \sigma(y) U_{xy} + \\
&(\alpha(m - y) - \beta\Lambda(x, y, t))U_y - r(xU_x - U) = 0 \ , \\
&U(x, y, T) = (x - K)^+ \ ,
\end{aligned} \tag{11}$$

where

$$\Lambda(x, y, t) = \rho\frac{\mu - r}{\sigma(y)} + \gamma(x, y, t)\sqrt{1 - \rho^2} \ .$$

Let us note in Equation (11) that now the price function $U(x, y, t, K, T)$ depends on the volatility level y at time t. In Equation (11), $\gamma(\cdot)$ is termed the *market price of volatility risk* or more succintly the *risk premium* [2]. For a justification of this terminology, in loose terms, (small) increases in the volatility risk β leads to increases in the rate of return on the option price by a factor proportional to γ. More precisely, we have, using Itô formula and Eq. (11), that

$$\begin{aligned}
dU(X_t, Y_t, t) = &\Big[\frac{\mu - r}{\sigma(Y_t)}\big(X_t \sigma(Y_t)\frac{\partial U}{\partial x} + \beta\rho\frac{\partial U}{\partial y}\big) + rU + \gamma\beta(1-\rho^2)^{1/2}\frac{\partial U}{\partial y}\Big]dt \\
&+ \big(X_t\sigma(Y_t)\frac{\partial U}{\partial x} + \beta\rho\frac{\partial U}{\partial y}\big)dW_t + \beta(1-\rho^2)^{1/2}\frac{\partial U}{\partial y} dZ_t.
\end{aligned}$$

[2] See Section 2.5 of [FPS00].

One of the difficulties with stochastic volatility models is the choice of a suitable risk premium function γ. One approach to circumvent such difficulty is to consider practical situations whereby one prices the options as *asymptotic* perturbations of the standard BS model and computes the corrections directly from market data without having to estimate the risk neutral measure. This approach was pursued by Fouque, Papanicolaou, and Sircar in a series of works [FPS00, FPSS03a, FPSS03b]. It relies heavily on the hypothesis of fast mean reversion, i.e., $\alpha >> 1$ of the Ornstein-Uhlenbeck process Y_t that models the volatility in (10). This hypothesis, although verified in a number of markets, is not always true. This leads to the natural question of how to estimate the risk premium γ from market data, as well as, how to compute the sensitivity of the prices with respect to its change. More precisely, we are naturally led to the following questions:

i How smoothly do the prices depend on γ?

ii How sensitive are the prices on γ and how to compute the sensitivity of the prices with respect to γ?

iii Does the knowledge of the prices determine γ uniquely, or, is there some local uniqueness?

iv How to estimate γ ?

In this thesis we settle Question 1 above by showing analyticity in suitable function spaces for the case where γ is independent of x. We also compute the functional derivative of the price w.r.t. to γ in suitable spaces by means of both PDE techniques and Malliavin calculus techniques, thus addressing Question 2.

Finally, we discuss Questions 3 and 4. In this part we are concerned with the practical problem of calibrating or estimating the *risk premium* γ in Equation (11) assuming that we have at our disposal knowledge of the solution $U(x, y, t)$.

One might be tempted to consider a direct approach to estimating the coefficients of (11). For example, by directly solving for γ. However, this has several drawbacks. In practice, we do not know the function $U(x, y, t)$ for all values of y. Furthermore, differentiation is an ill-posed operation and any noise in the data may contaminate the results. Thus, we are forced to apply regularization techniques to our problem within the scope of inverse problem theory [HH03, Cré03, BI99, CT06].

Let us state the main objectives of this work:

i To establish the main property of analyticity of the map $\Gamma : \gamma \to U$ in appropriate function spaces. This will be exploited, for example in order to devise iterative methods to handle the calibration problem.

ii To consider an iterative technique to tackle the inverse problem, we focused on the framework of classical *Landweber* regularization [BL05, EHN96, Sch95]. This involves computing the functional derivative of the map $\Gamma : \gamma \to U$.

iii To explore the use of a stochastic variational calculus, namely *Malliavin calculus*, to carry out the computation of the functional derivative mentioned above and compare this approach with that of classical calculations using partial differential equations. Let us remark the fact that Malliavin calculus has been applied recently by E. Fournié, J.-M. Lasry, J. Lebuchoux, N. Touzi and P.-L. Lions to compute the so-called *option greeks,*[3] for some option pricing models [FLLL01, FLL$^+$99], see also [Ben02, Lio00].

iv To design a thorough strategy to invert the operator Γ, that would be suitable to numerical implementation. To accomplish this, it may be helpful to have as much understanding as possible of the mechanism of the direct map Γ that sends the function γ to the corresponding solution $U(x, y, t)$ of (11). Other examples of applications of inverse problems to mathematical finance can be found in [HH03, Cré03, BI99, CT06].

Let us also give here a brief description of the contents of this thesis.

Chapter 1 consists of background material to keep this work as self contained as possible. It starts with a short introduction to Malliavin calculus. It will be used later to compute the Fréchet derivative of $\Gamma(\gamma)$. We introduce the notion of the derivative of random variables defined on a Gaussian probability space, and a few of its properties that will allow us to perform some calculations. In this chapter, there is also a reference to Girsanov's theorem and the relation between partial differential equations and stochastic differential equations (SDEs).

In Chapter 2, it is proved that the operator $\Gamma(\gamma)$ is differentiable, and indeed it is analytic in a suitable space. The rest of the chapter is devoted to finding a way of computing the Fréchet derivative $\frac{\partial \Gamma}{\partial \gamma}$ through the use of Malliavin calculus and to see if such approach would provide any advantage with respect to traditional methods of calculation using partial differential equations.

In Chapter 3, we try to design a strategy to handle the inversion of the map $\Gamma : \gamma \to U$ in the case when the Brownian motion of the asset price process X_t and that of the volatility process Y_t are uncorrelated, i.e, $\rho = 0$. We discuss whether the technique based on Malliavin calculus to compute functional derivatives developed in Chapter 2, can be used in the numerical treatment of the reconstruction problem. The chapter ends with a short note suggesting how to deal with inversion problem in the more difficult correlated case where $\rho \neq 0$.

We close with some conclusions and suggestions for further research in Chapter 4.

[3] The option greeks are derivatives of the option prices with respect to different quantities of interest.

CHAPTER 1

Some Analytical Tools

In this chapter we will give an introduction to some notions of *Malliavin Calculus*. We will focus on what is needed for the purpose of the sequel. We follow closely the exposition in [Nua06]. Good references for the prerequisites in stochastic calculus for this chapter and the others are [Øks03, Kar88, Fri75]. For applications of Malliavin calculus to mathematical finance see:[FLL$^+$99, Ben02, Lio00].

Consider the separable Hilbert space $H = L^2(T, \mathcal{B}, \lambda)$ where $(T, \mathcal{B})$ is a measurable space formed by an interval $T = [0, T] \subset \mathbb{R}$ with the Borel σ-algebra $\mathcal{B}$, and the Lebesgue measure λ defined on $\mathcal{B}$.

We consider the class of stochastic Itô integrals $W(h) := \int_0^T h(s)dW_s$ of the elements $h \in H$ with respect to the canonical Brownian motion W_s defined on $\Omega = \mathcal{C}_0([0, T]; \mathbb{R})$ and with associated probability measure P. Let $\mathcal{F}$ denote the σ-algebra generated by all the random variables of the form $W(h), \quad h \in L^2(T, \mathcal{B}, \lambda)$. Now let $W(A) := W(1_A)$, where $A \in \mathcal{B}$ and 1_A is its characteristic function. Note that we can think of W as an $L^2(\Omega, \mathcal{F}, P)$-valued measure on the parameter space $(T, \mathcal{B})$ which takes independent values (in the stochastic sense) on any family of disjoint Borel subsets of T, and such that any random variable $W(A)$ is distributed as $N(0, \lambda(A))$ if $\lambda(A) < \infty$. We will call W a Gaussian measure, (or Brownian) in $(T, \mathcal{B})$. This measure is also called *white noise* based on λ.

1.1 The Derivative Operator

We want to introduce the derivative DF of a square integrable random variable $F : \Omega \longrightarrow \mathbb{R}$. This means that we want to differentiate F with respect to the parameter $\omega \in \Omega$. In the usual applications of this theory, the random variables F are only defined P-almost-everywhere (a.e) and do not possess a continuous version. For this reason, we will introduce a notion of derivative in a weak sense. We denote by $\mathcal{C}_p^\infty(\mathbb{R}^n)$ the set of all continuously differentiable functions $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ such that f and all its partial derivatives have polynomial growth.

Let $\mathcal{S}$ denote the class of smooth random variables $F \in \mathcal{S}$ of the form

$$F = f(W(h_1), \ldots, W(h_n)), \tag{1.1}$$

where f belongs to $\mathcal{C}_p^\infty(\mathbb{R}^n)$, and $h_1, \ldots, h_n$ are in $H = L^2(T, \mathcal{B}, \lambda)$ for $n \geq 1$. We will denote by $\mathcal{S}_b$ and $\mathcal{S}_0$ the classes of smooth random variables of the form (1.1) such that the function f belongs to $\mathcal{C}_b^\infty(\mathbb{R}^n)$ and to $\mathcal{C}_0^\infty(\mathbb{R}^n)$ respectively. Here $f \in \mathcal{C}_b^\infty$ means that f is infinitely differentiable and is bounded together with all its partial derivatives, $f \in \mathcal{C}_0^\infty(\mathbb{R}^n)$ means that f is infinitely differentiable and has compact support. Moreover, we will denote by $\mathcal{P}$ the class of random variables of the form (1.1), such that f is a polynomial. Note that $\mathcal{P} \subset \mathcal{S}$, $\mathcal{S}_0 \subset \mathcal{S}_b \subset \mathcal{S}$ and that $\mathcal{P}$ and $\mathcal{S}_0$ are dense in $L^2(\Omega)$.

Definition 1. *The derivative of a smooth random variable of the form (1.1), is the stochastic process $\{D_tF\}_{t\in T}$ given by*

$$D_tF = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(W(h_1), \ldots, W(h_n))h_i(t). \tag{1.2}$$

For example, $D_tW(h) = h(t)$. We will consider DF as an element of $L^2(T \times \Omega) \cong L^2(\Omega; H)$. In order to interpret DF as a directional derivative, note that for any element $h \in H$ we have

$$\begin{aligned} \langle DF, h\rangle_H &= \lim_{\epsilon\to 0} \frac{1}{\epsilon}[f(W(h_1) + \epsilon\langle h_1, h\rangle_H, \cdots, W(h_n) + \epsilon\langle h_n, h\rangle_H) \\ &- f(W(h_1), \cdots, W(h_n))]. \end{aligned}$$

An important example: Consider the Brownian motion on the interval $[0, 1]$, so that Ω is the canonical space $\Omega = \mathcal{C}_0([0, 1]; \mathbb{R})$. In this case, $\langle DF, h\rangle_H$ can be interpreted as a directional *Fréchet derivative*. In fact, let us introduce the subspace H^1 of Ω which consists of all the absolutely continuous functions $x : [0, 1] \to \mathbb{R}$, with a square integrable density, i.e. $x(t) = \int_0^t \dot{x}(s)ds$, $\dot{x} \in H = L^2([0, 1]; \mathbb{R})$. The space H^1 is usually called the *Cameron-Martin space*. We can transport the Hilbert space structure of H to H^1 by putting

$$\langle x, y\rangle_{H^1} = \langle \dot{x}, \dot{y}\rangle_H = \int_0^t \dot{x}(s)\dot{y}(s)ds.$$

Thus, H^1 becomes a Hilbert space isomorphic to H. The injection of H^1 into H is continuous, since we have

$$\sup_{0\leq t\leq 1} |x(t)| \leq \int_0^1 |\dot{x}(s)|ds \leq \|\dot{x}\| = \|x\|_{H^1}.$$

Consider a smooth functional of the particular form $F = f(W(t_1), \ldots, W(t_n))$, $f \in \mathcal{C}_p^\infty(\mathbb{R}), 0 \leq t_1 \leq \cdots < t_n \leq 1$, where $W(t_i) = \int_0^{t_i} dW_s = W(1_{[0,t_i]})$. Notice that such functional is continuous in Ω. Then, for any function $h \in H$, the scalar product $\langle DF, h\rangle_H$ coincides with the directional derivative of F in the direction of the element $\int_0^{\cdot} h(s)ds$ which belongs to H^1. In fact

$$
\begin{aligned}
\langle DF, h\rangle_H &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(W(t_1), \ldots, W(t_n))\langle 1_{[0,t_i]}, h\rangle_H \\
&= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(W(t_1), \ldots, W(t_n)) \int_0^{t_i} h(s)ds \\
&= \frac{d}{d\epsilon}\left[F(\omega + \epsilon \int_0^{\cdot} h(s)ds)\right]_{\epsilon=0}.
\end{aligned}
$$

The following result is an integration-by-parts formula, which will play an important role.

Lemma 1. *Suppose that F is a smooth functional and $h \in H = L^2(T)$. Then,*

$$\mathbb{E}[\langle DF, h\rangle_H] = \mathbb{E}[FW(h)]. \tag{1.3}$$

Proof: First notice that we can normalize Eq (1.3) and assume that the norm of h is one. There exist orthonormal elements of H, say $e_1, \ldots, e_n$, such that $h = e_1$ and F is a random variable of the form

$$F = f(W(e_1), \ldots, W(e_n)),$$

where f is in $\mathcal{C}_p^\infty(\mathbb{R})$. Let $\phi(x)$ denote the density of the standard normal distribution on $\mathbb{R}^n$, that is,

$$\phi(x) = (2\pi)^{-\frac{n}{2}} \exp(-\frac{1}{2}\sum_{i=1}^n x_i^2).$$

Then, we have,

$$
\begin{aligned}
\mathbb{E}[\langle DF, h\rangle_H] &= \int_{\mathbb{R}^n} \frac{\partial f}{\partial x_1}(x)\phi(x)dx \\
&= \int_{\mathbb{R}^n} f(x)\phi(x)x_1 dx \\
&= \mathbb{E}[FW(e_1)] = \mathbb{E}[FW(h)]. \quad \square
\end{aligned}
$$

***Remark*:** We can see from Lemma 1 that the operator D is closable as an operator from $L^p(\Omega)$ into $L^p(\Omega; L^2(T))$ so it makes sense to use the following conventions:

We will denote the domain of D in $L^p(\Omega)$ by $\mathbb{D}^{1,p}$, meaning that $\mathbb{D}^{1,p}$ is the closure of the class of smooth random variables $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p} = \left[\mathbb{E}[|F|^p] + \mathbb{E}[\|DF\|^p_{L^2(T)}]\right]^{\frac{1}{p}}.$$

For $p = 2$, the space $\mathbb{D}^{1,2}$ is a Hilbert space with the scalar product

$$\langle F, G\rangle_{1,2} = \mathbb{E}[FG] + \mathbb{E}[\langle DF, DG\rangle_H].$$

More generally, we can introduce iterated derivatives of k-times weakly differentiable random variables. If F is a smooth random variable and k is an integer, we set

$$D^k_{t_1,\dots,t_k}F = D_{t_1}D_{t_2}\cdots D_{t_k}F.$$

Note that for a smooth random variable F, the derivative D^kF is considered as a measurable function on the product space $T^k \times \Omega$, which is defined a.e. with respect to the measure $\lambda^k \times P$. Then, for every $p \geq 1$ and any natural number k we introduce the seminorm on $\mathcal{S}$ defined by

$$\|F\|_{k,p} = \Big[\mathbb{E}[|F|^p] + \sum_{i=1}^{k} \mathbb{E}[\|D^iF\|^p_{L^2(T)}]\Big]^{\frac{1}{p}}. \tag{1.4}$$

This family of seminorms satisfies the following properties:

- *Monotonicity*: $\|F\|_{k,p} \leq \|F\|_{j,q}$, for any $F \in \mathcal{S}$, if $p \leq q$ and $k \leq j$.
- *Closability*: The operator D^k is closable from $\mathcal{S}$ into $L^p(\Omega; H^{\otimes k})$.
- *Compatibility*: Let $p, q \geq 1$ be real numbers and k, j be natural numbers. Suppose that F_n is a sequence of smooth random variables such that $\|F_n\|_{k,p}$ converges to zero as n tends to infinity, and $\|F_n - F_m\|_{j,q}$ tends to zero as n and m tend to infinity. Then $\|F_n\|_{j,q}$ tends to zero as n tends to infinity; This is a consequence of the closability of the operators $D^i, i \geq 1$ on $\mathcal{S}$.

  The completion of the family of smooth random variables $\mathcal{S}$ with respect to the norm $\|\cdot\|_{k,p}$ is denoted by $\mathbb{D}^{k,p}$. and it follows from the first property above that $\mathbb{D}^{k+1,p} \subset \mathbb{D}^{k,q}$ if $k \geq 0$ and $p > q$.

