THE GEOMETRY OF INTEREST RATE MODELS

Damir Filipović, University of Munich
www.math.lmu.de/∼filipo

Dimitsana Summer School
Stochastic Differential Geometry and Applications in Finance

10–17 July 2005

Bond Markets

1 EUR today is worth more than 1 EUR tomorrow. The time t value of 1 EUR 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 EUR to be paid at the maturity date T.

$\rightarrow$ future cashflows can be discounted, such as coupon-bearing bonds

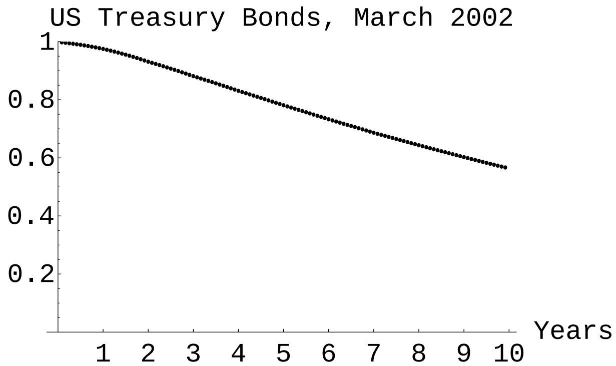
$$C_1 P(t,t_1) + \cdots + C_{n-1} P(t,t_{n-1}) + (1 + C_n) P(t,T).$$

In theory we will assume that

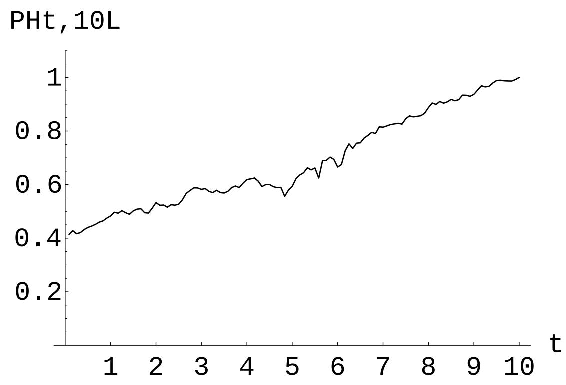
- there exists a frictionless market for T-bonds for every $T > 0$.
- $P(T,T) = 1$ for all T.
- $P(t,T)$ is continuously differentiable in T.

In reality this 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.



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



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.

Interest Rates

A prototypical forward rate agreement for $t < T < S$ is:

- 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.

Net effect: forward investment of 1 euro at time T yielding $\frac{P(t,T)}{P(t,S)}$ euros at S.

The **continuously compounded forward rate** for $[T, S]$ prevailing at t is

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

The (instantaneous) **forward rate** with maturity T prevailing at time t is

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

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

Bank Account and Short Rates

The (instantaneous) **short rate** at time t is $f(t,t)$.

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

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

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 $t \leq T$. This discount factor is stochastic!

HJM Methodology

Stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, d-dimensional Brownian motion W.

Heath–Jarrow–Morton (HJM, 1992): for all $T > 0$, let

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

follow an Itô process.

What are sufficient conditions such that the implied bond market

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

is arbitrage-free?

Arbitrage := self-financing strategy $(\phi_1, \dots, \phi_n)$ yielding a profit without risk:

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

for some $n \in \mathbb{N}$ and $0 < T_1 < \cdots < T_n \leq T$.

Exercise: in a deterministic world absence of arbitrage holds if and only if

$$P(t,S) = P(t,T)P(T,S), \quad \forall t \le T \le S.$$

This is equivalent to

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

The forward curve $x \mapsto f(t, t+x)$ is shifted!

Fundamental Theorem of Asset Pricing (Delbaen–Schachermayer 94):
There is no arbitrage if(f) there exists $\mathbb{Q} \sim \mathbb{P}$ such that

$$\left(\frac{P(t,T)}{B(t)}\right)_{t\in[0,T]} \quad \text{is a } \mathbb{Q}\text{-local martingale, for all } T > 0. \qquad (1)$$

Consequence: $P(0,T) = \mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B(T)}\right]$ = fair price of 1 EUR at T

