FORWARD RATES MODELS ON THE SPACE OF SQUARE INTEGRABLE FUNCTIONS

1 Introduction

The history of modeling forward rates goes back to the paper [8] by Heath, Jarrow and Morton, who made the assumption that for every \( T > 0 \) the forward rate process \( \{ f(t,T) : t \in [0,T] \} \) is an Itô process:

\[
  f(t,T) = f(0,T) + \int_0^t a(s,T)ds + \int_0^t \langle b(s,T), dZ_s \rangle_U,
\]

where \( \{a(t,T) : t \in [0,T]\} \) is an \( \mathbb{R} \)-valued process and \( \{b(t,T) : t \in [0,T]\} \) is a \( U \)-valued process. Although in [8] \( U = \mathbb{R}^d \) and \( Z \) is a \( d \)-dimensional Wiener process, rate models with \( Z \) being an infinite dimensional Lévy process seems to capture more of the relevant features of the markets. The absence of arbitrage on the market implies that the following dependence between the coefficients in (1.1) holds (see [3]):

\[
  \int_t^T a(t,\xi)d\xi = J \left( \int_t^T b(t,\xi)d\xi \right),
\]
where \( J(u) = \ln \mathbb{E} e^{-<u,Z>_t} \).

Musiela [9] proposed to define rates in terms of the remaining time to maturity \( x = T - t \). With

\[
\begin{align*}
  f_t(x) &= f(t,t + x), & a_t(x) &= a(t,t + x), & b_t(x) &= b(t,t + x), \\
(1.1)
\end{align*}
\]

(1.1) becomes

\[
\begin{align*}
  f_t(x) &= f_0(x + t) + \int_0^t a_s(x + t - s)ds + \int_0^t \langle b_s(x + t - s), dZ_s \rangle U. \\
(1.3)
\end{align*}
\]

For \( \alpha > 0 \) let \( L^2_{\alpha} \) denote the space of all \( f : \mathbb{R}_+ \to \mathbb{R} \) such that

\[
\|f\|_{L^2_{\alpha}} = \int_0^{+\infty} |f(x)|^2 e^{\alpha x} dx < +\infty.
\]

We prove that if \( a = \{a_s : s \in [0,t]\} \) and \( b = \{b_s : s \in [0,t]\} \) are predictable integrable \( L^2_{\alpha} \)-valued processes such that for some \( K(t) > 0 \), we have \( \|a_s\|_{L^2_{\alpha}}, \|b_s\|_{L^2_{\alpha}} \leq K(t) \) for all \( s \in [0,t] \), then \( f_t \) is a mild solution to the following equation:

\[
\begin{align*}
  f_t &= (Af_t + a_t)dt + b_t dZ_t, \\
(1.4)
\end{align*}
\]

where \( Af = f' \) (in Theorem 2.4 the result is formulated for an infinite dimensional Lévy process \( Z \)).

The difficulties of performing stochastic analysis in \( L^2_{\alpha} \) were mentioned already in [5] (see remarks after Example 3.16 therein), where the weighted Sobolev space \( \mathcal{H} \) was introduced and proposed as a state space for (1.4).

On the state space \( \mathcal{H} \) the transition from the original Heath-Jarrow-Morton-Musiela description (1.3) to the stochastic differential equation (1.4) was a consequence of the boundedness of the point evaluations \( J_x \), given by \( J_x f = f(x) \). In contrast, the point evaluations \( J_x \) on \( L^2_{\alpha} \) fail to be bounded. Various results regarding Musiela equation (1.4) has been presented for \( \mathcal{H} \) as well as \( L^2_{\alpha} \) (see for instance [1], [2], [6], [7], [10], [11] and [12]), although for \( L^2_{\alpha} \) the first step (the transition to a stochastic equation) of the research was missing.

Heath, Jarrow and Morton [8] present an example of a model with state dependent coefficient \( b \), namely

\[
\begin{align*}
  b(t,T) &= \sigma \min\{f(t,T), \lambda\}, & T \geq 0, t \in [0,T], \\
(1.5)
\end{align*}
\]

for some \( \sigma, \lambda > 0 \).