Now, we will state the chain rule. It can be easily proved by approximating the random variable F by smooth random variables and the function ϕ by $(\phi * \psi_\epsilon)c_M$, where $\{\psi_\epsilon\}$ is an approximation of the identity and c_M is a $\mathcal{C}^\infty$ function such that $0 \leq c_M \leq 1$, $c_M = 1$ if $|x| \leq M$ and $c_M = 0$ if $|x| \geq M + 1$.

Proposition 1. *Let $\phi : \mathbb{R}^m \to \mathbb{R}$ be a continuously differentiable function with bounded partial derivatives, and $p \geq 1$. Suppose that $F = (F^1, \dots, F^m)$ is a random vector whose components belong to the space $\mathbb{D}^{1,p}$. Then, $\phi(F) \in \mathbb{D}^{1,p}$ and*

$$D(\phi(F)) = \sum_{i=1}^{m} \frac{\partial \phi}{\partial x_i}(F)DF^i. \tag{1.5}$$

The next result concerns the derivative of a conditional expectation with respect to a σ-field generated by Gaussian stochastic integrals. Let $A \in \mathcal{B}$, denote by $\mathcal{F}_A$ the σ-field (completed with respect to P) generated by the random variables $\{W(B) : B \subset A \text{ and } B \in \mathcal{B}\}$.

Proposition 2. *Suppose that F belongs to $\mathbb{D}^{1,2}$, and let $A \in \mathcal{B}$. Then, the conditional expectation $\mathbb{E}[F|\mathcal{F}_A]$ also belongs to the space $\mathbb{D}^{1,2}$ and we have*

$$D_t(\mathbb{E}[F|\mathcal{F}_A]) = \mathbb{E}[D_t F|\mathcal{F}_A]1_A(t),$$

a.e. in $T \times \Omega$.

Remark: In particular, if F belongs to $\mathbb{D}^{1,2}$ and is $\mathcal{F}_A$-measurable, then $D_t F$ is zero a.e. in $A^c \times \Omega$.

1.2 The Skorohod Integral

In this section, we will consider the adjoint of the derivative operator and we will show that it coincides with a generalization of the stochastic integral introduced by Skorohod in [Sko75].

Definition 2. *We denote by δ the adjoint of the operator D. That means δ is an unbounded operator on $L^2(T \times \Omega)$ with values in $L^2(\Omega)$ such that:*

- *The domain of δ, denoted by $Dom(\delta)$, is the set of processes $u \in L^2(T \times \Omega)$ such that*

$$\left| \mathbb{E}[\int_T D_t F u_t \lambda(dt)] \right| \leq c\|F\|_2, \tag{1.6}$$

  for all $F \in \mathbb{D}^{1,2}$, where c is some constant depending on u.

- *If u belongs to $Dom(\delta)$, then $\delta(u)$ is the element of $L^2(\Omega)$ characterized by*

$$\mathbb{E}[F\delta(u)] = \mathbb{E}[\int_T D_t F u_t \lambda(dt)], \tag{1.7}$$

  for any $F \in \mathbb{D}^{1,2}$.

We refer to the operator δ as the Skorohod stochastic integral of the process u. It transforms square integrable processes into random variables. The following notation is usual:

$$\delta(u) = \int_T u_t \delta W_t.$$

1.2.1. Properties of the Skorohod Integral

We will now state some properties of the Skorohod integral that are common in calculations that apply Malliavin calculus to mathematical finance. We suggest the interested reader to go to [Nua06], for the ditails.

First notice that the Skorohod integral has zero mean, that is, $\mathbb{E}[\delta(u)] = 0$ if $u \in \mathrm{Dom}(\delta)$ and that δ is a linear operator on $\mathrm{Dom}(\delta)$.

i *Integration of Smooth Elementary Processes.*
We denote by $\mathcal{S}_H$ the class of smooth elementary processes of the form

$$u(t) = \sum_{i=1}^{n} F_i h_i(t),$$

where the F_i are smooth random variables, and the h_i are elements of $H = L^2(T)$. From the integration-by-parts formula established in Lemma 1, we deduce that a process of this type is Skorohod integrable and moreover that

$$\delta(u) = \sum_{i=1}^{n} F_i W(h_i) - \sum_{i=1}^{n} \int_T D_t F_i h_i(t) \lambda(dt). \tag{1.8}$$

We see here that the Skorohod integral of a smooth elementary process can be decomposed into two parts, one that can be considered as a pathwise integral and another one that involves the derivative operator. We remark that if for every i, the function h_i is an indicator 1_{A_i} of a set $A_i \in \mathcal{B}$, and F_i is $\mathcal{F}_{A_i^c}$-measurable, then by the remark after Proposition 2, the second summand of equation (1.8), vanishes and the Skorohod integral of u is just the first summand of (1.8).

Definition 3. *We denote by $\mathbb{L}^{1,2}$, the class of processes $u \in L^2(T \times \Omega)$ such that $u(t) \in \mathbb{D}^{1,2}$ for almost all t and there exists a measurable version of the two parameter process $D_s u_t$ verifying*

$$\mathbb{E}[\int_T \int_T (D_s u_t)^2 \lambda(ds)\lambda(dt)] < \infty$$

$\mathbb{L}^{1,2}$ is a Hilbert space with the norm

$$\|u\|_{1,2}^2 = \|u\|_{L^2(T\times\Omega)}^2 + \|Du\|_{L^2(T^2\times\Omega)}^2,$$

it can be seen that $\mathbb{L}^{1,2}$ is isomorphic to $L^2(T; \mathbb{D}^{1,2})$, and it is shown that $\mathbb{L}^{1,2} \subset \mathrm{Dom}(\delta)$.

ii *Commutativity Relation Between the Derivative and the Skorohod Integral.*
Suppose that u is a process in the space $\mathbb{L}^{1,2}$. We assume that for almost all t the process $\{D_t u_s\}_{s\in T}$ is Skorohod integrable and there is a version of the process

$\{\int_T (D_t u_s)\delta W_s\}_{t\in T}$ which is in $L^2(T\times\Omega)$. Then, the Skorohod integral $\delta(u)$ belongs to $\mathbb{D}^{1,2}$, and we have

$$D_t\delta(u) = u_t + \int_T (D_t u_s)\delta W_s. \tag{1.9}$$

iii *The Skorohod Integral of a Process Multiplied by a Random Variable.*
Suppose that u is a Skorohod integrable process. Let F be a random variable in the space $\mathbb{D}^{1,2}$ such that $\mathbb{E}[F^2\int_T u_t^2\lambda(dt)] < \infty$. Then it holds that

$$\int_T (Fu_t)\delta W_t = F\int_T u_t\delta W_t - \int_T (D_tF)u_t\lambda(dt), \tag{1.10}$$

in the sense that Fu_t is integrable if the right-hand side of (1.10) belongs to $L^2(\Omega)$.

iv ***The Itô Stochastic Integral as a Particular Case of the Skorohod Integral***.

Recall that for $A\in\mathcal{B}$ we denote by $\mathcal{F}_A$ the σ-algebra $\sigma(\{W(B) : B\subset A, B\in\mathcal{B}\})$.

Lemma 2. *Let $A\in\mathcal{B}$ and let F be a square integrable random variable which is measurable with respect to the σ-algebra $\mathcal{F}_{A^c}$. Then, the process $F1_A$ is Skorohod integrable and*

$$\delta(F1_A) = FW(A).$$

Proof: By (1.10) and the remark after Proposition 2 we have

$$\delta(F1_A) = FW(A) - \int_T D_tF1_A(t)\lambda(dt) = FW(A). \quad \square$$

Using this lemma we can show that the operator δ is an extension of the Itô integral in the case of Brownian motion. Let $W = \{W_t\}_{t\in[0,1]}$ be Brownian motion. We denote by L_a^2 the closed subspace of $L^2([0,1]\times\Omega;\mathbb{R}) \cong L^2(T\times\Omega)$ formed by the adapted process, i.e, processes measurable with respect to the σ-algebra $\mathcal{F}_{[0,t]}$, where $\mathcal{F}_{[0,t]}$ includes all the information until the time t. Then, we have the following:

Proposition 3. *$L_a^2\subset Dom(\delta)$, and the operator δ restricted to L_a^2 coincides with the Itô integral. That is,*

$$\delta(u) = \int_0^1 u_t dW_t. \tag{1.11}$$

Proof: Assume that u is an elementary adapted process of the form

$$u_t = \sum_{i=1}^{n} F_i 1_{(t_i, t_{i+1}]}(t),$$

where $F_j \in L^2(\Omega, \mathcal{F}_{t_i}, P; \mathbb{R})$, and $0 \leq t_1 < \cdots < t_{n+1} \leq 1$ with $\mathcal{F}_t = \mathcal{F}_{[0,t]}$. Then, from the above lemma we obtain $u \in \mathrm{Dom}(\delta)$ and

$$\delta(u) = \sum_{i=1}^{n} F_i(W(t_{i+1}) - W(t_i)). \tag{1.12}$$

We know that any process $u \in L_a^2$ can be approximated in the norm of $L^2(T \times \Omega)$ by a sequence u^n of elementary adapted processes. Then by (1.12), $\delta(u^n)$ is equal to the Itô integral of u^n and it converges in $L^2(\Omega)$ to the Itô integral of u. Since δ is closed we deduce that $u \in \mathrm{Dom}(\delta)$ and $\delta(u)$ is equal to the Itô integral of u.

v ***Stochastic Integral Representation of Wiener Functionals***.

Assume that $W = \{W_t\}_{t \in [0,1]}$ is the canonical Brownian motion, it is well known that any square integrable random variable F, measurable with respect to the σ-algebra generated by W, can be written as

$$F = \mathbb{E}[F] + \int_0^1 \phi(t) dW_t,$$

where the process ϕ belongs to L_a^2. When the variable F belongs to the space $\mathbb{D}^{1,2}$, it turns out that the process ϕ can be identified as the optimal projection of the derivative of F. This is called the *Clark-Ocone* representation formula.

Proposition 4. *Let $F \in \mathbb{D}^{1,2}$, and assume that W is one dimensional Brownian motion. Then*

$$F = \mathbb{E}[F] + \int_0^1 \mathbb{E}[D_t F | \mathcal{F}_t] dW_t. \tag{1.13}$$

1.3 Differentiability of Solutions to SDEs

Now we want to state some facts about the differentiability in the sense of Malliavin of solutions to stochastic differential equations. Let us consider the following SDE,

$$X(t) = x_0 + \int_0^t B(s, X_s) ds + \sum_{j=1}^{d} \int_0^t A_j(s, X_s) dW_s^j, \tag{1.14}$$

where $A_j, B : [0,T] \times \mathbb{R}^m \to \mathbb{R}^m$, for $1 \leq j \leq d$ are measurable functions. We assume that the functions A and B satisfy the suitable conditions for the existence of solutions to (1.14). Suppose, as usual, that $(\Omega, \mathcal{F}, P)$ is the canonical probability space associated with a d-dimensional Brownian motion $\{W_t^i, t \in [0,T], 1 \leq i \leq d\}$. We will use the convention of summation over repeated indices. The next result describes the "perturbations" of the solutions to (1.14).

Proposition 5. *Let X be the solution of the stochastic differential equation (1.14), and suppose that the coefficients A_j^i, B^i are continuously differentiable functions with bounded derivatives. Then, X_t belongs to $\mathbb{D}^{1,p}$ for $p \geq 1$, $i = 1, \ldots, m$, and all $t \in [0,T]$. The derivatives satisfy the following SDE:*

$$\begin{aligned} D_r^j X_t^i &= A_j^i(r, X_r) + \int_r^t (\partial_k B)(s, X_s) D_r^j X_s^k ds \\ &+ \int_r^t (\partial_k A_l^i)(s, X_s) D_r^j X_s^k dW_s^l \end{aligned} \tag{1.15}$$

As a consequence, the next result shows in the case of an Itô process how to compute its "first variation".

Proposition 6. *Let $\{X_t\}_{t\geq 0}$ be an m-dimensional Itô process whose dynamics is driven by the stochastic differential equation*

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t,$$

where b and σ are assumed to be continuously differentiable functions with bounded derivatives. Let $\{Y_t\}_{t\geq 0}$ be the associated first variation process defined by the stochastic differential equation[1]

$$dY_t = b'(X_t)Y_t dt + \sum_{i=1}^m \sigma_i'(X_t) Y_t dW_t^i, \qquad Y(0) = I_m,$$

where I_m is the identity matrix of $\mathbb{R}^{m\times m}$. Then, the process $\{X_t\}_{t\geq 0}$ belongs to $\mathbb{D}^{1,2}$ and its Malliavin derivative is given by

$$D_s X_t = Y_t Y_s^{-1} \sigma(X_s) 1_{s\leq t}, \qquad s \geq 0, P\text{-a.e.} \tag{1.16}$$

In connection to the chain rule, Proposition 1, we have that if $\psi \in C_b^1(\mathbb{R}^m)$ then,

$$D_s\psi(X_T) = \nabla\psi(X_T) Y_T Y_s^{-1} \sigma(X_s) 1_{s\leq T}, \qquad s \geq 0, \quad P\text{-a.e.}$$

and also

$$D_s \int_0^T \psi(X_t)dt = \int_s^T \nabla\psi(X_t) Y_t Y_s^{-1} \sigma(X_s) dt, \qquad P\text{-a.e.}$$

[1] Here, primes denote derivatives and σ_i is the i-th column vector of σ.

1.4 Girsanov's Theorem.

We include a short discussion about this theorem so as to be as self contained as possible. It will be used, in Proposition 7.Ch 2 and Proposition 9.Ch 3. The result concerns the change of probability measure. As motivation, let us consider a probability space $(\Omega, \mathcal{F}, P)$ and independent normal random variables $\xi_1, \xi_2, \ldots, \xi_n$ on it. For an arbitrary vector $\mu \in \mathbb{R}^n$, introduce a new measure $\tilde{P}$ on $(\Omega, \mathcal{F})$ by

$$\tilde{P}(d\omega) = \exp\left\{ \sum_{i=1}^{n} \mu_i \xi_i(\omega) - \frac{1}{2} \sum_{i=1}^{n} \mu_i^2 \right\} \cdot P(d\omega),$$

which is actually a probability since

$$\tilde{P}(\Omega) = e^{-\frac{1}{2}\sum_{i=1}^{n} \mu_i^2} \cdot \prod_{i=1}^{n} \mathbb{E}[e^{\mu_i \xi_i}] = 1.$$

The question now is, what is the distribution of $(\xi_1, \ldots, \xi_n)$ under $\tilde{P}$?.

We have

$$\begin{aligned}\tilde{P}[\xi_1 \in dz_1, \ldots, \xi_n \in dz_n] &= \exp(\sum_{i=1}^{n} \mu_i z_i - \frac{1}{2} \sum_{i=1}^{n} \mu_i^2) \cdot P[\xi_1 \in dz_1, \ldots, \xi_n \in dz_n] \\ &= (2\pi)^{-\frac{n}{2}} \exp\left\{ - \frac{1}{2} \sum_{i=1}^{n} (z_i - \mu_i)^2 \right\} dz_1 \ldots dz_n.\end{aligned}$$

In other words, under $\tilde{P}$ the random variables $(\xi_1, \ldots, \xi_n)$ are independent, and $\xi_i \sim \mathcal{N}(\mu_i, 1)$. Equivalently, setting $\tilde{\xi}_i = \xi_i - \mu_i$, for $1 \le i \le n$, the random variables $(\tilde{\xi}_1, \ldots, \tilde{\xi}_n)$ under $\tilde{P}$ have the same law as the random variables $(\xi_1, \ldots, \xi_n)$ under P (namely independent and standard normal). The following result extends this idea to processes, and it is of great importance in stochastic analysis.

Theorem 1 (Girsanov.)**.** *Let $W = \{W_t, \mathcal{F}_t; 0 \le t \le T\}$ be d-dimensional Brownian motion, $X = \{X_t, \mathcal{F}_t; 0 \le t \le T\}$ a measurable, adapted, $\mathbb{R}^d$-valued process with $\int_0^T \|X_t\|^2 dt < \infty$, $\mathbb{P}$-a.e., and suppose that the exponential super martingale Z,*

$$Z_t = \exp\Big(\int_0^t X_s dW_s - \frac{1}{2} \int_0^t \|X_s\|^2 ds \Big),$$

is actually a martingale, i.e,

$$\mathbb{E}[Z_T] = 1. \tag{1.17}$$

Then, under the measure

$$\tilde{P}(d\omega) = Z_T(\omega) P(d\omega), \tag{1.18}$$

which is actually a probability by virtue of (1.17), the process

$$\tilde{W}_t = - \int_0^t X_s ds + W_t, \quad \mathcal{F}_t; \quad 0 \le t \le T, \tag{1.19}$$

is a d-dimensional Brownian motion.