HJM Drift Condition: (1) holds if(f) under $\mathbb{Q}$

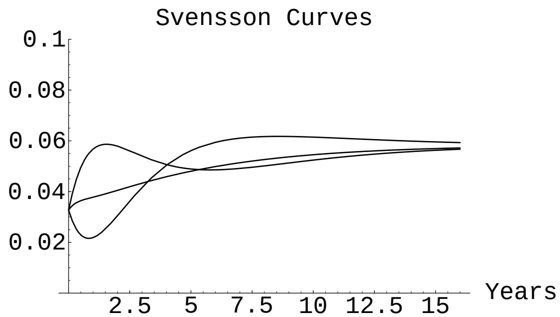
$$\alpha_f(t,T) = \sigma_f(t,T) \cdot \int_t^T \sigma_f(t,u)\,du, \quad \forall t \le T.$$

Estimating the Forward Curve

- General data: coupon-bearing bonds and related data
- $n \times N$ - cash flow matrix C: n instruments, N cash flows
- Observed prices $p \in \mathbb{R}^n$
- Term structure (discount curve) $d \in \mathbb{R}^N$
- Solve
$$\|p - C \cdot d\|^2 \to \min$$
- Problem: $n \ll N$ (too many zeros) $\to$ parametrized curve families

Estimating the Forward Curve

- Finite-dimensional state space $\mathcal{Z} \subset \mathbb{R}^m$
- Forward curves $x \mapsto G(x; z)$, for $G : \mathbb{R}_+ \times \mathcal{Z} \to \mathbb{R}$
- Nelson–Siegel (87): $G_{NS}(x; z) = z_1 + z_2 e^{-z_4 x} + z_3 x e^{-z_4 x}$
- Svensson (94): $G_S(x; z) = z_1 + (z_2 + z_3 x)\, e^{-z_5 x} + z_4 x e^{-z_6 x}$



$\rightarrow$ time series for $z \in \mathcal{Z}$ available $\rightsquigarrow$ stochastic model $Z(t)$

$\rightarrow$ $f(t, t+x) = G(x; Z(t))$ accurate factor model.
Problem: $f(t, t+x)$ is not necessarily an Itô process!

From HJM to Stochastic Equations

Let $S(t)$ be the semigroup of right shifts $S(t)g(x) := g(x+t)$ and rewrite HJM

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

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

$$r(t,x) := f(t,t+x)$$

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

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

$\to r(t)$ can be interpreted as mild solution of the stochastic equation

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

Stochastic Equations: Ingredients

- Stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, d-dimensional Brownian motion W
- H a separable Hilbert space
- $\{S(t) \mid t \geq 0\}$ a strongly continuous semigroup on H:

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

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

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

- $D(A)$ is dense in H: $\int_0^t S(u)h\,du \in D(A)$ and $\lim_{t \to 0^+} \frac{1}{t} \int_0^t S(u)h\,du = h$ for all $h \in H$
- $A^* :=$ adjoint of A: $D(A^*) := \{h \in H \mid g \mapsto \langle Ag, h\rangle \text{ continuous on } D(A)\}$,
  Hahn–Banach: $\exists! A^*h \in H$ with $\langle g, A^*h\rangle = \langle Ag, h\rangle\ \forall g \in D(A)$. $A^{**} = A$.
- $F : H \to H$, $B : H \to H^d$ continuous

Stochastic Equations

A stochastic equation in H is

$$\begin{aligned} dX(t) &= (AX(t) + F(X(t)))\,dt + B(X(t))\,dW(t) \\ X(0) &= h_0. \end{aligned} \tag{2}$$

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

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

The construction is just as in $\mathbb{R}^d$. It is possible to define infinite dimensional Brownian motion and stochastic integrals, which requires additional effort, see Da Prato and Zabczyk (92).

Write

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

Lemma 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 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

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

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*

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

Lemma 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. *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.