Let \( \tau > 0 \) and \( W \) be a one-dimensional Wiener process. It follows from
Propositions 4 and 5 of [8] that for an arbitrary initial forward curve $\eta$ such that $\eta(t) > 0$ for all $t \in [0, \tau]$, there exists a jointly continuous $f(t, T)$, $t \in [0, T]$, $T \in [0, \tau]$ which solves

$$df(t, T) = b(t, T) \left( \int_t^T b(t, x) ds \right) dt + b(t, T) dW_t,$$

$$f(0, T) = \eta(T),$$

and with probability one $f(t, T) \geq 0$ for all $T \in [0, \tau]$ and $t \in [0, T]$.

The above example is discussed further in [5]: when we move to the stochastic equation setting (1.4) the example does not work in $H$, but works in $L^2$ (see Example 3.16 of [5]).

We show that if $b_t$ in (1.4) is given by

$$b_t(x) = \max \{ f_t(x), \lambda(x) \}, \quad (1.6)$$

for some positive $\lambda \in L^2_\alpha$ and $Z$ is a square integrable Lévy process such that its jumps are bounded from below by $-1$, then for every positive $\eta \in L^2_\alpha$, there exists a unique positive solution to (1.4) with $f_0 = \eta$.

## 2 Stochastic integral with $L^2$-valued operators

Let $U$, $H$ be two separable Hilbert spaces. A linear operator $A \in L(U, H)$ is said to belong to the space of Hilbert-Schmidt operators, denoted by $L^2(U, H)$, if

$$\|A\|_{L^2(U, H)}^2 = \sum_{i=1}^{+\infty} \|Ae_i\|_H^2 < +\infty,$$

where $\{e_i\}_i$ is an orthonormal basis in $U$.

Let $L^2(Y, \mu, U)$ denote the Hilbert space of all functions $f : Y \to U$ such that

$$\int_0^{+\infty} \|f(y)\|^2_U \mu(dy) < +\infty,$$

with the standard inner product

$$\langle f, g \rangle_{L^2(Y, \mu, U)} = \int_Y \langle f(y), g(y) \rangle_U \mu(dy).$$

It is clear that every $\gamma \in L^2(Y, \mu, U)$ defines a Hilbert-Schmidt operator from $U$ into $L^2(Y, \mu, \mathbb{R})$ by $(Au)(y) = \langle \gamma(y), u \rangle_U$. The lemma below ensures that every $A \in L^2(U, L^2(Y, \mu, \mathbb{R}))$ admits such representation. The lemma can be found in [14] (see Theorem 6.12 therein).
Lemma 2.1. If $A \in \mathcal{L}^2(U, L^2(Y, \mu, \mathbb{R}))$, then there exists $\gamma \in L^2(Y, \mu, U)$ such that for almost every $y \in Y$

$$(Au)(y) = \langle \gamma(y), u \rangle_U.$$  

Furthermore,

$$\|A\|_{\mathcal{L}^2(U, L^2(Y, \mu, \mathbb{R}))} = \|\gamma\|_{L^2(Y, \mu, U)}.$$  

From the above lemma, for every $A \in \mathcal{L}^2(U, L^2(Y, \mu, \mathbb{R}))$, mapping $\mathcal{J}_\gamma \circ A : U \to \mathbb{R}$ is a bounded linear functional for almost every $y \in Y$. By $A^* \mathcal{J}_\gamma$ we shall denote the unique element of $U$ such that

$$\mathcal{J}_\gamma(Au) = \langle A^* \mathcal{J}_\gamma, u \rangle_U.$$  

Theorem 2.2. Suppose $M$ is a $U$-valued martingale and $\Psi = \{\Psi_s : s \in [0, t]\}$ is a predictable integrable $L^2(U, L^2(Y, \mu, \mathbb{R}))$-valued process such that

$$\|\Psi_s\|_{\mathcal{L}^2(U, L^2(Y, \mu, \mathbb{R}))} \leq K(t) \quad \text{for all } s \in [0, t] \text{ and some } K(t) > 0,$$