Remark: The *Novikov* condition

$$\exp\Big(\int_0^T \|X_s\|^2 ds\Big) < \infty,$$

is sufficient for Z_t to be a martingale. So in particular, this is the case if X_t is bounded. See [Øks03, KS91, Fri75] and the references therein for more material on Girsanov's theorem.

1.5 Parabolic Operators and their Relation with Stochastic Differential Equations.

In this section, we describe a theme that will play an important role in Chapter 3. References for this section are [Øks03, KS91, Fri64]. The main goal here is to give an intuitive idea of the results that will allow us to perform some formal calculations.

As motivation let us assume we are given an *Itô diffusion* in $\mathbb{R}^n$, $\{X_t^{x,t_0}\}$ starting at point x in the space, at time $t = t_0$. Since an Itô diffusion is a Markov process, we may consider the transition densities $\psi(X, t; x, t_0)$ and compute for any Borel set B in $\mathbb{R}^n$ the probability of X_T to belong to B as

$$\Pr(X_T \in B) = \int_B \psi(\bar{x}, T; x, t)d\bar{x}.$$

Due to the Markov property, it is also possible to define a semigroup of operators $\{\mathcal{G}_t\}_{t\geq 0}$ parametrized by $t \geq 0$, that act on functions $f \in C(\mathbb{R}^n)$ as

$$\mathcal{G}_t : f \to \mathbb{E}^{x,0}[f(X_t)],$$

so all the statements in this section also have interpretations on semigroup theory.

Our main goal in this section is to show that the Markov transitions $\psi(\bar{x}, T; x, t)$ are fundamental solutions to certain parabolic equations whose coefficients are linked with those of the stochastic differential equation satisfied by the diffusion X_T and thus are related with the generator of the one parameter family semigroup mentioned above.

Here, we will follow closely the exposition and notation of [Øks03]. We refer the reader to the books [Fri75, KS91] for more details on this topic.

Definition 4. *A (time-homogeneous)* Itô diffusion *is a stochastic process*

$$X_t(\omega) = X(t, \omega) : [0, \infty) \times \Omega \longrightarrow \mathbb{R}^n$$

satisfying a stochastic differential equation of the form

$$\begin{aligned} dX_t &= b(X_t)dt + \sigma(X_t)dW_t, \; t \geq t_0, \\ X(t_0) &= x. \end{aligned} \tag{1.20}$$

Here, W_t is the m-dimensional Brownian motion and the functions

$$b : \mathbb{R}^n \longrightarrow \mathbb{R}^n, \quad \sigma : \mathbb{R}^n \longrightarrow \mathbb{R}^{n\times m},$$

satisfy the following uniform Lipschitz condition

$$\begin{aligned} &|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq D|x - y| \quad x, y \in \mathbb{R}^n, \\ &\textit{where} \;\; |\sigma|^2 = \sum |\sigma_{i,j}|^2. \end{aligned} \tag{1.21}$$

Remarks:

- The condition (1.21) is sufficient to ensure existence and uniqueness of solutions to (1.20).
- Since the functions b and σ do not depend on t, the diffusion becomes homogeneous in time, in the sense that the probability density $\psi(\bar{x}, t_2 + h; x, t_1 + h)$ for $t_2 > t_1 > 0$ and $h > 0$ does not depend on h.

Definition 5. ***The Infinitesimal Generator of a Diffusion***.

Let $\{X_t^{x,0}\}$ be a time-homogeneous Itô diffusion in $\mathbb{R}^n$. The (infinitesimal) generator *A of $\{X_t\}$ is defined for f in its domain by*

$$Af(x) = \lim_{t \downarrow 0} \frac{\mathbb{E}^{x,0}[f(X_t)] - f(x)}{t}, \quad x \in \mathbb{R}^n. \tag{1.22}$$

Here the notation $\mathbb{E}^{x,0}$ means that we take the expectation of the diffusion $\{X_t^{x,0}\}$, such that $X(0) = x$, i.e, it starts at x at time $t = 0$.

The set of functions $f : \mathbb{R}^n \to \mathbb{R}$ such that the limit exists at x is denoted by $D_A(x)$, while D_A denotes the set of functions for which the limit exists for all $x \in \mathbb{R}^n$.

Now we have the first theorem:

Theorem 2. *Let $\{X_t\}$ be the Itô diffusion*

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

If $f \in C_0^2(\mathbb{R}^n)$, then $f \in D_A$ and

$$Af(x) = \sum_i b_i(x) \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,j} (\sigma\sigma^T)_{i,j}(x) \frac{\partial^2 f}{\partial x_i \partial x_j}. \tag{1.23}$$

Here, σ^T is the transpose matrix of σ.

Remarks:

- We also assume that the matrix $\sigma\sigma^T$ is positive definite.
- The result above and the following are both strongly connected to the Itô formula for stochastic calculus.

Theorem 3. ***(Dynkin's Formula)*** *Let $f \in C_0^2(\mathbb{R}^n)$ then we have*

$$\mathbb{E}^{x,0}[f(X_t)] = f(x) + \mathbb{E}^{x,0}[\int_0^t Af(X_s)ds].$$

1.5.1. Kolmogorov's Backward Equations.

In the sequel we let X_t be an Itô diffusion with generator A. If we chose $f \in C_0^2(\mathbb{R}^n)$, then from Dynkin's formula we get that the function

$$U(T,x) = \mathbb{E}^{x,t}[f(X_T)],$$

is differentiable with respect to T and

$$\frac{\partial U}{\partial T} = \mathbb{E}^x[Af(X_T)] \tag{1.24}$$

Here, we remark that in our notation $\mathbb{E}^{x,t}[f(X_T)]$ is in fact the conditional expectation of $f(X_T)$ given that $X_t = x$ i.e.

$$\mathbb{E}^{x,t}[f(X_T)] = \mathbb{E}[f(X_T)|X_t = x].$$

The right hand side of (1.24) can be also expressed directly in terms of $U(T,x)$.

Theorem 4. ***Kolmogorov's Backward Equation****. Let $f \in C_0^2$ and define*

$$U(T,x) = \mathbb{E}^{x,t}[f(X_T)], \quad t < T.$$

Then, $U(T,\cdot) \in D_A$ for each T and

$$\begin{aligned} \frac{\partial U}{\partial T} &= AU, \\ U(t,x) &= f(x), \quad x \in \mathbb{R}^n. \end{aligned} \tag{1.25}$$

Now, let us make some remarks about the preceding theorems. At this point we will develop some formal calculations to give an interpretation of the above facts. Let us begin with Dynkin's formula

$$U(t,x) = \mathbb{E}^{x,t}[f(X_T)] = f(x) + \mathbb{E}^{x,t}[\int_t^T Af(X_s)ds], \tag{1.26}$$

Here, we are considering an Itô diffusion $\{X_s^{x,t}\}$ that satisfies $X_t = x$. Let us assume that we have the transition density of the process X_s for $t \le s \le T$ namely $\psi(\bar{X}, s; x, t)$. Substituting ψ into equation (1.26) we see that

$$\mathbb{E}^{x,t}[f(X_T)] = \int f(\bar{X})\psi(\bar{X},T;x,t)d\bar{X} = f(x) + \int_t^T \int A_{\bar{X}} f(\bar{X})\psi(\bar{X},s;x,t)d\bar{X}ds, \tag{1.27}$$

The indefinite integral being performed over the support of ψ. Thinking of ψ as the transition probability, it is desirable also that

$$\lim_{s\downarrow r \text{ or } r\uparrow s} \psi(\cdot, s; x, r) = \delta_x(\cdot) \text{ for } t \le r < s \le T, \tag{1.28}$$

holds either we fixed r and make $s \downarrow r$ or fix s and make $r \uparrow s$. Here, $\delta_x(\bar{X})$ is the Dirac's delta measure supported at x. Differentiating (1.27) with respect to T and applying integration by parts to the right hand side we obtain formally

$$\int f(\bar{X})\frac{\partial}{\partial T}\psi(\bar{X},T;x,t)d\bar{X} = \int f(\bar{X})A^*_{\bar{X}}\psi(\bar{X},T;x,t)d\bar{X}, \tag{1.29}$$

where $A^*_{\bar{x}}$ is the formal adjoint of A, that is,

$$A^*_{\bar{x}}g(\bar{x}) = -\sum_i \frac{\partial}{\partial \bar{x}_i}(b_i g) + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial \bar{x}_i \partial \bar{x}_j}((\sigma\sigma^T)_{i,j}g).$$

Since equation (1.29) holds for every $f \in C_0^2$, it follows from it and (1.28) that ψ satisfies

$$\begin{aligned} &\frac{\partial \psi}{\partial T} - A^*_{\bar{x}}\psi = 0, \\ &\lim_{T\downarrow t}\psi(\bar{x},T;x,t) = \delta_x(\bar{x}). \end{aligned} \tag{1.30}$$

This is what we refered to at the beginning of this section: The Markovian transition kernel of the diffusion X_T namely $\psi(\bar{x},T;x,t)$, is the fundamental solution of the PDE (1.30). But there is more, this kernel also satisfies an equation that involves the operator A instead of its adjoint A^* this time in the pair of variables (x,T). To see this, we use the Kolmogorov's backward equation. Rewriting Equation (1.25) in terms of ψ, it takes the form

$$\frac{\partial}{\partial T}\int f(\bar{x})\psi(\bar{x},T;x,t)d\bar{x} = A_x\int f(\bar{x})\psi(\bar{x},T;x,t)d\bar{x}.$$

Proceeding formally, for all $f \in C_0^2$ we have

$$\int f(\bar{x})\frac{\partial}{\partial T}\psi(\bar{x},T;x,t) = \int f(\bar{x})A_x\psi(\bar{x},T;x,t)d\bar{x}.$$

Therefore, $\psi(\bar{x},T;x,t)$ must also satisfy

$$\begin{aligned} &\frac{\partial \psi}{\partial T} - A_x\psi = 0, \\ &\lim_{T\downarrow t}\psi(\bar{x},T,\cdot,t) = \delta_{\bar{x}}(\cdot). \end{aligned} \tag{1.31}$$

Now, let us conclude this section describing the solution to the non-homogeneous equation

$$\begin{aligned} &\frac{\partial U}{\partial t} - AU = V(x,t), \\ &U(x,t_0) = 0. \end{aligned} \tag{1.32}$$

Here, the operator A is assumed to be uniformly elliptic of the form given in (1.23), to this we refer to the *Duhamel's principle* [Fri64].

Let us solve the final value problem

$$\frac{\partial U(x,t;s)}{\partial t} + AU(x,t;s) = 0, \quad t < s,$$

$$U(x,s,s) = V(x,s).$$

Inspired by (1.31), we see that $U(x,t;s)$ can be written as

$$U(x,t,s) = \int V(\bar{x},t)\psi(\bar{x},s;x,t)d\bar{x}.$$

Keeping this in mind let us check that the function $\widetilde{U}(x,s) = \int_{t_0}^{s} U(x,t;s)dt$ is a candidate to solve (1.32). Indeed

$$\begin{aligned}
&(\frac{\partial}{\partial s} - A_x)\int_{t_0}^{s}\int V(\bar{x},t)\psi(\bar{x},s;x,t)d\bar{x}dt, \\
&= \int_{t_0}^{s}\int V(\bar{x},t)(\frac{\partial}{\partial s} - A_x)\psi(\bar{x},s;x,t)d\bar{x}dt \\
&+ \lim_{r\uparrow s}\int V(\bar{x},r)\psi(\bar{x},s;x,r)d\bar{x}\ .
\end{aligned} \tag{1.33}$$

On the right hand side of Equation (1.33) the first term vanishes and the second equals $V(x,s)$.

A similar formula can be found using the same ideas for solutions of equations with the adjoint A^*.

CHAPTER 2

Stochastic Volatility Models and Analytic Properties

Classical Black-Scholes models [BS73], although highly successfull for their simplicity and applicability, have a number of drawbacks [FPS00]. One of such drawbacks is the fact that the *implied volatility* determined from real data is not constant. Another one is that the distribution of real asset prices presents *fat* tails. There are several competing models to try to address such shortcomings. It is fair to say that one of the most competitive ones is the class of stochastic volatility models. Let us consider an asset with price X_t at the instant t, that evolves according to the following stochastic differential equation

$$\left\{ \begin{array}{rcl} dX_t & = & \mu X_t dt + \sigma_t X_t dW_t, \\ X_0 & = & x. \end{array} \right. \tag{2.1}$$

Here, μ is the *mean return rate*, W_t is the Brownian motion, and the volatility follows the process $\{\sigma_t\}$. We will refer to the model (2.1) as the *diffusion with stochastic volatility* model for the asset price dynamics. We can see that the model (2.1) is indeed a generalization of the Black-Scholes model where the volatility is now modeled by the process σ_t which has its own stochastic dynamics. We are going to consider a *mean reverting stochastic volatility* model, it consists of modeling the volatility process σ_t of (2.1) by a function of a *mean reverting Ornstein-Uhlenbeck* process Y_t, i.e., $\sigma_t = \sigma(Y_t)$ where:

$$dY_t = \alpha(m - Y_t)dt + \beta d\widehat{W}_t. \tag{2.2}$$

Here, m, α and β are positive constants, $\{\widehat{W}_t\}$ is a Brownian motion, and $\sigma(\cdot)$ is a positive function of a real variable.

It can be shown that the law of Y_t given Y_0 is

$$\mathcal{N}\bigg(m + (Y_0 - m)e^{-\alpha t}, \frac{\beta^2}{2\alpha}(1 - e^{-2\alpha t})\bigg).$$

In other words, it is normally distributed with mean $m + (Y_0 - m)e^{-\alpha t}$ and variance $\frac{\beta^2}{2\alpha}(1 - e^{-2\alpha t})$. Therefore, m is the limit of the mean value of Y_t as $t \to +\infty$, and $1/\alpha$ is the characteristic time of mean reversion The parameter α is called the *rate of mean reversion*. The ratio $\beta^2/2\alpha$ is the limit of the variance of Y_t as $t \to +\infty$. See [FPS00] for more information on this model.

The Brownian motion $\widehat{W}_t$ may be correlated with W_t of (2.1), and in the sequel, we will write:

$$\widehat{W}_t = \rho W_t + \sqrt{1-\rho^2} W_t^1.$$

Here, W_t and W_t^1 are two independent Brownian motions, and the *correlation factor* ρ lies in $[-1, 1]$.

We will then consider the following dynamics for the asset price in the so-called *physical measure* in contrast with the *equivalent martingale measure* (e.m.m for shortly) or also called the *risk neutral measure* in the mathematical finance literature:

$$\begin{aligned} dX_t &= \mu X_t dt + \sigma(Y_t) X_t dW_t, \\ dY_t &= \alpha(m - Y_t)dt + \beta(\rho\, dW_t + \sqrt{1-\rho^2} dW_t^1). \end{aligned} \tag{2.3}$$

It turns out that in the stochastic volatility context we do not have a complete market [Sch03, FPS00]. According to the *non-arbitrage principle*[Sch03, Duf01], one has to choose an equivalent measure under which the the discounted price of the asset stochastic process $\{e^{-rt}X_t\}_{t\geq 0}$, becomes a martingale for the Brownian filter.

We are interested on the price of a *European call option* written on the asset X_t. The following modification of (2.3), gives rise in a natural way to several of such equivalent martingale measures,

$$\begin{aligned} dX_t =& rX_t dt + \sigma(Y_t) X_t dW_t^*, \quad X_0 = x, \\ dY_t =& [\alpha(m - Y_t) - \beta(\sqrt{1-\rho^2}\gamma(X_t, Y_t, t) + \rho\frac{\mu - r}{\sigma(Y_t)})]dt \\ &+ \beta(\rho dW_t^* + \sqrt{1-\rho^2} dW_t^{1*}), \quad Y_0 = y, \end{aligned} \tag{2.4}$$

where

$$\begin{aligned} dW_t^* =& \frac{r - \mu}{\sigma(Y_t)}dt + dW_t, \\ dW_t^{1*} =& \gamma(X_t, Y_t, t)dt + dW_t^1. \end{aligned}$$

System (2.4) describes the dynamics of X_t in the risk neutral measure, here r is the *constant interest rate* and the function $\gamma(x, y, t)$ accounts for the so called *risk premium factor* from the second source of randomness $\widehat{W}_t$ that drives the volatility. Notice that its introduction in the equation for Y_t does not change the drift of X_t.

