D.E.A de Probabilités Appliquées - Filiére Finances
Université Pierre et Marie Curie - Paris VI

Finite dimensional approximation for the Musiela model in the Gaussian case

Tiziano Vargiolu

sous la direction de
Nicole El Karoui, Franco Flandoli et Jerzy Zabczyk

Année Académique 1996/'97

Contents

Introduction

The aim of this work is to analyze the Musiela model for the forward rates presented in [13], and in particular the Gaussian version of this model. The Musiela model is a reparametriziation of the HJM model [10], so it is essentially a set of relations which an interest rate model must satisfy in order not to have arbitrage opportunities in the market.

This work is intended to be an attempt to analyze this model, and in particular what happens if we try to approximate the forward rate dynamics with a finite or infinite dimensional Gaussian model. Various topics associated to this problem are touched, as the reduction of the model to a finite dimensional one, the calibration of the model to hystorical market datas in order to get the most possibly precise shape of the forward rate curve, and the so called "robustness" of the model, that is a kind of good dependence of the model from the extimated datas, in such a way that even if an agent extimates the coefficients of the model making errors, he (she) succeeds in obtaining a superhedging strategy, and that a little error on the extimation of the curve gives a little error on the price and the replication of the derivative assets. In this work we wanted to present ideas related to this topics, giving also open problems, without insisting too much on technicalities; this also because, while some parts (like the reduction of the model to a finite dimensional one) don't present technical difficulties, other ones (in particular the part on the robustness of the model) would require heavy mathematical tools, whose complexity would be parhaps sufficient for a Ph. D. thesis! Finally there is an appendix devoted to the aspect that is considered on of the main drawbacks of the Gaussian models, that is the possibility of negative rates; also in this appendix we try to present the main ideas and don't insist too much on the technical side.

In the work we proceed in this way. First we present the Musiela model: its core is the stochastic differential equation

$$\begin{cases} dr(t,x) = \left(\dfrac{\partial}{\partial x} r(t,x) + \displaystyle\sum_{n=1}^{\infty} \tau_n(t,x) \int_0^x \tau_n(t,u)\ du \right)\ dt + \displaystyle\sum_{n=1}^{\infty} \tau_n(t,x)\ dw_t^n, \quad t \geq 0 \\ \\ r(0,\cdot) \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^+) \text{ given} \end{cases}$$

where $r(t,\cdot) \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^+)$ $\mathbb{Q}$-a.s. $\forall t$, and $(w^n)_n$ are independent real Brownian motions; we often will write the equation in a shorter way:

$$\begin{cases} dr_t = (Ar_t + c_t)\ dt + \{\tau_t, dW_t\}, \quad t \geq 0 \\ r_0 \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^+) \text{ given} \end{cases}$$

We notice that the choice of τ determines entirely the evolution of the forward rate; in particular we can choose to consider a finite or infinite number of driving Brownian motions, or we can choose to consider a deterministic τ, thus obtaining a Gaussian model. We also notice that the general model is in the space $W^{1,1}_{\mathrm{loc}}(\mathbb{R}^+)$, which is a locally convex space; in this kind of spaces the theory of stochastic integration is rather difficult, while it becomes much simpler in Hilbert spaces. Though such a problem can be treated in its generality, here we present a simpler problem, that is to analyze the model supposing that the forward rates take values in some Hilbert spaces contained in $W^{1,1}_{\mathrm{loc}}(\mathbb{R}^+)$. We present some results from an existing paper [15], regarding the existence and uniqueness of a so called mild solution of the Musiela equation in the simplest case, that is when τ is deterministic and there is a finite number of driving Brownian motions; the problems is solved in two different Hilbert spaces, namely $H^1_\gamma(\mathbb{R}^+)$ and $H^1_\gamma(\mathbb{R}^+)$, in which the equation has different behaviour. We present also some results linked to the existence and (possibly) the uniqueness of an invariant measure for the equation. This last problem is linked to the calibration of the model: we will find out that an implicit calibration gives very few informations on the model, while a particular hystorical calibration succeeds in giving much more information. But each time that a calibration based on time series is used (as in our case), there is always the implicit hypothesis that the process is stationary; this is equivalent to the existence of an invariant measure for the driving equation.

This study introduces us to the main topic of this work: even if we restrict the model to the Gaussian case, the natural framework for it is $W^{1,1}_{\mathrm{loc}}(\mathbb{R}^+)$, and the choices of the two Hilbert spaces appears arbitrary; moreover, there is no apparent reason to impose that there is only a finite number of driving Brownian motions. The main question is then: is it possible to find a "natural" Hilbert space in which the equation "behaves well"? More formally, does it exist a Hilbert space and an orthonormal basis $(e_n)_n$ such that, if we project the solution of the equation on the first N components, we make a small error. The simplest case in which this happens is when the model is itself finite dimensional, so we first analyze how the Musiela equation becomes when we have the Vasicek model (the simplest Gaussian model) or more generally a finite dimensional Gaussian model in which the state variables are particular reference rates. Starting from this, we make some qualitative considerations on which could be the "right" Hilbert space for the equation, but these considerations are neither rigorous nor fully satisfactory! We finish this session with a first attempt to calculate the error we make when we approximate the "real" dynamics of the model with a finite dimensional approximation. This leaves a first series of open problems, linked to the choices of the space, of the basis, and of the criterium to use to calculate the error.

Now we come to the part about the calibration of the model: we present an implicit approach, that is classically used in the market, but doesn't give us much information; then we present two hystorical approaches, and we develop with more detail the first one; in particular, we apply the method known as "Karhunen–Loeve expansion". This method is classically used in fluidodynamics in order to find an orthonormal basis, whose vectors are known as "empirical orthonormal functions", which diagonalize the covariance operator of an infinite dimensional random variable; it turns out that this method gives nice numerical results, in the sense that few components of the basis are sufficient to describe much of the evolution of the system, and there doesn't exist another orthonormal basis that can improve the result. We present the method following [12], and we give a way to apply the method to our situation, considering a time series as a sample of an infinite dimensional random variable having as law an invariant measure of the equation. Here there is a hidden detail: in fact we always studied the Musiela model under the risk-neutral probability $\mathbb{Q}$, while we can obtain from the time series an estimation of the hystorical probability $\mathbb{P}$. This doesn't create problems, because since $\mathbb{Q}$ and $\mathbb{P}$ are equivalent, the Gaussian random variable we are studying has the same covariance both under $\mathbb{P}$ as under $\mathbb{Q}$, so we can proceed safely.

Finally we arrive to the part devoted to the robustness of the model. First of all we present a functional differential equation that gives, under no arbitrage assumptions, the price of an interest rate derivative under the hypothesis that the forward rates follow a (even non linear) Musiela model. We do this for two reason: the first is to make a link between this case and the classical finite dimensional cases (see [4], [16]) in which the price of interest rate derivatives can be derived via a PDE; besides, we can introduce the mathematical tools that we need later and an equation that is simpler, but similar, to the one known as Black–Scholes–Barenblatt equation, that we'll meet later. Then we consider the real robustness problem: we suppose that the dynamics of the forward rates is given by the Musiela equation, but that we aren't able to determine precisely the process τ, but we only know that τ takes values in some compact set $\Sigma \subseteq H^d$. This model is a quite general one, and it is an analogous of the stochastic volatility models for stock prices analyzed in literature in the last years (see for example [8] and [1]). Now we consider an operator in the market that wants to create and sell an European interest rate derivative asset with payoff $F(r_{\bar{T}})$ depending on the value in the maturity $\bar{T}$ of all the forward rates curve $r_{\bar{T}}$. In many relevant cases, the payoff F will depend not on all the curve $r_{\bar{T}}$, but only on the values of N particular bonds with maturity $T_i > \bar{T}$, so it will be of the form

$$F(B(\bar{T}, T_1), \dots, B(\bar{T}, T_N)), \quad \bar{T} < T_1 < \dots < T_N$$

From now on, we suppose that F is of the form above. In order to hedge this asset, the operator chooses to use the homogeneous Gaussian Musiela model described in Chapter 1. Then he (she) builds a self-financing strategy in this way: he (she) fixes the price C_t of the option at time t as

$$C_t = u(t, r_t)$$

where u is the solution of a nonlinear functional differential equation, namely a Hamilton-Jacobi-Bellmann type equation;

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,r) + \sup\limits_{\tau \in \Sigma} Lu(t,r) - r(0)u(t,r) = 0, \quad (t,r) \in [0,\bar{T}] \times H \\ \\ u(\bar{T}, r) = F(r), \quad r \in H \end{cases}$$

where L is a particular second order operator acting on functionals on the Hilbert space H. We remark that the finite-dimensional versions of this equation appears for the first time in one dimension in [1], and it is generalized to the finite-dimensional case in [14]; as in the previous works the Equation above will be called from now on the **Black-Scholes-Barenblatt** equation associated to the contingent claim F. To our knowledge, this is the first time that such an approach is used in hedging interest rates derivatives, so it is also the first time that such an equation is analyzed in this context. In order to hedge the claim F, the agent builds the hedging portfolio holding the quantity Δ_t^i of the zero-coupon bond $B(t,T_i)$ in the hedging portfolio Π defined in the classical delta-hedging way as the derivative of the solution of the equation with respect to the corresponding asset. Finally we define the tracking error $e_t = \Pi_t - C_t$, with $e_0 = 0$. Intuitively, the tracking error gives the error made by the operator in his (her) estimation, or better the difference between the hedging portfolio held by the agent and the option he (she) sold. We find out that if C is a $C^{1,2}([0,T) \times \mathbf{R}^n)$ solution of a suitable finite dimensional version of the BSB equation, then $(e_t)_t$ is an increasing process, that is the delta-hedging gives a superstrategy for our operator. This is the part where the open problems are more deep, not only because of the intrinsic difficulty of the study of the BSB equation in a Hilbert space, but also because all the cases of interest in Finance are such that the final payoff is not even C^1; it seems so that it would be necessary to develop some tools analogous to what viscosity solution are in finite dimensions.

The appendix is devoted to one of the main drawbacks of the Gaussian models, that is the possibility of negative rates with positive probability. First of all we concentrate on the one-dimensional Vasicek model, and calculate the probability that at some time there will be a negative forward rate; this probability being 1, we calculate the probability that there will be a negative forward rate before a fixed instant T. While we succeed to calculate exactly the first probability in terms of the solution of a ODE, we can't calculate explicitly the second probability, because it is given in terms of the solution to a PDE, that doesn't appear to have an explicit solution. Then we give two extimations of this probability, one using Kolmogorov's maximal inequality and the fact that the process $(e^{at}r_t)_t$ is a martingale, and the other one making a time change, seeing the stochastic integral appearing in the solution r_t as a Brownian motion with respect to the new time scale, and using a sort of Cameron-Martin theorem. Then, we try to do a similar work for the d-dimensional Gaussian model, even if the work in this direction is far to be complete. Anyway, the final aim of this appendix would be to extend all or some of the results found in the case of Vasicek's model to the general Gaussian Musiela model in a Hilbert space: from a mathematical point of view, we feel that this kind of problem is interesting and far from being trivial.

The work is organized as follows. In Chapter 1 there is an introduction to the Musiela model, with the main technical results concerning the equation in the spaces $H^1_\gamma(\mathbb{R}^+)$ and $H^1(\mathbb{R}^+)$. In Chapter 2 we want to evaluate when the homogeneous Gaussian Musiela model can be reduced to a finite factor model; moreover, we want to find a Hilbert space H which behaves well with respect to this interpretation. When this is not possible, we want to evaluate the error we do when we approximate a term structure having a Musiela dynamics with a finite factor model. In Chapter 3 we concentrate on the derivation of the forward curve from economical datas; in particular, we present an "implicit" approach, that derives bonds' volatility from the prices of derivatives in the market, and two "historical" ones, that derive bonds' volatility from the prices of bonds in the past. In particular, we develop one hystorical approach using the Karhunen–Loeve expansion. In Chapter 4 we present the functional differential equation satisfied by the bonds' prices under no arbitrage assumptions. Then we present the Black–Scholes–Barenblatt equation, and a finite dimensional version of this equation, and prove that in certain cases the strategy built starting from the BSB equation is a superhedging one. Finally we present the concept of extended duration and show the possible links of this concept with our model. At last in the Appendix 1 we concentrate on the positivity of interest rates; first we calculate the probability to have a negative rate for the Vasicek model, then we make the same calculations in the case of a Gaussian d-dimensional model.

The author wishes to thank Nicole El Karoui for having given the idea of finding a "right" Hilbert space for the Musiela model, which is the idea that inspired this work; Franco Flandoli for his precious advices and his suggestion to use Karhunen–Loeve expansion; the Laboratory of Probability of the University of Paris VI, which has hosted him during the revision of this work; Vincent Lacoste for interesting talkings about robustness; Marek Musiela, which in a talk suggested the idea of the robustness of his model; `root@linuz.sns.it` (that is, Massimiliano Ghilardi and Michele Mader), for having allowed him to finish this work in spite of various accidents happened to the Calculus Center of the Scuola Normale Superiore in the late June 1997; Jerzy Zabczyk for having assisted him in the part on the positivity of rates.

Chapter 1

The Musiela model

1.1 Presentation of the Musiela model

Let's introduce now the Musiela interest rate term structure model. We have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with the filtration $(\mathcal{F}_t)_{t\geq 0}$. We suppose to have only a single bond in the market, and we suppose that the price at time t of a bond expiring at time T is given by the process $(B(t,T))_{t\geq 0}$. We suppose to have a two parameters process $(r(t,x))_{t,x\geq 0}$, called **forward rate** process, in which $r(t,x)$ represents the rate at which at time t one can enter a contract (to borrow or lend money) at time $T = t + x$, for a short period of time. Then the price at time t of a bond expiring at time $T = t + x$ is given by:

$$B(t,T) = B(t,t+x) = \exp\left(-\int_0^x r(t,u)\ du\right)$$

We call **spot rate** process the process $(r(t,0))_t$; it represents the rate at which at time t one can enter a contract expiring immediately after. We also call **price progress of the savings account** the process $(\beta(t))_t$, given by:

$$\beta(t) = \exp\left(\int_0^t r(u,0)\ du\right)$$

The **actualized price** at time t of a bond expiring at time $t+x$ is given by:

$$\tilde{B}(t,T) = \frac{B(t,T)}{\beta(t)} = \exp\left(\int_0^t r(u,0)\ du - \int_0^x r(t,u)\ du\right)$$

Now we add the hypotheses that there exists a k-dimensional standard Brownian motion $(W_t)_t$ adapted to $(\mathcal{F}_t)_t$, and there exist two two-parameter progressively measurable processes $(\alpha(t,x))_{t,x\geq 0}$, $(\tau(t,x))_{t,x\geq 0}$ such that $\forall x \geq 0, (\alpha(t,x))_t$ has trajectories in $L^1([0,+\infty);\mathbb{R})$, $(\tau^*(t,x))_t$ has trajectories in $L^2([0,+\infty);\mathbb{R}^k)$ and such that the forward rate satisfies the following stochastic differential equation:

$$\begin{cases} dr(t,x) = \alpha(t,x)\ dt + \tau^*(t,x)\ dW_t \\ r(0,x) \text{ given} \end{cases} \tag{1.1}$$

A natural thing to ask is that the actualized bond price process $(\tilde{P}(t,T))_t$ is a martingale under a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$, which is called **risk-neutral** probability. To this purpose, we cite the:

Theorem 1.1 *: if we model the forward rate with an equation of the form (1.1), then the following facts are equivalent:*

1. *$\forall T > 0$ the process $(\tilde{B}(t,T))_t$ is a local martingale with respect to $\mathbb{Q}$.*
2. *the process $(r_t(x))_t$ is such that $\forall t \in \mathbb{R}^+$ the application $x \to r(t,x)$ $\mathbb{Q}$-a.s. belongs to $AC(\mathbb{R}^+)$ and:*

$$\alpha(t,x) = \frac{\partial}{\partial x}\left(r(t,x) + \frac{1}{2}\left| \int_0^x \tau(t,u)\ du \right|^2 \right)$$

If one of the conditions above is verified, then for all $T > 0$ the dynamics of $B(t,T)$ is given by:

$$\begin{cases} dB(t,T) &= B(t,T)\left(r(t,0)\ dt + \Gamma(t,T)\ dW_t\right), \qquad t \le T \\ B(T,T) &= 1 \end{cases}$$

where $\Gamma(t,T) = -\int_0^{T-t} \tau(t,u)\ du$.