Solutions

Let X be an H-valued predictable process and $\tau > 0$ a stopping time with

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

We call X a

1. local **mild solution** of (2) if
$$X(t) = S(t)h_0 + \int_0^t S(t-s)F(X(s))\,ds + \int_0^t S(t-s)B(X(s))\,dW(s) \quad \forall t \le \tau$$

2. local **weak solution** of (2) if, for all $\zeta \in D(A^*)$,
$$\langle \zeta, X(t)\rangle = \langle \zeta, h_0\rangle + \int_0^t (\langle A^*\zeta, X(s)\rangle + \langle \zeta, F(X(s))\rangle)\,ds + \int_0^t \langle \zeta, B(X(s))\rangle dW(s) \quad \forall t \le \tau$$

3. local **strong solution** of (2) if $X \in D(A)$ $dt \otimes d\mathbb{P}$-a.s., $\int_0^{t\wedge\tau} \|AX(s)\|\,ds < \infty$ a.s. and
$$X(t) = h_0 + \int_0^t (AX(s) + F(X(s)))\,ds + \int_0^t B(X(s))\,dW(s) \quad \forall t \le \tau.$$

τ is called **life time** of X. If $\tau = \infty$ then skip "local".

Solutions

Lemma 5. *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 6. *weak* $\Rightarrow$ *mild*

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

$$\begin{aligned} d\langle \zeta\phi(t), X(t)\rangle &= d\left(\langle \zeta, X(t)\rangle \phi(t)\right) \\ &= \Big(\langle \zeta\phi'(t) + A^*\zeta\phi(t), X(t)\rangle + \langle \zeta\phi(t), F(X(t))\rangle\Big)\,dt + \langle \zeta\phi(t), B(X(t))\rangle dW(t) \end{aligned}$$

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

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

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. $\square$

Solutions

Lemma 7. *If $B(X) \in \mathcal{L}^2_T$, then mild $\Rightarrow$ weak*

Proof. For simplicity $F = 0$. Write

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

By assumption the stochastic Fubini theorem 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 $\zeta \in D(A^*)$. $\square$

Existence and Uniqueness

Definition 8. $G : H \to E$ *is* **(locally) Lipschitz continuous** *if (for all $n \in N$)*

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

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

Theorem 9 (DPZ Theorem 7.4). *Suppose F and B are Lischitz continuous. Then, for all $h_0 \in H$, there exists a unique continuous weak solution $X = X^{h_0}$ of (2). Moreover, for every $p \geq 2$ and $T < \infty$, there exists a constant $K = K(p, T)$ with*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \|X(t)\|_H^p\right] \leq K\left(1 + \|h_0\|_H^p\right). \tag{3}$$

Corollary 10. *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 (2).*

Proof of Corollary 10. 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

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

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 (2) with lifetime τ.

If X is a continuous local weak solution of (2) 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 $\tau_n \uparrow \infty$. $\square$

Idea of Proof of Theorem 9.

Uniqueness: let X_1, X_2 be two mild solutions of (2). Fix $R > 0$ and define the stopping time

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

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 $p > 2$ and define the Banach space $\mathcal{H}_p$ with norm

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

One shows 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 $\|\mathcal{K}(Y_1) - \mathcal{K}(Y_2)\|_p \leq C(T)\|Y_1 - Y_2\|_p$.

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 in $\mathcal{H}_p$. One then proceeds for $[0, T]$, $[T, 2T]$,... (with random initial condition) to derive global existence.

Notice: $\mathbb{E}\left[\|X(t)\|^p\right] \leq C(T,p)\left(\|h_0\|^p + \int_0^t \mathbb{E}\left[\|X(s)\|^p\right] ds\right)$, $\forall t \leq T$. Hence Gronwall's Lemma implies (3).

For the continuity of X we need the following lemma:

Lemma 11. *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 $A_n := nA\int e^{-nt}S(t)\,dt = aA(n-A)^{-1}$ *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.*

Idea of proof. Let $\alpha \in (1/p, 1/2)$, and write (Fubini!)

$$Z(t) = \frac{\sin(\pi\alpha)}{\pi}\int_0^t \underbrace{\int_u^t (t-s)^{\alpha-1}(s-u)^{-\alpha}\,ds}_{=\frac{\pi}{\sin(\pi\alpha)}} \quad \underbrace{S(t-u)}_{=S(t-s)S(s-u)} \quad Y(u)\,dW(u)$$

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

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

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

Moreover (DPZ, Lemma 7.2 + Hölder, $p > 2$ essential)

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

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

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

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

Foward Curve 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 $h \in L^1_{\text{loc}}(\mathbb{R}_+)$. The weak derivative $h' \in L^1_{\text{loc}}(\mathbb{R}_+)$ 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

$$h(x) - h(y) = \int_y^x h'(u)\,du, \quad \forall x, y \in \mathbb{R}_+. \tag{4}$$

Accordingly, the following definition makes sense.

