The Geometry of Interest Rate Models

Lecture Notes from the
Dimitsana Summer School 2005

Damir Filipović
Department of Mathematics
University of Munich
80333 Munich, Germany
Damir.Filipovic@math.lmu.de

19 June 2006

Abstract

1 Introduction

These lecture notes give a brief overview of the geometric properties of interest rate models and their finite dimensional realizations. We consider interest rate models of the Heath–Jarrow–Morton (HJM) type, where the forward rates are driven by a finite dimensional Wiener process. We will see that such models can be realized as stochastic equations in an infinite dimensional Hilbert space H of forward curves. Within this framework it is possible to find necessary and sufficient conditions for the existence of finite dimensional realizations. With this we mean models which produce forward curves that lie in a given parametrized curve family with finite dimensional parameter. Such curve families arise in connection with the estimation of the forward curve and are thus an important object in practice.

From a geometric point of view, these parametrized curve families can be considered as finite dimensional submanifolds in the Hilbert space H. Thus we are led to the problem of characterization and existence of finite dimensional invariant submanifolds for stochastic equations in infinite dimension.

These notes are structured as follows: Section 2 gives a brief introduction to the main terminology of interest rate models. This part is by no means self-comprehensive. Readers without prior knowledge of the mathematics of financial markets are advised to consult a textbook such as [1] or [12].

In Section 3 we sketch the stochastic forward rate modelling methodology introduced by Heath, Jarrow and Morton [14]. The main result here is the particular form of the drift of the forward rate dynamics. We then discuss how forward curves are estimated in the market and how this requires a certain consistency with the stochastic HJM model. We give an interpretation of the HJM forward rate dynamics as a curve-valued process.

This requires some theory of stochastc integrals and equations in a Hilbert space, which is briefly reviewed in Section 4. Here we follow the lines of [6], where all proofs can be found in detail. Due to the infinite dimensionality of H, there are several concepts of a solution to a stochastic equation: mild, weak and strong. We review them and sketch an existence and uniqueness result for weak solutions.

In Section 5 we apply these results and see how the HJM equation can be interpreted in the context of stochastic equations in Hilbert spaces in a mild sense. As a consequence, we can restate the consistency problem for HJM models as a stochastic invariance problem.

In Section 6 we recall the definition and basic properties of submanifolds in a Banach space.

In Section 7 we provide the main characterization of invariant manifolds in our context.

In Section 8 we apply the general stochastic invariance results and solve the consistency problem for HJM models stated in Section 5. The obtained consistency conditions are explicitely checked for the Nelson–Siegel, Svensson and affine families.

In Section 9 we outline the theory that is used to prove existence of finite dimensional invariant manifolds. This part uses techniques involving differential calculus and the Frobenius theorem on Fréchet spaces, which is beyond the scope of these notes. We sketch the main ideas and results.

2 Bond Markets

In this section, we find a brief introduction to the financial terminology for interest rate models.

One euro today is worth more than one euro tomorrow. The time t value of 1 euro at time $T \geq t$ is expressed by the **zero-coupon bond** with **maturity** T, $P(t,T)$, for briefty also T**-bond**. This is a contract which guarantees the holder 1 euro to be paid at the maturity date T, see Figure 1.

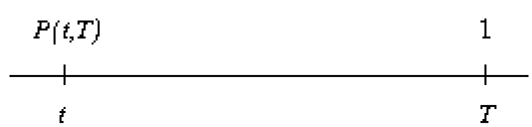

Figure 1: Cashflow of a T-bond

As a consequence, future cashflows can be discounted, such as coupon-bearing bonds

$$C_1 P(t,t_1) + \cdots + C_{n-1} P(t,t_{n-1}) + (1 + C_n) P(t,T). \tag{1}$$

In theory we will assume that

- there exists a frictionless market for T-bonds for every $T > 0$.
- $P(T, T) = 1$ for all T.
- $P(t, T)$ is continuously differentiable in T.

In reality these assumptions are not always satisfied: zero-coupon bonds are not traded for all maturities, and $P(T, T)$ might be less than one if the issuer of the T-bond defaults. Yet, this is a good starting point for doing the mathematics.

The third condition is purely technical and implies that the *term structure* of zero-coupon bond prices $T \mapsto P(t, T)$ is a smooth curve, see Figure 2 for an example.

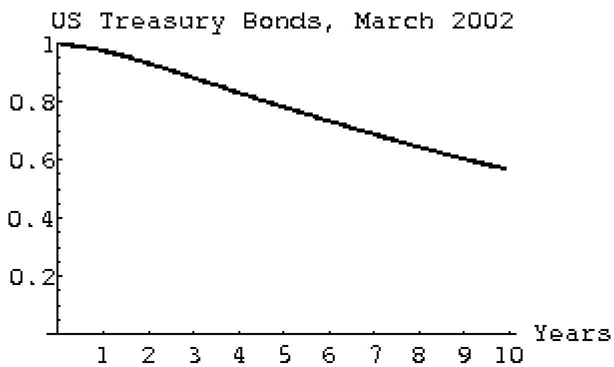

Figure 2: Term Structure $T \mapsto P(t, T)$

Note that $t \mapsto P(t, T)$ is a stochastic process since bond prices $P(t, T)$ are not known with certainty before t, see Figure 3.

A reasonable assumption would also be that $T \mapsto P(t, T) \leq 1$ is a decreasing curve (which is equivalent to positivity of interest rates). However, already classical interest rate models imply zero-coupon bond prices greater than 1. Therefore we leave away this requirement.

2.1 Interest Rates

A prototypical so-called forward rate agreement for $t < T < S$ is given by the following contractual terms:

- At t: sell one T-bond and buy $\frac{P(t,T)}{P(t,S)}$ S-bonds = zero net investment.
- At T: pay 1 euro.
- At S: obtain $\frac{P(t,T)}{P(t,S)}$ euros.

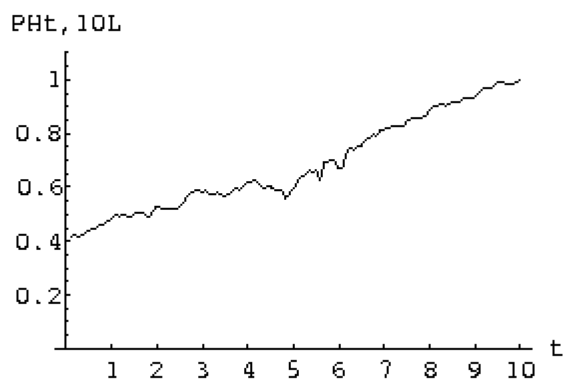

Figure 3: T-bond price process $t \mapsto P(t,T)$

The net effect is a forward investment of 1 euro at time T yielding $\frac{P(t,T)}{P(t,S)}$ euros at S. The corresponding **continuously compounded forward rate** for $[T,S]$ prevailing at t is defined as

$$e^{R(t;T,S)(S-T)} := \frac{P(t,T)}{P(t,S)} \Leftrightarrow R(t;T,S) = -\frac{\log P(t,S) - \log P(t,T)}{S-T}.$$

As we let S tend to T, we arrive at the (instantaneous) **forward rate** with maturity T prevailing at time t, which is defined as

$$f(t,T) := \lim_{S \downarrow T} R(t;T,S) = -\frac{\partial \log P(t,T)}{\partial T} \Leftrightarrow P(t,T) = \exp\left(-\int_t^T f(t,u)\,du\right).$$

The function $x \mapsto f(t,t+x)$ is called the **forward curve** at time t. We call $f(t,t)$ the (instantaneous) **short rate** at time t.

The **bank account** $B(t)$ is the asset which grows at time t instantaneously at short rate $f(t,t)$. That is,

$$dB(t) = f(t,t)B(t)dt.$$

With $B(0) = 1$ we obtain

$$B(t) = \exp\left(\int_0^t f(s,s)\,ds\right).$$

B is important for relating amounts of currencies available at different times: in order to have 1 euro in the bank account at time T we need to have

$$\frac{B(t)}{B(T)} = \exp\left(-\int_t^T f(s,s)\,ds\right)$$

euros in the bank account at time $t \leq T$. Note that this discount factor is stochastic.

3 HJM Methodology

In this section, we sketch the stochastic forward rate modelling methodology provided by Heath–Jarrow–Morton (HJM) in [14].

Throughout these notes, we fix a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, satisfying the usual conditions, carrying a d-dimensional Brownian motion W.

For every $T > 0$, we let the forward rate

$$f(t,T) = f(0,T) + \int_0^t \alpha_f(s,T)\, ds + \int_0^t \sigma_f(s,T)\, dW(s), \quad t \in [0,T] \tag{2}$$

follow an Itô process.

Heath, Jarrow and Morton [14] respond to the following question: what are sufficient conditions on the dynamics (2) such that the implied bond market

$$P(t,T) = \exp\left(-\int_t^T f(t,s)\, ds\right)$$

is arbitrage-free? With **arbitrage** we mean any self-financing strategy, that is, a predictable process

$$(\phi_1, \ldots, \phi_n),$$

yielding a profit without risk, that is,

$$V(T) := B(T) \sum_i \int_0^{T_i} \phi_i(t)\, d\left(\frac{P(t,T_i)}{B(t)}\right) \geq 0 \quad \text{and} \quad \mathbb{P}[V(T) > 0] > 0,$$

for some $n \in \mathbb{N}$ and $0 < T_1 < \cdots < T_n \leq T$. For more background on the stochastics of arbitrage-free financial markets, we refer to [1].

