University of Padova

DEPARTMENT OF MATHEMATICS

**Master's degree in Mathematical Sciences**

# The Geometry of Interest Rates in a Post-Crisis Framework

Supervisor:
**Prof. Claudio Fontana**

Candidate:
**Giacomo Lanaro, 1182152**

September 2019

Academic Year 2018-2019

# Contents

# Introduction

In this thesis we aim at some geometric properties of multi-curve interest rate models. We investigate problems which are not still solved in a post-crisis contest. In particular, we analyse an infinite-dimensional system of forward rate processes, each of them described by a stochastic differential equation, driven by a $d$-dimensional Brownian motion. We aim at conditions under which these processes are consistent with a given parameterized surface, defined on the infinite-dimensional domain of the solution. Therefore, we provide conditions on the forward rate processes which guarantee the existence of finite-dimensional realizations. We investigate these problems because after the last financial crisis the structure of forward rate processes has become more complex. In particular, from a single-curve model, we now have to manage a vector of forward rate processes, in which each component is related to the others.

This work is structured as follows. In the first chapter, we describe the profound changes which the interest rate market has suffered since the financial crisis of $2007-2008$ and we derive the system of stochastic differential equations (SDEs) which describe it. In particular, in the post crisis framework, the counterparty and liquidity risk are no longer negligible. As a consequence of this fact, it is no more possible to describe the complete interest-rate market by a unique fixed-income instrument, the zero-coupon bond (ZCB), whose price is denoted by $(B_t(T))_{t\in[0,T]}$, where $T$ is the maturity of the contract. Moreover, the equivalence between the simple spot rate $-\frac{B_T(T+\delta)-1}{\delta B_T(T+\delta)}$, computed for the time interval $[T, T+\delta]$ and the LIBOR rate $L(T;T,T+\delta)$, which is an interbank interest rate for lending and borrowing for a set of banks called LIBOR panel, does not hold any longer. Indeed, from market data we can notice that spreads between the LIBOR rates associated with different time interval's length $\delta$ emerged. In particular, while the pre-crisis equivalence is respected for $\delta = 1$ day, more $\delta$ is high, more the LIBOR rate associated with $\delta$ is higher than the simple spot rate.

To describe the interest rate market, several authors adopted a Heath-Jarrow-Morton (HJM) approach which consists in modeling not directly the price of a ZCB, but the instantaneous forward rate $f_t(T) = -\frac{\partial}{\partial T}\log B_t(T)$. We adopt the same approach to model the interest-rate market in the post crisis framework. By the presence of these spreads, it is necessary to describe separately each forward instantaneous LIBOR rates associated with a finite set of positive time intervals $\delta_0 < \cdots < \delta_m$. Hence, we introduce the LIBOR rates $L^\delta(T;T,T+\delta)$ for $\delta \in \{\delta_0,\dots,\delta_m\}$. Moreover, it is convenient to define positive multiplicative spread processes $S^\delta$ which connect the LIBOR rate associated with $\delta_0$ (risk-free) and the LIBOR rate associated with $\delta \in \{\delta_1,\dots,\delta_m\}$. For each $\delta \in \{\delta_1,\dots,\delta_m\}$, by $S^\delta$ a fictitious $\delta$-bond associated with $L^\delta(T;T,T+\delta)$ can be introduced. In conclusion, by non arbitrage conditions, we derive the following Heath-Jarrow-Morton system of Stochastic differential equations:

$$\begin{cases} dr_t(x) = \Big(\mathbf{F}r_t(x) + \widetilde{\sigma}_t(t+x)\mathbf{H}\widetilde{\sigma}\Big)dt + \widetilde{\sigma}_t(t+x)dW_t; \\ dr_t^\delta(x) = \Big[-\beta_t^\delta\sigma_t^{\delta*}(t+x) + \sigma_t^\delta(t+x)\mathbf{H}\sigma^\delta + \mathbf{F}r_t^\delta(x)\Big]dt + \sigma_t^\delta(t+x)dW_t; \\ dY_t^\delta = \Big(-r_t^\delta(0) - \frac{1}{2}||\beta_t^\delta||^2 + r_t(0)\Big)dt + \beta_t^\delta dW_t, \end{cases} \tag{1}$$

where $r_t(x) = f_t^{\delta_0}(t+x)$ and $r_t^\delta(x) := f_t^\delta(t+x)$ are the instantaneous forward rates associated with each $\delta$, whereas the finite-dimensional process $Y^\delta$ is the logarithm of the spread process $S^\delta$, defined for each $\delta \in \{\delta_1,\dots,\delta_m\}$. The $x$ variable stands for the time to maturity $T = t+x$. Finally, at the end of the first chapter we describe an analogy between the market model determined by (1) and a model for the multi-currency interest rate market.

In the second chapter, we describe the problem of consistency. First of all, adopting the geometric approach developed by Biörk in [5], we introduce a Banach space $\mathcal{H} \subset \mathcal{C}^{+\infty}(\mathbb{R}_+,\mathbb{R})$ in which the solution of each instantaneous forward rate $r^\delta$ lives. Therefore, the domain of the solution of system (1) is a Banach Space $\hat{\mathcal{H}} := \mathcal{H}^{m+1} \times \mathbb{R}^m$ satisfying suitable conditions. In this framework, we generalize the results proposed by Björk *et al.* in [5] and [2]. These results are related to the problem of consistency between a model $\mathcal{M}$ and a parameterized family $\mathcal{G} \subset \hat{\mathcal{H}}$, where we say that a model $\mathcal{M}$ is the solutions of the system (1), where the volatility terms $(\widetilde{\sigma},\sigma^{\delta_1},\dots,\sigma^{\delta_m},\beta^{\delta_1},\dots,\beta^{\delta_m})$ are specified. The consistency problem can be intuitively described as follows:

Take as given a model $\mathcal{M}$ and a parameterized family $\mathcal{G} \subset \hat{\mathcal{H}}$ of forward rate curves, we say that the couple $(\mathcal{M},\mathcal{G})$ is consistent if given an initial forward rate curve $r^M(x) \in \mathcal{G}$, the interest rate model $\mathcal{M}$ starting on $r^M(x)$ produces forward rate curves belonging to the family $\mathcal{G}$.

We provide a characterization of the consistency determined by the geometric concepts of vector fields and tangent space. Therefore, we analyse several examples of models $\mathcal{M}$ and parameterized families $\mathcal{G}$, in particular we provide results for the model Hull-White and Ho-Lee related to the family of Nelson-Siegel and Svensson and their generalizations. Differently from the pre-crisis framework, now we have to manage the presence of the spreads and how the spreads entangle the structure of the model $\mathcal{M}$. In particular, we construct a strategy for these examples which provides the conditions which have to be satisfied by the components of $\mathcal{G}$ related to the spreads.

In Chapter 3, we focus on the problem of the existence of finite-dimensional realization in particular cases. The problem can be introduced as follows:

Given a model $\mathcal{M}$, finite-dimensional realizations exist if the forward rate process

$$\hat{r}_t(x) = (\widetilde{r}_t(x), r_t^{\delta_1}(x), \dots, r_t^{\delta_m}(x), \beta^{\delta_1}, \dots, \beta^{\delta_m}),$$

describing the model $\mathcal{M}$, admits a suitable mapping $G : \mathbb{R}^n \longrightarrow \hat{\mathcal{H}}$ and a finite-dimensional process $Z$, such that:

$$\begin{aligned} dZ_t &= a(Z_t)dt + b(Z_t)dW_t, \\ r_t(x) &= G(Z_t)(x), \end{aligned}$$

where $W$ is the same Brownian motion of (1).

To solve it, we exploit an analogy between the post-crisis interest rate market and a multi-currency interest rate market. In particular, we generalize the results proposed in [21] for the finite-dimensional realization in multi-currency market context adapting it to our purposes. We provide an equivalent condition on the volatility term of the solution of the system (1) when the volatility term is not dependent on the entire solution $\hat{r}$ but only on the time-to-maturity $x$ and a sufficient condition when the volatility term has the following form:

$$\hat{\sigma}(\hat{r}, x) = (\varphi^0(\hat{r})\lambda^0(x), \dots, \varphi^m(\hat{r})\lambda^m(x), \beta^1(\hat{r}), \dots, \beta^m(\hat{r})),$$

where the mappings $\varphi^i$ are real-valued, $\varphi^i : \hat{\mathcal{H}} \longrightarrow \mathbb{R}$. In order to provide these results, we adopt a geometric approach deriving the conditions which guarantee the existence of finite-dimensional realizations by strong results of infinite-dimensional differential geometry related to the concepts of tangential manifold and Lie algebra generated by a given set of vector fields.

Finally, in Appendix A we briefly describe the pre-crisis context and the Heath-Jarrow-Morton approach, whereas in Appendix B, we introduce the main concepts of infinite-dimensional differential geometry and we prove the results we need for our purposes.

# Chapter 1

# Fixed-Income Markets in the Post-Crisis Framework

In this chapter we aim at presenting the main differences between the fixed-income market in a pre-crisis environment, described in Appendix A and the framework which has developed after the financial crisis of $2007-2008$. First of all, we will give a brief description of the problems generated by liquidity and credit risk and their consequences, related in particular with the inequality of the classical pre crisis relation between the interest rate and the price of a particular contract, the *Zero Coupon Bond*. These facts have led to the necessity to provide new conditions on the fixed-income market, which was described, after the crisis, by a system of forward rate equations different from the one used in the pre-crisis environment (see Appendix A). Finally, at the end of the chapter, we will show a connection between the forward rate system developed in this new context and a multi-currency interest rate market, described by Slinko in [21]. We will exploit this connection in the next chapters in order to analyse some properties of the fixed-income market, in the post-crisis framework.

## 1.1 Post-crisis framework

After the financial crisis of 2007-2008 the fixed-income market has undergone deep changes. This is due to the fact that, before the crisis, in the interbank market it was possible to neglect the counterpart and the liquidity risk. These concepts respectively represent the risk related to the impossibility for the counterpart to fulfil its obligations in a financial contract and the risk of excessive costs of funding a position in a financial contract due to the lack of liquidity in the market.
After the crisis it was necessary to take into account these problems, and this

necessity has led to many consequences also in the general fixed-income market. Indeed, many contracts pledged in the fixed-income market are determined by derivatives on interbank interest rates, for example Euribor or Libor.
The main consequence of this fact can be observed by comparing quoted prices of same contracts for different maturity dates. Market data have shown how the relations between prices quoted in the market with different maturity dates have no longer respected standard no arbitrage relation, which held in a pre-crisis environment (see (A.1)). In particular, we can observe spreads between LIBOR rates and the swap rate, based on the overnight indexed swaps (OIS), which have taken a crucial role in the framework that we are developing.
In conclusion, if we aim at describing the fixed-income market, we can not parameterize the interest-rate curve, as in the pre-crisis environment, with the instantaneous forward rate (described in (A.3)) of a Zero Coupon Bond, but it is necessary to distinguish all the interest-rate curves associated with the spreads introduce above, adopting an approach called *multi-curve.*

### 1.1.1 Interbank Rates

LIBOR is the acronym for *London InterBank Offered Rate*, we take the description of LIBOR rate by ICE Benchmark Administration IBA (from the website: https://www.theice.com/iba/libor), which is administering the LIBOR as of February 2014:
"ICE Libor is designed to reflect the short term funding costs of major banks active in London, [...]. The ICE Libor is a *polled* rate. This means that panel of representative banks submits rates which are then combined to give the ICE Libor rate. Panel banks are required to submit a rate in answer to the ICE Libor question: *At what rate could you borrow funds, were you to do so by asking for and then accepting inter-bank offers in a reasonable market size just prior to 11 a.m.?*. [...]. Reasonable market size is intentionally unquantified. The definition of an appropriate market size depends on the currency and tenor in question, as well as supply and demand.[...]".

Before the crisis, the spot LIBOR rate was assumed to be equal to the floating rate defined through *Zero Coupon Bond (ZCB)* prices. This expression of LIBOR rate (A.2) represented the rate at time $T$ for the interval $[T, T+\delta]$.

In this context, the LIBOR panel, which determined this rate, was composed by a set of banks, whose credit quality was guaranteed. Indeed, if one of these banks had had a deteriorated credit quality, it would have been replaced by a bank with a better credit quality. This condition had made possible to assume risk-freedom in the panel and this property is implicitly given supposing (A.1).

After the crisis, this mechanism is still valid, but the credit and liquidity risks described above are no longer negligible because they can affect also solid banks, which are in the panel, in short time. As a consequence of this fact it is no longer possible to suppose that LIBOR rates are not affected by interbank risks, thus the definition (A.1) does no longer hold:

$$L(T;T,T+\delta) \neq -\frac{B_T(T+\delta)-1}{\delta B_T(T+\delta)}. \tag{1.1}$$

### 1.1.2 Forward Rates Agreements

The problem described in the previous section has led to the consequence that a pre-crisis connection between *ZCB* (see (A.1.1)), LIBOR interest rate and a fixed-income contract, called *Forward-Rate-Agreements, (FRA)* does not hold anymore.

**Definition 1.1.1.** *A forward rate agreement, is an OTC (over the counter) derivative, which allows to the holder to lock at any date $0 \leq t \leq T$, the interest rate between the inception date $T$ and the maturity $T+\delta$, $\delta > 0$ at a fixed value $K$. At the maturity, a payment based on $K$ is made and the one based on the relevant floating rate (usually the spot LIBOR rate $L(T, T+\delta)$) is received. The notional amount is denoted by $N$.*
*The payoff of the FRA with notional amount $N$ and inception date $T$, at maturity $T+\delta$ is given by:*

$$\Pi^{FRA}(T+\delta;T,\delta,K,N) = N\delta(L(T,T+\delta)-K). \tag{1.2}$$

In the following we will consider, without loss of generality $N=1$. Therefore, we can use the following notation:

$$\Pi^{FRA}(T+\delta;T,\delta,K,1) \equiv \Pi^{FRA}(T+\delta;T,\delta,K).$$

We introduce now a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T^*]}, \mathbb{P})$, where $T^*$ is the time horizon. All the stochastic processes introduced below are supposed to be adapted processes, defined on this probability space, whereas $\mathbb{P}$ is supposed to be an objective probability measure.

Using general pricing approach, if we want to compute the price of a FRA at time $t \leq T$, we have to compute the conditional expectation with respect to the $(T+\delta)$-forward martingale measure $\mathbb{Q}^{T+\delta}$ , which is obtained using as numeraire the OIS price process $B_t^{OIS}(T+\delta)$ that in the following will be simply denoted by $B_t(T+\delta)$. The justification of this choice will be described in the next section. Under this condition, we obtain that:

$$\Pi^{FRA}(t;T,T+\delta;K) = \delta B_t(T+\delta)\mathbb{E}^{\mathbb{Q}^{T+\delta}}[L(T;T,T+\delta)-K|\mathcal{F}_t], \quad t \leq T. \tag{1.3}$$

Recalling that (1.1) holds, it is no longer possible to compute spot LIBOR rate, using only ZCB price processes. As a consequence of this fact we cannot describe the forward LIBOR rate as in the pre-crisis context ((A.1)). This implies that the price $\Pi^{FRA}(t;T,\delta;K)$ cannot be determined using a replicating portfolio of ZCBs, and thus we have to consider a new fixed-income market, different from the one described before the crisis, formed by all the ZCBs, but also all the FRAs.

To solve the non sustainability of classical definition of forward LIBOR rate, in general we need to give an alternative definition, which is in accord with spot LIBOR rate $L(T, T+\delta)$.

**Definition 1.1.2.** *The* forward LIBOR rate *with for the period* $[T, T+\delta]$ *at time* $t \leq T$*, is the value of* $K$*, such that* $\Pi^{FRA}(t;T,T+\delta;K) = 0$*. It is given by:*

$$L(t;T,T+\delta) := \mathbb{E}^{\mathbb{Q}^{T+\delta}}[L(T;T,T+\delta)|\mathcal{F}_t], \quad 0 \leq t \leq T. \tag{1.4}$$

In particular, we can observe that:

$$L(t;T,T+\delta) = \mathbb{E}^{\mathbb{Q}^{T+\delta}}[L(T;T,T+\delta)|\mathcal{F}_t] \neq \frac{1}{\delta}\Big(\frac{B_t(T)}{B_t(T+\delta)} - 1\Big). \tag{1.5}$$

In the previous definition, the forward LIBOR rate is dependent on the time interval $\delta$, also called *tenor*, which will play a crucial role in the approach that we are developing. Indeed, by the above inequality, it is no more possible to determine the connection between LIBOR rates associated with different tenors, simply through direct non arbitrage relations. In particular, each tenor determines the behaviour of the contracts associated with it, which evolve in a proper independent way. This fact leads to the necessity to define a set of forward interest rates, each of them associated with a given tenor $\delta$:

$$L^{\delta}(t;T,T+\delta) = \mathbb{E}^{\mathbb{Q}^{T+\delta}}[L^{\delta}(T;T,T+\delta)|\mathcal{F}_t]. \tag{1.6}$$

This implies that it is necessary to model separately each component of the market, associated with each tenor. To do that, we will follow a multi-curve approach, based on modeling spread processes, which will characterize the dynamics of the contracts associated with every tenor. As we will see below, a spread process associated with tenor $\delta$ will take into account both forward LIBOR rate (1.4) and the classical pre-crisis definition (A.1). In particular, we will describe multiplicative spreads given by the ratio between normalized forward rates, defined by a forward rate agreement (as in (1.6)) and associated with a finite family of tenors and normalized compounded forward rates associated with $\delta = 1$ day. This choice will be formally justified in the next sections and it is based on the fact that if a contract is associated with a tenor equal to one day, we can consider it risk free, thanks to its very short maturity.

# 1.2 The multi-curve approach

## 1.2.1 Tenor Structures

In the end of the previous section, we have seen how the classical structure of fixed-income market is not adapt to describe the current market environment. To do this, we need a new approach which takes into account different tenors.

First of all, we recall the notation for a time horizon $T^*$. Adopting the notation of [GR15], we define:

**Definition 1.2.1.** *A discrete tenor structure $\mathcal{T}^\delta$ with tenor $\delta$ is a finite sequence of dates:*

$$\mathcal{T}^\delta := \{0 \le T_0^\delta < T_1^\delta < \cdots < T_{M_\delta}^\delta \le T^*\}, \tag{1.7}$$

*where we consider $\delta := T_k^\delta - T_{k-1}^\delta$. It represents the year fraction corresponding to the length of the interval $(T_{k-1}^\delta, T_k^\delta]$, for $k = 1, \dots, M_\delta$.*

LIBOR rates produced by ICE are given each business day for seven maturities (1 day, 1 week, 1, 2, 3, 6 and 12 months). In accord with this choice, we consider tenor $\delta$ range from one day ($\delta = \frac{1}{360}$) to twelve months ($\delta = 1$). This approach has to manage many different tenor structures, hence we define a collection of tenors $\mathcal{D} := \{\delta_1 < \delta_2 \cdots < \delta_m\}$ and for each of them we consider the tenor structure $\mathcal{T}^{\delta_i} = \{0 \le T_0^{\delta_i} < T_1^{\delta_i} < \cdots < T_{M_{\delta_i}}^{\delta_i} \le T^*\}$. Moreover, we assume that $\mathcal{T}^{\delta_n} \subset \mathcal{T}^{\delta_{n-1}} \subset \cdots \subset \mathcal{T}^{\delta_1} \subseteq \mathcal{T}$, where $\mathcal{T} := \{0 \le T_0 < T_1 < \cdots < T_M \le T^*\}$ is the *reference tenor structure*. Finally, we suppose that $T_{M_{\delta_i}}^{\delta_i} = T_M$ for all $i$; in this way all the tenor structures have the same final date.

## 1.2.2 Overnight Indexed Swaps

In paragraph 1.1.2, we have seen that interest rates associated with different tenors does no more evolve equivalently, then one of the main problems is the choice of the discount curve.

In order to solve this problem, there are two possibilities. The first choice consists in considering a different discount curve for each tenor structure, and as a consequence, considering each market determined by tenor $\delta$, as a separate market. This is not an efficient choice because the complete fixed-income market has to be arbitrage free and, adopting that approach, it is very difficult to determine conditions (on the separated markets) which guarantee the absence of arbitrage on the entire market. The other choice is to choose a common discount curve, which is used to compute the discounted price of all instruments, whatever their tenor is.

Nowadays, the last possibility is obliged and we adopt it to develop our dissertation. The common discounting curve that it was chosen is the one associated to the *overnight indexed swap (OIS)* contract.

First of all, it is convenient to give the definition of an interest rate swap. Therefore, we briefly describe what an OIS contract is.

**Definition 1.2.2.** *An* interest rate swap *is a financial contract, in which a stream of future interest rate payments linked to a pre-specified fixed rate denoted by $K$, is exchanged for another one linked to a floating interest rate (generally it is used the Libor rate), based on a specified notional amount $N$ (which in our dissertation is supposed to be equal to* 1*).*
*The swap's inception date is $T_0 \geq 0$, and $T_1 < \cdots < T_n$ ($T_1 > T_0$) denote the payment dates with $\delta = T_k - T_{k-1}$, $\forall k \in \{1, \ldots, n\}$. The value of this contract at time $t \leq T_0$ (supposing N=1), is determined by a combination of FRA contracts. It holds indeed that:*

$$\begin{aligned}\Pi^{SWAP}(t; T_0, \ldots, T_n, K) &= \sum_{k_1}^{n} \delta_k B_t(T_k) \mathbb{E}^{\mathbb{Q}^k}\Big[L(T_{k-1}; T_{k-1}, T_k) - K|\mathcal{F}_t\Big] = \\ &= \sum_{k=1}^{n} \Pi^{FRA}(t; T_{k-1}, T_k, K),\end{aligned}$$

*where $\mathbb{Q}^{T_k}$ is the $T_k$-forward martingale measure.*

The OIS rate is a particular Swap contract, described as follows:
In a OIS contract the counterparties exchange a stream of fixed rate ($K$) payments for a stream of floating rate payments linked to a compounded overnight rate. In order to compute the value of this contract at time $t \leq T_0$, we follow the idea described in [11] (chapter 1, section 4.4).

First of all, we compute the fixed leg payments:

$$\Pi^{OIS}(t; T_0, \ldots, T_n, K)_{fix} = K \sum_{k=1}^{n} \delta_k B_t(T_k). \tag{1.8}$$

To obtain the floating leg payments we need to describe how the floating rate is computed. For the time $(T_{k-1}, T_k)$, it is get compounding the overnight rates between these dates:

$$F^{ON}(T_{k-1}, T_k) = \frac{1}{\delta_k}\Big(\prod_{j=1}^{n_k}[1 + \delta_{t^k_{j-1}, t^k_j} F^{ON}(t^k_{j-1}, t^k_j)] - 1\Big). \tag{1.9}$$

We have divided the considered time interval in this way:
$T_{k-1} = t^k_0 < t^k_1 < \cdots < t^k_{n_k} = T_k$, where $\delta_{t^k_{j-1}, t^k_j} = t^k_j - t^k_{j-1} = 1$ day, thus

$F^{ON}(t^k_{j-1}, t^k_j)$ denotes the overnight rate for the period $(t^k_{j-1}, t^k_j)$.
This overnight rate is supposed to be related to the bond price process, through the classical pre-crisis formula:

$$F^{ON}(t^k_{j-1}, t^k_j) = -\frac{B_{t^k_{j-1}}(t^k_j) - 1}{\delta_{t^k_{j-1}, t^k_j} B_{t^k_{j-1}}(t^k_j)}.$$

This is due to the fact that, since the time interval associated with this interest rate is 1 day, the liquidity and credit risks are almost negligible. Hence, the formula for the floating leg payments is:

$$\begin{aligned}
\Pi^{OIS}(t; T_0, \ldots, T_n, K)_{floating} &= \sum_{k=1}^{n} \delta_k B_t(T_k) F^{ON}(t; T_{k-1}, T_k) = \\
&\overset{\bigstar}{=} \sum_{k=1}^{n} \delta_k B_t(T_k) \Big[\frac{1}{\delta_k}\Big(\frac{B_t(T_{k-1})}{B_t(T_k)} - 1\Big)\Big] = \\
&= B_t(T_0) - B_t(T_n),
\end{aligned} \tag{1.10}$$

where the equality $\bigstar$ is due fact that, since the overnight rate is supposed to be risk free and we are assuming $T_k = t^k_{n_k}$, the following equivalence holds:

$$\begin{aligned}
F^{ON}(t; T_{k-1}, T_k) =& \mathbb{E}^{\mathbb{Q}^{T_k}}\Big[\frac{1}{\delta_k}\Big(\prod_{j=1}^{n_k}[1 + \delta_{t^k_{j-1}, t^k_j} F^{ON}(t^k_{j-1}, t^k_j)] - 1\Big)|\mathcal{F}_t\Big] = \\
=& \frac{1}{\delta_k}\mathbb{E}^{\mathbb{Q}^{T_k}}\Big[\prod_{j=1}^{n_k}\frac{B_{t^k_{j-1}}(t^k_{j-1})}{B_{t^k_{j-1}}(t^k_j)} - 1|\mathcal{F}_t\Big] = \\
\overset{B.T}{=}& \frac{1}{\delta_k}\left\{\frac{\mathbb{E}^{\mathbb{Q}^{t^k_{n_k-1}}}\Big[\prod_{j=1}^{n_k-1}\frac{B_{t^k_{j-1}}(t^k_{j-1})}{B_{t^k_{j-1}}(t^k_j)}|\mathcal{F}_t\Big]}{\mathbb{E}^{\mathbb{Q}^{t^k_{n_k-1}}}\Big[\frac{B_{t^k_{n_k-1}}(t^k_{n_k})}{B_{t^k_{n_k-1}}(t^k_{n_k-1})}|\mathcal{F}_t\Big]} - 1\right\} = \\
=& \frac{1}{\delta_k}\left\{\frac{B_t(t^k_{n_k-1})}{B_t(t_{n_k})}\mathbb{E}^{\mathbb{Q}^{t^k_{n_k-1}}}\Big[\prod_{j=1}^{n_k-1}\frac{B_{t^k_{j-1}}(t^k_{j-1})}{B_{t^k_{j-1}}(t^k_j)}|\mathcal{F}_t\Big] - 1\right\} = \\
=& \text{ Repeating the same procedure } = \\
=& \frac{1}{\delta_k}\Big[\prod_{j=1}^{n_k}\frac{B_t(t^k_{j-1})}{B_t(t^k_j)} - 1\Big] = \frac{1}{\delta_k}\Big[\frac{B_t(T_{k-1})}{B_t(T_k)} - 1\Big],
\end{aligned}$$

where $B.T$ stands for Abstract Bayes Theorem (for the proof see [3], Appendix B, Proposition B.41). Moreover, we have used the fact that the Lebesgue-Radon-

Nikodym derivative between the two forward measures is:

$$L_t^{t_{n_k}^k} = \frac{d\mathbb{Q}^{t_{n_k}^k}}{d\mathbb{Q}^{t_{n_{k-1}}^k}}\Big|_t = \frac{B_t(t_{n_k}^k)B_0(t_{n_{k-1}}^k)}{B_t(t_{n_k}^k)B_0(t_{n_k}^k)}. \tag{1.11}$$

In conclusion the value at time $t$ of an OIS payer (in which the floating rate is received and the fixed rate is payed) is:

$$\Pi^{OIS}(t; T_0, \ldots, T_n, K) = B_t(T_0) - B_t(T_n) - K \sum_{k=1}^{n} \delta_k B_t(T_k). \tag{1.12}$$

By analogy to the FRA rate definition, the OIS rate $K^{OIS}(t; T_0, T_n)$, for $t \leq T_0$ is defined imposing that the OIS's value is equal to zero at time $t$:

$$K^{OIS}(t; T_0, T_n) = \frac{B_t(T_n) - B_t(T_0)}{\sum_{k=1}^{n} \delta_k B_t(T_k)}.$$

If we consider a single payment date, we obtain the classical formula for the forward rate in the pre-crisis environment:

$$K^{OIS}(t; T, T+\delta) = -\frac{B_t(T+\delta) - B_t(T)}{\delta B_t(T+\delta)} = \frac{1}{\delta}\Big[\frac{B_t(T)}{B_t(T+\delta)} - 1\Big]. \tag{1.13}$$

In the following, we will denote the simply compounded forward rate $K^{OIS}(t; T, T+\delta)$ with $L^D(t; T, T+\delta)$, because, as we will see in the next section, it will be associated with the discount curve.

### 1.2.3 The choice of the discount curve

In (1.3) we have chosen the discount curve, used to compute the price of a fixed-income instrument in the post-crisis framework, as a money market account, which pays the OIS rate. We have followed this strategy, because, as we have seen, the overnight rate determines very low risk, thanks to its short maturity, and then we can consider it risk free. Moreover, in Subsection 1.1.2, we have denoted with $B_t^{OIS}(T)$ the OIS bond price processes, which are not necessarily traded in the market, but they are simply determined by the OIS rate (1.13) through bootstrap algorithms, as done in the pre-crisis environment (for more details, see [1]). Using OIS bond price processes is a good choice, also because $B_t^{OIS}(T)$ is associated with the reference tenor structure (that is the one which contains more dates) and it can be used to compare the Bonds associated with the other tenor structures.
From $B_t^{OIS}(T)$, which in the following will be simply denoted by $B_t(T)$, we define the instantaneous forward rate, as done in the pre-crisis setting (see (A.3)), $f_t(T) = -\frac{\partial \log B_t(T)}{\partial T}$. To do this, we assume that the prices curve $T \to B_t(T)$

is sufficiently regular to compute the forward rate $f_t(T)$. Finally, we define the instantaneous short rate: $r_t = f(t,t)$.

Given the OIS short rate $r_t$, we define money market account in the same way of (A.4):

$$B_t = exp\Big(\int_0^t r_s ds\Big). \tag{1.14}$$

Then we consider a martingale probability measure $\mathbb{Q}$, equivalent to the objective one $\mathbb{P}$, under which all discounted by $B_t$ traded assets are martingales. In particular we postulate the condition for the OIS bond price processes $B_t(T)$:

$$B_t(T) = \mathbb{E}^{\mathbb{Q}}\Big\{\frac{B_t}{B_T}|\mathcal{F}_t\Big\} = \mathbb{E}^{\mathbb{Q}}\Big\{exp\Big[-\int_t^T r_s ds\Big]|\mathcal{F}_t\Big\}. \tag{1.15}$$

Since the process $\left(\frac{B_t(T)}{B_t}\right)_{t\leq T}$ is a $\mathbb{Q}$-martingale, after a normalization with $B_0(T)$, we can use it as density process to change $\mathbb{Q}$, with the equivalent forward mesaure $\mathbb{Q}^T$, which will be used in order to compute prices of other market instruments.

# 1.3 Heath-Jarrow-Morton approach in post-crisis framework

## 1.3.1 The parameterization of spreads

In the context described in the previous sections, we aim at adopting an Heath-Jarrow-Morton approach (A.3) to describe all interest rate curves, each one associated to a different tenor $\delta$. We follow the article [8].

We have seen in section 1.2.2 that we can assume the OIS rate $L^D(t;T,T+\delta)$ (defined on (1.13)) to be risk-free, whereas, adopting the concept of tenor structure 1.2.1 associated with the set of tenors $\mathcal{D}$ we have to manage with a set of LIBOR forward rates, each of them associated with a tenor $\delta$ and defined as (1.6). As we have seen, the LIBOR forward rates no longer respect classical pre-crisis relation, but also the following inequality is typically verified:

$$L^\delta(t;T,T+\delta) > L^D(t;T,T+\delta). \tag{1.16}$$

Moreover, we can observe from market data that the Libor rate is an increasing function of tenor $\delta$. We can observe this property in figure 1.1

As a consequence of the inequality (1.16) it is convenient to follow a multi-curve approach. We can model the OIS rate $L^D$ and a family of multiplicative spread processes ,each of them associated with a tenor $\delta$. These spreads will be related to

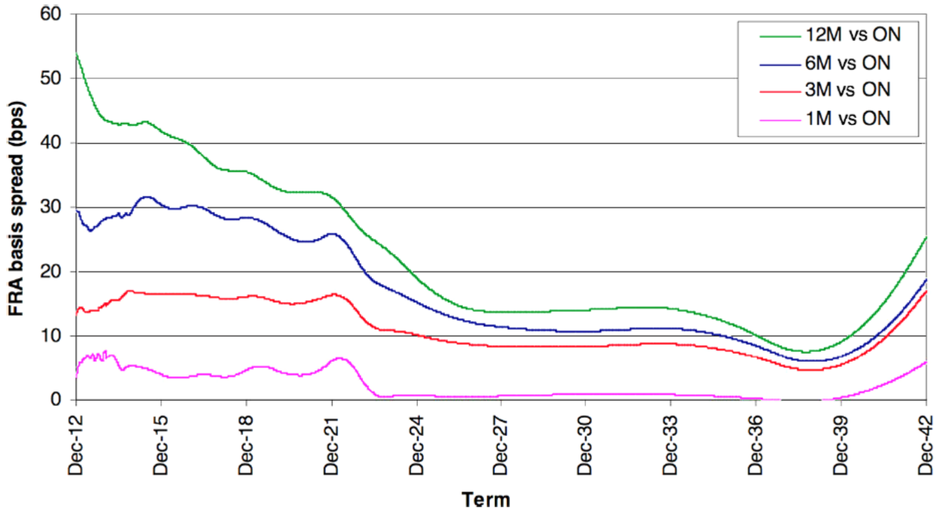


Figure 1.1: Term structure of additive spreads between FRA rates and OIS forward rates, on Dec. 11, 2012 for $\delta \in \{\frac{1}{12}, \frac{3}{12}, \frac{6}{12}, 1\}$. Source [8]

the credit and liquidity risk associated with the LIBOR forward rate $L^\delta$.directly the different LIBOR rates, but, chosen the OIS rate $L^D(t;T,T+\delta)$.

Hence we can give the following definition:

**Definition 1.3.1.** *The multiplicative forward spread rate between the LIBOR forward rate, defined on $\delta$-tenor structure, and the OIS rate is:*

$$S^\delta(t,T) := \frac{1+\delta L^\delta(t;T,T+\delta)}{1+\delta L^D(t;T,T+\delta)}, \tag{1.17}$$

*in particular, the spot spread rate, between the respective spot LIBOR rates respects:*

$$S^\delta(T,T) = \frac{1+\delta L^\delta(T;T,T+\delta)}{1+\delta L^D(T;T,T+\delta)}. \tag{1.18}$$

In this context the process $(S^\delta(T,T))_T$ represents the evaluation, given by the market, of the LIBOR panel credit and liquidity quality, at time $T$ and for the time interval $[T,T+\delta]$.

Recalling that the Lebesgue-Radon-Nykodim derivative between the $(T+\delta)$-forward measure and the $T$-forward measure is $L_t = \frac{d\mathbb{Q}^{T+\delta}}{d\mathbb{Q}^T}\Big|_{\mathcal{F}_t} = \frac{B_t(T+\delta)}{B_0(T+\delta)}\frac{B_0(T)}{B_t(T)}$

$\forall t \leq T$, we can derive some properties of the spread process $S^\delta(t,T)$:

$$
\begin{aligned}
S^\delta(t,T) =& \frac{1+\delta L^\delta(t;T,T+\delta)}{1+\delta L^D(t;T,T+\delta)} \\
=& \frac{B_t(T+\delta)}{B_t(T)} \mathbb{E}^{\mathbb{Q}^{T+\delta}}[1+\delta L(T;T,T+\delta)|\mathcal{F}_t] \\
\overset{B.T.}{=}& \frac{B_t(T+\delta)}{B_t(T)} \frac{\mathbb{E}^{\mathbb{Q}^T}[(1+\delta L(T;T,T+\delta))\cdot L_T|\mathcal{F}_t]}{L_t} \\
=& \frac{B_t(T+\delta)}{B_t(T)} \frac{B_0(T+\delta)\cdot B_t(T)}{B_t(T+\delta)\cdot B_0(T)} \mathbb{E}^{\mathbb{Q}^T}\left[(1+\delta L(T;T,T+\delta))\frac{B_T(T+\delta)}{B_0(T+\delta)}\frac{B_0(T)}{B_T(T)}\Big|\mathcal{F}_t\right] \\
=& \mathbb{E}^{\mathbb{Q}^T}\left[B_T(T+\delta)(1+\delta L(T;T,T+\delta))\Big|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}^T}\left[S^\delta(T,T)\Big|\mathcal{F}_t\right].
\end{aligned}
\tag{1.19}
$$

From the last equivalence, we can observe that the process $(S(t,T))_t$ is a $\mathbb{Q}^T$-martingale.

In order to develop the HJM framework, it is moreover convenient to split the spread process $S^\delta(t,T)$ in the spot component $S^\delta(t,t)$ and a forward component. In particular, we assume that:

**Assumption 1.3.2.** *In accord with [12], we assume that for each $t \leq T$, it holds:*

$$
S^\delta(t,T) = S^\delta(t,t)\frac{B_t^\delta(T)}{B_t(T)}, \tag{1.20}
$$

where the term $B_t^\delta(T)$ can be interpreted as a fictitious bond, since the classical terminal bond equivalence holds: $B_t^\delta(t) = 1$, $\forall\ t \in \mathbb{R}_+$.

**Remark 1.3.3.** *Through the previous assumption, we can observe that the fictitious bond's price curve is given by:*

$$
\begin{aligned}
B_t^\delta(T) =& \frac{S^\delta(t,T)}{S^\delta(t,t)}B_t(T) = \frac{1+\delta L^\delta(t;T,T+\delta)}{1+\delta L^D(t;T,T+\delta)}\cdot\frac{1+\delta L^D(t;t,t+\delta)}{1+\delta L^\delta(t;t,t+\delta)}B_t(T) = \\
=& \frac{1+\delta L^\delta(t;T,T+\delta)}{1+\delta L^\delta(t;t,t+\delta)}\cdot\frac{B_t(T+\delta)B_t(t)}{B_t(T)B_t(t+\delta)}B_t(T) = \\
=& \frac{1+\delta L^\delta(t;T,T+\delta)}{1+\delta L^\delta(t;t,t+\delta)}\cdot\frac{B_t(T+\delta)}{B_t(t+\delta)}.
\end{aligned}
$$

### 1.3.2 HJM approach description

In this paragraph we describe the Heath-Jarrow-Morton approach, which we will use in the following of the dissertation.

Our aim is to model an interest rates market composed by $m+1$ curves: one curve associated to the OIS curve, chosen as the discounting curve and one LIBOR rate for each given tenor $\delta \in \{1, \dots, m\}$. In order to adopt the HJM approach (see A.3), based on multiplicative spreads, we follow [8], using a slightly different (but equivalent) parameterization. To do this let us consider the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T^*]}, \mathbb{Q})$, defined on the first section, where $\mathbb{Q}$ is a martingale probability measure.

**OIS Curve** For the OIS curve, we use the same parameterization of [[8], Section 3.2], based on instantaneous forward rates $f_t(T)$, for $0 \leq t \leq T < T^*$:

$$f_t(T) = f_0(T) + \int_0^t \widetilde{\alpha}_s(T)ds + \int_0^t \widetilde{\sigma}_s(T)dW_s, \tag{1.21}$$

where $W = (W_t)_{t\geq 0}$ is and $\mathbb{R}^d$-valued Brownian motion and $\widetilde{\alpha}$ and $\widetilde{\sigma}$ satisfy the same conditions of A.2.1.
Moreover, we can pass to the Musiela parameterization (A.3.1): $r_t(x) := f_t(t+x)$. We obtain the following dynamics:

$$dr_t(x) = \Big(\frac{\partial}{\partial x} r_t(x) + \widetilde{\sigma}_t(t+x)\int_0^x \widetilde{\sigma}_t(t+u)^* du\Big)dt + \widetilde{\sigma}_t(t+x)dW_t, \tag{1.22}$$

where the volatility term is a row vector and, with $A^*$, we denote the transpose of the vector or the matrix $A$.

Finally, we denote with $B_t(T) = exp\big(-\int_0^{T-t} r_t(x)dx\big)$ the price of an OIS zero-coupon bond.

**Libor Curve** The Libor curve, associated with the tenor $\delta$, is obtained by the multiplicative spread process $(S^\delta(t,T))_{t\in[0,T]}$, for each $T \leq T^*$, defined as in (1.17). Moreover, we choose to adopt the parameterization of $(S^\delta(t,T))_{t\in[0,T]}$ described in (1.20).

The fictitious Bond, associated with tenor $\delta$ (also called $\delta$-bond) and introduced in (1.20) is supposed to have the following structure:

$$B_t^\delta(T) := exp\big(-\int_t^T f_t^\delta(u)du\big), \tag{1.23}$$

where the associated forward rate process $(f_t^\delta(T))_{t\in[0,T]}$ is given by:

$$f_t^\delta(T) = f_0^\delta(T) + \int_0^t \alpha_s^\delta(T)ds + \int_0^t \sigma_s^\delta(T)dW_s. \tag{1.24}$$

Finally we give the following assumption:

**Assumption 1.3.4.** *We impose that*

$$S^\delta(t,t) = e^{Y_t^\delta}, \tag{1.25}$$

*where* $(Y_t^\delta)_{t\geq 0}$ *is an adapted Itô process, which dynamics is driven by a* $\mathbb{Q}$*-Wiener process.*

In particular, the exponent process $(Y_t^\delta)_t$, is supposed to satisfy

$$Y_t^d = Y_0^d + \int_0^t \gamma_s^\delta ds + \int_0^t \beta_s^\delta dW_s. \tag{1.26}$$

We assume that all the processes introduced to define all the dynamics respect assumptions A.2.1.

**HJM drift condition** After the crisis, we have observed that FRA contracts have to be explicitly considered in the fixed-income market. Recalling the equivalence (1.3), we are going to describe the price of a FRA contract in terms of multiplicative spreads. By the formula of FRA value at time $t$ (1.3), we can observe that:

$$\begin{aligned}
\Pi^{FRA}(t;T,T+\delta,K) &= \delta B_t(T+\delta)\mathbb{E}^{\mathbb{Q}^{T+\delta}}[(L^\delta(T;T,T+\delta)-K)|\mathcal{F}_t] = \\
&= \delta B_t(T+\delta)(L^d(t;T,T+\delta)-K) = \\
&\overset{\bigstar}{=} \delta B_t(T+\delta)\Big[\frac{(1+\delta L_t^D(t;T,T+\delta))S^\delta(t,T)-1}{\delta} - K\Big] = \\
&= B_t(T+\delta)\Big[\frac{B_t(T)}{B_t(T+\delta)}S^\delta(t,T) - (\delta K+1)\Big] = \\
&= B_t(T)S^\delta(t,T) - B_t(T+\delta)(\delta K+1) = \\
&= B_t(T)S^\delta(t,t)\frac{B_t^\delta(T)}{B_t(T)} - B_t(T+\delta)(\delta K+1) = \\
&= S^\delta(t,t)B_t^\delta(T) - B_t(T+\delta)(\delta K+1),
\end{aligned} \tag{1.27}$$

where in equivalence $\bigstar$ we have used the definition of spread: $S^\delta(t,T) = \frac{1+\delta L^\delta(t;T,T+\delta)}{1+\delta L^D(t;T,T+\delta)}$ and the classical pre-crisis relation, which holds for OIS bonds:
$1+\delta L^D(t;T,T+\delta) = \frac{B_t(T)}{B_t(T+\delta)}$.

As we have observed in Section 1.2.3, the term $\left(\frac{B_t(T)}{B_t}\right)$ is already a $\mathbb{Q}$-martingale. In order to get absence of arbitrage in fixed-income market, we need to find conditions under which also the leg dependent on the spread is a $\mathbb{Q}$-martingale, when discounted by the bank account defined on (1.14).

We have thus to analyse the dynamics of the following process:

$$K_t^{T,\delta} = \frac{S^\delta(t,t)B_t^\delta(T)}{B_t} \tag{1.28}$$

Hence we compute:

$$\begin{aligned}
S^\delta(t,t)B_t^\delta(T) &= \exp\Big\{\overbrace{Y_0^\delta + \int_0^t \gamma_s^\delta ds + \int_0^t \beta_s^\delta dW_s}^{Y_t} - \int_t^T \overbrace{\Big[f_0^\delta(u) + \int_0^t \alpha_s^\delta(u)ds + \int_0^t \sigma_s^\delta(u)dW_s\Big]}^{f_t^\delta(u)} du\Big\} = \\
&\overset{F.T}{=} \exp\Big\{Y_0^\delta + \int_0^t \gamma_s^\delta ds + \int_0^t \beta_s^\delta dW_s - \int_0^T f_0^\delta(u)du + \int_0^t f_0^\delta(u)du + \\
&\quad - \int_0^t \Big(\int_t^T \alpha_s^\delta(u)du\Big)ds - \int_0^t \Big(\int_t^T \sigma_s^\delta(u)du\Big)dW_s\Big\} = \\
&= \exp\Big\{Y_0^\delta + \int_0^t \gamma_s^\delta ds + \int_0^t \beta_s^\delta dW_s - \int_0^T f_0^\delta(u)du + \int_0^t f_0^\delta(u)du + \\
&\quad - \int_0^t \Big(\int_s^T \alpha_s^\delta(u)du - \int_s^t \alpha_s^\delta(u)du\Big)ds - \int_0^t \Big(\int_s^T \sigma_s^\delta(u)du - \int_s^t \sigma_s^\delta(u)du\Big)dW_s\Big\} = \\
&= \exp\Big\{Y_0^\delta + \int_0^t \gamma_s^\delta ds + \int_0^t \beta_s^\delta dW_s - \int_0^T f_0^\delta(u)du + \int_0^t f_0^\delta(u)du + \\
&\quad + \int_0^t \Big(\int_s^t \alpha_s^\delta(u)du\Big)ds + \int_0^t \Big(\int_s^t \sigma_s^\delta(u)du\Big)dW_s + \int_0^t A_s(T)ds + \int_0^t \Sigma_s(T)dW_s\Big\},
\end{aligned}$$

where

$$\begin{cases} A_s(T) = -\int_s^T \alpha_s^\delta(u)du = -\int_0^{T-s} \alpha_s^\delta(s+u)du; \\ \Sigma_s(T) = -\int_s^T \sigma_s^\delta(u)du = -\int_0^{T-s} \sigma_s^\delta(s+u)du. \end{cases}$$

Moreover, using the stochastic version of Fubini Theorem (for the proof see [14], chapter 6, Theorem 6.2), we obtain:

$$\begin{aligned}
\int_0^t f_u^\delta(u)du &= \int_0^t \Big[f_0^\delta(u) + \int_0^u \alpha_s^\delta(u)ds + \int_0^u \sigma_s^\delta(u)dW_s\Big]du = \\
&= \int_0^t f_0^\delta(s)ds + \int_0^t \Big(\int_s^t \alpha_s^\delta(u)du\Big)ds + \int_0^t \Big(\int_s^t \sigma_s^\delta(u)du\Big)dW_s.
\end{aligned}$$

Finally:

$$S^\delta(t,t)B_t^\delta(T) = exp\Big\{Y_0^\delta - \int_0^T f_0^\delta(u)du + \int_0^t \Big[\gamma_s^\delta + f_s^\delta(s) + A_s(T)\Big]ds + \int_0^t \Big[\beta_s^\delta + \Sigma_s(T)\Big]dW_s\Big\}. \tag{1.29}$$

We recall that the process $(K_t^{T,\delta})_{t\in T}$ is a local martingale if it does not admit drift term. In particular, applying Itô formula: $dK_t^{T,\delta} = K_t^{T,\delta}(\mu_t dt + \nu_t dW_t)$, where:

$$\mu_t = \gamma_t^\delta + f_t^\delta(t) + A_t(T) + \frac{1}{2}||\beta_t^\delta + \Sigma_t(T)||^2 - r_t(0),$$

we have to impose that:

$$\gamma_t^\delta + f_t^\delta(t) + A_t(T) + \frac{1}{2}||\beta_t^\delta + \Sigma_t(T)||^2 - r_t(0) = 0, \quad \forall\, t \in [0,T], \quad \forall\, T \leq T^*. \quad (1.30)$$

In particular, if $t = T$ we get: $\gamma_t^\delta + f_t^\delta(t) + \frac{1}{2}||\beta_t^\delta||^2 - r_t(0) = 0$.
By the previous equivalence, we obtain the drift condition:

$$\begin{aligned}\alpha_t(T) &= -\frac{1}{2}||\beta_t^\delta + \Sigma_t(T)||^2 + \frac{1}{2}||\beta_t^\delta||^2 = \\ &= -\beta_t^\delta \Sigma_t(T)^* - \frac{1}{2}||\Sigma_t(T)||^2.\end{aligned}$$

differentiating with respect to the $T$ variable we get:

$$\gamma_t^\delta(T) = -\beta_t^\delta \sigma_t^{\delta *}(T) + \sigma_t^\delta(T)\Big(\int_0^{T-t} \sigma_t^\delta(t+u)du\Big)^*. \quad (1.31)$$

Then, if we consider the forward rate process $(f_t^\delta(T))_{t\in[0,T]}$ described through Musiela parameterization ($r_t^\delta(x) = f_t^\delta(t+x)$), we obtain

$$\begin{aligned}dr_t^\delta(x) =& df_t^\delta(t+x) + \frac{\partial}{\partial T} f_t^\delta(t+x)dt = \\ =& \alpha_t^\delta(t+x)dt + \sigma_t^\delta(t+x)dW_t + \frac{\partial}{\partial x} f_t^\delta(t+x)dt = \\ =& \Big[-\beta_t^\delta \sigma_t^{\delta *}(t+x) + \sigma_t^\delta(t+x)\Big(\int_0^x \sigma_t^\delta(t+u)du\Big)^* + \frac{\partial}{\partial x} r_t^\delta(x)\Big]dt + \sigma_t^\delta(t+x)dW_t.\end{aligned}$$

whereas the Itô process $(Y_t^\delta)_t$ which determines the exponent of the spot spread process satisfies the following dynamics:

$$dY_t^\delta = \Big(-f_t^\delta(t) - \frac{1}{2}||\beta_t^\delta||^2 + r_t(0)\Big)dt + \beta_t^\delta dW_t \quad (1.32)$$

**Conclusions** The HJM approach and the condition of arbitrage free market have determined the following system of SDEs:

$$\begin{cases} \text{[OIS Curve] } dr_t(x) = \Big(\mathbf{F} r_t(x) + \widetilde{\sigma}_t(t+x)\mathbf{H}\widetilde{\sigma}\Big)dt + \widetilde{\sigma}_t(t+x)dW_t; \\ \text{[Libor Curve] } dr_t^\delta(x) = \Big[-\beta_t^\delta \sigma_t^{\delta *}(t+x) + \sigma_t^\delta(t+x)\mathbf{H}\sigma^\delta + \mathbf{F} r_t^\delta(x)\Big]dt + \sigma_t^\delta(t+x)dW_t; \\ \text{[Log Spot Spread] } dY_t^\delta = \Big(-r_t^\delta(0) - \frac{1}{2}||\beta_t^\delta||^2 + r_t(0)\Big)dt + \beta_t^\delta dW_t. \end{cases} \quad (1.33)$$

where $\mathbf{F} := \frac{\partial}{\partial x}$, $\mathbf{H}\sigma = \int_0^x \sigma_t^*(t+u)du$. The previous system is composed by $2m+1$ stochastic differential equations, 2 for each tenor $\delta$ and one for the OIS curve.