We remark that in the model under consideration we do not have a unique equivalent martingale measure as in the case of the Black-Scholes model. Thus, option pricing becomes more difficult. In the model (2.4), the non-uniqueness of martingale measures is implicit in the fact that we need to choose a function $\gamma(x, y, t)$, and every such function gives rise to a particular equivalent martingale measure.
It is known [FPS00] that the price of an European Call option written on the asset X_t that follows the stochastic volatility dynamics of (2.4), is given by the solution $U(x, y, t)$ of the following *final* value problem:

$$\begin{aligned}
&U_t + \frac{\sigma^2(y)x^2}{2}U_{xx} + \frac{\beta^2}{2}U_{yy} + \rho\beta x\sigma(y)U_{xy} + \\
&(\alpha(m-y) - \beta\Lambda(y))U_y - r(xU_x - U) = 0 \ , \\
&U(x, y, T) = (x-K)^+ \ .
\end{aligned} \tag{2.5}$$

Here,

$$\Lambda(x, y, t) = \rho\frac{\mu - r}{\sigma(y)} + \sqrt{1-\rho^2}\gamma(x, y, t), \text{and} \quad 0 \le t < T, x > 0, y \in \mathbb{R}.$$

Now we are ready to state, at least informally, our main goal. In principle the ideal problem we would like to study is that of determining the function $\gamma(x, y, t)$ from data of the function $U(x, y, t)$. This is important because in many applications it is necessary to price other derivatives, such as exotic ones, in a way that is consistent with the ones that are available in the market. Actually, in this work we will focus on a particular case of this problem, namely, we are going to deal with the uncorrelated case. This means that $\rho = 0$. Furthermore, we will assume that γ depends only on y. See [SZ06] for some comments on this hypothesis.

Thus, our basic equations for asset dynamics in the risk neutral measure Q_γ become:

$$\begin{aligned}
dX_t &= rX_t dt + \sigma(Y_t)X_t dW_t^*, \quad X_0 = x, \\
dY_t &= (\alpha(m - Y_t) - \beta\gamma(Y_t))dt + \beta dW_t^{1*}, \quad Y_0 = y,
\end{aligned} \tag{2.6}$$

The price $U(x, y, t)$ of a European call option written on X_t, is

$$U_t + \frac{\sigma^2(y)x^2}{2}U_{xx} + \frac{\beta^2}{2}U_{yy} + (\alpha(m-y) - \beta\gamma(y))U_y - r(xU_x - U) = 0. \tag{2.7}$$

$$U(x, y, T) = (x-K)^+; \quad 0 \le t < T, \quad x > 0, \quad y \in \mathbb{R}. \tag{2.8}$$

The characterization of the function $\sigma(y)$ is another very important issue and has been the subject of extensive analysis [Dup97]. Here, we assume that it is given by the model of Stein and Stein [SS91] whereby $\sigma(y) = |y|$. Many of our ideas can be applied to other models as well.

2.1 Statement of the Problem.

First, we will assume that the function $\gamma(y)$ is Lipschitz continuous and that Equation (2.6) has a unique strong solution that is well behaved [Øks03]. The assumption on γ is a rather delicate point, indeed it should be investigated by appropriate financial arguments which properties the function γ must have in order to make the model consistent. We will not delve into this question here and rather we will try to leave the model as general as possible. For example, we would like the process Y_t to have moments of all orders bounded in the time interval under consideration. We are going to consider the following operator defined formally.

$$\Gamma : \mathcal{C}(\mathbb{R}) \cap A \longrightarrow \mathcal{C}(\mathbb{R}^+, \mathbb{R}, [0,T]) \cap B.$$

Here, A and B are spaces that will be specified later. The operator Γ is given by:

$$\Gamma(\gamma)(x,y,t) = U(x,y,t) = e^{-r(T-t)}\mathbb{E}^{Q_\gamma}[(X_T - K)^+ | X_t = x, Y_t = y]. \tag{2.9}$$

So, the operator Γ takes the coefficient $\gamma(y)$ and maps it to $U(x,y,t)$, the corresponding solution of (2.8) where $\gamma(y)$ is taken as a coefficient in the same equation. The last equality expresses the well known fact of the function $U(x,y,t)$ being the discounted expected value of the call payoff $(x-K)^+$ evaluated on X_T under a risk neutral measure Q_γ associated to the dynamics (2.6).

The inverse problem we are going to consider is then, that of inverting Γ or recovering as much information as posible about γ from $\Gamma(\gamma)$. In other words, we assume the function $U(x,y,t)$, of price data to be given, and we will try to recover a *suitable* coefficient $\tilde{\gamma}(y)$ for (2.8) such that $\Gamma(\tilde{\gamma})(x,y,t) = U(x,y,t)$. An important point concerning the inverse problem under consideration is the fact that in practice option prices are quoted according to their maturity T, and strike price K, and the values of X_t and Y_t would be known at a time t before the expiration date. However, in our approach we are considering the variables (x,y,t). We will exploit the fact that the option price function $U(x,y,t,K,T)$ depends on x and K, through their ratio $\frac{x}{K}$. By that we mean the following: (x_1, K_1) and (x_2, K_2) are two pairs such that

$$\frac{x_1}{K_1} = \frac{x_2}{K_2},$$

then

$$\frac{U(x_1,y,t,K_1,T)}{K_1} = \frac{U(x_2,y,t,K_2,T)}{K_2}\ .$$

This symmetry of the function U can be verified with the help of Eqs. (2.27), (2.33) and the *Hull-White* formula in Section 2.3 and is natural *assumption* from a dimensional point of view.

Furthermore, since the equation is homogeneous in time, then U depends on T and t through their difference $\tau = T - t$. Summing up, we have the following:

Remark 1. *The solution U to the final value problem (2.5) satisfies*

$$U(x,y,t,K,T) = KG(\frac{x}{K}, y, \tau), \text{ for some function } G.$$

These properties of U allow us to translate data given by quoted prices of (K,T) into values of (x,t). In fact, this procedure is equivalent to the standard practice of displaying the implied volatility as a function of the so-called moneyness ratio X/K.

The above remark leads us to formulate the following theoretical version of the calibration problem:

Inverse Problem 1. *Given the value of $U(x,y_0,t,K,T)$ for all values of x and a given y_0 determine the* risk premium γ.

We say this is a theoretical version because in practice we will never have a continuum of data.

The approach we will suggest to invert Γ in a regularized way will be trough the iterative method of *Landweber* [BL05, Sch95]. It is defined by the iteration

$$\gamma^{k+1}(y) = \gamma^k(y) - \frac{\partial\Gamma}{\partial\gamma}^*(\gamma^k)[\bar{U} - \Gamma(\gamma^k)](y), \tag{2.10}$$

where the linear operator

$$\frac{\partial\Gamma}{\partial\gamma}^*(\gamma^k) : \mathcal{C}(\mathbb{R}^+, \mathbb{R}, [0,T]) \cap B \longrightarrow \mathcal{C}(\mathbb{R}) \cap A$$

is the adjoint of the *Fréchet* derivative operator Γ, namely

$$\frac{\partial\Gamma}{\partial\gamma}(\gamma^k) : \mathcal{C}(\mathbb{R}) \cap A \longrightarrow \mathcal{C}(\mathbb{R}^+, \mathbb{R}, [0,T]) \cap B,$$

and $\bar{U} = \bar{U}(x, y_0, t, K, T)$ is the given data.

2.2 The Analyticity of Γ and $\frac{\partial\Gamma}{\partial\gamma}$.

In this section we will discus some theoretical estimates that grant the existence of $\frac{\partial\Gamma}{\partial\gamma}$ and the analyticity of Γ *for the case* $\sigma(y) = |y|$ in (2.5) for suitable function spaces.

Here, we will follow the recent work of Achdou, Franchi and Tchou [AFT05], concerning the analysis of the Equation (2.5). We will also make use of some classical results of Kato following [Paz83]. Among other things, in [AFT05] the authors give some estimates concerning the semigroup related with equation (2.5). In a previous work Achdou and Tchou [AT02], it was shown that it is convenient to use the following change of variables: Take a bounded solution U of (2.5) and consider the function

$$u(x,y,t) = e^{-(1-\eta)\frac{(y-m)^2}{2\nu^2}} U(x,y,t).$$

Here, $\eta \in (0,1)$ and $\nu^2 = \frac{\beta^2}{2\alpha}$. In terms of u, Equation (2.5) for $\sigma(y) = |y|$ takes the form

$$\frac{\partial u}{\partial t} - A_t u = 0, \tag{2.11}$$

with

$$\begin{aligned} A_t v &:= \frac{1}{2} y^2 x^2 \frac{\partial^2 v}{\partial x^2} + \frac{\beta^2}{2} \frac{\partial^2 v}{\partial y^2} + rx \frac{\partial v}{\partial x} + ((1-2\eta)\alpha(y-m) - \beta\gamma(x,y,t)) \frac{\partial v}{\partial y} \\ &- \left(r + 2\frac{\alpha^2}{\beta^2}\eta(1-\eta)(y-m)^2 + 2(1-\eta)\frac{\alpha}{\beta}(y-m)\gamma(x,y,t) - \alpha(1-\eta) \right) v. \end{aligned} \tag{2.12}$$

Let us denote by $\mathcal{Q}$ the half plane $\mathbb{R}^+ \times \mathbb{R}$. The domain of the operator A_t is defined by

$$D_t = \{v \in V : A_t v \in L^2(\mathcal{Q})\},$$

where the weighted Sovolev space V is

$$V := \left\{ v : \left(\sqrt{1+y^2} v, \frac{\partial v}{\partial y}, xy\frac{\partial v}{\partial x} \right) \in (L^2(\mathcal{Q}))^3 \right\}$$

and is endowed with the norm

$$\|v\|_V = \left(\int_{\mathcal{Q}} (1+y^2)|v|^2 + |\frac{\partial v}{\partial y}|^2 + x^2 y^2 |\frac{\partial v}{\partial x}|^2 \right)^{\frac{1}{2}}.$$

Clearly V is a Hilbert space. Let us write down three results from [AFT05] concerning the initial value problem

$$\left\{ \begin{array}{lll} \frac{d}{dt} u & = & A_t u \quad 0 < t < T, \\ u(t=0) & = & u_0 \ . \end{array} \right. \tag{2.13}$$

Theorem 5 ([AFT05])**.** *The domain of A_t does not depend on t, i.e., there exists $D = D_t$ for all t. If $\alpha^2/\beta^2 > 2$, then for suitable values of η (in particular such that $2\frac{\alpha^2}{\beta^2}\eta(1-\eta) > 1$),*

$$D = \left\{ v \in V : y^2 x^2 \frac{\partial^2 v}{\partial x^2}, \frac{\partial^2 v}{\partial y^2}, yx\frac{\partial^2 v}{\partial x \partial y}, x\frac{\partial v}{\partial x}, y\frac{\partial v}{\partial y}, y^2 v \in L^2(\mathcal{Q}) \right\}. \tag{2.14}$$

Theorem 6 ([AFT05])**.** *Assume that $\alpha > \beta$ for a suitable η (*see the details in the paper [AFT05]*). There exists a unique u in $L^2((0,T);V) \cap C^0([0,T];L^2(\mathcal{Q}))$, with $\frac{\partial u}{\partial t} \in L^2(0,T;V')$ such that in the sense of distributions in time,*

$$\left\{ \begin{array}{lll} \frac{d}{dt}(u,v) & = & \langle A_t u, v \rangle, \ \forall v \in V, \\ u(t=0) & = & u_0. \end{array} \right. \tag{2.15}$$

The mapping $u_0 \mapsto u$ is continuous from $L^2(\mathcal{Q})$ to $L^2([0,T];V) \cap C^0([0,T];L^2(\mathcal{Q}))$.

Theorem 7 ([AFT05]). *Assume that there exists $\xi \in (0,1]$, such that γ belongs to $C^\xi([0,T], L^\infty(\mathcal{Q}))$ and that $\frac{\alpha^2}{\beta^2} > 2$. Then, for each suitable η, if u_0 belongs to D defined by (2.14), then the solution of (2.15) given by Theorem 6, belongs also to $C^1((0,T); L^2(\mathcal{Q})) \cap C^0([0,T]; D)$ and satisfies the equation*

$$\frac{d}{dt}u - A_t u = 0,$$

for each $t \in [0,T]$. Furthermore, for $u_0 \in L^2(\mathcal{Q})$, the weak solution of (2.15) given by Theorem 6, belongs also to $C^1((\tau,T); L^2(\mathcal{Q})) \cap C^0([\tau,T]; D)$, for all $\tau > 0$ and we have that

$$\|\frac{du}{dt}(t)\|_{L^2(\mathcal{Q})} + \|A_t u(t)\|_{L^2(\mathcal{Q})} \leq \frac{C}{t} \ ,$$

for $t > 0$.

In [AFT05], the next three statements concerning the operator A_t are proved. Thus, one can apply Kato's theorem [Paz83]. This in turn implies Theorem 7.

Let us consider a family of operators $A(t) : D(A(t)) \subset X \to X$ where X is a Banach space and $0 \leq t \leq T$ satisfying the following properties:

i The domain $D(A_t) = D$ of $A(t)$ is dense in X and independent of t.

ii For $t \in [0,T]$ the resolvent $R(\lambda : A(t))$ exists for all λ with $\mathrm{Re}\lambda \leq 0$ and there is a constant M such that

$$\|R(\lambda : A(t))\| \leq \frac{M}{|\lambda| + 1} \quad \text{for} \quad \mathrm{Re}\lambda \leq 0, t \in [0,T]. \tag{2.16}$$

iii There exist constants L and $0 < \delta \leq 1$, such that

$$\|(A(t) - A(s))A(\tau)^{-1}\| \leq L|t-s|^\delta \quad \text{for } s,t,\tau \in [0,T]. \tag{2.17}$$

The next result remains valid if instead of considering the interval $[0,T]$ one considers the interval $[a,b]$, $0 \leq a < b < \infty$.

Theorem 8 (Kato). *Under the Assumptions 1-3 above, there is a unique evolution system $U(t,s)$ on $0 \leq s \leq t \leq T$, satisfying:*

i

$$\|U(t,s)\| \leq C \quad \textit{for } 0 \leq s \leq t \leq T,$$

ii For $0 \leq s \leq t \leq T$, $U(t,s) : X \to D$ and $t \to U(t,s)$ is strongly differentiable in X. The derivative $(\partial/\partial t)U(t,s) \in B(X)$ and it is strongly continuous on $0 \leq s \leq t \leq T$. Moreover,

$$\frac{\partial}{\partial t}U(t,s) + A(t)U(t,s) = 0, \qquad \textit{for } 0 \leq s \leq t \leq T,$$

$$\|\frac{\partial}{\partial t}U(t,s)\| = \|A(t)U(t,s)\| \leq \frac{C}{t-s},$$

and

$$\|A(t)U(t,s)A(s)^{-1}\| \le C, \quad \textit{for } 0 \le s \le t \le T.$$

iii For every $v \in D$ and $t \in (0,T)$, $U(t,s)v$ is differentiable with respect to s on $0 \le s \le t \le T$ and

$$\frac{\partial}{\partial s}U(t,s) = U(t,s)A(s)v.$$

Remark. Let us remark that the Assumption 2 and the fact that D is dense in X imply that for every $t \in [0,T]$, $A(t)$ is the infinitesimal generator of an analytic semigroup $S_t(s)$, $s \ge 0$, satisfying

$$\begin{aligned} \|S_t(s)\| &\le C \qquad \text{for } \ s \ge 0, \\ \|A(t)S_t(s)\| &\le \frac{C}{s} \qquad \text{for } \ s \ge 0, \end{aligned}$$

for a suitable constant C.

Now, let us see then how the above results imply the differentiability of $\Gamma(\gamma)$, in appropriate spaces. Here, and in the sequel, we consider the case of γ in (2.12) dependent only on y. In such case, A_t does not depend on t. Let us denote by $A_{\gamma+\epsilon h}$, the operator obtained making the substitution of $\gamma(y)$ by $\gamma(y)+\epsilon h(y)$, in the formula that defines A_t in Equation (2.12).