Proof. See [13]. $\diamond$

The forward rate model obtained in this way is then the following:

$$\begin{cases} dr(t,x) = \left(\frac{\partial}{\partial x} r(t,x) + \tau^*(t,x) \int_0^x \tau(t,u)\ du \right)\ dt + \tau^*(t,x)\ dW_t \\ \\ r(0,x) \text{ given} \end{cases} \tag{1.2}$$

We notice that this model is very general, and for particular choices of the process τ we can obtain models already studied in literature: for example if $\tau(x) = \sigma e^{-ax}$, we obtain the Vasicek model, if $\tau(x) = \sigma$, we obtain the Ho-Lee model, and if $\tau(x) = \sigma c(x)\sqrt{r(t,0)}$, for a particular deterministic function c, we obtain the CIR model; the first two models belong to the class of Gaussian models, which we will treat with more detail; we notice that it is possible to obtain also non-Gaussian models, as the CIR or others.

Now we will study this model using more powerful instruments. We now view the process $(r(t,x))_{t,x}$ no more as a real two parameter process, but as a one parameter process taking values in the absolutely continuous functions space on $\mathbb{R}^+$, which will be identified with the Sobolev space $W^{1,1}_{\text{loc}}(\mathbb{R}^+)$, and we write the equation (1.2) as a stochastic differential equation taking values in $W^{1,1}_{\text{loc}}(\mathbb{R}^+)$.

Taking:

$$A = \frac{\partial}{\partial x}, \quad \tau_t(x) = \tau(t,x), \quad c_t(x) = \tau_t^*(x) \int_0^x \tau_t(u)\ du$$

then equation (1.2) becomes the following equation taking values in $W^{1,1}_{\text{loc}}(\mathbb{R}^+)$:

$$\begin{cases} dr_t = (Ar_t + c_t)\ dt + \tau_t^*\ dW_t \\ r_0 \in W^{1,1}_{\text{loc}}(\mathbb{R}^+) \end{cases} \tag{1.3}$$

If we make the further hypotheses that $(\mathcal{F}_t)_t$ is the completion of the natural filtration of the brownian motion $(W_t)_t$, and that $|\tau_t(x)| \leq M(T)\ \forall t, x$ such that $t + x \leq T$ $\mathbb{Q}$-a.s., then the process $(\tilde{B}(t,T))_t$ is a martingale; then the hypotheses of the martingale representation theorem hold, and so every integrable random variable (that is every contingent claim) can be represented as the sum of his expectation, which is the arbitrage free price of the claim, and of a stochastic integral with respect to $(W_t)_t$. Besides, we can construct a self-financing portfolio strategy based on $B(\cdot, T)$ and on $\beta(\cdot)$, which simulates the claim (see [13]).

1.2 Formulation of the problem

Now we make some hypotheses under which this equation will have explicit solutions. We suppose that the process $(\tau_t)_t$ is identically equal to a real k-valued deterministic function $\tau(x)$ belonging to $W^{1,1}_{\text{loc}}(\mathbb{R}^+)$. Then the equation becomes:

$$\begin{cases} dr_t = (Ar_t + c)\ dt + \tau^*\ dW_t \\ r_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H) \end{cases} \tag{1.4}$$

This is a Langevin equation, which, under right hypotheses, has an explicit solution.

We notice that the general model is in the space $W^{1,1}_{\text{loc}}(\mathbb{R}^+)$, which is a locally convex space; in this kind of spaces the theory of stochastic integration is rather difficult, while it becomes much simpler in Hilbert spaces. Though such a problem can be treated in its generality, here we present a simpler problem, that is to analyze the model supposing that the forward rates take values in some Hilbert spaces contained in $W^{1,1}_{\text{loc}}(\mathbb{R}^+)$.

We now present a study of the equation in two particular separable Hilbert spaces contained in $W^{1,1}_{\text{loc}}(\mathbb{R}^+)$; this study is contained in ([15]), and here we present only the main results. We use the theory of the stochastic integration in separable Hilbert spaces contained in [5] which is somewhat complete. In particular, we will study the equation in the spaces:

$$H^1_\gamma(\mathbb{R}^+) = W^{1,2}_\gamma(\mathbb{R}^+) = \left\{u : \mathbb{R}^+ \to \mathbb{R} \mid u \in L^2_\gamma(\mathbb{R}^+), u' \in L^2_\gamma(\mathbb{R}^+)\right\}$$

$$H^1(\mathbb{R}^+) = W^{1,2}(\mathbb{R}^+) = \left\{u : \mathbb{R}^+ \to \mathbb{R} \mid u \in L^2(\mathbb{R}^+), u' \in L^2(\mathbb{R}^+)\right\}$$

where u' is the weak derivative of u with respect to x, and:

$$L^2_\gamma(\mathbb{R}^+) = \left\{u : \mathbb{R}^+ \to \mathbb{R}\ \middle|\ u \text{ meas. and s.t. } \int_0^{+\infty} u^2(x) e^{-\gamma x}\ dx < +\infty\right\}$$

$$L^2(\mathbb{R}^+) = \left\{u : \mathbb{R}^+ \to \mathbb{R}\ \middle|\ u \text{ meas. and s.t. } \int_0^{+\infty} u^2(x)\ dx < +\infty\right\}$$

H^1_γ is a Hilbert space with the scalar product:

$$\langle f, g\rangle_{H^1_\gamma} = \gamma \int_0^{+\infty} f(x)g(x)e^{-\gamma x}\ dx + \gamma \int_0^{+\infty} f'(x)g'(x)e^{-\gamma x}\ dx \tag{1.5}$$

and H^1 is a Hilbert space with the scalar product:

$$\langle f, g \rangle_{H^1} = \int_0^{+\infty} f(x)g(x)\ dx + \int_0^{+\infty} f'(x)g'(x)\ dx \tag{1.6}$$

In all the spaces cited above we will examine the following problems:

- to find an explicit mild solution
- to find the marginal law of the solutions r_t $\forall t > 0$ and the joint law of the process $(r_t)_{t>0}$
- existence, uniqueness (when it holds) and characterization of the invariant measures for the solution process

1.3 Preliminaries about Langevin equations in Hilbert spaces

Now we give a bit of general theory about the Langevin equation on separable Hilbert spaces.

We suppose we have a separable Hilbert space H, such that $r_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H)$, $c \in H$, that $A : D(A) \to H$ is the infinitesimal generator of a C^0 semigroup of linear continuous operators (see appendix), that $(W_t)_t$ is a Q-Wiener process in H, that is a $(\mathcal{F}_t)_t$-stochastic process such that:

1. $W_0 = 0$
2. $(W_t)_t$ has continuous trajectories
3. $\forall s < t$, $W_t - W_s$ is independent of $\mathcal{F}_s$
4. $\forall s < t$, $W_t - W_s$ has gaussian law equal to $N(0, (t-s)Q)$, where $Q : H \to H$ is linear, bounded, symmetric, positive and such that $\mathrm{Tr}\ Q < +\infty$

and that, if we call $Q_t = S_t Q S_t^*$, then $\int_0^t \mathrm{Tr}\ Q_u\ du < +\infty$. Then the equation

$$\begin{cases} dr_t = (Ar_t + c)\ dt + dW_t \\ r_0 \in L^2(\Omega, \mathcal{F}_0, \mathbb{P}; H) \end{cases} \tag{1.7}$$

has exactly one mild solution, which is the Ornstein-Uhlenbeck process in the Hilbert space H:

$$r_t = S_t r_0 + \int_0^t S_{t-u} c\ du + \int_0^t S_{t-u}\ dW_u$$

If r_0 is a Gaussian random variable, then the solution is a Gaussian process with functional mean:

$$\mathbb{E}[r_t] = S_t \mathbb{E}[r_0] + \int_0^t S_{t-u} c\ du \tag{1.8}$$

and functional covariance (if $t \leq v$):

$$\text{Cov }(r_t, r_v) = S_v \text{Cov }(r_0, r_0) S_t^* + \int_0^t S_{v-u} Q S_{t-u}^* \, du \tag{1.9}$$

This means that $\forall f, g \in H$:

$$\mathbb{E}[\langle f, r_t \rangle_H] = \langle f, S_t \mathbb{E}[r_0] \rangle_H + \langle f, \int_0^t S_{t-u} c \, du \rangle_H$$

$$\begin{aligned}
\text{Cov }(\langle f, r_t \rangle_H, \langle g, r_v \rangle_H) &= \langle (S_v \text{Cov }(r_0, r_0) S_t^*) f, g \rangle_H + \int_0^t \langle S_{v-u} Q S_{t-u}^* f, g \rangle_H = \\
&= \langle \text{Cov }(r_0, r_0) S_t^* f, S_v^* g \rangle_H + \int_0^t \langle Q S_{t-u}^* f, S_{v-u}^* g \rangle_H \, du
\end{aligned}$$

For a proof of these facts, see [5].

Applying these results to our equation, since $A = \frac{\partial}{\partial x}$, the natural semigroup is the translation semigroup $(S_t)_t$ such that:

$$S_t f(x) = f(t + x), \quad t, x \geq 0$$

Moreover, the Wiener process is of the kind $(\sum_{i=1}^k \tau_i W_t^i)_t$, where $\tau \in H^k$ and $(W_t)_t$ is a real k-dimenional Wiener process; then $Q = \sum_{i=1}^k \tau_i \otimes \tau_i$, in the sense that:

$$\langle Qf, g \rangle = \sum_{i=1}^k \langle \tau_i, f \rangle_H \langle \tau_i, g \rangle_H \quad \forall f \in H$$

Then Tr $Q = \sum_{i=1}^k \|\tau_i\|_H^2 < +\infty$, and Tr $Q_t = \sum_{i=1}^k \|S_t \tau_i\|^2$. Since, if $(S_t)_t$ is a C^0-semigroup in H, then it exist $M, \alpha \in \mathbb{R}^+$ such that $\|S_t\|_{L(H,H)} \leq M \mathbb{E}^{\alpha t}$, then the condition

$$\int_0^t \text{Tr } Q_u \, du = \sum_{i=1}^k \int_0^t \|S_u \tau_i\|^2 \, du \leq \sum_{i=1}^k \|\tau_i\|^2 \int_0^t M \mathbb{E}^{\alpha u} \, du < +\infty$$

is always verified. The solution becomes:

$$\begin{aligned}
r_t(x) &= S_t r_0(x) + \int_0^t S_{t-u} c(x) \, du + \sum_{i=1}^k \int_0^t S_{t-u} \tau_i(x) \, dW_u^i = \\
&= r_0(x+t) + \sum_{i=1}^k \int_0^t \tau_i(x+t-u) \left(\int_0^{x+t-u} \tau_i(v) \, dv \right) \, du + \\
&\quad + \int_0^t \tau_i(x+t-u) \, dW_u^i = \\
&= r_0(x+t) + \frac{1}{2} \sum_{i=1}^k \left(\left(\int_0^{x+t} \tau_i(u) \, du \right)^2 - \left(\int_0^x \tau_i(u) \, du \right)^2 \right) + \\
&\quad + \sum_{i=1}^k \int_0^t \tau_i(x+t-u) \, dW_u^i
\end{aligned} \tag{1.10}$$

The solution is a Gaussian process with functional mean (1.8) and functional covariance (1.9); this means that $\forall f, g \in H$:

$$\mathbb{E}[\langle r_t, f\rangle_H] = \langle r_0(\cdot + t), f\rangle_H + \frac{1}{2}\sum_{i=1}^{k}\left\langle \left(\left(\int_0^{\cdot + t} \tau_i(u)\ du\right)^2 - \left(\int_0^{\cdot} \tau_i(u)\ du\right)^2\right), f\right\rangle_H$$

$$\begin{aligned}
&\text{Cov }(\langle r_t, f\rangle_H, \langle r_v, g\rangle_H) = \\
&= \langle S_v \text{Cov }(r_0, r_0) S_t^* f, g\rangle_H + \sum_{i=1}^{k}\left\langle \left(\int_0^{s} \tau_i(\cdot + t - u)\otimes \tau_i(\cdot + v - u)\ du\right) f, g\right\rangle_H = \\
&= \langle \text{Cov }(r_0, r_0) S_t^* f, S_v^* g\rangle_H + \sum_{i=1}^{k}\int_0^{t} \langle \tau_i(\cdot + t - u), f\rangle_H \langle \tau_i(\cdot + v - u), g\rangle_H\ du
\end{aligned}$$

In some applications it can be useful to know only mean and covariance of particular rates and not necessarily all the forward rates curve (for example, to know the spot rate at time t, it is sufficient to know only $r_t(0)$). This can be done with the theory given before: if the Hilbert space H is a subset of the space of the continuous functions $C^0([0, +\infty))$, then $\forall t, v, x, y \geq 0$:

$$\mathbb{E}[r_t(x)] = \mathbb{E}[r_0(x+t)] + \frac{1}{2}\sum_{i=1}^{k}\left(\left(\int_0^{x+t} \tau_i(u)\ du\right)^2 - \left(\int_0^{x} \tau_i(u)\ du\right)^2\right)$$

$$\text{Cov }(r_t(x), r_v(y)) = \text{Cov }(r_0(t+x), r_0(v+y)) + \sum_{i=1}^{k}\int_0^{t} \tau_i(x+t-u)\tau_i(y+v-u)\ du$$

1.4 The space H^1_γ

Theorem 1.2 *: if $\tau_i \in H^1_\gamma \cap H^1 \cap L^4_\gamma$ $\forall i = 1, \ldots, k$ and $r_0 \in L^2_\gamma$, then it exists a unique mild solution of the equation (1.4), given by Eq. (1.10). The solution is a Gaussian process with functional mean (1.8) and functional covariance (1.9).*

Theorem 1.3 *: given the equation (1.4) in the Hilbert space H^1_γ, a necessary and sufficient condition to have an invariant measure is $\tau_i \in H^1_\gamma(\mathbb{R}^+) \cap H^1(\mathbb{R}^+)$ $\forall i = 1, \ldots, k$, and there exist infinitely many invariant measures. In particular, the measures of the kind:*

$$\delta_{b^*(\cdot)+b_0} * N(0, Q_\infty), \quad b_0 \in \mathbb{R}$$

$$\mu_f * N(0, Q_\infty), f \in H^1_\gamma \text{ periodical}$$

are invariant measures, where:

$$Q_\infty = \sum_{i=1}^{k}\int_0^{+\infty} \tau_i(\cdot + u)\otimes \tau_i(\cdot + u)\ du \tag{1.11}$$

$$b^*(x) = -\sum_{i=1}^{k} \int_0^x \tau_i(u) \int_0^u \tau_i(v)\ dv\ du = -\frac{1}{2}\sum_{i=1}^{k} \left(\int_0^x \tau_i(u)\ du\right)^2 \tag{1.12}$$

$f \in H^1_\gamma$ is periodical with period T and μ_f is the measure induced by the uniform density on $[0,T]$ via the application $t \to f(t+\cdot)$.