## 1.4 Foreign exchange analogy

It is possible to observe an analogy between spot spread processes defined on the previous section $(S^{\delta}(t,t))_t$ and an exchange rate, which characterizes a multi currency framework. Under this interpretation, we can represent model (1.33) in this way:

- Each fictitious bond $B_t^{\delta}(T)$, associated with LIBOR interest rates defined on the $\delta$-tenor structure, can be interpreted as a Zero-Coupon Bond traded in a foreign risky market;
- OIS ZCBs are associated with the domestic contracts.

We consider a market defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T^*}, \mathbb{P})$, where $\mathbb{P}$ is classical objective probability measure.

If $p_t^D(t+x)$, $p_t^F(t+x)$ are respectively the price processes of a domestic $ZCB$ and a foreign $ZCB$, with maturity date $t+x$. Defining the respective instantaneous forward rates $r_t^D(x), r_t^F(x)$, the classical pre-crisis HJM framework (obtained by Musiela parameterization) can be used:

$$\begin{cases} dr_t^D = \Big\{\mathbf{F}r_t^D(x) + \sigma_t^D(t+x)\mathbf{H}\sigma^D\Big\}dt + \sigma_t^D(t+x)dW_t^D, & r_0^D(x) = r_0^{D,0}(x); \\ dr_t^F = \Big\{\mathbf{F}r_t^F(x) + \sigma_t^F(t+x)\mathbf{H}\sigma^F\Big\}dt + \sigma_t^F(t+x)dW_t^F, & r_0^F(x) = r_0^{F,0}(x); \end{cases}$$

where the meaning $\mathbf{F}, \mathbf{H}$ is the same of (1.33).
In the previous system, the random sources are respectively driven by a $\mathbb{Q}^D$-Wiener process $W_t^D$ and a $\mathbb{Q}^F$-Wiener process $W_t^F$, where $\mathbb{Q}^D, \mathbb{Q}^F$ are martingale measures for the respective currency markets.
As done in section 1.2.3, we assume that the evolution of money account in each market $B_t^K = exp\Big\{\int_0^t r_s^K(0)ds\Big\}$ for $K \in \{D, F\}$. Then it holds:

$$\begin{cases} dB_t^F = B_t^F r_t^F(0)dt; \\ dB_t^D = B_t^D r_t^D(0)dt. \end{cases}$$

In order to have two arbitrage free markets, the martingale measure $\mathbb{Q}^K$, with $K \in \{D, F\}$ is obtained supposing that all discounted prices in each market are $\mathbb{Q}^K$-martingale. We assume that the exchange rate process $(S_t)_t$ follows this dynamics:

$$dS_t = S_t(\gamma_t dt + \eta_t dW_t^D). \tag{1.34}$$

This process represents the following equivalence: we can buy the foreign currency and invest in the foreign market (with the foreign short rate $r_t^F(0)$ and in an equivalent way we can invest in a domestic asset determined by the money account of the foreign market evaluated in the domestic currency through the exchange rate process $\hat{B}_t^F = S_t B_t^F$ . In particular, the dynamics of $\hat{B}_t^F$ is:

$$
\begin{aligned}
d\hat{B}_t^F =& d(S_t B_t^F) = dB_t^F \cdot S_t + B_t^F \cdot dS_t + \overbrace{d[B^F_t, S]}^{=0} = \\
=& \hat{B}_t^F[(r_t^F(0) + \gamma_t)dt + \eta_t dW_t^D].
\end{aligned}
$$

$\hat{B}_t^F$ is the price of a contract quoted in the domestic market, then the associated discounted price has to be a $\mathbb{Q}^D$-martingale. As done before, we compute the differential of the discounted price. Successively, we impose that the drift term of this process is null.

$$
d\Big(\frac{\hat{B}_t^F}{B_t^D}\Big) = \frac{\hat{B}_t^F}{B_t^D}[(r_t^F(0) - r_t^D(0) + \gamma_t)dt + \eta_t dW_t^D],
$$

then, the condition on the drift is: $\gamma_t = r_t^D(0) - r_t^F(0)$. Hence, we obtain:

$$
dS_t = S_t((r_t^D(0) - r_t^F(0))dt + \eta_t dW_t^D).
$$

Passing to the logarithm $Y = \log S$, we get:

$$
dY_t = \Big\{ r_t^D(0) - r_t^F(0) - \frac{1}{2}||\eta_t||^2 \Big\} dt + \eta_t dW_t^D. \tag{1.35}
$$

Moreover, the Lebesgue-Radon-Nikodym derivative between measure $\mathbb{Q}^F, \mathbb{Q}^D$ on $\mathcal{F}_t$ is:

$$
L_t = \frac{S_t}{S_0} \exp\Big\{ -\int_0^t (r_s^D(0) - r_s^F(0))ds \Big\},
$$

therefore, the relation between the two Wiener processes is:

$$
dW_t^F = dW_t^D - \eta_t dt.
$$

This condition allows us to describe the foreign forward rate dynamics in driven by the domestic Brownian motion.

In conclusion, the system composed by the domestic forward rate, the foreign forward rate and the exchange process is:

$$
\begin{cases}
dr_t^D = \Big\{ \mathbf{F} r_t^D(t+x) + \sigma_t^D(t+x)\mathbf{H}\sigma^D \Big\} dt + \sigma_t^D(t+x) dW_t^D, \\
r_0^D(x) = r_0^{D,0}(x); \\
dr_t^F = \Big\{ \mathbf{F} r_t^F(t+x) + \sigma_t^F(t+x)\mathbf{H}\sigma^F - \sigma_t^F(t+x)\eta_t^* \Big\} dt + \sigma_t^F(t+x) dW_t^F, \\
r_0^F(x) = r_0^{F,0}(x); \\
dY_t = \Big\{ r_t^D(0) - r_t^F(0) - \frac{1}{2}||\eta_t||^2 \Big\} dt + \eta_t dW_t^D.
\end{cases} \tag{1.36}
$$

We can see that this system is equivalent to (1.33). In the following of the dissertation, we will use this analogy to describe some properties of fixed-income market in post-crisis framework. Indeed, we will exploit the techniques developed by Slinko in [21], in order to find conditions under which the infinite-dimensional system (1.33) possesses finite dimensional realizations (we will describe these concept in details in chapter 3).

# Chapter 2

# The Geometric Approach and The Consistency Problem

At the end of Chapter 1, we introduced the forward rate system which describes the dynamics of instantaneous forward rates associated with each tenor $\delta$, belonging to a finite set of tenors $\mathcal{D} = \{\delta_1, \ldots, \delta_m\}$.

In this chapter, we aim at describing the problem of consistency in the post-crisis context. To this effect, we will adopt the geometric approach described by Björk in [5]. This approach provides a different interpretation of the system (1.33), which is interpreted as a finite-dimensional system of SDEs, each of them defined on an infinite-dimensional space. We aim at generalizing the strategy developed in [2], in order to find conditions which guarantee that couple $(\mathcal{M}, \mathcal{G})$ is consistent, where $\mathcal{M}$ and $\mathcal{G}$ denote respectively a forward rate model and a parameterized family of forward rates. The concept of consistency can be introduced as follows: we say that an interest rate model $\mathcal{M}$ and a parameterized family of forward rate curves $\mathcal{G}$ are consistent if $\mathcal{M}$ produces forward rate curves which belong to $\mathcal{G}$ for a strictly positive time interval.

Mathematical finance is interested in the previous concept because the problem of consistency is related to the problem of parameter recalibration of a concrete interest rate model. The parameter recalibration is essential in the analysis of a financial market through a model, because when we use a model $\mathcal{M}$ in order to describe the fixed-income market (i.e. we define a volatility term $\hat{\sigma}(\hat{r}_t)$ which determines a forward rate system as (1.33)) we have to take into account the fact that $\mathcal{M}$ is an approximation of the real financial market, hence, after a sufficient time interval, the comparison between the values provided by the model $\mathcal{M}$ and the market data will not coincide. Therefore, recalibrating the parameters of the model using the current market data, we can correct the behaviour of $\mathcal{M}$, adding the information given by the market data.

In order to recalibrate a model we have to develop the following strategy. First of all, we have to deal with the problem of production of a forward rate curve $\Gamma^M = \{r^M(x);\ x \geq 0\}$ from market data. Indeed, only a finite number of bonds are actually traded in the market, then we have to fit a finite set of points to obtain the entire term structure $\Gamma^M$. In order to do this, we can follow several approaches. The main strategies we can follow are described in [14] Chapter 3 and they consist in using splines or parameterized families of smooth forward rate curves, such as the Nelson-Siegel family or the Svensson family, which will be studied in details in Section 2.3.

When we have provided the term-structure from market data, we have to deal with the problem of recalibration. In order to face this problem, we can follow a strategy which takes into account times series combined with cross-section data. These strategies are justified only from a statistical point of view, hence, deeper theoretical motivations are related to the concept of consistency, between the dynamics of a given model $\mathcal{M}$ and the term structure determined by a parameterized forward rate family $\mathcal{G}$.

This chapter is structured as follows: in the first sections, we will provide a formal characterization of this concept of consistency in the post-crisis framework. Then, we will discuss the validity of the general consistency conditions in the context of several specific examples. The class of models and parameterized families which will be studied is inspired by [2].

## 2.1 The geometric approach

The system (1.33) is a system of SDEs depending on a positive real parameter $x$ (time to maturity). If we try to analyse the properties of this system directly, we have to deal with an infinite number of SDEs. In order to overcome this problem, we can interpret each equation of the system as a unique SDE, defined on an infinite-dimensional space. For ease of presentation, let us first consider only the OIS forward curve.

In order to formalize this idea, we use from now, this notation:

$$\begin{cases} r_t: \text{ forward rate curve at time } t\ , \\ r: \text{ the stochastic process } (r_t)_{t\geq 0} \text{ of forward rate curves}\ . \end{cases}$$

The stochastic process $r$ can be interpreted as a curve evolving on a infinite dimensional space:

$$\mathcal{H} \subset \mathcal{C}^{+\infty}(\mathbb{R}_+, \mathbb{R}).$$

Using this notation for $r: \mathbb{R}_+ \to \mathcal{H}$, $r_t$ can be interpreted as a point on $\mathcal{H}$.

In what follows, we will suppose that each equation of the system (1.33) respects some particular properties, which lead to the following definition of the space $\mathcal{H}$:

**Definition 2.1.1.** *For each $t \geq 0$, the solution of each forward rate equation of the system* (1.33) *at time $t$, $r_t$ and $r^i_t$ $i = 1, \dots, m$, belongs to the following infinite-dimensional space:*

$$\mathcal{H} := \{r : \mathbb{R}_+ \to \mathbb{R} \text{ infinite times differentiable, and s.t. } ||r||_\gamma < +\infty\},$$

*where the norm $|| \cdot ||_\gamma$ is defined as follows:*

$$||r||^2_\gamma = \sum_{n=0}^{+\infty} 2^{-n} \int_0^{+\infty} \Big(\frac{\partial^n}{\partial x^n} r(x)\Big)^2 e^{-\gamma x} dx, \qquad \gamma > 0.$$

*We have used the convention $\frac{\partial^0}{\partial x^0} r(x) \equiv r(x)$.*

The space $(\mathcal{H}, ||\cdot||_\gamma)$ is an Hilbert space for each $\gamma > 0$ (we refer to [4][Proposition 4.2] for the proof of this result), then we fix a value for $\gamma$ and in the following, for simplicity of notation, we denote the norm without the subscript.

**Remark 2.1.2.** *The choice of such a norm is necessary to guarantee the existence of a strong solution for the first $m + 1$ rows of the system* (1.33) *(associated with the infinite-dimensional dynamics). Indeed, the operator $\boldsymbol{F} : \mathcal{H} \longrightarrow \mathcal{H}$, defined by $\boldsymbol{F} := \frac{\partial}{\partial x}$ is bounded:*

$$\begin{aligned}
||\boldsymbol{F} r||^2 =& \sum_{n=0}^{+\infty} 2^{-n} \int_0^{+\infty} \Big(\frac{\partial^n}{\partial x^n}\Big(\frac{\partial}{\partial x} r(x)\Big)\Big)^2 e^{-\gamma x} dx \\
=& 2\sum_{j=0}^{+\infty} 2^{-j} \int_0^{+\infty} \Big(\frac{\partial^j}{\partial x^j} r(x)\Big)^2 e^{-\gamma x} dx - \overbrace{2\int_0^{+\infty} r^2(x) e^{-\gamma x} dx}^{\in[0,+\infty)} \\
\leq& 2||r||^2 < +\infty.
\end{aligned}$$

*Recalling that the operator norm is defined as:*

$$||\boldsymbol{F}|| := \sup_{r \in \mathcal{H}\setminus\{0\}} \Big\{\frac{||\boldsymbol{F} r||}{||r||}\Big\},$$

*we conclude that: $||\boldsymbol{F}|| \leq \sqrt{2}$.*

If we generalize this approach to multi-curve framework, we have to interpret each solution of the first $m+1$ equations of system (1.33), as a function on a space isomorphic to $\mathcal{H}$. We introduce the following notation:

$$r \longrightarrow r^0, \tag{2.1}$$
$$r^{\delta_i} \longrightarrow r^i, \qquad i \in \{1, \dots, m\}, \tag{2.2}$$
$$\beta^{\delta_i} \longrightarrow \beta^i, \qquad i \in \{1, \dots, m\}. \tag{2.3}$$

Under this notation, the entire solution of the system (1.33) can be interpreted as a vector forward rate process defined on the space:

$$\hat{\mathcal{H}} := \mathcal{H}^0 \times \cdots \times \mathcal{H}^m \times \mathbb{R}^m,$$

where $\mathcal{H}^i \equiv \mathcal{H}$ and we recall that $d$ is the dimension of the Brownian motion which drives the stochasticity of the model. In particular, each $\mathbb{R}$ component is associated with the spread process associated with a tenor $\delta_i$. $\hat{\mathcal{H}}$ is still an Hilbert space since it is a finite product of Hilbert spaces and the solution of (1.33) will be denoted in the following by:

$$\hat{r}_t = [\ \overbrace{r_t^0}^{OIS}, \overbrace{r_t^1, \ldots, r_t^m}^{LIBOR}, \overbrace{Y_t^1, \ldots, Y_t^m}^{\text{Log Spot spread}}\ ].$$

**Assumption 2.1.3.** *The dynamics describing system* (1.33) *are completely determined by the volatility terms* $\sigma_t^0(t+x)$, $\sigma_t^{\delta_i}(t+x)$, $\beta_t^{\delta_i}$ $\forall \delta \in \{\delta_1, \ldots, \delta_m\}$ *(this is due to the Heath-Jarrow-Morton drift condition* (1.31)*).*

*We introduce the same notation of* (2.1)*:* $\sigma_t^{\delta_i} \equiv \sigma_t^i$, $\widetilde{\sigma}_t \equiv \sigma_t^0$, $\beta_t^{\delta_i} \equiv \beta_t^i$. *In analogy to [21], we suppose that:*

- *The adapted processes describing the volatility of each component are defined as follows:*
$$\sigma_t^0(t+x) = \sigma^0(\hat{r}_t, t+x);$$
$$\sigma_t^i(t+x) = \sigma^i(\hat{r}_t^i, t+x);$$
$$\beta_t^i = \beta^i(\hat{r}_t),$$
*where* $\sigma^i, \beta^j$ $\ i \in \{0, \ldots, m\},\ j \in \{1, \ldots, m\}$ *are deterministic functions:*
$$\begin{cases} \sigma^0 : \hat{\mathcal{H}} \longrightarrow (\mathcal{H})^d, \\ \sigma^i : \hat{\mathcal{H}} \longrightarrow (\mathcal{H})^d & i \in \{1, \ldots, m\}, \\ \beta^j : \hat{\mathcal{H}} \longrightarrow \mathbb{R}^d & j \in \{1, \ldots, m\}, \end{cases}$$
*supposed to be smooth, in the sense of the Remark B.1.3.*

- *The following mappings are supposed to be smooth:*
$$\begin{cases} \hat{r} \longrightarrow \sigma^0(\hat{r})\boldsymbol{H}\sigma^0(\hat{r}) - \frac{1}{2}\frac{\partial \sigma^0}{\partial \hat{r}}(\hat{r}_t)\hat{\sigma}(\hat{r}_t), \\ \\ \hat{r} \longrightarrow \sigma^i(\hat{r})\boldsymbol{H}\sigma^i(\hat{r}) - \frac{1}{2}\frac{\partial \sigma^0}{\partial \hat{r}}(\hat{r}_t)\hat{\sigma}(\hat{r}_t) - \sigma^i(\hat{r}_t)\beta^{i*}(\hat{r}), \quad i = 1, \ldots, m. \end{cases}$$

In particular, we rewrite the system (1.33) as:

$$d\hat{r}_t = \mu(\hat{r}_t)dt + \hat{\sigma}(\hat{r}_t)dW_t, \tag{2.4}$$

where:

$$\mu(\hat{r}) = \begin{pmatrix} \mathbf{F}r^0 + \sigma^0(\hat{r})\mathbf{H}\sigma^0(\hat{r}) \\ \mathbf{F}r^1 + \sigma^1(\hat{r})\mathbf{H}\sigma^1(\hat{r}) - \beta^1\sigma^{1*}(\hat{r}) \\ \vdots \\ \mathbf{F}r^m + \sigma^m(\hat{r})\mathbf{H}\sigma^m(\hat{r}) - \beta^m\sigma^{m*}(\hat{r}) \\ \mathbf{B}r^0 - \mathbf{B}r^1 - \frac{1}{2}||\beta^1||^2 \\ \mathbf{B}r^0 - \mathbf{B}r^2 - \frac{1}{2}||\beta^2||^2 \\ \vdots \\ \mathbf{B}r^0 - \mathbf{B}r^m - \frac{1}{2}||\beta^m||^2 \end{pmatrix} \in \hat{\mathcal{H}},$$

where $\mathbf{B}$ denotes the mapping $\mathbf{B} : \hat{\mathcal{H}} \longrightarrow \mathbb{R}$, defined as follows:

$$\mathbf{B}(r) = r(0), \qquad \forall\ r \in \hat{\mathcal{H}}$$

and

$$\hat{\sigma}(\hat{r}) = \begin{pmatrix} \sigma^0(\hat{r}) \\ \sigma^1(\hat{r}) \\ \vdots \\ \sigma^m(\hat{r}) \\ \beta^1(\hat{r}) \\ \vdots \\ \beta^m(\hat{r}) \end{pmatrix} \in \hat{\mathcal{H}}^d. \tag{2.5}$$

For the details regarding infinite-dimensional Itô's formula we recall [9] and [10].

In order to adopt a classical differential approach, we need to use a slightly different notation, based on the Stratonovich integral definition:

**Definition 2.1.4.** *Given two semimartingales $X, Y$, the* Stratonovich *integral of $X$ with respect to $Y$ is defined by:*

$$\int_0^t X_s \circ dY_s = \int_0^t X_s dY_s + \frac{1}{2}\langle X, Y\rangle_t,$$

*where $\langle X_t, Y_t\rangle$ is the quadratic covariation process between $X_t$ and $Y_t$.*

The following proposition can be proved:

**Proposition 2.1.5** (Chain Rule)**.** *If $F(t, y)$ is a smooth function and $Y_t$ is an Itô process, then:*

$$dF(t, Y_t) = \frac{\partial}{\partial t}F(t, Y_t)dt + \frac{\partial}{\partial y}F(t, Y_t) \circ dY_y.$$

*Proof.* By Itô's formula:

$$dF(t,Y_t) = \frac{\partial}{\partial t}F(t,Y_t)dt + \frac{\partial}{\partial y}F(t,Y_t)dY_t + \frac{1}{2}\frac{\partial^2}{\partial y^2}F(t,Y_t)d\langle Y\rangle_t, \tag{2.6}$$

where

$$Y_t = Y_0 + \int_0^t \varphi_s ds + \int_0^t \psi_s dW_s,$$

where, for simplicity, we have supposed that the Brownian motion $(W_t)_{t\geq 0}$ is 1-dimensional.

Then, computing the Itô's derivative of $\frac{\partial}{\partial y}F(t,Y_t)$:

$$\begin{aligned} d\Big(\frac{\partial}{\partial y}F(t,Y_t)\Big) =& \frac{\partial}{\partial t}\frac{\partial}{\partial y}F(t,Y_t)dt + \frac{\partial^2}{\partial y^2}F(t,Y_t)dY_t + \frac{1}{2}\frac{\partial^3}{\partial y^3}F(t,Y_t)d\langle Y\rangle_t \\ =& \Big[\frac{\partial}{\partial t}\frac{\partial}{\partial y}F(t,Y_t)dt + \frac{\partial^2}{\partial y^2}F(t,Y_t)\varphi_t + \frac{1}{2}\frac{\partial^3}{\partial y^3}F(t,Y_t)\psi_t^2\Big]dt + \frac{\partial^2}{\partial y^2}F(t,Y_t)\psi_t dW_t. \end{aligned}$$

Hence

$$d\Big\langle \frac{\partial}{\partial y}F(\cdot,Y),Y\Big\rangle_t = \frac{\partial^2}{\partial y^2}F(t,Y_t)\psi_t^2 dt.$$

On the other hand, by definition of Stratonovich integral:

$$\begin{aligned} \frac{\partial}{\partial y}F(t,Y_t)\circ dY_t =& \frac{\partial}{\partial y}F(t,Y_t)dY_t + \frac{1}{2}d\Big\langle \frac{\partial}{\partial y}F(\cdot,Y),Y\Big\rangle_t \\ =& \frac{\partial}{\partial y}F(t,Y_t)dY_t + \frac{1}{2}\frac{\partial^2}{\partial y^2}F(t,Y_t)\psi_t^2 dt. \end{aligned}$$

Finally, by substituting in (2.6):

$$dF(t,Y_t) = \frac{\partial}{\partial t}F(t,Y_t)dt + \frac{\partial}{\partial y}F(t,Y_t)\circ dY_t.$$

□

Passing to the Stratonovich formulation, we rewrite (2.4) as follows:

$$\begin{aligned} d\hat{r}_t =& \mu(\hat{r}_t)dt + \hat{\sigma}(\hat{r}_t)dW_t \\ =& \mu(\hat{r}_t)dt - \frac{1}{2}d\langle\hat{\sigma}(\hat{r}),W\rangle_t + \hat{\sigma}(\hat{r}_t)\circ dW_t \end{aligned}$$

Recalling by [9][Theorem 4.17] how to compute the Itô's derivative of an infinite-dimensional SDE, we compute:

$$\begin{aligned} d\hat{\sigma}(\hat{r}_t) =& \frac{\partial\hat{\sigma}}{\partial\hat{r}}(\hat{r}_t)d\hat{r}_t + \frac{1}{2}\frac{\partial^2\hat{\sigma}(\hat{r}_t)}{\partial\hat{r}^2}d\langle\hat{r}\rangle_t \\ =& \frac{\partial\hat{\sigma}}{\partial\hat{r}}(\hat{r}_t)\Big[\mu(\hat{r}_t)dt + \hat{\sigma}(\hat{r}_t)dW_t\Big] + \frac{1}{2}\frac{\partial^2\hat{\sigma}}{\partial\hat{r}^2}(\hat{r}_t)\hat{\sigma}(\hat{r}_t)\cdot\hat{\sigma}(\hat{r}_t)dt \\ =& \Big[\frac{\partial\hat{\sigma}}{\partial\hat{r}}(\hat{r}_t)\mu(\hat{r}_t) + \frac{1}{2}\frac{\partial^2\hat{\sigma}}{\partial\hat{r}^2}(\hat{r}_t)\hat{\sigma}(\hat{r}_t)\cdot\hat{\sigma}(\hat{r}_t)\Big]dt + \frac{\partial\hat{\sigma}}{\partial\hat{r}}(\hat{r}_t)\hat{\sigma}(\hat{r}_t)dW_t, \end{aligned}$$

where $\frac{\partial}{\partial \hat{r}}$ denotes the Fréchét derivative.

Then, $d\langle \hat{\sigma}(\hat{r}), W\rangle_t = \frac{\partial \hat{\sigma}}{\partial \hat{r}} \hat{r}_t \hat{\sigma}(\hat{r}_t) dt$. Therefore, the solution of the forward rate system (2.4) can be rewritten as:

$$d\hat{r}_t = \overbrace{[\mu(\hat{r}_t) - \frac{1}{2}\frac{\partial \hat{\sigma}}{\partial \hat{r}}(\hat{r}_t)\hat{\sigma}(\hat{r}_t)]}^{\hat{\mu}(\hat{r}_t)} dt + \hat{\sigma}(\hat{r}_t) \circ dW_t, \tag{2.7}$$

where: $\hat{\mu} : \hat{\mathcal{H}} \longrightarrow \hat{\mathcal{H}}$ is given by:

$$\hat{\mu}(\hat{r}_t) = \begin{pmatrix} \mathbf{F}r^0 + \sigma^0(\hat{r}_t)\mathbf{H}\sigma^0(\hat{r}_t) \\ \mathbf{F}r^1 + \sigma^1(\hat{r}_t)\mathbf{H}\sigma^1(\hat{r}_t) - \sigma^1(\hat{r}_t)\beta^{1*} \\ \vdots \\ \mathbf{F}r^m + \sigma^m(\hat{r})\mathbf{H}\sigma^m(\hat{r}) - \sigma^m(\hat{r}_t)\beta^{m*} \\ \mathbf{B}r^0 - \mathbf{B}r^1 - \frac{1}{2}||\beta^1||^2 \\ \vdots \\ \mathbf{B}r^0 - \mathbf{B}r^m - \frac{1}{2}||\beta^m||^2 \end{pmatrix} - \frac{1}{2}\frac{\partial \hat{\sigma}}{\partial \hat{r}}(\hat{r}_t) \begin{pmatrix} \sigma^0(\hat{r}_t) \\ \sigma^1(\hat{r}_t) \\ \vdots \\ \sigma^m(\hat{r}_t) \\ \beta^1(\hat{r}_t) \\ \vdots \\ \beta^m(\hat{r}_t) \end{pmatrix}, \tag{2.8}$$

where

$$\frac{\partial \hat{\sigma}}{\partial \hat{r}}(\hat{r}_t) = \begin{pmatrix} \frac{\partial \sigma^0}{\partial r^0}(\hat{r}_t) & \dots & \frac{\partial \sigma^0}{\partial r^m}(\hat{r}_t) & \frac{\partial \sigma^0}{\partial Y^1}(\hat{r}_t) & \dots & \frac{\partial \sigma^0}{\partial Y^m}(\hat{r}_t) \\ \frac{\partial \sigma^1}{\partial r^0}(\hat{r}_t) & \dots & \frac{\partial \sigma^1}{\partial r^m}(\hat{r}_t) & \frac{\partial \sigma^1}{\partial Y^1}(\hat{r}_t) & \dots & \frac{\partial \sigma^1}{\partial Y^m}(\hat{r}_t) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial \sigma^m}{\partial r^0}(\hat{r}_t) & \dots & \frac{\partial \sigma^m}{\partial r^m}(\hat{r}_t) & \frac{\partial \sigma^m}{\partial Y^1}(\hat{r}_t) & \dots & \frac{\partial \sigma^m}{\partial Y^m}(\hat{r}_t) \\ \frac{\partial \beta^1}{\partial r^0}(\hat{r}_t) & \dots & \frac{\partial \beta^1}{\partial r^m}(\hat{r}_t) & \frac{\partial \beta^1}{\partial Y^1}(\hat{r}_t) & \dots & \frac{\partial \beta^1}{\partial Y^m}(\hat{r}_t) \\ \vdots & & \vdots & \vdots & & \vdots \\ \frac{\partial \beta^m}{\partial r^0}(\hat{r}_t) & \dots & \frac{\partial \beta^m}{\partial r^m}(\hat{r}_t) & \frac{\partial \beta^m}{\partial Y^1}(\hat{r}_t) & \dots & \frac{\partial \beta^m}{\partial Y^m}(\hat{r}_t) \end{pmatrix}.$$

**Remark 2.1.6.** *We observe that, since*

$$\hat{\sigma} = (\hat{\sigma}_1, \dots, \hat{\sigma}_d) : \hat{\mathcal{H}} \longrightarrow \hat{\mathcal{H}}^d,$$

*and*

$$\hat{\mu} : \hat{\mathcal{H}} \longrightarrow \hat{\mathcal{H}},$$

*are smooth mappings by Assumption 2.1.3, we can interpret $\hat{\mu}$, $\hat{\sigma}_j$ for each $j \in \{1, \dots, m\}$ (locally) as vector fields defined on the Banach space $\hat{\mathcal{H}}$.*

## 2.2 The consistency problem

In this section, we aim at providing a description of the property of consistency and a general characterization of the consistency between a model $\mathcal{M}$ and a parameterized $\mathcal{G}$. We generalize the results provided by Björk and Christensen in [2] to the multi-curve context.

Suppose that we have specified:

- A volatility $\hat{\sigma}$. In this sense we are representing an interest rate model $\mathcal{M}$, described by the SDE system (2.4).
- A mapping $G$, which determines a forward rate curve manifold $\mathcal{G} \subset \hat{\mathcal{H}}$.

In particular, in order to obtain a submanifold determined by $G$ we have to assume that:

**Assumption 2.2.1.**

$$G : \mathcal{Z} \longrightarrow \hat{\mathcal{H}}, \qquad \mathcal{Z} \subset \mathbb{R}^n \tag{2.9}$$

*is an injective function such that the differential of $G$ (in the sense of Definition B.1.8):*

$$dG|_z : \mathbb{R}^n \longrightarrow \hat{\mathcal{H}},$$

*for each $z \in \mathcal{Z}$.*

*For simplicity, we will use the following notation for the differential of a function: $dG|_z = G_z(z)$.*

Recalling Example B.1.11, the previous assumption allows to obtain that $G$ is an immersion. In particular, $\mathcal{G} := Im[G]$ is a submanifold of $\hat{\mathcal{H}}$.

The consistency problem consists in finding conditions under which a model $\mathcal{M}$ and a submanifold $\mathcal{G}$ are consistent in the sense described by the following definition:

**Definition 2.2.2.** *Given a forward rate dynamics, as* (2.4), *describing a model $\mathcal{M}$ and a family of forward rate curves, described by a submanifold $\mathcal{G} \subset \hat{\mathcal{H}}$, we say that the couple $(\mathcal{M}, \mathcal{G})$ is locally invariant under the action of $\hat{r}$ (solution of* (2.4)*) if for each $(r_s, s) \in \mathcal{G} \times \mathbb{R}_+$ there exists $\tau : \mathcal{G} \times \mathbb{R}_+ \longrightarrow \mathbb{R}_+$, stopping time, such that:*

$$\tau(r_s, s) > s, \qquad \mathcal{Q} - a.s.; \tag{2.10}$$

$$r_t \in \mathcal{G}, \qquad \text{for each } t \in [s, \tau(s, r_s)). \tag{2.11}$$

*If $\tau(s, r_s) = \infty$, for each $(r_s, s) \in \mathcal{G}$, $\mathcal{Q} - a.s.$ we say that the couple $(\mathcal{M}, \mathcal{G})$, is globally invariant.*

In order to prove a characterization of the previous definition in terms of the vector fields $\hat{\mu}(\hat{r})$, $\hat{\sigma}(\hat{r})$ and the mapping $G$, we give the following definition:

**Definition 2.2.3.** *We say that $\mathcal{G}$ is locally $\hat{r}$-invariant under the action of the forward rate process $\hat{r}$ if for each $\hat{r}_0 \in \mathcal{G}$ there exists a $\mathcal{Q}$-a.s. strictly positive* *stopping time $\tau(\hat{r}_0)$ and a stochastic process $(Z_t)_t$ taking values in $\mathbb{R}^n$, which has a Stratonovich differential of the form:*

$$dZ_t = a(Z_t)dt + b(Z_t) \circ W_t, \tag{2.12}$$

*such that for each $t \in [0, \tau(\hat{r_0}))$, $\quad r_t(x) = G(x, Z_t)$ for each $x$ $\mathcal{Q}$-a.s., where $G$ is assumed to be an immersion on $\hat{\mathcal{H}}$, such that $\mathcal{G} = Im[G]$.*

In what follows, we will prove local results, then we will use the term invariant or $\hat{r}$-invariant in order to denote the local invariance and the $\hat{r}$-local invariance respectively.

We now prove that, under the conditions given for $G$, the previous two definitions are equivalent. To this effect, we need classical results of functional analysis (see [6] and [18]).

**Proposition 2.2.4** (Local left inverse)**.** *Consider a mapping $g : \mathcal{X} \longrightarrow \mathcal{Y}$, where $\mathcal{X}$, $\mathcal{Y}$ are two Banach spaces. Let $h_0 \in \mathcal{X}$ and suppose that*

1. *$g$ is a differentiable function, with Fréchet derivative denoted by $\frac{\partial}{\partial h} g$;*
2. *the linear map $\frac{\partial}{\partial h} g$ is injective;*
3. *there exists a bounded left inverse of $\frac{\partial}{\partial h} g$, denoted with $A$ at the point $h_0$; in particular:*
$$A \frac{\partial}{\partial h} g \Big|_{h_0} = id_{\mathcal{X}},$$
*where $id_{\mathcal{X}}$ is the identity map on the Banach space $\mathcal{X}$.*

*Then:*

*There exists two open subset $U \subset \mathcal{X}$ and $W \subset \mathcal{Y}$, which respectively contain $h_0$ and $g(h_0)$ and a function $f : W \longrightarrow U$ such that $f(g(x)) = x$, for each $x \in W$.*

*Proof.* Define $\varphi : \mathcal{X} \longrightarrow \mathcal{X}$ by $\varphi(x) := Ag(x)$, then $\frac{\partial}{\partial h}\varphi(x) = A\frac{\partial}{\partial h} g(x) = id_{\mathcal{X}}$, (it is linear and bounded). Then, by the inverse function theorem there exists $U \subset \mathcal{X}$ open and a function $\psi_0 : U \longrightarrow U$ such that $\psi_0(\varphi(x)) = x, \quad x \in U$.

Then, we define $W := \varphi(U)$ and the function:

$$f : \begin{array}{ccc} W & \longrightarrow & U \\ y & \longrightarrow & \psi_0(Ay) \end{array}.$$

In particular, for each $x \in U$: $f(g(x)) = \psi_0(A(g(x))) = \psi_0(\varphi(x)) = x$. $\square$

We now need to show the regularity of the inverse function described in Proposition 2.2.4:

**Lemma 2.2.5.** *Let $\Psi : \mathcal{X} \longrightarrow \mathcal{Y}$ be a bounded injective linear mapping between two Banach spaces $\mathcal{X}, \mathcal{Y}$ with closed range.*

*If we denote with $\Psi^*$ the adjoint mapping, then the linear mapping:*

$$H_\Psi := (\Psi^* \Psi)^{-1} \Psi^*$$

*is a bounded left inverse of $\Psi$. Moreover, the operator $\Psi \longrightarrow H_\Psi$ is infinitely differentiable in the norm operator.*

*Proof.* $\Psi$ is injective with closed range, then $(\Psi^* \Psi)$ is invertible.

If we consider $y = \Psi x$, then:

$$H_\Psi y = (\Psi^* \Psi)^{-1} \Psi^* y = (\Psi^* \Psi)^{-1} (\Psi^* \Psi) x = x.$$

The smoothness descends from the fact that $\Psi \longrightarrow \Psi^*$ and $A \longrightarrow A^{-1}$ are smooth operators. □

**Remark 2.2.6.** *We can apply Proposition 2.2.4 to a function which satisfies the boundary condition on the local inverse of the Fréchet derivative. Therefore, if we want to apply this result to a function $G$ which satisfies Assumption 2.2.1, we have to assume that the local inverse of $G'$ is bounded.*

In the following proposition we will prove the equivalence between the concepts of equivalence and $\hat{r}$-equivalence under the Assumption 2.1.3 Assumption 2.2.1.

**Proposition 2.2.7.** *Let us consider a model $\mathcal{M}$, determined by* (2.4) *whose parameters satisfy Assumption 2.1.3 and a parameterized family $\mathcal{G} \subset \hat{\mathcal{H}}$, described as the image of a mapping $G$ which satisfies Assumption* (2.2.1).

*Then the couple $(\mathcal{M}, \mathcal{G})$ is invariant in the sense of Definition 2.2.2 if and only if $\mathcal{G}$ is $\hat{r}$-invariant, in the sense of Definition 2.2.3.*

*Proof.* $\hat{r}$-invariance $\Rightarrow$ invariance: It follows directly from the definitions.

invariance $\Rightarrow$ $\hat{r}$-invariance:
for an arbitrary fixed $\hat{r}_0 \in \mathcal{G}$, thanks to the hypothesis on $G : \mathcal{Z} \longrightarrow \mathcal{G} \subset \hat{\mathcal{H}}$, we have that: $\hat{r}_0 = G(z_0)$, for a unique $z_0 \in \mathcal{Z}$.

Moreover, $\frac{\partial}{\partial z} G(z_0)$ is injective, then it has left inverse, denoted by $\Psi(\hat{r}_0)$. We can also note that the left inverse $\Psi : \hat{\mathcal{H}} \longrightarrow \mathbb{R}^d$. Since the codomain is finite-dimensional, the mapping $\Psi(\hat{r}_0)$ is not only linear but also bounded.

We have thus shown that $G$ satisfies the hypotheses of Proposition 2.2.4. Therefore, $G$ has local left inverse, denoted by $F : U \longrightarrow W$ ($U$, $W$ are defined as the open subsets introduced in the proof of Proposition 2.2.4).

Let us define:

$$Z_t = F(\hat{r}_t),$$

around $\hat{r}_0 \in U$. Computing the Stratonovich dynamics on $Z$, we obtain:

$$dZ_t = \frac{\partial F}{\partial \hat{r}_t}(\hat{r}_t)\hat{\mu}(\hat{r}_t)dt + \frac{\partial F}{\partial \hat{r}_t}(\hat{r}_t)\hat{\sigma}(\hat{r}_t) \circ dW_t. \tag{2.13}$$

Thus, $(Z_t)_t$ is the solution of a finite dimensional system of SDEs like in (2.12), where:

$$a(z) = \frac{\partial F}{\partial \hat{r}}(G(z))\hat{\mu}(G(z)), \tag{2.14}$$

$$b(z) = \frac{\partial F}{\partial \hat{r}}(G(z))\hat{\sigma}(G(z)). \tag{2.15}$$

By construction, $F(\hat{r}_t) = Z_t$ and since $G$ is the local inverse of $F$ around $\hat{r}_0$, the following equation holds:

$$G(Z_t) = G(F(\hat{r}_t)) = \hat{r}_t.$$

□

We can prove now the central result of this section:

**Theorem 2.2.8** (Invariance). *If we consider the forward curve manifold $\mathcal{G} = Im[G]$ and the model $\mathcal{M}$, the couple $(\mathcal{M}, \mathcal{G})$ is invariant if and only if the following conditions hold:*

$$\hat{\mu}(G(z)) \in Im[G_z(z)] \equiv T_{G(z)}\mathcal{G}; \tag{2.16}$$

$$\hat{\sigma}_j(G(z)) \in Im[G_z(z)] \equiv T_{G(z)}\mathcal{G}, \qquad \forall j \in \{1, \dots, d\} \tag{2.17}$$

*where $\hat{r} = G(z)$, for each $z \in \mathcal{Z}$ domain of definition of $G$.*

*Proof.* $(\Rightarrow)$ We exploit the equivalence between $\hat{r}$-invariance and invariance, proved in Proposition 2.2.7.

By Itô's formula (with the correction term given by Stratonovich):

$$\begin{cases} d\hat{r}_t = G_z(Z_t)a(Z_t)dt + G_z(Z_t)b(Z_t) \circ dW_t \\ \hat{r}_0 = G(Z_0), \end{cases}$$

where $\hat{r}_0$ is chosen arbitrarily in $\mathcal{G}$.

Then, recalling that $\hat{r}$ satisfies (2.4) and equating the corresponding terms we obtain:

$$\hat{\mu}(\hat{r}_t) = G(Z_t)_z a(Z_t), \tag{2.18}$$

$$\hat{\sigma}(\hat{r}_t) = G(Z_t)_z b(Z_t); \tag{2.19}$$

these conditions are equivalent to: $\hat{\mu}(\hat{r}_t),\ \hat{\sigma}_j(\hat{r}_t) \in Im[G_z(Z_t)]$,
for each $j \in \{1,\dots,d\}$.

($\Leftarrow$) Let us suppose that $\hat{\mu}(\hat{r}_t),\ \hat{\sigma}(\hat{r}_t) \in Im[G_z(Z_t)]$. This means that there exists two vector fields $a(z),\ b(z) \in \mathbb{R}^n$, defined on the open subset $\mathcal{Z}$, such that:

$$\hat{\mu}(G(z)) = G_z(Z_t)a(Z_t), \tag{2.20}$$

$$\hat{\sigma}(G(z)) = G_z(Z_t)b(Z_t). \tag{2.21}$$

From the injectiveness of $G_z(z)$ $a(z), b(z)$ are uniquely determined.

Since $\mathbb{R}^n$ is finite-dimensional, $G_z(z)$ has closed range, then by Assumption 2.1.3 we can apply Lemma 2.2.5 to $G$.

Therefore, choosing an arbitrary point $z_0 \in \mathcal{Z}$ and denoting by $H$ the local inverse of $dG(z)$ around $z_0$ ($H : T\mathcal{G}|_{G(U)} \longrightarrow \mathbb{R}^n$, where $U$ is an open subset of $\mathcal{Z}$ containing $z_0$), we have that $H$ is smooth. This implies that:

$$a(z) = H(G(z))\hat{\mu}(G(z)), \tag{2.22}$$

$$b(z) = H(G(z))\hat{\sigma}(G(z)) \tag{2.23}$$

are smooth too.

Since $a(z),\ b_j(z)$, for each $j \in \{1,\dots,d\}$ are smooth vector fields defined on $U$, they are locally Lipschitz. This condition allows us to define a process $(Z_t)_t$ as the unique strong solution of the equation:

$$\begin{cases} dZ_t = a(Z_t)dt + b(Z_t) \circ dW_t \\ Z_0 = z_0. \end{cases}$$

Given the initial point $z_0$ the solution of the previous SDE is local on $U$. A priori, there could be no global solution of hte previous SDE on $\mathcal{Z}$.

Finally, we define the process $(y_t)_t \subset \hat{\mathcal{H}}$, as $y_t = G(Z_t)$ which satisfies the dynamics:

$$\begin{cases} dy_t = G_z(Z_t)a(Z_t)dt + G_z(Z_t)b(Z_t) \circ dW_t, \\ y_0 = G(z_0). \end{cases}$$

We can observe that $y_0 = \hat{r}_0 = G(z_0)$ and both the process $(y_t)_{t\geq 0}$ and $(\hat{r}_t)_{t\geq 0}$ solves the same SDE.

By the uniqueness of strong solution of SDEs, we conclude that $y_t = \hat{r}_t$. Since $\mathcal{G}$ is locally $\hat{r}$-invariant, then we can apply Proposition 2.2.7 in order to say that the couple $(\mathcal{M}, \mathcal{G})$ is locally invariant, which is the thesis. $\square$

The previous result is basically equivalent to Proposition 4.2 of [2], with a slight different notation, due to the multi-curve approach that we are developing.

The main change is due to the fact that $\hat{\mu}(\hat{r}),\ \hat{\sigma}(\hat{r})$ are vector fields defined on $\hat{\mathcal{H}}$, which is a product of Banach spaces. This fact implies that it is possible to determine relations between the components of the Fréchet derivative of the function $G$ computed on the vector fields $a(z)$ and $b(z)$, which guarantee that the couple $(\mathcal{M}, \mathcal{G})$ is invariant.

Using the notation:

$$G_z := \left(G_z^0, G_z^1, \ldots, G_z^m, G_z^{m+1}, \ldots, G_z^{2m}\right)^*, \tag{2.24}$$

condition (2.16) can be rewritten, emphasizing the relations among the different components. We obtain that:

$$G_z^0 a(z) = \mathbf{F}G^0(z) + \sigma^0(G(z))\mathbf{H}\sigma^0(G(z)) - \frac{1}{2}\frac{\partial\sigma^0}{\partial\hat{r}}(G(z))\hat{\sigma}(G(z)), \tag{2.25}$$

for $j \in \{1, \ldots, m\}$:

$$G_z^j a(z) = \mathbf{F}G^j(z) + \sigma^j(G(z))\mathbf{H}\sigma^j(G(z)) - \beta^j(G(z))\sigma^{j*}(G(z)) - \frac{1}{2}\frac{\partial\sigma^j}{\partial\hat{r}}(G(z))\hat{\sigma}(G(z)), \tag{2.26}$$

for $j \in \{m+1, \ldots, 2m\}$:

$$G_z^j a(z) = \mathbf{B}G^0(z) - \mathbf{B}G^{j-m}(z) - \frac{1}{2}||\beta^{j-m}(G(z))||^2 - \frac{1}{2}\frac{\partial\beta^{j-m}}{\partial\hat{r}}(G(z))\hat{\sigma}(G(z)). \tag{2.27}$$

The condition on the volatility term is:

$$\begin{cases} G_z^j b(z) = \sigma^j(G(z)), \quad j \in \{0, \ldots, m\}; \\ G_z^j b(z) = \beta^{j-m}(G(z)), \quad j \in \{m+1, \ldots, 2m\}. \end{cases} \tag{2.28}$$

Substituting conditions (2.28) in the conditions on the drift equation (2.25) becomes:

$$G_z^0 a(z) = \mathbf{F}G^0(z) + G_z^0 b(z)\mathbf{H}G_z^0 b(z) - \frac{1}{2}\frac{\partial\sigma^0}{\partial\hat{r}}(G(z))\hat{\sigma}(G(z)),$$

which can be rewritten as:

$$\mathbf{F}G^0(z) = G_z^0(z)[a(z) - b(z)\mathbf{H}G_z^0 b(z)] - \frac{1}{2}\frac{\partial\sigma^0}{\partial\hat{r}}(G(z))\hat{\sigma}(G(z)). \tag{2.29}$$

Equations (2.26) can be reinterpreted as follows:

$$\begin{aligned} G_z^j(a(z)) =& \mathbf{F}G^j(z) + [G_z^j b(z)]\mathbf{H}[G_z^j b(z)] + \\ & - G_z^{j+m} b(z)[G_z^j b(z)]^* - \frac{1}{2}\frac{\partial\sigma^j}{\partial\hat{r}}(G(z))\hat{\sigma}(G(z)). \end{aligned} \tag{2.30}$$

Recall that the following equivalence hold:

$$\begin{aligned} G_z^j b(z)\mathbf{H}[G_z^j b(z)](x) &= G_z^j b(z)(x)\int_0^x [G_z^j b(z)]^*(s)ds \\ &= \frac{1}{2}\frac{\partial}{\partial x}\Big|\Big|\int_0^x [G_z^j b(z)](s)ds\Big|\Big|^2 \end{aligned}$$

Therefore it holds that:

$$\begin{aligned} G_z^j a(z) =& \mathbf{F}G^j(z) + \frac{1}{2}\mathbf{F}\Big|\Big|\int_0^x [G_z^j b(z)](s)ds\Big|\Big|^2 + \\ & - G_z^{j+m} b(z)[G_z^j b(z)]^* - \frac{1}{2}\frac{\partial \sigma^j}{\partial \hat{r}}(G(z))\hat{\sigma}(G(z)), \end{aligned}$$

In conclusion, exploiting the linearity of $\mathbf{F}$:

$$\begin{aligned} G_z^j\{a(z) + b(z)[G_z^{j+m} b(z)]^*\} = \mathbf{F}\Big[G^j(z) + \frac{1}{2}\Big|\Big|\int_0^x [G_z^j b(z)](s)ds\Big|\Big|^2\Big] + \\ - \frac{1}{2}\frac{\partial \sigma^j}{\partial \hat{r}}(G(z))\hat{\sigma}(G(z)), \end{aligned} \tag{2.31}$$

where $j \in \{1, \dots, m\}$.