Then for almost every $y \in Y$,

$$\left(\int_0^t \Psi_s \, dM_s\right)(y) = \int_0^t \langle \Psi_s, \mathcal{J}_\gamma \rangle_U \, ds.$$  

Proof of Theorem 2.2. From the definition of the stochastic integral there exists a sequence of elementary processes $\{\Phi^n_s : s \in [0, t]\}_{n \in \mathbb{N}}$ such that the sequence $\|\Psi_s(\omega) - \Phi^n_s(\omega)\|_{L^2(Y, \mu, \mathbb{R})}$ decreases to 0, for all $\omega \in \Omega$, $s \in [0, t]$, and

$$\mathbb{E} \int_0^t \|\Psi_s - \Phi^n_s\|_{\mathcal{L}^2(U, L^2(Y, \mu, \mathbb{R}))}^2 \, ds \to 0. \tag{2.1}$$

It follows from Lemma 2.1 that the conclusion of the theorem holds for any elementary process, hence we only need to show that

$$\int_Y \mathbb{E} \left[ \int_0^t \langle \psi_s(y), dM_s \rangle_U - \int_0^t \langle \varphi^n_s(y), dM_s \rangle_U \right]^2 \mu(dy) \to 0,$$

where $\langle \varphi^n_s(y), u \rangle_U = (\Phi^n_s u)(y)$ and $\langle \psi_s(y), u \rangle_U = (\Psi u)(y)$. The existence of $\varphi^n_s, \psi_s$ follows from Lemma 2.1. It is enough to prove that

$$x_n = \int_Y \left( \mathbb{E} \int_0^t \|\psi_s(y) - \varphi^n_s(y)\|_U^2 \, ds \right) \mu(dy) \to 0,$$

and from Fubini’s theorem for $\sigma$-finite measures, we get

$$x_n = \mathbb{E} \int_0^t \int_Y \|\psi_s(y) - \varphi^n_s(y)\|_U^2 \mu(dy) \, ds.$$

But from Lemma 2.1 $\|\psi_s - \varphi^n_s\|_{L^2(Y, \mu, U)} = \|\Psi_s - \Phi^n_s\|_{\mathcal{L}^2(U, L^2(Y, \mu, \mathbb{R}))}$, so $x_n \to 0$ by (2.1).\[\square\]
Following the proof of Theorem 2.2 but using Bochner integral definition instead of stochastic integral definition we get the following lemma.

**Lemma 2.3.** If \( \phi = \{ \phi_s : s \in [0, t] \} \) is a predictable integrable \( L^2(Y, \mu, U) \)-valued process such that
\[
\| \phi_s \|_{L^2(Y, \mu, U)} \leq K(t) \quad \text{for all } s \in [0, t] \text{ and some } K(t) > 0,
\]
Then for almost every \( y \in Y \),
\[
\left( \int_0^t \phi_s ds \right)(y) = \int_0^t \phi_s(y) ds.
\]

The following result is a direct consequence of Theorem 2.2 and Lemma 2.3.

**Theorem 2.4.** Let \( t \geq 0 \). Suppose \( f_0 \in L^2_\alpha \), \( a = \{ a_s : s \in [0, t] \} \) is a predictable integrable \( L^2_\alpha \)-valued process and \( b = \{ b_s : s \in [0, t] \} \) is a predictable integrable \( L^2(U) \)-valued process. Assume that
\[
\| a_s \|_{L^2_\alpha} \leq K(t), \quad \| b_s \|_{L^2(U)} \leq K(t) \quad \text{for all } s \in [0, t] \text{ and some } K(t) > 0,
\]
(2.2)