For illustration of what arbitrage and its absence means, we consider a deterministic world. That is, all bond prices $P(t,T)$ are known (deterministic). In this case, absence of arbitrage holds if and only if

$$P(t,S) = P(t,T)P(T,S), \quad \forall t \leq T \leq S. \tag{3}$$

This is equivalent to

$$f(t,T) = f(0,T), \quad \forall t \in [0,T]. \tag{4}$$

Indeed, suppose that $P(t,S) > P(t,T)P(T,S)$. Then selling 1 S-bond and buying $P(T,S)$ T-bonds at t results in a strict positive net gain by time S without any risk of loss. This is an arbitrage strategy, which is exluded by assumption. By changing signs, one shows that also $P(t,S) < P(t,T)P(T,S)$ is impossible, whence (3).

The **Fundamental Theorem of Asset Pricing** (proved in full generality in [7], see e.g. [1, 12]) states that there is no arbitrage if and only if there exists an equivalent probability measure $\mathbb{Q} \sim \mathbb{P}$ such that

$$\left(\frac{P(t,T)}{B(t)}\right)_{t \in [0,T]} \quad \text{is a } \mathbb{Q}\text{-local martingale, for all } T > 0. \tag{5}$$

As a consequence, under some additional technical conditions

$$P(t,T) = \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B(T)} \mid \mathcal{F}_t\right]$$

is the fair price of 1 euro at T.

Heath, Jarrow and Morton show in [14] that – under some technical conditions – (5) is equivalent to the **HJM Drift Condition**

$$\alpha_f(t,T) = \sigma_f(t,T) \cdot \int_t^T \sigma_f(t,u)\, du, \quad \forall t \le T, \tag{6}$$

under $\mathbb{Q}$ and for the respective Girsanov transform of W in (2). For pricing purposes, one therefore usually assumes that $\mathbb{Q} = \mathbb{P}$ already satsifes (5).

3.1 From HJM to Stochastic Equations

The forward curve $x \mapsto f(t, t+x)$ cannot be directly observed on the market. The forward curve has to be estimated, on a day-by-day basis, from coupon-bearing bond prices (1) and other related data. Formally, the forward curve is estimated by a parametrized family of curves $x \mapsto G(x; z)$, for some deterministic smooth function $G : \mathbb{R}_+ \times \mathcal{Z} \to \mathbb{R}$ and a finite-dimensional state space $\mathcal{Z} \subset \mathbb{R}^m$. Most prominent examples are the Nelson–Siegel family ([17])

$$G_{NS}(x; z) = z_1 + z_2 e^{-z_4 x} + z_3 x e^{-z_4 x}$$

and Svensson family ([18])

$$G_S(x; z) = z_1 + (z_2 + z_3 x)\, e^{-z_5 x} + z_4 x e^{-z_6 x}.$$

That way, a time series for $z \in \mathcal{Z}$ is observed which can be used to calibrate a stochastic model $Z(t)$. This in turn implies an accurate factor model for the forward curve, $f(t, t+x) = G(x; Z(t))$. An important question arises: can this factor modelling be made consistent with the HJM methodology outlined above? In other words, what are the conditions on $G(x; z)$ and the dynamics of $Z(t)$ such that the resulting forward curve model $f(t,T)$ is arbitrage-free, i.e. satisfies the HJM drift condition (6)?

As a first step, we have to understand the HJM dynamics of $t \mapsto f(t, t+\cdot)$ as a curve-valued process. We therefore let $\{S(t) \mid t \ge 0\}$ denote the semigroup of right shifts, $S(t)g(x) := g(x+t)$, and rewrite (2)

$$f(t,x+t) = S(t)f(0,x) + \int_0^t S(t-s)\alpha_f(s,x+s)\, ds + \int_0^t S(t-s)\sigma_f(s,x+s)\, dW(s).$$

Hence the function valued process $r(t) = r(t,\cdot) : \mathbb{R}_+ \to \mathbb{R}$ defined by

$$r(t,x) := f(t,t+x) \tag{7}$$

satisfies

$$r(t) = S(t)r(0) + \int_0^t S(t-s)\alpha(s)\,ds + \int_0^t S(t-s)\sigma(s)\,dW(s)$$

where

$$\alpha(s,x) := \alpha_f(s, s+x), \quad \sigma(s,x) := \sigma_f(s, s+x). \tag{8}$$

As we will see in Section 4 below, it follows that $r(t)$ can be interpreted as mild solution of the stochastic equation

$$dr(t) = \left(\frac{d}{dx} r(t) + \alpha(t)\right) dt + \sigma(t)\,dW(t).$$

The parametrization (7) and the interpretation of r as solution of a stochastic PDE was first suggested by Musiela [16]. An infinite dimensional stochastic analysis perspective on HJM models is also given in the recent book by Carmona and Tehranchi [4].

4 Stochastic Equations in Infinite Dimensions

In this intermediary section, we provide a short introduction to stochastic equations in infinite dimensions. A thorough treatment can be found in [6]. Throughout, we let H be a separable Hilbert space, and $\{S(t) \mid t \geq 0\}$ denotes a strongly continuous semigroup on H, that is,

$$\begin{gathered} S(t) : H \to H \text{ bounded linear}, \quad S(t+s) = S(t)S(s), \quad S(0) = Id, \\ t \mapsto S(t)h \text{ continuous for all } h \in H, \end{gathered}$$

with infinitesimal generator $A : D(A) \to H$, defined as

$$Ah = \lim_{t \to 0^+} \frac{S(t)h - h}{t}, \quad D(A) := \{h \in H \mid Ah \text{ exists in } H\}.$$

It is well known that $D(A)$ is dense in H. In fact, for any $h \in H$ we have $\int_0^t S(u)h\,du \in D(A)$ with $\lim_{t\to 0^+} \frac{1}{t}\int_0^t S(u)h\,du = h$. We denote by $A^* : D(A^*) \to H$ the adjoint of A, defined as follows:

$$D(A^*) := \{h \in H \mid g \mapsto \langle Ag, h\rangle \text{ is continuous on } D(A)\}.$$

By the Hahn–Banach theorem, for every $h \in D(A^*)$, there exists then a unique element $A^*h \in H$ with $\langle g, A^*h\rangle = \langle Ag, h\rangle$ for all $g \in D(A)$. By reflexivity of H, we have $A^{**} = A$ and $D(A^*)$ is dense in H.

We let $F : H \to H$ and $B : H \to H^d$ be continuous mappings. A **stochastic equation in** H is

$$\begin{aligned} dX(t) &= (AX(t) + F(X(t)))\,dt + B(X(t))\,dW(t) \\ X(0) &= h_0. \end{aligned} \tag{9}$$

The **stochastic integral in** H can be defined for all $Y \in \mathcal{L}$ where

$$\mathcal{L} := \left\{ Y \ H^d\text{-valued predictable} \mid \int_0^T \|Y(t)\|_{H^d}^2 dt < \infty \text{ a.s. for all } T < \infty \right\}.$$

Indeed, the construction is just as in $\mathbb{R}^d$. We remark that it is also possible to define an infinite dimensional Brownian motion and its stochastic integrals, see [6]. This, however, requires additional effort and is beyond the scope of these notes.

Write

$$\mathcal{L}_T^2 := \left\{ Y \in \mathcal{L} \mid \mathbb{E}\left[\int_0^T \|Y(t)\|_{H^d}^2 \, dt \right] < \infty \right\}.$$

We now quote four technical lemmas. For a proof see [6].

Lemma 4.1. *For $Y \in \mathcal{L}_T^2$ we have*

$$\mathbb{E}\left[\left\| \int_0^T Y(t) \, dW(t) \right\|_H^2 \right] = \mathbb{E}\left[\int_0^T \|Y(t)\|_{H^d}^2 \, dt \right].$$

Lemma 4.2 (Stochastic Fubini Theorem)**.** *Let $(E, \mathcal{E}, \mu)$ be a probability space and let*

$$Y : ([0,T] \times \Omega \times E, \mathcal{P} \otimes \mathcal{E}) \to (H^d, \mathcal{B}(H^d)), \quad (t, \omega, x) \mapsto Y(t, \omega, x)$$

be a measurable mapping with

$$\int_0^T \int_E \|Y(t, \omega, x)\|_{H^d}^2 \, \mu(dx) \, dt < \infty \quad a.s.$$

Then there exists an $\mathcal{F}_T \otimes \mathcal{E}$-measurable version of the stochastic integral $\int_0^T Y(t,x) \, dW(t)$ which is μ-integrable a.s. and

$$\int_E \int_0^T Y(t,x) \, dW(t) \, \mu(dx) = \int_0^T \int_E Y(t,x) \, \mu(dx) \, dW(t) \quad a.s.$$

Lemma 4.3. *Let $Y \in \mathcal{L}$, then*

$$Z(t) = \int_0^t S(t-s) Y(s) \, dW(s)$$

has a predictable version.

Lemma 4.4. *Let Y be an H-valued predictable process. Then the random set $\{Y \in D(A)\}$ and*

$$Z(t, \omega) := \begin{cases} AY(t, \omega), & \text{if } Y(t, \omega) \in D(A) \\ 0, & \text{else} \end{cases}$$

are predictable.