Finally, for equations (2.27), we obtain the following equation:

$$G_z^j a(z) = \mathbf{B}G^0(z) - \mathbf{B}G^{j-m}(z) - \frac{1}{2}||G_z^j b(z)||^2 - \frac{1}{2}\frac{\partial \beta^{j-m}}{\partial \hat{r}}(G(z))\hat{\sigma}(G(z)), \tag{2.32}$$

where $j \in \{m+1, \dots, 2m\}$.

We can conclude that the family $\mathcal{G}$ has to satisfy the conditions imposed by equation (2.31) and (2.32), which represent the relations between the components of the forward rate equation satisfied by $\hat{r}$.

In the following remark we try to analyse if it is possible to divide the components of the solution of (1.33), in particular if there exists conditions under which we can check the consistency conditions only on the coordinates associated with a forward rate equation and which automatically guarantee those conditions on the components associated with the log-spread components.

**Remark 2.2.9.** *Looking at Definition 2.2.3 one could think that the existence of a process $Z_t$ and a mapping $G$ which guarantee the $\hat{r}$-invariance conditions, crucial in the introduction of the concept of consistency, has to be texted only for the equations of the system* (1.33) *associated with the forward rate $r^i$ for each $i = 1, \dots, d$, since only those components are infinite-dimensional. This problem can formalized as follows:*

*If there exists a finite-dimensional process $Z_t$ and a set of functions $G^i$ for $j = 0, \dots, m$ such that: $r_t^i(x) = G^i(x, Z_t)$ for each $i$, then we have the consistency condition:*

$$\hat{r}_t(x) = \hat{G}(x, \hat{Z}_t), \qquad \text{where } \hat{Z}_t = (Z_t, Y_t),$$

*and the process $Y_t = (Y_t^1, \dots, Y_t^m)$ is the log-spot spread process.*

*Such a condition allows us exploit the fact that the log-spot processes $Y_t^i$ are finite-dimensional in order to consider them as a part of the finite-dimensional process $Z_t$. Consequently, we can check the consistency conditions only on the infinite-dimensional equations of the system* (1.33)*, which solutions are $r^i$ for every $j = 0, \dots, m$.*

*Unfortunately, in the general framework this property does not hold. Indeed, analysing the conditions* (2.25) *and* (2.26) *it is possible to note that the components of the volatility term $\sigma^i$ and $\beta^i$ depend on the entire forward structure defined on $\hat{\mathcal{H}}$. This fact implies that, for each $j = 0, \dots, m$, the condition on $G_z^i a(z)$ is dependent on the entire function $G$, which is defined on $\hat{\mathcal{H}}$ and therefore, it is dependent also on the last $m$ components of $G$, associated with the log-spot spread processes. In particular, by the previous consideration we conclude that in order to check the consistency condition for a couple $(\mathcal{M}, \mathcal{G})$ we have to describe all the $2m+1$ components of $\mathcal{G}$ and text the conditions* (2.16) (2.17) *also for the last $m$ component. Since in the pre-crisis environment the spread processes were not defined, we do not have a family of functions which is used in the literature to parameterize the spreads (different from the forward rate curves associated with each tenor, for which several parameterized forward families have been introduced, for example the family of Nelson-Siegel or Svensson, which will be described in the next section). As a consequence of this fact, we try to find conditions for the components of $\mathcal{G}$ related to the log-spreads, in order to guarantee the consistency. In order to do this, we consider the function $G$ of Definition 2.2.3 and we assume that the function $\widetilde{G} = (G^0, \dots, G^m)^*$ is injective and its Fréchet derivative is injective too. Moreover, according to what we observed at the beginning of the remark, we suppose that the volatility term $\hat{\sigma}(\hat{r})$ does not depend on the log-spread processes, but it is only a function of $\widetilde{r}_t = (r_t^0, \dots, r_t^m)^*$. Under this assumption, which does not allow to consider very complex models, but it is respected by the models we will describe in Section 2.3, we can invert the conditions* (2.25) *and* (2.26)*:*

$$\widetilde{G}_z a(z) = \begin{pmatrix} \boldsymbol{F}G^0(z) + \sigma^0(\widetilde{G})\boldsymbol{H}\sigma^0(\widetilde{G}) - \frac{1}{2}\frac{\partial \sigma^0}{\partial \widetilde{r}}(\widetilde{G}(z))(\sigma^0(\widetilde{G}(z)), \dots, \sigma^m(\widetilde{G}(z)))^* \\ \boldsymbol{F}G^1(z) + \sigma^1(\widetilde{G})\boldsymbol{H}\sigma^1(\widetilde{G}) - \frac{1}{2}\frac{\partial \sigma^1}{\partial \widetilde{r}}(\widetilde{G}(z))(\sigma^0(\widetilde{G}(z)), \dots, \sigma^m(\widetilde{G}(z)))^* - \beta^1(\widetilde{G}(z))\sigma^{1*}(\widetilde{G}(z)) \\ \vdots \\ \boldsymbol{F}G^m(z) + \sigma^m(\widetilde{G})\boldsymbol{H}\sigma^m(\widetilde{G}) - \frac{1}{2}\frac{\partial \sigma^m}{\partial \widetilde{r}}(\widetilde{G}(z))(\sigma^0(\widetilde{G}(z)), \dots, \sigma^m(\widetilde{G}(z)))^* - \beta^m(\widetilde{G}(z))\sigma^{m*}(\widetilde{G}(z)) \end{pmatrix},$$

*After this computation we can determine the vector $a(z)$ and then we can use it in order to provide conditions on the differential $G_z^i$ for each $i = m+1, \dots, 2m$ such that the condition* (2.27) *is satisfied. Through this procedure, we determine the conditions on the functions which define the log-spot spreads which respect the consistency.*

## 2.3 Examples

In this section we shall use the Theorem 2.2.8 to determine if classical models such as the Ho-Lee model (1986) or, for instance, the Hull-White model (1990) and classical parameterized forward curves manifolds, such as the Nelson-Siegel family or the Svensson family (or their modifications) are consistent.

We proceed by generalizing the results obtained in [2] to the multi-curve framework. In order to do this, we consider a forward rate model $\mathcal{M}$, defined on the Banach space $\hat{\mathcal{H}}$. It will be determined by a system of SDEs, in which each component is described by a well known dynamics (for instance, the Ho-Lee or the Hull-White). On the other hand, we will introduce a vector forward parameterized family, denoted by $\mathcal{G}$, whose components are described by forward parameterized families such as the Nelson-Siegel or the Svensson family. We will provide explicit conditions for the consistency of the couple $(\mathcal{M}, \mathcal{G})$.

We will first consider the same forward rate family for each component and the same model for each component of $\mathcal{M}$. Afterwards, we will describe a model in which the first component (associated with the OIS curve) will be equipped with a richer structure than the components associated with the LIBOR forward rate.

In analogy to [2], we introduce a forward parameterized family, frequently used in literature, the Nelson-Siegel family (in the following, we denote it with $NS$).

### 2.3.1 The Nelson-Siegel family

The $NS$ forward curve manifold $\mathcal{G}$ was described for the first time in [17]. It is parameterized by $z \in \mathcal{Z} := \mathbb{R}^4$, through the mapping $G$, defined in the following way:

$$G(z, x) = z_1 + z_2 e^{-z_4 x} + z_3 x e^{-z_4 x} = z_1 + e^{-z_4 x}[z_2 + z_3 x]. \tag{2.33}$$

For a detailed description of this family we recall [13]. If we want to consider $G$ as a function defined on $\mathbb{R}$, and taking values on $\mathcal{H}_\gamma$, we need to suppose that: $z_4 > -\frac{\gamma}{2}$.

We consider now the Fréchet derivative of $G$:

- if $z_4 \neq 0$:

$$\frac{\partial G}{\partial z}(z, x) = \begin{pmatrix} 1 & e^{-z_4 x} & x e^{-z_4 x} & -x e^{-z_4 x}(z_2 + z_3 x) \end{pmatrix}. \tag{2.34}$$

- If $z_4 = 0$, the family is described by the mapping $G = z_1 + z_2 + x z_3$. The term $z_2$ is redundant, so that we impose that $z_2 = 0$ and $G$ becomes

$$G(z, x) = z_1 + z_3 x, \tag{2.35}$$

where $z = (z_1, z_3)$. In this case the Fréchet derivative of $G$ is:

$$\frac{\partial G}{\partial z}(z,x) = \begin{pmatrix} 1 & x \end{pmatrix}. \tag{2.36}$$

If $z_4 = 0$ the family $\mathcal{G} := Im[G]$ is called *degenerated NS family.*

We consider a NS family for each component of the multi-curve family. The parameters describing each row of this family are supposed to be independent row by row. Therefore, we have to consider a vector of parameters:

$$z = (z_1^0, \dots z_4^0, z_1^1, \dots, \dots, z_4^m) \tag{2.37}$$

Then, the first $m+1$ rows of the mapping $G$ are defined by:

$$G(z,x) = \begin{pmatrix} z_1^0 + z_2^0 e^{-z_4^0 x} + x z_3^0 e^{-z_4^0 x} \\ z_1^1 + z_2^1 e^{-z_4^1 x} + x z_3^1 e^{-z_4^1 x} \\ \dots \\ z_1^m + z_2^m e^{-z_4^m x} + x z_3^m e^{-z_4^m x} \end{pmatrix} \tag{2.38}$$

We can determine the Fréchet derivative of $G$, defined by the matrix:

$$\begin{pmatrix} 1 & [e^{-z_4 x}] & [xe^{-z_4 x}] & [-xe^{-z_4 x}(z_2 + z_3 x)] & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 1 & [e^{-z_4 x}] & [xe^{-z_4 x}] & [-xe^{-z_4 x}(z_2 + z_3 x)] & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \end{pmatrix}$$

For the degenerated case the Fréchet derivative is given by:

$$\frac{\partial G}{\partial z}(z,x) = \begin{pmatrix} 1 & x & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & x & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & x \end{pmatrix}$$

The Nelson-Siegel family is the main forward rate family analyzed by Björk and Christensen in [2]. In particular, the consistency is checked in relation with two models.

We briefly describe them in the following paragraphs.

### 2.3.2 The Ho-Lee model

The Ho-Lee model (in the following denoted by HL) is a short rate model, developed in 1986 in [15]. It is described by the following SDE:

$$dr_t = \Theta(t)dt + \sigma dW_t, \tag{2.39}$$

where $r_t$ is the short rate, $\sigma > 0$ is constant and the function $\Theta(t)$ is satisfying the conditions which guarantee the existence and uniqueness of strong solution.

We can derive the dynamics of the associated forward rate $f_t(T)$. The function $\Theta(t)$ will be determined by the HJM-drift condition. In order to compute $f_t(T)$, we introduce the following general dynamics:

$$df_t(T) = \alpha_t(T)dt + \sigma_t(T)dW_t, \tag{2.40}$$

where the drift term is supposed to respect the HJM-drift condition (A.18).

Let us consider constant volatility term: $\sigma_t(T) \equiv \sigma > 0$, then:

$$\begin{aligned} f_t(T) &= f_0(T) + \int_0^t \alpha_s(T)ds + \int_0^t \sigma dW_s \\ &= f_0(T) + \sigma^2 \int_0^t (T-s)ds + \sigma W_t = f_0(T) + \sigma^2 t\Big(T - \frac{t}{2}\Big) + \sigma W_t. \end{aligned}$$

If we compute the short rate associated with this forward rate, we obtain:

$$r_t = f_t(t) = f_0(t) + \sigma^2 \frac{t^2}{2} + \sigma W_t.$$

In conclusion, the dynamics of $r_t$ is described by:

$$dr_t = \Big(\frac{\partial}{\partial T} f_0(t) + \sigma^2 t\Big)dt + \sigma dW_t.$$

The previous dynamics corresponds to (2.39) with $\Theta(t) = \frac{\partial}{\partial T} f_0(t) + \sigma^2 t$.

The forward rate equation associated with the HL model is:

$$dr_t(x) = \Big[\sigma^2 t\Big(x + \frac{t}{2}\Big) + \frac{\partial r_0(x)}{\partial x} + \sigma^2 t\Big]dt + \sigma dW_t, \tag{2.41}$$

where $r_0(x) = f_0(x)$ for each $x \in \mathbb{R}_+$.

### 2.3.3 The Ho-Lee model and the Nelson-Siegel family

Similarly as in the single-curve approach the Ho-Lee model and the Nelson Siegel family are inconsistent. We focus only on the first $m+1$ rows of (2.7). In order to prove the inconsistency of the couple $(\mathcal{M}, \mathcal{G})$, where $\mathcal{M}$ is the model determined by a constant volatility term $\sigma > 0$ and $\mathcal{G} = Im[G]$, where $G$ is the mapping describing the NS family, we need to check the conditions of Theorem 2.2.8. Moreover, in analogy to the fact that $\sigma$ is constant, we will also assume that $\beta$, the volatility term of log spot spread process, is constant. The consistency condition is equivalent to:

$$\hat{\mu}^j(G(z)), \hat{\sigma}_i^j(G(z)) \in T_{G(z)}\mathcal{G}, \quad \forall\, j \in \{0, \dots, m\}, \quad \forall\, i \in \{1, \dots, d\}.$$

If we try to check this condition for $j = 0$ on the drift term, we obtain the following result.

First of all, we can make the following observation:

$$\frac{\partial}{\partial \hat{r}} \sigma^0(\hat{r}) = 0, \text{ (since } \sigma^0 \text{ is constant).}$$

Adopting the notation $\mathcal{G} = (\mathcal{G}_0, \ldots, \mathcal{G}_m, \mathcal{G}_{m+1}, \ldots, \mathcal{G}_{2m})^*$, the previous condition implies that:

$$\hat{\mu}(G(z))(x) = \mathbf{F}G^0(z)(x) + \sigma^0 \mathbf{H} \sigma^0 \in T_{G^0(z)} \mathcal{G}_0. \tag{2.42}$$

This is equivalent to the existence of a vector $\eta = (\eta_1, \ldots, \eta_4)$ such that, for every $x \in \mathbb{R}_+$ and for $z \in \mathcal{Z}$:

$$\frac{\partial}{\partial x} G^0(z,x) + \sigma^0 \int_0^x \sigma^0 ds = \eta_1 + \eta_2 e^{-z_4^0 x} + \eta_3 x e^{-z_4^0 x} - x e^{-z_4^0 x}(z_2^0 + z_3^0 x)\eta_4$$

$$e^{-z_4^0 x}(-z_4^0 z_2^0 - z_4^0 z_3^0 x + z_3) + (\sigma^0)^2 x = \eta_1 + e^{-z_4^0 x}(\eta_2 + x(\eta_3 - z_2^0 \eta_4) - x^2 z_3^0 \eta_4)$$

However, for $x \to +\infty$ and $z_4^0 > 0$, the left member tends to $+\infty$ whereas the right one is constant. Since the first condition does not hold, then we can say that the couple $(\mathcal{M}, \mathcal{G})$ is not consistent.

We can see that, recalling [2][ Proposition 5.3], in the single curve approach the NS degenerated family is consistent with the HL model, indeed:

$$\frac{\partial}{\partial x} G(z,x) + \sigma^2 x = \eta_1 + \eta_2 x \Longrightarrow z_3 + \sigma^2 x = \eta_1 + \eta_2 x.$$

If we aim at checking the consistency of $(\mathcal{M}, \mathcal{G}^2)$ where $\mathcal{G}^2$ is the parameterized family described by the degenerated NS mapping and $\mathcal{M}$ is the Ho-Lee model, we have to prove that: $\hat{\mu}(G(z)) \in T_{G(z)}\mathcal{G}^2$. By simplicity in the following of the dissertation we will assume that every component of the function $G$ is described by an independent set of parameters. In particular, if the coordinate $G^0$ is determined by the parameters $z_1^0, z_2^0$, the $i^{th}$ component $G^i$ will be described by $z_1^i, z_2^i$ which are different from $z_1^0, z_2^0$. For the equations associated with the OIS and LIBOR forward rates, this condition is equivalent to the existence of a vector $\eta = (\eta_1^0, \eta_2^0, \eta_1^1, \eta_2^1, \ldots, \eta_1^m, \eta_2^m)$:

$$\begin{cases} \mathbf{F}G^0(z) + (\sigma^0)^2 x = \eta_1^0 + \eta_2^0 x \\ \mathbf{F}G^1(z) + (\sigma^1)^2 x - \beta^1 \sigma^1 = \eta_1^1 + \eta_2^1 x \\ \vdots \\ \mathbf{F}G^m(z) + (\sigma^m)^2 x - \beta^m \sigma^m = \eta_1^m + \eta_2^m x \end{cases} \tag{2.43}$$

and this is equivalent to:

$$\begin{cases} z_3^0 + (\sigma^0)^2 x = \eta_1^0 + \eta_2^0 x \\ z_3^1 + (\sigma^1)^2 x - \beta^1 \sigma^1 = \eta_1^1 + \eta_2^1 x \\ \vdots \\ z_3^m + (\sigma^m)^2 x - \beta^m \sigma^m = \eta_1^m + \eta_2^m x \end{cases} \tag{2.44}$$

Therefore, choosing:

$$\begin{cases} \eta_2^j = (\sigma^j)^2, \quad \forall\ j = 0, \dots, m, \\ \eta_1^0 = z_3^0, \\ \eta_1^j = z_3^j - \beta^j \sigma^j, \Longrightarrow\ \eta_1^j = z_3^j - \beta^j \sqrt{\eta_2^j}, \quad \forall j = 0, \dots, m \end{cases} \tag{2.45}$$

the equivalence requested holds.

The condition on the volatility term, in the first $m+1$ coordinates is equivalent to the existence of a vector $(\xi_1^0, \xi_2^0, \xi_1^1, \xi_2^1, \dots, \xi_1^m, \xi_2^m)$, such that:

$$\sigma^j = \xi_1^j + \xi_2^j x, \quad \forall\ i \in \{0, \dots, m\}, \tag{2.46}$$

and such a condition can be obtained imposing that:

$$\begin{cases} \xi_1^j = \sigma^j, \\ \xi_2^j = 0, \end{cases} \tag{2.47}$$

for each $j \in \{0, \dots, m\}$. In particular, until now we have introduced $n := 2(m+1)$ parameters.

The condition on the log-spreads (the last $m$ components of the system) is, in this case:

$$G^0(z,0) - G^{j-m}(z,0) - \frac{1}{2}(\beta^{j-m})^2 = G^j(z)_z \eta(z), \quad \forall j = m+1, \dots, 2m. \tag{2.48}$$

Recalling the definition of $G$, and Remark 2.2.9 the previous condition becomes:

$$\begin{aligned} G_z^j(z)\eta(z) &= z_1^0 - z_1^{j-m} - \frac{1}{2}(\beta^{j-m})^2 \\ &= \sum_{h \in \{1,3\}} \sum_{k=0}^{m} G_{z_h^k}^j \eta_h^k = \sum_{k=1}^{m} G_{z_1^k}^j \left( z_3^k - \beta^k \sigma^k \right) + G_{z_1^0}^j z_3^0 + \sum_{k=0}^{m} G_{z_3^k}^j (\sigma^k)^2. \end{aligned} \tag{2.49}$$

By linearity, we can consider for instance:

$$z_1^0 = G_{z_1^0}^j z_3^0 + G_{z_3^0}^j (\sigma^0)^2. \tag{2.50}$$

If we consider $z_3^0 = 0$, the previous equation will lead to: $G^j(z_1^0, z_3^0) = \frac{z_1^0}{(\sigma^0)^2} z_3^0 + c(z_1^0)$, where $c$ is a suitable function on the $z_1^0$ variable. Hence, equation (2.50) becomes:

$$z_1^0 = \Big(\frac{z_3^0}{(\sigma^0)^2} + c'(z_1^0)\Big) z_3^0 + z_1^0 \Rightarrow \frac{z_3^0}{(\sigma^0)^2} + c'(z_1^0) = 0 \to c'(z_1^0) = -\frac{z_3^0}{(\sigma^0)^2}, \quad (2.51)$$

which is impossible, because $c$ is function only of $z_1^0$, by definition. Therefore, it is necessary to introduce other parameters in order to have the consistency. In particular, we consider the following vector:

$$\widetilde{z} = (z, u^1, \ldots, u^m)^* \in \mathbb{R}^{n+m} \equiv \mathbb{R}^{3m+2}, \quad (2.52)$$

where $u^j$ are additional parameters. Since $G^j$ for each $j = 0, \ldots, m$ are the degenerate Nelson Siegel family, then $\frac{\partial}{\partial u^i} G^j = 0$ for each $j$. If we introduce, with an abuse of notation the function $G^j(\widetilde{z}) = G^j(z)$ for $j = 0, \ldots, m$, $\eta(z) = \eta(\widetilde{z})$ and the vector field:

$$\widetilde{\eta}(\widetilde{z}) = (\eta(\widetilde{z}), \eta^{n+1}(\widetilde{z}), \ldots, \eta^{n+m}(\widetilde{z}))^*,$$

we obtain that: $G^j_{\widetilde{z}}(\widetilde{z})\widetilde{\eta}(\widetilde{z}) = G^j_z(z)\eta$ for each $j = 0 \ldots, m$. On the other hand, we define the functions $G^{j+m}(\widetilde{z}) = u^j$ for each $j = 1, \ldots, m$. This implies that:

$$\begin{cases} G^j_{u^h} \equiv 0, & \forall j \neq h + m, \\ G^{h+m}_{u^h} \equiv 1. \end{cases}$$

Moreover, we introduce the last $m$ coordinates of the vector field $\widetilde{\eta}$ as follows:

$$\eta^{n+j}(\widetilde{z}) = z_1^0 - z_1^j - \frac{1}{2}(\beta^j)^2, \quad j = 1, \ldots, m. \quad (2.53)$$

In particular, the equivalence $G^{j+m}_{\widetilde{z}}\widetilde{\eta}(\widetilde{z}) = z_1^0 - z_1^j - \frac{1}{2}(\beta^j)^2, \quad j = 1, \ldots, m$ holds. We have proved the consistency condition for the drift term. Now we consider the last $m$ coordinates of the volatility term. Since the vector field on $\xi(z)$ which respects the drift condition for the first $m+1$ components is given by (2.47), then adding the other $m$ components on $\xi$, we define the vector field:

$$\widetilde{\xi}(\widetilde{z}) = (\xi, \xi^{n+1}, \ldots, \xi^{n+m})^*.$$

Hence, we obtain that the condition on the spreads's volatility is:

$$G^{j+m}_{\widetilde{z}}(\widetilde{z})\widetilde{\xi}(\widetilde{z}) = \beta^j, \quad (2.54)$$

then, by the definition of the last $m$ component of $G$ given before:

$$G^{j+m}(\widetilde{z}) = u^j \quad (2.55)$$

and imposing that $\xi^{j+n} = \beta^j$, condition (2.54) is satisfied.

In conclusion, we have proved the following proposition:

**Proposition 2.3.1.** *If we consider a model $\mathcal{M}$ for the fixed-income market in which each forward rate equation is given by the Ho-Lee model and the parameterized family $\mathcal{G}$, defined by the mapping $G : \mathbb{R}^{3m+2} \longrightarrow \hat{\mathcal{H}}$,*

$$G(\widetilde{z}) = \begin{pmatrix} z_1^0 + z_3^0 x \\ z_1^1 + z_3^1 x \\ \vdots \\ z_1^m + z_3^m x \\ u^1 \\ \vdots \\ u^m \end{pmatrix}, \tag{2.56}$$

*where $\widetilde{z}$ is introduced in* (2.52)*, then the couple* $(\mathcal{M}, \mathcal{G})$ *is consistent.*

### 2.3.4 The Hull-White model

The second model analyzed by Björk and Christensen in Section 5.2 of [2] is the Hull-White model. This model is a generalization of the Vasicek model and it describes the short rate which satisfies the following dynamics:

$$dr_t = \{\Phi(t) - ar_t\}dt + \sigma dW_t, \tag{2.57}$$

where $a > 0$ and $\Phi$ is supposed to satisfy usual conditions which guarantee the existence and uniqueness of strong solution.

In order to analyse a multi-curve model determined by the SDE (2.57) it is necessary to provide the dynamics of the associated forward rate.

**Lemma 2.3.2.** *If $r_t$ satisfies the Hull-White equation, then the forward rate $f_t(T)$, such that $f_t(t) = r_t$ satisfies the following SDE:*

$$df_t(T) = \alpha_t(T)dt + \sigma e^{-a(T-t)}dW_t, \tag{2.58}$$

*where* $\alpha_t(T) = \frac{\sigma^2}{a} e^{-a(T-t)} \Big[1 - e^{-a(T-t)}\Big]$.

*Proof.* Section 2.4.1. □

Passing to the Musiela parameterization, we obtain that the forward rate equation (2.58) can be rewritten as:

$$dr_t(x) = \frac{\sigma^2}{a} e^{-ax} \Big[1 - e^{-ax}\Big] dt + \sigma e^{-ax} dW_t. \tag{2.59}$$

### 2.3.5 The Hull-White model and the Nelson-Siegel family

As done for the Ho-Lee model, we aim at checking the consistency of the couple $(\mathcal{M}, \mathcal{G})$, where $\mathcal{M}$ is determined by the vector forward rate equation (2.7) in which, for each component, we have chosen the volatility term as in (2.59): $\sigma^j(t,x) = \sigma^j e^{-a^j x}$. Moreover we have chosen $\mathcal{G} = Im[G]$, with $G$ determined by (2.33). Also in this case the volatility is constant on $\hat{\mathcal{H}}$.

We consider a model $\mathcal{M}$ associated with the forward rate equation (2.7), such that each component is given by the Hull-White forward rate equation associated with a couple of parameters $(a^j, \sigma^j)$ for each $j = 0, \dots, m$. We have to check the conditions provided by Theorem 2.2.8, in particular:

$$\hat{\mu}(G(z)), \sigma_i(G(z)) \in T_{G(z)}\mathcal{G}, \quad \forall\, i \in \{1, \dots, d\}. \tag{2.60}$$

For this example we consider a 1-dimensional Brownian Motion, i.e. $d = 1$.

If we test the condition on $\hat{\mu}^0(G(z))$, we observe that, since the volatility term is constant in $\hat{\mathcal{H}}$:

$$\hat{\mu}^0(G(z))(x) = \mathbf{F}(G^0(z))(x) + \sigma^0(t,x)\mathbf{H}\sigma^0(t,x), \tag{2.61}$$

whereas the first row of the family $\mathcal{G}$, is given by: $G^0(z,x) = z_1^0 + z_2^0 e^{-z_4^0 x} + x z_3^0 e^{-z_4^0 x}$.

Recalling the computation provided in (2.128), we obtain the following equivalence:

$$\sigma^0(t,x)\mathbf{H}\sigma^0(t,x) = (\sigma^0)^2 e^{-2a^0 x}\frac{1}{a^0}\Big[e^{a^0 x} - 1\Big], \tag{2.62}$$

Then, condition (2.61) amounts the existence of a vector $(\eta_1^0, \dots, \eta_4^0)$ (by simplicity of notation, we omit to the dependence of $z$) such that:

$$\begin{aligned} e^{-z_4^0 x}(-z_4^0 z_2^0 - z_4^0 z_3^0 x + z_3) + (\sigma^0)^2 e^{-2a^0 x}\frac{1}{a^0}\Big[e^{a^0 x} - 1\Big] &= \\ = \eta_1 + e^{-z_4^0 x}(\eta_2^0 + x(\eta_3^0 - z_2^0\eta_4^0) - x^2 z_3^0 \eta_4^0), & \end{aligned}$$

On the other hand, the condition on the volatility term $\hat{\sigma}^0(G(z)) = \sigma_0 e^{-a^0 x}$ is equivalent to the existence of a vector $\xi^0 = (\xi_1^0, \dots, \xi_4^0)$, omitting as done for $\eta$ the dependence on $z$, such that:

$$\sigma_0 e^{-a^0 x} = \xi_1^0 + e^{-z_4^0 x}[\xi_2^0 + x(\xi_3^0 - z_2^0\xi_4^0) - x^2 z_3^0 \xi_4^0], \quad \forall\, x \in \mathbb{R}_+,$$

which holds if and only if $z_4^0 = a^0$. Therefore, the couple $(\mathcal{M}, \mathcal{G})$ is not consistent.

Starting from the previous result, we can compute a parameterized family $\mathcal{G}$ such that the couple $(\mathcal{M}, \mathcal{G})$ is consistent. The strategy, developed in analogy to [2][Proposition 5.2], is given as follows: we try to modify $\mathcal{G}$, in order to impose the consistency condition.

The first step is imposing that $z_4^0 = a^0$. We can observe that, since $a^0 > 0$, the condition on $z_4$ which guarantees that $G(z) \in \hat{\mathcal{H}}$ is satisfied. Therefore, since $z_4^0$ is constant, then the condition on the drift (2.61) requires the existence of a vector $(\eta_1^0, \dots, \eta_3^0)$, such that:

$$e^{-a^0x}(-a^0 z_2^0 - a_0 z_3^0 x + z_3) + (\sigma^0)^2 e^{-2a^0x} \frac{1}{a^0}\left[e^{a^0x} - 1\right] = \eta_1^0 + e^{-a^0x}\eta_2^0 + \eta_3^0 x e^{-a^0x}.$$

To deal with the term $e^{-2a^0x}$, it can be convenient to expand the NS manifold adding an exponential of the form $e^{-2a^0x}$. Hence, we introduce the augmented NS family, defined by the following mapping:

$$G^{0A}(z,x) = z_1^0 + z_2^0 e^{-a^0x} + z_3^0 x e^{-a^0x} + z_4^0 e^{-2a^0x}, \tag{2.63}$$

where $z^0 = (z_1^0, \dots, z_4^0)$ and $z = (z^{0*}, \dots, z^{m*})$. The Fréchet derivative of this mapping is:

$$\frac{\partial G^{0A}}{\partial z}(z,x) = \begin{pmatrix} 1 & e^{-a^0x} & xe^{-a^0x} & e^{-2a^0x} \end{pmatrix}. \tag{2.64}$$

In particular, defining $\mathcal{G}^A := Im[G^A]$, we conclude that the consistency property (for the first row of $\hat{\mu}(G(z))$ is equivalent to the existence of a vector $(\eta_1^0, \dots, \eta_4^0)$ such that:

$$\begin{aligned} &e^{-a^0x}(-a^0 z_2^0 - a^0 z_3^0 x + z_3^0) - e^{-2a^0x}(2a^0 z_4^0) + (\sigma^0)^2 e^{-2a^0x}\frac{1}{a^0}\left[e^{a^0x} - 1\right] = \\ &\quad = \eta_1^0 + \eta_2^0 e^{-a^0x} + \eta_3^0 x e^{-a^0x} + \eta_4^0 e^{-2a^0x}. \end{aligned} \tag{2.65}$$

If we choose the parameters in the following way

$$\begin{cases} \eta_1^0 = 0, \\ \eta_2^0 = -a^0 z_2^0 + z_3^0 + \frac{(\sigma^0)^2}{a^0}, \\ \eta_3^0 = -a^0 z_3^0, \\ \eta_4^0 = -2a^0 z_4^0 - \frac{(\sigma^0)^2}{a^0}, \end{cases} \tag{2.66}$$

we prove the consistency condition on $\hat{\mu}^0(G(z))$.

Following the same strategy for the other forward rate components of $\hat{\mu}$, we introduce the following functions:

$$G^{jA}(z^j, x) = z_1^j + z_2^j e^{-a^jx} + z_3^j x e^{-a^jx} + z_4^j e^{-2a^jx}, \quad j = 1, \dots, m$$

where $z^j = (z_1^j, \dots, z_4^j)$. Then, we consider the vector mapping:

$$G^A := (G^{0A}, \dots, G^{mA}, G^{m+1}, \dots, G^{2m}).$$

We recall that:

$$\hat{\mu}^j(G^A(z)) = \mathbf{F}G^{jA}(z)(x) + \sigma^j(t,x)\mathbf{H}\sigma^j(t,x) - \beta^j\sigma^{j*}(t,x).$$

Therefore, $\hat{\mu}^j(G^A(z)) \in T_{G^{jA}(z)}\mathcal{G}_j^A$ is equivalent to the existence of a vector $(\eta_1^j,\ldots,\eta_4^j)$ such that:

$$\begin{aligned}
&e^{-a^jx}(-a^jz_2^j - a^jz_3^jx + z_3^j) - e^{-2a^jx}(2a^jz_4^j) + (\sigma^j)^2e^{-2a^jx}\frac{1}{a^j}\left[e^{a^jx}-1\right] - \beta^j\sigma^je^{-a^jx} = \\
&\quad = \eta_1^j + \eta_2^je^{-a^jx} + \eta_3^jxe^{-a^jx} + \eta_4^je^{-2a^jx},
\end{aligned} \tag{2.67}$$

which is equivalent to:

$$\begin{aligned}
&e^{-a^jx}\left(-a^jz_2^j + z_3^j + \frac{(\sigma^j)^2}{a^j} - \beta^j\sigma^j - a^jz_3^jx\right) + e^{-2a^jx}\left(-2a^jz_4^j - \frac{(\sigma^j)^2}{a^j}\right) = \\
&\quad = \eta_1^j + (\eta_2^j + \eta_3^jx)e^{-a^jx} + \eta_4^je^{-2a^jx}.
\end{aligned}$$

This condition is verified when the following equivalences hold:

$$\begin{cases}
\eta_1^j = 0, \\
\eta_2^j = -a^jz_2^j + z_3^j + \frac{(\sigma^j)^2}{a^j} - \beta^j\sigma^j, \\
\eta_3^j = -a^jz_3^j, \\
\eta_4^j = -2a^jz_4^j - \frac{(\sigma^j)^2}{a^j}.
\end{cases} \tag{2.68}$$

The condition on the volatility is easier to prove, since we have to check the existence of a vector $(\xi_1^j,\ldots,\xi_4^j)$ such that:

$$\sigma^je^{-a^jx} = \xi_1^j + \xi_2^je^{-a^jx} + \xi_3^jxe^{-a^jx} + \xi_4^je^{-2a^jx},$$

and this can be verified choosing:

$$\begin{cases}
\xi_1^j = 0, \\
\xi_2^j = \sigma^j, \\
\xi_3^j = 0, \\
\xi_4^j = 0.
\end{cases}$$

for each $j = 0,\ldots,m$.

Finally, we can assume that the volatility of the log-spot spread $\beta^j$ is constant. We try to exploit the conditions on the vector fields $\eta$ and $\xi$ in order to provide the components of the function $G$ associated with the spreads. If this procedure does not lead to a conclusion, we will follow the same strategy outlined in Section 2.3.3 adding an opportune number of parameters. First, we can observe that the components $\eta_1^j = 0$ for each $j = 0, \dots, m$. We can use this property in order to solve the problem.

At this point, we describe explicitly the conditions on the last $m$ components of $G$:

$$\begin{aligned} G_z^{m+j}(z)\eta(z) =& \mathbf{B}G^{0A}(z) - \mathbf{B}G^{jA}(z) - \frac{1}{2}(\beta^j)^2 \\ =& z_1^0 + z_2^0 + z_4^0 - z_1^j - z_2^j - z_4^j - \frac{1}{2}(\beta^j)^2, \quad i = 1, \dots, m. \end{aligned} \tag{2.69}$$

Let us suppose that the function $G^{m+j}$ is dependent $z_1^0, z_2^0, z_3^0, z_4^0, z_1^j, z_2^j, z_3^j, z_4^j$ because the other variables do not appear in equation (2.69). In these terms, the conditions on the components related to the drift become:

$$\begin{aligned} G_{z_2^0}^{m+j}(z)\Big(-a^0 z_2^0 + z_3^0 + \frac{(\sigma^0)^2}{a^0}\Big) + G_{z_3^0}^{m+j}(-a^0 z_3^0) + G_{z_2^j}^{m+j}\Big(-a^j z_2^j + z_3^j + \frac{(\sigma^j)^2}{a^j} - \beta^j\sigma^j\Big) + \\ + G_{z_3^j}^{m+j}(-a^j z_3^j) = z_1^0 + z_2^0 + z_4^0 - z_1^j - z_2^j - z_4^j - \frac{1}{2}(\beta^j)^2. \end{aligned}$$

Let us suppose that the real parameters $z_3^0, z_3^j > 0$, then we can consider the function:

$$\begin{aligned} G^{m+j}(z) =& \frac{1}{a^0}\Big[-z_2^0 + \Big(-z_1^0 - \frac{(\sigma^0)^2}{2(a^0)^2} + \frac{1}{2}(\beta^j)^2\Big)\log z_3^0 - \frac{z_3^0}{a^0} - \frac{1}{2}z_4^0\Big] + \\ & + \frac{1}{a^j}\Big[z_2^j + \Big(z_1^j + \frac{(\sigma^j)^2}{2(a^j)^2} - \frac{\beta^j\sigma^j}{a^j}\Big)\log z_3^j + \frac{z_3^j}{a^j} + \frac{1}{2}z_4^j\Big], \end{aligned} \tag{2.70}$$

In this case, by the vector $\eta$ is given (2.66) and (2.68), by the following equivalences hold:

$$\begin{cases} G_{z_2^0}^{m+j}(z)\Big(-a^0 z_2^0 + z_3^0 + \frac{(\sigma^0)^2}{a^0}\Big) = z_2^0 - \frac{z_3^0}{a^0} - \frac{(\sigma^0)^2}{(a^0)^2}, \\ G_{z_3^0}^{m+j}(z)(-a^0 z_3^0) = z_1^0 + \frac{(\sigma^0)^2}{2(a^0)^2} - \frac{1}{2}(\beta^j)^2 + \frac{z_3^0}{a^0}, \\ G_{z_4^0}^{m+j}(z)\Big(-2a^0 z_4^0 - \frac{(\sigma^0)^2}{a^0}\Big) = z_4^0 + \frac{(\sigma^0)^2}{2(a^0)^2}, \\ G_{z_2^j}^{m+j}(z)(-a^j z_2^j + z_3^j + \frac{(\sigma^j)^2}{a^j} - \beta^j\sigma^j) = -z_2^j + \frac{z_3^j}{a^j} + \frac{(\sigma^j)^2}{(a^j)^2} - \frac{\beta^j\sigma^j}{a^j}, \\ G_{z_3^j}^{m+j}(z)(-a^j z_3^j) = -z_1^j - \frac{(\sigma^j)^2}{2(a^j)^2} + \frac{\beta^j\sigma^j}{a^j} - \frac{z_3^j}{a^j}, \\ G_{z_4^j}^{m+j}(z)\Big(-2a^j z_4^j - \frac{(\sigma^j)^j}{a^j}\Big) = -z_4^j - \frac{(\sigma^j)^2}{2(a^j)^2}, \end{cases}$$

hence, summing the right members of the previous system we obtain the right member of (2.69), whereas, summing the left members of the previous system, we obtain the left member of (2.69), therefore, condition (2.69) is satisfied.

At this point it is necessary to find the conditions which guarantee that the function defined above are consistent also for the volatility term. We recall that the consistency condition for the forward rate equations leads to a vector field $\xi$ such that $\xi_2^j = \sigma^j$ and $\xi_k^j = 0$ for $j = 0, \ldots, m$ and $k \in \{1, 3, 4\}$. Hence, since $G^{m+j}$ is defined (2.70), we obtain that:

$$G_z^{m+j}(z)\xi = G_{z_2^0}^{m+j}(z)\xi_2^0 + G_{z_2^j}^{m+j}(z)\xi_2^j = \beta^j \quad \Leftrightarrow \quad -\frac{\sigma^0}{a^0} + \frac{\sigma^j}{a^j} = \beta^j, \quad j = 1, \ldots, m. \tag{2.71}$$

Therefore, it is necessary to assume that $\beta^j = \frac{\sigma^j}{a^j} - \frac{\sigma^0}{a^0}$ in order to have the consistency for the volatility term.

This implies that the functions $G^{m+j}$ defined in (2.70) satisfy the condition which guarantees the consistency. In conclusion, we have proved the following Proposition:

**Proposition 2.3.3.** *If we consider the model $\mathcal{M}$ given by the Hull-White model for each forward rate equation and the family $\mathcal{G}$ determined by the function $G : \mathbb{R}^{4(m+1)} \longrightarrow \hat{\mathcal{H}}$, where $\beta^j = \frac{\sigma^j}{a^j} - \frac{\sigma^0}{a^0}$ for each $j = 1, \ldots, m$:*

$$G(z) = \begin{pmatrix} z_1^0 + z_2^0 e^{-a^0 x} + z_3^0 x e^{-a^0 x} + z_4^0 e^{-2a^0 x} \\ z_1^1 + z_2^1 e^{-a^1 x} + z_3^1 x e^{-a^1 x} + z_4^1 e^{-2a^1 x} \\ \vdots \\ z_1^m + z_2^m e^{-a^m x} + z_3^m x e^{-a^m x} + z_4^m e^{-2a^m x} \\ \frac{1}{a^0}\left[-z_2^0 + \left(-z_1^0 - \frac{(\sigma^0)^2}{2(a^0)^2} + \frac{1}{2}(\beta^j)^2\right)\log z_3^0 - \frac{z_3^0}{a^0} - \frac{1}{2}z_4^0\right] + \\ +\frac{1}{a^1}\left[z_2^1 + \left(z_1^1 + \frac{(\sigma^1)^2}{2(a^1)^2} - \frac{\beta^1\sigma^1}{a^1}\right)\log z_3^1 + \frac{z_3^1}{a^1} + \frac{1}{2}z_4^1\right] \\ \vdots \\ \frac{1}{a^0}\left[-z_2^0 + \left(-z_1^0 - \frac{(\sigma^0)^2}{2(a^0)^2} + \frac{1}{2}(\beta^j)^2\right)\log z_3^0 - \frac{z_3^0}{a^0} - \frac{1}{2}z_4^0\right] + \\ +\frac{1}{a^m}\left[z_2^m + \left(z_1^m + \frac{(\sigma^m)^2}{2(a^m)^2} - \frac{\beta^m\sigma^m}{a^m}\right)\log z_3^m + \frac{z_3^m}{a^m} + \frac{1}{2}z_4^m\right] \end{pmatrix}, \tag{2.72}$$

*then the couple $(\mathcal{M}, \mathcal{G})$ is consistent.*

*On the other hand, recalling the same strategy of Proposition 2.3.1, we have that the model $\mathcal{M}$ is consistent with the family $\widetilde{\mathcal{G}}$ where $\widetilde{\mathcal{G}}$ is determined by the* *mapping* $\widetilde{G} : \mathbb{R}^{5m+4} \longrightarrow \hat{\mathcal{H}}$*:*

$$\widetilde{G}(\widetilde{z}) = \begin{pmatrix} z_1^0 + z_2^0 e^{-a^0 x} + z_3^0 x e^{-a^0 x} + z_4^0 e^{-2a^0 x} \\ z_1^1 + z_2^1 e^{-a^1 x} + z_3^1 x e^{-a^1 x} + z_4^1 e^{-2a^1 x} \\ \vdots \\ z_1^m + z_2^m e^{-a^m x} + z_3^m x e^{-a^m x} + z_4^m e^{-2a^m x} \\ u^1 \\ \vdots \\ u^m \end{pmatrix}, \tag{2.73}$$

*where* $\widetilde{z} = (z, u^1, \ldots, u^m)^*$.

We can make a last observation for the Hull-White model. In Remark 5.1 of [2] the following result is shown:

**Remark 2.3.4.** *The augmented manifold* $\mathcal{G}^A$ *is not the smallest possible manifold consistent with Hull-White. The minimal manifold satisfying the consistency property is given by:*

$$G(z, x) = z_1 e^{-ax} + z_2 e^{-2ax}. \tag{2.74}$$

This remark still holds in the multi-curve framework. Indeed, choosing a parameterized family described by (2.74) for each component $G^j$ with $i \in \{0, \ldots, m\}$, for the first coordinate of the drift $\hat{\mu}^0(G(z))$, there exists a vector $(\eta_1^0, \eta_2^0)$ such that:

$$-a^0 z_1^0 e^{-a^0 x} - 2a^0 z_2 e^{-2a^0 x} + (\sigma^0)^2 e^{-2a^0 x} \frac{1}{a^0} \left[ e^{a^0 x} - 1 \right] = \eta_1^0 e^{-a^0 x} + \eta_2^0 e^{-2a^0 x}, \tag{2.75}$$

Indeed, choosing $(\eta_1^0, \eta_2^0)$ as follows:

$$\begin{cases} \eta_1^0 = -a^0 z_1^0 + \frac{(\sigma^0)^2}{a^0}, \\ \eta_2^0 = -2a^0 z_2^0 - \frac{(\sigma^0)^2}{a^0}, \end{cases} \tag{2.76}$$

the condition (2.75) is verified.

For the other coordinates, the consistency condition is given by the existence of a vector $(\eta_1^j, \eta_2^j)$ such that, for each $j = 1, \ldots, m$:

$$-a^j z_1^j e^{-a^j x} - 2a^j z_2^j e^{-2a^j x} + (\sigma^j)^2 e^{-2a^j x} \frac{1}{a^j} \left[ e^{a^j x} - 1 \right] - \beta^j \sigma^j e^{-a^j x} = \eta_1^j e^{-a^j x} + \eta_2^j e^{-2a^j x},$$

which is equivalent to

$$\left[ -a^j z_1^j + \frac{(\sigma^j)^2}{a^j} - \beta^j \sigma^j \right] e^{-a^j x} + \left[ -2a^j z_2^j - \frac{(\sigma^j)^2}{a^j} \right] e^{-2a^j x} = \eta_1^j e^{-a^j x} + \eta_2^j e^{-2a^j x}. \tag{2.77}$$

Also in this case, in order to impose the consistency condition, it is sufficient to choose:

$$\begin{cases} \eta_1^j = -a^j z_1^j + \frac{(\sigma^j)^2}{a^j} - \beta^j \sigma^j, \\ \eta_2^j = -2a^j z_2^j - \frac{(\sigma^j)^2}{a^j}. \end{cases}$$

For the volatility term, we recall that $\sigma^j(\hat{r}_t) = \sigma^j e^{-a^j x}$ for each $j = 0, \ldots, m$. Hence, the condition is:

$$\sigma^j e^{-a^j x} = \xi_1^j e^{-a^j x} + \xi_2^j e^{-2a^j x},$$

therefore, the solution is given by $\xi^j = (\sigma^j, 0)^*$, for each $j = 0, \ldots, m$. In particular, we have constructed two vector fields defined on $\mathbb{R}^{2(m+1)}$, one for the drift term and one for the volatility term:

$$\begin{aligned} \eta(z) &= \begin{pmatrix} \eta_1^0 & \eta_2^0 & \eta_1^1 & \eta_2^1 & \cdots & \eta_m^1 & \eta_m^2 \end{pmatrix}, \\ \xi(z) &= \begin{pmatrix} \xi_1^0 & \xi_2^0 & \xi_1^1 & \xi_2^1 & \cdots & \xi_m^1 & \xi_m^2 \end{pmatrix}. \end{aligned}$$

Now it is necessary to consider the last $m$ components of $G$. As done before, first, we try to find suitable conditions using the parameters already introduced, if a solution can not be found, we exploit the procedure outlined in the previous subsection and we add an opportune number of parameters.