Let \( f_t : \mathbb{R}^+ \to \mathbb{R} \) be given by (1.3) and let \( \bar{f}_t : \mathbb{R}^+ \to \mathbb{R} \) be given by
\[
\bar{f}_t = S(t)f_0 + \int_0^t S(t-s)a_s ds + \int_0^t S(t-s)b_s dZ_s,
\]
where \( (B_t u)(x) = \langle b_t(x), u \rangle_u, \ x, t \geq 0, \ u \in U \) and \( (S(t)f)(x) = f(x+t), \ x, t \geq 0 \). Then for almost every \( (\omega, x) \in \Omega \times \mathbb{R}^+ \), we have
\[
\bar{f}_t(x) = f_t(x).
\]

## 3 Short rate

In financial applications the concept of the so-called short rate, given by \( r_t = f_t(0) \), plays an important role. For instance, in proofs regarding the absence of arbitrage in the market, it is often shown that the process of discounted bond prices \( \{ \hat{P}(t, T) : t \in [0, T] \} \) given by
\[
\hat{P}(t, T) = \exp \left( \int_0^t f_s(0) ds \right) \exp \left( - \int_0^{T-t} f_t(x) dx \right),
\]
is a local martingale. Although \( f_s(0) \) may not exists for \( f_s \in L^2_\alpha \), our next result ensures that the process \( \left\{ \int_0^t f_s(0) ds : t \in [0, T] \right\} \) is well-defined.
Proposition 3.1. Let $t \geq 0$. Suppose $f_0 \in L^2_\alpha$, $a = \{a_s : s \in [0, t]\}$ is a predictable integrable $L^2_\alpha$-valued process and $b = \{b_s : s \in [0, t]\}$ is a predictable integrable $L^2(U)$-valued process. Assume that the condition (2.2) holds. Let $f_t : \mathbb{R}_+ \to \mathbb{R}$ be given by (1.3). Then for every $x \geq 0$ the integral

$$\int_0^t f_s(x)ds$$

is well-defined. In particular, the short rate $r_s = f_s(0)$ is well-defined for almost every $s \in [0, t]$.

Proof of Proposition 3.1. It is clear that $g, \varphi_s : \mathbb{R}_+ \to \mathbb{R}$, $\phi_s : \mathbb{R}_+ \to U$, given by

$$g(x) = \int_x^{x+t-s} f_0(\xi)d\xi, \quad \varphi_s(x) = \int_x^{x+t-s} a_s(\xi)d\xi, \quad \phi_s(x) = \int_x^{x+t-s} b_s(\xi)d\xi,$$

are continuous functions. In fact, they are Lipschitz continuous. Indeed, from the Hölder inequality and the Lagrange mean value theorem, if $f \in L^2_\alpha$, $y \geq x \geq 0$, then

$$\int_x^y f(\xi)d\xi \leq \frac{\alpha^{-\alpha y} - \alpha^{-\alpha x}}{\alpha} \|f\|_{L^2_\alpha} \leq \frac{\|f\|_{L^2_\alpha}}{\alpha} |x - y|.$$

Hence, by (2.2), for all $s \in [0, t]$

$$|g(x) - g(y)| \leq \frac{\|f_0\|_{L^2_\alpha}}{\alpha} |x - y|,$n
$$|\varphi_s(x) - \varphi_s(y)| \leq \frac{2K(t)}{\alpha} |x - y|,$n
$$|\phi_s(x) - \phi_s(y)| \leq \frac{2K(t)}{\alpha} |x - y|.$$

Note that the Lipschitz constant of $\varphi_s$ and $\phi_s$ does not depend on $s$, hence for every $x \geq 0$, $\int_0^t \langle \phi_s(x), dZ_s \rangle_U$ can be defined as a $L^2(\Omega)$ limit of $\int_0^t \langle \phi_s(x_n), dZ_s \rangle_U$ for some sequence $x_n \to x$.