Lemma 1.4 *If $\tau \in H^1_\gamma \cap H^1$, and the initial data is a deterministic $r_0 \in H^1_\gamma$, such that $\lim_{x\to+\infty} r_0(x) = r_0(\infty) \in \mathbb{R}$, $\lim_{x\to+\infty} r_0'(x) = 0$, then the marginal law of r_t converges for $t \to +\infty$ to the invariant measure $N\left(r_0(\infty) + b_0 + b(x),\ Q_\infty\right)$, where b and Q_∞ are given respectively by (1.12) and (1.11), and b_0 is given by*

$$b_0 = \frac{1}{2}\sum_{i=1}^{k} \left(\int_0^{+\infty} \tau_i(u)\ du\right)^2 . \tag{1.13}$$

1.5 The space H^1

Theorem 1.5 *: if $\forall i = 1,\dots,k$, $\tau_i(x) \in H^1(\mathbb{R}^+) \cap L^4(\mathbb{R}^+)$, $\sqrt{x}\tau_i(x) \in L^2(\mathbb{R}^+)$, $\sqrt{x}\tau_i'(x) \in L^2(\mathbb{R}^+)$ and $r_0 \in H^1(\mathbb{R}^+)$, then the equation (1.4) has an unique solution in $H^1(\mathbb{R}^+)$, given by (1.10), which is a gaussian process with functional mean (1.8) and functional covariance (1.9).*

Theorem 1.6 *: given the equation (1.4) in the Hilbert space H^1, necessary and sufficient conditions to have an invariant measure are:*

$$\forall i = 1,\dots,k : \begin{cases} \tau_i(x) \in H^1(\mathbb{R}^+) \cap L^1(\mathbb{R}^+) \\ \sqrt{x}\tau_i(x) \in L^2(\mathbb{R}^+), \qquad \sqrt{x}\tau_i'(x) \in L^2(\mathbb{R}^+) \\ \int_x^{+\infty} \tau_i(u)\ du \in L^2(\mathbb{R}^+) \end{cases}$$

in these hypotheses only the following invariant measure exists:

$$\delta_{b^*(\cdot)+b_0} * \delta_0 * N(0, Q_\infty) = N(b^*(\cdot) + b_0, Q_\infty)$$

where b^* is given by (1.12), b_0 is given by (1.13), and Q_∞ is given by (1.11).

Chapter 2

The interest rates

2.1 The Musiela model in the Gaussian case

We suppose that the term structure of interest rates has the dynamics described from the following SPDE:

$$\begin{cases} dr_t(x) = (Ar_t(x) + c_t(x))\ dt + \tau_t(x)\ dW_t \\ r_0(\cdot) \in H \end{cases} \tag{2.1}$$

where $A = \frac{\partial}{\partial x}$, $(W_t)_t$ is a d-dimensional Brownian motion, H is a separable Hilbert space of functions from $\mathbb{R}^+$ to $\mathbb{R}$ (for example, $L^2(\mathbb{R}^+)$ or $H^1(\mathbb{R}^+)$), $\tau_t(x) = (\tau_t^1(x), \ldots, \tau_t^d(x))$ is a H^d-valued adapted process, and $c(x) = \sum_{i=1}^d \tau_t^i(x) \int_0^x \tau_t^i(u)\ du$ is a H-valued process. A solution for Eq. (2.1) has to be intended in the mild sense, as presented in [5].

If we make the hypothesis that τ is a deterministic function of x, and it belongs to H, then Eq. (2.1) has the unique mild solution:

$$r_t(x) = r_0(t+x) + \int_0^t c(x+t-s)\ ds + \int_0^t \tau(x+t-s)\ dW_s \tag{2.2}$$

2.2 From the Musiela model to the Vasicek model

We suppose that the interest rates follow a Vasicek dynamics: this means that the spot rate $r_t(0)$ is Markov and follows the dynamics:

$$dr_t(0) = a(b - r_t(0))\ dt + \sigma\ dW_t \tag{2.3}$$

where $(W_t)_t$ is a real Brownian motion, and $a, b, \sigma \in \mathbb{R}^+$.

Theorem 2.1 *In order for the spot rate to follow a Vasicek dynamic, it is necessary and sufficient that $\tau(\cdot)$ and $r_0(\cdot)$ are of the form:*

$$\tau(x) = e^{-ax}\sigma, \quad r_0(t) = e^{-at}r_0(0) + (1 - e^{-at})b \tag{2.4}$$

Proof. If the spot rate $r_t(0)$ follows a Vasicek dynamics, then we have two possible expressions for it:

$$\begin{aligned} r_t(0) &= e^{-at}r_0(0) + \int_0^t e^{-a(t-s)}ab\ ds + \int_0^t e^{-a(t-s)}\sigma\ dW_s = \\ &= r_0(t) + \int_0^t \tau(t-s)\ dW_s \end{aligned} \tag{2.5}$$

In order for the two expressions to give the same quantity, we must impose that $\tau(x) = e^{-ax}\sigma$; moreover, $r_0(\cdot)$ can assume only the particular form given in the thesis. In fact:

$$\begin{aligned} r_0(t) &= e^{-at}r_0(0) + \int_0^t e^{-a(t-s)}ab\ ds = e^{-at}r_0(0) + \left[-be^{-as}\right]_0^t = \\ &= e^{-at}r_0(0) + b(1-e^{-at}) \end{aligned}$$

Conversely, if Eqs. (2.4) hold, then $r_t(0)$ is given by (2.5), which is the solution of (2.3). Since the knowledge of the spot rate is sufficient to derive the entire term structure, whe have the thesis. ◇

2.3 From the Musiela model to a state variables model

Now we suppose that the interest rates follow a particular state variables model: we suppose that all the instantaneous forward rates can be written in the form:

$$r_t(\theta) = \sum_{i=1}^{n} \alpha_i(\theta)\cdot r_t(\theta_{i-1}) + \beta(\theta) = \alpha(\theta)\cdot R_t + \beta(\theta) \tag{2.6}$$

where $R_t = (r_t(0), r_t(\theta_1), \ldots, r_t(\theta_{n-1}))$ are particular reference rates, with $\theta_i = i\theta_1$, and where R_t has the dynamics:

$$dR_t = (C - AR_t)\ dt + \sigma\ dW_t \tag{2.7}$$

where $C \in \mathbb{R}^n$, and $A, \sigma \in \mathbb{R}^{n\times n}$. This corresponds to say that the only relevant quantities to know are the rates in R_t, and all the other rates can be derived from these through the weight coefficients $\alpha_i(\theta)$. It can be proved (see [6]) that the matrix A satisfies:

$$e^{-A\theta_1} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & 1 \\ \mu_1 & \ldots & \ldots & \ldots & \mu_n \end{pmatrix}$$

Theorem 2.2 *In order for the rates $(R_t)_t$ to be state variables as in the model (2.6), it is necessary and sufficient that $\tau(\cdot)$ and $r_0(\cdot)$ are of the form:*

$$\tau_j(x) = (e^{-Ax}\sigma)_{1j} \tag{2.8}$$

$$r_0(t) = e^{-At}R_0|_1 + \int_0^t e^{-A(t-s)}C|_1 \, ds - \int_0^t c(t-s) \, ds \tag{2.9}$$

Besides, the coefficients $\alpha(\cdot)$ must be of the form:

$$\alpha(\theta) = (1, 0, \ldots, 0)e^{-A\theta}$$

Proof. If the rates in R are state variables which satisfy (2.6) and (2.7), then we have two expressions for the solution:

$$\begin{aligned} r_t(\theta_i) &= e^{-At}R_0|_{i+1} + \int_0^t (e^{-A(t-s)}C)|_{i+1} \, ds + \left(\int_0^t e^{-A(t-s)}\sigma \, dW_s\right)|_{i-1} = \quad (2.10) \\ &= r_0(\theta_i + t) + \int_0^t c(\theta_i + t - s) \, ds + \int_0^t \tau(\theta_i + t - s) \, dW_s \end{aligned}$$

where the symbol $x|_i$ indicates the i-th component of the vector $x \in \mathbb{R}^n$. In order for the two expressions to give the same quantity, we must impose that $\tau(\theta_i + x) = (e^{-ax}\sigma)|_i$, that is:

$$\tau_j(x) = (e^{-Ax}\sigma)_{1j}$$

Now, let's consider the rates with a maturity different from θ_i: we have

$$\begin{aligned} r_t(\theta) &= r_0(\theta + t) + \int_0^t c(\theta + t - s) \, ds + \int_0^t \tau(\theta + t - s) \, dW_s = \\ &= \alpha(\theta)e^{-At}R_0 + \int_0^t \alpha(\theta)e^{-A(t-s)}C \, ds + \int_0^t \alpha(\theta)e^{-A(t-s)}\sigma \, dW_s + \beta(\theta) \end{aligned}$$

From this expression, we can derive the right weight coefficients:

$$\tau(\theta + t - s) = \alpha(\theta)e^{-A(t-s)}\sigma = e^{-A(\theta+t-s)}\sigma|_1$$

so:

$$\alpha(\theta) = (1, 0, \ldots, 0)e^{-A\theta}$$

Conversely, if τ and r_0 satisfy (2.8) and (2.9), then R is given by (2.10), which is the solution of (2.7). Once again, since the knowledge of the spot rate is sufficient to derive the entire term structure, we have the thesis. ◇

We notice that in this model, as in the Vasicek model, $r_0(\cdot)$ can assume only particular forms, though if it can have a more complex structure than in the Vasicek model.

2.4 Considerations on the Hilbert space H

If we want to use the Musiela model in a state variable framework, then from the comparison of the models above we can derive qualitative considerations on the Hilbert space H and on its norm:

- the best kind of norm is an "homogeneous" one, in which no particular maturity has a particular weight (for example, the norms defined by the measures dx or $\gamma e^{-\gamma x}\,dx$, with $\gamma > 0$). In fact, while for the state variables model the relevant maturities are θ_i, for the Musiela model all the maturities are relevant, because of the particular semigroup involved in the equation (the translation semigroup $S_t f = f(t + \cdot)$): if the norm were not homogeneous, the semigroup would move this weight in the Musiela model in a way that would be incoherent with the way the state variables model behave.

- if we have a infinite horizon (that is, if we search a solution for all $t \in \mathbb{R}^+$), then we need all the forward curve $r_0(t), t \in \mathbb{R}^+$; we notice that a flat structure of interest rates $r_0(t) = r\ \forall t$ doesn't belong to L^2 (or H^1), while it belongs to L^2_γ (or H^1_γ). Moreover, we have to choose well γ: in fact, for the equation to have sense, we must assume this condition on γ:

$$\gamma > \|e^{-A\theta_1}\|_\infty = \max\{\Re\lambda \mid \lambda \text{ eigenvalue of } e^{-A\theta_1}\}$$

- if we have a finite horizon ($t \in [0, T]$), then we only need the forward curve $r_0(t), t \in [0, 2T]$, and the space L^2 (as H^1) allows a flat structure of interest rates (such as the spaces L^2_γ, H^1_γ), so we don't have particular limitations.

2.5 Structuring the Musiela model with respect to state variables

Now we want to impose on the Hilbert space a structure that gives a good interpretation in terms of the state variables models seen above.

2.5.1 The Vasicek model

First of all we notice that the particular τ found for the Musiela model is stable with respect to a translation; this tells us that the noise in the Hilbert space will be always in the component $[e^{-au}]$, where the symbol $[\cdot]$ indicates the linear subspace generated by the argument in H.

We can reduce the Musiela model to the Vasicek model in this way: we can find a complete orthonormal basis $(e_i)_i$ in H, such that $[e_1(u)] = [e^{-au}]$, and $[e_1(u), e_2(u)] = [e^{-au}, r_0]$: so with the first two components of our Hilbert space we can represent all the structure given by the Vasicek model. Then we can define a Hilbert seminorm on H that gives weight 1 to e_1 and e_2, and weight 0 to all the other components:

$$\|\!|\cdot|\!\| = \langle e_1, \cdot\rangle + \langle e_2, \cdot\rangle$$

If we quotientate H with respect to the seminorm $\|\cdot\|$, we obtain a new Hilbert space $\bar{H}$, endowed with the norm $\|\cdot\|$, which is isomorphic to $\mathbb{R}^2$; then the initial data r_0 can be decomposed in its component in $\bar{H}$.

2.5.2 The general state variables model

We notice that, if A is diagonalizable, then the particular τ_j found for the Musiela model in this case are linear combinations of exponential functions:

$$\tau_j(u) = \sum_{i=1}^{N} a_{ij} e^{-\lambda_i u}, \qquad \text{where } (\lambda_i)_i \text{ are the eigenvalues of } A$$

so the τ_j are stable with respect to a translation; this tells us that the noise in the Hilbert space will be always in the subspace $[e^{-\lambda_i u}]_{i=1,\dots,N}$.

As before, we can reduce the Musiela model to the state variables model in this way: we can find a complete orthonormal basis $(e_i)_i$ in H, such that

$$\begin{aligned}
[e_i(u)]_{i=1,\dots,n} &= [e^{-\lambda_i u}]_{i=1,\dots,n} \quad \forall n \leq N \\
[e_i(u)]_{i=1,\dots,N+k} &= [e^{-\lambda_1 u}, \dots, e^{-\lambda_N u}, e^{-2\lambda_1 u}, \dots, e^{-2\lambda_k u}] \quad \forall k = 1, \dots, N \\
[e_i(u)]_{i=1,\dots,2N+1} &= [e^{-\lambda_1 u}, \dots, e^{-\lambda_N u}, e^{-2\lambda_1 u}, \dots, e^{-2\lambda_N u}, r]
\end{aligned}$$

We add the functions $e^{-2\lambda_i u}$ in order to be able to represent the vector $c(u)$.