The conditions for the coordinates of the drift are:

$$\begin{aligned} G_z^{m+j}(z)\eta(z) =& \mathbf{B}G^0(z) - \mathbf{B}G^j(z) - \frac{1}{2}(\beta^j)^2, \\ =& z_1^0 + z_1^0 - z_1^j - z_2^j - \frac{1}{2}(\beta^j)^2, \qquad j = 1, \ldots, m, \end{aligned} \tag{2.78}$$

where the vector field $\eta$ is given by (2.77). If we assume that the function $G$ satisfies:

$$G^{m+j}(z) = -\frac{z_1^0 + \frac{1}{2}z_2^0}{a^0} + \frac{z_1^j + \frac{1}{2}z_2^j}{a^j}, \quad j = 1, \ldots, m,$$

therefore, the left member of (2.77) is given by:

$$\begin{aligned} G_z^{m+j}(z)\eta(z) = & -\frac{1}{a^0}\Big(-a^0 z_1^0 + \frac{(\sigma^0)^2}{a^0}\Big) - \frac{1}{2a^0}\Big(-2a^0 z_2^0 - \frac{(\sigma^0)^2}{a^0}\Big) + \Big(-a^j z_1^j + \\ & + \frac{(\sigma^j)^2}{a^j}\Big)\frac{1}{a^j} + \Big(-2a^j z_2^j - \frac{(\sigma^j)^2}{a^j}\Big)\frac{1}{2a^j} - \frac{\beta^j \sigma^j}{a^j} \\ = & \; z_1^0 + z_2^0 - z_1^j - z_2^j - \frac{\beta^j \sigma^j}{a^j} - \frac{1}{2}\Big(\frac{(\sigma^0)^2}{(a^0)^2} - \frac{(\sigma^j)^2}{(a^j)^2}\Big), \end{aligned}$$

hence, condition (2.78) on the drift is satisfied if and only if:

$$\begin{aligned} & \frac{1}{2}(\beta^j)^2 = \frac{\beta^j \sigma^j}{a^j} + \frac{1}{2}\Big(\frac{(\sigma^0)^2}{(a^0)^2} - \frac{(\sigma^j)^2}{(a^j)^2}\Big), \qquad \Longleftrightarrow \\ & (\beta^j)^2 - 2\frac{\beta^j \sigma^j}{a^j} - \Big(\frac{(\sigma^0)^2}{(a^0)^2} - \frac{(\sigma^j)^2}{(a^j)^2}\Big) = 0, \end{aligned} \tag{2.79}$$

whose solutions are:

$$\begin{aligned}\beta^j_{1,2} =& \frac{\sigma^j}{a^j} \pm \sqrt{\frac{(\sigma^j)^2}{(a^j)^2} + \Big(\frac{(\sigma^0)^2}{(a^0)^2} - \frac{(\sigma^j)^2}{(a^j)^2}\Big)} \\ =& \frac{\sigma^j}{a^j} \pm \frac{\sigma^0}{a^0}, \quad j = 1, \dots, m.\end{aligned} \tag{2.80}$$

On the other hand, for the volatility term the condition is:

$$\beta^j = G_z^{m+j}\xi(z) = -\frac{\sigma^0}{a^0} + \frac{\sigma^j}{a^j},$$

Recalling the conditions provided in (2.80), we obtain that the unique condition on $\beta^j$ is:

$$\beta^j = \frac{\sigma^j}{a^j} - \frac{\sigma^0}{a^0}, \quad j = 1, \dots, m, \tag{2.81}$$

which is the same provided in (2.71).

The previous equivalence determines a dependence of the volatility of the LIBOR forward rates, expressed by the ratio $\frac{\sigma^j}{a^j}$ from the volatility of the associated spread $\beta^j$ and the volatility of the OIS forward rate, expressed by the ratio $\frac{\sigma^0}{a^0}$. In conclusion, the following Proposition is proved:

**Proposition 2.3.5.** *If we consider the model $\mathcal{M}$ determined by the Hull-White model for each forward rate equation and the family $\mathcal{G}$ described by the function $G : \mathbb{R}^{2(m+1)} \longrightarrow \hat{\mathcal{H}}$, where $\beta^j$ satisfies* (2.81) *for each $i = 1, \dots, m$:*

$$G(z) = \begin{pmatrix} z_1^0 e^{-a^0 x} + z_2^0 e^{-2a^0 x} \\ z_1^1 e^{-a^1 x} + z_2^1 e^{-2a^1 x} \\ \vdots \\ z_1^m e^{-a^m x} + z_2^m e^{-2a^m x} \\ -\frac{z_1^0 + \frac{1}{2} z_2^0}{a^0} + \frac{z_1^1 + \frac{1}{2} z_2^1}{a^1} \\ \vdots \\ -\frac{z_1^0 + \frac{1}{2} z_2^0}{a^0} + \frac{z_1^m + \frac{1}{2} z_2^m}{a^m} \end{pmatrix}, \tag{2.82}$$

*then, the couple $(\mathcal{M}, \mathcal{G})$ is consistent.*

*On the other hand, recalling the same strategy of Proposition 2.3.1, we have that the model $\mathcal{M}$ is consistent with the family $\widetilde{\mathcal{G}}$ where $\widetilde{\mathcal{G}}$ is determined by the* *mapping* $\widetilde{G} : \mathbb{R}^{3m+2} \longrightarrow \hat{\mathcal{H}}$*:*

$$\widetilde{G}(\widetilde{z}) = \begin{pmatrix} z_1^0 e^{-a^0 x} + z_2^0 e^{-2a^0 x} \\ z_1^1 e^{-a^1 x} + z_2^1 e^{-2a^1 x} \\ \vdots \\ z_1^m e^{-a^m x} + z_2^m e^{-2a^m x} \\ u^1 \\ \vdots \\ u^m \end{pmatrix}, \tag{2.83}$$

*where* $\widetilde{z} = (z, u^1, \ldots, u^m)^*$.

### 2.3.6 The Svensson family

The Svensson family is one of the most widely employed forward parameterized families. The Svensson family is described by a six-dimensional vector of parameters: $z = (z_1, \ldots, z_6)$ (for the details we refer to [14] Chapter 9.5.2). The mapping which describes this submanifold is:

$$G(z, x) := z_1 + (z_2 + z_3 x)e^{-z_5 x} + z_4 x e^{-z_6 x}. \tag{2.84}$$

In particular, the Fréchet derivative of this mapping is:

$$G_z(z, x) = \begin{pmatrix} 1 & e^{-z_5 x} & x e^{-z_5 x} & x e^{-z_6 x} & -x(z_2 + z_3 x)e^{-z_5 x} & -z_4 x^2 e^{-z_6 x} \end{pmatrix}.$$

**The Svensson family and the Hull-White model** We consider the Hull-White model (2.59). Then, the couple $(\mathcal{M}, \mathcal{G})$ we aim at studying is determined by:

$$\mathcal{M} : \quad d\hat{r}_t = \hat{\mu}(\hat{r}_t)dt + \hat{\sigma}(\hat{r}_t) \circ dW_t,$$

where each row associated with the above SDE is described by the volatility term $\sigma^i e^{-a^i x}$ (Hull-White volatility term), whereas we suppose that the volatility term of the spread processes is constant.

On the other hand, the submanifold $\mathcal{G} \subset \hat{\mathcal{H}}$ is defined componentwise by the Svensson family: $\mathcal{G} := Im[G]$, where $G$ is:

$$G(z, x) = \begin{pmatrix} z_1^0 + (z_2^0 + z_3^0 x)e^{-z_5^0 x} + z_4^0 x e^{-z_6^0 x} \\ z_1^1 + (z_2^1 + z_3^1 x)e^{-z_5^1 x} + z_4^1 x e^{-z_6^1 x} \\ \vdots \\ z_1^m + (z_2^m + z_3^m x)e^{-z_5^m x} + z_4^m x e^{-z_6^m x} \\ G^{m+1}(z) \\ \vdots \\ G^{2m}(z) \end{pmatrix}, \tag{2.85}$$

where $z = (z_1^0, \ldots, z_6^m)$. First of all, we focus on the first $m+1$ coordinates and when we find the consistency conditions for those components (associated with the forward rate equations) we can characterize the conditions on the last $m$ components of the functions $G$.

By construction, $G$ is defined on an open subset of a finite-dimensional real vector space, $\mathcal{Z}$, and in order to assure that $G(z,x) \in \hat{\mathcal{H}}$, for every $z$, we need to impose that $z_5^j, z_6^j > -\frac{\gamma}{2}$ for each $j = 0, \ldots, m$.

In order to determine whether the couple $(\mathcal{M}, \mathcal{G})$ we exploit the invariance Theorem 2.2.8. As before, we need to check conditions (2.60). Starting from the first coordinate of the drift term $\hat{\mu}^0(G(z))$ we get:

$$\begin{aligned}\hat{\mu}^0(G(z)) &= \mathbf{F}(G^0(z))(x) + \sigma^0 \mathbf{H} \sigma^0 \\ &= e^{-z_5^0}[z_3^0 - z_5^0(z_2^0 + z_3^0 x)] + e^{-z_6^0 x}[z_4^0 - z_6^0 z_4^0 x] + \frac{(\sigma^0)^2}{a^0} e^{-2a^0 x}\Big[e^{a^0 x} - 1\Big].\end{aligned} \tag{2.86}$$

Condition (2.60) is equivalent to the existence of a vector $\eta^0 = (\eta_1^0, \ldots, \eta_6^0) \in \mathbb{R}^6$ such that:

$$\hat{\mu}^0(G(z)) = \eta_1^0 + e^{-z_5 x}[\eta_2^0 + x\eta_3^0 - x\eta_5^0(z_2^0 + z_3^0 x)] + xe^{-z_6 x}[\eta_4^0 - z_4^0 x \eta_0^6]. \tag{2.87}$$

Clearly, the previous equivalence does not hold if: $\{z_5^0, z_6^0\} \neq \{a^0, 2a^0\}$. Then we suppose that: $z_5^0 = a^0$, $z_6^0 = 2a^0$. Under these assumptions, the Svensson forward rate function associated with $i =$ becomes:

$$G^0(z,x) = z_1^0 + z_2^0 e^{-a^0 x} + z_3^0 x e^{-a^0 x} + z_4^0 x e^{-2a^0 x}. \tag{2.88}$$

The Fréchet derivative of $G$ is given by:

$$G_z^0(z,x) = \begin{pmatrix} 1 & e^{-a^0 x} & xe^{-a^0 x} & xe^{-2a^0 x} \end{pmatrix}, \tag{2.89}$$

whereas:

$$\mathbf{F}G(z,x) = e^{-a^0 x}[(z_2^0 + z_3^0 x)(-a^0) + z_3^0] + e^{-2a^0 x}[z_4 - 2a^0 z_4 x]. \tag{2.90}$$

Therefore, the consistency condition $\mu^0(G(z)) = G_z^0(z)\eta^0(z)$ is equivalent to:

$$\begin{aligned}e^{-a^0 x}&[(z_2^0 + z_3^0 x)(-a^0) + z_3^0] + e^{-2a^0 x}[z_4 - 2a^0 z_4 x] + \frac{(\sigma^0)^2}{a^0} e^{-a^0 x} - \frac{(\sigma^0)^2}{a^0} e^{-2a^0 x} = \\ &= \eta_1^0 + e^{-a^0 x}[\eta_2^0 + \eta_3^0 x] + \eta_4^0 x e^{-2a^0 x}.\end{aligned} \tag{2.91}$$

Imposing that the coefficients of the exponential terms are equal

$$\eta_1^0 = 0 \tag{2.92}$$

$$z_3^0 - a^0(z_2^0 + z_3^0 x) + \frac{(\sigma^0)^2}{a^0} = \eta_2^0 + \eta_3^0 x \tag{2.93}$$

$$z_4^0(1 - 2a^0 x) - \frac{(\sigma^0)^2}{a^0} = \eta_4^0 x. \tag{2.94}$$

If the last equivalence holds for each $x$ then $z_4^0 = \frac{(\sigma^0)^2}{a^0}$, then we have the inconsistency of the couple $(\mathcal{M}, \mathcal{G})$. This implies that the Svensson family is inconsistent with the Hull-White model.

We can enlarge the previous family adding a term $z_5 e^{-2a^0 x}$, where $z_5$ is a new parameter, in order obtain the consistency. Indeed, if we consider the mapping:

$$G^0(z, x) = z_1^0 + [z_2^0 + z_3^0 x]e^{-a^0 x} + [z_4^0 x + z_5^0]e^{-2a^0 x}. \tag{2.95}$$

The previous mapping is an extension of the mapping $G^A$, defined on (2.63). The associated submanifold: $\mathcal{G} := Im[G]$ forms with the Hull-White model a couple whose first row respects condition (2.60) on the drift. In particular, if we compute the drift term on a function $G$ which takes value on $\mathcal{H}$ and such that its first coordinate is $G^0$, we obtain:

$$\begin{aligned}
\mu^0(G(z))(x) &= \mathbf{F}G^0(z, x) - \sigma^0(x)\mathbf{H}\sigma^0(x) \\
&= [-a^0(z_2^0 + z_3^0 x) + z_3^0]e^{-a^0 x} + [-2a^0(z_4^0 x + z_5^0) + z_4^0]e^{-2a^0 x} + \\
&\quad + \frac{(\sigma^0)^2}{a^0}(e^{-a^0 x} - e^{-2a^0 x}) \\
&= \Big(\frac{(\sigma^0)^2}{a^0} - a^0 z_2^0 + z_3^0 - z_3^0 a^0 x\Big)e^{-a^0 x} + \Big(-\frac{(\sigma^0)^2}{a^0} - 2^0 z_5^0 + z_4^0 - 2z_4^0 a_x^0\Big)e^{-a^0 x},
\end{aligned}$$

whereas,

$$\begin{aligned}
G_z^0(z, x)\eta^0(z) =& \eta_1^0(z) + e^{-a^0 x}\eta_2^0(z) + xe^{-a^0 x}\eta_3^0(z) + xe^{-2a^0 x}\eta_4^0(z) + e^{-2a^0 x}\eta_5^0(z) \\
=& e^{-a^0 x}\Big[\eta_2^0(z) + x\eta_3^0(z)\Big] + e^{-2a^0 x}\Big[x\eta_4^0(z) + \eta_5^0(z)\Big] + \eta_1^0(z),
\end{aligned}$$

Which implies that:

$$\begin{cases}
\eta_1^0(z) = 0, \\
\eta_2^0(z) = \frac{(\sigma^0)^2}{a^0} - a^0 z_2^0 + z_3^0, \\
\eta_3^0(z) = -z_3^0 a^0, \\
\eta_4^0(z) = -2z_4^0 a^0, \\
\eta_5^0(z) = -\frac{(\sigma^0)^2}{a^0} - 2a^0 z_5^0 + z_4^0.
\end{cases} \tag{2.96}$$

For the condition on the drift terms for $j = 1, \dots, m$, we need to follow the same strategy adopted for $j = 0$. In particular, the Svensson family (2.84) is inconsistent with the model Hull-White model. The same discussion can be done for the family (2.95) since the drift term associated to the $j-th$ coordinate of the forward rate equation is:

$$\hat{\mu}^j(G(z)) = \mathbf{F}(G^j(z))(x) + \sigma^j \mathbf{H} \sigma^j(x) - \beta^j \sigma^{i*}(x),$$

where $G^j(z)$ is given by:

$$G^j(z,x) = z_1^j + [z_2^j + z_3^j x]e^{-a^j x} + [z_4^j x + z_5^j]e^{-2a^j x}.$$

The Fréchet derivative of this mapping is:

$$G_z^j(z,x) = \begin{pmatrix} 1 & e^{-a^j x} & xe^{-a^j x} & xe^{-2a^j x} & e^{-2a^j x} \end{pmatrix},$$

and the consistency condition for this component becomes:

$$\begin{aligned} & e^{-a^j x}[-a^j(z_2^j + z_3^j x) + z_3^j] + e^{-2a^j x}[-2a^j(z_4^j x + z_5^j) + z_4^j] + \frac{(\sigma^j)^2}{a^j} e^{-2a^j x}\left[e^{a^j x} - 1\right] + \\ & - \beta^j \sigma^j e^{-a^j x} \in T_{G^j(x)} \mathcal{G}^j. \end{aligned}$$

Rewriting the previous expression:

$$e^{-a^j x}\left[-a^j(z_2^j + z_3^j x) + z_3^j + \frac{(\sigma^j)^2}{a^j} - \beta^j \sigma^j\right] + e^{-2a^j x}\left[-2a^j(z_4^j x + z_5^j) - \frac{(\sigma^j)^2}{a^j} + z_4^j\right] \in T_{G^j(x)} \mathcal{G}^j.$$

If we consider the vector:

$$\eta^j = \left(0 \quad \left[-a^j z_2^j + z_3^j + \tfrac{(\sigma^j)^2}{a^j} - \beta^j \sigma^j\right] \quad [-a^j z_3^j] \quad [-2a^j z_4^j] \quad \left[-2a^j z_5^j - \tfrac{(\sigma^j)^2}{a^j} + z_4^j\right]\right),$$

where we omitted the dependence on the $z$ variable of $\eta$. Hence, we get that $\hat{\mu}^j(G(z))(x) = G_z^j(z,x)\eta(z)$.

Finally, we have to check that the volatility term of the model $\mathcal{M}$ satisfies condition (2.60). In particular, by the form of the mapping which defines the extension of the Svensson family (2.95), the condition on the volatility of the first $m+1$ components of the forward rate equation is:

$$\sigma^j e^{-a^j x} = G_z^j(z,x)\xi,$$

where $\xi = \begin{pmatrix} 0 & \sigma^j & 0 & 0 & 0 \end{pmatrix}^*$.

For the coordinates related to the spreads components $G^{m+1}, \dots, G^{2m}$ we observe that, for every $j = 1, \dots, m$ the condition is:

$$\begin{aligned} G_z^{m+j}(z)\eta^j(z) =& \mathbf{B}G^0(z) - \mathbf{B}G^j - \frac{1}{2}(\beta^j)^2 \\ =& z_1^0 + z_2^0 + z_5^0 - z_1^j - z_2^j - z_5^j - \frac{1}{2}(\beta^j)^2, \qquad j = 1, \dots, m. \end{aligned} \tag{2.97}$$

For simplicity, we can assume that the function $G^{m+j}$, for every $j = 1, \dots, m$ depends only on the variables $z_1^0, z_2^0, z_3^0, z_4^0, z_5^0, z_1^j, z_2^j, z_3^j, z_4^j, z_5^j$ because the other coordinates do not appear in equation (2.97). Adopting a similar strategy to Proposition 2.3.3 and assuming that $z_3^0, z_4^0, z_3^j, z_4^j$ we introduce the function, for every $j = 1, \dots, m$:

$$\begin{aligned} G^{m+j}(z) = &-\frac{1}{a^0}\Big(z_2^0 + z_1^0 \log z_3^0 + \frac{z_3^0}{a^0} + \frac{z_5^0}{2}\Big) - \frac{\log z_4^0}{2a^0}\Big(\frac{(\sigma^j)^2}{2(a^0)^2}\Big) - \frac{z_4^0}{4(a^0)^2} + \\ &+ \frac{1}{a^j}\Big(z_2^j + z_1^j \log z_3^j + \frac{z_3^j}{a^j} + \frac{z_5^j}{2}\Big) + \frac{\log z_4^j}{2a^j}\Big(\frac{(\sigma^j)^2}{2(a^j)^2}\Big) + \frac{z_4^j}{4(a^j)^2} \end{aligned} \tag{2.98}$$

We observe that the following equivalences hold:

$$\begin{cases} G^{m+j}_{z_2^0}\Big(-a^0 z_2^0 + z_3^0 + \frac{(\sigma^0)^2}{a^0}\Big) = z_2^0 - \frac{z_3^0}{a^0} - \frac{(\sigma^0)^2}{(a^0)^2}, \\ G^{m+j}_{z_3^0}\Big(-a^0 z_3^0\Big) = z_1^0 + \frac{z_3^0}{a^0}, \\ G^{m+j}_{z_4^0}\Big(-2a^0 z_4^0\Big) = \frac{(\sigma^0)^2}{2(a^0)^2} + \frac{z_4^0}{2a^0}, \\ G^{m+j}_{z_5^0}\Big(-\frac{(\sigma^0)^2}{a^0} - 2a^0 z_5^0 + z_4^0\Big) = +\frac{(\sigma^0)^2}{2(a^0)^2} + z_5^0 - \frac{z_4^0}{2a^0}, \\ G^{m+j}_{z_2^j}\Big(-a^j z_2^j + z_3^j + \frac{(\sigma^j)^2}{a^j} - \beta^j \sigma^j\Big) = -z_2^j + \frac{z_3^j}{a^j} + \frac{(\sigma^j)^2}{(a^j)^2} - \frac{\beta^j \sigma^j}{a^j}, \\ G^{m+j}_{z_3^j}\Big(-a^j z_3^j\Big) = -z_1^j - \frac{z_3^j}{a^j}, \\ G^{m+j}_{z_4^j}\Big(-2a^j z_4^j\Big) = -\frac{(\sigma^j)^2}{2(a^j)^2} - \frac{z_4^j}{2a^j}, \\ G^{m+j}_{z_5^j}\Big(-\frac{(\sigma^j)^2}{a^j} - 2a^j z_5^j + z_4^j\Big) = -\frac{(\sigma^j)^2}{2(a^j)^2} - z_5^j + \frac{z_4^j}{2a^j}. \end{cases}$$

In particular, condition (2.98) becomes:

$$G^{m+j}_z(z)\eta(z) = z_2^0 + z_1^0 + z_5^0 - z_2^j - \frac{\sigma^j}{a^j}\beta^j - z_1^j - z_5^j,$$

which is satisfies if and only if $\beta^j = \frac{\sigma^j}{2a^j}$. Therefore, it is necessary to control the condition on the volatility. Recalling that the vector field $\xi = (\xi^0, \dots, \xi^m)$, the condition is:

$$G^{m+j}_z(z)\xi(z) = \beta^j \ \Rightarrow -\frac{\sigma^0}{a^0} + \frac{\sigma^j}{a^j} = \beta^j \ \Rightarrow \frac{\sigma^0}{a^0} = \beta^j,$$

for each $j = 1, \dots, m$. In conclusion, we have proved the following Proposition:

**Proposition 2.3.6.** *If we consider the model $\mathcal{M}$ determined by the Hull-White model or each forward rate equation and the family $\mathcal{G} = Im[G]$ described by the* *function:* $G : \mathbb{R}^{5(m+1)} \longrightarrow \hat{\mathcal{H}}$, *where* $\beta^j = \frac{\sigma^j}{2a^j} = \frac{\sigma^0}{a^0}$, *for each* $j = 1, \dots, m$*:*

$$G(z) = \begin{pmatrix} z_1^0 + (z_2^0 + z_3^0 x)e^{-a^0 x} + (z_4^0 x + z_5^0)e^{-2a^0 x} \\ \vdots \\ z_1^m + [z_2^m + z_3^m x]e^{-a^m x} + [z_4^m x + z_5^m]e^{-2a^m x} \\ -\frac{1}{a^0}\Big(z_2^0 + z_1^0 \log z_3^0 + \frac{z_3^0}{a^0} + \frac{z_5^0}{2}\Big) - \frac{\log z_4^0}{2a^0}\Big(\frac{(\sigma^0)^2}{2(a^0)^2}\Big) - \frac{z_4^0}{4(a^0)^2} + \\ +\frac{1}{a^1}\Big(z_2^1 + z_1^1 \log z_3^1 + \frac{z_3^1}{a^1} + \frac{z_5^1}{2}\Big) + \frac{\log z_4^1}{2a^1}\Big(\frac{(\sigma^1)^2}{2(a^1)^2}\Big) + \frac{z_4^1}{4(a^1)^2} \\ \vdots \\ -\frac{1}{a^0}\Big(z_2^0 + z_1^0 \log z_3^0 + \frac{z_3^0}{a^0} + \frac{z_5^0}{2}\Big) - \frac{\log z_4^0}{2a^0}\Big(\frac{(\sigma^0)^2}{2(a^0)^2}\Big) - \frac{z_4^0}{4(a^0)^2} + \\ +\frac{1}{a^m}\Big(z_2^m + z_1^m \log z_3^m + \frac{z_3^m}{a^m} + \frac{z_5^m}{2}\Big) + \frac{\log z_4^m}{2a^m}\Big(\frac{(\sigma^m)^2}{2(a^m)^2}\Big) + \frac{z_4^m}{4(a^m)^2} \end{pmatrix}, \tag{2.99}$$

where $z = (z_1^0, \dots, z_5^0, z_1^1, \dots, \dots, z_5^m)^*$, hence the couple $(\mathcal{M}, \mathcal{G})$ is consistent.

On the other hand, recalling the same strategy of Proposition 2.3.3, we have that the model $\mathcal{M}$ is consistent with the family $\widetilde{G}$ where $\widetilde{G} = Im[\widetilde{G}]$ and the mapping $\widetilde{G} : \mathbb{R}^{6m+5} \longrightarrow \hat{\mathcal{H}}$ is:

$$\widetilde{G}(z) = \begin{pmatrix} z_1^0 + (z_2^0 + z_3^0 x)e^{-a^0 x} + (z_4^0 x + z_5^0)e^{-2a^0 x} \\ \vdots \\ z_1^m + [z_2^m + z_3^m x]e^{-a^m x} + [z_4^m x + z_5^m]e^{-2a^m x} \\ u^1 \\ \vdots \\ u^m \end{pmatrix}, \tag{2.100}$$

where $\widetilde{z} = (z, u^1, \dots, u^m)^*$.

### 2.3.7 Hybrid models

We can also consider hybrid models, where each component is described by different model. For instance, we can consider the following forward rate model $\mathcal{M}$:

$$d\hat{r}_t = \hat{\mu}(\hat{r}_t)dt + \hat{\sigma}(\hat{r}_t) \circ dW_t,$$

where

$$d\hat{r}_t^0(x) = \frac{(\sigma^0)^2}{a^0} e^{-a^0 x}\Big[1 - e^{-a^0 x}\Big]dt + \sigma^0 e^{-a^0 x} \circ dW_t, \tag{2.101}$$

$$d\hat{r}_t^i(x) = \Big[(\sigma^i)^2 t\Big(x + \frac{t}{2}\Big) + \frac{\partial r_0(x)}{\partial x} + (\sigma^i)^2 t\Big]dt + \sigma^i \circ dW_t, \qquad \forall\, i \in \{1, \dots, m\}. \tag{2.102}$$

In other words, we model the OIS forward rate with the Hull-White model while the forward rates associated with each LIBOR rates are described by the Ho-Lee model. We suppose moreover that the log-spread processes have constant volatility.

In the previous paragraphs we have determined conditions which guarantee the consistency of a couple $(\mathcal{M}', \mathcal{G}')$ where $\mathcal{G}'$ is a forward rate family and $\mathcal{M}'$ is a multi-curve model in which each forward rate equation is described by the same model. We can exploit those results in order to understand which structure a consistent forward rate family has to respect. First of all, denoting with $\mathcal{G}$ the candidate forward rate submanifold, we have to impose that the first coordinate of $\mathcal{G}$ forms with the Hull-White model associated with the first row of $\mathcal{M}$, a consistent couple. For instance, in analogy to (2.74), we can choose:

$$G^0(z,x) = z_1^0 e^{-a^0 x} + z_2^0 e^{-2a^0 x}.$$

Therefore, we have to find conditions for the equations associated with the LIBOR forward rates. In particular, we can observe that it is no longer possible to choose the same family:

$$G^j(z,x) = z_1^j e^{-a^j x} + z_2^j e^{-2a^j x}, \quad j = 1, \dots, m,$$

because that family is not consistent with the Ho-Lee model. Indeed, for the invariance Theorem 2.2.8, the condition on the drift term is given by:

$$\hat{\mu}^j(G(z)) = \mathbf{F}G^j(z) + \sigma^j \mathbf{H} \sigma^j - \beta^j \sigma^j \in T_{G^j(z)} \mathcal{G}^j, \qquad m \in \{1, \dots, m\}.$$

If we write explicitly the previous equation, we can observe that the condition is equivalent to the existence of a vector $\eta^j = \begin{pmatrix} \eta_1^j & \eta_2^j \end{pmatrix}$, such that:

$$-a^j z_1^j e^{-a^j x} - 2a^2 z_2 e^{-2a^j x} + (\sigma^j)^2 x - \beta^j \sigma^j = \eta_1^j e^{-a^j x} + \eta_2^j e^{-2a^j x}.$$

Clearly the previous equation has no solution, so that the entire couple $(\mathcal{M}, \mathcal{G})$ is inconsistent.

One possibility is to build a linear combination of the family determined by the function (2.74) and the degenerated NS family (2.35):

$$G^j(z) = z_1^j + z_2^j x + z_3^j e^{-a^j x} + z_4^j e^{-2a^j x}, \tag{2.103}$$

but it seems not to be an efficient strategy to follow, because we introduce to many parameters. Therefore, in order to solve the problem of consistency for such a forward rate model, we can build a sub-manifold with a different structure for each row. We proved in Proposition 2.3.1 that the degenerate NS family is consistent with the Ho-Lee model, therefore we construct the following sub-manifold:

$$G(z,x) = \begin{pmatrix} z_1^0 e^{-a^0 x} + z_2^0 e^{-2a^0 x} \\ z_1^1 + z_2^1 x \\ \vdots \\ z_1^m + z_2^m x \\ G^{m+1}(z) \\ \vdots \\ G^{2m}(z) \end{pmatrix},$$

where as usual we do not focus at the moment on the last $m$ coordinates of the mapping $G$.

Recalling (2.76) and (2.45), the condition on the drift $G_z(z)\eta(z) = \mu(G(z))$ is satisfied by:

$$\eta(z) = \begin{pmatrix} -a^0 z_1^0 + \frac{(\sigma^0)^2}{a^0} & -2a^0 z_2^0 - \frac{(\sigma^0)^2}{a^0} & z_2^1 - \beta^1\sigma^1 & (\sigma^1)^2 & \cdots & z_2^m - \beta^m\sigma^m & (\sigma^m)^2 \end{pmatrix},$$

whereas the condition for the first coordinate is: $G_z^0\xi(z) = \sigma^0 e^{-a^0 x}$, which is equivalent to impose that the vector

$$\xi(z) = \begin{pmatrix} \xi_1^0(z) & \xi_2^0(z) & \xi_1^1(z) & \xi_2^1(z) & \ldots & \xi_1^m(z) & \xi_2^m(z) \end{pmatrix},$$

satisfies $\xi_1^0(z)e^{-a^0 x} + \xi_2^0 e^{-2a^0 x} = \sigma^0 e^{-a^0 x}$, which implies that $\xi_1^0 = \sigma^0$ and $\xi_2^0 = 0$. On the other hand, computing the equivalence $G_z^j(z)\xi(z) = \sigma^j$ for every $j = 1, \ldots, m$ we obtain: $\xi_1^j + \xi_2^j x = \sigma^j$. The previous equivalence is satisfied by imposing that $\xi_1^j = \sigma^j$ and $\xi_2^j = 0$ for each $j = 1, \ldots, m$. In conclusion:

$$\xi(z) = \begin{pmatrix} \sigma^0 & 0 & \sigma^1 & 0 & \sigma^2 & 0 & \cdots & \sigma^m & 0 \end{pmatrix}.$$

Now, we analyze the condition on the last $m$ component in order to understand if it necessary to add a suitable number of parameters or it is possible to exploit the form of the vector fields $\eta$ and $\xi$ in order to characterize the functions $G^{m+j}$. In particular, we observe that the condition on the $m+j$ components, for $j = 1, \ldots, m$ is:

$$\begin{aligned} G_z^{m+j}(z)\eta(z) =& \mathbf{B}G^0(z) - \mathbf{B}G^j(z) - \frac{1}{2}(\beta^j)^2 \\ =& z_1^0 + z_2^0 - z_1^j - \frac{1}{2}(\beta^j)^2, \end{aligned}$$

Differently from the previous cases, we assume that $G^{m+j}$ is dependent on the variables $z_1^0, z_2^0, z_1^j, z_2^j, z_2^{j+1}$, for every $j = 1, \ldots, m-1$ and the function $G^{2m}$ is dependent on the variables $z_1^0, z_2^0, z_1^j, z_2^j, z_2^1$. For each $j = 1, \dots, m-1$, the condition becomes:

$$G^{m+j}_{z_1^0}(z)\Big(-a^0 z_1^0 + \frac{(\sigma^0)^2}{a^0}\Big) + G^{m+j}_{z_2^0}(z)\Big(-2a^0 z_2^0 - \frac{(\sigma^0)^2}{a^0}\Big) + G^{m+j}_{z_1^j}(z)(z_2^j - \beta^j\sigma^j) + \\ + G^{m+j}_{z_2^j}(z)((\sigma^j)^2) + G^{m+j}_{z_2^{j+1}}(z)((\sigma^{j+1})^2) = z_1^0 + z_2^0 - z_1^j - \frac{1}{2}(\beta^j)^2. \tag{2.104}$$

If we consider the function:

$$G^{m+j}(z) = -\frac{z_1^0 + \frac{1}{2}z_2^0}{a^0} - \frac{z_2^{j+1}z_1^j}{(\sigma^{j+1})^2} + \frac{(z_2^j)^2 z_2^{j+1}}{2(\sigma^{j+1})^2(\sigma^j)^2} - \frac{(z_2^j)^3}{6(\sigma^j)^4} + \\ - \frac{(z_2^{j+1})^2\beta^j\sigma^j}{2(\sigma^{j+1})^4} - \frac{z_2^{j+1}}{(\sigma^{j+1})^2}\frac{1}{2}\Big((\beta^j)^2 - \frac{(\sigma^0)^2}{(a^0)^2}\Big), \tag{2.105}$$

the condition (2.104) is satisfied, indeed if $j = 1, \dots, m-1$:

$$\begin{cases} G^{m+1}_{z_1^0}(z)\Big(-a^0 z_1^0 + \frac{(\sigma^0)^2}{a^0}\Big) = z_1^0 - \frac{(\sigma^0)^2}{(a^0)^2}, \\ G^{m+j}_{z_2^0}(z)\Big(-2a^0 z_2^0 - \frac{(\sigma^0)^2}{a^0}\Big) = z_2^0 + \frac{(\sigma^0)^2}{2(a^0)^2}, \\ G^{m+j}_{z_1^j}(z)(z_2^j - \beta^j\sigma^j) = -\frac{z_2^{j+1}z_2^j}{(\sigma^{j+1})^2} + \frac{z_2^{j+1}\beta^j\sigma^j}{(\sigma^{j+1})^2}, \\ G^{m+j}_{z_2^j}(z)((\sigma^j)^2) = \frac{z_2^j z_2^{j+1}}{(\sigma^{j+1})^2} - \frac{(z_2^j)^2}{2(\sigma^j)^2}, \\ G^{m+j}_{z_2^{j+1}}(x)((\sigma^{j+1})^2) = -z_1^j + \frac{(z_2^j)^2}{2(\sigma^j)^2} - \frac{z_2^{j+1}\beta^j\sigma^j}{(\sigma^{j+1})^2} - \frac{1}{2}\Big((\beta^j)^2 - \frac{(\sigma^0)^2}{(a^0)^2}\Big), \end{cases} \tag{2.106}$$

the condition (2.104) is

$$G^{m+j}_z(z)\eta(z) = z_1^0 - \frac{(\sigma^0)^2}{(a^0)^2} + z_2^0 + \frac{(\sigma^0)^2}{2(a^0)^2} - \frac{z_2^{j+1}z_2^j}{(\sigma^{j+1})^2} + \frac{z_2^{j+1}\beta^j\sigma^j}{(\sigma^{j+1})^2} + \\ + \frac{z_2^j z_2^{j+1}}{(\sigma^{j+1})^2} - \frac{(z_2^j)^2}{2(\sigma^j)^2} - z_1^j + \frac{(z_2^j)^2}{2(\sigma^j)^2} - \frac{z_2^{j+1}\beta^j\sigma^j}{(\sigma^{j+1})^2} - \frac{1}{2}\Big((\beta^j)^2 - \frac{(\sigma^0)^2}{(a^0)^2}\Big).$$

Whereas, the conditions for the function $G^{2m}$ are the same of (2.106), but $z_2^{j+1} \to z_2^1$. Now, it is sufficient to find conditions for the functions $G^{m+j}$ and the vector field $\xi$. In particular, the conditions which has to be respected is:

$$G^{m+j}_z(z)\xi(z) = \beta^j, \quad j = 1, \dots, m,$$

where $G^{m+1}$ is described in (2.105). The conditions is explicitly:

$$-\frac{\sigma^0}{a^0} - \frac{z_2^{j+1}\sigma^j}{(\sigma^{j+1})^2} = \beta^m, \quad j = 1, \dots, m,$$

which has no solutions. In conclusion, it seems very difficult to find conditions which guarantee the consistency also for the components associated with the spreads without adding new parameters. Hence, as in Proposition 2.3.1, only the following result is proved:

**Proposition 2.3.7.** *If we consider the model $\mathcal{M}$ for the post-crisis interest rates market, given by the Hull-White model* (2.59) *for the OIS forward rate equations and the Ho-Lee model* (2.41) *for the forward rate equations associated with the LIBOR rates and the parameterized family $\mathcal{G}$ given by the function $G : \mathbb{R}^{3m+2} \to \hat{\mathcal{H}}$,*

$$G(\widetilde{z}) = \begin{pmatrix} z_1^0 e^{-a^0 x} + z_2^0 e^{-2a^0 x} \\ z_1^1 + z_2^1 x \\ \vdots \\ z_1^m + z_2^m x \\ u^1 \\ \vdots \\ u^m \end{pmatrix}, \tag{2.107}$$

*where $\widetilde{z} = (z, u^1, \ldots, u^m)^*$ and $z = (z_1^0, z_2^0, z_1^1, z_2^1, \ldots, z_1^m, z_2^m)^*$, therefore the couple $(\mathcal{M}, \mathcal{G})$ is consistent.*

### 2.3.8 Vector Brownian motion examples

In this subsection we aim at describing the case where the Brownian motion $W$, which drives the dynamics, is characterized by different correlation structures. First, we will consider the case in which the Brownian motion is 1-dimensional and the volatility term is the same for each forward rate equation, afterwords we will analyse the case of $d$-dimensional Brownian motion and a volatility term $\hat{\sigma}$ such that each forward rate equation is driven by a specific Brownian motion independent from the others.

#### Common volatility for all forward rate equation

We consider the Hull-White model introduced in (2.59). For each forward rate equation, we consider a 1-dimensional Brownian motion and a volatility term given by:

$$\sigma^j(x) = \sigma e^{-ax}, \quad \forall\, j = 0, \ldots, m,$$

where $\sigma, a > 0$. The volatilities of log-spot spread processes are given by a constant $\beta^j = \beta$, for each $j = 1, \ldots, m$. In particular, there is one 1-dimensional Brownian motion $W$ which drives every equation.

We check the consistency of this model coupled with a parameterized forward rate family $\mathcal{G}$ determined by the function $G$ introduced in (2.74). Differently from the previous examples, we can use less parameters. Indeed, if we introduce the following family:

$$G^0(z) = z_1 e^{-ax} + z_2 e^{-2ax}, \tag{2.108}$$

$$G^j(z) = (z_3 + z_1) e^{-ax} + z_2 e^{-2ax}, \quad j = 1, \dots, m, \tag{2.109}$$

the conditions which guarantee the consistency (2.16) and (2.17) are given by:

$$\begin{aligned}
\mu^0(G(z))(x) =& \mathbf{F}G^0(z)(x) + \frac{\sigma^2}{a} e^{-ax}\Big[1 - e^{-ax}\Big] = (G^0_z(z)\eta)(x),\\
\mu^j(G(z))(x) =& \mathbf{F}G^j(z)(x) + \frac{\sigma^2}{a} e^{-ax}\Big[1 - e^{-ax}\Big] - \beta^\sigma e^{-ax} = (G^j_z(z)\eta(z))(x),\\
\sigma^j(G(z)) =& (G^j_z(z)\xi(z)),
\end{aligned}$$

for suitable vectors $\xi, \eta \in \mathbb{R}^3$ and for each $j = 0, \dots, m$.

Explicitly, the conditions on the drift become:

$$\begin{aligned}
&- az_1 e^{-ax} - 2az_2 e^{-2ax} + \frac{\sigma^2}{a} e^{-2ax}(e^{ax} - 1) = \eta_1 e^{-ax} + \eta_2 e^{-2ax},\\
&- a(z_1 + z_3) e^{-ax} - 2az_2 e^{-2ax} + \frac{\sigma^2}{a} e^{-2ax}(e^{ax} - 1) - \beta\sigma e^{-ax} = \eta_1 e^{-ax} + \eta_2 e^{-2ax} + \eta_3 e^{-ax}.
\end{aligned}$$

In the following of the subsection, we will omit the dependence on the $z$ variable for the vector fields $\eta$ and $\xi$. The previous equations are satisfied imposing that:

$$\eta = (-az_1 + \frac{\sigma^2}{a}, -2az_2 - \frac{\sigma^2}{a}, -az_3 - \beta\sigma)^*. \tag{2.110}$$

On the other hand, the conditions on the volatility terms become:

$$\sigma e^{-ax} = \xi_1 e^{-ax} + \xi_2 e^{-2ax}, \tag{2.111}$$

$$\sigma e^{-ax} = \xi_1 e^{-ax} + \xi_2 e^{-2ax} + \xi_3 e^{-ax} \tag{2.112}$$

which is satisfied by the vector

$$\xi = (\sigma, 0, 0)^*. \tag{2.113}$$

At this point, we have to determine functions $G^{m+j} : \mathbb{R}^n \longrightarrow \mathbb{R}$ for $j = 1, \dots, m$ where $n$ is an opportune natural number, such that:

$$\begin{aligned}
G^{m+j}_z(z)\eta(z) =& \mathbf{B}G^0 - \mathbf{B}G^j - \frac{1}{2}\beta^2\\
=& z_1 + z_2 - (z_3 + z_1) - z_2 - \frac{1}{2}\beta^2 = -z_3 - \frac{1}{2}\beta^2, \qquad j = 1, \dots, m.
\end{aligned}$$

In particular, we can observe that the previous condition does not depend on $j$. This implies that it sufficient to add one parameter, which will be denote by $u$ in analogy to the examples of the previous sections.

Explicitly, we introduce the following finite-dimensional vector: $z = (z_1, z_2, z_3, u)^* \in \mathbb{R}^4$. We introduce the following vector fields:

$$\begin{cases} \eta(z) = \begin{pmatrix} -az - 1 + \frac{\sigma^2}{a} & -2az_2 - \frac{\sigma^2}{a} & -az_3 - \beta\sigma & -z_3 - \frac{1}{2}\beta^2 \end{pmatrix}, \\ \xi(z) = \begin{pmatrix} \sigma & 0 & 0 & \beta \end{pmatrix}, \end{cases}$$

and the function $G : \mathbb{R}^4 \longrightarrow \hat{\mathcal{H}}$:

$$G(z) = \begin{pmatrix} z_1 e^{-ax} + z_2 e^{-2ax} \\ z_1 e^{-ax} + z_2 e^{-2ax} \\ \vdots \\ z_1 e^{-ax} + z_2 e^{-2ax} \\ u \\ \vdots \\ u \end{pmatrix}. \tag{2.114}$$

Therefore, $G_z^j(z)\eta(z) = \mu^j(G(z))$ and $G_z^j(z)\xi(z) = \sigma^j(G(z))$ for every $j = 0, \ldots, m$. Moreover, the coordinates related to the spreads satisfy:

$$\begin{cases} G_z^{m+j}(z)\eta(z) = -z_3 - \frac{1}{2}\beta^2, & j = 1, \ldots, m; \\ G_z^{m+j}(z)\xi(z) = \beta, & j = 1, \ldots, m. \end{cases}$$

In conclusion, the following result is proved:

**Proposition 2.3.8.** *We consider the model $\mathcal{M}$ determined by the volatility vector:*

$$\sigma(x) = \begin{pmatrix} \sigma e^{-ax} & \cdots & \sigma e^{-ax} \end{pmatrix},$$

*and the family $\mathcal{G} = Im[G]$, where $G$ is given by* (2.114). *Hence, the couple $(\mathcal{M}, \mathcal{G})$ is consistent.*

### Independent Brownian motion for each forward rate equation

In this paragraph, we aim at analysing the case of a $\mathbb{R}^d$-valued Brownian motion $W$, where $d \geq m$. We introduce the volatility term of each forward rate equation as follows:

$$\sigma^j(x) = (0, \cdots, 0, \overbrace{\sigma^j e^{-a^j x}}^{\text{j-th component}}, 0, \cdots, 0), \quad j = 1, \ldots, m, \tag{2.115}$$

and similarly:

$$\beta^j = (0,\cdots,0,\overbrace{\beta^j}^{\text{i-th component}},0,\cdots,0),\quad j=1,\ldots,m, \tag{2.116}$$

where $\sigma^j, \beta^j, a^j$ are positive constants. Under this assumption, we have that each component of the forward rate equation is driven by a component of the vector Brownian motion $W$. As a consequence, all forward rates are independent processes. If we want to check the consistency condition for the model $\mathcal{M}$ determined by the previous volatility terms and the parameterized families introduced in the previous subsections, we can make the following observation. The term $\sigma^j \mathbf{H} \sigma^j$ is given by:

$$\sigma^j(x)\mathbf{H}\sigma^j(x) = \begin{pmatrix} 0 & 0 & \cdots & \sigma^j e^{-a^j x} & 0 & \ldots & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \int_0^x \sigma^j e^{-a^j s} ds \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \sigma^j e^{-a^j x} \int_0^x \sigma^j e^{-a^j s} ds,$$

which is equivalent to the one introduced in (2.120) for the drift condition of the Hull-White model. Through the same computations we can observe that $\sigma^j \beta^{j*} = \sigma^j e^{-a^j x} \beta^j$ as the term introduced in (2.67) for the 1-dimensional Brownian motion case. On the basis of these considerations, we can conclude that the consistency results associated with this model corresponds exactly to the ones demonstrated in sections 2.3.5, 2.3.6, 2.3.7.

#### Independent Brownian motion for each forward rate equation and common volatility term

In this paragraph, we construct a trade off between the previous two examples. In particular, we consider the case of a $\mathbb{R}^d$-valued Brownian motion $W$, where $d \geq 0$. The volatility term of each forward rate equation is defined as follows:

$$\sigma^j(x) = (0,\cdots,0,\sigma e^{-ax},0,\cdots,0), \tag{2.117}$$

and

$$\beta^j = (0,\cdots,0,\beta,0,\cdots,0), \tag{2.118}$$

where $\sigma, a, \beta$ are positive constants. As in the last example, for each $i = 1,\ldots,d$ the term:

$$\sigma^j(x)\mathbf{H}\sigma^j(x) = \sigma e^{-ax} \int_0^x \sigma e^{-as} ds,$$

and the term $\sigma^j\beta^{j*} = \sigma\beta$ for every $j$.

In particular we obtain the same structure of the first paragraph of this subsection.

In conclusion, we have proved the following proposition:

**Proposition 2.3.9.** *If we consider the vector volatility term associated with the Hull-White model, the following statements hold:*

1. *The Brownian motion $W$ is $R^d$-valued, where $d \geq m$, and the volatility term which determines $\mathcal{M}$ is given by equations* (2.115) *and* (2.116)*. Then, the consistency conditions are analogous to Sections 2.3.5, 2.3.6, 2.3.7.*

2. *If the Brownian motion $W$ is $R^d$-valued, where $d \geq m$, and the volatility term which determines $\mathcal{M}$ is given by equations* (2.117) *and* (2.118)*, then we have the consistency between the model $\mathcal{M}$ and the forward family described by the function $G$ given by* (2.108) *and* (2.109).

### The general case

We consider a $d$-dimensional Brownian motion which drives a forward rate model $\mathcal{M}$ determined by the following volatility term:

$$\hat{\sigma}(\hat{r}_t) = \begin{pmatrix} \sigma_1^0 e^{-a_1^0 x} & \sigma_2^0 e^{-a_2^0 x} & \cdots & \sigma_d^0 e^{-a_d^0 x} \\ \sigma_1^1 e^{-a_1^1 x} & \sigma_2^1 e^{-a_2^1 x} & \cdots & \sigma_d^1 e^{-a_d^1 x} \\ \vdots & \vdots & & \vdots \\ \sigma_1^m e^{-a_1^m x} & \sigma_2^m e^{-a_2^m x} & \cdots & \sigma_d^m e^{-a_d^m x} \\ \beta_1^1 & \beta_2^1 & \cdots & \beta_d^1 \\ \vdots & \vdots & & \vdots \\ \beta_1^m & \beta_2^m & \cdots & \beta_d^m \end{pmatrix} \equiv \begin{pmatrix} \sigma^0 \\ \sigma^1 \\ \vdots \\ \sigma^m \\ \beta^1 \\ \vdots \\ \beta^m \end{pmatrix}, \tag{2.119}$$

where $\sigma_i^j, a_i^j$ and $\beta_i^j$ are positive real constants, for ever $i = 1, \ldots, d$ and $j = 0, \ldots, m$.