The Hölder inequality will also imply that

$$\|f\|_{L^1} \leq \alpha^{-\frac{1}{2}} \|f\|_{L^2_\alpha}.$$  \hspace{1cm} (3.1)
Hence, by (2.2), we get
\[ \int_0^t \int_{x}^{x+t-s} |a_s(\xi)| d\xi ds < \int_0^t \|a_s\|_{L_2^\alpha} \alpha^{-\frac{1}{2}} ds < tK(t)\alpha^{-\frac{1}{2}}, \]
and
\[ \mathbb{E} \int_0^t \int_x^{x+t-s} \|b_s(\xi)\|^2_{L_2(\mathbb{U})} d\xi ds < \mathbb{E} \int_0^t \|b_s\|^2_{L_2^2(\mathbb{U})} ds < t(K(t))^2. \]
Applying Fubini's theorem for \( \sigma \)-finite measures to the function \( g(s,v) = a_s(x+v-s) \) and the stochastic Fubini's theorem (see [13]) to the process \( \Phi(s,v) = \langle b_s(x+v-s), u \rangle_{\mathbb{U}} \), we obtain
\[ \int_0^t f_v(x) dv = \int_0^t f_0(x+v) dv + \int_0^t \int_0^v a_s(x+v-s) ds dv + \int_0^t \int_0^v \langle b_s(x+v-s), dZ_s \rangle_{\mathbb{U}} dv = g(x) + \int_0^t \varphi_s(x) ds + \int_0^t \langle \phi_s(x), dZ_s \rangle_{\mathbb{U}}. \]

\section{State dependent coefficients}

Let \( Z \) be an \( \mathbb{R} \)-valued square integrable Lévy process and let \( b_t = G(f_t) \), for some \( G : L_2^\alpha \rightarrow L_2^\alpha \). Then (1.4) reads as
\[ f_t = (Af_t + F(f_t)) dt + G(f_t) dZ_t, \quad (4.1) \]
with \( F : L_2^\alpha \rightarrow L_2^\alpha \) given by
\[ F(f)(x) = \mathcal{S}(G(f))(x), \]
where \( \mathcal{S} : L_2^\alpha \rightarrow L_2^\alpha \) is the so-called HJM mapping,
\[ \mathcal{S}(h)(x) = J'(\int_0^x h(\xi) d\xi) h(x). \]
The dependence between \( G \) and \( F \) is a consequence of (1.2).

We wish to discuss the positivity of solutions to the equation (4.1) with the coefficient \( G(f)(x) = \min\{f(x), \lambda(x)\} \) (this is in fact the example from the Introduction, where coefficient \( \{b_t : t \geq 0\} \) is given by (1.6)). We collect a few existing results regarding Musiela’s equation (4.1). First we restate Lemma 4.4 and Theorem 3.7 of [10].
Lemma 4.1. Assume that $f, g \in L^2_\alpha$ are positive functions, $\|f\|_{L^2_\alpha}, \|g\|_{L^2_\alpha} \leq M$ and $\|f\|_{L^1}, \|g\|_{L^1} \leq R$. Then

$$\|S(f) - S(g)\|_{L^2_\alpha} \leq (J'(R) + \max\{E|Z_1|^2, J''(R)\} \alpha^{-\frac{1}{2}} M) \|f - g\|_{L^2_\alpha}.$$ 

Lemma 4.2. Assume $G : L^2_\alpha \to L^2_\alpha$ in (4.1) is given by

$$G(f)(x) = \min\{|f(x)|, \lambda(x)\},$$

for some positive $\lambda \in L^2_\alpha$. If $J''(\|\lambda\|_{L^1}) < +\infty$, then for every $\eta \in L^2_\alpha$, there exists a unique solution $(f_t)_{t \geq 0}$ to (4.1) with $f_0 = \eta$.