4.1 Solution Concepts

Due to the infinite dimensionality of H and the unboundedness of A, there are several concepts of a solution to (9). Indeed, let X be an H-valued predictable process and $\tau > 0$ a stopping time with

$$\int_0^{t\wedge\tau} \left(\|X(s)\|_H + \|F(X(s))\|_H + \|B(X(s))\|_{H^d}^2\right) ds < \infty \quad \text{a.s.} \quad \forall t < \infty.$$

We call X a

(i) local **mild solution** of (9) if

$$X(t) = S(t)h_0 + \int_0^t S(t-s)F(X(s))\,ds + \int_0^t S(t-s)B(X(s))\,dW(s) \quad \forall t \leq \tau$$

(ii) local **weak solution** of (9) if, for all $\zeta \in D(A^*)$,

$$\langle \zeta, X(t)\rangle = \langle \zeta, h_0\rangle + \int_0^t \left(\langle A^*\zeta, X(s)\rangle + \langle \zeta, F(X(s))\rangle\right) ds \\ + \int_0^t \langle \zeta, B(X(s))\rangle dW(s) \quad \forall t \leq \tau$$

(iii) local **strong solution** of (9) if $X \in D(A)$ $dt\otimes d\mathbb{P}$-a.s., $\int_0^{t\wedge\tau} \|AX(s)\|\,ds < \infty$ a.s. and

$$X(t) = h_0 + \int_0^t (AX(s) + F(X(s)))\,ds + \int_0^t B(X(s))\,dW(s) \quad \forall t \leq \tau.$$

The stopping time τ is called the **life time** of X. If $\tau = \infty$ then we skip the word "local" in the above definitions.

Remark 4.5. *In view of Lemma 4.3 we understand that the implicitly t-dependent stochastic integral in* (i) *is predictable. Moreover, the integrand AX in* (iii) *is to be interpreted in the sense of Lemma 4.4.*

The next lemmas state how these concepts of a solution are related.

Lemma 4.6. *strong $\Rightarrow$ weak.*

Proof. Follows from $U\int Y\,dW = \int UY\,dW$ if $U \in L(H;E)$, $Y \in \mathcal{L}$. $\square$

Lemma 4.7. *weak $\Rightarrow$ mild.*

Proof. Let $\zeta \in D(A^*)$, $\phi \in C^1([0,T];\mathbb{R})$. Then

$$d\langle \zeta\phi(t), X(t)\rangle = d\left(\langle \zeta, X(t)\rangle\phi(t)\right) \\ = \left(\langle \zeta\phi'(t) + A^*\zeta\phi(t), X(t)\rangle + \langle \zeta\phi(t), F(X(t))\rangle\right) dt + \langle \zeta\phi(t), B(X(t))\rangle dW(t)$$

Since the elements $\zeta(t) = \zeta\phi(t)$ lie dense in $C^1([0,T]; D(A^*))$ we have

$$\langle \zeta(t), X(t)\rangle = \int_0^t (\langle \zeta'(s) + A^*\zeta(s), X(s)\rangle + \langle \zeta(s), F(X(s))\rangle)\, ds + \int_0^t \langle \zeta(s), B(X(s))\rangle dW(s)$$

for all $\zeta \in C^1([0,T]; D(A^*))$. In particular, for $\zeta(s) := S^*(t-s)\zeta$ with $\zeta \in D(A^*)$, we have

$$\zeta'(s) = -A^*\zeta(s)$$

and hence

$$\langle \zeta, X(t)\rangle = \int_0^t \langle \zeta, S(t-s)F(X(s))\rangle ds + \int_0^t \langle \zeta, S(t-s)B(X(s))\rangle dW(s).$$

Since $D(A^*)$ is dense in H, the claim follows. □

Lemma 4.8. *If $B(X) \in \mathcal{L}_T^2$, then mild $\Rightarrow$ weak.*

Proof. For simplicity we assume $F = 0$. Write

$$Y(t) := \int_0^t S(t-s)B(X(s))\, dW(s).$$

By assumption the stochastic Fubini theorem 4.2 applies:

$$\begin{aligned}
\int_0^t \langle A^*\zeta, Y(s)\rangle\, ds &= \int_0^t \int_0^s \langle A^*\zeta, S(s-u)B(X(u))\rangle\, dW(u)ds \\
&= \int_0^t \left\langle A^*\zeta, \int_u^t S(s-u)B(X(u))\, ds \right\rangle dW(u) \\
&= \int_0^t \left\langle \zeta, A\int_0^{t-u} S(s)B(X(u))\, ds \right\rangle dW(u) \\
&= \int_0^t \langle \zeta, S(t-u)B(X(u)) - B(X(u))\rangle\, dW(u) \\
&= \langle \zeta, Y(t)\rangle - \int_0^t \langle \zeta, B(X(u))\rangle\, dW(u),
\end{aligned}$$

for all $\zeta \in D(A^*)$. □

In view of Lemmas 4.6–4.8 it is obvious that the above solution concepts coincide if A is bounded, which is in particular the case if H is finite-dimensional.

4.2 Existence and Uniqueness

For proving existence of a solution to (9) we will use a fix point argument.

Definition 4.9. *$G : H \to E$ is* **(locally) Lipschitz continuous** *if (for all $n \in \mathbb{N}$)*

$$\|G(x) - G(y)\|_E \leq C\|x - y\|_H$$

for all $x, y \in H$ (with $\|x\| \leq n$, $\|y\| \leq n$) and a constant C ($C = C(n)$).

The following result is from Theorem 7.4 in [6]. We will sketch the proof below.

Theorem 4.10. *Suppose F and B are Lipschitz continuous. Then, for all $h_0 \in H$, there exists a unique continuous weak solution $X = X^{h_0}$ of (9). Moreover, for every $p \geq 2$ and $T < \infty$, there exists a constant $K = K(p, T)$ with*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X(t)\|_H^p\right] \leq K\left(1 + \|h_0\|_H^p\right). \tag{10}$$

This existence result can be localized.

Corollary 4.11. *Suppose F and B are locally Lipschitz continuous. Then, for all $h_0 \in H$, there exists a unique continuous local weak solution $X = X^{h_0}$ of (9).*

Proof of Corollary 4.11. Let $h_0 \in H$. Set $R := 2\|h_0\|_H$ and define

$$\widetilde{F}(h) := F((R/\|h\|_H \wedge 1)h), \quad \widetilde{B}(h) := B((R/\|h\|_H \wedge 1)h).$$

Then $\widetilde{F}$ and $\widetilde{B}$ are Lipschitz continuous. Hence there exists a unique continuous weak solution $\widetilde{X}$ of

$$dX = \left(AX + \widetilde{F}(X)\right)dt + \widetilde{B}(X)\,dW, \quad X(0) = h_0.$$

Define the stopping time $\tau := \inf\{t \geq 0 \mid \|\widetilde{X}(t)\|_H \geq R\}$. Then $\tau > 0$ and $X(t) := \widetilde{X}(t \wedge \tau)$ is a continuous local weak solution of (9) with lifetime τ.

If X is a continuous local weak solution of (9) then, by the above arguments, it is unique on $[0, \tau_n]$ for $n \geq 2$ where $\tau_n := \inf\{t \geq 0 \mid \|X(t)\|_H \geq n\|h_0\|\}$. Now use that $\tau_n \uparrow \infty$. □

Sketch of Proof of Theorem 4.10. Uniqueness: let X_1, X_2 be two mild solutions of (9). Fix $R > 0$ and define the stopping time

$$\tau := \inf\Big\{t \leq T \mid \int_0^t \|F(X_i(s))\|_H\,ds \geq R \text{ or}$$
$$\int_0^t \|B(X_i(s))\|_{H^d}^2\,ds \geq R \text{ for } i = 1 \text{ or } i = 2\Big\}.$$

Then $X_i^\tau(t) := X_i(t\wedge\tau)$ satisfy

$$X_1^\tau(t) - X_2^\tau(t) = \int_0^{t\wedge\tau} S(t\wedge\tau - s)\left(F(X_1^\tau(s)) - F(X_2^\tau(s))\right) ds \\ + \int_0^{t\wedge\tau} S(t\wedge\tau - s)\left(B(X_1^\tau(s)) - B(X_2^\tau(s))\right) dW(s)$$

and hence

$$\begin{aligned}\mathbb{E}\left[\|X_1^\tau(t) - X_2^\tau(t)\|_H^2\right] &\le C\mathbb{E}\left[\left(\int_0^{t\wedge\tau} \|F(X_1^\tau(s)) - F(X_2^\tau(s))\|_H\, ds\right)^2\right] \\ &\quad + C\mathbb{E}\left[\int_0^{t\wedge\tau} \|B(X_1^\tau(s)) - B(X_2^\tau(s))\|_{H^d}^2\, ds\right] \\ &\le C\int_0^t \mathbb{E}\left[\|X_1^\tau(s) - X_2^\tau(s)\|_H^2\right] ds.\end{aligned}$$

Gronwall's Lemma ($0 \le f(t) \le \epsilon + M\int_0^t f(s)\,ds \Rightarrow f(t) \le \epsilon e^{Mt}$) implies $\mathbb{E}\left[\|X_1^\tau(t) - X_2^\tau(t)\|_H^2\right] = 0$. This is true for all $R > 0$, hence $X_1(t) = X_2(t)$ a.s. for all t.

Existence: Let $p \ge 2$ and define the Banach space of H-valued predictable processes $\mathcal{H}_p$ with norm

$$\|Y\|_p^p := \sup_{t\in[0,T]} \mathbb{E}\left[\|Y_t\|_H^p\right].$$

One can show that

$$\mathcal{K}(Y)(t) := S(t)h_0 + \int_0^t S(t-s)F(Y(s))\,ds + \int_0^t S(t-s)B(Y(s))\,dW(s)$$

maps $\mathcal{H}_p$ into $\mathcal{H}_p$ and $\|\mathcal{K}(Y_1) - \mathcal{K}(Y_2)\|_p \le C(T)\|Y_1 - Y_2\|_p$.