First, it is necessary to generalize the computation provided in (2.128), in order to manage the term $\sigma^j(t,x)\mathbf{H}\sigma^j(t,x)$, where $\sigma^j(t,x)$ is a vector. In particular, we obtain the following equivalence:

$$\begin{aligned} \sigma^j(t,x)\mathbf{H}\sigma^j(t,x) &= \begin{pmatrix} \sigma_1^j e^{-a_1^j x} & \sigma_2^j e^{-a_2^j x} & \cdots & \sigma_d^j e^{-a_d^j x} \end{pmatrix} \begin{pmatrix} \int_0^x \sigma_1^j e^{-a_1^j s} ds \\ \int_0^x \sigma_2^j e^{-a_2^j s} ds \\ \vdots \\ \int_0^x \sigma_d^j e^{-a_d^j s} ds \end{pmatrix} \\ &= \sum_{i=1}^d \frac{(\sigma_i^j)^2}{a_i^j} e^{-2a_i^j x}\left(e^{a_i^j x} - 1\right), \qquad j = 0, \ldots, m, \end{aligned} \tag{2.120}$$

whereas the term related to the presence of the spread on the LIBOR forward rate dynamics is given by:

$$\sigma^j(x)\beta^{i*} = \sum_{i=1}^{d} \sigma_i^j e^{-a_i^j x}\beta_i^j.$$

Hence, we analyse the problem of consistency between the model $\mathcal{M}$, previously defined and a suitable parameterized family. We recall the conditions on the drift and volatility terms, $\hat{\mu}(G(z)), \hat{\sigma}_i(G(z)) \in T_{G(z)}\mathcal{G}$, where $\mathcal{G} = Im[G]$ and for every $i = 1, \ldots, d$. The drift term is given by:

$$\hat{\mu}(G(z)) = \begin{pmatrix} \mathbf{F}G^0(z) + (\sigma^0\mathbf{H}\sigma^0)(z) \\ \mathbf{F}G^1(z) + (\sigma^1\mathbf{H}\sigma^1)(z) - \sum_{i=1}^d \sigma_i^1(G(z))\beta_i^1(G(z)) \\ \vdots \\ \mathbf{F}G^m(z) + (\sigma^m\mathbf{H}\sigma^m)(z) - \sum_{i=1}^d \sigma_i^m(G(z))\beta_i^m(G(z)) \\ \mathbf{B}G^0(z) - \mathbf{B}G^1(z) - \frac{1}{2}\sum_{i=1}^d (\beta_i^1)^2 \\ \vdots \\ \mathbf{B}G^0(z) - \mathbf{B}G^m(z) - \frac{1}{2}\sum_{i=1}^d (\beta_i^m)^2 \end{pmatrix}.$$

At this point, we consider the function defined on (2.74). We observe that it is not possible to use that function in order to have the consistency, because the element $\sigma^j\mathbf{H}\sigma^j$ involves the sum of $2d$ exponential terms. Hence, we propose the following generalization of the function introduced in (2.74):

$$G^j(z,x) = \sum_{i=1}^{d}\Big(z_i^j e^{-a_i^j x} + w_i^j e^{-2a_i^j x}\Big), \qquad j = 0, \ldots, m, \tag{2.121}$$

where the vector $z \in \mathbb{R}^{2d(m+1)}$ is introduced by the following notation:

$$z = (z_1^0, z_2^0, \ldots, z_d^0, w_1^0, \ldots, w_d^0, z_1^1, \ldots, \ldots, w_d^m)^*. \tag{2.122}$$

If we consider the previous function the existence of a vector field $\eta$ defined on the domain $\mathcal{Z} \subset \mathbb{R}^{2d(m+1)}$ such that $\mu^0(G(z))(x) = G_z^0(z,x)\eta(z)$ is satisfied. Indeed:

$$\begin{aligned} \mu^0(G(z)) &= \mathbf{F}G^0(z) + (\sigma^0\mathbf{H}\sigma^0)(z) \\ &= \sum_{i=1}^{d}\Big(-a_i^0 z_i^0 e^{-a_i^0 x} - 2a_i^0 w_i^0 e^{-2a_i^0 x}\Big) + \sum_{i=1}^{d}\frac{(\sigma_i^0)^2}{a_i^0}e^{-2a_i^0 x}\Big(e^{a_i^0 x} - 1\Big) \\ &= \sum_{i=1}^{d}\Big\{e^{-a_i^0 x}\Big[-a_i^0 z_i^0 + \frac{(\sigma_i^0)^2}{a_i^0}\Big] - e^{-2a_i^0}\Big[2a_i^0 w_i^0 + \frac{(\sigma_i^0)^2}{a_i^0}\Big]\Big\}, \end{aligned}$$

On the other hand, we compute the Fréchet derivative of $G^0(z)$ against the vector field $\eta$. In order to simplify the computation, we introduce the following notation for $\eta$:

$$\eta(z) = (\eta_{z_1^0}, \eta_{z_2^0}, \dots, \eta_{z_d^0}, \eta_{w_1^0}, \dots, \eta_{w_d^0}, \eta_{z_1^1}, \dots, \dots, \eta_{z_d^m}) \in \mathbb{R}^{2d(m+1)}.$$

In particular, the consistency condition is satisfied by:

$$\begin{cases} \eta_{z_i^0} = -a_i^0 z_i^0 + \frac{(\sigma_i^0)^2}{a_i^0}, & i = 1, \dots, d, \\ \eta_{w_i^0} = -2a_i^0 w_i^0 - \frac{(\sigma^0)^2}{a_i^0}, & i = 1, \dots, d, \end{cases}$$

We have not already provided conditions for $\eta_{z_i^j}, \eta_{w_i^j}$ for $j = 1, \dots, m$. In order to do this, we consider the other conditions for the drift term:

$$\mu^j(G(z)) = \mathbf{F}G^j(z) + \sum_{i=1}^d \frac{(\sigma_i^j)^2}{a_i^j} e^{-2a_i^j x}\left(e^{a_i^j x} - 1\right) - \sum_{i=1}^d \sigma_i^j e^{-a_i^j x}\beta_i^j = G_z^j(z)\eta(z).$$

Recalling the shape of the function $G^j$, $j = 1, \dots, m$ given by (2.121), the consistency condition on the drift is:

$$\begin{aligned} \sum_{i=1}^d \Big(\eta_{z_i^j}(z)e^{-a_i^j x} + \eta_{w_i^j}(z)e^{-2a_i^j x}\Big) &= \sum_{i=1}^d \Big(-a_i^j z_i^j e^{-a_i^j x} - 2a_i^j w_i^j e^{-2a_i^j x}\Big) + \\ &+ \sum_{i=1}^d \frac{(\sigma_i^j)^2}{a_i^j} e^{-2a_i^j x}\Big(e^{-a_i^j x} - 1\Big) - \sum_{i=1}^d \sigma_i^j e^{-a_i^j x}\beta_i^j \\ &= \sum_{i=1}^d \Big[e^{-a_i^j x}\Big(-a_i^j z_i^j + \frac{(\sigma_i^j)^2}{a_i^j} - \sigma_i^j\beta_i^j\Big) + e^{-2a_i^j x}\Big(-2a_i^j w_i^j - \frac{(\sigma_i^j)^2}{a_i^j}\Big)\Big], \end{aligned}$$

which implies that:

$$\begin{cases} \eta_{z_i^j}(z) = -a_i^j z_i^j + \frac{(\sigma_i^j)^2}{a_i^j} - \sigma_i^j \beta_i^j, & i = 1, \dots, d, \\ \eta_{w_i^j}(z) = -2a_i^j w_i^j - \frac{(\sigma_i^j)^2}{a_i^j}, & i = 1, \dots, d, \end{cases}$$

for every $j = 1, \dots, m$.

On the other hand, the consistency condition for the volatility is equivalent to the existence of a vector field $\xi_i(z) \in \mathbb{R}^{2d(m+1)}$, $i = 1, \dots, d$, such that $\sigma_i^j(G(z)) = G_z^{m+i}(z)\xi_i(z)$. In particular, we recall that $\sigma_i^j(G(z)) = \sigma_i^j e^{-a_i^j x}$, hence, for each $i = 1, \dots, d$, we can choose

$$\begin{cases} \xi_{i,z_i^j}(z) = \sigma_i^j, & j = 0, \dots, m, \\ \xi_{i,\alpha}(z) = 0, & \text{otherwise.} \end{cases}$$

It necessary to control the conditions on the coordinates related to the spreads. In particular, we have to determine the shape of $G^{m+j}$ such that $G_z^{m+j}(z)\eta(z) = \mu^{m+j}(G)$ and $G^{m+j}z\xi_i(z) = \beta_i^j(z)$ for every $j = 1, \dots, m$.

We base on Proposition 2.3.5 and we exploit the linearity of the Fréchet derivative in order to construct a suitable function $G^{m+j}$. Explicitly, the conditions on the drift term are:

$$\begin{aligned} G_z^{m+j}(z)\eta(z) =& \mathbf{B}G^0(z) - \mathbf{B}G^j(z) - \frac{1}{2}\sum_{i=1}^{d}(\beta_i^j)^2 \\ =& \sum_{i=1}^{d}\big[z_i^0 + w_i^0 - z_i^j - w_i^j - \frac{1}{2}(\beta_i^j)^2\big]. \end{aligned} \tag{2.123}$$

We introduce the family:

$$G^{m+j}(z) = \sum_{i=1}^{d}\Big(-\frac{z_i^0 + \frac{1}{2}w_i^0}{a_i^0} + \frac{z_i^j + \frac{1}{2}w_i^j}{a_i^j}\Big). \quad j = 1, \dots, m.$$

By the previous definition, $G_z^{m+j}(z)\eta(z)$ becomes:

$$\begin{aligned} G_z^{m+j}(z)\eta(z) &= \sum_{i=1}^{d}\Big(z_i^0 - \frac{(\sigma_i^0)^2}{(a_i^0)^2} + w_i^0 + \frac{(\sigma_i^0)^2}{2(a_i^0)^2} - z_i^j + \frac{(\sigma_i^j)^2}{(a_i^j)^2} - \frac{\sigma_i^j}{a_i^j}\beta_i^j - w_i^j - \frac{(\sigma_i^j)^2}{2(a_i^j)^2} \\ &= \sum_{i=1}^{d}(z_i^0 + w_i^0 - z_i^j - w_i^j) - \frac{1}{2}\sum_{i=1}^{d}\left[\frac{(\sigma_i^0)^2}{(a_i^0)^2} + 2\frac{\sigma_i^j}{a_i^j}\beta_i^j - \frac{(\sigma_i^j)^2}{(a_i^j)^2}\right], \end{aligned}$$

hence, condition (2.123) is equivalent to:

$$\begin{aligned} &\sum_{i=1}^{d}(\beta_i^j)^2 = \sum_{i=1}^{d}\Big[\frac{(\sigma_i^0)^2}{(a_i^0)^2} - \frac{(\sigma_i^j)^2}{(a_i^j)^2} + 2\frac{\sigma_i^j}{a_i^j}\beta_i^j\Big] \\ &\sum_{i=1}^{d}\left[(\beta_i^j)^2 - 2\frac{\sigma_i^j}{a_i^j}\beta_i^j + \left(\frac{(\sigma_i^j)^2}{(a_i^j)^2} - \frac{(\sigma_i^0)^2}{(a_i^0)^2}\right)\right] = 0. \end{aligned} \tag{2.124}$$

The following condition on the matrix $\beta = (\beta_i^j)_{j=1,\dots,m;\ i=1,\dots,d}$ is a sufficient condition, such that (2.124) holds:

$$\beta_i^j = \frac{\sigma_i^j}{a_i^j} \pm \frac{\sigma_i^0}{a_i^0}. \tag{2.125}$$

If we test the consistency condition on the volatility term of the spreads, we obtain:

$$
\begin{aligned}
\beta_i^j =& G_z^{m+j}(z)\xi_i(z) = \sum_{\bar{j}=1}^{m}\sum_{\bar{i}=1}^{d} G_{z_i^{\bar{j}}}^{m+j}(z)\xi_i(z) \\
=& \sum_{\bar{j}=1}^{m}\sum_{\bar{i}=1}^{d}\Big(-\frac{\delta_{i\bar{i}}}{a_{\bar{i}}^0}\xi_{i,z_i^0} + \frac{\delta_{j\bar{j}}\delta_{i\bar{i}}}{a_{\bar{i}}^{\bar{j}}}\xi_{i,z_i^{\bar{j}}}\Big) = -\frac{\xi_{i,z_i^0}}{a_i^0} + \frac{\xi_{i,z_i^j}}{a_i^j} \\
=& -\frac{\sigma_i^0}{a_i^0} + \frac{\sigma_i^j}{a_i^j},
\end{aligned}
\tag{2.126}
$$

where $\delta_{hk}$ stands for the Kronecker delta between the indeces $h$ and $k$. Hence, by (2.125) and (2.126) the condition on $\beta_i^j$ is:

$$
\beta_i^j = -\frac{\sigma_i^0}{a_i^0} + \frac{\sigma_i^j}{a_i^j}, \quad j = 1, \dots, m, \; i = 1, \dots, d. \tag{2.127}
$$

In conclusion the following proposition is proved:

**Proposition 2.3.10.** *If we consider the model $\mathcal{M}$ for the fixed-income market determined by the volatility term described in* (2.119) *where the volatility terms of the spreads satisfy* (2.127) *and the family $\mathcal{G} = Im[G]$ given by the function $G : \mathbb{R}^{2d(m+1)} \longrightarrow \hat{\mathcal{H}}$*

$$
G(z) = \begin{pmatrix}
\sum_{i=1}^d \Big(z_i^0 e^{-a_i^0 x} + w_i^0 e^{-2a_i^0 x}\Big) \\
\vdots \\
\sum_{i=1}^d \Big(z_i^m e^{-a_i^m x} + w_i^m e^{-2a_i^m x}\Big) \\
\sum_{i=1}^d \Big(-\frac{z_i^0 + \frac{1}{2}w_i^0}{a_i^0} + \frac{z_i^1 + \frac{1}{2}w_i^1}{a_i^1}\Big) \\
\vdots \\
\sum_{i=1}^d \Big(-\frac{z_i^0 + \frac{1}{2}w_i^0}{a_i^0} + \frac{z_i^m + \frac{1}{2}w_i^m}{a_i^m}\Big)
\end{pmatrix},
$$

*where $z$ is given by* (2.119)*, then the couple $(\mathcal{M}, \mathcal{G})$ is consistent.*

*On the other hand, recalling the same strategy of Proposition 2.3.3, we have that the model $\mathcal{M}$ is consistent with the family $\widetilde{\mathcal{G}}$ where $\widetilde{\mathcal{G}} = Im[\widetilde{G}]$ with the mapping $\widetilde{G} : \mathbb{R}^{(2d+1)m+2d} \longrightarrow \hat{\mathcal{H}}$ given by*

$$
\widetilde{G}(\widetilde{z}) = \begin{pmatrix}
\sum_{i=1}^d \Big(z_i^0 e^{-a_i^0 x} + w_i^0 e^{-2a_i^0 x}\Big) \\
\vdots \\
\sum_{i=1}^d \Big(z_i^m e^{-a_i^m x} + w_i^m e^{-2a_i^m x}\Big) \\
u^1 \\
\vdots \\
u^m
\end{pmatrix},
$$

*where* $\widetilde{z} = (z, u^1, \ldots, u^m)^*$.

# 2.4 Appendix

## 2.4.1 Hull-White forward rate

In this section we will provide the proof of Lemma 2.3.2.

*Proof.* Let us consider the following SDE:

$$df_t(T) = \alpha_t(T)dt + \sigma e^{-a(T-t)}dW_t.$$

If $f_t(T)$ represents the forward rate associated with the short rate $r_t$, solution of the Hull-White equation, $f_t(T)$ has to satisfy the HJM-drift condition (A.18). In particular:

$$\begin{aligned}\alpha_t(T) &= \sigma e^{-a(T-t)} \int_t^T \sigma e^{-a(T-s)}ds = \sigma e^{-a(T-t)}\sigma e^{-aT} \int_t^T e^{as}ds \\ &= \frac{\sigma^2}{a}\Big[e^{-a(T-t)} - e^{-2a(T-t)}\Big] = \frac{\sigma^2}{a}e^{-a(T-t)}\Big\{1 - e^{-a(T-t)}\Big\}.\end{aligned} \tag{2.128}$$

Then:

$$\begin{aligned}f_t(T) &= f_0(T) + \frac{\sigma^2}{a}\int_0^t e^{-a(T-s)}ds - \frac{\sigma^2}{a}\int_0^t e^{-2a(T-s)}ds + \sigma\int_0^t e^{-a(T-s)}dW_s \\ &= f_0(T) + \frac{\sigma^2}{a}e^{-aT}\int_0^t e^{as}ds - \frac{\sigma^2}{a}e^{-2aT}\int_0^t e^{2as}ds + \sigma\int_0^t e^{-a(T-s)}dW_s \\ &= f_0(T) + \frac{\sigma^2}{a^2}e^{-aT}\Big(e^{at} - 1\Big) - \frac{\sigma^2}{2a^2}e^{-2aT}\Big(e^{2at} - 1\Big) + \sigma\int_0^t e^{-a(T-s)}dW_s\end{aligned}$$

Computing the short rate:

$$r_t = f_t(t) = f_0(t) + \frac{\sigma^2}{a^2}\Big(1 - e^{-at}\Big) - \frac{\sigma^2}{2a^2}\Big(1 - e^{-2at}\Big) + \sigma\int_0^t e^{-a(t-s)}dW_s,$$

and differentiating it:

$$dr_t = \Big\{\frac{\partial}{\partial T}f_0(t) + \frac{\sigma^2}{a}e^{-at} - \frac{\sigma^2}{a}e^{-2at} + \sigma\int_0^t -ae^{-a(t-s)}dW_s\Big\}dt + \sigma dW_t.$$

Recall now that:

$$\sigma\int_0^t e^{-a(t-s)}dW_s = r_t - \frac{\sigma^2}{a^2}\Big[\Big(1 - e^{-at}\Big) - \frac{1}{2}\Big(1 - e^{-2at}\Big)\Big] - f_0(t),$$

In conclusion:

$$\begin{aligned}
dr_t =&\Big\{\frac{\partial}{\partial T}f_0(t)+\frac{\sigma^2}{a}e^{-at}-\frac{\sigma^2}{a}e^{-2at}-ar_t+\frac{\sigma^2}{a}\Big[\Big(1-e^{-at}\Big)-\frac{1}{2}\Big(1-e^{-2at}\Big)\Big]-f_0(t)\Big\}dt+\sigma dW_t\\
=&\Big\{\frac{\partial}{\partial T}f_0(t)-\frac{1}{2}\frac{\sigma^2}{a}e^{-2at}-ar_t+\frac{\sigma^2}{2a}-f_0(t)\Big\}dt+\sigma dW_t\\
=&\Big\{\frac{\partial}{\partial T}f_0(t)+\frac{\sigma^2}{2a}\Big[1-e^{-2at}\Big]-f_0(t)-ar_t\Big\}dt+\sigma dW_t.
\end{aligned}$$

The previous SDE is the Hull-White equation, if

$$\Phi(t)=\frac{\partial}{\partial T}f_0(t)+\frac{\sigma^2}{2a}\Big[1-e^{-2at}\Big]-f_0(t).$$

□

# Chapter 3

# Finite-dimensional Realizations

In this chapter we will exploit the concept of invariance developed in Chapter 2, in order to understand if the solution of the system (1.33) can be described as the image of a process, whose dynamics given by a finite-dimensional SDE. Moreover, if it is the case, we will provide a strategy to construct this process and the mapping which associates it to the forward rate $\hat{r}$. To this effect, we will exploit the geometric theory developed in Appendix B, applying it to the geometric interpretation of equation (1.33) described in Section 2.1. The general conditions will be applied to the study of particular cases: first, we will analyse the case of deterministic volatility (constant in the space $\hat{\mathcal{H}}$), then we will study the case of constant direction volatility.

The main references for this chapter are represented by [4], [5] and [21].

## 3.1 The general result

In the previous chapter we developed a geometric interpretation of system (1.33), representing the infinite-dimensional system of SDEs as a unique SDE of the form:

$$d\hat{r}_t = \hat{\mu}(\hat{r}_t)dt + \hat{\sigma}(\hat{r}_t) \circ dW_t, \tag{3.1}$$

$$\hat{r}_0 = \hat{r}^M. \tag{3.2}$$

We have seen that, since $\hat{\mu} : U_{\hat{r}^M} \longrightarrow \hat{\mathcal{H}}$ and $\hat{\sigma} : U_{\hat{r}^M} \longrightarrow \hat{\mathcal{H}}^d$ are smooth functions, they can be interpreted as local vector fields on $\hat{\mathcal{H}}$, as described in Definition B.1.14, denoting by $U_{\hat{r}^M}$ a neighborhood of $\hat{r}^M$ in $\hat{\mathcal{H}}$.

We define as follows the main concept of this chapter, the finite-dimensional realization. First of all we need the following concept:

**Definition 3.1.1.** *We say that $\hat{r}$, given an initial point $\hat{r}^M$, has the* local representation *around $\hat{r}^M$ given by a function $G : \mathcal{Z} \subseteq \mathbb{R}^n \longrightarrow \hat{\mathcal{H}}$ for a suitable $n$ and a finite-dimensional stochastic process $Z_t$ defined on $\mathcal{Z}$, if there exists a strictly positive stopping time $\tau(\hat{r}^M)$ such that $\hat{r}_t = G(Z_t)$ for each $t \in [0, \tau(\hat{r}^M))$.*

Now we say that a the system (1.33) possesses finite-dimensional realizations if it has $n$-dimensional realizations, defined as follows, for a suitable $n$:

**Definition 3.1.2.** *We say that* (3.1) *has a $n$-dimensional realization if for each $\hat{r}_0^M \in \hat{\mathcal{H}}$, there exists $z_0 \in \mathbb{R}^n$ and $(d+1)-$smooth vector fields $a, b_1 \dots, b_d$, defined on a neighborhood of $z_0$ denoted with $\mathcal{Z}$ and a smooth mapping $G : \mathcal{Z} \longrightarrow \hat{\mathcal{H}}$, such that $\hat{r}$ has the local representation:*

$$\hat{r}_t = G(Z_t),$$

*where*

$$\begin{aligned} dZ_t &= a(Z_t)dt + b(Z_t) \circ dW_t, \\ Z_0 &= z_0, \end{aligned}$$

*where $b := (b_1, \dots, b_d)$.*

**Remark 3.1.3.** *The definition of finite-dimensional realization is strictly related to the concept of $\hat{r}$-invariance. In particular, the existence of a finite-dimensional realization for a model described by* (3.1) *is equivalent to the existence of a $\hat{r}$-invariant $\mathcal{G}$ for $\hat{r}$.*

By the invariance Theorem 2.2.8, given a forward rate model $\mathcal{M}$, a submanifold $\mathcal{G} \subset \hat{\mathcal{H}}$ is such that the couple $(\mathcal{M}, \mathcal{G})$ is invariant if and only if $\hat{\mu}(G(z)), \hat{\sigma}(G(z)) \in T_{G(z)}\mathcal{G}$, for each $G(z) \in U$, where $U$ is a neighborhood of $\hat{r}^M$ and $\hat{r}^M \in \mathcal{G}$.

The condition $\hat{\mu}(G(z)), \hat{\sigma}(G(z)) \in T_{\hat{r}^M}\mathcal{G}$ is equivalent to assume that the distribution (see Definition B.1.13) generated by $\hat{\mu}$ and $\hat{\sigma}$ is a subset of $T\mathcal{G}$, where $T\mathcal{G}$ is the tangent bundle of $\mathcal{G}$ (B.1.12). In other words, we are looking for a tangential submanifold $\mathcal{G}$ of the distribution $F = Span\{\hat{\mu}, \hat{\sigma}\}$. We recall Theorem B.3.2, which guarantees the existence of a tangential sub-manifold for a smooth distribution $F$ if and only if $F$ is involutive (see Definition B.1.22). We will use the Frobenius theorem (Theorem B.2.4) in order to construct a tangential submanifold when the distribution $F$ generated by $\hat{\mu}$ and $\hat{\sigma}$ is involutive.

Unfortunately, given equation (3.1), do not exist a priori conditions under which the distribution generated by $\hat{\mu}, \hat{\sigma}$ is involutive. As we observed in Appendix $B$, given a distribution $F$ generated by $n$-vector fields, the smallest involutive distribution which contains $F$ is the Lie algebra of $F$ (see Definition B.3.4). Therefore, denoting by $\mathcal{L} := \{\hat{\mu}, \hat{\sigma}_1, \dots \hat{\sigma}_d\}_{LA}$ the Lie algebra of $F$, we obtain that the existence of finite-dimensional realizations is equivalent to the existence of a finite-dimensional tangential submanifold. By Theorem B.3.2, this is equivalent to the condition:

$$dim[\mathcal{L}] = dim\{\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_d\}_{LA} < +\infty. \tag{3.3}$$

The analogous condition for the pre-crisis environment is provided in [5][Theorem 4.2].

If the above condition holds, a finite-dimensional realization can be provided. To this end, we provide a strategy based on [5][Chapter 5]. This strategy is described in the following steps:

1. Choose a finite number of vector fields $\xi_1, \ldots \xi_n$, which span $\{\hat{\mu}, \hat{\sigma}_1, \ldots \hat{\sigma}_d\}_{LA}$;

2. Compute the invariant manifold

$$G(z_1, \ldots, z_n) = e^{\xi_n z_n} \cdots e^{\xi_1 z_1} \hat{r}^M,$$

where $e^{\xi_n z_n}$ denotes the integral curve of $\xi_n$ at time $z_n$ (as in Proposition B.3.3);

3. Through the mapping $G$ defined in the previous step, define the state space process $Z$, such that $\hat{r} = G(Z)$. $Z$ is a $\mathbb{R}^n$- valued process determined by:

$$dZ_t = a(Z_t)dt + b(Z_t) \circ dW_t, \tag{3.4}$$

where

$$G_* a(G(z)) = G_z(z)a(z) = \hat{\mu}(G(z)), \tag{3.5}$$
$$G_* b_j(G(z)) = G_z(z)b_j(z) = \hat{\sigma}_j(G(z)), \quad j \in \{1, \ldots, d\}, \tag{3.6}$$

where, as before, we have denoted with $G_z(z)$ the Fréchet derivative of $G$. We recall that the symbol $G_* a$ stands for the $G$-related vector field to the vector field $a$, defined on $\hat{\mathcal{H}}$, introduced in Definition B.1.20.

The uniqueness of $a$ and $b$ is guaranteed since $G$ respects Assumption 2.2.1, then it is a local diffeomorphism. Therefore applying Definition B.1.20, there exists a unique vector field defined on $\mathcal{Z}$ $a$ for $\hat{\mu}$ and $b_i$ for $\hat{\sigma}_i$ for every $i = 1, \ldots, d$ such that conditions (3.5) and (3.6) are satisfied.

In the following section we will analyze the problem of the existence of finite-dimensional realizations for model whose volatility has a certain structure.

## 3.2 Constant volatility

We now examine the case in which the volatility vector field $\hat{\sigma}(\hat{r})$ is constant. In particular, $\hat{\sigma}$ does not depend on $\hat{r}$. Equivalently, this assumption means that $\sigma^0, \sigma^1, \ldots, \sigma^m$ are all constant vector fields and $\beta^1, \ldots, \beta^m$ are constant on $\mathbb{R}^d$. In this section we generalize the results provided in [21][Section 3].

The logarithm of the spread process associated with the tenor $\delta_i$ is given by:

$$dY_t = \Big\{\mathbf{B}r^0_t - \mathbf{B}r^i_t - \frac{1}{2}||\beta^i||^2\Big\}dt + \beta^i dW_t,$$

whereas the drift and volatility terms of equation (3.1) are respectively given by:

$$\hat{\mu}(\hat{r}) = \begin{pmatrix} \mathbf{F}r^0 + \sigma^0\mathbf{H}\sigma^0 \\ \mathbf{F}r^1 + \sigma^1\mathbf{H}\sigma^1 - \beta^1\sigma^1 \\ \vdots \\ \mathbf{F}r^m + \sigma^m\mathbf{H}\sigma^m - \beta^m\sigma^m \\ \mathbf{B}r^0 - \mathbf{B}r^1 - \frac{1}{2}||\beta^1||^2 \\ \vdots \\ \mathbf{B}r^0 - \mathbf{B}r^m - \frac{1}{2}||\beta^m||^2 \end{pmatrix},$$

$$\hat{\sigma}(\hat{r}) = \begin{pmatrix} \sigma^0 \\ \vdots \\ \sigma^m \\ \beta^1 \\ \vdots \\ \beta^m \end{pmatrix}.$$

We aim now at computing the successive Lie brackets between $\hat{\mu}$ and $\hat{\sigma}$, in order to determine suitable conditions under which (3.3) holds. Recalling Definition B.1.18 we have to compute:

$$[\hat{\mu}, \hat{\sigma}](\hat{r}) = d\hat{\mu}(\hat{r})(\hat{\sigma}(\hat{r})) - d\hat{\sigma}(\hat{r})(\hat{\mu}(\hat{r})),$$

where $d\hat{\mu}$ denotes the differential of $\hat{\mu}$, which is locally represented by the Fréchet derivative of the local representation of $\hat{\mu}$ (with an abuse of notation we will denote by $\hat{\mu}$ the local representation of $\hat{\mu}$ since all the properties that we are studying are local). Therefore, recalling that:

$$\hat{r} = (r^0, \ldots, r^m, Y^1, \ldots, Y^m),$$

Fréchet derivatives of $\hat{\mu}$ and $\hat{\sigma}$ are respectively given by:

$$\frac{\partial}{\partial \hat{r}}\hat{\mu}(\hat{r}) = \begin{pmatrix} \mathbf{F} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{F} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{F} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{F} & 0 & \cdots & 0 \\ \mathbf{B} & -\mathbf{B} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \mathbf{B} & 0 & -\mathbf{B} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{B} & 0 & 0 & \cdots & -\mathbf{B} & 0 & \cdots & 0 \end{pmatrix}, \tag{3.7}$$

whereas

$$\frac{\partial}{\partial \hat{r}}\hat{\sigma}_i(\hat{r}) = \mathbb{O},$$

where we have denote with $\mathbb{O}$ the matrix of the same dimension of (3.7) such that $\mathbb{O}_{ij} = 0$ for each $i$ and $j$.

Computing the Lie bracket of $\hat{\mu}$ and $\hat{\sigma}_i$, for $i \in \{1, \ldots, d\}$, we get:

$$[\hat{\mu}, \hat{\sigma}_i] = \frac{\partial}{\partial \hat{r}}\hat{\mu}(\hat{r})\hat{\sigma}_i(\hat{r}) - \overbrace{\frac{\partial}{\partial \hat{r}}\hat{\sigma}_i(\hat{r})}^{=\mathbb{O}}\hat{\mu}(\hat{r}) = \begin{pmatrix} \mathbf{F}\sigma_i^0 \\ \mathbf{F}\sigma_i^1 \\ \vdots \\ \mathbf{F}\sigma_i^m \\ \mathbf{B}\sigma_i^0 - \mathbf{B}\sigma_i^1 \\ \vdots \\ \mathbf{B}\sigma_i^0 - \mathbf{B}\sigma_i^m \end{pmatrix}.$$

We can observe that the Lie brackets is constant on $\hat{\mathcal{H}}$. This means that in $\{\hat{\mu}, \hat{\sigma}\}_{LA}$, the only vector field which is not constant is $\hat{\mu}$. Therefore, it is sufficient to find a law which describes the Lie bracket between $\hat{\mu}$ and the successive Lie bracket between $\hat{\mu}$ and $\hat{\sigma}$. The other Lie brackets will be null by definition. For instance, if we compute:

$$[\hat{\mu}, [\hat{\mu}, \hat{\sigma}_i]] = d\hat{\mu}(\hat{r})([\hat{\mu}, \hat{\sigma}_i](\hat{r})) - \overbrace{d[\hat{\mu}, \hat{\sigma}_i]}^{:=\mathbb{O}}(\hat{r})(\hat{\mu}(\hat{r})) = \begin{pmatrix} \mathbf{F}^2\sigma_i^0 \\ \mathbf{F}^2\sigma_i^1 \\ \vdots \\ \mathbf{F}^2\sigma_i^m \\ \mathbf{B}\mathbf{F}\sigma_i^0 - \mathbf{B}\mathbf{F}\sigma_i^1 \\ \vdots \\ \mathbf{B}\mathbf{F}\sigma_i^0 - \mathbf{B}\mathbf{F}\sigma_i^m \end{pmatrix},$$

where the Fréchet derivative of the Lie bracket is null, since $[\hat{\mu}, \hat{\sigma}]$ is constant.

If we generalize the previous procedure inductively, we obtain the following result:

$$\mathcal{L} := \{\hat{\mu}, \hat{\sigma}\}_{LA} = Span\Big\{\hat{\mu}, \hat{\sigma}_1, \dots, \hat{\sigma}_d, \nu_i^k \mid k \in \mathbb{N}, \quad i = 1, \dots, d\Big\}, \tag{3.8}$$

where

$$\nu_i^k = \begin{pmatrix} \mathbf{F}^k \sigma_i^0 \\ \mathbf{F}^k \sigma_i^1 \\ \vdots \\ \mathbf{F}^k \sigma_i^m \\ \mathbf{B}\mathbf{F}^{k-1}\sigma_i^0 - \mathbf{B}\mathbf{F}^{k-1}\sigma_i^1 \\ \vdots \\ \mathbf{B}\mathbf{F}^{k-1}\sigma_i^0 - \mathbf{B}\mathbf{F}^{k-1}\sigma_i^m \end{pmatrix}. \tag{3.9}$$

In order to find sufficient conditions under which $dim[\mathcal{L}] < +\infty$ we introduce the following definition:

**Definition 3.2.1.** *A quasi-exponential function (QE) is a function of the form:*

$$f(x) = \sum_i e^{\lambda_i x} + \sum_j e^{\alpha_j x}[p_j(x)\cos\omega_j x + q_j(x)\sin\omega_j x],$$

*where $\lambda_i, \alpha_j$ and $\omega_j$ are real numbers and $p_i, q_j$ are real polynomials.*

For a detailed description of quasi-exponential functions we refer to [19] and [7]. The following characterization of $QE$ functions is crucial for our purposes:

**Lemma 3.2.2.** *A function $f$ is QE if and only if it is a component of the solution of a vector valued linear ODE with constant coefficients:*

$$\frac{\partial^n}{\partial x^n} f = \sum_{i=0}^{n-1} \gamma_i \frac{\partial^i}{\partial x^i} f$$

We prove now the main result of this section, which characterizes condition (3.3). It is based on [21][Proposition 3.2].

**Theorem 3.2.3.** *System* (3.1) *with constant volatility possesses finite-dimensional realization (FDR) (i.e. equivalence* (3.3) *holds) if and only if: $\sigma_j^i(x)$ are QE functions for each $j \in \{1, \dots, d\}$ and $i \in \{0, \dots, m\}$.*

*Proof.* As a preliminary, we can observe that:

$$\mathcal{L} = Span\{\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_d\} + \overbrace{Span\{(\nu_i^k)_{i=1,\ldots,d},\ k \in \mathbb{N}\}}^{N},$$

where $+$ denotes the sum between vector spaces. This implies that $dim[\mathcal{L}] < +\infty$ if and only if $dim[N] < +\infty$. Therefore, to prove the theorem it suffices to prove that $N$ is finite-dimensional if and only if $\hat{\sigma}_i^j$ are QE functions for each $i = 1, \ldots, d$ and for each $j = 0, \ldots, m$.
($\Longrightarrow$) Suppose that $N$ is finite-dimensional. Then, for each $i$ $Span\{\nu_i^n, n \in \mathbb{N}\}$ is finite-dimensional. In particular, this fact means that the following condition holds:

$$\forall\ i = 1, \ldots, d\ \exists\ n_i \in \mathbb{N}: \quad \nu_i^{n_i+1} = \sum_{k=1}^{n_i} \alpha_{k,i} \nu_i^k, \tag{3.10}$$

where $\alpha_{k,i} \in \mathbb{R}$ for each $i$ and $k$. The first $m+1$ rows of the system (3.10) imply that:

$$\begin{cases} \mathbf{F}^{n_i+1}\sigma_i^0 = \sum_{k=1}^{n_i} \alpha_{k,i}\mathbf{F}^k\sigma_i^0 \\ \vdots \\ \mathbf{F}^{n_i+1}\sigma_i^m = \sum_{k=1}^{n_i} \alpha_{k,i}\mathbf{F}^k\sigma_i^m \end{cases}$$

By Lemma 3.2.2, for the previous system all the constant vector fields $\sigma_i^0, \ldots, \sigma_i^m$ are QE-functions.
($\Longleftarrow$) Let us suppose that $\sigma_i^j(x)$ are QE functions for each $i = 1, \ldots, d$ and $j = 0, \ldots, m$. By Lemma 3.2.2, for each $j$ and $i$ there exists a polynomial:

$$P_i^{(j)}(\lambda) = \lambda^{n_i^j+1} - \alpha^{(j)}_{n_i^j,i}\lambda^{n_i^j} - \alpha^{(j)}_{n_i^j-1,i}\lambda^{n_i^j-1} - \cdots - \alpha^{(j)}_{0,i},$$

such that $P_i^{(j)}(\mathbf{F})\sigma_i^{(j)} = 0$, where $\mathbf{F}^n \equiv \frac{\partial^n}{\partial x^n}$.
If we consider now the polynomial:

$$M_i(\lambda) = \prod_{j=0}^{m} P_i^{(j)}(\lambda), \tag{3.11}$$

then, the following conditions hold:

$$\begin{cases} M_i(\mathbf{F})\sigma_i^0 = 0, \\ \vdots \\ M_i(\mathbf{F})\sigma_i^m = 0, \end{cases} \tag{3.12}$$

The degree of $M_i$ is $n_i = \sum_{j=0}^{m} n_i^j + 1$.

Denoting the polynomial $M_i$ with: $M_i(\lambda) = \lambda^{n_i} + \widetilde{\alpha}^i_{n_i-1}\lambda^{n_i-1} + \cdots + \widetilde{\alpha}^i_1\lambda + \widetilde{\alpha}^i_0$, we obtain:

$$\begin{cases} \mathbf{F}^{n_i}\sigma_i^0 + \widetilde{\alpha}^i_{n_i-1}\mathbf{F}^{n_i-1}\sigma_i^0 + \cdots + \widetilde{\alpha}^i_1\mathbf{F}\sigma_i^0 + \widetilde{\alpha}^i_0\sigma_i^0 = 0, \\ \vdots \\ \mathbf{F}^{n_i}\sigma_i^m + \widetilde{\alpha}^i_{n_i-1}\mathbf{F}^{n_i-1}\sigma_i^m + \cdots + \widetilde{\alpha}^i_1\mathbf{F}\sigma_i^m + \widetilde{\alpha}^i_m\sigma_i^m = 0. \end{cases}$$

We can also observe that:

$$0 = \mathbf{F}0 = \mathbf{F}M_i(\mathbf{F})\sigma_i^0 = M_i(\mathbf{F})[\mathbf{F}\sigma_i^0], \tag{3.13}$$

$$0 = \mathbf{F}0 = \mathbf{F}M_i(\mathbf{F})\sigma_i^j = M_i(\mathbf{F})[\mathbf{F}\sigma_i^j], \quad \text{for } j = 1, \ldots, m, \tag{3.14}$$

By the linearity of $\mathbf{F}$ and applying a reduction between the equations of (3.12), we get $M_i(\mathbf{F})(\sigma_i^0 - \sigma_i^j) = 0$. This means that $M_i(\mathbf{F})(\sigma_i^0 - \sigma_i^j)(x) = 0$ for each $x \in \mathbb{R}_+$. In turn, this implies that:

$$\mathbf{B}M_i(\mathbf{F})(\sigma_i^0 - \sigma_i^j) = M_i(\mathbf{F})(\sigma_i^0 - \sigma_i^j)(0) = 0. \tag{3.15}$$

Writing (3.13),(3.14),(3.15) in expanded form, we have that:

$$\begin{cases} \mathbf{F}^{n_i+1}\sigma_i^0 + \widetilde{\alpha}^i_{n_i-1}\mathbf{F}^{n_i}\sigma_i^0 + \cdots + \widetilde{\alpha}^i_1\mathbf{F}^2\sigma_i^0 + \widetilde{\alpha}^i_0\mathbf{F}\sigma_i^0 = 0, \\ \mathbf{F}^{n_i+1}\sigma_i^j + \widetilde{\alpha}^i_{n_i-1}\mathbf{F}^{n_i}\sigma_i^j + \cdots + \widetilde{\alpha}^i_1\mathbf{F}^2\sigma_i^j + \widetilde{\alpha}^i_0\mathbf{F}\sigma_i^j = 0, \\ \mathbf{B}\mathbf{F}^n\sigma_i^0 - \mathbf{B}\mathbf{F}^n\sigma_i^j + \widetilde{\alpha}^i_{n_i-1}(\mathbf{B}\mathbf{F}^{n_i-1}\sigma_i^0 - \mathbf{B}\mathbf{F}^{n_i-1}\sigma_i^j) + \cdots + \widetilde{\alpha}^i_1(\mathbf{B}\mathbf{F}\sigma^0 - \mathbf{B}\mathbf{F}\sigma^j) + \\ \qquad + \widetilde{\alpha}^i_0(\mathbf{B}\sigma_i^0 - \mathbf{B}\sigma_i^j) = 0 \end{cases}$$

In conclusion, we obtain:

$$\begin{pmatrix} \mathbf{F}^{n_i+1}\sigma_i^0 \\ \mathbf{F}^{n_i+1}\sigma_i^1 \\ \vdots \\ \mathbf{F}^{n_i+1}\sigma_i^m \\ \mathbf{B}\mathbf{F}^{n_i}\sigma_i^0 - \mathbf{B}\mathbf{F}^{n_i}\sigma_i^1 \\ \vdots \\ \mathbf{B}\mathbf{F}^{n_i}\sigma_i^0 - \mathbf{B}\mathbf{F}^{n_i}\sigma_i^m \end{pmatrix} = -\widetilde{\alpha}^i_{n_i-1}\begin{pmatrix} \mathbf{F}^{n_i}\sigma_i^0 \\ \mathbf{F}^{n_i}\sigma_i^1 \\ \vdots \\ \mathbf{F}^{n_i}\sigma_i^m \\ \mathbf{B}\mathbf{F}^{n_i-1}\sigma_i^0 - \mathbf{B}\mathbf{F}^{n_i-1}\sigma_i^1 \\ \vdots \\ \mathbf{B}\mathbf{F}^{n_i-1}\sigma_i^0 - \mathbf{B}\mathbf{F}^{n_i-1}\sigma_i^m \end{pmatrix} - \cdots - \widetilde{\alpha}^i_0\begin{pmatrix} \mathbf{F}\sigma_i^0 \\ \mathbf{F}\sigma_i^1 \\ \vdots \\ \mathbf{F}\sigma_i^m \\ \mathbf{B}\sigma_i^0 - \mathbf{B}\sigma_i^1 \\ \vdots \\ \mathbf{B}\sigma_i^0 - \mathbf{B}\sigma_i^m \end{pmatrix},$$

which is equivalent to:

$$\nu_i^{n_i+1} = -\sum_{k=0}^{n_i} \alpha_k^i \nu_i^k.$$

Since the previous equivalence holds for each $i \in \{1, \ldots, d\}$, the vector space $Span\{\nu_i^n | n \in \mathbb{N}\}$ is finite dimensional for each $i$. Therefore, $N$ is the sum of $d$ finite-dimensional vector spaces, then it is finite dimensional too. In conclusion, for what we said at the beginning of the proof, the Lie algebra $\{\hat{\mu}, \hat{\sigma}\}_{LA}$ is finite-dimensional. Moreover, this result implies that if the condition on volatility $\hat{\sigma}$ holds, then the dimension of $\{\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_d\}_{LA}$ is dominated by:

$$dim\{\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_d\}_{LA} \leq 1 + d + \sum_{i=1}^{d} n_i = 1 + \sum_{i=1}^{d} (1 + n_i). \tag{3.16}$$

□

### 3.2.1 Construction of finite-dimensional realizations

In order to construct explicitly the finite-dimensional realizations, we can apply the strategy outlined at the end of Section 3.1. We have to compute the integral curve of each vector field which span the Lie algebra generated $\mathcal{L} = \{\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n\}$. We have seen that if the volatility is constant then the Lie algebra $\mathcal{L}$ is determined by equation (3.8). Therefore, by Theorem 3.2.3, it is sufficient to compute $e^{\xi t} x_0$ for:

$$\xi \in \{\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_d, \nu_i^k |\ i = 1, \ldots, d;\ k = 1, \ldots, n_i\},$$

because these vector fields generate the entire Lie algebra. We now introduce the following notation:

$$\begin{aligned}\hat{\mu} =& (\mu_0, \mu_1, \ldots, \mu_m, \mu_{m+1}, \ldots, \mu_{2m})^*,\\ \nu_i^k =& (\nu_{i,0}^k, \nu_{i,1}^k, \ldots, \nu_{i,2m}^k)^*, \quad k = 0, \ldots, n_i, \quad i = 1, \ldots, d,\end{aligned}$$

where, with a slight abuse of notation we denote: $\hat{\sigma}_i = \nu_i^0$. Moreover, we use the following notation for the initial value $\hat{r}^M$:

$$\hat{r}^M = \begin{pmatrix} r_0^M & \cdots & r_m^M & y_1^M & \cdots & y_m^M \end{pmatrix}.$$

Now, we compute the integral curves of all these vector fields. We do this componentwise:

$\boxed{\mu_0}$ The integral curve of $\mu_0$ is a curve, denoted by $\vartheta_{r_0^M}$, solution to the following ODE:

$$\begin{cases} \frac{d}{dt}\vartheta_{r_0^M}(t) = \mu_0(\vartheta_{r_0^M}(t)) = \mathbf{F}\vartheta_{r_0^M} + \sigma^0 \mathbf{H} \sigma^0, \\ \vartheta_{r_0^M}(0) = r_0^M. \end{cases}$$

By assumption, $\sigma^0 \mathbf{H} \sigma^0$ is constant on $\mathcal{H}^0$. Therefore, the solution to this ODE can be computed in analogy to the finite-dimensional case:

$$\vartheta_{r_0^M}(t) = e^{\mathbf{F}t} r_0^M + \int_0^t e^{\mathbf{F}(t-s)} \sigma^0 \mathbf{H} \sigma^0 ds,$$

where:

$$e^{\mathbf{F}t}r_0^M(x) = \sum_{n=0}^{+\infty} \frac{(\mathbf{F}t)^n}{n!} r_0^M(x) = \sum_{n=0}^{+\infty} \frac{1}{n!}\frac{\partial^n}{\partial x^n} r_0^M(x)(t+x-x)^n = r_0^M(t+x).$$

The previous equivalence is obtained by the Taylor expansion of $r_0^M$ around $x$. We can follow this strategy because it is possible to prove that if $r_0^M \in \mathcal{H}$, then $r_0^M$ is an analytic function.