We define a Hilbert seminorm on H that gives weight 1 to $e_i, i = 1, \dots, 2N+1$, and weight 0 to all the other components:

$$\|\cdot\| = \sum_{i=1}^{2N+1} \langle e_i, \cdot \rangle$$

If we quotientate H with respect to the seminorm $\|\cdot\|$, we obtain a new Hilbert space $\bar{H}$, endowed with the norm $\|\cdot\|$, which is isomorphic to $\mathbb{R}^{2N+1}$; then the initial data r_0 can be decomposed in its components in $\bar{H}$. Moreover, the covariance σ of the rates R_t influences τ and the a_i.

2.6 Approximating the Musiela model with state variables

Now we suppose that the term structure of interest rates has the dynamics described from the following SPDE:

$$\begin{cases} dr_t(x) = (Ar_t(x) + c(x))\ dt + \tau(x)\ dW_t \\ \\ r_0(\cdot) \in H\ \ (= L^2_{(\gamma)}, H^1_{(\gamma)}) \end{cases} \tag{2.11}$$

where:

- $A = \dfrac{\partial}{\partial x}$,

- the object $\tau(x)\ dW_t$ represents the integration with respect to

$$\sum_{n=1}^{\infty} \tau_n e_n(\cdot) w_t^n$$

where $\sum_{n=1}^{\infty} \tau_n^2 < +\infty$, $\{(w_t^n)_t\}_n$ are real independent Brownian motions, $(e_n)_n$ is a complete orthonormal basis of H such that:

$$[e_i(u)]_{i=1,\dots,n} = [e^{-\lambda_i u}]_{i=1,\dots,n}\ \forall n \in \mathbb{N}$$

where the λ_i are particular numbers (see section 7),

- c is given by:

$$c(x) = \sum_{n=1}^{\infty} \tau_n e_n(x) \int_0^x \tau_n e_n(u)\ du$$

then the initial data r_0 can be decomposed on the ONB $(e_n)_n$:

$$r_0(x) = \sum_{n=1}^{\infty} r_n e_n(x)$$

If $r_k = \tau_k = 0\ \forall k \geq n$, then we obtain the state variables model seen before.

Now we suppose to modelize (making a mistake) the forward rates with the model given by Eq. (2.7). Then we try to approximate the term structure given by the Musiela model by a term structure $(\tilde{r}_t)_t$ in this way:

$$\begin{aligned} \tilde{r}_t(\theta_i) &= R_t^{i+1} \quad \forall i = 0, \dots, n-1 \\ \tilde{r}_t(\theta) &= \sum_{i=1}^{n} \alpha(\theta)\tilde{r}_t(\theta_i) + \beta(\theta) = \alpha(\theta) \cdot R_t + \beta(\theta) \quad \forall \theta \neq \theta_i \end{aligned}$$

Now we want to compare the "real" rates $(r_t)_t$ with the approximating ones $(\tilde{r}_t)_t$.

Theorem 2.3 *If we build σ, A, R_0 and C such that*

$$\sigma_{ij} = \tau_j(\theta_i) \tag{2.12}$$

$$-AR_0 + C = (r_0'(\theta_i) + c(\theta_i))_i \tag{2.13}$$

then the error $r_t(\theta) - \tilde{r}_t(\theta)$ goes to 0 as $N \to \infty$ $\mathbb{Q}$-a.s. $\forall t, \theta$.

Proof. The relations above are obtained "inverting" the relations of Theorem 2.2; in fact, relation (2.8) is equivalent to $\tau_j(\theta_i + x) = (e^{-Ax}\sigma)_{ij}$, and letting $x = 0$, we obtain relation (2.12). Relation (2.9) is equivalent to

$$e^{-At}R_0|_{i+1} + \int_0^t e^{-As}C|_{i+1}\ ds = r_0(\theta_i + t) + \int_0^t c(\theta_i + s)\ ds$$

Deriving these relations with respect to t we obtain:

$$-Ae^{-At}R_0 + e^{-At}C = (r_0'(\theta_i + t) + c(\theta_i + t))_i$$

which, letting $t = 0$, gives us relation (2.13). Under these conditions, the error we make on the rates with maturity θ_i is:

$$\begin{aligned}
r_t(\theta_i) - \tilde{r}_t(\theta_i) &= r_0(\theta_i + t) + \int_0^t c(\theta_i + t - s)\ ds + \int_0^t \tau(\theta_i + t - s)\ dW_s + \\
&\quad - e^{-At}R_0|_{i-1} - \int_0^t (e^{-A(t-s)}C)|_{i-1}\ ds - \left(\int_0^t e^{-A(t-s)}\sigma\ dW_s\right)|_{i-1} = \\
&= \sum_{n=N+1}^{\infty} r_n e_n(\theta_i + t) + \int_0^t \sum_{n=N+1}^{\infty} \tau_n^2 e_n(\theta_i + t - s) \int_0^{\theta_i + t - s} e_n(u)\ du + \\
&\quad + \int_0^t \sum_{n=N+1}^{\infty} \tau_n e_n(\theta_i + t - s)\ dw_s^n
\end{aligned}$$

and for the others is:

$$\begin{aligned}
& r_t(\theta) - \tilde{r}_t(\theta) = r_0(\theta + t) + \int_0^t c(\theta + t - s)\ ds + \int_0^t \tau(\theta + t - s)\ dW_s - \beta(\theta) + \\
& \quad - \sum_{i=1}^{N} \alpha_i(\theta) \left(e^{-At}R_0|_{i-1} - \int_0^t (e^{-A(t-s)}C)|_{i-1}\ ds - \left(\int_0^t e^{-A(t-s)}\sigma\ dW_s\right)|_{i-1} \right) = \\
& = \sum_{n=N+1}^{\infty} r_n e_n(\theta_i + t) + \int_0^t \sum_{n=N+1}^{\infty} \tau_n^2 e_n(\theta_i + t - s) \int_0^{\theta_i + t - s} e_n(u)\ du + \\
& \quad + \int_0^t \sum_{n=N+1}^{\infty} \tau_n e_n(\theta_i + t - s)\ dw_s^n
\end{aligned}$$

this doesn't surprise us, because we can see the state variables model as a particular Musiela model driven by a finite dimensional Brownian motion. The consequence is that the error $|r_t(\theta) - \tilde{r}_t(\theta)|$ goes to 0 as N goes to infinity, and the speed of convergence depends from the coefficients $(r_n)_n$ and $(\tau_n)_n$. $\diamond$

2.7 Open problems

- how to choose the "right" numbers λ_i? In the state variables model, they were the eigenvalues of A, so a natural choice would be that the λ_i are the eigenvalues of the matrix A in the "wrong" model.
- for a given orthonormal basis $e_1, e_2, \dots$ in a given Hilbert space H (for example $L^2(\mathbb{R}^+)$, $L^2_\gamma(\mathbb{R}^+)$, $H^1(\mathbb{R}^+)$, $H^1_\gamma(\mathbb{R}^+)$, $H^1([0,T])$, etc.), analyse the "distance" between

$r_t, t \leq T$ and $r_t^N, t \leq T$, where

$$r_t^N = \sum_{n=1}^{N} \langle r_t, e_n \rangle e_n, \quad t \leq T$$

The distance can be for example

1. $\sup_{t \leq T} \mathbb{E}\left[\|r_t - r_t^N\|_H^2\right]$
2. $\mathbb{E}\left[\sup_{t \leq T} \|r_t - r_t^N\|_H^2\right]$
3. $\sup_{t \leq T} \mathbb{P}\left\{\|r_t - r_t^N\|_H^2 \geq \delta\right\}$
4. the prices of some interest rate derivatives (such as call options on zero coupon bonds, caps, swaptions or futures contracts)

The distance will depend on the choice of the basis $e_1, e_2, \ldots$ and on the coefficients of the equation c, τ, that is, due to the particular structure of the model, only on τ.

- analyse the distance with respect to the choice of the basis; consider specific basis, such as trigonometric basis, Haar basis, Karhunen-Loeve basis or other ones.
- choose basis which is "good" for the volatility function τ with respect to observed datas.

Still are to be chosen:

- $T > 0$ (even if we could fix $T = 10$, 20 or 30 years)
- the basis $e_1, e_2, \ldots$
- the number N of finite components (a "good" number would be $N = 3$ or 4)

Instead of a separable Hilbert space H, one could consider Schauder basis in the space $C([0, T])$ and do a similar analysis.

Chapter 3

Calibrating the model

In this chapter we try to calibrate the model to the market datas, that is to the assets traded in the market. Tipically market datas are the prices of some coupon bonds, from which one can derive prices of zero-coupon bonds with typical maturities $T = j \cdot 3$ months, $j = 1, \ldots, N$ (for example $N = 40$, that is 10 years), and prices of some interest rate derivatives, such as futures or options on some bonds, caps, floors and swaptions.

3.1 An "implied" approach

Now we try to find the diffusion function τ implicitly from the prices of the caps in the market. We will find some general finite dimensional constraints for τ, which it must satisfy in order to price the caps.

A **cap** is the sum of several caplets, each one having a different maturity $T_i = i \cdot \theta, i = 1, \ldots, N$, where θ usually is equal to 0.25 or 0.5 (three or six months), and the same exercise rate.

A **caplet** is a claim giving the right to exchange the linear rate $R_L(T_{i-1}, \theta)$, defined by

$$1 + \theta R_L(T_{i-1}, \theta) = \frac{1}{B(T_{i-1}, T_i)}$$

with the fixed rate R^f, multiplied by θ, in the maturity T_i; hence the caplet gives the payoff

$$\left(\theta R_L(T_{i-1}, \theta) - \theta R^f\right)^+$$

at time T_i, so it has the price:

$$\text{Caplet}_i(t) = \mathbb{E}_{\mathbb{Q}}^t \left[e^{\int_t^{T_i} r_u \; du} \left(\theta R_L(T_{i-1}, \theta) - \theta R^f\right)^+ \right]$$

where we indicate with $\mathbb{E}_{\mathbb{Q}}^t$ the conditional expectation with respect to the probability $\mathbb{Q}$ and the σ-algebra $\mathcal{F}_t$. We notice that the caplet will be exercised if and only if it gives a positive payoff, that is if and only if we are in the so called **exercise set**

$$\mathcal{E}_i = \left\{ \frac{B(T_{i-1}, T_i)}{B(T_i, T_i)} \leq \frac{1}{1 + \theta R^f} \right\}$$

We also notice that the set $\mathcal{E}_i$ is $\mathcal{F}_{T_{i-1}}$-measurable, so one can know in T_{i-1} whether to exercise the cap in T_i or not. Now, introducing the forward neutral probabilities $\mathbb{Q}_{T_{i-1}}$ and $\mathbb{Q}_{T_i}$ (see [7]), we calculate explicitly the price of a caplet:

$$\begin{aligned}
\mathrm{Caplet}_i(t) &= \mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_i} r_u\ du}\left(\frac{1}{B(T_{i-1},T_i)} - (1+\theta R^f)\right)^+\right] = \\
&= \mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_i} r_u\ du}\frac{1}{B(T_{i-1},T_i)}\mathbf{1}_{\mathcal{E}_i} - (1+\theta R^f)\mathbf{1}_{\mathcal{E}_i}\right] = \\
&= \mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_{i-1}} r_u\ du}\mathbb{E}^{T_{i-1}}_{\mathbb{Q}}\left[e^{\int_{T_{i-1}}^{T_i} r_u\ du}\right]\frac{1}{B(T_{i-1},T_i)}\mathbf{1}_{\mathcal{E}_i}\right] - \\
&\quad -(1+\theta R^f)\mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_i} r_u\ du}\mathbf{1}_{\mathcal{E}_i}\right] = \\
&= \mathbb{E}^t_{\mathbb{Q}}\left[e^{\int_t^{T_{i-1}} r_u\ du}\mathbf{1}_{\mathcal{E}_i}\right] - (1+\theta R^f)B(t,T_i)\mathbb{Q}^t_{T_i}(\mathcal{E}_i) = \\
&= B(t,T_{i-1})\mathbb{Q}^t_{T_{i-1}}(\mathcal{E}_i) - (1+\theta R^f)B(t,T_i)\mathbb{Q}^t_{T_i}(\mathcal{E}_i)
\end{aligned}$$

Now we suppose that

$$\frac{dB(t,T)}{B(t,T)} = r_t\ dt + \Gamma(t,T)\ dW_t$$

for all $T = T_i$, where W is a d-dimensional (infinite dimensional?) Brownian motion under $\mathbb{Q}$, and Γ is deterministic; this is equivalent to suppose that

$$dr_t(x) = (Ar_t(x) + c(t,x))\ dt + \tau(t,x)\ dW_t$$

with τ deterministic function of t and x such that

$$\tau(t,x) = -\frac{\partial}{\partial T}\Gamma(t,t+x), \quad \text{that is} \quad \Gamma(t,T) = -\int_0^{T-t}\tau(t,u)\ du \tag{3.1}$$

Then the price of a caplet is given by:

$$\mathrm{Caplet}_i(t) = B(t,T_{i-1})N(d_{0,i}) - B(t,T_i)(1+\theta R^f)N(d_{1,i})$$

where

$$d_{0,i} = \frac{1}{\sigma_{t,T_i}}\log\left(\frac{B(t,T_{i-1})}{(1+\theta R^f)B(t,T_i)}\right) + \frac{1}{2}\sigma_{t,T_i}, \quad d_{1,i} = d_{0,i} - \sigma_{t,T_i}$$

and

$$\sigma^2_{t,T_i} = \int_t^{T_{i-1}} \|\Gamma(s,T_i) - \Gamma(s,T_{i-1})\|^2\ ds$$

where $\|\cdot\|$ is the norm on $\mathbb{R}^d$.

We notice that this formula is very similar to the Black-Scholes formula; in particular, $\mathrm{Caplet}_i(t)$ it is an increasing function of σ_{t,T_i}. This means that if one knows the price of a caplet and all the variables in the price, except σ_{t,T_i}, one can derive it inverting the function.

The problem is that in the market caplets are not priced, but one can read only the price of caps, which are given by

$$\text{Cap}(t) = \sum_{i=1}^{N} \text{Caplet}_i(t)$$

so it is not well clear how to derive the σ_{t,T_i} for the single caplet. One method could be to ricave a volatility for each cap priced in the market, and then to make a linear interpolation for the volatility of each caplet; this method is presented in [3] as being the market practise.

Once that we have the quantities σ_{t,T_i}, we obtain a constraint for the function τ:

$$\frac{\partial}{\partial t}\sigma^2_{t,T_i} = -\|\Gamma(t,T_i) - \Gamma(t,T_{i-1})\|^2 = -\left\| \int_{T_{i-1}-t}^{T_i-t} \tau(t,u)\ du \right\|^2$$

however, these are only finite dimensional constraints, and it is still to be specified:

- the "dimension" of the model, that is the number d of Brownian motions to use
- the shape of the functions τ: for example, in [3] the authors use cubic splines, but say that the market practise is linear interpolation.