The constant $C(T)$ is independent of the initial condition h_0 and can be made < 1 for T small enough. Then $\mathcal{K}$ has a unique fix point X in $\mathcal{H}_p$. One then proceeds for $[0,T]$, $[T,2T]$,... (with random initial condition) to derive global existence.

Notice that $\mathbb{E}\left[\|X(t)\|^p\right] \le C(T,p)\left(\|h_0\|^p + \int_0^t \mathbb{E}\left[\|X(s)\|^p\right] ds\right)$, $\forall t \le T$. Hence Gronwall's Lemma implies (10). From Lemma 4.8 we deduce that X is a weak solution.

Finally, from Lemma 4.12 below we obtain a continuous modification of X, whence the theorem is proved. □

Lemma 4.12. *Let $Y \in \mathcal{L}_T^2$ with $\mathbb{E}\left[\int_0^T \|Y(t)\|_{H^d}^p\, dt\right] < \infty$, and write*

$$Z(t) := \int_0^t S(t-s)Y(s)\,dW(s), \quad Z_n(t) := e^{A_n t}\int_0^t e^{-A_n s}Y(s)\,dW(s)$$

where the bounded linear operators $A_n := nA \int e^{-nt} S(t)\,dt = nA(n-A)^{-1}$ are the Yosida approximations ($\lim_n A_n x = Ax$ if $x \in D(A)$). Then

$$\lim_n \mathbb{E}\left[\sup_{t\in[0,T]} \|Z(t) - Z_n(t)\|^p\right] = 0.$$

Hence Z has a continuous modification.

Sketch of Proof of Lemma 4.12. Let $\alpha \in (1/p, 1/2)$, and write (stochastic Fubini)

$$\begin{aligned} Z(t) &= \frac{\sin(\pi\alpha)}{\pi} \int_0^t \underbrace{\int_u^t (t-s)^{\alpha-1}(s-u)^{-\alpha}\,ds}_{=\frac{\pi}{\sin(\pi\alpha)}} \underbrace{S(t-u)}_{=S(t-s)S(s-u)} Y(u)\,dW(u) \\ &= \frac{\sin(\pi\alpha)}{\pi} \int_0^t (t-s)^{\alpha-1} S(t-s) \underbrace{\int_0^s (s-u)^{-\alpha} S(s-u) Y(u)\,dW(u)}_{=:U(s)}\,ds \end{aligned}$$

Using Hölder's inequality ($\alpha > 1/p \Rightarrow (\alpha-1)p/(p-1) > -1$), we deduce

$$\sup_{t\in[0,T]} \|Z(t)\|^p \le C \sup_{t\in[0,T]} \left(\int_0^t (t-s)^{\frac{(\alpha-1)p}{p-1}}\,ds\right)^{p-1} \int_0^T \|U(s)\|_H^p\,ds.$$

Moreover, from [6, Lemma 7.2] and Young's convolution inequality,

$$\int_0^T \mathbb{E}\left[\|U(s)\|_H^p\,ds\right] \le C\mathbb{E}\left[\int_0^T \|Y(u)\|_{H^d}^p\,du\right].$$

Define $U_n(s)$ for Z_n as above, and decompose

$$\begin{aligned} Z(t) - Z_n(t) &= \frac{\sin(\pi\alpha)}{\pi} \int_0^t \left(S(t-s) - e^{(t-s)A_n}\right)(t-s)^{\alpha-1} U(s)\,ds \\ &\quad + \frac{\sin(\pi\alpha)}{\pi} \int_0^t e^{(t-s)A_n} (t-s)^{\alpha-1} \left(U(s) - U_n(s)\right) ds =: I_n(t) + J_n(t). \end{aligned}$$

Eventually, one can show that

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|I_n(t)\|_H^p\right] \to 0 \text{ and } \mathbb{E}\left[\sup_{t\in[0,T]} \|J_n(t)\|_H^p\right] \to 0.$$

□

5 Forward Curve Space

In order to carry over the HJM methodology to the above stochastic equation framework, we need to define a reasonable Hilbert space. Since in practice the forward curve is obtained by smoothing data points using smooth fitting methods it is reasonable to assume

$$\int_{\mathbb{R}_+} |\frac{d}{dx} r(t,x)|^2 \, dx < \infty.$$

Moreover, the curve flattens for large time to maturity x. There is no reason to believe that the forward rate for an instantaneous loan that begins in 10 years differs much from one which begins one day later. We take this into account by penalizing irregularities of $r_t(x)$ for large x by some increasing weighting function $w(x) \geq 1$, that is,

$$\int_{\mathbb{R}_+} |\frac{d}{dx} r(t,x)|^2 \, w(x) dx < \infty.$$

However, this does not define a norm yet since constant functions are not distinguished. So we add the square of the short rate $|r(t,0)|^2$.

Let us recall a few facts from real analysis. Let $h \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$. The weak derivative $h' \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$ of h, if it exists, is uniquely specified by the property

$$\int_{\mathbb{R}_+} h(x)\varphi'(x) \, dx = -\int_{\mathbb{R}_+} h'(x)\varphi(x) \, dx, \quad \forall \varphi \in C^1_c((0,\infty)).$$

If h has a weak derivative h' then there exists an absolutely continuous representative of h, still denoted by h, such that

$$h(x) - h(y) = \int_y^x h'(u) \, du, \quad \forall x, y \in \mathbb{R}_+. \tag{11}$$

Accordingly, the following definition makes sense.

Definition 5.1. *Let $w \in C^1(\mathbb{R}_+; [1,\infty))$ be increasing such that*

$$w^{-\frac{1}{3}} \in L^1(\mathbb{R}_+). \tag{12}$$

We write

$$\|h\|_w^2 := |h(0)|^2 + \int_{\mathbb{R}_+} |h'(x)|^2 \, w(x) dx$$

and define

$$H_w := \left\{ h \in L^1_{loc}(\mathbb{R}_+) \mid \exists h' \in L^1_{loc}(\mathbb{R}_+) \text{ and } \|h\|_w < \infty \right\}.$$

The choice of H_w is established by the next theorem.

Theorem 5.2. *H_w equipped with $\|\cdot\|_w$ is a separable Hilbert space satisfying*

(H1) *$H_w \subset C(\mathbb{R}_+; \mathbb{R})$ and $\mathcal{J}_x(h) := h(x)$ is continuous*

(H2) *$S(t)f(x) := f(x+t)$, $t \geq 0$, is a strongly continuous semigroup with infinitesimal generator d/dx and $D(d/dx) = \{h \in H_w \mid h' \in H_w\}$*

(H3) *$\|\mathcal{S}h\|_w \le K\|h\|_w^2$ $\forall h \in H_{w,0}$ for some constant K where*

$$H_{w,0} := \{h \in H_w \mid h(\infty) = 0\}, \quad \mathcal{S}f(x) := f(x)\int_0^x f(y)\,dy. \tag{13}$$

Moreover, $\mathcal{S} : H_{w,0} \to H_{w,0}$ is locally Lipschitz continuous.

Remark 5.3. *The definition* (13) *of $\mathcal{S}$ becomes clear in view of the HJM drift condition* (6) *and* (8)*. Indeed,* (6) *simply translates to*

$$\alpha = \sum_{j=1}^{d} \mathcal{S}\sigma_j. \tag{14}$$

Examples of admissible weighting functions w which satisfy condition (12):

Example 1 $w(x) = e^{\alpha x}$, for $\alpha > 0$.

Example 2 $w(x) = (1+x)^\alpha$, for $\alpha > 3$.

5.1 HJM Revisited

Now let $\sigma : H_w \to H_{w,0}^d$ be locally Lipschitz continuous. From Theorem 5.2 we conclude that

$$\alpha := \sum_j \mathcal{S}\sigma_j : H_w \to H_w$$

is locally Lipschitz continuous. Hence σ fully determines a HJM model.

Theorem 5.4. *The continuous weak solution r (if it exists globally) of*

$$\begin{aligned} dr(t) &= \left(\frac{d}{dx}r(t) + \alpha(r(t))\right)dt + \sigma(r(t))\,dW(t) \\ r(0) &= r_0 \end{aligned} \tag{15}$$

induces an arbitrage-free bond market

$$P(t,T) = \exp\left(-\int_0^{T-t} r(t,x)\,dx\right)$$

with initial term structure $P(0,T) = \exp\left(-\int_0^T r_0(x)\,dx\right)$.