Moreover:

$$\int_0^t \left(e^{\mathbf{F}(t-s)}\sigma^0 \mathbf{H}\sigma^0\right)(x)ds = \int_0^t (\sigma^0\mathbf{H}\sigma^0)(x+t-s)ds.$$

If we define $S^0(x) = \int_0^x \sigma^0(s)ds$, then:

$$\frac{1}{2}\frac{\partial}{\partial x}||S^0(x)||^2 = \frac{1}{2}\frac{\partial}{\partial x}\left|\left|\int_0^x \sigma^0(s)ds\right|\right|^2 = \sigma^0(x)\int_0^x \sigma^{0*}(s)ds = (\sigma^0\mathbf{H}\sigma^0)(x).$$

In conclusion, the solution is given by:

$$\begin{aligned}
\vartheta_{r_0^M}(t)(x) &= r_0^M(t+x) + \int_0^t e^{\mathbf{F}(t-s)}\sigma^0\mathbf{H}\sigma^0 ds \\
&= r_0^M(t+x) + \int_0^t \frac{1}{2}\frac{\partial}{\partial x}||S^0(x+t-s)||^2 ds \\
&= r_0^M(t+x) + \int_x^{x+t} \frac{1}{2}\frac{\partial}{\partial u}||S^0(u)||^2 du \\
&= r_0^M(t+x) + \frac{1}{2}\Big[||S^0(x+t)||^2 - ||S^0(x)||^2\Big].
\end{aligned} \tag{3.17}$$

$\boxed{\mu_j,\ j=1,\dots,m}$ The integral curve of the drift term of each component $r^j,\ j=1,\dots,m$ satisfies the following ODE:

$$\begin{cases} \phi'_{r_j^M}(t) = \mu_j(\phi_{r_j^M}(t)) = \mathbf{F}\phi_{r_j^M}(t) + \sigma^j\mathbf{H}\sigma^j - \beta^j\sigma^{j*}, \\ \phi_{r_j^M}(0) = r_j^M. \end{cases}$$

Similarly as in the previous case we can notice that $\sigma^j\mathbf{H}\sigma^j - \beta^j\sigma^{i*}$ is constant on $\hat{\mathcal{H}}$; therefore:

$$\begin{aligned}
\phi_{r_j^M}(t)(x) &= r_j^M(t+x) + \frac{1}{2}\Big[||S^j(t+x)||^2 - ||S^j(x)||^2\Big] - \int_0^t \left(e^{\mathbf{F}(t-s)}\beta^j\sigma^{j*}\right)(x)ds \\
&= r_j^M(t+x) + \frac{1}{2}\Big[||S^j(t+x)||^2 - ||S^j(x)||^2\Big] - \int_0^t \beta^j\sigma^{j*}(t+x-s)ds \\
&= r_j^M(t+x) + \frac{1}{2}\Big[||S^j(t+x)||^2 - ||S^j(x)||^2\Big] - \Big(S^j(t+x) - S^j(x)\Big)\beta^{j*},
\end{aligned} \tag{3.18}$$

where $S^j(x) = \int_0^x \sigma^j(s)ds$.

$\boxed{\mu_j,\ j = m+1, \dots, 2m}$ The integral curve of the drift term of the last $m$ components is the solution of the following ODE:

$$\begin{cases} \frac{d}{dt}\psi_{y^M_{j-m}}(t) = \mu(\psi_{y^M_{j-m}}(t)) = \mathbf{B}\vartheta_{r^M_0}(t) - \mathbf{B}\phi_{r^M_{j-m}}(t) - \frac{1}{2}\|\beta^{j-m}\|^2, \\ \psi_{y^M_{j-m}}(0) = y^M_{j-m}. \end{cases}$$

then

$$\psi_{y^M_{j-m}}(t) = y^M_{j-m} + \int_0^t (\vartheta_{r^M_0}(s)(0) - \phi_{r^M_{j-m}}(s)(0))ds - \frac{1}{2}\|\beta^{j-m}\|^2 t.$$

If we exploit (3.17) and (3.18), we obtain:

$$\begin{aligned} \psi_{y^M_{j-m}}(t) =& y^M_{j-m} + \int_0^t \Big\{ r^M_0(s) + \frac{1}{2}\Big[\|S^0(0+s)\|^2 - \overbrace{\|S^0(0)\|^2}^{=0}\Big] - r^M_{j-m}(s) - \frac{1}{2}\Big[\|S^{j-m}(s)\|^2 + \\ & - \|S^{j-m}(0)\|^2\Big] - \Big(S^{j-m}(s) - \overbrace{S^{j-m}(0)}^{=0}\Big)\beta^{j-m*}\Big\} ds - \frac{1}{2}\|\beta^{j-m}\|^2 t \\ =& y^M_{j-m} + \int_0^t (r^M_0(s) - r^M_{j-m}(s))ds + \frac{1}{2}\int_0^t \Big[\|S^0(s)\|^2 - \|S^{j-m}(s)\|^2\Big] ds + \\ & - \int_0^t S^{j-m}(s)ds \cdot \beta^{j-m*} - \frac{1}{2}\|\beta^{j-m}\|^2 t. \end{aligned} \tag{3.19}$$

We now compute the integral curves of each component of the vector filed $\nu_i^k$ for each $k = 0, \dots, n_i$ and $i = 1, \dots, d$. In this case, the vector fields $\nu_i^k$ defined in (3.9) are constant, then their integral curves are lines defined on $\hat{\mathcal{H}}$. In particular, the integral curve of $\nu_i^k$ at time $t$, denoted by $e^{\nu_i^k t}\hat{r}^M$, has the following form:

$$e^{\nu_i^k t} r^M = \hat{r}^M + \nu_i^k t, \quad i = 1, \dots, d;\ k = 0, \dots, n_i.$$

Recalling Proposition B.3.3 and the notation $e^{\xi t}$, in order to describe the integral curve of a vector field $\xi$ we can compute the tangential manifold of the involutive distribution $\{\hat{\mu}, \hat{\sigma}_1, \dots, \hat{\sigma}_d\}_{LA}$ which contains the initial point $\hat{r}^M$ as the image of the mapping $G : \mathbb{R}^n \longrightarrow \hat{\mathcal{H}}$,

$$G(z^0, z^j_k :\ i = 1, \dots, d;\ k = 0, \dots, n_i)(x) = \Big(\Big(\prod_{i=1,\dots,d\ k=0,\dots,n_i} e^{\nu_j^k}\Big) e^{\hat{\mu}}\hat{r}^M\Big)(x) \tag{3.20}$$

**Remark 3.2.4.** *For the state-space vector $z \in \mathcal{Z}$, we introduce the following notation. Remembering the condition* (3.16)*, we have that the dimension of the Lie algebra $\mathcal{L}$ is $n = 1 + \sum_{i=1}^{d}(1+n_j)$. Therefore, we use $(z^0, (z_i^k))$ to denote:*

$$z^* = (z^0, (z_i^k)^*_{i=1,\ldots,d\ \ k=0,\ldots,n_i})^*,$$

*where $(z_i^k)^*_{i=1,\ldots,d\ \ k=0,\ldots,n_i} \in \mathbb{R}^{n-1}$ is given by:*

$$(z_i^k)^*_{i=1,\ldots,d\ \ k=0,\ldots,n_i} = (z_1^0, \ldots, z_d^0, z_0^1, \ldots, \ldots, \ldots, z_d^{n_i})^*.$$

*Moreover, we will use the same notation for the finite-dimensional vector field $a$ and $b_i$ $i = 1, \ldots, d$ which define the process $Z_t$ introduced in* (3.4)*.*

In particular, the coordinates $G^j$ for $j = 0, \ldots, 2m$ of the function defined in (3.20) are given by:
if $j = 0$:

$$G^0(z^0, z_i^k)(x) = r_0^M(x+z^0) + \sum_{i=1}^{d}\sum_{k=0}^{n_i} \mathbf{F}^k \sigma_i^0(x) z_i^k + \frac{1}{2}(||S^0(x+z^0)||^2 - ||S^0(x)||^2), \tag{3.21}$$

for $j = 1, \ldots, m$:

$$\begin{aligned} G^j(z^0, z_i^k)(x) =& r_j^M(x+z^0) + \sum_{i=1}^{d}\sum_{k=0}^{n_i} \mathbf{F}^k \sigma_i^j(x) z_i^k + \frac{1}{2}(||S^j(x+z^0)||^2 + \\ & - ||S^j(x)||^2) - (S^j(x+z^0) - S^j(x))\beta^{i*} \\ =& r_j^M(x+z^0) + \sum_{i=1}^{d}\sum_{k=0}^{n_i} \mathbf{F}^k \sigma_i^j(x) z_i^k + \frac{1}{2}(||S^j(x+z^0)||^2 - ||S^j(x)||^2) + \\ & - \int_x^{x+z^0} \sigma^j(s)ds \cdot \beta^{j*}, \end{aligned} \tag{3.22}$$

and finally, for $j = m+1, \ldots, 2m$:

$$\begin{aligned} G^j(z^0, z_i^k) =& \sum_{i=1}^{d}\sum_{k=1}^{n_i} (\mathbf{B}\mathbf{F}^{k-1}\sigma_i^0 - \mathbf{B}\mathbf{F}^{k-1}\sigma_i^{j-m}) z_i^k + \sum_{i=1}^{d} \beta_i^j z_i^0 + y_{j-m}^M + \int_0^{z^0} \Big( r_0^M(s) + \\ & - r_j^M(s)\Big) ds + \frac{1}{2}\int_0^{z^0} \Big[||S^0(s)||^2 - ||S^{j-m}(s)||^2\Big] ds + \\ & - \int_0^{z^0} S^{j-m}(s) ds \cdot \beta^{j-m*} - \frac{1}{2}||\beta^{j-m}||^2 z^0. \end{aligned} \tag{3.23}$$

Afterword, we have to perform step 3 of the strategy outlined at the end of Section 3.1. In particular, we have to find two vector fields on $\mathbb{R}^n$ where $a$ and $b$ which satisfy respectively conditions (3.5), (3.6). For brevity of notation, in the following we will not indicate the argument $z \in \mathbb{R}^n$ on the coordinates $a^h, b^l$ for each $h, l \in \{1, \ldots, n\}$. First of all, we search for a vector field $a$ such that: $G^0_z(z)a = \mu_0(G(z))$. Explicitly, the members of the previous equation are given by:

$$G^0_z(z)a = a^0\Big(\mathbf{F}r^M_0(x+z^0) + \sigma^0\mathbf{H}\sigma^0(x+z^0)\Big) + \sum_{i=1}^{d}\sum_{k=0}^{n_i} a^k_i \mathbf{F}^k \sigma^0_i(x)$$

$$\mu_0(G(z)) = \mathbf{F}G^0(z) + \sigma^0\mathbf{H}\sigma^0$$

In particular, since $G^0$ is given by (3.21), the following equation holds:

$$\mathbf{F}G^0 = \sum_{i=1}^{d}\sum_{k=0}^{n_i} \mathbf{F}^{k+1}\sigma^0_i(x)z^k_i + \mathbf{F}r^M_0(x+z^0) + \overbrace{\mathbf{F}\left[\frac{1}{2}(||S^0(x+z^0)||^2 - ||S^0(x)||^2)\right]}^{\sigma^0\mathbf{H}\sigma^0},$$

then:

$$\mu_0(G(z)) = \sum_{i=1}^{d}\sum_{k=0}^{n_i} \mathbf{F}^{k+1}\sigma^0_i(x)z^k_i + \mathbf{F}r^M_0(x+z^0) + \sigma^0\mathbf{H}(x+z^0) - \sigma^0\mathbf{H}\sigma^0(x) + \sigma^0\mathbf{H}\sigma^0(x).$$

Hence, the condition is given by:

$$a^0\Big(\mathbf{F}r^M_0(x+z^0) + \sigma^0\mathbf{H}\sigma^0(x+z^0)\Big) + \sum_{i=1}^{d}\sum_{k=0}^{n_j} a^k_j \mathbf{F}^k \sigma^0_j(x) =$$
$$= \sum_{i=1}^{d}\sum_{k=0}^{n_i} \mathbf{F}^{k+1}\sigma^0_i(x)z^k_i + \mathbf{F}r^M_0(x+z^0) + \sigma^0\mathbf{H}(x+z^0).$$

Since the last equivalence must hold for every $x \in \mathbb{R}_+$, the conditions on the vector field $a$ are:

$$a^0 = 1, \tag{3.24}$$
$$a^0_i = 0, \qquad i = 1, \ldots, d, \tag{3.25}$$
$$\sum_{i=1}^{d}\sum_{k=1}^{n_i} a^k_i \mathbf{F}^k \sigma^0_i(x) = \sum_{i=1}^{d}\sum_{k=0}^{n_i} z^k_i \mathbf{F}^{k+1}\sigma^0_i(x). \tag{3.26}$$

Since we are assuming the existence of a finite-dimensional realization, Theorem 3.2.3 must hold. Hence, the functions $\sigma^0_i$ has to be QE for each $i$, then by Lemma 3.2.2 there exists $\alpha = (\alpha^0, (\alpha^k_i)^*)^* \in \mathbb{R}^n$ such that:

$$\mathbf{F}^{n_i+1}\sigma^0_i(x) = \sum_{k=1}^{n_i} \mathbf{F}^k \sigma^0_i(x)\alpha^k_i,$$

Therefore:

$$\begin{aligned}
\sum_{i=1}^{d}\sum_{k=1}^{n_i} a_i^k \mathbf{F}^k \sigma_i^0(x) &= \sum_{i=1}^{d}\sum_{k=0}^{n_i} z_i^k \mathbf{F}^{k+1}\sigma_i^0(x) \\
&= \sum_{i=1}^{d}\sum_{k=1}^{n_i+1} z_i^{k-1}\mathbf{F}^k\sigma_i^0(x) \\
&= \sum_{i=1}^{d}\Big(\sum_{k=1}^{n_i} z_i^{k-1}\mathbf{F}^k\sigma_i^0(x)\Big) + \sum_{i=1}^{d}\Big(z_i^{n_i}\sum_{k=1}^{n_i}\alpha_i^k\mathbf{F}^k\sigma_i^0(x)\Big) \\
&= \sum_{i=1}^{d}\left(\sum_{k=1}^{n_i}(z_i^{k-1} + z_i^{n_i}\alpha_i^k)\mathbf{F}^k\sigma_i^0(x)\right).
\end{aligned}$$

This implies that:

$$a_i^k = z_i^{k-1} + z_i^{n_i}\alpha_i^k, \qquad k = 1,\dots,n_i,\ i = 1,\dots,d. \tag{3.27}$$

The vector field $a$ is uniquely determined by the injectiveness of $dG$, so that we have described the solution of the condition (3.5). Moreover, we can observe that the other coordinates of condition (3.5) lead an analogous conclusion. Indeed, the following statements hold:

- $\boxed{G_z^j(z)a = \mu_j(G(z))\ j = 1,\dots,m}$ The two members of the equation are given by:

$$\begin{aligned}
G_z^j(z)a =& a^0\Big(\mathbf{F}r_j^M(x+z^0) + \sigma^j\mathbf{H}\sigma^j(x+z^0) - \sigma^j(x)\beta^{j*}\Big) + \sum_{i=1}^{d}\sum_{k=0}^{n_i}\mathbf{F}^k\sigma_i^j(x)a_i^k, \\
\mu_j(G(z)) =& \mathbf{F}G^j(z) + \sigma^j\mathbf{H}\sigma^{j*} - \sigma^j\beta^{j*} = \sum_{i=1}^{d}\sum_{k=0}^{n_i}\mathbf{F}^{k+1}\sigma_i^j(x)z_i^k + \\
& + \mathbf{F}r_j^M(x+z^0) + \sigma^j\mathbf{H}\sigma^j(x+z^0).
\end{aligned}$$

  Comparing the previous expressions, we obtain again conditions (3.24), (3.25), (3.27).

- $\boxed{G_z^j(z)a = \mu_j(G(z))\ j = m+1,\dots,2m}$ Also in this case we get:

$$\begin{aligned}
G_z^j(z)a =& a^0\Bigg(-\frac{1}{2}||\beta^{j-m}||^2 - S^{j-m}(z^0)\beta^{j-m*} + r_0^M(z^0) - r_j^M(z^0) + \\
&+\frac{1}{2}\Big(||S^0(z^0)||^2 - ||S^{j-m}(z^0)||^2\Big)\Bigg) + \sum_{i=1}^{d}\Big(\sum_{k=1}^{n_i} a_i^k(\mathbf{BF}^{k-1}\sigma_i^0 + \\
&- \mathbf{BF}^{k-1}\sigma_i^{j-m})\Big) - \sum_{i=1}^{d}\beta_i^j a_i^0, \\
\mu_j(G(z)) =& \sum_{i=1}^{d}\Bigg(\sum_{k=0}^{n_i} z_k^i\Big(\mathbf{BF}^k\sigma_i^0 - \mathbf{BF}^k\sigma_i^{j-m}\Big)\Bigg) + \frac{1}{2}\Big(||S^0(z^0)||^2 - ||S^{j-m}(z^0)||^2\Big) + \\
&+ (r_0^M(z^0) - r_j^M(z^0)) + S^{j-m}(z^0)\beta^{j-m*} - \frac{1}{2}||\beta^{j-m}||^2;
\end{aligned}$$

Also in this case, comparing the previous expressions we obtain again conditions (3.24), (3.25), (3.27) indeed the following condition

On the other hand, analysing the behaviour of the volatility term, we can compute the value of the coordinates $b$. In particular, we have to solve the condition: $G_z(z)b(z) = \hat{\sigma}(G(z))$, for each $z \in \mathcal{Z}$. This condition corresponds to the following system:

$$\begin{cases}
G_z^0(z)b_i(z) = \sigma_i^0(G(z)), \quad i = 1,\dots,d; \\
G_z^j(z)b_i(z) = \sigma_i^j(G(z)), \quad i = 1,\dots,d; \quad j = 1,\dots,m; \\
G_z^{j+m}(z)b_i(z) = \beta_i^j(G(z)), \quad i = 1,\dots,d; \quad j = 1,\dots,m.
\end{cases}$$

For each $i = 1,\dots,d$ the first condition of the previous system is explicitly given by:

$$b_i^0\Big(r_0^0(x+z^0) + (\sigma^0\mathbf{H}\sigma^0)(x+z^0)\Big) + \sum_{h=1}^{d}\sum_{k=0}^{n_i} b_{h,i}^k\mathbf{F}^k\sigma_j^0(G(z)) = \sigma_i^0(G(z)),$$

which implies that:

$$b_{i,i}^0 = 1; \tag{3.28}$$
$$b_i^0 = 0; \tag{3.29}$$
$$b_{h,i}^k = 0, \quad h = 1,\dots,d,\ j \neq i;\ k = 1,\dots,n_j. \tag{3.30}$$

If we analyse the other conditions, we obtain the same result as in the drift term. Indeed:

- $\boxed{G_z^j(z)b_i = \sigma_i^j(G(z)) \ j = 1, \dots, m}$

$$G_z^j(z)b_i = b_i^0\Big(\mathbf{F}r_j^M(x+z^0) + \sigma^j\mathbf{H}\sigma^j(x+z^0) - \sigma^j(x)\beta^{j*}\Big) + \\ + \sum_{h=1}^{d}\sum_{k=0}^{n_h}\mathbf{F}^k\sigma_h^j(x)b_{h,i}^k = \sigma_i^j(G(z)),$$

which is satisfied if conditions (3.28), (3.29), (3.30);

- $\boxed{G_z^j(z)b_i = \beta_i^{j-m}(G(z)) \ j = m+1, \dots, 2m}$

$$G_z^j(z)b_i = b_i^0\Bigg(-\frac{1}{2}||\beta^{j-m}||^2 - S^{j-m}(z^0)\beta^{j-m*} + r_0^M(z^0) - r_j^M(z^0) + \frac{1}{2}\Big(||S^0(z^0)||^2 + \\ - ||S^{j-m}(z^0)||^2\Big)\Bigg) + \sum_{h=1}^{d}\Big(\sum_{k=1}^{n_h} b_{i,h}^k(\mathbf{B}\mathbf{F}^{k-1}\sigma_h^0 - \mathbf{B}\mathbf{F}^{k-1}\sigma_h^{j-m})\Big) + \\ - \sum_{h=1}^{d}\beta_h^{j-m}b_{i,h}^0 = \beta_i^{j-m}(G(z)).$$

which is satisfied if conditions (3.28), (3.29), (3.30), too.

In conclusion, we have proved the following proposition:

**Proposition 3.2.5.** *If the model $\mathcal{M}$, described by the equation* (1.33) *is determined by a constant volatility term $\hat{\sigma}$, then there exists finite-dimensional realizations if and only if the function $\sigma_j^i(x)$ are QE functions for each $i = 1, \dots, d$ and $j = 0, \dots, m$. In this case, the existence of finite-dimensional realizations is guaranteed and the coefficients of the $\mathbb{R}^n$-valued process $dZ_t = a(Z_t)dt + b(Z_t) \circ dW_t$, are determined by conditions* (3.25), (3.25), (3.27), (3.28), (3.29), (3.30).

## 3.3 Constant direction volatility

In this section we aim at analysing the existence of finite-dimensional realizations for a model $\mathcal{M}$ determined by a volatility term given by:

$$\hat{\sigma}_i(\hat{r}, x) = \begin{pmatrix} \varphi_i^0(\hat{r})\lambda_i^0(x) \\ \varphi_i^1(\hat{r})\lambda_i^1(x) \\ \vdots \\ \varphi_i^m(\hat{r})\lambda_i^m(x) \\ \beta_i^1(\hat{r}) \\ \vdots \\ \beta_i^m(\hat{r}) \end{pmatrix}, \quad i = 1, \dots, d, \tag{3.31}$$

where $\lambda_i^j(x)$ are elements of $\mathcal{H}$ for each $i = 1, \ldots, d$; $j = 0, \ldots, m$ and $\varphi_i^j(\hat{r})$ are smooth scalar vector fields defined on $\hat{\mathcal{H}}$, i.e.: $\varphi_i^j(\hat{r}) \in \mathcal{C}^\infty(\hat{\mathcal{H}}, \mathbb{R})$. This condition implies that we can divide the dependence on the variable time to maturity $(x)$ and the dependence on the entire solution $\hat{r}$. The contents of this section are based on [21][Section 4].

In order to find conditions which guarantee finite-dimensional realizations we need the following assumption:

**Assumption 3.3.1.** *For each $i = 1, \ldots, d$ and $j = 0, \ldots, m$ we suppose that: $\varphi_i^j(\hat{r}) \neq 0$ and $\beta_i^j(\hat{r}) \neq 0$ for each $\hat{r} \in \hat{\mathcal{H}}$.*

In what follows we characterize the drift term introduced in (2.8) when the volatility is given by (3.31). First of all, it is convenient to introduce the following notation for the volatility:

$$\hat{\sigma}(\hat{r}) = \begin{pmatrix} \varphi_1^0(\hat{r})\lambda_1^0(x) & \cdots & \varphi_d^0(\hat{r})\lambda_d^0(x) \\ \varphi_1^1(\hat{r})\lambda_1^1(x) & \cdots & \varphi_d^1(\hat{r})\lambda_d^1(x) \\ \vdots & & \vdots \\ \varphi_1^m(\hat{r})\lambda_1^m(x) & \cdots & \varphi_d^m(\hat{r})\lambda_d^m(x) \\ \beta_1^1(\hat{r}_t) & \cdots & \beta_d^1(\hat{r}_t) \\ \vdots & & \vdots \\ \beta_1^m(\hat{r}_t) & \cdots & \beta_d^m(\hat{r}_t) \end{pmatrix} \equiv \begin{pmatrix} \sigma^0 \\ \vdots \\ \sigma^m \\ \beta^1 \\ \vdots \\ \beta^m \end{pmatrix}.$$

Recalling that the Stratonovich dynamics of $\hat{r}$ is given by (2.7), we have to compute the term related to the Fréchet derivative of the volatility involved by the Stratonovich correction term. In particular, if $j = 0, \ldots, m$:

$$\begin{aligned} \frac{\partial \sigma^j}{\partial \hat{r}}(\hat{r}_t, x)\hat{\sigma}(\hat{r}_t) &= \sum_{h=0}^m \frac{\partial \sigma^j(\hat{r}_t)}{\partial r^h}\sigma^h(\hat{r}_t) + \sum_{h=1}^m \frac{\partial \sigma^j(\hat{r}_t)}{\partial Y^h}\beta^h(\hat{r}_t) \\ &= \sum_{h=0}^m \sum_{i=1}^d \frac{\partial \sigma_i^j(\hat{r}_t)}{\partial r^h}\sigma_i^h(\hat{r}_t) + \sum_{h=1}^m \sum_{i=1}^d \frac{\partial \sigma_i^j(\hat{r}_t)}{\partial Y^h}\beta_i^h(\hat{r}_t) \\ &= \sum_{i=1}^d \left( \sum_{h=0}^m \lambda_i^j(x) \frac{\partial \varphi_i^j(\hat{r}_t)}{\partial r^h}\varphi_i^h(\hat{r}_t)\lambda_i^h(\hat{r}_t) + \sum_{h=1}^m \lambda_i^j(x)\frac{\partial \varphi_i^j(\hat{r}_t)}{\partial Y^h}\beta_i^h(\hat{r}_t) \right). \end{aligned}$$

It is also necessary compute the term $\sigma^j(\hat{r}_t)\mathbf{H}\sigma^j(\hat{r}_t)$:

$$\sigma^j(\hat{r}_t)\mathbf{H}\sigma^j(\hat{r}_t) = (\varphi^j(\hat{r}_t)\lambda^j(x)) \cdot \int_0^x (\varphi^j(\hat{r}_t)\lambda^j(s))^* ds = \sum_{i=1}^d (\varphi_i^j(\hat{r}_t))^2 \lambda_i^j(x) \int_0^x \lambda_i^j(s) ds.$$

Moreover, we introduce the notation $\frac{\partial \varphi_i^j(\hat{r})}{\partial \hat{r}^h}[\lambda_i^h]$ in order to denote the Fréchet derivative of $\varphi_i^j$ on the variable $r^h$ computed on $\hat{r}$ acting on the vector $\lambda_i^h$, for each

$h, j = 0, \ldots, 2m$ and $i = 1, \ldots, d$. It is convenient to introduce also the following notation:

$$D_i^j(x) := \lambda_i^j(x) \int_0^x \lambda_i^j(s)ds, \quad j = 0, \ldots, m, \; i = 1, \ldots, d.$$

Therefore, in the Stratonovich form, the dynamics of each equation of system (1.33) is given as follows. If $j = 0$:

$$\begin{aligned} dr_t^0 = & \Bigg[ \mathbf{F}r^0 + \sum_{i=1}^d (\varphi_i^0(\hat{r}_t))^2 D_i^0 - \frac{1}{2} \sum_{i=1}^d \lambda_i^0 \Bigg( \sum_{h=0}^m \varphi_i^h(\hat{r}_t) \frac{\partial \varphi_i^0(\hat{r}_t)}{\partial r^h} [\lambda_i^h] + \\ & + \sum_{h=1}^m \frac{\partial \varphi_i^0(\hat{r}_t)}{\partial Y^h} \beta_i^h(\hat{r}_t) \Bigg) \Bigg] dt + \sum_{i=1}^d \varphi_i^0(\hat{r}_t) \lambda_i^0(x) \circ dW_t, \end{aligned}$$

whereas for $j = 1, \ldots, m$:

$$\begin{aligned} dr_t^j = & \Bigg[ \mathbf{F}r^j + \sum_{i=1}^d (\varphi_i^j(\hat{r}_t))^2 D_i^j - \frac{1}{2} \sum_{i=1}^d \lambda_i^j \Bigg( \sum_{h=0}^m \varphi_i^h(\hat{r}_t) \frac{\partial \varphi_i^j(\hat{r}_t)}{\partial r^h} [\lambda_i^h] + \\ & + \sum_{h=1}^m \frac{\partial \varphi_i^j(\hat{r}_t)}{\partial Y^h} \beta_i^h(\hat{r}_t) - 2\varphi_i^j(\hat{r}_t) \beta_i^j(\hat{r}_t) \Bigg) \Bigg] dt + \sum_{i=1}^d \varphi_i^j(\hat{r}_t) \lambda_i^j(x) \circ dW_t, \end{aligned}$$

and, finally, the log-spread spot processes are determined by the following dynamics:

$$\begin{aligned} dY_t^j = & \Bigg\{ \mathbf{B}r^0 - \mathbf{B}r^j - \frac{1}{2} \sum_{i=1}^d (\beta_i^j(\hat{r}_t))^2 - \frac{1}{2} \Bigg[ \sum_{i=1}^d \sum_{h=0}^m \frac{\partial \beta_i^j(\hat{r}_t)}{\partial \hat{r}^h} [\lambda_i^h] \varphi_i^h(\hat{r}_t) + \\ & + \sum_{h=1}^m \frac{\partial \beta_i^j(\hat{r}_t)}{\partial Y^h} \beta_i^h(\hat{r}_t) \Bigg] \Bigg\} dt + \sum_{i=1}^d \beta_h^j(\hat{r}_t) \circ dW_t, \end{aligned}$$

for each $j = 1, \ldots, m$. We aim at determining conditions under which the Lie algebra generated by $\hat{\mu}$ and $\hat{\sigma}_i$ $i=1,\dots,d$ is finite-dimensional, where:

$$\hat{\mu}(\hat{r}_t)=\begin{pmatrix} \mathbf{F}r^0+\sum_{i=1}^d(\varphi_i^0(\hat{r}_t))^2D_i^0-\frac{1}{2}\sum_{i=1}^d\lambda_i^0\Big(\sum_{j=0}^m\varphi_i^j(\hat{r}_t)\frac{\partial\varphi_i^0(\hat{r}_t)}{\partial r^j}[\lambda_i^j]+\sum_{j=1}^m\frac{\partial\varphi_i^0(\hat{r}_t)}{\partial Y^j}\beta_i^j(\hat{r}_t)\Big)\\ \mathbf{F}r^1+\sum_{i=1}^d(\varphi_i^1(\hat{r}_t))^2D_i^1-\frac{1}{2}\sum_{i=1}^d\lambda_i^1\Big(\sum_{j=0}^m\varphi_i^j(\hat{r}_t)\frac{\partial\varphi_i^1(\hat{r}_t)}{\partial r^j}[\lambda_i^j]+\sum_{j=1}^m\frac{\partial\varphi_i^1(\hat{r}_t)}{\partial Y^j}\beta_i^j(\hat{r}_t)+\\ +2\varphi_i^1(\hat{r}_t)\beta_i^1(\hat{r}_t)\Big)\\ \vdots\\ \mathbf{F}r^m+\sum_{i=1}^d(\varphi_i^m(\hat{r}_t))^2D_i^m-\frac{1}{2}\sum_{i=1}^d\lambda_i^m\Big(\sum_{j=0}^m\varphi_i^j(\hat{r}_t)\frac{\partial\varphi_i^m(\hat{r}_t)}{\partial r^j}[\lambda_i^j]+\sum_{j=1}^m\frac{\partial\varphi_i^m(\hat{r}_t)}{\partial Y^j}\beta_i^j(\hat{r}_t)+\\ +2\varphi_i^m(\hat{r}_t)\beta_i^m(\hat{r}_t)\Big)\\ \mathbf{B}r^0-\mathbf{B}r^1-\frac{1}{2}\sum_{i=1}^d(\beta_i^1(\hat{r}_t))^2-\frac{1}{2}\Big[\sum_{i=1}^d\Big(\sum_{j=0}^m\frac{\partial\beta_i^1(\hat{r}_t)}{\partial r^j}[\lambda_i^j]\varphi_i^j(\hat{r}_t)+\sum_{j=1}^m\frac{\partial\beta_i^1(\hat{r}_t)}{\partial Y^j}\beta_i^j(\hat{r}_t)\Big)\Big]\\ \vdots\\ \mathbf{B}r^0-\mathbf{B}r^m-\frac{1}{2}\sum_{i=1}^d(\beta_i^m(\hat{r}_t))^2-\frac{1}{2}\Big[\sum_{i=1}^d\Big(\sum_{j=0}^m\frac{\partial\beta_i^m(\hat{r}_t)}{\partial r^j}[\lambda_i^j]\varphi_i^j(\hat{r}_t)+\sum_{j=1}^m\frac{\partial\beta_i^m(\hat{r}_t)}{\partial Y^j}\beta_i^j(\hat{r}_t)\Big)\Big] \end{pmatrix}, \tag{3.32}$$

and

$$\hat{\sigma}_i(\hat{r}_t)(x)=\begin{pmatrix}\varphi_i^0(\hat{r}_t)\lambda_i^0(x)\\ \varphi_i^1(\hat{r}_t)\lambda_i^1(x)\\ \vdots\\ \varphi_i^m(\hat{r}_t)\lambda_i^m(x)\\ \beta_i^1(\hat{r}_t)\\ \beta_i^m(\hat{r}_t)\end{pmatrix},\qquad i=1,\dots,d. \tag{3.33}$$

Differently from the case described in the previous section, the drift term is more complex and it seems very difficult to compute the integral curve of $\hat{\mu}$. This implies that we cannot compute the integral curve directly. In order to overcome this problem, we exploit Lemma B.3.5 and, to apply it, Assumption 3.3.1. Therefore we provide conditions such that a larger distribution than $\{\hat{\mu},\hat{\sigma}_1,\dots,\hat{\sigma}_d\}$ is finite-dimensional. In this way, we will determine a sufficient condition which guarantees the existence of finite-dimensional realizations.

We consider a model $\mathcal{M}$ determined by a drift term in (3.32) and a volatility term in (3.33). We introduce now the following set of vector fields:

$$\mathcal{N}:=\{\xi^0,\xi_i^j,\eta_i^j,\gamma_k|\ j=0,\dots,m,\ i=1,\dots,d,\ k=1,\dots,m\},$$

where

$$\xi^0=\begin{pmatrix}\mathbf{F}r^0\\ \mathbf{F}r^1\\ \vdots\\ \mathbf{F}r^m\\ \mathbf{B}r^0-\mathbf{B}r^1\\ \vdots\\ \mathbf{B}r^0-\mathbf{B}r^m\end{pmatrix},\quad \xi_i^j=\lambda_i^jE_j,\quad \eta_i^j=D_i^jE_j,\quad \gamma_k=E_{m+k},$$

where $E_j = (0, \cdots, \overbrace{1}^{j^{th}\ place}, 0, \cdots, 0)^* \in \hat{\mathcal{H}} = \mathcal{H}^{m+1} \times \mathbb{R}^m$ for $j = 0, \ldots, 2m$.

We can see that:

$$\hat{\mu}(\hat{r}) = \xi^0 + \sum_{j=0}^{m} \sum_{i=1}^{d} \Big( (\varphi_i^j(\hat{r}_t))^2 \eta_i^j - \kappa_i^j \xi_i^j \Big) - \sum_{j=1}^{m} \zeta^j \gamma_j, \tag{3.34}$$

where

$$\kappa_i^j = \frac{1}{2} \Bigg\{ \sum_{h=0}^{m} \varphi_i^h(\hat{r}_t) \frac{\partial \varphi_i^j(\hat{r}_t)}{\partial r^h} [\lambda_i^h] + \sum_{h=1}^{m} \left[ \frac{\partial \varphi_i^j(\hat{r})}{\partial Y^h} \beta_i^h(\hat{r}_t) \right] + 2(1 - \delta_h^0) \varphi_i^j(\hat{r}_t) \beta_i^j(\hat{r}_t) \Bigg\},$$

and $\delta_h^0$ is the Kronecker delta of indeces 0 and $h$, whereas

$$\zeta^j = \frac{1}{2} \Bigg\{ \sum_{i=1}^{d} (\beta_i^j(\hat{r}_t))^2 + \sum_{i=1}^{d} \left( \sum_{h=0}^{m} \frac{\partial \beta_i^j(\hat{r}_t)}{\partial r^h} [\lambda_i^h] \varphi_i^h(\hat{r}_t) + \sum_{h=1}^{m} \frac{\partial \beta_i^j(\hat{r}_t)}{\partial Y^h} \beta_i^h(\hat{r}_t) \right) \Bigg].$$

Moreover, for each $i = 1, \ldots, d$ the following holds:

$$\hat{\sigma}_i(\hat{r}_t) = \sum_{h=0}^{m} \varphi_i^h(\hat{r}_t) \xi_i^h + \sum_{h=1}^{m} \beta_i^h(\hat{r}_t) \gamma_h. \tag{3.35}$$

Conditions (3.34) and (3.35) imply that

$$\mathcal{L} := \{\hat{\mu}, \hat{\sigma}^i,\ i = 1, \ldots, d\}_{LA} \subseteq \mathcal{L}^1 := \{\xi^0, \xi_i^j, \eta_i^j, \gamma_k |\ j = 0, \ldots, m,\ i = 1, \ldots, d,\ k = m{+}1, \ldots, 2m\}_{LA}.$$

Therefore, if we provide conditions such that $\mathcal{L}^1$ is finite-dimensional, we will determine sufficient conditions which guarantee that $\mathcal{L}$ is finite-dimensional.

To this effect, we prove the following result, closely related to [21][ Proposition 4.2 ].

**Theorem 3.3.2.** *If $\lambda_i^j(x)$ is a quasi-exponential function for each $j = 0, \ldots, m$ and $i = 1, \ldots, d$, then the Lie algebra $\mathcal{L}_1$ is finite-dimensional.*

*Proof.* We can observe that all the vector fields which generate $\mathcal{L}_1$ are constant except $\xi^0$. This implies that, to compute $\mathcal{L}_1$ it is necessary to compute only the Lie brackets $[\xi^0, \phi]$ for each $\phi \in \mathcal{N} \setminus \{\xi^0\}$. Indeed, the Lie brackets between all the other couples of vector fields in $\mathcal{N}$ are 0, since $[v, w] = 0$ if $v, w$ are constant vector fields and $[\xi^0, \xi^0] = 0$, by definition. Moreover, since $\xi$ is linear as a function of $\hat{r}$, the Lie brackets $[\xi^0, \phi]$ are constant vector fields on $\hat{\mathcal{H}}$. Therefore, it is sufficient to compute $[\xi^0, \phi]$ for each $\phi \in \mathcal{N} \setminus \{\xi^0\}$.

If $\phi = \xi_j^j$, for each $j = 0, \dots, m$ and $i = 1, \dots, d$:

$$[\xi^0, \xi_i^j] = \frac{\partial \xi^0}{\partial \hat{r}} \xi_i^j \overbrace{\frac{\partial \xi_i^j}{\partial \hat{r}}}^{=\mathbb{O}} \xi^0 = \begin{pmatrix} \mathbf{F} & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{F} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{F} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & \mathbf{F} & 0 & \cdots & 0 \\ \mathbf{B} & -\mathbf{B} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \mathbf{B} & 0 & -\mathbf{B} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{B} & 0 & 0 & \cdots & -\mathbf{B} & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \lambda_i^j \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{F}\lambda_i^j \\ 0 \\ \vdots \\ 0 \end{pmatrix};$$

whereas if $\phi = \eta_i^j$:

$$[\xi^0, \eta_i^j] = \Big(\mathbf{F} D_i^j\Big) E_j,$$

and, finally:

$$[\xi^0, \gamma_k] = 0.$$

Iterating this procedure in order to compute the successive Lie brackets, we achieve a similar result to the one obtained in Section 3.2. In particular we can conclude that:

$$\begin{aligned} \mathcal{L}_1 =& Span\Big\{\xi^0,\ (\mathbf{F}^n \lambda_i^j) E_j,\ (\mathbf{F}^n D_i^j) E_j,\ \gamma_k \mid j = 0, \dots, m, \\ & i = 1, \dots, d\ k = 1, \dots, m,\ n \in \mathbb{N}\Big\}. \end{aligned} \tag{3.36}$$

Therefore, by Lemma 3.2.2 and through the strategy proposed in the proof of Theorem 3.2.3, if the functions $\lambda_i^j$ are QE functions, then $\{\mathbf{F}^n \lambda_i^j |\ n \in \mathbb{N}\}$ is finite-dimensional for each $i$ and $j$. Moreover, in this case, also $(\varphi_i^j(\hat{r}_t))^2 D_i^j$ are QE functions for each $i$ and $j$. Indeed, if a function is QE, also its integral function is QE and, also, the product between two QE functions is still a QE function. Hence, we obtain that $\{\mathbf{F}^n D_i^j |\ n \in \mathbb{N}\}$ is finite-dimensional for each $i$ and $j$. In conclusion, $\mathcal{L}_1$ is finite dimensional if and only if both $\{\mathbf{F}^n \lambda_i^j |\ n \in \mathbb{N}\}$ and $\{\mathbf{F}^n D_i^j |\ n \in \mathbb{N}\}$ are finite-dimensional and this holds if the functions $\lambda_j^i$ are QE. □

The previous theorem determines sufficient conditions on the functions $\lambda_i^j$ such that the Lie algebra $\mathcal{L}$ is finite-dimensional, in particular, the following proposition holds:

**Corollary 3.3.3.** *If the functions $\lambda_i^j$ defined on (3.31) are QE functions for each $i = 1, \dots, d$ and $j = 0, \dots, m$, then:*

$$dim[\mathcal{L}] := dim[\{\hat{\mu}, \hat{\sigma}_i,\ i = 1, \dots, d\}_{LA}] < +\infty.$$

### 3.3.1 Construction of finite-dimensional realizations

In the previous subsection we proved a general sufficient set of conditions which guarantees the existence of finite-dimensional realizations for a model $\mathcal{M}$, determined by a volatility term as in (3.31). In this subsection we will use Corollary 3.3.3 in order to describe the finite-dimensional realizations in the constant direction volatility case.

We assume that each function $\lambda_i^j(x)$ is QE for each $i = 1, \dots, d$ and $j = 0, \dots, m$. Under this assumption, also the functions $D_i^j(x)$ are QE. Then, by Lemma 3.2.2 there exists natural numbers $n_i^j$ and $p_i^j$ for each $i = 1, \dots, d$ and $j = 0, \dots, m$ such that the following ODEs are satisfied:

$$\mathbf{F}^{n_i^j} \lambda_i^j = \sum_{k=0}^{n_i^j - 1} c_{ki}^j \mathbf{F}^k \lambda_i^j(x), \tag{3.37}$$

$$\mathbf{F}^{p_i^j} D_i^j(x) = \sum_{k=0}^{p_i^j - 1} d_{ki}^j \mathbf{F}^k D_i^j(x), \tag{3.38}$$

for a suitable real constants $c_{ki}^j$ and $d_{ki}^j$.

In this case, the dimension $n$ of the Lie-algebra $\mathcal{L}_1$ is dominated by

$$n \leq m + 1 + \sum_{i=1}^{d} \sum_{j=0}^{m} (n_i^j + p_i^j).$$

In order to build an invariant manifold we introduce the following notation to denote a vector of the state-space $z \in \mathbb{R}^{m+1+\sum_{i=1}^d \sum_{j=0}^m (n_i^j + p_i^j)}$:

$$z = (w^0, w^1, \dots, w^m, (z_{ki}^j)^*, (x_{ki}^j)^*)^*, \text{ where} \tag{3.39}$$

$$(z_{ki}^j)^* = (z_{01}^0, z_{11}^0, \dots, z_{n_i^0 - 1,1}^0, z_{02}^0, \dots, \dots, z_{n_d^d - 1,d}^d)^* \in \mathbb{R}^{\sum_{i=1}^d \sum_{j=0}^m n_i^j}, \tag{3.40}$$

$$(x_{ki}^i)^* = (x_{01}^0, x_{11}^0, \dots, x_{p_i^0 - 1,1}^0, \dots, \dots, x_{p_d^d - 1,d}^d)^* \in \mathbb{R}^{\sum_{i=1}^d \sum_{j=0}^m p_i^j}. \tag{3.41}$$

Using this notation we construct the tangential manifold of $\mathcal{L}_1$ following the same strategy developed in Section 3.1., based on Proposition B.3.3.

This manifold is described by the following mapping:

$$G(z) = \prod_{\substack{i = 1, \dots, d \\ j = 0, \dots, m \\ k = 0, \dots, n_i^j \\ h = 0, \dots, p_i^j \\ l = 1, \dots, m}} (e^{\mathbf{F}^k \lambda_i^j E_j z_{ki}^j})(e^{\mathbf{F}^h D_i^j E_j x_{hi}^j})(e^{\gamma_l w^l})(e^{\xi_0 w^0})\hat{r}^M,$$

for an arbitrary point $\hat{r}^M = (r_0^M, \ldots, r_m^M, y_1^M, \ldots, y_m^M)^* \in \hat{\mathcal{H}}$.

We have to compute the integral curve of each vector field which determines $\mathcal{L}_1$, introduced in (3.36). The integral curve of $\xi^0(r) = (\xi_0^0, \xi_1^0, \ldots, \xi_{2m}^0)^*$ is given componentwise by:

$$\begin{cases} \frac{d}{dt}\varphi^0_{j,r_j^M}(t) = \xi_j^0(\varphi^0_{j,r_j^M(x)}) = \mathbf{F}\varphi^0_{j,r_j^M}(t), \\ \varphi^0_{j,r_j^M(x)}(0) = r_j^M \end{cases}$$

and the solution is given by:

$$\varphi^0_{j,r_j^M(x)}(t) = e^{\mathbf{F}t} r_j^M(x) = r_j^M(t+x), \qquad j = 0, \ldots, m. \tag{3.42}$$

For the last $m$ components the integral curve is:

$$\begin{cases} \frac{d}{dt}\varphi^0_{j,y^M_{j-m}}(t) = \xi_j^0(\varphi^0_{j,y^M_{j-m}}) = \mathbf{B}\varphi^0_{0,r_0^M(x)}(t) - \mathbf{B}\varphi^0_{j-m,r^M_{j-m}(x)}(t) \\ \varphi^0_{j,y^M_{j-m}}(0) = y^M_{j-m}, \end{cases}$$

and the solution of this system is given by:

$$\varphi^0_{j,y^M_{j-m}}(t) = y^M_{j-m} + \int_0^t \left[\mathbf{B}\varphi^0_{0,r_0^M(x)}(s) - \mathbf{B}\varphi^0_{j-m,r^M_{j-m}(x)}(s)\right] ds = y^M_{j-m} + \int_0^t r_0^M(s) - r^M_{j-m}(s) ds, \tag{3.43}$$

for each $j = m+1, \ldots, 2m$.

Since $(\mathbf{F}^k \lambda_i^j)$ is constant, the integral curve of $\xi_i^j$, denoted by $e^{\xi_j^j t} r^M$ is given by:

$$e^{\xi_i^j t} r^M = r_j^M + t\mathbf{F}^j \lambda_i^j, \quad j = 0, \ldots, m. \tag{3.44}$$

In the same way, the integral curve of $\eta_i^j$, $e^{\eta_i^j t} r^M$ is given by:

$$e^{\eta_i^j t} r^M(t) = r^M + t\mathbf{F}^j D_i^j, \quad j = 0, \ldots, m. \tag{3.45}$$

Finally, the integral curve of the constant vector field $\gamma_k$, $e^{\gamma_k t} r^M$ is given by:

$$\begin{cases} (e^{\gamma_k t} r^M)_j = r_j^M, & j = 0, \ldots, m, \\ (e^{\gamma_k t} r^M)_j = y_j^M + \delta_k^j t, & j = m+1, \ldots, 2m, \end{cases} \tag{3.46}$$

where $\delta_k^j$ is the Kronecker delta between indexes $j$ and $k$.