Definition 12. *Let $w \in C^1(\mathbb{R}_+; [1,\infty))$ be increasing such that*

$$w^{-\frac{1}{3}} \in L^1(\mathbb{R}_+). \tag{5}$$

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

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

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

Examples of admissible weighting functions w which satisfy condition (5):

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

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

HJM Revisited

Let $\sigma : H_w \to H_{w,0}^d$ be locally Lipschitz continuous. Then (Theorem 13)

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

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

Theorem 14. *The continuous weak solution r (if it exists globally) of*

$$\begin{aligned} dr(t) &= \left(\frac{d}{dx} r(t) + \alpha(r(t))\right) dt + \sigma(r(t))\, dW(t) \\ r(0) &= r_0 \end{aligned} \tag{6}$$

induces an arbitrage-free bond market

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

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

Proof. Show $f(t,T) = r(t, T-t)$ and $P(t,T)$, $0 \le t \le T$, are Itô processes. □

Remark: Remember deterministic case: $\frac{d}{dt} r(t,x) = \frac{d}{dx} r(t,x)$.

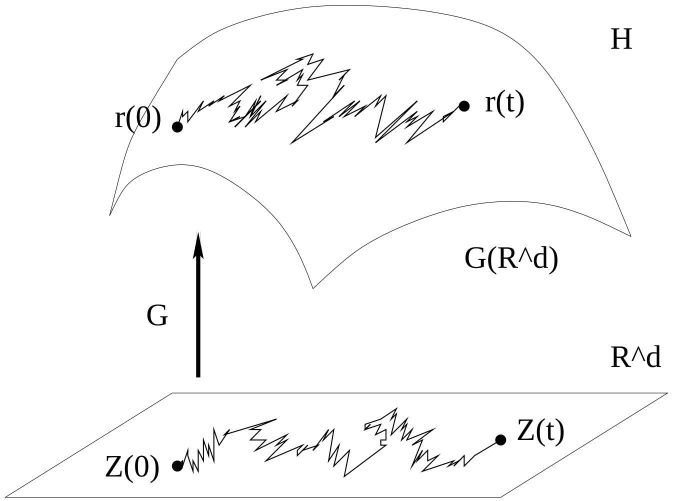
Back to the Consistency Problem

Is there a HJM model σ which is consistent with Nelson–Siegel?

That is,

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

for some $\mathbb{R}^4$-valued diffusion process Z.



Björk et al. (99): Consider $\mathcal{G} := G(\mathcal{Z})$ as submanifold in H_w $\rightarrow$ **Stochastic invariance problem**

Submanifolds in Banach Spaces

Let E denote a 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.

Let $k, m \in \mathbb{N}$. We begin with an important corollary of the inverse mapping theorem.

Proposition 15. *Let $\phi \in C^k(V; E)$, for some open set $V \subset \mathbb{R}^m$. Suppose $D\phi(y_0)$ is one to one for some $y_0 \in V$. Then $D\phi(y_0)\mathbb{R}^m$ is m-dimensional and complemented in E*

$$E = D\phi(y_0)\mathbb{R}^m \oplus E_2.$$

Moreover, there exist two open neighborhoods V' of $(y_0, 0)$ in $V \times E_2$ and U of $\phi(y_0)$ in E, and a C^k diffeomorphism $\Psi : U \to V'$ such that

$$\Psi \circ \phi(y) = (y, 0), \quad \forall y \in V' \cap (\mathbb{R}^m \times \{0\}). \tag{7}$$

Furthermore, $D\phi(y)$ is one to one and

$$D\phi(y)^{-1} = D\Psi(\phi(y))|_{D\phi(y)\mathbb{R}^m}, \quad \forall y \in V' \cap (\mathbb{R}^m \times \{0\}). \tag{8}$$

Definition 16. *The mapping ϕ from Proposition 15 is called a C^k* **immersion at** y_0*. If ϕ is a C^k immersion at each $y_0 \in V$, we just say ϕ is a C^k* **immersion***.*

*If $\phi : V \to E$ is an injective C^k immersion, $V \subset \mathbb{R}^m$ open, we call $\mathcal{M} := \phi(V)$ an m-***dimensional immersed** C^k **submanifold** *of E.*

Definition 17. *A subset $\mathcal{M} \subset E$ is an m-***dimensional (regular)** C^k **submanifold** *of E, if for all $h \in \mathcal{M}$ there is a neighborhood U in E, an open set $V \subset \mathbb{R}^m$ and a C^k map $\phi : V \to E$ such that*

1. *$\phi : V \to U \cap \mathcal{M}$ is a homeomorphism*
2. *$D\phi(y)$ 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 of E ($k \geq 2$). Then $\mathcal{M}$ shares the characterizing property of a C^k manifold:

Lemma 18. *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 19. *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.