3.2 Two "historical" approaches

Now we present two historical approaches, each one based on the observation of the prices of zero-coupon bonds, which we suppose to have dynamics

$$\frac{dB(t,T)}{B(t,T)} = r_t(0)\ dt + \Gamma(t,T)\ dW_t$$

To this purpose we define $R(t,T)$ and $\tilde{R}(t,T)$ in this way:

$$B(t,T) = \exp\left(-\int_0^{T-t} r_t(u)\ du \right) = \exp\left(-(T-t)R(t,T)\right) = \exp\left(\tilde{R}(t,T)\right)$$

then, since $\tilde{R}(t,T) = \log B(t,T)$, we have:

$$d\tilde{R}(t,T) = (r_t(0) + 1/2\Gamma^2(t,T))\ dt + \Gamma(t,T)\ dW_t$$

and since $R(t,T) = \frac{\tilde{R}(t,T)}{T-t}$, we have:

$$\begin{aligned} dR(t,T) &= -\frac{\tilde{R}(t,T)}{(T-t)^2}\ dt + \frac{d\tilde{R}(t,T)}{T-t}\frac{d\langle \tilde{R}(\cdot,T)\rangle_t}{2(T-t)} = \\ &= -\frac{R(t,T)}{T-t}\ dt + \frac{r_t(0) + 1/2\Gamma^2(t,T)}{T-t}\ dt + \frac{\Gamma(t,T)}{T-t}\ dW_t \end{aligned}$$

3.2.1 With less datas

In this approach we suppose to have datas in the form $B(t,T)$, $T = j\cdot 3$ months, $j = 1,\dots,N$ (for example $N = 40$, that is 10 years), $t = i$ days, $i = 1,\dots,90$. We can get the rates in this way:

$$B(t,T) = \exp\left(-\int_t^T f(t,u)\ du\right) = \exp\left(-\int_0^{T-t} r(t,u)\ du\right)$$

$$f(t,T) = -\frac{\partial}{\partial T}\log B(t,T), \qquad r(t,x) = -\frac{\partial}{\partial T}\log B(t,t+x)$$

(we can get only approximate values)

If we put $x = T - t$, then for each day t the values of $(B(t,t+x))_x$ have a 1 day shift to the left, so if we have 90 samples we have a discrete scheme with $90\cdot N$ values. If we interpolate each curve (which way? linear? spline? exponential?), then we could have 90 samples of forward rates' curves. Then we can use Karhunen-Loewe method to find a "good" basis, and try to remake the interpolation using Karhunen-Loewe basis...

3.2.2 With more datas

In this approach we suppose to have datas in the form $B(t,T_i)$, $T_i = i\cdot 3$ months, $i = 1,\dots,N$ (for example $N = 40$, that is 10 years), and datas in t with a difference in t very near (for example, if we have tick by tick datas: this datas are taken in each transaction made in the stock market, and tipically the difference between two t's near is 5-10 seconds). From these datas, we can get the volatility Γ of the bonds

$$dB(t,T) = B(t,T)(r_t\ dt + \Gamma(t,T)\ dW_t)$$

where Γ is a d-dimensional vector.

A possible way is the following: we fix the j-th day, and let t vary only into the j-th day. We suppose that $\Gamma(t,T_i) \simeq \Gamma(T_i)$ for $t \in [i \text{ day}, i{+}1 \text{ day}]$, and we pose $X_t^i = \log B(t,T_i)$, so we have

$$X_t^i = X_0^i + (r_t - 1/2\Gamma^2(T_i))t + \Gamma(T_i)W_t$$

Now we discretize $[i \text{ day}, i+1 \text{ day}]$ with a step $h = 1 \text{ day}/n$, and we pose:

$$\hat{\sigma}_n^{lm} = \sum_{i=0}^{n-1}(X_{(i+1)h}^l - X_{ih}^l)(X_{(i+1)h}^m - X_{ih}^m)$$

Then we have:

$$\begin{aligned}
\hat{\sigma}_n^{lm} &= \sum_{i=0}^{n-1}\big((r_t - 1/2\Gamma^2(T_l))h + \Gamma(T_l)(W_{(i+1)h} - W_{ih})\big)\cdot \\
&\quad \cdot\big((r_t - 1/2\Gamma^2(T_m))h + \Gamma(T_m)(W_{(i+1)h} - W_{ih})\big) = \\
&= (r_t - 1/2\Gamma^2(T_l))(r_t - 1/2\Gamma^2(T_m))\sum_{i=0}^{n-1}h^2 + \Gamma(T_l)\Gamma(T_m)\sum_{i=0}^{n-1}(W_{(i+1)h} - W_{ih})^2 +
\end{aligned}$$

$$+\left((r_t - 1/2\Gamma^2(T_l))\Gamma(T_m) + (r_t - 1/2\Gamma^2(T_m))\Gamma(T_l)\right) h \sum_{i=0}^{n-1}(W_{(i+1)h} - W_{ih}) =$$

$$= \quad (r_t - 1/2\Gamma^2(T_l))(r_t - 1/2\Gamma^2(T_m))\frac{t^2}{n} + \Gamma(T_l)\Gamma(T_m)\sum_{i=0}^{n-1}(W_{(i+1)h} - W_{ih})^2 +$$

$$+(r_t - 1/2\Gamma(T_l)\Gamma(T_m))(\Gamma(T_l) + \Gamma(T_m))\frac{t}{n}W_t$$

As n tends to infinity, the first and the third addend tend to zero, while the second tends to $\Gamma(T_l)\Gamma(T_m)t$; so the estimator $\hat{\sigma}_n = (\hat{\sigma}_n^{lm})_{lm}$ tends to $\Gamma \cdot \Gamma^t$, where Γ is the matrix having as rows $\Gamma(T_l), l = 1, \dots, d$. We can then get Γ making the square root of the resulting matrix.

Once we have $\Gamma(t, T)$, we get:

$$\tau(t, x) = -\frac{\partial}{\partial T}\Gamma(t, t + x)$$

(as before, we can get only approximate values)

3.3 Karhunen–Loeve expansion

In this part we follow the approach of [12]; the Karhunen–Loeve expansion is a tool that is classically used in fluidodynamics giving nice numerical results, and we believe that it succeeds in treating well also our problem.

We suppose we have a random variable u with values in a function space H, dual of another space K.

Example 3.1 *In our case, we could take* $H = K = L^2(\mathbb{R}^+), H^1(\mathbb{R}^+), L^2([0, T])$ *or* $H^1([0, T])$.

From now on, we suppose that $\mathbb{E}[u] = 0$ (we can always reduce the problem to this case by defining $u' = u - \mathbb{E}[u]$). We suppose then that the bilinear function defined by the variance R of u from H^2 into $\mathbb{R}^+$

$$R(k, k) = \mathbb{E}[\langle u, k\rangle^2] \quad \forall k \in K$$

is continuous; then, for the **kernel representation theorem**, we can find a function $R(t, t')$ such that

$$R(k, k) = \int R(t, t')k(t)k(t')\ dt\ dt'$$

then we can interpret R as an operator going from K to H.

3.3.1 The representation problem

Now we want to maximize the expression

$$\lambda = \frac{\mathbb{E}[\langle u, k\rangle^2]}{\langle k, k\rangle} = \frac{R(k, k)}{\langle k, k\rangle}$$

in $k \in K \cap H$. From a standard procedure by calculus of variations, the extremal points are the k such that

$$R(k, k') = \lambda \langle k, k' \rangle \quad \forall k' \in K \cap H$$

that is, the extremal k are the eigenvectors of R:

$$Rk = \lambda k$$

3.3.2 The "finite energy" case: Karhunen–Loeve expansion

Now we suppose that $(f_n)_n$ is a complete (non necessary orthonormal) system in K (in L^2), decreasing exponentially to infinity, and that

$$\sum_{n=1}^{\infty} \mathbb{E}[\langle u, f_n \rangle^2] < +\infty$$

this is equivalent to $\sum_{n=1}^{\infty} \langle u, f_n \rangle^2 < +\infty$ $\mathbb{P}$-a.s.; then $u \in L^2$ $\mathbb{P}$-a.s.; besides, R is symmetric and nonnegative-definite, so it exists a complete orthonormal set $(k_n)_n$ such that $Rk_n = \lambda_n k_n$, with $\lambda_n \geq 0$.

3.3.3 Calculation of the eigenvalues of R

Now we present an iterative method to find the first eigenvalue λ_1 of R and the corresponding unitary eigenvector k_1.

We fix $k^0 \in K$, and we define:

$$k^{n+1} = Rk^n$$

then we have that:

$$\frac{k^{n+1}}{k^n} \to \lambda_1, \qquad \frac{k^n}{\lambda_1^n} \to \langle k_1, k^0 \rangle k_1$$

In order to find the other eigenvalues and eigenvectors, we can apply the method to the operator $R_1 = R - \lambda k_1 \otimes k_1$.

3.3.4 Rate of convergence of the series

We search for a number N that would be in some way an index of how much of the terms of the series are "significative"; we define

$$N = \frac{\sum_{n=1}^{\infty} \lambda_n}{\lambda_1}$$

this means that *at least* N terms of the series are needed in order that the sum is significative. If we suppose that L is large enough to include in $[0, L]^2$ "most" of the support of the function $R(\cdot, \cdot)$, then we have:

$$\lambda_1 \geq \frac{1}{L} \int_0^L \int_0^L R(t, t') \; dt \; dt'$$

and we have the estimation:

$$N \leq L \frac{\int_0^\infty R(t,t)\ dt}{\int_0^L \int_0^L R(t,t')\ dt\ dt'}$$

3.3.5 Application: the Musiela model

Now we want to apply this expansion to the process $(r_t)_t$ in this way: we suppose that $(r_t)_t$ is a stationary solution of the Musiela equation:

$$\begin{cases} dr_t(x) = (Ar_t(x) + c_t(x))\ dt + \tau_t(x)\ dW_t \\ r_0(\cdot) \in L^2(\Omega, \mathcal{F}, \mathbb{P}; H) \end{cases}$$

where τ is a priori an element of H^∞ such that $\sum_{n=1}^\infty \|\tau_n\|^2 < +\infty$. Then the marginal law of r_t is

$$\mathbb{P}(r_t) = N(b^*(\cdot) + b_0, Q_\infty) \quad \forall t \geq 0$$

with $b_0 \in \mathbb{R}$ and

$$Q_\infty = \int_0^\infty \sum_{n=1}^\infty S_u\tau_n \otimes S_u\tau_n\ du, \qquad b^*(\cdot) = -\int_0^x c(u)\ du$$

Since the process is stationary, in order to know its law it is sufficient to know Q_∞; in fact, once we know Q_∞, we are able to get the τ_n and to get c from them.

Here there is a hidden detail: in fact we always studied the Musiela model under the risk-neutral probability $\mathbb{Q}$, while we can obtain from law of the time series an estimation of the hystorical probability $\mathbb{P}$. This doesn't create problems, because since $\mathbb{Q}$ and $\mathbb{P}$ are equivalent, the Gaussian random variable we are studying has the same covariance both under $\mathbb{P}$ as under $\mathbb{Q}$, so we can proceed without troubles.

Now we suppose to have $\bar{N}$ observations $r_{t_i}(x)$, $i = 1, \ldots, \bar{N}$ of the forward curve for $x \in [0, T]$. Since the process is stationary, we can consider a realization of r_t consisting of $\bar{N}$ observations r_{t_i}, $i = 1, \ldots, \bar{N}$ as $\bar{N}$ realizations of the random variable r_t for a given $t \geq 0$.

From now on, we indicate $r_i = r_{t_i}$ for brevity. Then we can approximate:

$$\mathbb{E}[r_t] \simeq \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} r_i, \quad \text{Cov}\ (r_t, r_t) \simeq \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} r_i \otimes r_i$$

In order to implement the algorithm, we follow these procedure:

1st step we center r, passing to $r'_t = r_t - \mathbb{E}[r_t]$:

$$r'_j = r_j - \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} r_i$$

then $r'_t \circ N(0, Q_\infty)$, that is $R = Q_\infty$.

2nd step we apply the algorithm: we fix an initial k^0 and we define:

$$k^{n+1} = \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} \langle r_i, k^n \rangle r_i$$

and we find λ_1 and k_1.

3rd step we define R_1:

$$R_1 = \frac{1}{\bar{N}} \sum_{i=1}^{\bar{N}} r_i \otimes r_i - \lambda_1 k_1 \otimes k_1$$

and we come back to the 2nd step (we can stop to the suitable N).

Once we obtain the $(k_n)_n$, we can obtain the $(\tau_n)_n$ in this way: we have two expressions of the covariance Q_∞:

$$Q_\infty = \sum_{n=1}^{\infty} \int_0^\infty S_u \tau_n \otimes S_u \tau_n \; du = \sum_{n=1}^{\infty} \lambda_n k_n \otimes k_n$$

each term in the two sums corresponding to a different component of the Brownian motion W. Since the components of W are independent, then we have:

$$\int_0^\infty S_u \tau_n \otimes S_u \tau_n \; du = \lambda_n k_n \otimes k_n \quad \forall n \in \mathbb{N}$$

so we have:

$$\int_0^\infty \tau_n(x+u) \; du = \sqrt{\lambda_n} k_n(x) \quad \forall x \in \mathbb{R}^+, n \in \mathbb{N}$$

Deriving under the integral sign, we obtain

$$\sqrt{\lambda_n} k_n'(x) = \int_0^\infty \tau_n'(x+u) \; du = [\tau_n(x+u)]_0^{+\infty} = -\tau_n(x) \quad \forall x \in \mathbb{R}^+, n \in \mathbb{N}$$

because $\lim_{x \to +\infty} \tau_n(x) = 0$; so we have that

$$\tau_n = -\sqrt{\lambda_n} k_n'$$

Chapter 4

Robustness of the Musiela model

4.1 A functional differential equation for interest rates derivatives

We start from the Musiela equation

$$\begin{cases} dr_t(x) = (Ar_t(x) + c(t, r_t(\cdot))(x)\ dt + \tau(t, r_t(\cdot))(x)\ dW_t, & t \in [0,T] \\ r_0(\cdot) \in H \end{cases} \tag{4.1}$$

where $A = \frac{\partial}{\partial x}$, $(W_t)_t$ is a d-dimensional Brownian motion, H is a separable Hilbert space of functions from $\mathbb{R}^+$ to $\mathbb{R}$ (for example, $L^2(\mathbb{R}^+)$ or $H^1(\mathbb{R}^+)$), $\tau : \mathbb{R}^+ \times H \to H^d$ and $c : \mathbb{R}^+ \times H \to H$, defined by $c(t, r(\cdot))(x) = \sum_{i=1}^d \tau_i(t, r(\cdot))(x) \int_0^x \tau_i(t, r(\cdot))(u)\ du$, satisfy the following conditions:

(A1) the mapping $c : \mathbf{R}_+ \times H \to H$ is measurable and

$$\|c(t,r) - c(t,r')\| \leq K\|r - r'\| \quad \forall r, r' \in H, \forall t \in [0,T]$$

(A2) the mapping $\tau : \mathbf{R}_+ \times H \to H^d$ is measurable and

$$\|\tau(t,r) - \tau(t,r')\| \leq K\|r - r'\| \quad \forall r, r' \in H, \forall t \in [0,T]$$

These conditions comprehend the stationary Gaussian case we analyzed in Chapter 1 as the particular case in which τ is a constant function.