Now, we observe that we can compute the tangential manifold $G$ starting by the integral cuve $\xi^0$, (3.42), (3.43) and then the integral curves of (3.44), (3.45) and (3.46) because they have a simpler shape. Following this strategy, we obtain that each component of $G = (G^0, \ldots, G^{2m})^*$ is given by:

$$G^j(z,x) = r_j^M(w^0 + x) + \sum_{i=1}^d \left\{ \sum_{k=0}^{n_i^j - 1} z_{ki}^j \mathbf{F}^k \lambda_i^j(x) + \sum_{k=0}^{p_i^j - 1} x_{ki}^j \mathbf{F}^k \left( \lambda_i^j(x) \int_0^x \lambda_i^j(s) ds \right) \right\}, \quad j = 0, \ldots, m, \tag{3.47}$$

$$G^j(z,x) = y_j^M + \int_0^{w^0} (r_0^M(s) - r_j^M(s)) ds + w^j, \quad j = m+1, \ldots, 2m. \tag{3.48}$$

At this point, we have to compute the step 3 outlined at the end of Section 3.1. In particular, we will determine the coefficients of a finite dimensional process of the form:

$$dZ_t = a(Z_t)dt + b(Z_t) \circ dW_t,$$

such that $G_*a = \hat{\mu}$ and $G_*b_i = \hat{\sigma}_i$ for each $i = 1, \dots, d$. For simplicity we will omit the $z$ variable on the functions $a$ and $b$ and we use for those functions a notation similar to the one introduced in (3.39):

$$\begin{aligned} a =&(a^0, a^1, \dots, a^m, (a^j_{ki})^*, (\widetilde{a}^j_{ki})^*)^*, \\ b =&(b^0, b^1, \dots, b^m, (b^j_{ki})^*, (\widetilde{b}^j_{ki})^*)^*. \end{aligned}$$

We observe that, since the Brownian motion $W$ which drives the dynamics is $d$ dimensional, the term $b$, similarly as the volatility $\hat{\sigma}$ is a $d$-dimensional vector. Hence it is necessary to introduce an additional parameter $h = 1, \dots, d$ in order to compute consistency condition:

$$\begin{aligned} b &= (b_1, \dots, b_d)^*, \\ b_h &= (b^0_h, b^1_h, \dots, b^m_h, (b^j_{ki,h})^*, (\widetilde{b}^j_{ki,h})^*)^*, \quad h = 1, \dots, d. \end{aligned}$$

Therefore, for each $j = 0, \dots, m$:

$$(G^j_z(z)a)(x) = \mathbf{F}r^M_j(w^0 + x)a^0 + \sum_{i=1}^{d}\left[\sum_{k=0}^{n^j_i-1} \mathbf{F}^k \lambda^j_i(x)a^j_{ki} + \sum_{k=0}^{p^j_i-1} \mathbf{F}^k D^j_i(x)\widetilde{a}^j_{ki}\right], \quad (3.49)$$

whereas for $j = m+1, \dots, 2m$:

$$G^j_z(z)a = a^0(r^M_0(w^0) - r^M_j(w^0)) + a^j.$$

In order to obtain explicitly the condition $G^0_z(z)a = \mu^0(G(z))$, we have now to compute $\hat{\mu}(G(z))$:

$$\begin{aligned} \mu^0(G(z))(x) =&\mathbf{F}G^0(z,x) + \sum_{i=1}^{d}(\varphi^0_i(G(z)))^2 D^0_i(x) - \frac{1}{2}\sum_{i=1}^{d}\lambda^0_i(x)\left\{\sum_{j=0}^{m}\varphi^j_i(G(z))\frac{\partial\varphi^0_i}{\partial r^j}(G(z))[\lambda^j_i] + \right. \\ &\left. + \sum_{j=1}^{m}\frac{\partial\varphi^0_i}{\partial Y^j}(G(z))\beta^j_i(G(z))\right\}, \end{aligned}$$

where we exploit hypotheses (3.37) and (3.38) in order to provide the following computation:

$$
\begin{aligned}
\mathbf{F}G^0(z,x) &= \mathbf{F}r_0^M(x+w^0) + \sum_{i=1}^{d}\left\{\sum_{k=0}^{n_i^0-1} z_{ki}^0 \mathbf{F}^{k+1}\lambda_i^0(x) + \sum_{k=0}^{p_i^0-1} x_{ki}^0 \mathbf{F}^{k+1} D_i^0(x)\right\} \\
&= \mathbf{F}r_0^M(x+w^0) + \sum_{i=1}^{d}\left\{\sum_{k=1}^{n_i^0} z_{k-1,i}^0 \mathbf{F}^{k}\lambda_i^0(x) + \sum_{k=1}^{p_i^0} x_{k-1,i}^0 \mathbf{F}^{k} D_i^0(x)\right\} \\
&= \mathbf{F}r_0^M(x+w^0) + \sum_{i=1}^{d}\left\{\sum_{k=1}^{n_i^0-1} z_{k-1,i}^0 \mathbf{F}^{k}\lambda_i^0(x) + \sum_{k=1}^{p_i^0-1} x_{k-1,i}^0 \mathbf{F}^{k} D_i^0(x) + \right. \\
&\quad \left. + z_{n_i^0-1,i}^0 \sum_{k=0}^{n_i^0-1} c_{ki}^0 \mathbf{F}^k \lambda_i^0(x) + x_{p_i^0-1,i}^0 \sum_{k=0}^{p_i^0-1} d_{ki}^0 \mathbf{F}^k D_i^0(x)\right\} \\
&= \mathbf{F}r_0^M(x+w^0) + \sum_{i=1}^{d}\left\{\sum_{k=1}^{n_i^0-1} (z_{k-1,i}^0 + z_{n_i^0-1,i}^0 c_{ki}^0)\mathbf{F}^{k}\lambda_i^0(x) + \sum_{k=1}^{p_i^0-1} (x_{k-1,i}^0 + \right. \\
&\quad \left. + x_{p_i^0-1,i}^0 d_{ki}^0)\mathbf{F}^k D_i^0(x) + z_{n_i^0-1,i}^0 c_{0i}^0 \lambda_i^0(x) + x_{p_i^0-1,i}^0 d_{0i}^0 D_i^0(x)\right\},
\end{aligned}
$$

Therefore the first component of the drift term is given by:

$$
\begin{aligned}
\mu^0(G(z)) = \mathbf{F}r_0^M(x+w^0) &+ \sum_{i=1}^{d}\left\{\lambda_i^0(x)\left[z_{n_i^0-1,i}^0 c_{0i}^0 - \frac{1}{2}\left(\sum_{h=0}^{m}\varphi_i^h(G(z))\frac{\partial \varphi_i^0}{\partial r^h}(G(z))[\lambda_i^h] + \right.\right.\right. \\
&\left.\left. + \sum_{h=1}^{m}\frac{\partial \varphi_i^0}{\partial Y^h}(G(z))\beta_i^h(G(z))\right)\right] + D_i^0(x)\left[(\varphi_i^0(G(z)))^2 + x_{p_i^0-1,i}^0 d_{0i}^0\right] + \\
&\left. + \sum_{k=1}^{n_i^0-1}(z_{k-1,i}^0 + z_{n_i^0-1,i}^0 c_{ki}^0)\mathbf{F}^k\lambda_i^0(x) + \sum_{k=1}^{p_i^0-1}(x_{k-1,i}^0 + x_{p_i^0-1,i}^0 d_{ki}^0)\mathbf{F}^k D_i^0(x)\right\}.
\end{aligned} \tag{3.50}
$$

In conclusion equating (3.50) and (3.49) for $j=0$, we obtain $\mu^0(G(z)) = G_z^0(z)a$. The comparison of each term of this equation leads to:

$$
\begin{cases}
a^0 = 1, \\
a_{0i}^0 = z_{n_i^0-1,i}^0 c_{0i}^0 - \frac{1}{2}\left(\sum_{h=0}^m \varphi_i^h(G(z))\frac{\partial\varphi_i^0}{\partial r^h}(G(z))[\lambda_i^h] + \sum_{h=1}^m \frac{\partial\varphi_i^0}{\partial Y^h}(G(z))\beta_i^h(G(z))\right), \\
a_{ki}^0 = z_{k-1,i}^0 + z_{n_i^0-1,i}^0 c_{ki}^0, \qquad k=1,\dots,n_i^0-1, \\
\widetilde{a}_{0i}^0 = (\varphi_i^0(G(z)))^2 + x_{p_i^0-1,i}^0 d_{0i}^0, \\
\widetilde{a}_{ki}^0 = x_{k-1,i}^0 + x_{p_i^0-1,i}^0 d_{ki}^0, \qquad k=1,\dots,p_i^0-1
\end{cases} \tag{3.51}
$$

Following the same strategy, we compute the other coordinates of the drift term:

$$\begin{aligned}\mu^j(G(z)) =&\mathbf{F}r_j^M(x+w^0)+\sum_{i=1}^d\Bigg\{\lambda_i^j(x)\Bigg[z^j_{n_i^j-1,i}c^j_{0i}-\frac{1}{2}\Bigg(\sum_{h=0}^m\varphi_i^h(G(z))\frac{\partial\varphi_i^j}{\partial r^h}(G(z))[\lambda_i^h]+\\&+\sum_{h=1}^m\beta_i^h(G(z))\frac{\partial\varphi_i^j}{\partial Y^h}(G(z))+2\varphi_i^j(G(z))\beta_i^j(G(z))\Bigg)\Bigg]+\\&+D_i^j(x)\Big[(\varphi_i^j(G(z))^2+x^j_{p_{i-1,i}^j}d^j_{0i}\Big]+\sum_{k=1}^{n_i^j-1}(z^j_{k-1,i}+z^j_{n_i^j-1,i}c^j_{ki})\mathbf{F}^k\lambda_i^j(x)+\\&+\sum_{k=1}^{p_i^j-1}(x^j_{k-1,i}+x^j_{p_i^j-1,i}d^j_{ki})\mathbf{F}^kD_i^j(x)\Bigg\}.\end{aligned} \tag{3.52}$$

Therefore, by the comparison between (3.52) and (3.49) we obtain the equation $\mu^j(G(z))=G^j_*a$, which leads to:

$$\begin{cases}a^0=1,\\ a^j_{0i}=z^j_{n_i^j-1,i}c^j_{0i}-\frac{1}{2}\Big(\sum_{h=0}^m\varphi_i^h(G(z))\frac{\partial\varphi_i^j}{\partial r^h}(G(z))[\lambda_i^j]+\\ \qquad+\sum_{h=1}^m\beta_i^h(G(z))\frac{\partial\varphi_i^j}{\partial Y^h}(G(z))+2\varphi_i^j(G(z))\beta_i^j(G(z))\Big),\\ a^j_{ki}=z^j_{k-1,i}+z^j_{n_i^j-1,i}c^j_{ki},\quad k=1,\dots,n_i^j-1,\\ \widetilde{a}^j_{0i}=(\varphi_i^j(G(z)))^2+x^j_{p_i^j-1,i}d^j_{0i},\\ \widetilde{a}^j_{ki}=x^j_{k-1,i}+x^j_{p_i^j-1,i}d^j_{ki}\quad k=1,\dots,p_i^j-1.\end{cases} \tag{3.53}$$

Finally, for $j=m+1,\dots,2m$:

$$\begin{aligned}\mu^j(G(z)) =&\mathbf{B}G^0(z)-\mathbf{B}G^{j-m}(z)-\frac{1}{2}\sum_{i=1}^d(\beta_i^{j-m}(G(z)))^2-\frac{1}{2}\Bigg[\sum_{i=1}^d\Bigg(\sum_{h=0}^m\frac{\partial\beta_i^{j-m}}{\partial r^h}(G(z))[\lambda_i^h]\varphi_i^h(G(z))+\\&+\sum_{h=1}^m\frac{\partial\beta_i^{j-m}(G(z))}{\partial Y^h}\beta_i^h(G(z))\Bigg)\\ =&r_0^M(w^0)+\sum_{i=1}^d\sum_{k=0}^{n_i^0-1}z^0_{ki}\mathbf{F}^k\lambda_i^0(0)-r^M_{j-m}(w^0)-\sum_{i=1}^d\sum_{k=0}^{n_i^{j-m}-1}z^{j-m}_{ki}\mathbf{F}^k\lambda_i^{j-m}(0)+\\&-\frac{1}{2}\sum_{i=1}^d(\beta_i^{j-m}(G(z)))^2-\frac{1}{2}\Bigg[\sum_{i=1}^d\sum_{h=0}^m\frac{\partial\beta_i^{j-m}}{\partial r^h}(G(z))[\lambda_i^h]\varphi_i^h(G(z))+\\&+\sum_{h=1}^m\frac{\partial\beta_i^{j-m}}{\partial Y^h}(G(z))\beta_i^h(G(z))\Bigg].\end{aligned}$$

Therefore, comparing with the last $m$ components of $G_* a$, we obtain:

$$\begin{cases} a^0 = 1, \\ a^{j-m} = \sum_{i=1}^{d}\Big[\sum_{k=0}^{n_i^0-1} z_{ki}^0 \mathbf{F}^k \lambda_i^0(0) - \sum_{k=0}^{n_i^{j-m}-1} z_{ki}^j \mathbf{F}^k \lambda_i^j(0) - \frac{1}{2}(\beta_i^j(G(z)))^2 + \\ \qquad -\frac{1}{2}\Big(\sum_{h=0}^{m} \frac{\partial \beta_i^{j-m}}{\partial r^h}(G(z))[\lambda_i^h]\varphi_i^h(G(z)) + \sum_{h=1}^{m} \frac{\partial \beta_i^{j-m}}{\partial Y^h}(G(z))\beta_i^h(G(z))\Big)\Big], \end{cases} \tag{3.54}$$

for each $j = m+1, \dots, 2m$.

We follow the same procedure in order to compute the value of $b$.Therefore we analyse the equation $G_z b_h = \hat{\sigma}_h(G(z))$, for each $h = 1, \dots, d$.

Then, for $j = 0, \dots, m$:

$$\begin{aligned}(G_z^j(z)b_h)(x) &= \mathbf{F} r_j^M(w^0 + x)b_h^0 + \sum_{i=1}^{d}\Big(\sum_{k=0}^{n_i^j-1} \mathbf{F}^k \lambda_i^j(x) b_{ki,h}^j + \sum_{k=0}^{p_i^j-1} \mathbf{F}^k D_i^j(x)\widetilde{b}_{ki,h}^j\Big) \\ &= \varphi_h^j(G(z))\lambda_h^j(x).\end{aligned}$$

Therefore:

$$\begin{cases} b_h^0 = 0, \\ b_{0h,h}^j = \varphi_h^j(G(z)), \\ b_{0i,h}^j = 0, \quad i \neq h, \quad i = 1, \dots, d, \\ b_{ki,h}^j = 0, \quad k = 1, \dots, n_i^j - 1, \\ \widetilde{b}_{ki,h}^j = 0, \quad k = 0, \dots, p_i^j - 1. \end{cases} \tag{3.55}$$

On the other hand, for $j = m+1, \dots, 2m$ and $h = 1, \dots, d$:

$$G_z^j(z)b_h = b_h^0(r_0^M(w^0) - r_j^M(w^0)) + b_{i,h}^j = \beta_h^j(G(z)),$$

which leads to:

$$\begin{cases} b^0 = 0, \\ b_{h,h}^j = \beta_h^j(G(z)), \qquad j = 1, \dots, m. \\ b_{i,h}^j = 0, \qquad i \neq h, \quad j = 1, \dots, m. \end{cases} \tag{3.56}$$

In conclusion, we have proved the following result:

**Proposition 3.3.4.** *Let us consider a forward rate model $\mathcal{M}$, described by:*

$$d\hat{r}_t = \hat{\mu}(\hat{r}_t)dt + \hat{\sigma}(\hat{r}_t) \circ dW_t,$$

*where $\hat{\mu}$ and $\hat{\sigma}_i$, $i = 1, \dots, d$ are respectively determined by* (3.32) *and* (3.33) *and the functions $\lambda_i^j$ are QE for each $i = 1, \dots, d$ and $j = 0, \dots, m$. Hence, $\mathcal{M}$ possesses finite-dimensional realizations. In particular, the equation $\hat{r}_t(x) = G(Z_t, x)$ holds in a neighborhood of an initial point $\hat{r}^M$, where $G$ is defined in* (3.47)*and the finite-dimensional process $Z_t$, such that:*

$$dZ_t = a(Z_t)dt + b(Z_t) \circ dW_t,$$

*where the drift $a$ and volatility $b$ terms are described by the conditions* (3.51), (3.53), (3.54), (3.55), (3.56).

### 3.3.2 Necessary and sufficient conditions for a simplified constant direction volatility model

Let us consider a volatility term for a model as (1.33) of the form:

$$\hat{\sigma}_i(\hat{r}_t) = \varphi_i(\widetilde{r}_t) \begin{pmatrix} \lambda_i^0(x) \\ \vdots \\ \lambda_i^m(x) \\ \beta_i^1 \\ \vdots \\ \beta_i^m \end{pmatrix}, \qquad i = 1, \dots, d, \tag{3.57}$$

where $\widetilde{r}_t = (r^0, \dots, r^m)^*$ and $\beta_i^j$ are real constants. This is a simplified case of the volatility term introduced in (3.31). In particular, the scalar vector field $\varphi_i$ is the same for each component of $\hat{\sigma}_i$ and differently from the previous section, in $\varphi_i$ there is no dependence on the last $m$ components of the forward rate structure $\hat{r}$ (the ones associated with the spreads). This assumption allows us to separate the components associated with the forward rate equations to the components associated with the spreads, following the strategy outlined in Remark 2.2.9.

Under this assumption, the drift term introduced in (3.32) has the following form:

$$\hat{\mu}(\hat{r}_t) = \begin{pmatrix} \mathbf{F}r^0 + \sum_{i=1}^d \Big[(\varphi_i(\widetilde{r}_t))^2 D_i^0(x) - \frac{1}{2}\lambda_i^0(x)\varphi_i(\hat{r}_t) \sum_{h=0}^m \frac{\partial \varphi_i(\hat{r}_t)}{\partial r^h}[\lambda_i^h]\Big] \\ \mathbf{F}r^1 + \sum_{i=1}^d \Big[(\varphi_i(\widetilde{r}_t))^2 D_i^1(x) - \frac{1}{2}\lambda_i^1(x)\Big(\varphi_i(\widetilde{r}_t) \sum_{h=0}^m \frac{\partial \varphi_i(\hat{r}_t)}{\partial r^h}[\lambda_i^h] + 2(\varphi_i(\widetilde{r}_t))^2 \beta_i^1\Big)\Big] \\ \vdots \\ \mathbf{F}r^m + \sum_{i=1}^d \Big[(\varphi_i(\widetilde{r}_t))^2 D_i^m(x) - \frac{1}{2}\lambda_i^m(x)\Big(\varphi_i(\widetilde{r}_t) \sum_{h=0}^m \frac{\partial \varphi_i(\hat{r}_t)}{\partial r^h}[\lambda_i^h] + 2(\varphi_i(\widetilde{r}_t))^2 \beta_i^m\Big)\Big] \\ \mathbf{B}r^0 - \mathbf{B}r^1 - \frac{1}{2}\sum_{i=1}^d \Big[(\varphi_i(\widetilde{r}_t))^2(\beta_i^1)^2 + \beta_i^1 \varphi_i(\widetilde{r}_t) \sum_{h=0}^m \frac{\partial \varphi_i(\widetilde{r}_t)}{\partial r^h}[\lambda_i^h]\Big] \\ \vdots \\ \mathbf{B}r^0 - \mathbf{B}r^m - \frac{1}{2}\sum_{i=1}^d \Big[(\varphi_i(\widetilde{r}_t))^2(\beta_i^m)^2 + \beta_i^m \varphi_i(\widetilde{r}_t) \sum_{h=0}^m \frac{\partial \varphi_i(\widetilde{r}_t)}{\partial r^h}[\lambda_i^h]\Big] \end{pmatrix}, \tag{3.58}$$

where $D_i^j(x) := \lambda_i^j(x) \int_0^x \lambda_i^j(s)ds$, for each $j = 0, \dots, m$ and $i = 1, \dots, d$.

We aim at providing equivalent conditions such that the Lie algebra:

$$\mathcal{L} := \{\hat{\mu}, \hat{\sigma}_1, \dots, \hat{\sigma}_d\}_{LA}$$

is finite-dimensional. We consider the following vector fields:

$$\xi^i = \begin{pmatrix} \lambda_i^0(x) \\ \vdots \\ \lambda_i^m(x) \\ \beta_i^1 \\ \vdots \\ \beta_i^m \end{pmatrix}, \qquad i = 1, \dots, d, \tag{3.59}$$

and

$$\xi^0 = \begin{pmatrix} \mathbf{F}r_t^0 + \sum_{i=1}^d (\varphi_i(\widetilde{r}_t))^2 D_i^0(x) \\ \mathbf{F}r_t^1 + \sum_{i=1}^d (\varphi_i(\widetilde{r}_t))^2 D_i^1(x) - \sum_{i=1}^d \lambda_i^1(x)(\varphi_i(\widetilde{r}_t))^2 \beta_i^1 \\ \vdots \\ \mathbf{F}r_t^m + \sum_{i=1}^d (\varphi_i(\widetilde{r}_t))^2 D_i^m(x) - \sum_{i=1}^d \lambda_i^m(x)(\varphi_i(\widetilde{r}_t))^2 \beta_i^m \\ \mathbf{B}r_t^0 - \mathbf{B}r_t^1 - \frac{1}{2}\sum_{i=1}^d (\varphi_i(\widetilde{r}_t))^2 (\beta_i^1)^2 \\ \vdots \\ \mathbf{B}r_t^0 - \mathbf{B}r_t^m - \frac{1}{2}\sum_{i=1}^d (\varphi_i(\widetilde{r}_t))^2 (\beta_i^m)^2 \end{pmatrix}. \tag{3.60}$$

Therefore, by the comparison of the vector fields defined in (3.58),(3.59) and(3.60), we obtain:

$$\hat{\mu}(\hat{r}_t) = \xi^0 - \frac{1}{2}\sum_{i=1}^d \varphi_i(\widetilde{r}_t)\Big(\sum_{j=0}^m \frac{\partial \varphi_i(\widetilde{r}_t)}{\partial r^j}[\lambda_i^j]\Big)\xi^i,$$

and

$$\hat{\sigma}_i(\hat{r}_t) = \varphi_i(\widetilde{r}_t)\xi^i, \qquad i = 1, \dots, d,$$

and since $-\frac{1}{2}\sum_{i=1}^d \varphi_i(\widetilde{r}_t)\Big(\sum_{j=0}^m \frac{\partial \varphi_i(\widetilde{r}_t)}{\partial r^j}[\lambda_i^j]\Big)$ and $\varphi_i(\widetilde{r}_t)$ are scalar vector fields for each $i = 1, \dots, d$, we can conclude that:

$$\mathcal{L} = \{\xi^0, \xi^1, \dots, \xi^d\}_{LA} =: \mathcal{L}_1. \tag{3.61}$$

Therefore, the Lie algebra $\mathcal{L}$ is finite-dimensional if and only if $\mathcal{L}_1$ is finite-dimensional. In order to compute the conditions under which $dim[\mathcal{L}_1] < \infty$, we need the successive Lie brackets between the vector fields which determine $\mathcal{L}_1$.

First of all, we describe the Fréchet derivative of $\xi^i$ for $i = 1, \dots, d$, that will be denoted with $\xi^i_{\hat{r}}$, and we use the notation:

$$\Phi(\widetilde{r}_t) = ((\varphi_1(\widetilde{r}_t))^2, \dots, (\varphi_d(\widetilde{r}_t))^2)^*, \tag{3.62}$$

$$D^j(x) = (D_1^j(x), \dots, D_d^j(x))^*, \tag{3.63}$$

$$(\beta^j)^2 = ((\beta_1^j)^2, \dots, (\beta_d^j)^2)^*, \tag{3.64}$$

$$\lambda^j(x)\beta^j = (\lambda_1^j(x)\beta_1^j, \dots, \lambda_d^j(x)\beta_d^j)^*. \tag{3.65}$$

Therefore, we can denote:

$$\sum_{i=1}^{d}(\varphi_i(\widetilde{r}_t))^2 D_i^j(x) = \Phi(\widetilde{r}_t)D^j(x),$$
$$\sum_{i=1}^{d}(\varphi_i(\widetilde{r}_t))^2 (\beta_i^j)^2 = \Phi(\widetilde{r}_t)(\beta^j)^2,$$
$$\sum_{i=1}^{d}(\varphi_i(\widetilde{r}_t))^2 \lambda_i^j(x)\beta_i^j = \Phi(\widetilde{r}_t)\lambda^j(x)\beta^j.$$

Using the same notation introduced for the Fréchet derivative of $\xi^i$ for the scalar vector field $\Phi$, $\frac{\partial\Phi}{\partial r^j}(\widetilde{r}) = \Phi_{r^j}(\widetilde{r})$, we compute the Lie brackets of $\xi^0, \xi^i$ for each $i = 1, \ldots, d$. First of all, it is necessary to compute:

$$\xi^0_{\hat{r}} = \begin{pmatrix} \mathbf{F} + \Phi_{r^0}(\widetilde{r}_t)D^0 & \Phi_{r^1}(\widetilde{r}_t)D^0 & \cdots & \Phi_{r^m}(\widetilde{r}_t)D^0 & 0 & \cdots & 0 \\ \Phi_{r^0}(\widetilde{r}_t)(D^1 - \lambda^1\beta^1) & \mathbf{F} + \Phi_{r^1}(\widetilde{r}_t)(D^1 - \lambda^1\beta^1) & \cdots & \Phi_{r^m}(\widetilde{r}_t)(D^1 - \lambda^1\beta^1) & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{B} - \frac{1}{2}\Phi_{r^0}(\widetilde{r})(\beta^1)^2 & -\mathbf{B} - \frac{1}{2}\Phi_{r^1}(\widetilde{r})(\beta^1)^2 & \cdots & -\frac{1}{2}\Phi_{r^m}(\widetilde{r})(\beta^1)^2 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ \mathbf{B} - \frac{1}{2}\Phi_{r^0}(\widetilde{r})(\beta^m)^2 & -\frac{1}{2}\Phi_{r^1}(\widetilde{r})(\beta^m)^2 & \cdots & -\mathbf{B} - \frac{1}{2}\Phi_{r^m}(\widetilde{r})(\beta^m)^2 & 0 & \cdots & 0 \end{pmatrix},$$

whereas $\xi^i_{\hat{r}} = \mathbb{O}$, where $\mathbb{O}$ means that $\xi^i_{\hat{r}}$ is 0 in each element of the matrix, since the vector field $\xi^i$ is constant in $\hat{\mathcal{H}}$, for each $i = 1, \ldots, d$. Then, we compute the following Lie bracket:

$$\eta^i := [\xi^0, \xi^i] = \xi^0_{\hat{r}}(\xi^i) - \overbrace{\xi^i_{\hat{r}}(\xi^0)}^{=0} = \begin{pmatrix} \mathbf{F}\lambda_i^0 + \sum_{h=0}^m \Phi_{r^h}(\widetilde{r})[\lambda_i^h]D^0 \\ \mathbf{F}\lambda_i^1 + \sum_{h=0}^m \Phi_{r^h}(\widetilde{r})[\lambda_i^h](D^1 - \lambda^1\beta^1) \\ \vdots \\ \mathbf{F}\lambda_i^m + \sum_{h=0}^m \Phi_{r^h}(\widetilde{r})[\lambda_i^h](D^m - \lambda^m\beta^m) \\ \lambda_i^0(0) - \lambda_i^1(0) - \frac{1}{2}(\beta^1)^2\Big(\sum_{h=0}^m \Phi_{r^h}(\widetilde{r})[\lambda_i^h]\Big) \\ \vdots \\ \lambda_i^0(0) - \lambda_i^m(0) - \frac{1}{2}(\beta^m)^2\Big(\sum_{h=0}^m \Phi_{r^h}(\widetilde{r})[\lambda_i^h]\Big) \end{pmatrix}. \tag{3.66}$$

Therefore, we introduce the following notation for the second order derivative of the function $\Phi$ on the variables $r^h, r^l$, computed on the couple of vector fields $\lambda_i^h$,

$\lambda_k^l \in \mathcal{H}$: $\Phi_{r^h r^l}[\lambda_i^h, \lambda_k^l]$. We can observe that:

$$\eta^i = \begin{pmatrix} \mathbf{F}\lambda_i^0 \\ \mathbf{F}\lambda_i^1 \\ \vdots \\ \mathbf{F}\lambda_i^m \\ \mathbf{B}\lambda_i^0 - \mathbf{B}\lambda_i^1 \\ \vdots \\ \mathbf{B}\lambda_i^0 - \mathbf{B}\lambda_i^m \end{pmatrix} + \sum_{h=0}^{m} \Phi_{r^h}(\widetilde{r})[\lambda_i^h] \begin{pmatrix} D^0 \\ D^1 - \lambda^1\beta^1 \\ \vdots \\ D^m - \lambda^m\beta^m \\ -\frac{1}{2}(\beta^1)^2 \\ \vdots \\ -\frac{1}{2}(\beta^m)^2 \end{pmatrix}, \quad i = 1, \ldots, d.$$

Hence, if we compute the Fréchet derivative of $\eta^i$, we obtain:

$$\eta^i_{\hat{r}} = \begin{pmatrix} D^0 \sum_{h=0}^m \Phi_{r^h r^0}(\widetilde{r})[\lambda_i^h, \cdot] & \cdots & D^0 \sum_{h=0}^m \Phi_{r^h r^m}(\widetilde{r})[\lambda_i^h, \cdot] & 0 & \cdots & 0 \\ (D^1 - \lambda^1\beta^1) \sum_{h=0}^m \Phi_{r^h r^0}(\widetilde{r})[\lambda_i^h, \cdot] & \cdots & (D^1 - \lambda^1\beta^1) \sum_{h=0}^m \Phi_{r^h r^m}(\widetilde{r})[\lambda_i^h, \cdot] & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ (D^m - \lambda^m\beta^m) \sum_{h=0}^m \Phi_{r^h r^0}(\widetilde{r})[\lambda_i^h, \cdot] & \cdots & (D^m - \lambda^m\beta^m) \sum_{h=0}^m \Phi_{r^h r^m}(\widetilde{r})[\lambda_i^h, \cdot] & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ -\frac{1}{2}(\beta^1)^2 \sum_{h=0}^m \Phi_{r^h r^0}(\widetilde{r})[\lambda_i^h, \cdot] & \cdots & -\frac{1}{2}(\beta^1)^2 \sum_{h=0}^m \Phi_{r^h r^m}(\widetilde{r})[\lambda_i^h, \cdot] & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & \vdots & & \vdots \\ -\frac{1}{2}(\beta^m)^2 \sum_{h=0}^m \Phi_{r^h r^0}(\widetilde{r})[\lambda_i^h, \cdot] & \cdots & -\frac{1}{2}(\beta^m)^2 \sum_{h=0}^m \Phi_{r^h r^m}(\widetilde{r})[\lambda_i^h, \cdot] & 0 & \cdots & 0 \end{pmatrix}. \tag{3.67}$$

Using (3.67), we can compute the vector field $\kappa^{ik} = [\eta^i, \xi^k] = \eta^i_{\hat{r}}(\xi^k)$ for each $i, k = 1, \ldots, d$:

$$\kappa^{ik} = \begin{pmatrix} D^0 \sum_{l=0}^m \sum_{h=0}^m \Phi_{r^h r^l}[\lambda_i^h, \lambda_k^l] \\ (D^1 - \lambda^1\beta^1) \sum_{l=0}^m \sum_{h=0}^m \Phi_{r^h r^l}[\lambda_i^h, \lambda_k^l] \\ \vdots \\ (D^m - \lambda^m\beta^m) \sum_{l=0}^m \sum_{h=0}^m \Phi_{r^h r^l}[\lambda_i^h, \lambda_k^l] \\ -\frac{1}{2}(\beta^1)^2 \sum_{l=0}^m \sum_{h=0}^m \Phi_{r^h r^l}[\lambda_i^h, \lambda_k^l] \\ \vdots \\ -\frac{1}{2}(\beta^m)^2 \sum_{l=0}^m \sum_{h=0}^m \Phi_{r^h r^l}[\lambda_i^h, \lambda_k^l] \end{pmatrix} = \sum_{l=0}^m \sum_{h=0}^m \Phi_{r^h r^l}[\lambda_i^h, \lambda_k^l] \begin{pmatrix} D^0 \\ D^1 - \lambda^1\beta^1 \\ \vdots \\ D^m - \lambda^m\beta^m \\ -\frac{1}{2}(\beta^1)^2 \\ \vdots \\ -\frac{1}{2}(\beta^m)^2 \end{pmatrix}. \tag{3.68}$$

We introduce the following assumption:

**Assumption 3.3.5.** *We suppose that:*

$$\sum_{l=0}^m \sum_{h=0}^m \Phi_{r^h r^l}[\lambda_i^h, \lambda_k^l] \neq 0, \tag{3.69}$$

*for each* $i, k = 1, \ldots, d$.

We denote with $(\varphi_n(\widetilde{r}_t))^2_{r^h r^l}[\lambda_i^h, \lambda_k^l]$ the second order derive of $(\varphi_n(\widetilde{r}_t))^2$ on the variables $r^h, r^l$ acting on the couple of vector fields $\lambda_i^h$ and $\lambda_k^l$, for each $n = 1, \ldots, d$. Therefore, we can make the following observation, recalling that $\Phi$ is given by (3.62):

$$\kappa^{ik} = \sum_{l=0}^{m}\sum_{h=0}^{m} \Phi_{r^h r^l}[\lambda_i^h, \lambda_k^l] \begin{pmatrix} D^0 \\ D^1 - \lambda^1\beta^1 \\ \vdots \\ D^m - \lambda^m\beta^m \\ -\frac{1}{2}(\beta^1)^2 \\ \vdots \\ -\frac{1}{2}(\beta^m)^2 \end{pmatrix} = \sum_{n=1}^{d}\sum_{l=0}^{m}\sum_{h=0}^{m} (\varphi_n(\widetilde{r}_t))^2_{r^h r^l}[\lambda_i^h, \lambda_k^l] \begin{pmatrix} D_n^0 \\ D_n^1 - \lambda_n^1\beta_n^1 \\ \vdots \\ D_n^m - \lambda_n^m\beta_n^m \\ -\frac{1}{2}(\beta_n^1)^2 \\ \vdots \\ -\frac{1}{2}(\beta_n^m)^2 \end{pmatrix},$$

for each $i, k = 1, \ldots, d$. We introduce now the vector fields:

$$\zeta^n = \begin{pmatrix} D_n^0 \\ D_n^1 - \lambda_n^1\beta_n^1 \\ \vdots \\ D_n^m - \lambda_n^m\beta_n^m \\ -\frac{1}{2}(\beta_n^1)^2 \\ \vdots \\ -\frac{1}{2}(\beta_n^m)^2 \end{pmatrix}, \quad n = 1, \ldots, d, \tag{3.70}$$

In particular, we can observe that $\kappa^{ik} = c_n^{ik}\zeta_n$ for each $i, k = 1, \ldots, d$ where

$$c_n^{ik} := \sum_{l=0}^{m}\sum_{h=0}^{m} (\varphi_n(\widetilde{r}_t))^2_{r^h r^l}[\lambda_i^h, \lambda_k^l],$$

are real constants. Hence, the $d^2$ vector fields $\kappa^{ik}$ can be written as linear combination of the $d$ vector fields $\zeta_n$. Afterwards, we introduce the following vector fields:

$$\varrho^i = \begin{pmatrix} \mathbf{F}\lambda_i^0 \\ \mathbf{F}\lambda_i^1 \\ \vdots \\ \mathbf{F}\lambda_i^m \\ \mathbf{B}\lambda_i^0 - \mathbf{B}\lambda_i^1 \\ \vdots \\ \mathbf{B}\lambda_i^0 - \mathbf{B}\lambda_i^m \end{pmatrix}, \quad i = 1, \ldots, d. \tag{3.71}$$

We observe that the vector fields $\zeta^n$ and $\varrho^i$ are constant in the space $\hat{\mathcal{H}}$, for each

$n, i = 1, \ldots, d$. Moreover, we consider the vector field $\nu^0$, defined as follows:

$$\nu^0 = \begin{pmatrix} \mathbf{F}r^0 \\ \mathbf{F}r^1 \\ \vdots \\ \mathbf{F}r^m \\ \mathbf{B}r^0 - \mathbf{B}r^1 \\ \vdots \\ \mathbf{B}r^0 - \mathbf{B}r^m \end{pmatrix}. \tag{3.72}$$

Now we are able to prove the following result:

**Proposition 3.3.6.** *The vector fields $\zeta^n, \varrho^i, \xi^i, \nu^0$, respectively introduced in* (3.70), (3.71), (3.59) *and* (3.72), *for each $n, i = 1, \ldots, d$, determine the Lie algebra $\mathcal{L}$, i.e.*

$$\mathcal{L} = \{\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_d\}_{LA} = \{\zeta^i, \varrho^i, \xi^i, \nu^0 |\ i, = 1, \ldots, d\}_{LA} =: \mathcal{L}_2. \tag{3.73}$$

*Proof.* We proved that $\mathcal{L} = \{\xi^0, \ldots, \xi^d\}_{LA}$ in (3.61). We observe that:

$$\xi^0 = \nu^0 + \sum_{n=1}^{d} (\varphi_n(\widetilde{r}))^2 \zeta^n, \tag{3.74}$$

$$\eta^i = \varrho^i + \sum_{n=1}^{d} \sum_{h=0}^{m} (\varphi_n(\widetilde{r}))^2_{r^h} [\lambda_i^h] \zeta^n, \quad i = 1, \ldots, d, \tag{3.75}$$

$$\kappa^{ik} = \sum_{n=1}^{d} \sum_{l=0}^{m} \sum_{h=0}^{m} (\varphi_n(\widetilde{r}))^2_{r^h r^l} [\lambda_i^h, \lambda_k^l] \zeta^n, \quad i, k = 1, \ldots, d. \tag{3.76}$$

Hence, if we compute the successive Lie brackets between two elements of $\mathcal{L}_1$, we can exploit the bilinearity of the Lie brackets, equations (3.74), (3.75), (3.76) and Lemma B.3.5, in order to substitute

$$\begin{aligned} \kappa^{ik}, \quad i, k = 1, \ldots, d &\longrightarrow \zeta^n, \quad n = 1, \ldots, d, \\ \eta^i, \quad i = 1, \ldots, d &\longrightarrow \varrho^i, \quad i = 1, \ldots, d, \\ \xi^0 &\longrightarrow \nu^0, \end{aligned}$$

in $\{\xi^0, \xi^1, \ldots, \xi^d\}_{LA} = \{\xi^0, \xi^i, \eta^i, \kappa^{ki} |\ i, k = 1, \ldots, d\}_{LA} = \{\nu^0, \xi^i, \zeta^n, \varrho^i |\ i, n = 1, \ldots, d\}_{LA}$. □

By Proposition 3.3.6, $\mathcal{L}$ is finite-dimensional if and only if the Lie algebra $\mathcal{L}_2$ is finite-dimensional.

We observe that all the vector fields in $\mathcal{L}_2$ are constants except $\nu^0$. Hence, by an analogous strategy to the one provided in the proof of Theorem 3.3.2, we determine equivalent conditions under which $\dim[\mathcal{L}_2] < +\infty$. Indeed, if we compute $[v, w]$ where $v, w$ are constant vector fields we obtain $[v, w] = 0$ and by definition $[\nu^0, \nu^0] = 0$. Therefore, it suffices to compute $[\nu^0, \phi]$ where $\phi \in \{\xi^i, \zeta^n, \varrho^i |\ i, n = 1, \dots, d\}$.

In particular, in analogy to the strategy developed at the beginning of Section 3.2 for the constant volatility case, we can observe that the Fréchet derivative of $\nu^0$ is the same of the one computed for the drift term in (3.7). Hence, the Lie brackets involving $\nu^0$ are given by:

$$[\nu^0, \xi^i] = \begin{pmatrix} \mathbf{F}\lambda_i^0 \\ \vdots \\ \mathbf{F}\lambda_i^m \\ \mathbf{B}\lambda_i^0 - \mathbf{B}\lambda_i^1 \\ \mathbf{B}\lambda_i^0 - \mathbf{B}\lambda_i^2 \\ \vdots \\ \mathbf{B}\lambda_i^0 - \mathbf{B}\lambda_i^m \end{pmatrix}, \quad [\nu^0, \varrho^i] = \begin{pmatrix} \mathbf{F}^2\lambda_i^0 \\ \vdots \\ \mathbf{F}^2\lambda_i^m \\ \mathbf{BF}\lambda_i^0 - \mathbf{BF}\lambda_i^1 \\ \vdots \\ \mathbf{BB}\lambda_i^0 - \mathbf{BF}\lambda_i^m \end{pmatrix},$$

and the analogous result is obtain for $\zeta^n$, $n = 1, \dots, d$:

$$[\nu^0, \zeta^n] = \begin{pmatrix} \mathbf{F}D_n^0 \\ \vdots \\ \mathbf{F}D_n^m - \beta_n^m \mathbf{F}\lambda_n^m \\ \mathbf{B}D_n^0 - \mathbf{B}D_n^1 + \beta_n^1 \mathbf{B}\lambda_n^1 \\ \vdots \\ \mathbf{B}D_n^0 - \mathbf{B}D_n^m + \beta_n^m \mathbf{B}\lambda_n^m \end{pmatrix}.$$

Therefore, iterating this procedure, we obtain that the Lie algebra $\mathcal{L}_2$ is given by:

$$\mathcal{L}_2 = Span\left\{ \nu^0, \varrho^i, \zeta^n, \xi^i, \phi^{i,k} = \begin{pmatrix} \mathbf{F}^k\lambda_i^0 \\ \vdots \\ \mathbf{F}^k\lambda_i^m \\ \mathbf{BF}^{k-1}\lambda_i^0 - \mathbf{BF}^{k-1}\lambda_i^1 \\ \vdots \\ \mathbf{BF}^{k-1}\lambda_i^0 - \mathbf{BF}^{k-1}\lambda_i^m \end{pmatrix}, \right.$$

$$\left. \psi^{i,k} = \begin{pmatrix} \mathbf{F}^k D_i^0 \\ \mathbf{F}^k D_i^1 - \beta_i^1 \mathbf{F}^k \lambda_i^1 \\ \vdots \\ \mathbf{F}^k D_i^m - \beta_i^m \mathbf{F}^k \lambda_i^m \\ \mathbf{BF}^{k-1} D_i^0 - \mathbf{BF}^{k-1} D_i^1 + \beta_i^1 \mathbf{BF}^{k-1}\lambda_i^1 \\ \vdots \\ \mathbf{BF}^{k-1} D_i^0 - \mathbf{BF}^{k-1} D_i^m + \beta_i^m \mathbf{BF}^{k-1}\lambda_i^m \end{pmatrix} |i, n = 1, \dots, d,\ k \in \mathbb{N} \right\}.$$

Hence, a necessary condition for

$$dim[\mathcal{L}_2] < +\infty \tag{3.77}$$

is:

$$dim[Span\{\nu^0, \xi^i, \phi^{i,k} |\ i = 1, \ldots, d,\ k \in \mathbb{N}\}] < +\infty, \tag{3.78}$$

and this is equivalent to

$$dim[Span\{\phi^{i,k} |\ k \in \mathbb{N}\}] < +\infty, \quad i = 1, \ldots, d. \tag{3.79}$$

In equivalence, we have to prove that:

$$\forall\ i = 1, \ldots, d\ \exists p^i \in \mathbb{N} :\ \phi^{i,p^i} = \sum_{k=0}^{p^i-1} \alpha_{k,i} \phi^{i,k}, \tag{3.80}$$

for a suitable set of real coefficients $\{\alpha_{k,i}\}$. At this point, we recall the proof of ($\Rightarrow$) part of Theorem 3.2.3. In particular, by firs $m + 1$ components of (3.80), we can affirm:

$$\begin{cases} \mathbf{F}^{p^i} \lambda_i^0 = \sum_{k=0}^{p^i-1} \alpha_{k,i} \mathbf{F}^k \lambda_i^0, \\ \vdots \\ \mathbf{F}^{p^i} \lambda_i^m = \sum_{k=0}^{p^i-1} \alpha_{k,i} \mathbf{F}^k \lambda_i^m. \end{cases}$$

By the previous system and applying Lemma 3.2.2, we conclude that

$$\lambda_i^j \text{ are QE functions} \quad \forall\ i = 1, \ldots, d,\ j = 0, \ldots, m. \tag{3.81}$$

Vice versa, if $\lambda_i^j$ are QE function for each $i = 1, \ldots, d$ and $j = 0, \ldots, m$, we can follow the strategy outlined for $\sigma_i^j(x)$ in the proof of ($\Leftarrow$) of Theorem 3.2.3, in order to conclude that $Span[\phi^{i,k} | k \in \mathbb{N}]$ is finite-dimensional for every $i = 1, \ldots, d$.

At the moment, we have shown that condition (3.81) is equivalent to a necessary condition for (3.77). Therefore, we aim at proving that (3.81) is also a sufficient condition for (3.77). In particular, if $\lambda_i^j(x)$ are QE functions also

$$dim[Span\{\psi^{i,k} |\ k \in \mathbb{N}\}] < \infty. \tag{3.82}$$

Indeed, if $\lambda_i^j(x)$ are QE functions, $D_i^j(x)$ and $D_i^j(x) - \beta_i^j \lambda_i^j(x)$ are QE functions for each $i = 1, \ldots, d$ and $j = 0, \ldots, m$, because the integral of a QE function is a QE function too and a linear combination of QE functions is a QE function too. By Lemma 3.2.2, for each $i = 1, \ldots, d$, we can provide the common minimal annihilator $M_i$ for $D_i^j$ and $D_i^j - \beta_i^j \lambda_i^j$, for each $j = 0, \ldots, m$. To do this, we exploit an analogous strategy to he one provided in the proof of Theorem 3.2.3 part ($\Leftarrow$). In particular, the following equations hold:

$$M_i(\mathbf{F})D_i^j = 0, \quad j = 0, \dots, m \tag{3.83}$$
$$M_i(\mathbf{F})(D_i^j - \lambda_i^j \beta_i^j) = 0, \quad j = 1, \dots, m, \tag{3.84}$$

for each $i = 1, \dots, d$. Moreover (3.83) implies that $M_i(\mathbf{F})D_i^j(x) = 0$ for each $x \in \mathbb{R}_+$, therefore:

$$\mathbf{B}(M_i(\mathbf{F})D_i^j(x)) = 0, \quad j = 0, \dots, m, \tag{3.85}$$
$$\mathbf{B}(M_i(\mathbf{F})(D_i^j(x) - \beta_i^j \lambda_i^j(x))) = 0, \quad j = 1, \dots, m. \tag{3.86}$$

Finally, we observe that:

$$M_i(\mathbf{F})D_i^j = 0 \Rightarrow M_i(\mathbf{F})(\mathbf{F}D_i^j) = 0, \quad j = 0, \dots, m, \tag{3.87}$$
$$M_i(\mathbf{F})(D_i^j(x) - \beta_i^j \lambda_i^j(x)) = 0 \Rightarrow M_i(\mathbf{F})(\mathbf{F}D_i^j(x) - \beta_i^j \mathbf{F}\lambda_i^j(x)) = 0, \quad j = 1, \dots, m. \tag{3.88}$$

Rewriting in components equations (3.85),(3.86),(3.87),(3.88), we obtain:

$$\begin{cases} \mathbf{F}^{q^i+1} D_i^0 = \sum_{k=0}^{q^i} \gamma_{k,i} \mathbf{F}^k D_i^0, \\ \mathbf{F}^{q^i+1}(D_i^1 - \beta_i^1 \lambda_i^1) = \sum_{k=0}^{q^i} \gamma_{k,i} \mathbf{F}^k (D_i^1 - \beta_i^1 \lambda_i^1), \\ \vdots \\ \mathbf{F}^{q^i+1}(D_i^m - \beta_i^m \lambda_i^m) = \sum_{k=0}^{q^i} \gamma_{k,i} \mathbf{F}^k (D_i^m - \beta_i^m \lambda_i^m), \\ \mathbf{B}\mathbf{F}^{q^i} D_i^0 - \mathbf{B}\mathbf{F}^{q^i} D_i^0 - \mathbf{B}\mathbf{F}^{q^i} D_i^1 + \beta_i^1 \mathbf{F}^{q^i} \lambda_i^1 = \sum_{k=0}^{q^i} \gamma_{k,i} (\mathbf{B}\mathbf{F}^{k-1} D_i^0 - \mathbf{B}\mathbf{F}^{k-1} D_i^1 + \\ \qquad + \beta_i^1 \mathbf{B}\mathbf{F}^{k-1} \lambda_i^1), \\ \vdots \\ \mathbf{B}\mathbf{F}^{q^i} D_i^0 - \mathbf{B}\mathbf{F}^{q^i} D_i^0 - \mathbf{B}\mathbf{F}^{q^i} D_i^m + \beta_i^m \mathbf{B}\mathbf{F}^{q^i} \lambda_i^m = \sum_{k=0}^{q^i} \gamma_{k,i} (\mathbf{B}\mathbf{F}^{k-1} D_i^0 - \mathbf{B}\mathbf{F}^{k-1} D_i^m + \\ \qquad + \beta_i^m \mathbf{B}\mathbf{F}^{k-1} \lambda_i^m), \end{cases}$$

for a suitable set of real coefficients $\{\gamma_{k,i}\}$. The last $m$ components are obtained computing the difference between (3.85) for $j = 0$ and (3.86). The previous system is equivalent to:

$$\psi^{i,q^i+1} = \sum_{k=0}^{q^i} \gamma_{k,i} \psi^{i,k}, \quad i = 1, \dots, d,$$

which implies that (3.82) holds. In conclusion, if (3.81) holds, both (3.79) and (3.82) hold. But in this case (3.77) holds. Hence, we have shown that $\lambda_i^j(x)$ is a QE function for each $i = 1, \dots, d$ and $j = 0, \dots, m$ is equivalent to $dim[\mathcal{L}_2] < \infty$.

In conclusion, we have proved the following proposition:

**Proposition 3.3.7.** *Given a model $\mathcal{M}$ described by a forward rate system of SDEs as* (1.33) *and determined by a forward volatility term $\hat{\sigma}$ of the form* (3.57)*, such that Assumption 3.3.5 holds, the Lie algebra $\mathcal{L} = \{\hat{\mu}, \hat{\sigma}_1, \ldots, \hat{\sigma}_d\}$ is finite-dimensional if and only if the functions $\lambda^i_j$ are QE functions for each $i = 1, \ldots, d$ and $j = 0, \ldots, m$.*

# Appendix A

# Interest Rate Models in a Pre-crisis Framework

This Appendix aims at describing the structure of fixed-income market models, before the last financial crisis. We need this, because we want to understand why the framework adopted until $2007-2008$ is no longer appropriate. To develop these contents we based on [3].

## A.1 Zero-Coupon-Bonds and interest rate processes

Fixed-income instruments are contracts, which form the fixed-income market, that guarantee to the holder a fixed (deterministic) amount of money at a given date $T$, called maturity date.