By Lemma 18, 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

$$X(h) = D\phi(y)\alpha(y), \quad y = \phi^{-1}(h), \quad \forall h \in U \cap \mathcal{M}, \tag{9}$$

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 20. *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 (9) is of class C^r.*

Again by Lemma 18 this is a well defined concept.

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 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 on weak solutions to stochastic equations viable in $\mathcal{M}$.

Proposition 21. *Let $D \subset E'$ be a dense subset. Then for any $h \in \mathcal{M}$ there exist elements $f'_1, \dots, f'_m$ in D and a parametrization $\phi : V \to U \cap \mathcal{M}$ in h such that*

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

If $\mathcal{M}$ is linear, then ϕ is linear: $\phi(v) = e_0 + \sum_{i=1}^m (\sum_{j=1}^m N_{ij} v_j) e_i$, for $v \in V$.

Proof. The idea is to find a decomposition $E = F_1 \oplus F_2$, $\dim F_1 = m$, such that F_1 is "not too far" from $T_h\mathcal{M}$ and such that

$$\Pi_{(F_2,F_1)} = \langle f'_1, \cdot\rangle f_1 + \cdots + \langle f'_m, \cdot\rangle f_m$$

with $f'_1 \dots, f'_m \in D$. Thereby the expression "not too far" means that $\Pi_{(F_2,F_1)|T_h\mathcal{M}} : T_h\mathcal{M} \to F_1$ is an isomorphism. □

Lemma 22. *Let $\phi : V \to U \cap \mathcal{M}$ be a parametrization. If there exist elements $e'_1, \ldots, e'_m$ in E' with the property that*

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

then $e'_1, \ldots, e'_m$ are linearly independent in E' and

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

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

$$\Pi_{(E_2, T_h\mathcal{M})} = D\phi(y)(\langle e'_1, \cdot\rangle, \ldots, \langle e'_m, \cdot\rangle), \quad y = \phi^{-1}(h), \quad \forall h \in U \cap \mathcal{M}. \tag{10}$$

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

$$\frac{d}{dt}B(c(t))_{|t=0} = DB(h)B(h). \tag{11}$$

On the other hand, in view of (10) we have

$$\begin{aligned}\frac{d}{dt}B(c(t))_{|t=0} &= \frac{d}{dt}D\phi(\langle e', c(t)\rangle)\langle e', B(c(t))\rangle_{|t=0} \\ &= D^2\phi(y)(\langle e', B(h)\rangle, \langle e', B(h)\rangle) + D\phi(y)\langle e', DB(h)B(h)\rangle. \end{aligned} \tag{12}$$

Proposition 23.

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

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

Invariant Manifolds

Let $\mathcal{M}$ denote an m-dimensional C^2 submanifold of H.

Since H is separable, by Lindelöf's Lemma 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 $\phi_k \in C_b^2(\mathbb{R}^m; H)$.

Since $D(A^*)$ is dense in H, by Proposition 21 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

$$\phi_k(\langle \zeta_{k,1}, h \rangle, \dots, \langle \zeta_{k,m}, h \rangle) = h, \quad \forall h \in U_k \cap \mathcal{M}. \tag{13}$$

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

Consider the stochastic equation in H

$$\begin{aligned} dX(t) &= (AX(t) + F(X(t)))\, dt + B(X(t))\, dW(t) \\ X(0) &= h_0. \end{aligned} \tag{14}$$

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

Corollary 10: $\exists!$ continuous local weak solution X of (14) with life time $\tau > 0$.

Theorem 24 (Regularity). *Suppose*

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

Then X is a local strong solution of (14).