Proposition 4.1 *If (A1) and (A2) hold then for every H-valued random variable r_0 independent of the Wiener process W there exists a unique continuous solution $(r_t)_t$ to (1.1) which is a strong Markov process on H. Moreover, if $\mathbb{E}[\|r_0\|^p] < \infty$ for a certain $p \geq 2$ then for any $T > 0$:*

$$\sup_{t \leq T} \mathbb{E}[\|r_t\|^p] \leq C_{T,p} \left(1 + \mathbb{E}[\|r_0\|^p]\right).$$

Moreover, if τ and c are Lipschitz transformations on $H^1(0,\infty)$ and $r_0 \in H^1(0,\infty)$ then $r_t \in H^1(0,\infty)$ $\mathbb{Q}$-almost surely for all $t \geq 0$.

The proof is an immediate consequence of the general theory developed in [5].

From now on we assume that our system is homogeneous in time, that is c and τ don't depend explicitly on time, but only on the rate $r \in H$. We know that the bonds' prices are deterministic functions of the spot forward rates:

$$B(t,T) = \mathcal{B}(t, r_t) = \exp\left(-\int_t^T r_t(u)\, du\right)$$

More in general, we can consider a zero coupon bond as the simplest case of an European contingent claim. We say that $F : H \to \mathbb{R}$ defines an **European interest rate derivative** or **European interest rate contingent claim** with maturity T if $\mathbb{E}[F(r_T)] < +\infty$; the claim gives to his owner the final payoff $F(r_T)$. In particular the case $F \equiv 1$ corresponds to the zero-coupon bond $B(t,T)$ with maturity T. Due to the Markov property of the process r the price C_T of the claim F at time $t \leq T$ is

$$\begin{aligned} C_T(t) &= \mathbb{E}\left[\exp\left(-\int_t^T r_u(0)\, du\right) F(r(T)) \middle| \mathcal{F}_t\right] \\ &= \mathbb{E}\left[\exp\left(-\int_t^T r_u(0)\, du\right) F(r(T)) \middle| r(t)\right]. \end{aligned} \tag{4.2}$$

As in the finite dimensional case, we have the following result, that links the European contingent claim F with the equation

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,r) + Lu(t,r) - r(0)u(t,r) = 0 \\ \\ u(T,r) = F(r). \end{cases} \tag{4.3}$$

where L is defined on $D(A)$ by

$$LF(r) = \frac{1}{2}\langle D^2F(r)\tau(r), \tau(r)\rangle + \langle Ar + c(r), DF(r)\rangle \tag{4.4}$$

Equation (4.3) is a Feynman-Kac equation in the Hilbert space H. To make our statements more precise, for $m \geq 0$ we define the space

$$C_m(H) = \left\{F : H \to \mathbf{R} \mid \frac{F(\cdot)}{1 + \|\cdot\|^m} \text{ bounded and uniformly continuous on } H\right\}$$

The space $C_m(H)$ is endowed with the norm

$$\|F\|_m = \sup_{r \in H}\left(\frac{|F(r)|}{1 + \|r\|^m}\right).$$

Let $C_m^n(H)$ denote the subspace of $C_m(H)$ containing all functions F which are n times Fréchet differentiable on H and

$$\sup_{r \in H}\left(\frac{\|D^kF(r)\|}{1 + \|r\|^m}\right) < \infty, \quad k = 1, \ldots, n,$$

where $D^k F(\phi)$ denotes the k-th Fréchet derivative of F. If $F \in C^1_m(H)$ then the derivative $DF(\phi)$ of F at $\phi \in H$ can be identified with an element of H and therefore $DF : H \to H$ is a continuous mapping. If $F \in C^2_m(H)$ then the second derivative $D^2F(\phi)$ is a symmetric linear operator in H and we can define a continuous mapping $D^2F : H \to \mathcal{L}(H)$ from H into the algebra of bounded operators on H. Finally by $C^n(H)$ we denote the vector space of all real-valued continuous functions on H with n continuous Fréchet derivatives. Let $\psi_1, \dots, \psi_n \in \mathrm{dom}\,(A^*)$ and let P_n denote the orthogonal projection on the linear span H_n of the vectors $\psi_1, \dots, \psi_n$. Moreover, let

$$\mathcal{D}_0 = \{F \in C_\alpha(H) : F = f \circ P_n,\ f \in C^2_m(R^n), n = 1, \dots\}.$$

Note that every $F \in C_m(H)$ is a pointwise limit of elements of $\mathcal{D}_0$.

Theorem 4.2 *If τ and c are twice differentiable on H and such that*

$$\|\tau(\phi)\| + \|c(\phi)\| \leq C(1 + \|\phi\|).$$

then for every $F \in C^2_m(H)$ the function C_T is a unique solution of the Feynman-Kac equation (4.3) in the following sense.

i) The function $C_T : \mathbf{R}_+ \times H \to \mathbf{R}$ is bounded and continuous with respect to each variable.

ii) For every $t \geq 0$ we have $C_T(t, \cdot) \in C^2(H)$.

iii) $C_T \in C^1([0,T], H^1)$.

iv) Equation (4.3) holds for every $\phi \in$ dom(A) and $t \geq 0$. Moreover, C_T is given by (4.2).

For a proof, see [9].

This theorem is a first result, but it is not enough: in fact the principal examples of European interest rates derivatives (calls, puts, swaptions) are of class C^0 but not of class C^1.

4.2 Robustness of the Musiela model

We suppose that the dynamics of the forward rates is given by Equation

$$\begin{cases} dr_t(x) = (Ar_t(x) + \bar{c}_t(x))\ dt + \bar{\tau}_t(x)\ dW_t \\ r_0(\cdot) \in H \end{cases} \tag{4.5}$$

where $A = \frac{\partial}{\partial x}$, $(W_t)_t$ is a d-dimensional Brownian motion, H is a separable Hilbert space of functions from $\mathbb{R}^+$ to $\mathbb{R}$ (for example, $L^2(\mathbb{R}^+)$ or $H^1(\mathbb{R}^+)$), $\bar{\tau}_t(x)$ is a H^d-valued adapted process, and $\bar{c}(x) = \sum_{i=1}^d \bar{\tau}^i_t(x) \int_0^x \bar{\tau}^i_t(u)\ du$ is a H-valued process; moreover, the process τ is supposed to take values in some compact set $\Sigma \in H^d$. This model is a quite general one, and it is an analogous of stochastic volatility models for stock prices analyzed in literature in the last years (see for example [8] and [1]).

Now we consider an operator in the market that wants to create and sell an European interest rate derivative asset with payoff $F(r_{\bar{T}})$ depending on the value in the maturity $\bar{T}$ of all the forward rates curve $r_{\bar{T}}$. In many relevant cases, the payoff F will depend not on all the curve $r_{\bar{T}}$, but only on the values of N particular bonds with maturity $T_i > \bar{T}$, so it will be of the form

$$F(B(\bar{T},T_1),\ldots,B(\bar{T},T_N)), \quad \bar{T} < T_1 < \ldots < T_N \tag{4.6}$$

From now on, we suppose that F is of the form (4.6). In order to hedge this asset, the operator chooses to use the homogeneous Gaussian Musiela model described in Chapter 1. Then he (she) builds a self-financing strategy in this way: he (she) fixes the price C_t of the option at time t as

$$C_t = u(t, r_t)$$

where u is the solution of the following nonlinear functional differential equation:

$$\begin{cases} \dfrac{\partial u}{\partial t}(t,r) + \sup\limits_{\tau\in\Sigma} Lu(t,r) - r(0)u(t,r) = 0, \quad (t,r) \in [0,\bar{T}] \times H \\ \\ u(\bar{T},r) = F(r), \quad r \in H \end{cases} \tag{4.7}$$

where L is the operator (4.4). Equation (4.7) is a Hamilton-Jacobi-Bellmann type equation; we notice that its finite-dimensional versions appear for the first time in one dimension in [1], and it is generalized to the finite-dimensional case in [14]; as in the previous works Equation (4.7) will be called from now on the **Black-Scholes-Barenblatt** equation associated to the contingent claim F.

Lemma 4.3 *If F is of the form (4.6), and the function $u(t,r)$ is a solution of Equation (4.7) and is of the form*

$$u(t,r) = C(t, r_t(0), B(t,T_1), \ldots, B(t,T_N))$$

where $C : [0,\bar{T}] \times \mathbb{R} \times \mathbb{R}^N$, $r_t(0)$ is the spot rate at time t and $B(t,T_i)$ are the values in t of the zero-coupon bonds maturing in T_i, $i = 1,\ldots,N$, then C is solution of the following equation

$$\begin{cases} \dfrac{\partial C}{\partial t}(t,r_0,B) + r_0\langle DC(t,r_0,B),B\rangle + \dfrac{1}{2}\sup\limits_{\tau\in\Sigma} \mathrm{tr}\ \left(D^2C(t,r_0,B)(\bar{B}\Gamma)(\bar{B}\Gamma)^*\right) = r_0C(t,B), \\ \\ C(\bar{T},r_0,B) = F(B), \quad r_0 \in \mathbb{R}, B \in \mathbb{R}^N \end{cases} \tag{4.8}$$

where $DC(t,r_0,B) = (\frac{\partial C}{\partial B_1}(t,r_0,B),\ldots,\frac{\partial C}{\partial B_n}(t,r_0,B))$, and $D^2C(t,r_0,B) = (\frac{\partial^2 C}{\partial B_i B_j}(t,r_0,B))_{ij}$, $\Gamma(t,T)$ is a $d \times N$ matrix which the i-th row is the volatility $\Gamma(t,T_i)$ of the bond $B(t,T_i)$, given by Equation (3.1), and

$$\bar{B} = \mathrm{diag}\ (B) = \begin{pmatrix} B_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & B_n \end{pmatrix}$$

Proof. We notice that

$$\frac{\partial u}{\partial r} = \frac{\partial C}{\partial B}\frac{\partial B}{\partial r}, \qquad \frac{\partial^2 u}{\partial r^2} = \left(\frac{\partial B}{\partial r}\right)^* \frac{\partial^2 C}{\partial B^2}\left(\frac{\partial B}{\partial r}\right) + \frac{\partial C}{\partial B}\frac{\partial^2 B}{\partial r^2}$$

and that for all maturities T_i we have:

$$\frac{\partial B(t,T_i)}{\partial r} = -\mathbf{1}_{[0,T_i-t]}B, \qquad \frac{\partial^2 B(t,T_i)}{\partial r^2} = \mathbf{1}^*_{[0,T_i-t]}B\mathbf{1}_{[0,T_i-t]}$$

then from Equation (4.7) we get:

$$\begin{aligned}
r(0)u(t,r) &= \frac{\partial u}{\partial t}(t,r) + \max_{\tau\in\Sigma} Lu(t,r) = \\
&= \frac{\partial u}{\partial t}(t,r) + \max_{\tau\in\Sigma}\left(\left\langle \frac{\partial u}{\partial r}(t,r), Ar + c(r)\right\rangle + \frac{1}{2}\left\langle \frac{\partial^2 u}{\partial r^2}(t,r)\tau(r),\tau(r)\right\rangle\right) = \\
&= \frac{\partial C}{\partial t}(t,B) + \max_{\tau\in\Sigma}\left(-\sum_{i=1}^N \frac{\partial C}{\partial B_i}(t,B)B(t,T_i)\langle \mathbf{1}_{[0,T_i-t]}, Ar + c(r)\rangle + \right. \\
&\quad + \frac{1}{2}\sum_{i=1}^N \tau(r)\left(\mathbf{1}^*_{[0,T_i-t]}B(t,T_i)\frac{\partial^2 C}{\partial B^2}(t,B)B(t,T_i)\mathbf{1}_{[0,T_i-t]} + \right. \\
&\quad \left.\left. + \frac{\partial C}{\partial B}\mathbf{1}^*_{[0,T_i-t]}B_i\mathbf{1}_{[0,T_i-t]}\right)\tau(r)\right) = \\
&= \frac{\partial C}{\partial t}(t,B) + \max_{\tau\in\Sigma}\left(-\sum_{i=1}^N \frac{\partial C}{\partial B_i}(t,B)B(t,T_i)(r(T_i-t) - r(0) + \langle \mathbf{1}_{[0,T_i-t]}, c(r)\rangle) + \right. \\
&\quad + \frac{1}{2}\mathrm{Tr}\left(\sum_{i=1}^N \Gamma(t,T_i)^* B(t,T_i)\frac{\partial^2 C}{\partial B^2}(t,B)B(t,T_i)\Gamma(t,T_i) + \right. \\
&\quad + \left.\left. \Gamma(t,T_i)^*\frac{\partial C}{\partial B_i}B(t,T_i)\Gamma(t,T_i)\right)\right) = \\
&= \frac{\partial C}{\partial t}(t,B) + \max_{\tau\in\Sigma}\left(\sum_{i=1}^N \frac{\partial C}{\partial B_i}(t,B)B(t,T_i)r(0) + \right. \\
&\quad \left. + \frac{1}{2}\mathrm{Tr}\left(\sum i = 1^N \Gamma(t,T_i)^* B(t,T_i)\frac{\partial^2 C}{\partial B^2}(t,B)B(t,T_i)\Gamma(t,T_i)\right)\right) \\
&\quad \left(-\sum_{i=1}^N \frac{\partial C}{\partial B_i}(t,B)B(t,T_i)(r(T_i-t))\right)
\end{aligned}$$

and Equation (4.8) follows. ◇

Equation (4.8) will be called from now on the **reduced Black-Scholes-Barenblatt equation** for F.

In order to hedge the claim F, the agent builds the hedging portfolio holding the quantity Δ_t^i of the zero-coupon bond $B(t,T_i)$ in the hedging portfolio Π as

$$\Delta_t^i = \frac{\partial C}{\partial B_i}(t, B(t,T))$$

where we indicated as $B(t,T)$ the vector $(B(t,T_1),\dots,B(t,T_N))$; Π follows the dynamics:

$$\begin{cases} d\Pi_t = r_t(0)(\Pi_t - \langle \Delta_t, B(t,T)\rangle)\ dt + \langle \Delta_t, dB(t,T)\rangle \\ \Pi_0 = P_0 \end{cases}$$

(for a proof, see [11]).

We define the **tracking error**:

$$e_t = \Pi_t - C_t$$

with $e_0 = 0$. Intuitively, the tracking error gives the error made by the operator in his (her) estimation, or better the difference between the hedging portfolio held by the agent and the option he (she) sold.

We want to find a superstrategy where e_t is increasing $\forall t$ $\mathbf{Q}$-a.s.