Let us also use the notation

$$A_{\gamma+\epsilon h} = A_\gamma + B_{\epsilon h} \ .$$

We write $\bar{u} = u + \delta u$ as the solution to the problem

$$\begin{aligned} \frac{d}{dt}\bar{u} - A_{\gamma+\epsilon h}\bar{u} &= 0, \quad \text{for } t_0 \le t \le T, \\ \bar{u}(T) &= (x-K)^+. \end{aligned}$$

Here u is the solution of

$$\begin{aligned} \frac{d}{dt}u - A_\gamma u &= 0, \quad \text{for } t_0 \le t \le T, \\ u(T) &= (x-K)^+. \end{aligned}$$

Thus, we can see that

$$\left(\frac{d}{dt} - A_{\gamma+\epsilon h}\right)(u+\delta u) = \left(\frac{d}{dt} - A_\gamma\right)\delta u + B_{\epsilon h}u + B_{\epsilon h}\delta u,$$

so the difference δu satisfies

$$\begin{aligned}\left(\frac{d}{dt} - A_\gamma\right)\delta u - B_{\epsilon h}\delta u &= B_{\epsilon h} u, \\ \delta u(T) = 0,\end{aligned} \tag{2.18}$$

which is the same as

$$\begin{aligned}\frac{\partial}{\partial t}\delta u + A_\gamma \delta u &= -\epsilon\beta h(y)u_y - 2\epsilon\frac{\alpha}{\beta}h(y)(1-\eta)(y-m)u \\ &\quad -\epsilon h(y)\beta\delta u_y - 2\epsilon\frac{\alpha}{\beta}h(y)(1-\eta)(y-m)\delta u, \\ &= -\epsilon h(y)\left[\beta u_y + 2\frac{\alpha}{\beta}(1-\eta)(y-m)u + \beta\delta u_y + 2\frac{\alpha}{\beta}(1-\eta)(y-m)\delta u\right].\end{aligned} \tag{2.19}$$

So, thanks to the work on [AFT05], we can consider the semigroup $S(t)$ associated to A_γ,

$$S(t) : L^2(\mathcal{Q}) \to D, \quad D = \left\{ v \in V : A_\gamma v \in L^2(\mathcal{Q})\right\},$$

and we can integrate Equation (2.19) to get

$$\begin{aligned}\delta u[h](t) &= -\int_{t_0}^{t} S(t-r)h(y)[\beta u_y + 2\frac{\alpha}{\beta}(1-\eta)(y-m)u](r)dr \\ &\quad -\int_{t_0}^{T} S(t-r)[\beta\delta u_y + 2\frac{\alpha}{\beta}(1-\eta)(y-m)\delta u](r)dr.\end{aligned} \tag{2.20}$$

The results also grants that $S(t)$ is uniformly bounded for t in $[t_0, T]$ by a constant C, therefore the linear operator δu, is bounded and,

$$\begin{aligned}\|\delta u[h]\|_V &\leq \epsilon C(T-t_0)\|h\|_{L^\infty(\mathbb{R})}\|u_y + 2\frac{\alpha}{\beta}(1-\eta)(y-m)u\|_{L^2(\mathcal{Q})}, \\ &\quad +\epsilon C(T-t_0)\|h\|_{L^\infty(\mathbb{R})}\|\beta\delta u_y + 2\frac{\alpha}{\beta}(1-\eta)(y-m)\delta u\|_{L^2(\mathcal{Q})}, \\ &\quad \leq \epsilon C(T-t_0)\|h\|_{L^\infty}\|u\|_{H^1(\mathcal{Q})} + \epsilon C_1\|\delta u\|_V.\end{aligned} \tag{2.21}$$

Then taking ϵ small enough we have

$$\|\delta u\|_V \leq \frac{\epsilon}{1-\epsilon C_1}C(T-t_0)\|h\|_{L^\infty}\|u\|_{H^1(\mathcal{Q})}. \tag{2.22}$$

This proves that $\Gamma : L^\infty(R) \to V$ is differentiable. In fact, using an idea that can be traced back at least to the ground breaking work of Calderon [Cal80], we can show that Γ is analytic. Let us denote by $\mathcal{L}(D)$ the algebra of bounded linear operators on D.

Theorem 9. *Assume that $\alpha^2/\beta^2 > 2$ and $T > t_0$. Then, $\Gamma : L^\infty(\mathbb{R}) \to \mathcal{L}(D)$ is analytic in the sense that it can be written as a convergent series of multilinear operators .*

Proof: Due to the existence of the semigroup S_t associated with the operator A_γ. We can consider the map $\mathcal{G}_t : C^0([0,t], D) \to D$ that sends $v \in C^0([0,t], D)$ into a solution u to the initial value problem

$$\begin{cases} \frac{d}{dt}u - A_t u = v(t) \\ u|_{t=0} = 0. \end{cases} \tag{2.23}$$

$\mathcal{G}_t$ is bounded because

$$\mathcal{G}_t v = \int_0^t S(t-s)v(s)ds.$$

So,

$$\|\mathcal{G}_t v\|_V \leq C(t - t_0)\|v\|_{\mathcal{L}^2(\mathcal{Q})}.$$

Thus, applying $\mathcal{G}_t$ to (2.18) we obtain

$$\delta u - \mathcal{G}_t B_{\epsilon h}\delta u = \mathcal{G}_t B_{\epsilon h} u, \tag{2.24}$$

since for $w \in V$ we have

$$B_{\epsilon h} w = -\epsilon h(y)(\beta w_y + 2\frac{\alpha}{\beta}(1-\eta)(m-y)w), \tag{2.25}$$

then we can estimate that,

$$\|B_{\epsilon h} w\|_{L^2(\mathcal{Q})} \leq \epsilon C\|w\|_V \|h\|_{L^\infty}.$$

Therefore, taking ϵ small enough, for example such that,

$$\|\mathcal{G}_t B_{\epsilon h}\| < 1,$$

so we can consider the inverse of $I - \mathcal{G}_t B_{\epsilon h}$, to obtain

$$\delta u = \sum_{j=0}^{\infty} (\mathcal{G}_t B_{\epsilon h})^j (\mathcal{G}_t B_{\epsilon h})[u]. \tag{2.26}$$

We finally note that u does not depent on h. Thus looking at (2.25), we see that in (2.26) we have δu as a power series of the linear operator,

$$\mathcal{G}_t[-\epsilon(\cdot)(\beta u_y + 2\frac{\alpha}{\beta}(1-\eta)(m-y)u)],$$

applied to h. $\square$

2.3 The computation of $\partial\Gamma/\partial\gamma$.

Our main result here is the derivation of a fairly explicit formula for the operator $\frac{\partial\Gamma}{\partial\gamma}^*(\gamma)$. This would allow an implementation of *Landweber's* method as a route to attack our callibration problem. We are going to use two different techniques to carry out such computations. The first one is based on *Malliavin Calculus* introduced in Chapter 1. The second will be a more classical derivation dealing directly with Equation (2.8) and making use of a perturbation argument.

The first calculation assumes that the functions involved are well behaved so that all the steps in the calculation are permitted.

We start by introducing some notation. We denote by U_{BS} the solution to the classical Black-Scholes model whereby the volatility $\sigma(y)$, is considered constant. More precisely, for $S = \sigma^2$ and $\tau = T - t$ we have

$$U_{\mathrm{BS}}[S](x,\tau;K,r) = U_{\mathrm{BS}}(x,K,r,\tau,S) = \begin{cases} x\Phi(d_1) - Ke^{-r\tau}\Phi(d_2) & (S>0) \\ \max(x - Ke^{-r\tau}, 0) & (S=0), \end{cases} \tag{2.27}$$

with

$$d_1 = \frac{\ln\left(\frac{x}{K}\right) + r\tau + \frac{S}{2}}{\sqrt{S}}, \qquad d_2 = d_1 - \sqrt{S},$$

and Φ is the cumulative density function of the normal distribution, i.e,

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{x^2}{2}} dx.$$

Proposition 7. *The operator $\partial\Gamma/\partial\gamma$ is given by*

$$\begin{aligned} &\tfrac{\partial\Gamma}{\partial\gamma}(\gamma)[h](y) = \\ &-\mathbb{E}^{Q_\gamma}\left[\tfrac{\partial U_{\mathrm{BS}}}{\partial S}(\xi_T) \textstyle\int_t^T\left(\int_s^T 2\sigma(Y_v)\sigma'(Y_v)e^{\int_s^v f'(Y_u)du} dv\right) h(Y_s)ds \,\middle|\, Y_t = y\right] , \end{aligned} \tag{2.28}$$

where

$$\begin{aligned} \xi_T &= \int_t^T \sigma^2(Y_s)ds \quad \text{and} \\ f(y) &= \alpha(m-y) - \beta\gamma(y) \ . \end{aligned}$$

Proof. Let us assume that the dynamics of the stock prices X_t and the volatility process Y_t in the risk neutral measure follow the following system of SDEs.

$$\begin{aligned} dX_t &= rX_t dt + \sigma(Y_t)X_t dW_t^1, \\ dY_t &= f(Y_t)dt + \beta dW_t^*. \end{aligned}$$

Let us also denote

$$
\begin{aligned}
dX_t^\epsilon &= rX_t^\epsilon dt + \sigma(Y_t)X_t^\epsilon dW_t^1, \quad X_0^\epsilon = x, \\
dY_t &= (f(Y_t) + \epsilon h(Y_t))dt + \beta dW_t^\epsilon, \quad Y_0 = y,
\end{aligned} \tag{2.29}
$$

where the process

$$
dW_t^\epsilon = -\frac{\epsilon}{\beta}h(Y_t)dt + dW_t^*,
$$

is by *Girsanov's* theorem Brownian motion under the probability measure

$$
dQ^\epsilon = \exp\left(\frac{\epsilon}{\beta}\int_t^T h(Y_s)dW_s^* - \frac{\epsilon^2}{2\beta^2}\int_t^T h(Y_s)^2 ds\right)dQ_\gamma. \tag{2.30}
$$

Then, let us consider

$$
U^\epsilon(x,y,t) = e^{-r(T-t)}\mathbb{E}^{Q^\epsilon}\left[(X_T^\epsilon - K)^+ \middle| X_t^\epsilon = x, Y_t = y\right], \tag{2.31}
$$

We can see that the last expectation is taken under Q^ϵ, which is the measure associated with (2.29) and note that when $\epsilon = 0$, we recover (2.9). It can be shown by Itô calculus that

$$
X_T^\epsilon = x\exp\left(\int_t^T \left[r - \frac{\sigma^2(Y_s)}{2}\right]ds + \int_t^T \sigma(Y_s)dW_s^1\right). \tag{2.32}
$$

We use this expression to find a way of conditioning the expectation in (2.31), on the whole path trajectory of Y_s, with $t \le s \le T$, getting by means of this trick the *Hull-White* formula

$$
U^\epsilon(x,y,t,K,T) = \mathbb{E}^{Q^\epsilon}\left[U_{BS}[\xi_T]\right], \tag{2.33}
$$

where U_{BS} is given by Equation (2.27) and we are leaving all the arguments but ξ_T implicit. The variable ξ_T is given by

$$
\xi_s := \int_t^s \sigma^2(Y_v)dv .
$$

Now, we will get rid off the dependence of the measure in (2.31) on the parameter ϵ applying Girsanov's theorem. This yields

$$
\begin{aligned}
&\mathbb{E}^{Q^\epsilon}\left[U_{\text{BS}}(\textstyle\int_t^T \sigma^2(Y_s)ds)\right] = \\
&\mathbb{E}^{Q_\gamma}\left[e^{(\frac{\epsilon}{\beta}\int_t^T h(Y_s)dW_s^* - \frac{\epsilon^2}{2\beta^2}\int_t^T h^2(Y_s)ds)}U_{\text{BS}}(\textstyle\int_t^T \sigma^2(Y_s)ds)\right].
\end{aligned} \tag{2.34}
$$

In what follows, all the expected values will be assumed to be evaluated under the measure Q_γ. So we can write

$$
\begin{aligned}
&\tfrac{1}{\epsilon}(U^\epsilon - U)_{(x,y,t)} = \\
&\tfrac{1}{\epsilon}\mathbb{E}^{Q_\gamma}\left[(e^{\{\frac{\epsilon}{\beta}\int_t^T h(Y_s)dW_s^* - \frac{\epsilon^2}{2\beta^2}\int_t^T h^2(Y_s)ds\}} - 1)U_{\text{BS}}\big(\textstyle\int_t^T \sigma^2(Y_s)ds\big)\right].
\end{aligned}
$$

Taking the limit when $\epsilon \to 0$ we get

$$\frac{1}{\beta}\mathbb{E}^{\mathbb{Q}_\gamma}\Big[(\int_t^T h(Y_s)dW_s^*)U_{\mathrm{BS}}(\int_t^T \sigma^2(Y_s)ds)\Big]. \tag{2.35}$$

Now we use Malliavin calculus, in particular the fact that the Malliavin derivative is the adjoint of the stochastic integral, to write the last expression as

$$\frac{1}{\beta}\mathbb{E}^{\mathcal{Q}_\gamma}\Big[\int_t^T D_s^*\big[U_{\mathrm{BS}}(\int_t^T \sigma^2(Y_u)du)\big]h(Y_s)ds\Big], \tag{2.36}$$

with D_s^* being the Malliavin derivative with respect to the Brownian motion W_t^*. We will continue performing the derivations on (2.36), taking into account that we are working with the measure Q_γ, for which the dynamics of Y_t takes the form

$$dY_t = (\alpha(m - Y_t) - \beta\gamma(Y_t))dt + \beta dW_t^* = f(Y_t)dt + \beta dW_t^*. \tag{2.37}$$

So,

$$\begin{aligned}&\tfrac{1}{\beta}\mathbb{E}^{\mathbb{Q}_\gamma}\Big[\textstyle\int_t^T D_s^*\big[U_{\mathrm{BS}}(\int_t^T \sigma^2(Y_u)du)\big]h(Y_s)ds\Big] = \\ &\tfrac{1}{\beta}\mathbb{E}^{\mathbb{Q}_\gamma}\Big[\tfrac{\partial U_{\mathrm{BS}}}{\partial S}(\textstyle\int_t^T \sigma^2(Y_u)du)\int_t^T\Big(\int_s^T 2\sigma(Y_v)\sigma'(Y_v)D_s^*Y_v dv\Big)h(Y_s)ds\Big]\,.\end{aligned} \tag{2.38}$$

Using that

$$d(D_s^*Y_v) = f'(Y_v)D_s^*Y_v dv, \quad \text{so} \quad D_s^*Y_v = \beta e^{(\int_s^v f'(Y_u)du)}1_{\{s\le v\}}, \tag{2.39}$$

(C.f. Proposition 6 of Chapter 2), and substituting in (2.38), we finally get our desired formula (2.28). □Let us now to write formula (2.28) in a more comfortable form. Let us introduce the following notation

$$\begin{aligned}\eta_s &:= 2\int_t^s \sigma(Y_v)\sigma'(Y_v)e^{\theta_v}dv. \\ \theta_s &:= \int_t^s f'(Y_v)dv,\end{aligned} \tag{2.40}$$

and as before we have

$$\xi_s = \int_t^s \sigma^2(Y_v)dv.$$

Thus, we can see that (2.28), or equivalently (2.38), becomes

$$\begin{aligned}&\mathbb{E}^{\mathbb{Q}_\gamma}\Big[\tfrac{\partial U_{\mathrm{BS}}}{\partial S}(\xi_T)\textstyle\int_t^T e^{-\theta_s}(\eta_T - \eta_s)h(Y_s)ds\Big] = \\ &\textstyle\iint_t^T \tfrac{\partial U_{\mathrm{BS}}}{\partial S}(\xi_T)e^{-\theta_s}(\eta_T - \eta_s)h(Y_s)\Psi(\xi_T, \eta_T, T; \theta_s, \eta_s, Y_s, s; y, t)d\theta_s d\eta_s dY_s ds d\eta_T d\xi_T.\end{aligned} \tag{2.41}$$

Here, the function $\Psi(\xi_T, \eta_T, T; \theta_s, \eta_s, Y_s, s; y, t)$, is the joint probability distribution of the functionals in (2.40) and Y_s in the considered times. It is clear the dependence of Ψ on $Y_t = y$ at time t, since all the above functionals depend on the whole trajectory of Y_s.

An interesting point concerning (2.41) is that it can be used to carry out Monte Carlo simulations. Later on we will return to expression (2.41) to get an expression for $\partial\Gamma^*/\partial\gamma$ so as to connect these ideas with Landweber's method and talk a little bit about numerical calculations.

We will continue with the derivation of another formula for $\partial\Gamma/\partial\gamma$ however different, via classical calculations dealing directly with the partial differential equation (2.8).