Theorem 25 (Consistency Conditions). *Suppose $\mathcal{M}$ is closed, $\mathcal{M} \subset D(A)$, $h_0 \in \mathcal{M}$ and*

$$Ah + F(h) - \frac{1}{2} \sum_j DB_j(h) B_j(h) \in T_h \mathcal{M} \tag{15}$$

$$B_j(h) \in T_h \mathcal{M} \tag{16}$$

for all $h \in \mathcal{M}$. Then $X(t) \in \mathcal{M}$ for all $t \geq 0$.

Definition 26. *$\mathcal{M}$ is called* **locally invariant** *for* (14) *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 $t \leq \tau$.*

Theorem 27 (Main Characterization). *The following are equivalent:*

1. *$\mathcal{M}$ is locally invariant for* (14)

2. *$\mathcal{M} \subset D(A)$ and* (15)–(16) *hold for all $h \in \mathcal{M}$.*

A key step in proving Theorems 25 and 27 is the following

Lemma 28. *Suppose $U \cap \mathcal{M} \subset D(A)$, and let $\phi : V \to U \cap \mathcal{M}$ be a parametrization satisfying* (13)*. Then* (15) *and* (16) *hold for all $h \in U \cap \mathcal{M}$ if and only if*

$$Ah + F(h) = D\phi(h) \left(\langle A^* \zeta, h \rangle + \langle \zeta, F(h) \rangle \right) + \frac{1}{2} \sum_j D^2 \phi(y) \left(\langle \zeta, B_j(h) \rangle \langle \zeta, B_j(h) \rangle \right) \tag{17}$$

$$B_j(h) = D\phi(y) \langle \zeta, B_j(h) \rangle, \tag{18}$$

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

Consistency Conditions in Local Coordinates

Assume $\mathcal{M}$ is locally invariant for (14). Let $\phi : V \to U \cap \mathcal{M}$ be a parametrization, and define

$$D\phi(y)\beta(y) := A\phi(y) + F(\phi(y)) - \frac{1}{2}\sum_j DB_j(\phi(y))B_j(\phi(y)), \quad D\phi(y)\rho_j(y) := B_j(\phi(y)). \quad (19)$$

As shown above

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

Plug this in (19), we obtain

Theorem 29. *Consistency conditions* (15)–(16) *hold for all* $h \in U \cap \mathcal{M}$ *if and only if*

$$A\phi(y) + F(\phi(y)) - \frac{1}{2}\sum_j D^2\phi(y)\left(\rho_j(y), \rho_j(y)\right) = D\phi(y)b(y) \quad (20)$$

$$B_j(\phi(y)) = D\phi(y)\rho_j(y) \quad (21)$$

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* (14) *in* $U \cap \mathcal{M}$ *if and only if* $X = \phi(Y)$ *where*

$$dY(t) = b(Y(t))\,dt + \rho(y(t))\,dW(t), \quad Y(0) = \phi^{-1}(X(0)). \quad (22)$$

Consistent Forward Curve Families

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: $\alpha = \sum_j \mathcal{S}(\sigma_j)$

$$\begin{aligned} dr(t) &= \left(\frac{d}{dx} r(t) + \alpha(r(t))\right) dt + \sigma(r(t))\, dW(t) \\ r(0) &= r_0. \end{aligned} \tag{23}$$

Theorem 30. *G is consistent with the HJM model σ (:⇔ $G(V)$ is locally invariant for (23)) if and only if there exist $b : V \to \mathbb{R}^m$ and $\rho : V \to \mathbb{R}^{m\times d}$ continuous such that*

$$\partial_x G(x,z) = b(z) \cdot \nabla_z G(x,z) + \sum_{k,l} a_{kl}(z) \left(\frac{1}{2} \partial_{z_k} \partial_{z_l} G(x,z) - \partial_{z_k} G(x,z) \int_0^x \partial_{z_l} G(y,z)\, dy \right) \tag{24}$$

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

The consistency condition (24) can be explicitely checked!