Next, let $\nu$ be the Lévy measure of $Z$, i.e.

$$\nu(\Gamma) = \mathbb{E}\left( \sum_{0 < t \leq 1} 1_{\Gamma}(Z(t) - Z(t^-)) \right),$$

where $Z(t^-) = \lim_{t \to t^-} Z(s)$, $\Gamma$ is a Borel subset of $U$ such that $\Gamma \subset U \setminus \{0\}$, and $\nu(\{0\}) = 0$. It is well-known that

$$\int_U \min\{1, y^2\} \nu(dy) < +\infty.$$

The support of an $\mathbb{R}$-valued Lévy process $Z$ with the Lévy measure $\nu$ is defined as

$$S_Z = \{z \in \mathbb{R} : \forall \epsilon > 0 \quad \nu([z - \epsilon, z + \epsilon]) > 0\}.$$ 

The function $J''$ can be written in terms of $\nu$ as

$$J''(z) = \int_{-\infty}^{+\infty} y^2 e^{-zy} \nu(dy).$$

Note that if $S_Z \in [-1, +\infty)$, then

$$J''(z) = \int_{-1}^{+\infty} y^2 e^{-zy} \nu(dy) \leq e^{|z|} \int_{-1}^{1} y^2 \nu(dy) + \int_{1}^{+\infty} y^2 \nu(dy).$$

We conclude that

$$S_Z \in [-1, +\infty) \quad \Rightarrow \quad |J''(z)| < +\infty, \quad \forall z > 0. \quad (4.2)$$

The following result can be found in [1] (see Section 4 therein).
Lemma 4.3. Assume $G : L^2_\alpha \rightarrow L^2_\alpha$ in (4.1) is given by

$$G(f)(x) = g(x, f(x)),$$

for some $g : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$. The following conditions ensures the positivity of solutions to (4.1),

$$g(x, 0) = 0, \quad y + g(x, y)u \geq 0, \quad \forall x, y \geq 0, \quad \forall u \in S_Z.$$  \hfill (4.3)

Now, let $Z$ be square integrable with $S_Z \in [-1, +\infty)$ and let $G(f)(x) = \min\{f(x), \lambda(x)\}$ for a positive $\lambda \in L^2_\alpha$. First note that for $g : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}$ given by $g(x, y) = \min\{y, \lambda(x)\}$ condition (4.3) is fulfilled, hence by Lemma 4.2 and (4.2) for any positive $\eta \in L^2_\alpha$ there exists a unique positive solution to (4.1). Further if $f, g \in L^2_\alpha$ are positive, so is $G(f)$ and $G(g)$. Further for all $f, g \in L^2_w$

$$\|G(f) - G(g)\|_{L^2_\alpha} \leq \|f - g\|_{L^2_\alpha}, \quad \|G(f)\|_{L^2_\alpha} \leq \|\lambda\|_{L^2_\alpha},$$

and from (3.1),

$$\|G(f)\|_{L^1} \leq \alpha^{-\frac{1}{2}} \|\lambda\|_{L^2_\alpha}.$$

Hence by Lemma 4.1

$$\|F(f) - F(g)\|_{L^2_\alpha} \leq 2C\alpha^{-\frac{1}{2}} \|\lambda\|_{L^2_\alpha} \|f - g\|_{L^2_\alpha},$$

since $J'(r) \leq Cr$, $r \geq 0$, from the Lagrange mean value theorem. The Lagrange mean value theorem implies also that $|F(f)(x)| \leq C \|G(f)\|_{L^1} |G(f)(x)|$, hence

$$\|F(f)\|_{L^2_\alpha} \leq C\alpha^{-\frac{1}{2}} \|\lambda\|^2_{L^2_\alpha}.$$

Therefore the assumptions of Theorem 2.4 are satisfied for $a_t = F(f_t)$ and $b_t = G(f_t)$. The predictability of $\{a_t : t \geq 0\}$ and $\{b_t : t \geq 0\}$ follows from the Lipschitz continuity of $F$ and $G$, the strong continuity of the semigroup $S$ and the predictability of $\{f_t : t \geq 0\}$.

Acknowledgments

The author is greatly indebted to Ben Goldys, Francesco Russo and Jerzy Zabczyk for helpful suggestions.
References