Naturally associated to equation (2.8), we have a function $\phi(\bar{x}, \bar{y}, s; x, y, t)$ defined for $t \leq s$ as the density function of the homogeneous diffusion (X_s, Y_s) whose dynamics is driven by

$$\begin{aligned} dX_t &= rX_t dt + \sigma(Y_t) X_t dW_t^1, \quad X_0 = x, \\ dY_t &= (\alpha(m - Y_t) - \beta\gamma(Y_t))dt + \beta dW_t^*, \quad Y_0 = y. \end{aligned}$$

So, we interpret $\phi(\bar{x}, \bar{y}, s; x, y, t) d\bar{x} d\bar{y}$ as the probability density

$$\Pr\left[X_s \in (\bar{x}, \bar{x} + dx), Y_s \in (\bar{y}, \bar{y} + dy) | X_t = x, Y_t = y\right],$$

Yet another view of ϕ is that of the Markov transition probability of the diffusion (X_s, Y_s), the same interpretation holds for Ψ in (2.41) considering the diffusions $\xi_s, \eta_s, \theta_s, Y_s$. Let us denote the differential operator in (2.8) by $\mathcal{L}$, i.e.,

$$\mathcal{L} := \frac{\partial}{\partial t} + \frac{x^2\sigma^2(y)}{2}\frac{\partial^2}{\partial x^2} + \frac{\beta^2}{2}\frac{\partial^2}{\partial y^2} + rx\frac{\partial}{\partial x} + (m(\alpha - y) - \beta\gamma(y))\frac{\partial}{\partial y} \tag{2.42}$$

The formal adjoint of $\mathcal{L}$ is

$$\begin{aligned} &\mathcal{L}^* U := \\ &\tfrac{\partial^2}{\partial x^2}(\tfrac{x^2\sigma^2(y)}{2}U) + \tfrac{\beta^2}{2}\tfrac{\partial^2 U}{\partial y^2} - r\tfrac{\partial}{\partial x}(xU) - \tfrac{\partial}{\partial y}((m(\alpha - y) - \beta\gamma(y))U). \end{aligned} \tag{2.43}$$

Taking into account that $\mathcal{L}$ is the generator of the diffusion (X_s, Y_s), it is a well known result that under some regularity conditions on the coefficients $\sigma(y)$, $\gamma(y)$, then the function $\phi(\bar{x}, \bar{y}, s; x, y, t)$ is just the fundamental solution in PDE terminology of both (2.42) and (2.43). See Section 1.5 of Chapter 1. Therefore, they satisfy the following relations:

$$\begin{aligned} (\frac{\partial}{\partial t} + \mathcal{L})_{(x,y,t)}[\phi(\bar{x}, \bar{y}, s; x, y, t)] &= 0, \\ \lim_{t \to s^-} \phi(\bar{x}, \bar{y}, s; x, y, t) = \delta_{(\bar{x},\bar{y})}(x, y). \end{aligned} \tag{2.44}$$

$$\begin{aligned} (-\frac{\partial}{\partial s} + \mathcal{L}^*)_{(\bar{x},\bar{y},s)}[\phi(\bar{x}, \bar{y}, s; x, y, t)] &= 0 \\ \lim_{s \to t^+} \phi(\bar{x}, \bar{y}, s; x, y, t) = \delta_{(x,y)}(\bar{x}, \bar{y}), \end{aligned} \tag{2.45}$$

for all $s > t$ where $\delta_{(u,v)}(x,y)$ is the Dirac delta in the variables (x,y).

We are now ready to state and prove the following result:

Proposition 8 (An alternative expression for $\partial\Gamma/\partial\gamma$.). *Under sufficiently regular coefficients $\sigma(y)$ and $\gamma(y)$, we have*

$$\begin{array}{l} \frac{\partial \Gamma}{\partial \gamma}(\gamma)[h(s)] = \\ e^{-r(T-t)} \iint_t^T (u-K)^+ \phi_{\bar{y}}(u,v,T;\bar{x},\bar{y},s) h(\bar{y}) \phi(\bar{x},\bar{y},s;x,y,t) d\bar{y} d\bar{x} ds dv du. \end{array} \tag{2.46}$$

Proof. Let us work with the adjoint equation

$$\begin{array}{l} -\phi_s + \frac{x^2\sigma^2(y)}{2}\phi_{\bar{x}\bar{x}} + \frac{\beta^2}{2}\phi_{\bar{y}\bar{y}} \\ +(2\sigma^2(y)-r)x\phi_{\bar{x}} - (\alpha(m-y)-\beta\gamma(y))\phi_{\bar{y}} + (\sigma^2(y)-r+m-\beta\gamma'(y))\phi = 0 \end{array} \tag{2.47}$$

performing on it the following perturbation on the coefficient γ,

$$\gamma(y) \longrightarrow \gamma(y) + \epsilon h(y)\ .$$

So that $\phi \longrightarrow \phi + \epsilon(\delta\phi)$. We now substitute this into Equation (2.47), equate terms on ϵ and we get the following equation for the first variation $\delta\phi$.

$$\begin{array}{rcl} -\delta\phi_s + \frac{x^2\sigma^2(y)}{2}\delta\phi_{\bar{x}\bar{x}} + \frac{\beta^2}{2}\delta\phi_{\bar{y}\bar{y}} & + & (2\sigma^2(y)-r)x\delta\phi_{\bar{x}} - (\alpha(m-y)-\beta\gamma(y))\delta\phi_{\bar{y}} \\ +(\sigma^2(y)-r+m-\beta\gamma'(y))\delta\phi & = & \beta h(y)\phi_{\bar{y}}(y) + \beta h'(y)\phi. \\ & = & (h(\bar{y})\phi)_{\bar{y}}. \end{array} \tag{2.48}$$

This looks exactly like (2.47), but with the forcing term $(h(\bar{y})\phi)_{\bar{y}}$. Then, using *Duhamel's principle* we can write the solution of (2.48), as

$$\begin{array}{l} \delta\phi_{[h]}(u,v,T;x,y,t) = \\ -\iint_t^T \phi(u,v,T;\bar{x},\bar{y},s) \frac{\partial}{\partial \bar{y}}[h(\bar{y})\phi(\bar{x},\bar{y},s;x,y,t)] d\bar{y} d\bar{x} ds = \\ \iint_t^T \phi_{\bar{y}}(u,v,T;\bar{x},\bar{y},s) h(\bar{y}) \phi(\bar{x},\bar{y},s;x,y,t) d\bar{y} d\bar{x} ds, \end{array} \tag{2.49}$$

where the last equality was obtained by integration by parts. Since ϕ is the fundamental solution of the operator $\frac{\partial}{\partial t} + \mathcal{L}$, we may write the solution of (2.8) as

$$\begin{array}{rcl} U(x,y,t) & = & e^{-r(T-t)} E^{Q_\gamma}\left[(X_T-K)^+ \middle| X_t = x, Y_t = y\right] \\ & = & e^{-r(T-t)} \int (u-K)^+ \phi(u,v,T;x,y,t) dv du. \end{array} \tag{2.50}$$

Then, the variation $\partial\Gamma/\partial\gamma$ should be

$$\frac{\partial\Gamma}{\partial\gamma}(\gamma)[h(s)] =: \delta U[h(s)]_{(x,y,t)} = e^{-r(T-t)}\int (u-K)^+\delta\phi_{[h]}(u,v,T;x,y,t)dvdu. \tag{2.51}$$

Finally, we take the last expression in (2.49) and plug it into (2.51), to get the desired formula (2.46).

$$\frac{\partial\Gamma}{\partial\gamma}(\gamma)[h(s)] = \tag{2.52}$$

$$e^{-r(T-t)}\iint_t^T (u-K)^+\phi_{\bar{y}}(u,v,T;\bar{x},\bar{y},s)h(\bar{y})\phi(\bar{x},\bar{y},s;x,y,t)dvdud\bar{y}d\bar{x}ds. \qquad \square$$

Despite the fact that Equation (2.52) could be of some use, we see that in contrast to (2.41), Monte Carlo method might not help us too much here since the density derivative $\phi_{\bar{y}}(u,v,T;\bar{x},\bar{y},s)$ is not easy at all to be numerically simulated.

CHAPTER 3

The Inverse Problem

In this chapter we shall sketch a complete plan to handle the Inverse Problem 1. This is needed since Equations (2.41) and (2.46) from Chapter 2 are still too complex for numerical implementation.

We start by giving some intuition on how the forward map $\Gamma : \gamma \mapsto U$ acts in the case of zero correlation ($\rho = 0$). Using this we will design a strategy to invert this operator. In the final part of the chapter we give a rough idea that could be used to begin the work with the uncorrelated case.

Let us come back to the formula of $\partial\Gamma/\partial\gamma$ in (2.41) and see how much of it could be treated numerically. As we saw in Chapter 2

$$\frac{\partial \Gamma}{\partial \gamma}[h](y,t) = \mathbb{E}\left[\frac{\partial U_{\mathrm{BS}}}{\partial S}(\xi_T)\int_t^T e^{-\theta_s}(\eta_s - \eta_T)h(Y_s)ds\right] = \tag{3.1}$$
$$\iint_t^T \frac{\partial U_{\mathrm{BS}}}{\partial S}(\xi_T)e^{-\theta_s}(\eta_s-\eta_T)h(Y_s)\Psi(\xi_T,\eta_T,T;\theta_s,\eta_s,Y_s,s;y,t)d(\theta_s,\eta_s,Y_s,s,\eta_T,\xi_T)\ ,$$

where

$$\begin{aligned} dY_s &= f(Y_s)dt + \beta dW_s^2, \\ \eta_s &= 2\int_t^s \sigma(Y_v)\sigma'(Y_v)e^{\theta_v}dv, \\ \theta_s &= \int_t^s f'(Y_v)dv, \\ \xi_s &= \int_t^s \sigma^2(Y_v)dv, \end{aligned}$$

and

$$d(\theta_s,\eta_s,Y_s,s,\eta_T,\xi_T) = d\theta_s d\eta_s dY_s ds d\eta_T d\xi_T.$$

The joint density Ψ for this system of diffusions solves the following initial value problem

$$-\frac{\partial\Psi}{\partial t}+\frac{\beta^2}{2}\frac{\partial^2\Psi}{\partial y^2}-2\sigma(y)\sigma'(y)e^{\theta}\frac{\partial\Psi}{\partial\eta}-f'(y)\frac{\partial\Psi}{\partial\theta}-\sigma^2(y)\frac{\partial\Psi}{\partial\xi}=0\ , \tag{3.2}$$
$$\Psi(y,\eta,\theta,\xi,t_0;y_0,t_0)=\delta_{(y_0,0,0,0)}(y,\eta,\theta,\xi)\ .$$

From (3.1), we see that Malliavin calculus helped us to avoid the inconvenient presence of the derivative of the diffusion density $\phi_{\bar{y}}$ in the formula (2.52)

$$\frac{\partial\Gamma}{\partial\gamma}[h]=e^{-r(T-t)}\iint_t^T(u-K)^+\phi_{\bar{y}}(u,v,T;\bar{x},\bar{y},s)h(\bar{y})\phi(\bar{x},\bar{y},s;x,y,t)dvdud\bar{y}d\bar{x}ds.$$

So, instead of $\phi_{\bar{y}}$, by making use of Malliavin calculus we can consider a diffusion in a higher dimensional space. To wit, a four dimensional space in the variables $(\xi_s,\theta_s,\eta_s,Y_s)$.

The question we address now is: How to apply Equation (3.1) in Landweber's method in the reconstruction of $\gamma(y)$?. We recall the Landweber iteration

$$\gamma_{k+1}=\gamma_k-\frac{\partial\Gamma^*}{\partial\gamma}[\gamma_k](\Gamma(\gamma_k)-U). \tag{3.3}$$

Here, as before, U is the given data. Through this iterative method we are looking for a least square solution of

$$\|\Gamma(\gamma)-U\|^2\ ,$$

and $\partial\Gamma^*/\partial\gamma$ is the adjoint of the operator given in (3.1). Let us set $V(x,y,t):=(\Gamma(\gamma)-U)$. Then, also from (3.1) it can be seen that

$$\frac{\partial\Gamma^*}{\partial\gamma}[\gamma](V)(\omega)\ = \tag{3.4}$$

$$\iint_{t_0}^T\frac{\partial U_{\mathrm{BS}}}{\partial S}(\ \xi_T\ )e^{-\theta_s}(\eta_s-\eta_T)\Psi(\xi_T,\eta_T,T;\theta_s,\eta_s,\omega,s;y,t)V(x,y,t)d(\theta_s,\eta_s,s,\eta_T,\xi_T,x,y,t).$$

This last equation displays the difficulty in implementing Landweber's method. We must work numerically essentially with Equation (3.2) and its adjoint, both four dimensional PDEs, and the error should be kept under control in a bounded domain, the integral in the last expression also must be approximated.

So, we will seek another technique that would reduce the computational complexity of the equations involved. The key is the Hull-White formula that we already met in Chapter 2.

$$U(x,y,t,K,T)=\mathbb{E}\left[U_{\mathrm{BS}}[\xi_T](x,T-t,K,r)\right]\ , \tag{3.5}$$

where

$$\xi_T=\int_t^T\sigma^2(Y_s)ds\ . \tag{3.6}$$

Here, U represents the prices of European call options in the stochastic volatility model that we are considering with maturity T, strike price K, when the states of the stock price and volatility level were (x, y) at the initial time t. Furthermore, U_{BS} is the Black-Scholes function, so it represents the prices of European call options in the classic model where the volatility is considered constant. Let us write the Hull-White formula in terms of the density Φ of the process ξ_T,

$$U(x, y, t, K, T) = \int_0^\infty U_{\mathrm{BS}}[\xi](x, T-t, K, r)\Phi(\xi, T; y, t)d\xi, \tag{3.7}$$

Here, $\Phi(\xi, T; y, t)$ is the distribution of $\xi_T = \int_t^T \sigma^2(Y_s)ds$. It clearly depends on the initial value y at initial time t.

One idea to reduce the computational complexity associated to solving the PDE in Equation (3.2) would be to first "invert" for Φ in the linear Equation (3.7).

Remark 2. *A important theoretical problem, which may be investigated numerically, is that of computing the singular value decomposition of the integral operator*

$$\begin{aligned} \mathcal{U} \ : \ & A \to B \\ & \phi \mapsto \int_0^\infty U_{\mathrm{BS}}[\xi](\cdot/K, T-t, K, r)\phi(\xi)d\xi \end{aligned} \tag{3.8}$$

in appropriate spaces A and B.

In fact, this is a classical integral equation whose literature is pretty extensive [BL05, EHN96]. The next step is to use the distribution of the process Y_t that follows the dynamics

$$\begin{aligned} dY_t &= f(Y_t)dt + \beta dW_t^2, \\ Y_{t_0} &= y_{t_0}. \end{aligned} \tag{3.9}$$

From now on, we will denote the distribution of Y_t by $\rho(y, t; y_0, t_0)$ for $t > t_0$. It turns out that ρ has a more direct relation with the function γ. This link is given by the simpler PDE associated to the process of Equation (3.9)

$$\begin{aligned} \frac{\partial \rho}{\partial t} &= \frac{\beta^2}{2}\frac{\partial^2 \rho}{\partial y^2} - \frac{\partial}{\partial y}(f(y)\rho) \ , \\ \rho(y, t_0; y_0, t_0) &= \delta_{y_0}(y) \ , \end{aligned} \tag{3.10}$$

where $f(y) = \alpha(m - y) - \beta\gamma(y)$ and, as usual, $\delta_{y_0}(y)$ is the Dirac measure supported at y_0.

3.1 The Action of Γ.

In this part we would like to describe intuitively the mechanisms by which the operator $\Gamma : \gamma \mapsto U$, acts. We start with equations (3.9) and (3.10). Thus, the function

f controls the evolution of the mean of the paths of the process Y_t, by a discretization of (3.9). It can be seen that for a small time increment Δt, the transition density $\rho(y, t_0 + \Delta t; y_0, t_0)$ has the following form

$$\rho(y, t_0 + \Delta t; y_0, t_0) \approx \frac{1}{\sqrt{2\pi\beta(\Delta t)}} \exp\left[- \frac{(y - y_0 - f(y_0)\Delta t)^2}{2\beta(\Delta t)} \right].$$

The process Y_t is then mapped to $w_t = \sigma^2(Y_t)$ and we assume that the function $p(w_{t_2}; w_{t_1})$ is the transition density function from w_{t_1} to w_{t_2}.

Next we consider the distribution of the functional ξ_T defined on the paths of the process Y_v by

$$\xi_T = \int_t^T \sigma^2(Y_v) dv.$$

The joint density $\bar{\rho}(Y, \xi, s; y, 0, t)$ of (Y_s, ξ_s) for $t \leq s \leq$ given that $Y_t = y, \xi_t = 0$ at the initial time t, satisfies the following initial value problem

$$\begin{aligned} &-\tfrac{\partial \bar{\rho}}{\partial T} + \tfrac{\beta^2}{2} \tfrac{\partial^2 \bar{\rho}}{\partial y^2} - \tfrac{\partial}{\partial y}(f(y)\bar{\rho}) - \sigma^2(y) \tfrac{\partial \bar{\rho}}{\partial \xi} = 0, \\ &\bar{\rho}(Y, \xi, t; y, 0, t) = \delta_{(y,0)}(Y, \xi). \end{aligned} \tag{3.11}$$

However, we prefer to work with (3.10) instead of (3.11). One of the reasons for this is that we want to reduce the dimension of our equations as much as possible. Another reason is that we want to focus in the specific case of the Stein model [SS91] where the function $\sigma(y)$ is taken to be $|y|$. From now on we consider $\sigma(y) = |y|$. Finally, we note that the distribution $\Phi(\xi, T; y, t)$ of ξ_T is given by

$$\Phi(\xi, T; y, t) = \int_{-\infty}^{\infty} \bar{\rho}(Y, \xi, s; y, 0, t) dY.$$

Once the function Φ is determined, we have that the value of our operator Γ in γ is given by

$$\Gamma(\gamma) = U(x, y, t, T, K) = \int_0^{\infty} U_{\mathrm{BS}}[\xi](x, T - t, K, r) \Phi(\xi, T; y, t) d\xi,$$

This completes our intuitive description of the action of the operator Γ.