Nelson–Siegel Family

Recall the form of the Nelson–Siegel curves

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

Proposition 31. *There is no non-trivial diffusion process Z that is consistent with the Nelson–Siegel family. In fact, the unique solution to* (24) *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. Exercise. □

Svensson Family

Here the forward curve is

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

Proposition 32. *The only non-trivial HJM model that is consistent with the Svensson family is the Hull–White extended Vasicek short rate model*

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

where $(z_1, \ldots, z_5)$ *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. The consistency equation (24) becomes

$$q_1(x) + q_2(x)e^{-z_5 x} + q_3(x)e^{-z_6 x} + q_4(x)e^{-2z_5 x} + q_5(x)e^{-(z_5+z_6)x} + q_6(x)e^{-2z_6 x} = 0, \tag{25}$$

for some polynomials $q_1, \ldots, q_6$. Indeed, we assume for the moment that

$$z_5 \neq z_6, \quad z_5 + z_6 \neq 0 \quad \text{and } z_i \neq 0 \text{ for all } i = 1, \ldots, 6. \tag{26}$$

Then the terms involved in (24) are

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

$$\nabla_z G_S(x,z) = \begin{pmatrix} 1 \\ e^{-z_5 x} \\ x e^{-z_5 x} \\ x e^{-z_6 x} \\ (-z_2 x - z_3 x^2)e^{-z_5 x} \\ -z_4 x^2 e^{-z_6 x} \end{pmatrix}, \quad \partial_{z_i}\partial_{z_j} G_S(x,z) = 0 \quad \text{for } 1 \le i,j \le 4,$$

$$\nabla_z \partial_{z_5} G_S(x,z) = \begin{pmatrix} 0 \\ -x e^{-z_5 x} \\ -x^2 e^{-z_5 x} \\ 0 \\ (z_2 x^2 + z_3 x^3)e^{-z_5 x} \\ 0 \end{pmatrix}, \quad \nabla_z \partial_{z_6} G_S(x,z) = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -x^2 e^{-z_6 x} \\ 0 \\ z_4 x^3 e^{-z_6 x} \end{pmatrix},$$

$$\int_0^x \nabla_z G_S(u,z)\,du = \begin{pmatrix} x \\ -\frac{1}{z_5} e^{-z_5 x} + \frac{1}{z_5} \\ \left(-\frac{x}{z_5} - \frac{1}{z_5^2}\right) e^{-z_5 x} + \frac{1}{z_5^2} \\ \left(-\frac{x}{z_6} - \frac{1}{z_6^2}\right) e^{-z_6 x} + \frac{1}{z_6^2} \\ \left(\frac{z_3}{z_5} x^2 + \left(\frac{z_2}{z_5} + \frac{2z_3}{z_5^2}\right) x + \frac{z_2}{z_5^2} + \frac{2z_3}{z_5^3}\right) e^{-z_5 x} - \frac{z_2}{z_5^2} - \frac{z_3}{z_5^3} \\ \left(\frac{z_4}{z_6} x^2 + \frac{2z_4}{z_6^2} x + \frac{2z_4}{z_6^3}\right) e^{-z_6 x} - \frac{z_4}{z_6^3} \end{pmatrix}.$$

Straightforward calculations lead to

$$q_1(x) = -a_{11}(z)x + \cdots ,$$
$$q_2(x) = a_{55}(z)\frac{z_3^2}{z_5}x^4 + \cdots ,$$
$$q_3(x) = a_{66}(z)\frac{z_4^2}{z_6}x^4 + \cdots ,$$
$$\deg q_4,\ \deg q_5\ \deg q_6 \le 3,$$

where $\cdots$ stands for lower order terms in x. Because of (26) we conclude that

$$a_{11}(z) = a_{55}(z) = a_{66}(z) = 0.$$

But a is a positive semi-definite symmetric matrix. Hence

$$a_{1j}(z) = a_{j1}(z) = a_{5j}(z) = a_{j5}(z) = a_{6j}(z) = a_{j6}(z) = 0 \quad \forall j = 1\ldots,6.$$

Taking this into account, expression (25) simplifies considerably. We are left with

$$q_1(x) = b_1(z),$$
$$\deg q_2(x),\ \deg q_3 \le 1,$$
$$q_4(x) = a_{33}(z)\frac{1}{z_5}x^2 + \cdots ,$$
$$q_5(x) = a_{34}(z)\left(\frac{1}{z_5} + \frac{1}{z_6}\right)x^2 + \cdots ,$$
$$q_6(x) = a_{44}(z)\frac{1}{z_6}x^2 + \cdots .$$

Because of (26) we know that the exponents $-2z_5$, $-(z_5+z_6)$ and $-2z_6$ are mutually different. Hence

$$b_1(z) = a_{3j}(z) = a_{j3}(z) = a_{4j}(z) = a_{j4}(z) = 0 \quad \forall j = 1, \dots 6.$$

Only $a_{22}(z)$ is left as strictly positive candidate among the components of $a(z)$. The remaining terms are

$$\begin{aligned} q_2(x) &= (b_3(z) + z_3 z_5)x + b_2(z) - z_3 - \frac{a_{22}(z)}{z_5} + z_2 z_5, \\ q_3(x) &= (b_4(z) + z_4 z_6)x - z_4, \\ q_4(x) &= a_{22}(z)\frac{1}{z_5}, \end{aligned}$$

while $q_1 = q_5 = q_6 = 0$.