Idea of Proof. Show that $f(t,T) = r(t, T-t)$ and $P(t,T)$, $0 \le t \le T$, are Itô processes, and apply the HJM drift condition (14). See [8, Section 4]. $\square$

Remark 5.5. *Recall the deterministic case ($\sigma = 0$) in* (4)*: $\frac{d}{dt}r(t,x) = \frac{d}{dx}r(t,x)$.*

5.2 Back to the Consistency Problem

We can now restate the consistency problem of Section 3.1: is there a HJM model σ which is consistent with the Nelson–Siegel family $G = G_{NS}$? In other words, can we find an $\mathbb{R}^4$-valued diffusion process Z and a volatility map σ such that

$$r(t,x) = G(x; Z(t))$$

solves the corresponding HJM equation (15)? Björk et al. [2, 3] translate this

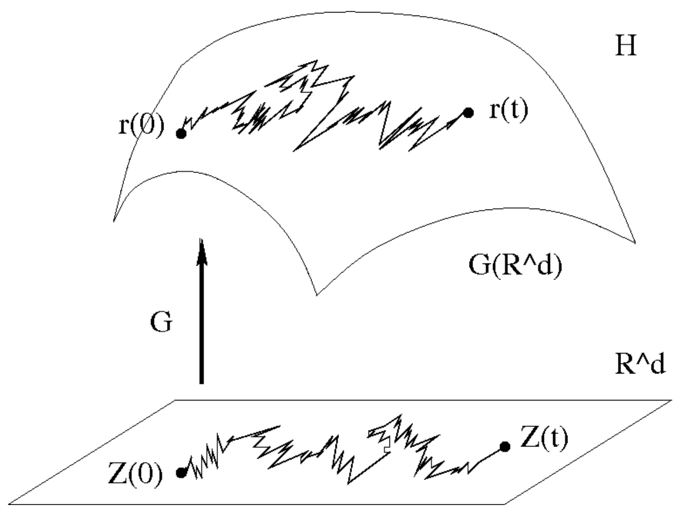

Figure 4: Geometric View of the Consistency Problem

in a geometric problem: consider $\mathcal{G} := G(\mathcal{Z})$ as submanifold in H_w and check whether $\mathcal{G}$ is invariant under the stochastic dynamic equation (15), see Figure 4. This is a **stochastic invariance problem**.

6 Submanifolds in Banach Spaces

For the convenience of the reader, we give a brief introduction to submanifolds. We follow [8, Section 6.1]. All proofs can be found there.

Let E denote a reflexive Banach space, E' its dual space, $\langle e', e\rangle$ the duality pairing. For a direct sum decomposition $E = E_1 \oplus E_2$ we denote by $\Pi_{(E_2,E_1)}$ the induced projection onto E_1.

Definition 6.1. *A subset $\mathcal{M} \subset E$ is an m-**dimensional (regular)** C^k **submanifold with boundary** of E, if for all $h \in \mathcal{M}$ there is a neighborhood U in E, an open set $V \subset \mathbb{R}^m_{\geq 0} := \{y_m \geq 0\}$ and a C^k map $\phi : V \to E$ such that*

(i) *$\phi : V \to U \cap \mathcal{M}$ is a homeomorphism, and $\phi(V \cap \{y_m = 0\}) = U \cap \partial\mathcal{M}$ (hence $\partial\mathcal{M}$ is a submanifold)*

(ii) *$D\phi(y)$ is one to one for all $y \in V$.*

The map ϕ is called a **parametrization in** *h.*

$\mathcal{M}$ is a **linear submanifold** *if for all $h \in \mathcal{M}$ there exists a linear parametrization of the form $\phi(y) = h + \sum_{i=1}^m y_i e_i$ in h.*

In what follows, $\mathcal{M}$ denotes an m-dimensional C^k submanifold with boundary of E ($k \geq 2$). Using the inverse mapping theorem, one can show that $\mathcal{M}$ shares the characterizing property of a C^k manifold:

Lemma 6.2. *Let $\phi_i : V_i \to U_i \cap \mathcal{M}$, $i = 1, 2$, be two parametrizations such that $W := U_1 \cap U_2 \cap \mathcal{M} \neq \emptyset$. Then the change of parameters*

$$\phi_1^{-1} \circ \phi_2 : \phi_2^{-1}(W) \to \phi_1^{-1}(W)$$

is a C^k diffeomorphism.

Definition 6.3. *For $h \in \mathcal{M}$ the* **tangent space** *to $\mathcal{M}$ at h is the subspace*

$$T_h\mathcal{M} := D\phi(y)\mathbb{R}^m, \quad y = \phi^{-1}(h),$$

where $\phi : V \subset \mathbb{R}^m \to \mathcal{M}$ is a parametrization in h. Moreover,

$$T_h\partial\mathcal{M} = D\phi(y)\{z_m = 0\}, \quad T_h\mathcal{M}_{\geq 0} := D\phi(y)\mathbb{R}^m_{\geq 0}, \quad h \in \partial\mathcal{M}.$$

By Lemma 6.2, the definition of $T_h\mathcal{M}$ is independent of the choice of the parametrization.

A **vector field** $X : \mathcal{M} \ni h \mapsto X(h) \in T_h\mathcal{M}$ can be represented locally as

$$X(h) = D\phi(y)\alpha(y), \quad y = \phi^{-1}(h), \quad \forall h \in U \cap \mathcal{M}, \tag{16}$$

where $\phi : V \to U \cap \mathcal{M}$ is a parametrization and α is an $\mathbb{R}^m$-valued vector field on V (uniquely determined by ϕ).

Definition 6.4. *The vector field X is* **of class** *C^r, $0 \leq r < k$, if for any parametrization ϕ the corresponding $\mathbb{R}^m$-valued vector field α in (16) is of class C^r.*

Again by Lemma 6.2 this is a well defined concept.

Remark 6.5. *The above definitions and properties of submanifolds carry over to Fréchet spaces E, which we will consider in Section 9 below. We note, however, that differential calclulus is more difficult on such general spaces.*

For further use, we may and will assume that any parametrization $\phi : V \to U \cap \mathcal{M}$ extends to $\phi \in C_b^k(\mathbb{R}^m; E)$: let $h \in U \cap \mathcal{M}$ and $y = \phi^{-1}(h)$. There exists $\epsilon > 0$ such that the open ball $B_{2\epsilon}(y) = \{v \in \mathbb{R}^m \mid |y - v| < 2\epsilon\}$ is contained in V. On $B_{2\epsilon}(y)$ one can define a function $\psi \in C^\infty(\mathbb{R}^m; [0, 1])$ satisfying $\psi \equiv 1$ on $\overline{B_\epsilon(y)}$ and $\mathrm{supp}(\psi) \subset B_{2\epsilon}(y)$. Since ϕ is a homeomorphism there exists an open neighborhood U' of h in E with $\phi(B_\epsilon(y)) = U' \cap \mathcal{M}$. Set $\tilde{\phi} := \psi\phi$. Then $\tilde{\phi} \in C_b^k(\mathbb{R}^m; E)$ and $\tilde{\phi}_{|B_\epsilon(y)} = \phi_{|B_\epsilon(y)} : B_\epsilon(y) \to U' \cap \mathcal{M}$ is a parametrization in h.

The following result is crucial for our discussion of weak solutions to stochastic equations viable in $\mathcal{M}$.

Proposition 6.6. *Let $D \subset E'$ be a dense subset. Then for any $h \in \mathcal{M}$ there exist linearly independent elements $e'_1, \ldots, e'_m$ in D and a parametrization $\phi : V \to U \cap \mathcal{M}$ in h such that*

$$\phi(\langle e'_1, z\rangle, \ldots, \langle e'_m, z\rangle) = z, \quad \forall z \in U \cap \mathcal{M}.$$

Moreover,

$$E = T_h\mathcal{M} \oplus E_2, \quad \forall h \in U \cap \mathcal{M},$$

where $E_2 := \bigcap_{i=1}^m \ker(e'_i)$, and the induced projections are given by

$$\Pi_{(E_2, T_h\mathcal{M})} = D\phi(y)(\langle e'_1, \cdot\rangle, \ldots, \langle e'_m, \cdot\rangle), \quad y = \phi^{-1}(h), \quad \forall h \in U \cap \mathcal{M}. \tag{17}$$

Idea of Proof. The idea is to find a decomposition $E = E_1 \oplus E_2$, $\dim E_1 = m$, such that E_1 is "not too far" from $T_h\mathcal{M}$ and such that

$$\Pi_{(E_2,E_1)} = \langle e'_1, \cdot\rangle e_1 + \cdots + \langle e'_m, \cdot\rangle e_m$$

with $e'_1 \ldots, e'_m \in D$. Thereby the expression "not too far" means that $\Pi_{(E_2,E_1)|T_h\mathcal{M}} : T_h\mathcal{M} \to E_1$ is an isomorphism. $\square$

Let $B \in C^1(E; E)$ be such that $B(h) \in T_h\mathcal{M}$; that is, $B_{|\mathcal{M}}$ is a C^1 vector field on $\mathcal{M}$. Let $h \in \mathcal{M}$ and $y = \phi^{-1}(h)$. Then

$$c(t) := \phi(y + tD\phi(y)^{-1}B(h))$$

satisfies

$$\frac{d}{dt}B(c(t))_{|t=0} = DB(h)B(h). \tag{18}$$

On the other hand, in view of (17) we have

$$\begin{aligned}\frac{d}{dt}B(c(t))_{|t=0} &= \frac{d}{dt}D\phi(\langle e', c(t)\rangle)\langle e', B(c(t))\rangle_{|t=0} \\ &= D^2\phi(y)(\langle e', B(h)\rangle, \langle e', B(h)\rangle) + D\phi(y)\langle e', DB(h)B(h)\rangle. \end{aligned} \tag{19}$$

We have thus proved:

Lemma 6.7.

$$DB(h)B(h) = D\phi(y)\langle e', DB(h)B(h)\rangle + D^2\phi(y)(\langle e', B(h)\rangle, \langle e', B(h)\rangle)$$

is the decomposition according to $E = T_h\mathcal{M} \oplus E_2$, for all $h \in U \cap \mathcal{M}$.

7 Invariant Manifolds

In this section we provide the main characterization of invariant manifolds in our context. In what follows, we let $\mathcal{M}$ denote an m-dimensional C^2 submanifold of H. We follow [8, Section 6.2]. All proofs can be found there.