Thus, we can think of the operator Γ as a composition of operators as follows: $\mathcal{G}_1 : \gamma \mapsto \rho$, $\mathcal{G}_2 : \rho \mapsto \bar{\rho}$, $\mathcal{G}_3 : \bar{\rho} \mapsto \Phi$ and $\mathcal{G}_4 : \Phi \mapsto U$ so that

$$\Gamma = \mathcal{G}_4 \circ \mathcal{G}_3 \circ \mathcal{G}_2 \circ \mathcal{G}_1. \tag{3.12}$$

Remark Another way of thinking of the relation between the distributions Φ and ρ is as follows. The process Y_t is mapped to $w_t = \sigma^2(Y_t)$ and we assume that the function $p(w_{t_2}; w_{t_1})$ is the transition density function from w_{t_2} to w_{t_1}. So we would like to discuse the following intuitive idea.

The function $\Phi(\xi, T; y, t)$ is given formally by a limit of the form:

$$\Phi(\xi, T; y, t) =$$

$$\lim_{\|\Delta t\| \to 0} 1/(\Delta t)^{n-1} \int p(\frac{\xi - z_{n-1}}{\Delta t}; \frac{z_{n-1} - z_{n-2}}{\Delta t}) p(\frac{z_{n-1} - z_{n-2}}{\Delta t}; \frac{z_{n-2} - z_{n-3}}{\Delta t}) \cdots p(\frac{z_1}{\Delta t}; y, t) dz_{n-1} \cdots dz_1. \tag{3.13}$$

Here, $p(w_{t_2}, w_{t_1})$ is the transition kernel (Markov transition) of the processes $w_t = \sigma^2(Y_t)$,i.e., the conditional probability

$$\Pr(\sigma^2(Y_{t_2}) = w_{t_2} | \sigma^2(Y_{t_1}) = w_{t_1}),$$

Remark: Once we know $p(w_{t_2}, w_{t_1})$, it is necessary a change of variable to get the density ρ of Y_t.

We now provide an heuristic argument for this statement. We think of the random variable $\xi_T = \int_t^T \sigma^2(Y_s) ds$ as being approximated by a finite sum

$$\sum_{i=1}^{n} w_{t_i} = \sum_{i=1}^{n} \sigma^2(Y_{t_i}) \Delta t,$$

Then, let us inquire about the distribution of the last sum,

$$\Pr(w_1 + w_2 + \cdots w_n = \xi) = \int_{\{w_1 + w_2 + \cdots w_n = \xi\}} p(w_1, w_2, \ldots, w_{n-1}) dw_1 dw_2 \cdots dw_{n-1}, \tag{3.14}$$

Here $p(w_1, w_2, \ldots, w_{n-1})$ is the joint density of the variables $\{w_{t_i}\}_{i=1}^{n-1}$. Now we make the following change of variables in (3.14),

$$\begin{aligned} w_1 &= \frac{z_1}{\Delta t}, \\ w_2 &= \frac{z_2 - z_1}{\Delta t}, \\ w_3 &= \frac{z_3 - z_2}{\Delta t}, \\ &\vdots \\ w_{n-1} &= \frac{\xi - z_{n-1}}{\Delta t}. \end{aligned}$$

After this (3.14) will look like:

$$\begin{aligned} &= \frac{1}{(\Delta t)^{n-1}} \int_{\mathbb{R}^{n-1}} p(\frac{\xi - z_{n-1}}{\Delta t}, \frac{z_2 - z_1}{\Delta t}, \ldots, \frac{z_1}{\Delta t}) dz_{n-1} dz_{n-2} \cdots dz_1, \\ &= \frac{1}{(\Delta t)^{n-1}} \int_{\mathbb{R}^{n-1}} p(\frac{\xi - z_{n-1}}{\Delta t}; \frac{z_{n-1} - z_{n-2}}{\Delta t}) p(\frac{z_{n-1} - z_{n-2}}{\Delta t}; \frac{z_{n-2} - z_{n-3}}{\Delta t}) \cdots p(\frac{z_1}{\Delta t}; y, t) dz_{n-1} dz_{n-2} \cdots dz_1. \end{aligned} \tag{3.15}$$

The last equality is obtained by splitting the joint density

$$p(\frac{\xi - z_{n-1}}{\Delta t}, \frac{z_2 - z_1}{\Delta t}, \dots, \frac{z_1}{\Delta t}),$$

into the transitions $p(w_i, w_{i-1})$, using the Markov property.

3.2 A Strategy to Invert Γ in the Uncorrelated Case ($\rho = 0$) and γ linear

Through-out this section we assume γ to be linear. Let us recall the main equations

$$\begin{aligned} &U_t + \frac{|y|^2 x^2}{2} U_{xx} + \frac{\beta^2}{2} U_{yy} + (\alpha(m-y) - \beta\gamma(y))U_y - r(xU_x - U) = 0 \ , \\ &U(x,y,T) = (x-K)^+; \quad 0 \le t < T, \quad x > 0, \quad y \in \mathbb{R} \ . \end{aligned} \tag{3.16}$$

We would like to reconstruct $\gamma(y)$ based on the knowledge of the solution U. The SDE that controls the dynamics of the asset price and the volatility processes, in the risk neutral measure, is

$$\begin{aligned} dX_t &= rX_t dt + |Y_t| X_t dW_t^1, \\ X_0 &= x, \\ dY_t &= (\alpha(m - Y_t) - \beta\gamma(Y_t))dt + \beta dW_t^2, \\ Y_0 &= y. \end{aligned} \tag{3.17}$$

Here, as usual, W_t^1 and W_t^2 are two independent Brownian motions. Now, we will explain a possible methodology to handle our inverse problem which is based on trying to invert each of the operators in (3.12).

i Solve for the function Φ (the distribution density of the process $\xi_T = \int_t^T |Y_v|^2 dv$), assuming we are given as data the function $U(x, y, t, T, K)$ in the following linear problem

$$\Gamma(\gamma) = U(x, y, t, T, K) = \int_0^\infty U_{\mathrm{BS}}[\xi](x, T-t, K, r)\Phi(\xi, T; y, t)d\xi,$$

ii Determine the distribution $\rho(Y, v; y, t)$ of the process Y_v. For this let us note that this distribution is Gaussian whose parameters we will be talking a little about in a while.

Let us assume that $\gamma(y) = a + by$, for two constants a and b, then we have the equation for Y_t in the risk neutral measure

$$\begin{aligned} dY_t =& [\alpha(m - Y_t) - \beta(a + bY_t)]dt + \beta dW_t^2, \\ =& \alpha_1(m^* - Y_t)dt + \beta dW_t^2, \end{aligned} \tag{3.18}$$

here

$$\alpha_1 = \alpha + \beta b \ , \ \ m^* = \frac{\alpha m - \beta a}{\alpha + \beta b} \ .$$

Then we have for Y_t

$$d(e^{\alpha_1 t} Y_t) = \alpha_1 m^* e^{\alpha_1 t} dt + \beta e^{\alpha_1 t} dW_t^2,$$

and thus

$$Y_T = y e^{-\alpha_1 (T-t)} + e^{-\alpha_1 T} [\alpha_1 m^* \int_t^T e^{\alpha_1 s} ds + \beta \int_t^T e^{\alpha_1 s} dW_s^2].$$

Therefore, the mean value M_T and the variance Σ_T of Y_T are given by:

$$M_T = y e^{-\alpha_1 (T-t)} + m^* (1 - e^{-\alpha_1 (T-t)}), \ \ \Sigma_T = \frac{\beta^2}{2\alpha_1} (1 - e^{-2\alpha_1 (T-t)}) \ .$$

Once we have estimated Φ and taking into account that

$$\xi_s = \xi_t + \int_t^s Y_u^2 du,$$

we deduce that

$$\frac{d}{ds} \mathbb{E}[\xi_s] = \mathbb{E}[Y_s^2] \iff \frac{d}{ds} \int \xi \Phi(\xi, s; y, t) d\xi = \Sigma_s + M_s^2 \tag{3.19}$$

Finally, we need another equation that relates M_T and Σ_T that together with Eq. (3.19) determine the unknows a and b based on the knowledge of the distribution Φ. One posibility to obtain such equation could be computing $\mathbb{E}[Y_s^2]$ through Itô formula as follows:

Since

$$\begin{aligned} d(Y_s^2) =& 2 Y_s dY_s + d\langle Y \rangle_s, \\ =& 2 Y_s [\alpha_1 (m^* - Y_s) ds + \beta dW_s^2] + \beta^2 ds, \end{aligned}$$

we have

$$\mathbb{E}[Y_T^2] = y^2 + \int_t^T \mathbb{E}[2\alpha_1 Y_s (m^* - Y_s) + \beta^2] ds,$$

and thus

$$\frac{d}{dT} \big(\Sigma_T + M_T^2 \big) = 2\alpha_1 m^* M_T - 2\alpha_1 (\Sigma_T + M_T^2) + \beta^2.$$

Therefore, the equation that relates M_T and Σ_T is

$$2\alpha_1\Sigma_T + \frac{d}{dT}\Sigma_T = 2\alpha_1 m^* M_T - \frac{d}{dT}M_T^2 - 2\alpha_1 M_T^2 + \beta^2. \tag{3.20}$$

Finally, let us observe that we have the following estimation for the distribution $\rho(Y, v; y, t)$ of $\{Y_v\}_{v\geq t}$.

$$\rho(Y, v) = \begin{cases} \frac{1}{\sqrt{2\pi\Sigma_v}} e^{-\frac{(Y-M_v)^2}{2\Sigma_v}}, & v > t, \\ \delta_y(Y), & v = t. \end{cases}$$

iii Next, through the technique of Malliavin calculus we estimate γ from $\mathcal{G}_1 : \gamma(y) \mapsto \rho(Y, v)$ via Landweber's method.

We now explain Step 3: Consider the equation

$$\begin{aligned} &\frac{\partial\rho}{\partial v} = \frac{\beta^2}{2}\frac{\partial^2\rho}{\partial y^2} - \frac{\partial}{\partial y}(f(y)\rho), \quad t_0 < v \ , \\ &\rho(y, t_0; y_0, t_0) = \delta_{y_0}(y) \ , \end{aligned} \tag{3.21}$$

We now want to reconstruc the function $f(y)$ from the knowledge of ρ. However, instead of ρ will use an average of it. For that, fix a smooth function $\phi(y)$ and consider the function $\Upsilon(y, t)$ given by

$$\Upsilon(y, t) = \int_{-\infty}^{\infty} \phi(Y)\rho(Y, v; y, t)dY = \mathbb{E}^{y,t}[\phi(Y_v)].$$

Let us denote by $\mathcal{G}(f) = \Upsilon(y, t)$ the operator that maps f to Υ. We consider a technique based on Malliavin calculus of Chapter 2 applying Landweber's method.

$$f_{k+1} = f_k - \frac{\partial\mathcal{G}}{\partial f}^*[f_k](\mathcal{G}[f_k] - \Upsilon),$$

Furthermore, we denote by

$$\frac{\partial\mathcal{G}}{\partial f}[f](h)$$

the derivative of $\mathcal{G}$ evaluated at the point f and applied to h.

Proposition 9. *The Fréchet derivative* $\frac{\partial\mathcal{G}}{\partial f}[f](h)$ *is given by*

$$\frac{\partial\mathcal{G}}{\partial f}[f](h) = \iint_t^v \phi'(Y_v)e^{\theta_v-\theta_s}h(Y_s)\Psi(Y_v, \theta_v, v; Y_s, \theta_s, s; y, t)dY_v d\theta_v dY_s d\theta_s ds, \tag{3.22}$$

where Ψ is the joint density at the relevant times of the two-dimensional diffusion (Y_s, θ_s) that follows the dynamics

$$\begin{aligned} dY_s &= f(Y_s)ds + \beta dW_s^2, \\ d\theta_s &= f'(Y_s)ds. \end{aligned}$$

Furthermore,

$$\frac{\partial \mathcal{G}^*}{\partial f}[f](V)(\omega) = \iint_{t_0}^{r} \phi'(Y_v) e^{\theta_v - \theta_s} \Psi(Y_v, \theta_v, v; \omega, \theta_s, s; y, t) V(y,t) dY_v d\theta_v d\theta_s ds dy dt. \tag{3.23}$$

Proof: We basically follow the same route as in the proof of Proposition 7 of Chapter 2. Consider the SDE

$$dY_s = f(Y_s)ds + \beta dW_s^2 \ .$$

Then, let us consider also the following perturbation of the SDE above

$$\begin{aligned} dY_s &= (f(Y_s) + \epsilon h(Y_s))ds + \beta d\widetilde{W}_s, \\ d\widetilde{W}_s &= -\frac{\epsilon}{\beta} h(Y_s)dt + dW_s^2, \end{aligned} \tag{3.24}$$

By Girsanov's theorem, $\widetilde{W}_s$ is a Brownian motion under the probability measure

$$dP^\epsilon = e^{\frac{\epsilon}{\beta}\int_t^v h(Y_s)dW_s^2 - \frac{\epsilon^2}{\beta^2}\int_t^v h^2(Y_s)ds} dP.$$

Then,

$$\mathcal{G}(f + \epsilon h) = \mathbb{E}^{P^\epsilon}[\phi(Y_v)].$$

So,

$$\begin{aligned} \frac{\partial \mathcal{G}}{\partial \gamma} &= \frac{d}{d\epsilon}\mathbb{E}^{P^\epsilon}[\phi(Y_v)]_{\epsilon=0}, \\ &= \frac{d}{d\epsilon}\mathbb{E}^{P}[e^{\frac{\epsilon}{\beta}\int_t^v h(Y_s)dW_s^2 - \frac{\epsilon^2}{\beta^2}\int_t^v h^2(Y_s)ds}\phi(Y_v)]_{\epsilon=0}, \\ &= \frac{1}{\beta}\mathbb{E}\left[\left(\int_t^v h(Y_s)dW_s^2\right)\phi(Y_v)\right], \\ &= \frac{1}{\beta}\mathbb{E}\left[\int_t^v \phi'(Y_v)(D_sY_v)h(Y_s)ds\right], \end{aligned}$$

The last equality can be justified as follows:
Since

$$dY_s = f(Y_s)ds + \beta dW_s^2,$$

we have for the Malliavin derivative D_sY_v that

$$dD_sY_v = f'(Y_v)(D_sY_v)ds\ .$$

Therefore,

$$D_sY_v = \beta\chi_{\{s<v\}}e^{\int_s^v f'(Y_u)du},$$

and thus

$$\frac{\partial \mathcal{G}}{\partial \gamma} = -\mathbb{E}\Big[\int_t^v \phi'(Y_v)e^{\theta_v-\theta_s}h(Y_s)ds\Big],$$

where we have used the notation

$$\theta_s := \int_t^s f'(Y_u)du.$$

Finally, we write the above expectation in terms of the joint density of the diffusions in the respective times to get the formula (3.22). □

The next proposition shows how to compute (3.23) in practice by solving some PDEs that have Ψ as fundamental solution. See Section 1.5 of Chapter 1.