If $2z_5 \neq z_6$ then also $a_{22}(z) = 0$. If $2z_5 = z_6$ then the condition $q_3 + q_4 = q_2 = 0$ leads to

$$\begin{aligned} a_{22}(z) &= z_4 z_5, \\ b_2(z) &= z_3 + z_4 - 2_5 z_2, \\ b_3(z) &= -z_5 z_3, \\ b_4(z) &= -2 z_5 z_4. \end{aligned}$$

We derived the above results under the assumption (26). But the set of z where (26) holds is dense $\mathcal{Z}$. By continuity of $a(z)$ and $b(z)$ in z, the above results thus extend for all $z \in \mathcal{Z}$. In particular, all Z_i's but Z_2 are deterministic; Z_1, Z_5 and Z_6 are even constant.

Thus, since

$$a(z) = 0 \quad \text{if } 2z_5 \neq z_6,$$

we only have a non-trivial process Z if

$$Z_6(t) \equiv 2Z_5(t) \equiv 2Z_5(0).$$

In that case we have, writing shortly $z_i = Z_i(0)$,

$$\begin{aligned} Z_1(t) &\equiv z_1, \\ Z_3(t) &= z_3 e^{-z_5 t}, \\ Z_4(t) &= z_4 z^{-2z_5 t} \end{aligned}$$

and

$$dZ_2(t) = \left(z_3 e^{-z_5 t} + z_4 z^{-2z_5 t} - z_5 Z_2(t) \right) dt + \sum_{j=1}^{d} \rho_{2j}(t)\, dW_j(t),$$

where $\rho_{2j}(t)$ (not necessarily deterministic) are such that

$$\sum_{j=1}^{d} \rho_{2j}^2(t) = a_{22}(Z(t)) = z_4 z_5 e^{-2z_5 t}.$$

By Lévy's characterization theorem we have that

$$W^*(t) := \sum_{j=1}^{d} \int_0^t \frac{\rho_{2j}(s)}{\sqrt{z_4 z_5} e^{-z_5 s}} \, dW_j(s)$$

is a real-valued standard Brownian motion (→ exercise). Hence the corresponding short rate process

$$r(t,0) = G_S(0, Z(t)) = z_1 + Z_2(t)$$

satisfies

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

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 (24) reduces to

$$\partial_x g_0(x) + \sum_{i=1}^{m} z_i \partial_x g_i(x) = \sum_{i=1}^{m} b_i(z) g_i(x) - \frac{1}{2}\partial_x \left(\sum_{i,j=1}^{m} a_{ij}(z) G_i(x) G_j(x) \right), \tag{27}$$

where

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

Integrating (27) yields

$$g_0(x) - g_0(0) + \sum_{i=1}^{m} z_i(g_i(x) - g_i(0)) = \sum_{i=1}^{m} b_i(z) G_i(x) - \frac{1}{2} \sum_{i,j=1}^{m} a_{ij}(z) G_i(x) G_j(x). \tag{28}$$

If $G_1, \ldots, G_m$, G_1G_1, $G_1G_2, \ldots, G_mG_m$ are linearly independent functions, we can invert and solve the linear equation (28) 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 (28) and matching constant terms and terms containing z_ks we obtain a system of Riccati equations

$$\partial_x G_0(x) = g_0(0) + \sum_{i=1}^{m} b_i G_i(x) - \frac{1}{2} \sum_{i,j=1}^{m} a_{ij} G_i(x) G_j(x) \tag{29}$$

$$\partial_x G_k(x) = g_k(0) + \sum_{i=1}^{m} \beta_{ki} G_i(x) - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_{k;ij} G_i(x) G_j(x), \tag{30}$$

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.