Theorem 4.4 *If C is a $C^{1,2}([0,T)\times\mathbf{R}^n)$ solution of Equation (4.8), then $(e_t)_t$ is an increasing process.*

Proof. The evolution of $(e_t)_t$ is the following:

$$\begin{aligned} de_t &= r_t(0)(\Pi_t - \langle \Delta_t, B(t,T)\rangle)\ dt + \langle \Delta_t, dB(t,T)\rangle - \frac{\partial C}{\partial t}(t,B(t,T))\ dt - \\ &\quad -\langle DC(t,B(t,T)), dB(t,T)\rangle - \\ &\quad -\frac{1}{2}\mathrm{tr}\ \left(D^2C(t,B(t,T))(\bar{B}(t,T)\Gamma(t,T))(\bar{B}(t,T)\Gamma(t,T))^*\right)\ dt = \\ &= r_t(0)(\Pi_t - \langle \Delta_t, B(t,T)\rangle)\ dt + r_t(0)\langle B(t,T), DC(t,B(t,T))\rangle\ dt - r_t(0)C_t\ dt + \\ &\quad +\frac{1}{2}\max_{\tau\in\Sigma}\mathrm{tr}\ \left(D^2C(t,B(t,T))(\bar{B}(t,T)\Gamma(t,T))(\bar{B}(t,T)\Gamma(t,T))^*\right)\ dt - \\ &\quad -\frac{1}{2}\mathrm{tr}\ \left(D^2C(t,B(t,T))(\bar{B}(t,T)\Gamma(t,T))(\bar{B}(t,T)\Gamma(t,T))^*\right)\ dt = \\ &= r_t(0)e_t\ dt + \frac{1}{2}\max_{\tau\in\Sigma}\mathrm{tr}\ \left(D^2C(t,B(t,T))(\bar{B}(t,T)\Gamma(t,T))(\bar{B}(t,T)\Gamma(t,T))^*\right) - \\ &\quad -\frac{1}{2}\mathrm{tr}\ \left(D^2C(t,B(t,T))(\bar{B}(t,T)\sigma(t,T))(\bar{B}(t,T)\sigma(t,T))^*\right) = \\ &= r_t(0)e_t\ dt + R_t\ dt \end{aligned}$$

where

$$\begin{aligned} R_t &= \frac{1}{2}\max_{\tau\in\Sigma}\mathrm{tr}\ \left(D^2C(t,B(t,T))(\bar{B}(t,T)\Gamma(t,T))(\bar{B}(t,T)\Gamma(t,T))^*\right) - \\ &\quad -\frac{1}{2}\mathrm{tr}\ \left(D^2C(t,B(t,T))(\bar{B}(t,T)\sigma(t,T))(\bar{B}(t,T)\sigma(t,T))^*\right) \geq 0 \end{aligned}$$

We call the quantity R_t **consumption rate**. We have:

$$e_t = M_t\int_0^t \frac{1}{M_u}R_u\ du$$

where $M_t = \exp(\int_0^t r_u(0)\ du)$. Then, to have an increasing e_t is equivalent to have $R_t \geq 0$. Since this is verified, we have the thesis. ◇

4.3 Sensibility of the model

First of all we suppose that the rates evolve following the Vasicek model:

$$dr_t = a(b - r_t)\ dt + \sigma\ dW_t$$

The zero coupon bond is

$$B(t,T) = e^{-(T-t)Y(t,T)} = \mathbb{E}\left[e^{-\int_t^T r_u\ du}|\mathcal{F}_t\right]$$

The explicit solution for the spot rate is:

$$r_u = e^{a(u-t)} + \int_0^{u-t} e^{a(u-t-v)} ab\ dv + \int_0^{u-t} e^{a(u-t-v)} \sigma\ dW_v$$

that is, r_u depends linearly on r_t, on a deterministic term and on a (Gaussian) random variable independent on $\mathcal{F}_t$, so we can represent the zero coupon bond in this way:

$$B(t,T) = \exp\left(-r_t(T-t)\beta(T-t) - (T-t)d(T-t)\right)$$

where

$$\beta(T-t) = \frac{1}{T-t}\int_t^T e^{a(u-t)}\ du = \frac{e^{a(T-t)} - 1}{a(T-t)}$$

Then we have:

$$\begin{aligned} Y(t,T) &= r_t\beta(T-t) + d(T-t) \\ dY(t,T) &= \beta(T-t)\ dr_t - \beta'(T-t)r_t\ dt - d'(T-t)\ dt \end{aligned}$$

By applying Itô formula, we obtain that the volatility of $B(t,T)$ is

$$\Gamma(t,T) = (T-t)\beta(T-t)\sigma$$

Now we take a portfolio having the deterministic payoffs F_i in the dates T_i, $i = 1, \ldots, n$; then its evolution is given by:

$$\begin{aligned} P_t &= \sum_{i=1}^n F_i B(t,T_i) \\ dP_t &= \sum_{i=1}^n F_i dB(t,T_i) = r_t P_t\ dt + \sum_{i=1}^n F_i B(t,T_i)\Gamma(t,T_i)\ dW_t \end{aligned}$$

Then the volatility of P_t is

$$\frac{\sum_{i=1}^n F_i B(t,T_i)\Gamma(t,T_i)}{\sum_{i=1}^n F_i B(t,T_i)} = \frac{\sum_{i=1}^n F_i B(t,T_i)(T_i - t)\beta(T_i - t)\sigma}{P_t}$$

Now we define the **duration** of the portfolio P with respect to the direction f in this way:

$$D_f P = \frac{\sum_{i=1}^n F_i B(t,T_i)(T_i - t) f(T_i - t)}{P_t}$$

then the volatility of P_t can be interpreted as the duration of the portfolio with respect to the direction $f(x) = \sigma\beta(x)$.

There is a link with the classical definition of duration: in fact we have that

$$\frac{1}{P_t}\frac{\partial P_t}{\partial r_t} = \frac{1}{P_t}\sum_{i=1}^{n} F_i \frac{\partial B(t,T_i)}{\partial r_t} = \frac{\sum_{i=1}^{n} F_i B(t,T_i)(T_i - t)\beta(T_i - t)\sigma}{P_t}$$

If more in general we have that $Y(t,T)$ is a deterministic function of k reference quantities (which can be rates or other variables)

$$Y(t,T) = \beta_1(t,T)R_t^1 + \ldots + \beta_k(t,T)R_t^k + \alpha(t,T)$$

then:

$$\frac{1}{P_t}\frac{\partial P_t}{\partial R_t^j} = \frac{1}{P_t}\sum_{i=1}^{n} F_i \frac{\partial B(t,T_i)}{\partial R_t^j} = \frac{\sum_{i=1}^{n} F_i B(t,T_i)(T_i - t)\beta_j(t,T_i)}{P_t}$$

that can be interpreted as the **sensibility** of P_t in the j-th direction.

4.4 Open problems

- existence and uniqueness of the solution of the BSB equation (4.7); while in the case in which the final condition $F \in C^2(H)$ the proof would probably be quite standard, in the case $F \in C^0$ is not clear; probably one has to introduce a notion like viscosity solutions in infinite dimensions.

- is it possible to reduce the sup operator in both the equations (4.7) and (4.8) to a max? Moreover, when is it possible to project the $\Sigma \in H^\infty$ in a compact set in $\mathbb{R}^N$? Finally, is it possible to reduce the reducted BSB equation (4.8) to a partial differential equation in $\mathbb{R}^{N+2}$ (the $N+2$ variables are N bonds, the short rate and the time)?

- if $Y(t,T) = G(r_t(\cdot))$ is a linear functional of all the forward rate curve $r_t(\cdot)$, is it possible to write something like

$$\frac{1}{P_t}\frac{\partial P_t}{\partial f} = \frac{\sum_{i=1}^{n} F_i B(t,T_i)(T_i - t)\frac{\partial G}{\partial f}}{P_t}$$

  where $\frac{\partial G}{\partial f}$ is the Frechet derivative of G with respect to F?

Appendix A

Positivity of rates

One of the main weaknesses of Gaussian interest rates models is the possibility of negative rates with positive probability. This appendix partially investigates this point. First of all we concentrate on the one-dimensional Vasicek model, and calculate the probability that there will be a negative forward rate at a certain time; this probability being 1, we calculate the probability that there will be a negative forward rate before a fixed instant T. Later, we do all this in the more general finite dimensional Gaussian model.

A.1 The Vasicek model

We take $R \geq 0$, and we define

$$\tau_R = \inf\{t \mid r_t < 0 \text{ or } r_t > R\}, \quad \tau = \inf\{t \mid r_t < 0\}$$

We notice that $\tau_R \to \tau$ a.s. as R tends to $+\infty$.

Lemma A.1 *$\tau < +\infty$ a.s.*

Proof. First of all we search for a function h solution of $Ah = 0$, where A is the second order operator associated to $(r_t)_t$, that is:

$$Ah(x) = a(b - x)h'(x) + \frac{1}{2}\sigma^2 h''(x) \tag{A.1}$$

Lemma A.2 *A solution of $Ah = 0$ is of the form*

$$h(r) = D + C \int_0^r e^{\frac{a}{\sigma^2}(y-b)^2}\, dy$$

where C, D are arbitrary real numbers.

Proof. We define $g = h'$; then the equation $Ah = 0$ becomes:

$$g'(x) = \frac{2a(x-b)}{\sigma^2} g(x)$$

then

$$g(x) = Ce^{\int \frac{2a(y-b)}{\sigma^2}\, dy} = Ce^{\frac{a(x-b)^2}{\sigma^2}}$$

since $h = \int v(x)\, dx$, we have the thesis. $\diamond$

If h is of the above form, then $M_t = h(r_{t\wedge\tau_R})$ is a bounded martingale (because $0 \leq r_{t\wedge\tau_R} \leq R$ and h is continuous), so it converges a.s.; but, if $\tau_R = +\infty$, then $h(r_{t\wedge\tau_R}) = h(r_t)$ cannot converge, so $\tau_r < +\infty$ a.s.; besides, since $(M_t)_t$ is a bounded martingale, we have $\mathbb{E}[M_t] = \mathbb{E}[M_0]\ \forall t \geq 0$, so:

$$h(x) = \mathbb{E}[h(r_0)] = \lim_{t\to+\infty} \mathbb{E}[h(r_{t\wedge\tau_R})] = \mathbb{E}\left[\lim_{t\to+\infty} h(r_{t\wedge\tau_R})\right] = \mathbb{E}[h(r_{\tau_R})]$$

where the third equality follows from Lebesgue theorem, so:

$$h(x) = h(0)\mathbb{P}\{r_{\tau_R} = 0\} + h(R)\mathbb{P}\{r_{\tau_R} = R\} = h(0)\mathbb{P}\{r_{\tau_R} = 0\} + h(R)(1 - \mathbb{P}\{r_{\tau_R} = 0\})$$

and we obtain:

$$\mathbb{P}\{r_{\tau_R} = 0\} = \frac{h(x) - h(R)}{h(0) - h(R)} = \frac{\displaystyle\int_x^R e^{\frac{a}{\sigma^2}(y-b)^2}\, dy}{\displaystyle\int_0^R e^{\frac{a}{\sigma^2}(y-b)^2}\, dy}$$

We call $v(x, R) = \mathbb{P}\{r_{\tau_R} = 0\}$ under the condition $r_0 = x$. Then we have the intuitive conditions $v(0, R) = 1$ and $v(R, R) = 0$. Moreover we have that

$$\{\tau < +\infty\} = \{\exists R \text{ such that } r_{\tau_R} = 0\} = \bigcup_{R>0} \{r_{\tau_R} = 0\}$$

so

$$\mathbb{P}\{\tau < +\infty\} = \lim_{R\to+\infty} \{r_{\tau_R} = 0\} = \lim_{R\to+\infty} v(x, R) = 1$$

$\diamond$

We have seen that $\mathbb{P}\{\tau < +\infty\} = 1$, that is there exists almost surely a T such that $r_T < 0$. Now, let's fix $T > 0$, and calculate $\mathbb{P}\{\tau \leq T\}$, that is the probability that the spot rate is positive up to the time T. First of all we give the result in terms of the solution of a parabolic PDE with boundary condition.

Lemma A.3 *If there exist a solution $v(t, x) \in C_b^1$ of the parabolic partial differential equation*

$$\begin{cases} \dfrac{\partial v}{\partial t} + Av = 0 & (t, x) \in [0, T) \times (0, +\infty) \\ \\ v(T, x) = 0 & x \in (0, +\infty) \\ v(t, 0) = 1 & t \in [0, T) \end{cases}$$

where A is given by Eq. (A.1), then $\mathbb{P}\{\tau \leq T\} = v(0, r_0)$.

Proof. We apply Itô formula to the function v:

$$dv(t,r_t) = \left(\frac{\partial v}{\partial t} + Av\right)(t,r_t)\ dt + \frac{\partial v}{\partial x}(t,r_t)\sigma\ dW_t$$

then, because of the PDE, we have:

$$v(T\wedge\tau, r_{T\wedge\tau}) = v(0,r_0) + \int_0^{T\wedge\tau} \frac{\partial v}{\partial x}(t,r_t)\sigma\ dW_t$$

making the expectation, since $v \in C_b^1$, then the expectation of the stochastic integral is 0, and we get:

$$\begin{aligned} v(0,r_0) &= \mathbb{E}[v(T\wedge\tau, r_{T\wedge\tau})] = \mathbb{E}[v(T,r_T)\mathbf{1}_{\tau>T} + v(\tau,r_\tau)\mathbf{1}_{\tau\leq T}] = \\ &= \mathbb{E}[0\cdot\mathbf{1}_{\tau>T} + 1\cdot\mathbf{1}_{\tau\leq T}] = \mathbb{P}\{\tau\leq T\} \end{aligned}$$

◇

The problem with this approach is that there is not an explicit solution to this PDE, so we have no idea of the behaviour of the solution. The following proposition partially answers this question, giving an estimation from above to the searched probability.

Proposition A.4

$$\mathbb{P}\{\tau\leq T\} \leq \frac{\sigma\sqrt{e^{2aT}-1}}{2x\sqrt{\pi a}}$$

Proof. We notice that

$$\begin{aligned} \mathbb{P}\{\tau\leq T\} &= \mathbb{P}\left\{\min_{t\in[0,T]} r_t \leq 0\right\} = \mathbb{P}\left\{\max_{t\in[0,T]}(-r_t)\geq 0\right\} = \mathbb{P}\left\{\max_{t\in[0,T]}(-e^{at}r_t)\geq 0\right\} = \\ &= \mathbb{P}\left\{\max_{t\in[0,T]}\left(-b(e^{at}-1) - \int_0^t e^{as}\sigma\ dW_s\right)\geq r_0\right\} \leq \\ &\leq \mathbb{P}\left\{\max_{t\in[0,T]}\left(-\int_0^t e^{as}\sigma\ dW_s\right)\geq r_0\right\} \end{aligned}$$

the process in the last expression is a martingale, so we can apply Kolmogorov's maximal inequality, and we get:

$$\mathbb{P}\{\tau\leq T\} \leq \frac{1}{r_0}\mathbb{E}\left[\left(\int_0^T e^{as}\ dW_s\right)^+\right]$$

Now we need the

Lemma A.5 *If X is a random variable with law $N(0,1)$, then* $\mathbb{E}[X^+] = (2\pi)^{-1/2}$.