Since H is separable, there exists a countable open covering $(U_k)_{k\in\mathbb{N}}$ of $\mathcal{M}$ and for each k a parametrization $\phi_k : V_k \subset \mathbb{R}^m \to U_k \cap \mathcal{M}$, where $\phi_k \in C_b^2(\mathbb{R}^m; H)$.

Moreover, since $D(A^*)$ is dense in H, by Proposition 6.6 we can assume that for each k there exists a linearly independent set $\{\zeta_{k,1}, \dots, \zeta_{k,m}\}$ in $D(A^*)$ such that

$$\phi_k(\langle \zeta_{k,1}, h\rangle, \dots, \langle \zeta_{k,m}, h\rangle) = h, \quad \forall h \in U_k \cap \mathcal{M}. \tag{20}$$

Notation: we write $\langle \zeta_k, h\rangle$ instead of $(\langle \zeta_{k,1}, h\rangle, \dots, \langle \zeta_{k,m}, h\rangle)$.

Now consider the stochastic equation in H

$$\begin{aligned} dX(t) &= (AX(t) + F(X(t)))\, dt + B(X(t))\, dW(t) \\ X(0) &= h_0. \end{aligned} \tag{21}$$

Assumption: $B \in C^1(H; H^d)$ and F locally Lipschitz continuous.

In view of Corollary 4.11 there exists a unique continuous local weak solution X of (21) with life time $\tau > 0$.

We know that in finite-dimensional spaces, weak solutions are strong solutions. This fact carries over to the non-linear case of a finite-dimensional manifold.

Theorem 7.1 (Regularity)**.** *Suppose*

$$X(t) \in \mathcal{M} \quad \forall t \le \tau.$$

Then X is a local strong solution of (21)*.*

Conversely, it turns out that an invariant manifold has to lie in $D(A)$. We first give a rigorous definition of stochastic invariance.

Definition 7.2. *$\mathcal{M}$ is called* **locally invariant** *for* (21) *if, for all $h_0 \in \mathcal{M}$, there exists a stopping time $\tau = \tau(h_0) > 0$ with $X^{h_0}(t) \in \mathcal{M}$ for all $t \le \tau$.*

We now obtain the familiar tangent conditions for stochastic invariance as in the finite-dimensional case.

Theorem 7.3 (Main Characterization)**.** *The following are equivalent:*

(i) *$\mathcal{M}$ is locally invariant for* (21)

(ii) *$\mathcal{M} \subset D(A)$ and*

$$Ah + F(h) - \frac{1}{2}\sum_j DB_j(h)B_j(h) \in T_h\mathcal{M} \quad (T_h\mathcal{M}_{\ge 0}) \tag{22}$$

$$B_j(h) \in T_h\mathcal{M} \quad (T_h\partial\mathcal{M}) \tag{23}$$

for all $h \in \mathcal{M}$ ($h \in \partial\mathcal{M}$).

If, in addition, $\mathcal{M}$ is closed then $X(t) \in \mathcal{M}$ for all $t \ge 0$.

A key step in proving Theorem 7.3 is the following lemma, wich is based on Lemma 6.7.

Lemma 7.4. *Suppose $U \cap \mathcal{M} \subset D(A)$, and let $\phi : V \to U \cap \mathcal{M}$ be a parametrization satisfying* (20). *Then* (22) *and* (23) *hold for all $h \in U \cap \mathcal{M}$ if and only if*

$$\begin{aligned} Ah + F(h) = D\phi(h) \left(\langle A^* \zeta, h \rangle + \langle \zeta, F(h) \rangle \right) & \\ + \frac{1}{2} \sum_j D^2 \phi(y) \left(\langle \zeta, B_j(h) \rangle \langle \zeta, B_j(h) \rangle \right) & \quad (24) \\ B_j(h) = D\phi(y) \langle \zeta, B_j(h) \rangle, & \quad (25) \end{aligned}$$

where $y = \langle \zeta, h \rangle$, for all $h \in U \cap \mathcal{M}$.

7.1 Consistency Conditions in Local Coordinates

For applications it is convenient to express the consistency conditions (22) and (23) in local coordinates.

Assume $\mathcal{M}$ is locally invariant for (21). Let $\phi : V \to U \cap \mathcal{M}$ be any parametrization, and define

$$\begin{aligned} D\phi(y)\beta(y) &:= A\phi(y) + F(\phi(y)) - \frac{1}{2} \sum_j DB_j(\phi(y)) B_j(\phi(y)), \\ D\phi(y)\rho_j(y) &:= B_j(\phi(y)) \end{aligned} \quad (26)$$

($\beta(y) \in \mathbb{R}^m_{\geq 0}$ and $\rho_j(y) \in \{z_m = 0\}$ if $y_m = 0$). As shown above

$$\begin{aligned} DB_j(\phi(y)) B_j(\phi(y)) &= \frac{d}{dt} B_j \left(\phi(y + t\rho_j(y)) \right) |_{t=0} \\ &= \frac{d}{dt} \left(D\phi(y + t\rho_j(y)) \rho_j(y + t\rho_j(y)) \right) |_{t=0} \\ &= D^2 \phi(y) \left(\rho_j(y), \rho_j(y) \right) + D\phi(y) \left(D\rho_j(y) \rho_j(y) \right) \end{aligned}$$

Plugging this in (26), we obtain

Theorem 7.5. *Consistency conditions* (22)–(23) *hold for all $h \in U \cap \mathcal{M}$ if and only if*

$$A\phi(y) + F(\phi(y)) - \frac{1}{2} \sum_j D^2 \phi(y) \left(\rho_j(y), \rho_j(y) \right) = D\phi(y) b(y) \quad (27)$$

$$B_j(\phi(y)) = D\phi(y) \rho_j(y) \quad (28)$$

for all $y \in V$, where $b(y) := \beta(y) + \frac{1}{2} \sum_j D\rho_j(y) \rho_j(y)$.

Moreover, X is a continuous local strong solution of (21) *in $U \cap \mathcal{M}$ if and only if $X = \phi(Y)$ where*

$$dY(t) = b(Y(t))\, dt + \rho(y(t))\, dW(t), \quad Y(0) = \phi^{-1}(X(0)). \quad (29)$$

8 Consistent Forward Curve Families

We now apply the above general results on stochastic invariance and solve the consistency problem for HJM models stated in Section 5.2.

Let $G \in C^2(\mathbb{R}^m; H_w)$ be a parametrized forward curve family, and suppose that $G : V \subset \mathbb{R}^m \to G(V)$ is a parametrization.

Assume $\sigma \in C^1(H_w; H^d_{w,0})$, and remember the HJM equation, with $\alpha = \sum_j \mathcal{S}(\sigma_j)$:

$$\begin{aligned} dr(t) &= \left(\frac{d}{dx} r(t) + \alpha(r(t))\right) dt + \sigma(r(t))\, dW(t) \\ r(0) &= r_0. \end{aligned} \tag{30}$$

We say that G *is consistent with the HJM model* σ if $G(V)$ is locally invariant for (30). From Theorem 7.5 we can deduce the following explicit consistency condition.

Theorem 8.1. *G is consistent with the HJM model σ if and only if there exist $b : V \to \mathbb{R}^m$ and $\rho : V \to \mathbb{R}^{m\times d}$ continuous such that*

$$\begin{aligned} \partial_x G(x,z) &= b(z) \cdot \nabla_z G(x,z) \\ &\quad + \sum_{k,l} a_{kl}(z) \left(\frac{1}{2} \partial_{z_k} \partial_{z_l} G(x,z) - \partial_{z_k} G(x,z) \int_0^x \partial_{z_l} G(y,z)\, dy \right) \end{aligned} \tag{31}$$

for all $(x,z) \in \mathbb{R}_+ \times V$, where $a := \rho^T \cdot \rho$ is the diffusion matrix.

The consistency condition (31) can be explicitely checked, and we will do this below for the Nelson–Siegel, Svensson and affine families.

8.1 Nelson–Siegel Family

Recall the form of the Nelson–Siegel curves

$$G_{NS}(x,z) = z_1 + (z_2 + z_3 x) e^{-z_4 x}.$$

The consistency conditions (31) turn out to be very restrictive in this case.

Proposition 8.2. *There is no non-trivial diffusion process Z that is consistent with the Nelson–Siegel family. In fact, the unique solution to* (31) *is*

$$a(z) = 0, \quad b_1(z) = b_4(z) = 0, \quad b_2(z) = z_3 - z_2 z_4, \quad b_3(z) = -z_3 z_4.$$

The corresponding state process is

$$\begin{aligned} Z_1(t) &\equiv z_1, \\ Z_2(t) &= (z_2 + z_3 t)\, e^{-z_4 t}, \\ Z_3(t) &= z_3 e^{-z_4 t}, \\ Z_4(t) &\equiv z_4, \end{aligned}$$

where $Z(0) = (z_1, \ldots, z_4)$ denotes the initial point.

Proof. Left to the reader. □

8.2 Svensson Family

Here the forward curve is

$$G_S(x,z) = z_1 + (z_2 + z_3 x)e^{-z_5 x} + z_4 x e^{-z_6 x}.$$

G_S has more degrees of freedom than G_{NS}. It turns out that there exists a non-trivial consistent HJM model.