Proposition 10. *An alternative representation for (3.23) is*

$$\frac{\partial \mathcal{G}^*}{\partial f}[f](V)(\omega) = \iint_{t_0}^{v} Z(\omega,\theta,s)e^{-\theta}W(\omega,\theta,s)d\theta ds\ , \tag{3.25}$$

where the functions $Z(\omega,\theta,s)$ and $W(\omega,\theta,s)$ satisfy the following problems:

$$\frac{\partial Z}{\partial s} + \frac{\beta^2}{2}\frac{\partial^2 Z}{\partial \omega^2} + f(\omega)\frac{\partial Z}{\partial \omega} - f'(\omega)\frac{\partial Z}{\partial \theta} \quad = \quad 0, \quad t_0 \le s \le v, \tag{3.26}$$

$$Z(\omega,\theta,v) = \phi'(\omega)e^{\theta},$$

and

$$\begin{gathered} -\frac{\partial W}{\partial s} + \frac{\beta^2}{2}\frac{\partial^2 W}{\partial \omega^2} - \frac{\partial}{\partial \omega}(f(\omega)W) + f'(\omega)\frac{\partial W}{\partial \theta} = V(\omega,s), \quad t_0 \le s \le v, \\ W(\omega,\theta,t_0) = 0. \end{gathered} \tag{3.27}$$

Proof: Just remark that

$$Z(\omega,\theta,s) = \int \phi'(Y_v)e^{\theta_v}\Psi(Y_v,\theta_v,v;\omega,\theta,s)dY_v d\theta_v, \tag{3.28}$$

solves (3.26), and that

$$W(\omega,\theta,s)=\int_{t_0}^{v}\int \Psi(\omega,\theta,s;y,t)V(y,t)dydt \tag{3.29}$$

satisfies (3.27). Finally substituing (3.28) and (3.29) in (3.25) we get (3.23). $\square$

Remark. The approach to attack the inverse problem that lead us to equations (3.26) and (3.27) by computing $\partial\mathcal{G}^*/\partial f[f](V)$ through (3.25) seems simpler than that of Equation (2.41) in Chapter 2. This is due to the fact that the role of the boundary values at truncated domains is simplified when performing numerical simulations. We also remark that the method suggested by the Proposition 10 to compute (3.23) may be used to compute the formulas (2.41) and (2.46) in Chapter 2.

3.3 A Note in The Correlated Case.

Our aim here is to sketch a possible starting point to work out the reconstruction problem in the case of correlation $\rho \neq 0$. We continue assuming that the coefficient γ depends only on y, in a linear form, i.e., $\gamma(y)=a+by$.

Let us return to the general equations of the asset dynamics in the risk neutral measure.

$$\begin{aligned} dX_t &= rX_tdt+|Y_t|X_t((1-\rho^2)^{1/2}dW_t^1+\rho dW_t^2), \\ X_0 &= x, \\ dY_t &= (\alpha(m-Y_t)-\beta\Lambda(Y_t))dt+\beta dW_t^2, \\ Y_0 &= y. \end{aligned} \tag{3.30}$$

Here, as usual, W_t^1 and W_t^2 are both independent Brownian motions and $\Lambda(y) = \rho\frac{\mu-r}{|y|}+\sqrt{1-\rho^2}\gamma(y)$. We also consider the solution $U(x,y,t,T,K)$ of the final value problem for the price of European call options based on X_t,

$$\begin{aligned} &U_t+\frac{|y|^2x^2}{2}U_{xx}+\frac{\beta^2}{2}U_{yy}+\rho\beta x|y|U_{xy} \\ &+(\alpha(m-y)-\beta\Lambda(y))U_y-r(xU_x-U)=0, \\ &U(x,y,T)=(x-K)^+, \end{aligned} \tag{3.31}$$

where

$$\Lambda(y)=\rho\frac{\mu-r}{|y|}+\sqrt{1-\rho^2}\gamma(y), \text{and} \quad 0\le t<T, x>0, y\in\mathbb{R}.$$

From the expression for Λ we see that the closer we get of total correlation, (i.e. $\rho \approx \pm 1$) the less it makes sense to invert for the function γ since in such case the solution of (3.31) does not depend significantly on γ. Therefore, let us assume in this section that $|\rho| \in (0,1)$ is far enough from the endpoints. We will try to imitate the procedure followed in the case of zero correlation. First, we will see that there is a generalization of the Hull-White formula for the correlated case. Then, we would like to estimate the distribution of the volatility process Y_t. Finally, we will try to apply Malliavin calculus.

Let us discuss then the Hull-White formula keeping in mind its derivation in the uncorrelated case.

Now, in the presence of correlation, from Equation (3.30) and Itô calculus we can see that

$$\begin{aligned} X_T &= x e^{\int_t^T (r - \frac{|Y_s|^2}{2})ds + \int_t^T |Y_s|(\rho dW_s^2 + (1-\rho^2)^{1/2} dW_s^1)}, \\ &= x e^{-\rho^2 \int_t^T \frac{|Y_s|^2}{2} ds + \rho \int_t^T |Y_s| dW_s^2} e^{\int_t^T (r - (1-\rho^2)\frac{|Y_s|^2}{2})ds + (1-\rho^2)^{1/2} \int_t^T |Y_s| dW_s^1}, \\ &= x e^{\eta_T} e^{\int_t^T (r - (1-\rho^2)\frac{|Y_s|^2}{2})ds + (1-\rho^2)^{1/2} \int_t^T |Y_s| dW_s^1}. \end{aligned} \tag{3.32}$$

Here,

$$\eta_T = -\rho^2 \int_t^T \frac{|Y_s|^2}{2} ds + \rho \int_t^T |Y_s| dW_s^2.$$

At this point, we should compare equation (3.32) with (2.32) of Chapter 2. Note, that they are basically the same after substituting x and the process $|Y_s|$ in (2.32) by xe^{η_T} and the process $(1-\rho^2)^{1/2}|Y_s|$, respectively, getting (3.32). Then, in this case, the Hull-White formula yields

$$U(x, y, t, T, K) = E\Big[U_{\mathrm{BS}}(xe^{\eta_T}, t, T, K, \xi_T)\Big], \tag{3.33}$$

where

$$\xi_T = (1-\rho^2) \int_t^T |Y_s|^2 ds.$$

Let us write (3.33) in terms of the joint density $\Phi(\eta, \xi, T; y, t)$ of η_T and ξ_T:

$$U(x, y, t, T, K) = \int U(xe^{\eta}, t, T, K, \xi)\Phi(\eta, \xi, T; y, t) d\eta d\xi, \tag{3.34}$$

Now, the first thing to do is to solve (3.34) for Φ from the data $U(x, y, t, T, K)$, and as in the uncorrelated case we want to estimate the distribution of the process $|Y_s|$ from Φ. Now, differently from the uncorrelated case the distribution function of Y_s is not normal despite the fact that it continue being determinated by three constant parameters together with the correlation coefficient ρ. Let us note that since we are assuming that $\gamma(y)$ is linear we can write the SDE for Y_s in (3.30) as

$$dY_s = [\alpha_2(m_2 - Y_s) - \frac{\rho\beta}{|Y_s|}]ds + \beta dW_s^2. \tag{3.35}$$

Comparing (3.35) with (3.18), we easily see that we are forced to first take care of estimating ρ and then to search a way of estimating the parameters α_2 and m_2. At this point we would proceed with Malliavin calculus and Landweber iteration as we described at the end of the previous section to reconstruct the function $\gamma(y)$ from the density $\rho(Y, s)$ of the process Y_s.

CHAPTER 4

Conclusions and Suggestions for Further Research

In this work we have proved that the forward map $\Gamma : \gamma \to U$ is Frechét differentiable at bounded Hölder continuous functions $\gamma(\cdot)$ in the direction of functions belonging to an L^∞-type space. Even more, we proved that Γ is locally analytic, using the regularization effect of the semigroup associated to the map $\Gamma : \gamma \to U$.

Then, we have computed by two methods the derivative $\partial\Gamma/\partial\gamma$. The first one, which used classical techniques from partial differential equations, led to the formula

$$\begin{aligned}&\frac{\partial\Gamma}{\partial\gamma}(\gamma)[h] = \\ &e^{-r(T-t)}\iint\int_t^T (u-K)^+ \phi_{\bar{y}}(u,v,T;\bar{x},\bar{y},s)h(\bar{y})\phi(\bar{x},\bar{y},s;x,y,t)dvdud\bar{y}d\bar{x}ds\ .\end{aligned} \tag{4.1}$$

The second one, which used Malliavin calculus, led to

$$\begin{aligned}&\frac{\partial\Gamma}{\partial\gamma}(\gamma)[h] = \\ &\iint_t^T \frac{\partial U_{\mathrm{BS}}}{\partial S}(\xi_T)e^{-\theta_s}(\eta_s-\eta_T)h(Y_s)\Psi(\xi_T,\eta_T,T;\theta_s,\eta_s,Y_s,s;y,t)d(\theta_s,\eta_s,Y_s,s,\eta_T,\xi_T)\ .\end{aligned} \tag{4.2}$$

The expression obtained with the help of Malliavin calculus seems to be better suited to numerical purposes. In formula (4.2), in contrast to (4.1), we have the possibility of performing the calculations through Monte Carlo methods. This is due to the fact that in the Malliavin formula case (4.2) we do not have the inconvenient presence of the derivative of the Markov kernel for the underlying diffusions. Furthermore, equations associated to (4.2), such as for example (3.2), although higher dimensional[1],

[1]The increase in the dimension of these equations is due to the consideration of the distribution of some Malliavin derivative process in the calculation following this approach.

are less complex to be treated numerically than the ones associated to (4.1). Indeed, recalling Equation (3.2)

$$-\frac{\partial\Psi}{\partial t}+\frac{\beta^2}{2}\frac{\partial^2\Psi}{\partial y^2}-2\sigma(y)\sigma'(y)e^{\theta}\frac{\partial\Psi}{\partial\eta}-f'(y)\frac{\partial\Psi}{\partial\theta}-\sigma^2(y)\frac{\partial\Psi}{\partial\xi}=0\ , \tag{4.3}$$
$$\Psi(y,\eta,\theta,\xi,t_0;y_0,t_0)=\delta_{(y_0,0,0,0)}(y,\eta,\theta,\xi)\ ,$$

we see that we avoid the singular diffusion term $|y|^2x^2\partial^2/\partial x^2$. Furthermore, the boundary values for (4.3) can be considered as homogeneous for values of (y,η,θ,ξ) big enough.

We proposed the formula in (4.2) for $\partial\Gamma/\partial\gamma$ to attack the reconstruction problem in the context of Landbewer regularization. We discussed some issues concerning practical implementations like for example the use of option prices quoted both in the maturity T and the strike price K.

We presented a decomposition of the forward operator Γ that seems to simplify the analysis and the numerics of the inverse problem. This decomposition splits the inversion procedure first in a linear problem to reconstruct the distribution Φ of the process $\xi_T=\int_t^T\sigma^2(Y_s)ds$ (time-averaged volatility process).

In a second part it is proposed an alternative approach to a finite difference treatment of some four-dimensional partial differential equations like Equation (4.3), that consists of determining the distribution of the Ornstein-Uhlenbeck process Y_t, by estimating its mean value using the distribution Φ

Finally, we addressed the method to compute functional derivatives using Malliavin calculus to solve the nonlinear problem that links the distribution $\rho(Y,t)$ of Y_t and our target function f through[2] Landbewer regularization.

Finally, we suggested an approach to the reconstruction problem when $\rho\neq 0$, trying to adapt the ideas proposed in the uncorrelated case.

A natural continuation of this work is the numerical implementation and verification of the ideas presented here. Another direction would be to consider the reconstruction problem of the risk premium in option pricing models where the underlying asset is modeled by fractional Brownian motion or more general Levy processes so as to take into account the presence of jumps in the asset dynamics.

[2]Recall that we are actually interested on the function $\gamma(y)$, but $f(y)=\alpha(m-y)-\beta\gamma(y)$.

Bibliography

[AFT05] Yves Achdou, Bruno Franchi, and Nicoletta Tchou. A partial differential equation connected to option pricing with stochastic volatility: regularity results and discretization. *Math. Comp.*, 74(251):1291–1322 (electronic), 2005.

[AT02] Yves Achdou and Nicoletta Tchou. Variational analysis for the Black and Scholes equation with stochastic volatility. *M2AN Math. Model. Numer. Anal.*, 36(3):373–395, 2002.

[Ben02] Eric Benhamou. Smart Monte Carlo: various tricks using Malliavin calculus. *Quant. Finance*, 2(5):329–336, 2002.

[BI99] Ilia Bouchouev and Victor Isakov. Uniqueness, stability and numerical methods for the inverse problem that arises in financial markets. *Inverse Problems*, 15(3):R95–R116, 1999.

[BL05] Johann Baumeister and Antonio Leitão. *Topics in inverse problems*. Publicações Matemáticas do IMPA. [IMPA Mathematical Publications]. Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2005. 25º Colóquio Brasileiro de Matemática. [25th Brazilian Mathematics Colloquium].

[BS73] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *The Journal of Political Economy*, 81(3):637–654, 1973.

[Cal80] Alberto-P. Calderón. On an inverse boundary value problem. In *Seminar on Numerical Analysis and its Applications to Continuum Physics (Rio de Janeiro, 1980)*, pages 65–73. Soc. Brasil. Mat., Rio de Janeiro, 1980.

[Cré03] S. Crépey. Calibration of the local volatility in a generalized Black-Scholes model using Tikhonov regularization. *SIAM J. Math. Anal.*, 34(5):1183–1206 (electronic), 2003.

[CT06] Rama Cont and Peter Tankov. Retrieving Lévy processes from option prices: regularization of an ill-posed inverse problem. *SIAM J. Control Optim.*, 45(1):1–25 (electronic), 2006.

[Duf01] Darrell Duffie. *Dynamic asset pricing theory. 3rd ed.* Princeton, NJ: Princeton University Press. xix, 465 p., 2001.

[Dup97] Bruno Dupire. Pricing and hedging with smiles. In *Mathematics of derivative securities (Cambridge, 1995)*, volume 15 of *Publ. Newton Inst.*, pages 103–111. Cambridge Univ. Press, Cambridge, 1997.

[EHN96] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.

[FLL+99] Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, Pierre-Louis Lions, and Nizar Touzi. Applications of Malliavin calculus to Monte Carlo methods in finance. *Finance Stoch.*, 3(4):391–412, 1999.

[FLLL01] Eric Fournié, Jean-Michel Lasry, Jérôme Lebuchoux, and Pierre-Louis Lions. Applications of Malliavin calculus to Monte-Carlo methods in finance. II. *Finance Stoch.*, 5(2):201–236, 2001.

[FPS00] Jean-Pierre Fouque, George Papanicolaou, and K. Ronnie Sircar. *Derivatives in financial markets with stochastic volatility.* Cambridge University Press, Cambridge, 2000.

[FPSS03a] J.-P. Fouque, G. Papanicolaou, R. Sircar, and K. Solna. Singular perturbations in option pricing. *SIAM J. Appl. Math.*, 63(5):1648–1665 (electronic), 2003.

[FPSS03b] Jean-Pierre Fouque, George Papanicolaou, Ronnie Sircar, and Knut Solna. Multiscale stochastic volatility asymptotics. *Multiscale Model. Simul.*, 2(1):22–42 (electronic), 2003.

[Fri64] Avner Friedman. *Partial differential equations of parabolic type.* Prentice-Hall Inc., Englewood Cliffs, N.J., 1964.

[Fri75] Avner Friedman. *Stochastic differential equations and applications. Vol. 1.* Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975. Probability and Mathematical Statistics, Vol. 28.

[HH03] Torsten Hein and Bernd Hofmann. On the nature of ill-posedness of an inverse problem arising in option pricing. *Inverse Problems*, 19(6):1319–1338, 2003.

[Kar88] Ioannis Karatzas. A tutorial introduction to stochastic analysis and its applications. page 57, 1988. Department of Statistics Columbia University New York, N.Y. 10027.

[KK01] Ralf Korn and Elke Korn. *Option pricing and portfolio optimization*, volume 31 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Modern methods of financial mathematics, Translated from the 1999 German original by the authors.

[KS91] Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[Lio00] P. L. Lions. On mathematical finance. *Boll. Unione Mat. Ital. Sez. B Artic. Ric. Mat. (8)*, 3(3):553–572, 2000.

[Mer73] Robert C. Merton. Theory of rational option pricing. *Bell J. Econom. and Management Sci.*, 4:141–183, 1973.

[Nua06] David Nualart. *The Malliavin calculus and related topics*. Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[Øks03] Bernt Øksendal. *Stochastic differential equations*. Universitext. Springer-Verlag, Berlin, sixth edition, 2003. An introduction with applications.

[Paz83] A. Pazy. *Semigroups of linear operators and applications to partial differential equations*, volume 44 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1983.

[Sch95] Otmar Scherzer. Convergence criteria of iterative methods based on Landweber iteration for solving nonlinear problems. *J. Math. Anal. Appl.*, 194(3):911–933, 1995.

[Sch03] Walter Schachermayer. Introduction to the mathematics of financial markets. In *Lectures on probability theory and statistics (Saint-Flour, 2000)*, volume 1816 of *Lecture Notes in Math.*, pages 107–179. Springer, Berlin, 2003.

[Sko75] A. V. Skorohod. On a generalization of the stochastic integral. *Teor. Verojatnost. i Primenen.*, 20(2):223–238, 1975.

[SS91] EM Stein and JC Stein. Stock price distributions with stochastic volatility: an analytic approach. *Rev. Financ. Stud.*, 4(4):727–752, 1991.

[SZ06] Max Oliveira Souza and Jorge P. Zubelli. On the asymptotics of a stochastic volatility option price model. To appear: International Journal of Theoretical and Applied Finances, 2006.