Proof.

$$E[X^+] = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} x e^{-\frac{1}{2}x^2}\, dx = \frac{1}{\sqrt{2\pi}} [-e^{-\frac{1}{2}x^2}]_0^{+\infty} = \frac{1}{\sqrt{2\pi}}$$

$\diamond$

Since $\int_0^T e^{as}\, dW_s$ has law $N(0, \int_0^T e^{2as}\, ds)$, that is $N(0, \frac{\sigma^2}{2a}(e^{2aT}-1))$, then we can calculate the last expectation, and finally we get:

$$\mathbb{P}\{\tau \le T\} \le \frac{1}{r_0}\sqrt{\frac{\sigma^2}{2a}(e^{2aT}-1)} = \frac{\sigma\sqrt{e^{2aT}-1}}{2r_0\sqrt{\pi a}}$$

$\diamond$

Now we try a different approach. We see the stochastic integral in the solution as a Brownian motion with respect to another time scale, that is the one of the local times.

We define:

$$c(t) = \int_0^t (e^{as}\sigma)^2\, ds = \frac{\sigma^2}{2a}(e^{2at}-1)$$

and the **local time**:

$$\tau(u) = c^{-1}(u) = \frac{1}{2a}\log\left(1 + \frac{2a}{\sigma^2}u\right)$$

then $\tau(u)$ are (deterministic) stopping times such that

$$B(u) = \int_0^{\tau(u)} e^{as}\sigma\, dW_s$$

is a Brownian motion with respect to the filtration $(\mathcal{F}_{\tau(u)})_t$; then we can write:

$$b(e^{a\tau(u)}-1) = b\left(\exp\left(a\frac{1}{2a}\log\left(1+\frac{2a}{\sigma^2}u\right)\right)-1\right) = b\left(\sqrt{1+\frac{2a}{\sigma^2}u}-1\right)$$

and we finally obtain:

$$e^{a\tau(u)}r_{\tau(u)} = r_0 + b\left(\sqrt{1+\frac{2a}{\sigma^2}u}-1\right) + B(u), \quad 0 \le u \le c(T) = \frac{\sigma^2}{2a}(e^{2at}-1)$$

Now we define:

$$\begin{aligned} X_u &= r_0 + b\left(\sqrt{1+\frac{2a}{\sigma^2}u}-1\right) + B(u) = r_0 + \phi(u) + B(u), \\ Y_u &= r_0 + B(u) \end{aligned}$$

then the processes X and Y induce the laws $\mathbb{P}_X$ and $\mathbb{P}_Y$ on $C_0([0, c(T)])$, which are equivalent and which mutual density is

$$\begin{aligned}\frac{d\mathbb{P}_X}{d\mathbb{P}_Y} &= \exp\left(\int_0^{c(T)} \phi'(t)\ dY_t - \frac{1}{2}\int_0^{c(T)} \phi'(t)^2\ dt\right) = \\ &= \exp\left(\int_0^{c(T)} \phi'(t)\ dB(t) - \frac{1}{2}\int_0^{c(T)} \phi'(t)^2\ dt\right)\end{aligned}$$

Applying Hölder's inequality, we have the following extimation:

$$\begin{aligned}\mathbb{P}\{\tau \leq T\} &= \mathbb{P}\left\{\min_{u\in[0,c(T)]} X_u \leq 0\right\} = \int_\Omega \mathbf{1}_{\min_{u\in[0,c(T)]} X_u \leq 0}\ d\mathbb{P}_X = \\ &= \int_\Omega \mathbf{1}_{\min_{u\in[0,c(T)]} Y_u \leq 0} \frac{d\mathbb{P}_X}{d\mathbb{P}_Y}\ d\mathbb{P}_Y \leq \mathbb{E}[\mathbf{1}^p_{\min_{u\in[0,c(T)]} Y_u \leq 0}]^{1/p} \mathbb{E}\left[\left(\frac{d\mathbb{P}_X}{d\mathbb{P}_Y}\right)^q\right]^{1/q} = \\ &= \left(\mathbb{P}\left\{\min_{u\in[0,c(T)]} Y_u \leq 0\right\}\right)^{1/p} \mathbb{E}\left[\left(\frac{d\mathbb{P}_X}{d\mathbb{P}_Y}\right)^q\right]^{1/q}\end{aligned}$$

for all $p, q > 1$ such that $1/p + 1/q = 1$. We have that:

$$\begin{aligned}\mathbb{E}\left[\left(\frac{d\mathbb{P}_X}{d\mathbb{P}_Y}\right)^q\right] &= \mathbb{E}\left[\exp\left(q\int_0^{c(T)} \phi'(t)\ dB(t) - \frac{q}{2}\int_0^{c(T)} \phi'(t)^2\ dt\right)\right] = \\ &= \exp\left(-\frac{q}{2}\int_0^{c(T)} \phi'(t)^2\ dt\right) \mathbb{E}\left[\exp\left(q\int_0^{c(T)} \phi'(t)\ dB(t)\right)\right] = \\ &= \exp\left(\left(\frac{q^2}{2} - \frac{q}{2}\right)\int_0^{c(T)} \phi'(t)^2\ dt\right)\end{aligned}$$

and:

$$\begin{aligned}\int_0^{c(T)} \phi'(t)^2\ dt &= \int_0^{c(T)} \frac{b^2a^2\ du}{\sigma^4(1+\frac{2a}{\sigma^2}u)} = \int_0^{c(T)} \frac{b^2\ du}{\frac{2\sigma^2}{a}(\frac{\sigma^2}{2a}+u)} = \\ &= \left[\frac{b^2a}{2\sigma^2}\log\left(u+\frac{\sigma^2}{2a}\right)\right]_0^{c(T)} = \frac{b^2a}{2\sigma^2}\log\left(\frac{\frac{\sigma^2}{2a}e^{2aT}}{\frac{\sigma^2}{2a}}\right) = \frac{b^2a^2T}{\sigma^2}\end{aligned}$$

Moreover, if we call $\tau_a = \inf\{t \mid B(t) = a\}$ the stopping time of the first passage of the Brownian motion from a, then

$$\begin{aligned}\mathbb{P}\left\{\min_{u\in[0,c(T)]} Y_u \leq 0\right\} &= \mathbb{P}\left\{\min_{u\in[0,c(T)]} B(u) \leq -r_0\right\} = \mathbb{P}\{\tau_{-r_0} \leq c(T)\} = \\ &= \frac{2}{\sqrt{2\pi c(T)}}\int_{-r_0}^{+\infty} e^{-\frac{x^2}{2c(T)}}\ dx = \frac{2}{\sqrt{2\pi}}\int_{-\infty}^{\frac{r_0}{c(T)}} e^{-\frac{y^2}{2}}\ dy\end{aligned}$$

then we arrive to:

$$\begin{aligned}\mathbb{P}\{\tau \le T\} &= \left(2N\left(\frac{r_0}{c(T)}\right)\right)^{1/p} \exp\left((q-1)\frac{b^2a^2T}{2\sigma^2}\right) = \\ &= \left(2N\left(\frac{r_0}{c(T)}\right)\right)^{1/p} \exp\left(\frac{1}{p-1}\frac{b^2a^2T}{2\sigma^2}\right)\end{aligned}$$

for all $p > 1$, where N is the distribution function of the normal law $N(0,1)$:

$$N(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-\frac{x^2}{2}}\, dx$$

We can try to find the optimal exponent p_1, that is the exponent $p_1 > 1$ such that the quantity on the right hand side is minimal.

Lemma A.6 *The optimal exponent is*

$$p_1 = \frac{\log A - \sqrt{-B\log A}}{B + \log A}$$

where

$$A = 2N\left(\frac{r_0}{c(T)}\right), \quad B = \frac{b^2a^2T}{2\sigma^2}$$

Proof. We have to minimize the function

$$f(p) = A^{1/p}\exp\left(\frac{B}{p-1}\right) = \exp\left(\frac{\log A}{p} + \frac{B}{p-1}\right)$$

We notice that minimizing f is equivalent to minimize its logarithm, then we calculate $(\log f)'$:

$$\frac{d}{dp}\left(\frac{\log A}{p} + \frac{B}{p-1}\right) = -\frac{\log A}{p^2} - \frac{B}{(p-1)^2}$$

Imposing $(\log f)' \ge 0$, we find:

$$p^2(\log A + B) - 2p\log A + \log A \le 0$$

$(\log f)'$ is decreasing from 1 to p_1, then it is increasing until a point p_2, then it is decreasing, so p_1 is the unique local minimum. ◇

A.2 The d-dimensional Gaussian model

As before, we take $L \ge 0$, and we define

$$\tau_L = \inf\{t \mid R_t \notin [0,L]^d\}, \quad \tau = \inf\{t \mid R_t \notin (\mathbb{R}^+)^d\}$$

We notice that $\tau_L \to \tau$ a.s. as L tends to $+\infty$. First of all we search for a function h solution of $\mathcal{A}h = 0$, where $\mathcal{A}$ is the second order operator associated to $(R_t)_t$, that is:

$$\mathcal{A}h(x) = \langle A(c-x), Dh(x)\rangle + \frac{1}{2}\mathrm{Tr}\,(\sigma\sigma^* D^2 h(x)) \tag{A.2}$$

where Dh is the gradient of h and D^2h indicates the Hessian matrix of h. If h is such a function, then $M_t = h(R_{t\wedge\tau_L})$ is a bounded martingale (because $R_{t\wedge\tau_L} \in [0,L]^d$ and h is continuous), so it converges a.s.; but, if $\tau_L = +\infty$, then $h(R_{t\wedge\tau_L}) = h(R_t)$ cannot converge, so $\tau_L < +\infty$ a.s.; besides, since $(M_t)_t$ is a bounded martingale, we have $\mathbb{E}[M_t] = \mathbb{E}[M_0]$ $\forall t \geq 0$, so:

$$h(x) = \mathbb{E}[h(R_0)] = \lim_{t\to+\infty}\mathbb{E}[h(R_{t\wedge\tau_L})] = \mathbb{E}\left[\lim_{t\to+\infty} h(R_{t\wedge\tau_L})\right] = \mathbb{E}[h(R_{\tau_L})]$$

where the third equality follows from Lebesgue theorem. We call $v(x,L) = \mathbb{P}\{R_{\tau_L}|_i = 0 \text{ for some } i\}$ under the condition $R_0 = x$. Then we have the intuitive conditions $v(0,L) = 1$ and $v(L,L) = 0$. Moreover we have that

$$\{\tau < +\infty\} = \{\exists L \text{ such that } R_{\tau_L}|_i = 0 \text{ for some } i\} = \bigcup_{L>0}\{R_{\tau_L}|_i = 0 \text{ for some } i\}$$

so

$$\mathbb{P}\{\tau < +\infty\} = \lim_{L\to+\infty}\{R_{\tau_L}|_i = 0 \text{ for some } i\} = \lim_{L\to+\infty} v(x,L) = 1$$

We have seen that $\mathbb{P}\{\tau < +\infty\} = 1$, that is there exists almost surely a T such that $r_T < 0$. Now, let's fix $T > 0$, and calculate $\mathbb{P}\{\tau \leq T\}$, that is the probability that the rates are positive up to the time T. First of all we give the result in terms of the solution of a parabolic PDE with boundary condition.

Lemma A.7 *If there exist a solution $v(t,x) \in C_b^1$ of the parabolic partial differential equation*

$$\begin{cases} \dfrac{\partial v}{\partial t} + \mathcal{A}v = 0 & (t,x) \in [0,T) \times (0,+\infty)^d \\ \\ v(T,x) = 0 & x \in (0,+\infty)^d \\ v(t,x) = 1 & t \in [0,T), x \in \partial(0,+\infty)^d \end{cases}$$

where $\mathcal{A}$ is given by Eq. (A.2), then $\mathbb{P}\{\tau \leq T\} = v(0,R_0)$.

Proof. We apply Itô formula to the function v:

$$dv(t,R_t) = \left(\frac{\partial v}{\partial t} + \mathcal{A}v\right)(t,R_t)\;dt + \frac{\partial v}{\partial x}(t,R_t)\sigma\;dW_t$$

then, because of the PDE, we have:

$$v(T\wedge\tau, R_{T\wedge\tau}) = v(0,r_0) + \int_0^{T\wedge\tau}\frac{\partial v}{\partial x}(t,R_t)\sigma\;dW_t$$

making the expectation, since $v \in C_b^1$, then the expectation of the stochastic integral is 0, and we get:

$$\begin{aligned} v(0, R_0) &= \mathbb{E}[v(T \wedge \tau, R_{T\wedge\tau})] = \mathbb{E}[v(T, R_T)\mathbf{1}_{\tau > T} + v(\tau, R_\tau)\mathbf{1}_{\tau \le T}] = \\ &= \mathbb{E}[0 \cdot \mathbf{1}_{\tau > T} + 1 \cdot \mathbf{1}_{\tau \le T}] = \mathbb{P}\{\tau \le T\} \end{aligned}$$

$\diamond$

As before, the problem with this approach is that there is not an explicit solution to this PDE, so we have no idea of the behaviour of the solution.

A.3 Open problems

- generalize the two kinds of estimation of $\mathbb{P}\{\tau \le T\}$ made for the Vasicek model to the d-dimensional Gaussian model
- make a similar work for the general Gaussian Musiela model, which is infinite dimensional

Bibliography

[1] M. Avellaneda, A. Levy, and A. Parás, *Pricing and hedging derivative securities in markets with uncertain volatilities*, preprint

[2] A. Brace, D. Gatarek, and M. Musiela, *The market model of interest rate dynamics*, preprint

[3] A. Brace, and M. Musiela, *A multifactor Gauss-Markov implementation of Heath, Jarrow, and Morton*, Mathematical Finance, vol. 4, no. 3 (July 1994), 259–283

[4] J. Cox, J. Ingersoll, S. Ross, *A theory of the term structure of the interest rates*, Econometrica, **53** (1985), 385–407

[5] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, 1992

[6] N. El Karoui, H. Geman, and V. Lacoste, *On the role of state variables in interest rates models*, preprint

[7] N. El Karoui, H. Geman, and J. C. Rochet, *Changes of numeraire, changes of probability measures and option pricing*, Journal of Applied Probability **32**, 1995, 443–458.

[8] N. El Karoui, M. Jeanblanc-Pique, and S. E. Shreve, *Robustness of the Black and Scholes formula*, submitted to Math. Finance, 1995

[9] B. Goldys, and M. Musiela, *On partial differential equations related to term structure models*, preprint

[10] D. Heath, R. Jarrow and A. Morton, *Bond pricing and the term structure of interest rates: a new methodology*, Econometrica **61**, 1992, 77–105.

[11] D. Lamberton, and B. Lapeyre, *Introduction au calcul stochastique appliquée à la Finance*, Société de Mathématiques Appliquées et Industrielles, 1991

[12] J. L. Lumley, *Stochastic tools in turbulence*, New York Academic Press, 1970

[13] M. Musiela, *Stochastic PDEs and term structure models*, Journees Internationales de Finance, IGR-AFFI, La Baule, 1993

[14] S. Romagnoli, and T. Vargiolu, *Robustness of the Black–Scholes approach in the case of options on several assets*, preprint

[15] T. Vargiolu, *Invariant measures for the Musiela equation with deterministic diffusion term*, to appear in Finance and Stochastics

[16] O. Vasicek, *An equilibrium characterisation of the term structure*, Journal of Financial Economics, **5** (1977), 177–188