Proposition 8.3. *The only non-trivial HJM model that is consistent with the Svensson family is the Hull–White extended Vasicek short rate model*

$$dr(t,0) = \left(z_1 z_5 + z_3 e^{-z_5 t} + z_4 z^{-2z_5 t} - z_5 r(t,0)\right) dt + \sqrt{z_4 z_5} e^{-z_5 t}\, dW^*(t),$$

where $(z_1,\dots,z_5)$ *are given by the initial forward curve*

$$f(0,x) = z_1 + (z_2 + z_3 x)e^{-z_5 x} + z_4 x e^{-2z_5 x}$$

and W^* *is some Brownian motion. The form of the corresponding state process* Z *is given in the proof below.*

Proof. See [8, Section 3.7.2] or [9]. □

8.3 Affine Term Structures

We now look at the simplest, namely the *affine* case:

$$G(x,z) = g_0(x) + g_1(x) z_1 + \cdots g_m(x) z_m.$$

Here the second order z-derivatives vanish, and (31) reduces to

$$\partial_x g_0(x) + \sum_{i=1}^m z_i \partial_x g_i(x) = \sum_{i=1}^m b_i(z) g_i(x) - \frac{1}{2}\partial_x \left(\sum_{i,j=1}^m a_{ij}(z) G_i(x) G_j(x) \right), \quad (32)$$

where

$$G_i(x) := \int_0^x g_i(u)\, du.$$

Integrating (32) yields

$$g_0(x) - g_0(0) + \sum_{i=1}^m z_i (g_i(x) - g_i(0)) = \sum_{i=1}^m b_i(z) G_i(x) - \frac{1}{2} \sum_{i,j=1}^m a_{ij}(z) G_i(x) G_j(x). \quad (33)$$

If $G_1,\dots,G_m$, G_1G_1, $G_1G_2,\dots,G_mG_m$ are linearly independent functions, we can invert and solve the linear equation (33) for b and a.

Since the left hand side is affine is z, we obtain that also b and a are affine

$$b_i(z) = b_i + \sum_{j=1}^m \beta_{ij} z_j$$

$$a_{ij}(z) = a_{ij} + \sum_{k=1}^m \alpha_{k;ij} z_k,$$

for some constant vectors and matrices b, β, a and α_k. Plugging this back into (33) and matching constant terms and terms containing z_ks we obtain a system of Riccati equations

$$\partial_x G_0(x) = g_0(0) + \sum_{i=1}^m b_i G_i(x) - \frac{1}{2} \sum_{i,j=1}^m a_{ij} G_i(x) G_j(x) \tag{34}$$

$$\partial_x G_k(x) = g_k(0) + \sum_{i=1}^m \beta_{ki} G_i(x) - \frac{1}{2} \sum_{i,j=1}^m \alpha_{k;ij} G_i(x) G_j(x), \tag{35}$$

with initial conditions $G_0(0) = \cdots = G_m(0) = 0$.

Notice that we have the freedom to choose $g_0(0), \ldots, g_m(0)$, which are related to the short rates by

$$r(t,0) = f(t,t) = g_0(0) + g_1(0) Z_1(t) + \cdots + g_m(0) Z_m(t).$$

A typical choice is $g_1(0) = 1$ and all the other $g_i(0) = 0$, whence $Z_1(t)$ is the (non-Markovian) short rate process.

9 Towards Existence of Invariant Manifolds

So far we have characterized the consistent HJM models for a given forward curve family. What can we say about the existence of consistent forward curve families for a given HJM model? In more general terms: given a stochastic equation (21) in H, can we find a finite-dimensional submanifold $\mathcal{M}$ satisfying the consistency conditions (22) and (23)?

The existence problem of consistent forward curve families was initiated and first solved by Björk and Svensson [3] using the Frobenius Theorem. They work on a particular Hilbert space $\mathcal{H}$ on which $A = d/dx$ is bounded, so that they can avoid weak solutions. It turns out that the space $\mathcal{H}$ consists solely of entire analytic functions, and therefore does not contain some important forward curve families, such as the Cox–Ingersoll–Ross [5] forward curves, see [10].

In [10] we thus considered the larger Fréchet spaces $D(A^\infty) := \cap_{k\geq 0} D(A^k)$ with seminorms

$$p_n(h) := \sum_{k=0}^n \|A^k h\|_H.$$

These seminorms turn $D(A^\infty)$ into a complete metric space, and A is bounded and continuous on $D(A^\infty)$. For the example $A = d/dx$ on H_w we have

$$D((\frac{d}{dx})^\infty) = \{h \in C^\infty \cap H_w \mid \frac{d^k h}{dx^k} \in H_w \text{ for all } k \geq 0\},$$

with metric

$$d(f,g) = \sum_{k\geq 0} 2^{-k} \left(\|d^k h/dx^k\|_w \wedge 1 \right).$$

But A (and any smooth function on $D(A^\infty)$) is not a contraction in general. Hence there is no fixed point theorem, and hence no existence for differential equations on $D(A^\infty)$, in general. A only generates a smooth **semi**flow.

There are more difficulties with calculus on Fréchet spaces. Indeed, $L(E,F)$ and $C^\infty(E;F)$ are no Fréchet spaces if E,F are so, in general. Kriegl and Michor [15] developped the so-called “convenient calculus” on a “convenient space” E which is by definition a locally convex vector space such that

$$c \in C^\infty(\mathbb{R};E) \quad \Leftrightarrow \quad \ell \circ c \in C^\infty(\mathbb{R}) \quad \forall \ell : E \to \mathbb{R} \text{ linear } \textit{bounded}.$$

A thorough treatment of this calculus is beyond the scope of these notes. In any case, Fréchet spaces are convenient, and the following useful facts hold true for convenient spaces E,F,G:

- $L(E,F)$ and $C^\infty(E;F)$ are convenient
- $C^\infty(E\times F;G) \tilde{=} C^\infty(E;C^\infty(F;G))$
- Taylor's formula holds on E
- the evaluation $C^\infty(E;F)\times E \to F, \quad (P,f)\mapsto P(f)$ is smooth
- the composition $C^\infty(E;F)\times C^\infty(F;G)\to C^\infty(E;G), \quad (R,Q)\mapsto Q\circ R$ is smooth.

9.1 (Semi)Flows

We provide sufficient conditions for the existence of integral curves of vector fields, such as (22) and (23) in $D(A^\infty)$. Let E be a Fréchet space.

Definition 9.1. *$P : U \subset E \to E$ is a* **Banach map** *if $P = Q\circ R$, for some $R : U \to B$ and $Q : V \subset B \to E$ smooth, for a Banach space B.*

Theorem 9.2 (Hamilton [13], Banach Map Principle)**.** *Let $P : U \subset E \to E$ be a Banach map. Then, for all $g \in U$, there exists $V = V(g) \subset U$ open and a unique smooth* **local flow** *$Fl : (-\epsilon,\epsilon)\times V \to E$ with*

$$Fl(0,g) = g, \quad \frac{d}{dt}Fl(t,g) = P(Fl(t,g)), \quad Fl(t,Fl(s,g)) = Fl(s+t,g)$$

where defined.

The following extension is proved in [10]:

Theorem 9.3. *Let $P : U \subset E \to E$ be a Banach map, A be the generator of a smooth semigroup. Then, for all $g \in U$, there exists $V = V(g) \subset U$ open and a unique smooth* **local semiflow** *$Fl : [0,\epsilon)\times V \to E$ with*

$$Fl(0,g) = g, \quad \frac{d}{dt}Fl(t,g) = AFl(t,g)+P(Fl(t,g)), \quad Fl(t,Fl(s,g)) = Fl(s+t,g)$$

where defined.

9.2 Frobenius Theorem

For stating the Frobenius theorem we need some terminology from differential geometry, which we briefly recall here. Let E be a Fréchet space, $U \subset E$ open. Let $X, Y \in C^\infty(U; E)$ be smooth vector fields on U and suppose X admits a local flow Fl^X. Then the **Lie bracket** of X and Y is

$$[X, Y] = DX \cdot Y - DY \cdot X = \frac{d}{dt}(Fl^X_{-t})^* Y|_{t=0}.$$

Recall the definition of the **pull back** $(F^*Y)(h) := DF(f)^{-1}(Y(F(h)))$ and the **push forward** $F_* = (F^*)^{-1}$ of a diffeomorphism $F : U \to F(U) \subset E$.

Definition 9.4. *A n-dimensional* **distribution** *on U is a collection of linear subspaces $\mathcal{D} = \{\mathcal{D}_h\}_{h \in U}$ such that*

$$\mathcal{D}_h = span\{X_1(h), \ldots, X_n(h)\} \quad \forall h \in U,$$

for some linearly independent smooth vector fields $X_1, \ldots, X_n$.

$\mathcal{D}$ is **involutive** *if $[X, Y] \in \mathcal{D}$ for all $X, Y \in \mathcal{D}$.*

Definition 9.5. *A n-dimensional* **weak foliation** *on U is a collection of n-dimensional C^∞ submanifolds with boundary $\Phi = \{\mathcal{M}_h\}_{h \in U}$ such that*

(i) *$r \in \mathcal{M}_r$ for all $r \in U$*

(ii) *the* **tangent distribution** *$\mathcal{D}(\Phi)(h) := span\{T_h\mathcal{M}_r \mid r \in U \text{ with } h \in \mathcal{M}_r\}$ has dimension n*

Note that a weak foliation Φ can have "gaps" and its leafs can touch.

We now can state the Frobenius theorem in our context. For a full proof we refer to [10].

Theorem 9.6. *Consider n-dimensional distribution $\mathcal{D} = span\{X_1, \ldots, X_n\}$ where $X_1, \ldots, X_{n-1}$ admit local flows and X_n admits local semiflow on U. Then the following are equivalent*

(i) *$\mathcal{D}$ is involutive*

(ii) *there exists a n-dimensional weak foliation with $\mathcal{D}(\Phi) = \mathcal{D}$.*

Idea of Proof. (i)$\Rightarrow$(ii): the map

$$\alpha : W \subset \mathbb{R}^n_{\geq 0} \times V \subset D((d/dx)^\infty) \to D((d/dx)^\infty),$$
$$\alpha(u, r) = Fl^{X_1}_{u_1} \circ \cdots \circ Fl^{X_n}_{u_n}(r)$$

is a parametrization of a weak foliation Φ with

$$\frac{\partial}{\partial u_i}\alpha(u, r) = ((Fl^{X_1}_{u_1})_* \cdots (Fl^{X_{i-1}}_{u_{i-1}})_* X_i)(\alpha(u, r)) \in \mathcal{D}_{\alpha(u,r)}.$$

Hence $\mathcal{D} = \mathcal{D}(\Phi)$.

(ii)⇒(i): we have $Fl^{X_i}(t,r) \in \mathcal{M}_r$ for $t \in (-\epsilon, \epsilon)$ ($[0,\epsilon)$ if $i = n$). Hence, for all $Y \in \mathcal{D}$, one can show that

$$[X_i, Y](h) = \frac{d}{dt}(Fl^{X_i}_{-t})^* Y(f)|_{t=0} = \frac{d}{dt} \underbrace{DFl^{X_i}_t(Fl^{X_i}_{-t}(h)) \cdot Y(Fl^{X_i}_{-t}(h))}_{\in \mathcal{D}_h} \in \mathcal{D}_h,$$

for all h in the interior $\mathcal{M}_r \setminus \partial\mathcal{M}_r$. If r itself is a boundary point, $r \in \partial\mathcal{M}_r$, one has to approximate r by a sequence of interior points $r_n \in \mathcal{M}_r \setminus \partial\mathcal{M}_r$. This way, we conclude that $\mathcal{D}$ is involutive. □

9.3 Application to HJM

In this last section, we apply the Frobenius theorem to the HJM equation (30) in H_w with $A = d/dx$ and volatility map $\sigma = (\sigma_1, \dots, \sigma_d)$. We assume

(A1) $\sigma_j = \phi_j \circ \ell$, where $\ell \in L(H_w; \mathbb{R}^p)$ and $\phi_j \in C^\infty(\mathbb{R}^p; D((d/dx)^\infty_0))$ are linearly independent

(A2) for all $q \geq 0$, the map $(\ell, \ell \circ d/dx, \dots, \ell \circ (d/dx)^q) : D((d/dx)^\infty) \to \mathbb{R}^{p(q+1)}$ is open

(A3) d/dx is unbounded.

As consequence of these assumptions one can show, see [10, Section 4]:

(i) $\alpha = \sum_j \mathcal{S}(\sigma_j)$ is a Banach map on $D((d/dx)^\infty)$

(ii) d/dx is not a Banach map on $D((d/dx)^\infty)$

(iii) $\mu := \frac{d}{dx} + \alpha - \frac{1}{2} D\sigma \cdot \sigma$ is not a Banach map on $D((d/dx)^\infty)$, but generates a smooth local semiflow Fl^μ by Theorem 9.3

(iv) If X is a Banach map then $[\frac{d}{dx}, X]$ is a Banach map. Hence the singular set
$$\mathcal{N} := \{h \mid \mu(h) \in span\{\sigma_i(h)\}\}$$
is closed and nowhere dense in $D((d/dx)^\infty)$, and the Lie algebra $\mathcal{D}_{LA}$ generated by $\mu, \sigma_1, \dots, \sigma_d$ has dimension $\dim \mathcal{D}_{LA} \geq d+1$ on $D((d/dx)^\infty) \setminus \mathcal{N}$.

Hence fix a number $N \geq d$ and some open set $U \subset D((d/dx)^\infty) \setminus \mathcal{N}$. By the Frobenius theorem 9.6 there exists a $(N+1)$-dimensional weak foliation $\Phi = \{\mathcal{M}_r\}_{r \in U}$ of locally invariant submanifolds $\mathcal{M}_r$ for the HJM model (30) if and only if

$$\dim \mathcal{D}_{LA} \equiv N + 1 \quad \text{on } U. \tag{36}$$

Remark 9.7. *The weak foliation Φ consists of submanifolds $\mathcal{M}_r$ in $D((d/dx)^\infty)$, albeit our starting point were forward curves in the larger space H_w. Is this a severe restriction? The answer is no. Indeed, we show in [11] that if $\mathcal{M}$ is locally invariant for* (30) *then necessarily $\mathcal{M} \subset D((d/dx)^\infty)$. Moreover, if $N = d$ then $\mathcal{M}$ is even a C^∞ submanifold of $D((d/dx)^\infty)$.*

It turns out that the Frobenius condition (36) has some remarkable consequences on the form of σ:

Theorem 9.8 (Characterization of invariant submanifolds for HJM)**.** *Suppose* (36) *holds. Then there exist linearly independent* $\lambda_1, \ldots, \lambda_N \in D((d/dx)^\infty)$ *such that*

$$\mathcal{D}_{LA} = span\{\mu, \lambda_1, \ldots, \lambda_N\} \quad and \quad \sigma_j \in span\{\lambda_1, \ldots, \lambda_N\} \quad on\ U.$$

Proof. See [10]. □

As a consequence, the maximal possible choice for $U \subset D((d/dx)^\infty) \setminus \mathcal{N}$ is $U = D((d/dx)^\infty) \setminus \Sigma$, where $\Sigma := \{h \mid \mu(h) \in span\{\lambda_1, \ldots, \lambda_N\}\}$. In this case, we can even say more about the structure of the HJM model:

Theorem 9.9 (Full classification of finite dimensional HJM models)**.** *Suppose* (36) *holds on* $U = D((d/dx)^\infty) \setminus \Sigma$*. Then, for all* $r_0 \in U$*,*

$$r(t) = Fl^\mu(t, r_0) + \sum_{i=1}^N Y_i(t)\lambda_i$$

is the unique continuous local solution of (30)*, for some time-inhomogeneous* $\mathbb{R}^N$*-valued diffusion process* $Y = Y^{r_0}$ *with* $Y(0) = 0$*.*

Moreover, for all $r_0 \in \Sigma$*,*

$$r(t) = r_0 + \sum_{i=1}^N Y_i(t)\lambda_i$$

for some time-homogeneous $\mathbb{R}^N$*-valued diffusion process* $Y = Y^{r_0}$ *with* $Y(0) = 0$*.*

In particular, Σ *is locally invariant for* (30)*.*

Proof. See [10]. □

As an important and striking corollary of the above theorem we can state that all generically finite dimensional HJM models are **affine models!**

References

[1] Björk, T. (2004) *Arbitrage Theory in Continuous Time*, 2nd Edition, Oxford University Press.

[2] Björk, T. and Christensen, B.J. (1999), "Interest rate dynamics and consistent forward curves," *Mathematical Finance*, 9, 323–348.

[3] Björk, T. and Svensson, L. (2001), "On the existence of finite-dimensional realizations for nonlinear forward rate models," *Mathematical Finance*, 11, 205–243.

[4] Carmona, R. and Tehranchi, M. (2006), *Interest Rate Models: an Infinite Dimensional Stochastic Analysis Perspective*, Springer.

[5] Cox, J., Ingersoll, J. and Ross, S. (1985), "A theory of the term structure of interest rates," *Econometrica*, 53, 385–408.

[6] Da Prato, G. and Zabczyk, J. (1992), *Stochastic Equations in Infinite Dimensions*, Cambridge University Press.

[7] Delbaen, F. and Schachermayer, W. (1994), "A General Version of the Fundamental Theorem on Asset Pricing," *Mathematische Annalen*, 300, 463–520.

[8] Filipovic, D. (2001), *Consistency Problems for Heath–Jarrow–Morton Interest Rate Models*, Lecture Notes in Mathematics 1760, Springer.

[9] Filipovic, D. (2005), "Interest Rate Models," Lecture Notes, University of Munich.

[10] Filipovic, D. and Teichmann, J. (2003), "Existence of invariant manfiolds for stochastic equations in infinite dimension," *Journal of Functional Analysis*, 197, 398–432.

[11] Filipovic, D. and Teichmann, J. (2003), "Regularity of finite-dimensional realizations for evolution equations," *Journal of Functional Analysis*, 197, 433–446.

[12] Föllmer, H. and Schied, A. (2002), *Stochastic Finance, An Introduction in Discrete Time*, de Gruyter Studies in Mathematics 27.

[13] Hamilton, R.S. (1982), "The inverse function theorem of Nash and Moser," *Bull. Amer. Math. Soc.*, 7, 65–222.

[14] Heath, D., Jarrow, R. and Morton, A. (1992), "Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation," *Econometrica*, 60, 77–105.

[15] Kriegl, A. and Michor, P.W. (1997), *The Convenient Setting for Global Analysis*, Surveys and Monographs, Vol. 53, Amer. Math. Soc., Providence.

[16] Musiela, M. (1993), "Stochastic PDEs and term structure models," Journées Internationales de Finance, IGR-AFFI, La Baule.

[17] Nelson, C. and Siegel, A. (1987), "Parsimonious modeling of yield curves," *J. of Business*, 60, 473–489.

[18] Svensson, L.E.O. (1994), "Estimating and interpreting forwrad interest rates: Sweden 1992–1994," IMF Working Paper No. 114.