In a pre-crisis environment every fixed-income contract can be determined, through no-arbitrage considerations, by a portfolio composed of Zero-Coupon-Bond contracts. These instruments are defined as follows:

**Definition A.1.1.** *A* Zero-Coupon-Bond (ZCB) *with a maturity date $T$, is a contract which guarantees to the holder* 1 *unit of currency to be paid at date $T$. We will denote the price at time $t \leq T$ of this contract as $B_t(T)$.*

In this framework the fixed-income market is formed by all the *ZCBs*. This market is supposed to respect the following assumptions:

**Assumption A.1.2.**

- *The relation $B_t(t) = 1$ holds $\forall t \geq 0$;*
- *For each $t \in [0, T]$ the price $B_t(T)$ is a differentiable function with respect to the time maturity $T$.*

Then, we denote $\Gamma_t := \{(T, B_t(T)), T \in \mathbb{R}_+\}$ as the bond price curve at $t$. $\Gamma_t$ represents the term structure of the bond price process.
Therefore, in our dissertation, we suppose that for each $T$ fixed the price $p(\cdot, T)$ is a scalar stochastic process, whose trajectory is driven by a $d$-dimensional Brownian motion, $W_t$.

This market is composed by an infinite number of assets (one for each maturity time $T$), therefore, one of the main problems which we have to face is to find relations between prices associated with different maturities, in order to ensure the absence of arbitrage opportunity. Hence, it is convenient to introduce the concept of interest rate, which describes the relation between $T$ bonds computed by different maturities, $T$ and $T+\delta$. Recalling the notation of [3], we define:

**Definition A.1.3.**

1. *The* ***simple forward rate*** *for* $[T, T+\delta]$*, contracted at time* $t$*, is defined as:*
$$L(t; T, T+\delta) = -\frac{B_t(T+\delta) - B_t(T)}{\delta B_t(T+\delta)}. \tag{A.1}$$

2. *The* ***simple spot rate*** *for* $[T, T+\delta]$*, is defined as:*
$$L(T, T+\delta) := -\frac{B_T(T+\delta) - 1}{\delta B_T(T+\delta)}. \tag{A.2}$$

3. *The* ***instantaneous forward rate*** *with maturity* $T$*, contracted at* $t$*, is defined by:*
$$f_t(T) = -\frac{\partial \log B_t(T)}{\partial T}. \tag{A.3}$$

4. *The* ***instantaneous short rate*** *at time* $t$ *is defined as:*
$$r(t) = f_t(t).$$

**Remark A.1.4.** *Before the last financial crisis, the simple forward rate and the simple spot rate denoted the LIBOR rate (forward and spot respectively), but, as it is described in Chapter* 1*, these equivalences, in general, do not hold anymore.*

Instead of analyzing directly the evolution of prices it is convenient to study the evolution of forward rate processes. Indeed, bond prices can be determined by instantaneous forward rate as described in the following lemma:

**Lemma A.1.5.** $\forall t \le S \le T$ *it holds:*
$$B_t(T) = B_t(S) \cdot exp\Big(-\int_S^T f_t(u)du\Big)$$
*In particular, if* $S = t$*, it holds that:* $B_t(T) = exp\Big(-\int_t^T f_t(u)du\Big)$.

Finally we define the money account $B_t$ as:

$$B_t = exp\Big\{\int_0^t r_s ds\Big\} \tag{A.4}$$

which is equivalent to:

$$\begin{cases} dB_t = r_t B_t dt \\ B_0 = 1 \end{cases}$$

We will use the money account as the numeraire for a martingale measure $\mathbb{Q}$.

## A.2 Relation between interest rates and $ZCB$ prices

In the previous section we have introduced the structure which characterizes a fixed-income market. Now we want to show explicit relations between the processes defined before, under suitable assumptions.

First of all, let us consider a stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}, W_t)$, where $(W_t)_t$ is a $\mathbb{Q}$-Wiener process. On this space we define the following processes:

$$\begin{aligned} \text{[Short rate] } dr_t &= a(t)dt + b(t)dW_t; && \text{(A.5)} \\ r_0 &= r^M && \text{(A.6)} \\ \text{[Price] } dB_t(T) &= B_t(T)m(t,T)dt + B_t(T)v(t,T)dW_t; && \text{(A.7)} \\ B_0(T) &= B_0^M(T) && \text{(A.8)} \\ \text{[Forward rate] } df_t(T) &= \alpha_t(T)dt + \sigma_t(T)dW_t. && \text{(A.9)} \\ f_0(T) &= f_0^M(T) && \text{(A.10)} \end{aligned}$$

where we assume that the initial conditions can be determined by market data.

In order to respect the assumption A.1.2, it holds:

**Assumption A.2.1.**

- *$a(t), b(t)$ are scalar adapted processes: $a(t), b(t) \in \mathcal{F}_t,\ \forall t \geq 0$;*
- *$m(t,T), v(t,T), \alpha_t(T), \sigma_t(T)$ are a 1-parametric family (on T-variable) of adapted processes, such that each of them is $\mathcal{C}^1(\mathbb{R})$ on T-variable (we will use $m_T(t,T)$ to denote the partial T-derivative).*
- *It is supposed that each dynamics allows to differentiate under the integral.*

The next proposition analyzes how those processes are related each other.

**Proposition A.2.2.** *Under Assumption A.2.1 the following hold:*

1. *If $f_t(T)$ satisfies* (A.9)*, then the short rate satisfies* (A.5)*, where:*

$$\begin{cases} a(t) = \frac{\partial}{\partial T} f(t,t) + \alpha(t,t); \\ b(t) = \sigma(t,t). \end{cases}$$

2. *If $f_t(T)$ satisfies* (A.9)*, then the price satisfies*

$$dB_t(T) = B_t(T)\big\{r(t) + A(t,T) + \frac{1}{2}||S(t,T)||^2\big\}dt + B_t(T)S(t,T)dW_t, \quad \text{(A.11)}$$

*where*

$$A(t,T) = -\int_t^T \alpha_t(s)ds; \quad \text{(A.12)}$$

$$S(t,T) = -\int_t^T \sigma_t(s)ds; \quad \text{(A.13)}$$

*Proof.* **Part 1**

By definition we have that $r_t = f_t(t)$ then, by definition:

$$r_t = f_0(t) + \int_0^t \alpha_s(t)ds + \int_0^t \sigma_s(t)dW_s.$$

Hence, by the fundamental theorem of calculus, we have that:

$$\begin{cases} \alpha_s(t) = \alpha_s(s) + \int_s^t \frac{\partial}{\partial T}\alpha_s(u)du \\ \\ \sigma_s(t) = \sigma_s(s) + \int_s^t \frac{\partial}{\partial T}\sigma_s(u)du \end{cases}$$

Therefore:

$$\begin{aligned} r_t =& f_0(t) + \int_0^t \alpha_s(s)ds + \int_0^t \Big(\int_s^t \frac{\partial}{\partial T}\alpha_s(u)du\Big)ds + \int_0^t \sigma_s(s)dW_s + \\ &+ \int_0^t \Big(\int_s^t \frac{\partial}{\partial T}\sigma_s(u)du\Big)dW_s = \\ =& \int_0^t \alpha_s(s)ds + \int_0^t \sigma_s(s)dW_s + \int_0^t \frac{\partial}{\partial T}f_0(s)ds + \overbrace{f_0(0)}^{r_0} + \int_0^t \int_0^u \frac{\partial}{\partial T}\alpha_s(u)dsdu + \\ &+ \int_0^t \int_0^u \frac{\partial}{\partial T}\sigma_s(u)dW_s ds = \\ \overset{F.T.}{=}& r_0 + \int_0^t \Big[\alpha_u(u) + \Big(\frac{\partial}{\partial T}f_0(u) + \int_0^u \frac{\partial}{\partial T}\alpha_s(u)ds + \int_0^u \frac{\partial}{\partial T}\sigma_s(u)dW_s\Big)\Big]du + \int_0^t \sigma_s(s)dW_s = \\ =& r_0 + \int_0^t \Big(\alpha_u(u) + \frac{\partial}{\partial T}f_u(u)\Big)du + \int_0^t \sigma_s(s)dW_s, \end{aligned}$$

where we have used the stochastic version of Fubini theorem (for the proof see [14], chapter 6, Theorem 6.2), and the possibility to differentiate under the integral sign. We can conclude that:

$$dr_t = \Big(\alpha_t(t) + \frac{\partial}{\partial T} f_t(t)\Big)dt + \sigma_t(t)dW_t.$$

**Part 2**

First of all, we define the following process: $Y_t(T) = -\int_t^T f_t(s)ds$, which means: $B_t(T) = \exp Y_t(T)$.
Using Itô's formula:

$$\begin{cases} dB_t(T) = exp(Y_t(T))d(Y_t(T))) + \frac{1}{2}exp(Y_t(T))d\langle Y_t(T)\rangle^2 \\ dY_t(T) = d\Big(-\int_t^T f_t(s)ds\Big) \end{cases}$$

Therefore, using Itô's formula and the integral version of A.9, we can compute the differential of $Y_t(T)$, and we can use it to compute the dynamics of $B_t(T)$:

$$\begin{aligned} Y_t(T) &= -\Big\{\int_t^T f_0(s)ds + \int_t^T \int_0^t \alpha_u(s)duds + \int_t^T \int_0^t \sigma_u(s)dW_u ds\Big\} = \\ &= -\int_t^T f_0(s)ds - \int_t^T \int_0^T \alpha_u(s)duds - \int_t^T \int_0^t \sigma_u(s)dW_u ds = \\ &\overset{F.T}{=} -\int_0^T f_0(s)ds - \int_0^t \int_u^T \alpha_u(s)dsdu - \int_0^t \int_u^T \sigma_u(s)dsdW_u + \\ &\quad + \int_0^t f_0(s)ds + \int_0^t \int_u^t \alpha_u(s)dsdu + \int_0^t \int_u^t \sigma_u(s)dsdW_u = \\ &= -\int_0^T f_0(s)ds - \int_0^t \int_u^T \alpha_u(s)dsdu - \int_0^t \int_u^T \sigma_u(s)dsdW_u + \\ &\quad + \int_0^t f_0(s)ds + \int_0^t \int_0^s \alpha_u(s)duds + \int_0^t \int_0^s \sigma_u(s)dW_u ds = \\ &= -\int_0^T f_0(s)ds - \int_0^t \int_u^T \alpha_u(s)dsdu - \int_0^t \int_u^T \sigma_u(s)dsdW_u + \\ &\quad + \int_0^t \underbrace{\Big\{f_0(s)ds + \int_0^s \alpha_u(s)du + \int_0^s \sigma_u(s)dW_u\Big\}}_{r_s} ds = \\ &= -\int_0^T f_0(s)ds + \int_0^t \Big\{r_s + A(s,T)\Big\}ds + \int_0^t \Sigma(s,T)dW_s, \end{aligned}$$

where:

$$\begin{cases} A(s,T) = -\int_s^T \alpha_s(u)du \\ \Sigma(s,T) = -\int_s^T \sigma_s(u)du \end{cases}$$

In conclusion, we have:

$$d(Y_t(T)) = (r_t + A(t,T))dt + \Sigma(t,T)dW_t$$

Hence the quadratic covariation is:

$$d\langle Y_t(T)\rangle^2 = ||\Sigma(t,T)||^2 dt$$

Therefore we can conclude:

$$dB_t(T) = B_t(T)\Big[r_t + A(t,T) + \frac{1}{2}||\Sigma(t,T)||^2\Big]dt + B_t(T)\Sigma(t,T)dW_t$$

The third equivalence is determined by a stochastic version of Fubini Theorem. This results is provided by Filipović in [14], chapter 6, Theorem 6.2. □

## A.3 Heath-Jarrow-Morton Framework

In the previous sections we have analyzed the theoretical results concerning fixed-income-market. If we want to model the market, we have to specify the processes which have been introduced above.

In order to do this, we can follow different strategies. The main approaches are:

- Short rate models;
- Forward rate models.
- LIBOR market models;

The first method consists in defining parameters for the short rate process (A.5), whereas the second one is obtained specifying the dynamics of (A.9). Finally, given a set of tenor $\mathcal{D}$, a LIBOR market model is formed by a discrete family of log-normal stochastic processes, each of them describing the dynamics of a forward rate associated with a tenor $\delta \in \mathcal{D}$.

We follow the second approach, because it is too restrictive to assume that the whole money market is governed by only one stochastic differential equation. Hence, we describe the market with an infinite system of stochastic differential equations (one for each maturity). Using this construction, we can define:

**Definition A.3.1.** *The* Heath-Jarrow-Morton (HJM) *framework is a family of models for the fixed-income market, built assuming that: for every $T > 0$, the* *dynamics of forward rate $f(\cdot, T)$ is described by the following stochastic differential equation, defined on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{Q})$ where $\mathbb{Q}$ is a martingale measure:*

$$df_t(T) = \alpha_t(T)dt + \sigma_t(T)dW_t \tag{A.14}$$

$$f_0(T) = f_0^M(T) \tag{A.15}$$

*where $W$is a $\mathbb{Q}$ Brownian motion the processes $\alpha.(T), \sigma.(T)$ are adapted.*

One of the main problem is the choice of the parameters in the previous system, in order to have an arbitrage-free fixed-income market.

To solve this problem, we remember that a ZCB is a contract which guarantees 1 at maturity date $T$. By the pricing formula. we get that:

$$B_0(T) = \mathbb{E}^{\mathbb{Q}}\Big[exp\Big\{-\int_0^T r_s ds\Big\}\Big] \tag{A.16}$$

where $\mathbb{Q}$ is a martingale measure.
Then, if we recall Lemma A.1.5, we have that:

$$B_t(T) = exp\Big\{-\int_t^T f_t(s)ds\Big\} \tag{A.17}$$

where it holds that: $r_s = f_s(s)$.

Comparing (A.16) and (A.17), we can determine a condition on the drift of the price process, called $HJM$ drift condition. We describe this condition in the following Proposition:

**Proposition A.3.2** (HJM drift condition). *Under the martingale measure $\mathbb{Q}$, the processes $\alpha$ and $\sigma$ must satisfy the following relation, for every $t$ and every $T \geq t$:*

$$\alpha_t(T) = \sigma_t(T)\int_t^T \sigma_t(s)^* ds \tag{A.18}$$

*where $A^*$ denote the transpose of the vector (or the matrix) $A$.*

*Proof.* First of all we recall that, by Proposition A.2.2 we have that:

$$dB_t(T) = B_t(T)\Big[r_t + A(t,T) + \frac{1}{2}||\Sigma(t,T)||^2\Big]dt + B_t(T)S(t,T)dW_t$$

Therefore, thanks to the fact that $\mathbb{Q}$ is a martingale measure, the drift term of the previous equation has to be equal to the short rate $r_t$. Thus we get:

$$r_t + A(t,T) + \frac{1}{2}||\Sigma(t,T)||^2 = r_t$$

this means that:

$$-\int_s^T \alpha_s(u)du + \frac{1}{2}\Big\|\int_s^T \sigma_s(u)du\Big\|^2 = 0$$

If we differentiate the previous equation in the T-variable we get the thesis. □

### A.3.1 Musiela parameterization

For our results it is more convenient to adopt an equivalent parameterization to describe the forward rate. Instead of describing the dynamics as an infinite family of SDEs, parameterized with the T-variable, we choose the *Musiela parameterization*, which does not consider the maturity time $T$, but the time to maturity $x := T - t$. In terms of $x$ the forward rate will become:

$$r_t(x) = f_t(t + x), \quad x \geq 0. \tag{A.19}$$

In order to analyze the dynamics of the forward rate parameterized in this way, we recall the following result:

**Proposition A.3.3** (Musiela equation)**.** *Assume that $f_t(T)$ is specified as in* (A.9)*. Then:*

$$dr_t(x) = \{\boldsymbol{F} r_t(x) + \sigma_t(t + x) \int_0^x \sigma_t(t + s)^* ds\}dt + \sigma_t(t + x)dW_t, \tag{A.20}$$

*where* $\boldsymbol{F} = \frac{\partial}{\partial x}$.

*Proof.* Using Itô's formula for processes which is stochastic in t-variable, but also it has a component which is a differential function in that variable, we have that:

$$dr_t(x) = df_t(t + x) + \frac{\partial}{\partial T} f_t(t + x)dt.$$

Then, computing the Itô differential for the first term, we obtain:

$$\begin{aligned} dr_t(x) &= \alpha_t(t + x)dt + \sigma_t(t + x)dW_t + \frac{\partial}{\partial x} r_t(x)dt \\ \text{Drift condition} &= \sigma_t(t + x) \int_t^{t+x} \sigma_t(s)^* ds + \sigma_t(t + x)dW_t + \frac{\partial}{\partial x} r_t(x)dt \qquad \text{(A.21)} \\ &= \{\mathbf{F} r_t + \sigma_t(t + x) \int_0^x \sigma_t(t + s)^* ds\}dt + \sigma_t(t + x)dW_t. \end{aligned}$$

□

# Appendix B

# Differential Geometry On An Infinite Dimensional Vector Space

In this chapter we aim to describe a geometric theory necessary to provide some results, on the geometric properties of forward interest rate curves.

In the first section, we will provide the main concepts of a general theory of varieties, defined on a Banach space. In particular, we will give the definition of $\mathcal{H}$-variety, where $\mathcal{H}$ is a Banach space. Then we will introduce the concepts of tangent space and tangent bundle, which are essential to understand the most important class of objects we need: the distributions. Then, by relying on the concept of Lie bracket, we will studying the notion of involutive distribution.
In the second section, we will show some preliminary propositions and remarks necessary to prove the Frobenius theorem. Finally, we will introduce the concept of Lie algebra, which will be foundamental in order to describe how to provide final dimensional realizations for a forward rate model.

## B.1 A brief introduction on infinite dimensional differential geometry

In this section, we recall some basic notions of differential geometry. Our presentation is based on [16].

Let us consider a Banach space $(\mathcal{H}, ||\cdot||)$, where $||\cdot||$ denotes the norm defined on the $\mathbb{R}$-vector space $\mathcal{H}$. In this dissertation, we admit the case in which $\mathcal{H}$ is infinite-dimensional.

### B.1.1 $\mathcal{H}$-manifolds

To introduce the concept of manifold defined on a Banach space $\mathcal{H}$, it is necessary to give the definition of *compatible Atlas* of a topological space $\mathcal{X}$:

**Definition B.1.1.** *An* atlas *on a connected topological space $\mathcal{X}$ is a collection of pairs $\{(U_i, \varphi_i)\}_{i\in\mathcal{I}}$ ($\mathcal{I}$ is an arbitrary set of indexes), which satisfies:*

- *Each $U_i$ is a subset of $\mathcal{X}$, $\forall i$ and $\{U_i\}_{i\in\mathcal{I}}$ cover $\mathcal{X}$;*
- *Each $\varphi_i$ is a bijection between $U_i$ and an open subset of a Banach space $\mathcal{H}$. Moreover, we suppose that for any $i, j$ : $\varphi_i(U_i \cap U_j)$ is open in $\mathcal{H}$.*
- *the map $\varphi_i\varphi_j^{-1} : \varphi_j(U_i \cap U_j) \longrightarrow \varphi_i(U_i \cap U_j)$ is a differentiable function for each pair $i, j$.*

*The element $(U_i, \varphi_i)$ is denoted as* chart.
*Moreover, we say that two atlases $\mathcal{A}^1, \mathcal{A}^2$ are compatible if, given $(U_i, \varphi_i) \in \mathcal{A}^1$, for every chart $(U, \varphi) \in \mathcal{A}^2$ the differentiable condition is satisfied by $\varphi_i\varphi^{-1}$.*

In particular the connection of $\mathcal{X}$ implies that the second property of atlas holds with the same Banach space $\mathcal{H}$ (modulo isomorphism).
We can see that the compatibility condition is an equivalence relation, so that we can consider the set of all the equivalence classes of atlases. Given one of these classes we can formulate the following definition:

**Definition B.1.2.** *A structure of $\mathcal{H}$-*manifold *(simply denoted with* manifold*), on a connected topological space $\mathcal{X}$, is a class of equivalent atlases.*

For example, every open subset of a Banach space $\mathcal{H}$ is a manifold.

**Remark B.1.3.** *From now, when we talk about the differential of a function $f : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ defined between two Banach spaces, we intend differential in the sense of Fréchet derivative:*
*The Fréchet derivative of a function $f$ is a bounded linear operator $L : \mathcal{H}_1 \longrightarrow \mathcal{H}_2$ such that:*

$$\lim_{||h||_{\mathcal{H}_1}\to 0} \frac{||f(x+h) - f(x) - Lh||_{\mathcal{H}_2}}{||h||_{\mathcal{H}_1}} = 0,$$

*where $||\cdot||_{\mathcal{H}_i}$ denotes the norm of $\mathcal{H}_i$ for $i = 1, 2$.*
*Then we can say that a function $f : \mathcal{Z} \to \mathcal{X}$ between two $\mathcal{H}$-manifolds $\mathcal{Z}$, $\mathcal{X}$ $f$, is differentiable if $\forall$ $z \in \mathcal{Z}$, when we consider two charts $(V, \psi)$ of $z$ and $(W, \varphi)$ of $f(z)$ such that $f(V) \subset W$, the so called* local representation *of $f$:*
*$\widetilde{f} := \varphi \circ f \circ \psi^{-1} : \psi(V) \to \varphi(W)$ is a differentiable function of Banach spaces.*

One of the crucial points now is how to determine those conditions which guarantee the structure of manifold on a subset of an $\mathcal{H}$-manifold. In order to solve this, we provide the following definition:

**Definition B.1.4.** *Let $\mathcal{X}$ be a $\mathcal{H}$-manifold. A subset $\mathcal{Y} \subset \mathcal{X}$ is a* submanifold *of $\mathcal{X}$ if $\forall\ y \in \mathcal{Y}$ exists a chart $(V_y, \psi_y)$ at $y$ satisfying the following properties:*

- *$V_y = V_1 \times V_2$ with $V_1, V_2 \leq \mathcal{H}$ and $\psi_y(\mathcal{Y} \cap V_y) = V_1 \times v_2$, where $v_2 \in V_2$;*
- *$\psi_y$ induces a bijection: $\psi_{y,1} : \mathcal{Y} \cap V_y \to V_1$.*

*The collection of pairs $(\mathcal{Y} \cap V_y, \psi_{1,y})_{y\in\mathcal{Y}}$ constitutes an atlas for $\mathcal{Y}$.*

Now, we introduce the concept of *immersion* on an $\mathcal{H}$-manifold $\mathcal{X}$.

**Definition B.1.5.** *Let $f : \mathcal{Z} \to \mathcal{X}$ be a differentiable function between two $\mathcal{H}$-manifolds $\mathcal{Z}$, $\mathcal{X}$. We say that $f$ is an* immersion *at $z$ if $U \subset \mathcal{Z}$ open and containing $z$ exists, such that $f_{|U}$ is an isomorphism between $U$ and a submanifold of $\mathcal{X}$. If $f$ is an immersion at each point, it is called global immersion (or simply immersion).*

## B.1.2 Distributions

The definition of immersion is strictly related to the concept of tangent space. In order to give the definition of tangent space at a point $x$ of an $\mathcal{H}$-manifold $\mathcal{X}$, it is necessary to introduce the definition of tangent vector.

**Definition B.1.6.** *Let $\mathcal{X}$ be an $\mathcal{H}$-manifold and let $x$ be a point of $\mathcal{X}$. We consider triples $\bar{v} := (U, \varphi, v)$ where $(U, \varphi)$ is a chart at $x$ and $v$ is an element of the vector space ($\mathcal{H}$) in which $\varphi(U)$ lies.*

*We say that two triples $(U, \varphi, v)$ and $(V, \psi, w)$ are equivalent if $(\psi\varphi^{-1})'_{\varphi(x)}(v) = w$. Clearly, the previous equivalence describes an equivalence relation. We call* tangent vector *an equivalence class of triples, as defined above.*

**Definition B.1.7.** *The tangent space at a point $x$ of $\mathcal{X}$ is the set of all tangents vectors of $\mathcal{X}$ at $x$, denoted by $T_x(\mathcal{X})$.*

Through the concept of tangent space we can generalize the differential of a function, defined between two manifolds:

**Definition B.1.8.** *If $f : \mathcal{X} \to \mathcal{Y}$ is a differentiable function between two $\mathcal{H}$-manifolds, we define the differential of a function: $df(x) : T_x(\mathcal{X}) \to T_{f(x)}(\mathcal{Y})$ as the unique linear function satisfying:*
*$\forall\ (U, \varphi)$ chart at $x \in \mathcal{X}$ and $\forall\ (V, \psi)$ chart at $y \in \mathcal{Y}$ such that $\varphi(U) \subset V$, given a tangent vector $\bar{v} := (U, \varphi, v)$ it holds:*

$$[df(x)](\bar{v}) = \bar{w} := [(V, \psi, w)] \text{ where } [(\psi f \varphi^{-1})'(\varphi(x))](v) = w.$$

**Remark B.1.9.** *It can be proved that $T_x(\mathcal{X})$ is a vector space. Moreover, and choosing a chart at $x$, $(U, \varphi)$, we can provide an isomorphism between $T_x(\mathcal{X})$ and $\mathcal{H}$. Indeed, chosen a point $x \in U \subset \mathcal{X}$, the differential of $\varphi$, computed on $x$ is an isomorphism with its image:*

$$d\varphi_x : T_x(U) \to T_{\varphi(x)}(\varphi U) = \mathcal{H} \tag{B.1}$$

We are now able to provide a proposition, which characterizes the definition of immersion, in terms of tangent spaces:

**Proposition B.1.10.** *Let $\mathcal{X}, \mathcal{Y}$ be manifolds and let $f$ be a differentiable function between those manifolds. Then, the function $f$ is an immersion at $x$ if and only if the map $df(x)$ is injective and splits $T_{f(x)}\mathcal{Y} \equiv F_1 \times \{0\}$ (this means that $T_x\mathcal{X}$ is isomorphic to $F_1 \leq T_{f(x)}\mathcal{Y}$).*

*Proof.* see [16][Chapter $II$]. □

**Example B.1.11.** *Recalling that open subsets of Banach spaces are manifolds, we can consider a differential function $f : D \to \mathcal{H}$, where $D$ is an open subset of $\mathbb{R}^n$ and $\mathcal{H}$ is a Banach space. If $f$ is injective and $df(x)$ is injective too, then $im[f]$ is a submanifold of $\mathcal{H}$.*

We can generalize the concept of tangent space, introducing a new object: the *Tangent Bundle.*

**Definition B.1.12.** *Denoted with $T\mathcal{X}$, the* tangent bundle *is determined by the disjointed union of tangent spaces $T_x(\mathcal{X})$:*

$$T\mathcal{X} = \{(x, \bar{v}) : \ x \in \mathcal{X}, \ \bar{v} \in T_x\mathcal{X}\}.$$

In order to visualize it, we can observe that $T\mathcal{X}$ is the set of $T_x\mathcal{X}$ and each of them is isomorphic to $\mathcal{H}$. Then, using a chart $(U, \varphi)$, we build an isomorphism (this operation is called *trivialization*) between: $T\mathcal{X}|_U = U \times \mathcal{H}$.

As usual, if $\mathcal{X} = V \subset \mathcal{H}$, we can build a global trivialization of $V$: $T\mathcal{X} = V \times \mathcal{H}$.

Generally, a map $\pi$ is paired to $T\mathcal{X}$. This map is the projection of $T\mathcal{X}$ on the first coordinate: $\pi : \underset{(x,\bar{v})}{T\mathcal{X}} \underset{\to}{\to} \underset{x}{\mathcal{X}}$ and each set $\pi^{-1}(x)$ is called the *fiber* of $x$.

Now, we can give the central definition of this subsection:

**Definition B.1.13.** *A* distribution *$S$ is a subset of $T\mathcal{X}$, which satisfies the following property: each fiber of $S$ is a vector subspace of dimension $n$ of $\mathcal{H}$. In particular we can associate to $S$ a map $F : \underset{x}{\mathcal{X}} \underset{\to}{\to} \underset{S_x}{T\mathcal{X}}$, where $S_x \leq T_x\mathcal{X} \equiv \mathcal{H}$ and $dim(S_x) = n$.*

A distribution represents a finite dimensional vector subspace associated with each point. If we want to provide a basis for each subspace, we need the following definition:

**Definition B.1.14.** *Given an $\mathcal{H}$-manifold $\mathcal{X}$, we define a* vector field $\xi$ *as a function:* $\xi : \mathcal{X} \to T\mathcal{X}$*, where* $\xi(x) = \bar{v} \in T_x\mathcal{X}$*, for each* $x \in \mathcal{X}$*. We assume that a vector field satisfies the following property:* $\pi(\xi(x)) = x$*,* $\forall x \in \mathcal{X}$ *if and only if* $\pi \circ \xi = id_{\mathcal{X}}$.

If $\mathcal{X} = U$ is an open subset of $\mathcal{H}$, then: $\xi : U \to TU = U \times \mathcal{H}$. In particular, thanks to the fact that $T_xU \cong \mathcal{H} \ \forall \ x \in U$, we can describe a vector field as $\xi : U \to \mathcal{H}$.

**Definition B.1.15.** *We say that a vector field* $\xi$ lies *on a distribution* $S$*, if* $\xi(x) \in S_x \ \forall x \in X$.

Recalling that $S_x$ is a n-dimensional subspace of $T_x\mathcal{X} \ \forall x_\in \mathcal{X}$, we can find $n$ vectors which form a base of $S_x$, $\forall x \in \mathcal{X}$. In particular, we aim to find a set of vector fields $\xi_1, \xi_2, \dots, \xi_n$, which generate $S$ in this sense:

$$Span\{\xi_1(x), \dots, \xi_n(x)\} = S_x, \quad \forall x \in \mathcal{X},$$

where $Span$ denotes the vector space generated by vectors in argument.

**Definition B.1.16.** *Moreover, we say that a distribution* $S$ *is* smooth*, if* $x \in \mathcal{X}$ $\exists \ U \subset \mathcal{X}$ *open neighborhood of* $x$ *such that* $\exists \ \xi_1, \dots, \xi_n$ *smooth vector fields defined on* $U$ *and* $S_x = Span\{\xi_1(x), \dots, \xi_n(x)\} \ \forall x \in U$.

**Remark B.1.17.** *In the previous definition we introduced smooth vector fields. With the term smooth, we intend that* $\xi$ *is supposed to be a smooth function between Banach spaces (locally). We use this interpretation of the term smooth in all the dissertation.*

### B.1.3 Lie Bracket

The last concept we need in order to develop a self-contained geometric theory is the *Lie bracket*.

We consider an $\mathcal{H}$-manifold $\mathcal{X}$ and $U \subset \mathcal{X}$, open. We consider a smooth function $\varphi : U \to \mathbb{R}$. Observing that $\varphi$ is a function between manifolds, we get that $d\varphi(x) : T_x(U) \to T_{\varphi(x)}(\mathbb{R}) = \mathbb{R}$ is a continuous linear map.

Given a smooth vector field $\xi : \mathcal{X} \to T\mathcal{X}$, we can define the function: $(\xi\varphi) : U \to \mathbb{R}$, defined as: $(\xi\varphi)(x) = d\varphi(x)\Big(\xi(x)\Big)$. We can provide this definition because we have seen that, at least locally, $\xi$ can be treated as a function between the open subset $U$ and $\mathcal{H}$. In particular, if $\xi$ is already defined on $U \subset \mathcal{H}$ open, then: $(\xi\varphi)(x) = \varphi'(x)(\xi(x))$.

Through this new function, we can develop a sort of composition of vector fields, the so called *Lie Bracket*, defined as follows:

**Definition B.1.18.** *Let $\xi, \eta$ be two vector fields on $\mathcal{X}$. Then there exists a unique vector field $[\xi, \eta]$ on $\mathcal{X}$, such that $\forall\ \varphi \in \mathcal{C}^\infty(U, \mathbb{R})$ with $U \subset \mathcal{X}$ open, we get:*

$$[\xi, \eta]\varphi = \eta(\xi\varphi) - \xi(\eta\varphi), \tag{B.2}$$

**Remark B.1.19.** *In particular, if $\xi, \eta$ are defined on $U \subset \mathcal{H}$ open, then:*

$$\Big([\xi, \eta]\varphi\Big)(x) = \varphi'(x)(\xi'(x)\eta(x) - \eta'(x)\xi(x)).$$

*Then, locally:*

$$[\xi, \eta](x) = \xi'(x)\eta(x) - \eta'(x)\xi(x). \tag{B.3}$$

Now, we use the concept of Lie algebra in order to define a particular class of distributions, called involutive distributions. First of all, we introduce the concept of *f-relation*:

**Definition B.1.20.** *We consider a diffeomorphism $f$ between two manifolds $\mathcal{X}, \mathcal{Y}$, defined on the same Banach space $\mathcal{H}$, i.e. $f : \mathcal{X} \tilde{\to} \mathcal{Y}$. We consider a vector field $\xi$, defined on $\mathcal{X}$. Through $f$ we can induce in a unique way a vector field on $\mathcal{Y}$, $\eta$, defined as follows:*

$$\eta(f(x)) = (f_*\xi)(f(x)) = df_x(\xi(x)),$$

*$\eta$ and $\xi$ are in this case called* f-related.

**Remark B.1.21.** *There is an interesting connection between the concept of f-relation and the Lie bracket. In particular, we can observe that, if $f : \mathcal{X} \to \mathcal{Y}$ is a function between two $\mathcal{H}$-manifolds and $\xi_1$ and $\xi_2$ are two vector fields on $\mathcal{X}$, the following equivalence holds:*

$$f_*[\xi_1, \xi_2] = [f_*\xi_2, f_*\xi_2]. \tag{B.4}$$

*We prove this remark locally, supposing then that $\mathcal{X} = U, \mathcal{Y} = V$ are open subset of $\mathcal{H}$. Then:*

$$(f_*[\xi_1, \xi_2])(x) = f'(x)(\xi_1'(x)\xi_2(x) - \xi_2'(x)\xi_1(x)).$$

*Recalling that: $\eta_i(f(x)) = f_*\xi_i(f(x)) = f'(x)\xi_i(x)$, we get:*

$$\begin{aligned}
[\eta_1, \eta_2](f(x)) &= \eta_1'(f(x))\eta_2(f(x)) - \eta_2'(f(x))\eta_1(f(x)) = \\
&= \eta_1'(f(x))f'(x)\xi_2(x) - \eta_2'(f(x))f'(x)\xi_1(x) = \\
&= (\eta_1 \circ f)'(x)(\xi_2(x)) - (\eta_2 \circ f)'(x)(\xi_1(x)) = \\
&= f''(x)\xi_1(x)\xi_2(x) + f'(x)\xi_1'(x)\xi_2(x) - f''(x)\xi_2(x)\xi_1(x) + f'(x)\xi_2'(x)\xi_1(x) = \\
&= f'(x)[\xi_1'(x)\xi_2(x) - \xi_2'(x)\xi_1(x)],
\end{aligned}$$

*where the last equality follows due to the fact that $f''(x)$ is symmetric.*

**Definition B.1.22.** *We say that a distribution $S$ is* involutive *if, given a two vector fields $\xi, \eta$ which lie on $S$, also $[\xi, \eta]$ lies on $S$.*

If we consider a distribution $S$, it is possible to determine a mapping $F$ which describes $S$. If we consider a diffeomorphism $f : \mathcal{X} \to \mathcal{Y}$, we can also compute $f_* F$. This mapping is clearly associated in a unique way with a distribution, denoted with $f_* F$. By Remark B.1.21, we can conclude that $S$ is involutive if and only if $f_* F$ is involutive, for each diffeomorphism $f$.

In the following sections, we will denote with distribution, both $S$ and its associated mapping $F$.

## B.2 Frobenius Theorem

In the previous section we have developed a consistent geometric theory on $\mathcal{H}$-manifolds, where $\mathcal{H}$ is a Banach space. In particular, we have described the concepts of involutive and smooth distribution. In this section we aim to prove a useful characterization of involutive distributions, which allows us to introduce the so called *tangential submanifolds*. The main result is the Frobenius Theorem. This section is based on [4] and [16].

We need the following preliminary definitions:

**Definition B.2.1.** *Given a smooth vector field $\xi$, defined on a $\mathcal{H}$-manifold $\mathcal{X}$, we define the* integral curve *of $\xi$ at $x_0$ as a function $\sigma_{x_0} : J \to \mathcal{X}$, where $J \subset \mathbb{R}$ is an open interval containing $0$, and the following equivalence holds:*

$$\sigma'_{x_0}(t) = \xi(\sigma_{x_0}(t)), \quad \forall t \in J, \text{ such that } \sigma(0) = x_0. \tag{B.5}$$

*In particular, $\sigma_{x_0}(t)$ has the following form:*

$$\sigma_{x_0}(t) = x_0 + \int_0^t \xi(\sigma_{x_0}(s)) ds. \tag{B.6}$$

*In several contexts, we will denote the integral curve of a smooth vector field $\xi$ with $\sigma_x(t) = e^{\xi t} x$.*

The definition of integral curve allows us to introduce the concept of *local flow*:

**Definition B.2.2.** *The local flow of a vector field $\xi$, restricted on an open subset $U \subset X$, is a function $\Theta : J \times U \to \mathcal{X}$, defined as:*

$$\Theta(t, x_0) = \sigma_{x_0}(t), \tag{B.7}$$

*where $\sigma_x$ is the integral curve of $\xi$ at $x$.*

Now, we prove that, given a smooth vector field $\xi$, defined as before, a continuous local flow for $\xi$ exists. We prove this result for vector fields already defined on open subset of the Banach space, since this property is local. It can be proved that if a function $\xi : \mathcal{H} \to \mathcal{H}$ is smooth, then an open subset $U \subset \mathcal{H}$, such that $\xi_{|U}$ is bounded and Lipschitz, can be found ([16] chapter $I$§4, corollary 4.2). Therefore, we can show the following result:

**Proposition B.2.3.** *Let $I$ be an interval of $\mathbb{R}$ containing $0$ and let $U \subset \mathcal{H}$ be open. Let us consider $x_0 \in U$ and $a \in (0,1)$ such that: $\bar{B}_{3a}(x_0) \subset U$, where*

$$\bar{B}_{3a}(x_0) := \{x \in \mathcal{H} : \ ||x - x_0|| \leq 3a\}.$$

*Let us suppose to have a smooth vector field $\xi : U \to \mathcal{H}$, which is bounded by a constant $L \geq 1$ on $U$ and satisfies a Lipschitz condition on $U$ with constant $K \geq 1$. If we consider $b < \frac{a}{LK}$, then*

$$\forall x \in \bar{B}_a(x_0) \ \ \exists\, \Theta : J_b \times B_a(x_0) \to U,$$

*where $J_b := [-b, b] \subset \mathbb{R}$.*

*Proof.* $\forall x \in \bar{B}_a(x_0)$, let us consider the set of functions:

$$\mathcal{M} := \{\alpha : J_b \to \bar{B}_{2a}(x_0) : \quad \alpha \text{ is continuous and } \alpha(0) = x\}.$$

Clearly $\mathcal{M} \neq \emptyset$.
$\mathcal{M}$ is a complete metric space, if we define the usual uniform metric:

$$\delta(\alpha, \beta) := sup_{t \in J_b} |\alpha(t) - \beta(t)|, \quad \forall \alpha, \beta \in \mathcal{M},$$

We define a mapping $S : \mathcal{M} \to \mathcal{M}$ as follows:

$$(S\alpha)(t) := x + \int_0^t \xi(\alpha(s))ds. \tag{B.8}$$

$S\alpha$ is continuous and $S\alpha(0) = x$, moreover:

$$\begin{aligned} ||(S\alpha)(t) - x_0|| =& ||x - x_0 + \int_0^t \xi(\alpha(s))ds|| \\ \leq& ||x - x_0|| + \int_0^t ||\xi(\alpha(s))||ds \\ \leq& a + Lt \leq a + Lb < a + \frac{a}{K} \leq 2a, \end{aligned}$$

so that, $S\alpha \in \mathcal{M}$. Moreover, we can observe that $\forall\ \alpha, \beta \in \mathcal{M}$:

$$\begin{aligned}\delta(S\alpha, S\beta) =& \sup_{t\in J_b} |S\alpha(t) - S\beta(t)| \\ =& \sup_{t\in J_b} \Big| \int_0^t [\xi(\alpha(s)) - \xi(\beta(s))] \Big| ds \\ \leq& \sup_{t\in J_b} \int_0^t |\xi(\alpha(s)) - \xi(\beta(s))| ds \\ \leq& \sup_{t\in J_b} \int_0^t K|\alpha(s) - \beta(s)| ds \\ \leq& \sup_{t\in J_b} \int_0^t K \underbrace{sup_{s\in[0,t]}|\alpha(s) - \beta(s)|}_{\delta(\alpha,\beta)} ds \\ \leq& bK\delta(\alpha, \beta).\end{aligned}$$

Then, choosing $b$ in an appropriate way, we get that $S$ is a shrinking mapping. By contractions lemma there exists $\alpha \in \mathcal{M}$ such that $S\alpha = \alpha$. This fact implies that:

$$\alpha_{x_0}(t) = x_0 + \int_0^t \xi(\alpha(s)) ds.$$

In particular, the mapping $\alpha : J_b \to \bar{B}_a(x_0)$ is continuous in the $t$-variable.

We can also note that the mapping $x_0 \to \alpha_{x_0}(t) = x_0 + \left(\int_0^t \xi(\alpha(s))ds\right)$ is continuous, $\forall\ t \in [-b, b]$. Actually, we will show that is Lipschitz. Let us consider the mapping $S_x : \mathcal{M} \to \mathcal{M}$, defined before, where the subscript $x$ emphasizes that the initial condition depends on on $x$.

Let $x, y$ be point on $\bar{B}_a(x_0)$:

$$||\alpha_x - S_y\alpha_x|| = ||S_x\alpha_x - S_y\alpha_x|| \leq bK||x - y||.$$

Now, denoting $C = bK$ with $0 < C < 1$ (choosing $b$ in a suitable way), we use the following notation: $S_y^n = S_y \overbrace{\circ \cdots \circ}^{n\ times} S_y$. Therefore:

$$\begin{aligned}||\alpha_x - S_y^n\alpha_x|| &\leq ||\alpha_x - S_y\alpha_x|| + ||S_y\alpha_x - S_y^2\alpha_x|| + \cdots + ||S_y^{n-1}\alpha_x - S_y^n\alpha_x|| = \\ &\leq (1 + C + C^2 + \cdots + C^{n-1})|x - y|.\end{aligned}$$

Since $\lim_{n\to+\infty} S_y^n\alpha_x = \alpha_y$, by the continuity of the norm, we obtain:

$$||\alpha_x - \alpha_y|| = \lim_{n\to+\infty} ||\alpha_x - S_y^n\alpha_x|| \leq \lim_{n\to+\infty} \left(\sum_{i=0}^{n-1} C^i\right)|x - y| \leq K_C|x - y|.$$

The integral curve is a Lipschitz function of the initial condition and therefore it is continuous. $\square$

We do not provide the proof of the uniqueness of local flow and we recall [16][Chapter $IV$, §1, Theorem 1.3.].

We are now ready to prove the Frobenius theorem:

**Theorem B.2.4** (Frobenius)**.** *Let $S$ be a smooth distribution and let $F$ be the associated function, defined on a open set $V$ of a Banach space $\mathcal{H}$. Let $x$ be an arbitrary point in $V$. Then, there exists a diffeomorphism $\Phi : U \to \mathcal{H}$ defined on some neighborhood $U \subset V$ of $x$, such that $\Phi_* F$ is constant on $\Phi(U)$ if and only if $F$ is involutive.*

*Proof.* **Part 1**

($\Rightarrow$) We suppose that a function satisfying the property described in the statement exists. We note that, if

$$F(x) = Span\{\xi_1(x), \dots \xi_n(x)\}, \quad x \in U, \quad \text{then} \tag{B.9}$$

$$\varphi_* F(\varphi(x)) = Span\{\varphi_* \xi_1(\varphi(x)), \dots, \varphi_* \xi_n(\varphi(x))\}, \quad \forall\, x \in U. \tag{B.10}$$

In particular $F$ is involutive if and only if $\varphi_* F$ is involutive.

By assumption, it also holds that: $\varphi_* F(\varphi(x)) = \varphi'(x) F(x) = \omega$ for each $x \in X$, where $\omega$ is a vector. This implies that:

$$\varphi_*[\xi, \eta](f(x)) = [\varphi_* \xi, \varphi_* \eta](f(x)) = [\omega, \omega](f(x)) = 0 \;\; \forall\, x \in U.$$

This fact implies that $\varphi_* F$ is involutive. In conclusion, we get that $F$ is involutive

**Part 2**

($\Leftarrow$) To prove this implication we adopt an inductive procedure, on the dimension $n$ of the distribution $F$.

**If n = 1**

Suppose that the distribution is generated by one vector field, denoted by $\xi$. Clearly, this distribution is involutive and, without loss of generality, we can assume that $0 \in V$. Let us define the vector $v = \xi(0)$, and write $\mathcal{H}$ as the direct sum $\mathcal{H} = \langle v \rangle \oplus Y$, where $\langle v \rangle \equiv Span\{v\}$. Note that, since $\langle v \rangle$ is finite dimensional, then the space $Y$ always exists. Let us now consider the function: $\Psi : U \to U$:

$$\Psi(tv + x_0) = \Theta(t, x_0) = x_0 + \int_0^t \xi(\Psi(sv + x_0))ds, \tag{B.11}$$

with $t \in J$, where $J$ is a open interval of $\mathbb{R}$, and $x_0 \in Y$.

By Proposition B.2.3, given a smooth vector field, a continuous local flow $\Theta(t, x_0)$ exists. From the existence of $\Theta(t, x_0)$ the existence and of the (continuous) mapping $\Psi(tv + x_0)$ follows by definition.

By the smoothness of $\xi$, we can also show the smoothness of $\Psi$. We do not provide this proof (see [16], chapter IV §I, theorem 1.14).

Note that:

$$\xi(\Psi(tv + x_0)) = \frac{\partial}{\partial t}(\Psi(tv + x_0)) = \Psi'(tv + x_0)(v); \tag{B.12}$$

$$\Psi(y) = y. \tag{B.13}$$

Moreover, we can exploit the smoothness of $\xi$ and $\Psi$, in order to show the following equivalence:

$$\Psi'(0)(tv + x_0) = t\xi(0) + y = tv + x_0, \tag{B.14}$$

indeed:

$$\begin{aligned}
\Psi(tv + x_0) &= x_0 + \int_0^t \xi(\Psi(sv + x_0))ds = \\
&= x_0 + \int_0^t \xi(\overbrace{\Psi(0)}^{=0} + \Psi'(0)(sv + x_0) + o(sv + x_0))ds = \\
&= x_0 + \int_0^t \Big(\xi(0) + \xi'(0)(\Psi'(0)(sv + x_0) + o(sv + x_0)) + \\
&\quad + o(\Psi'(0)(sv + x_0) + o(sv + x_0))\Big)ds = \\
&= x_0 + \xi(0)t + \xi'(0)\Psi'(0)\Big(\frac{1}{2}t^2 v + x_0 t\Big) + o(tv + x_0) = \\
&= x_0 + \xi(0)t + o(tv + x_0).
\end{aligned}$$

Then, developing the Taylor expansion of the function $\Psi$:

$$\Psi(tv + x_0) = \overbrace{\Psi(0)}^{=0} + \Psi'(0)(tv + x_0) + o(tv + x_0).$$

Hence, substituting in the previous equation, we obtain:

$$\Psi'(0)(tv + x_0) = x_0 + \xi(0)t = vt + x_0 + o(vt + x_0).$$

This means that, near 0, $\Psi'(0)$ is invertible (it is the identity). By the smoothness of $\Psi$, we get that $\Psi'(0)$ is a local diffeomorphism. By the theorem of inverse function, we can provide a local inverse $\Phi = \Psi^{-1}$. Moreover, we can restrict $U$ until we get: $\Phi : U \to U$.

Let $x = \Psi(tv + x_0)$ if and only if $\Phi(x) = tv + x_0$. Recalling the concept of $\Phi$-relation, we can write:

$$\Big(\Phi_* \xi(\overbrace{tv + x_0}^{\Phi(x)})\Big) = \Phi'(x)(\xi(x)),$$

On the other hand, we have seen that: $\Psi'(tv+x_0)v = \xi(\Psi(tv+x_0)) = \xi(x)$, hence:

$$\begin{aligned}(\Phi_*\xi)(tv+x_0) =& \Phi'(x)(\xi(x)) = \\ =& \Phi'(\Psi(tv+x_0))\Psi'(tv+x_0)v = \\ =& \underbrace{(\Phi\circ\Psi)'}_{\text{id on U}}(tv+x_0)v = v.\end{aligned}$$

Then, $\Phi_*\xi$ is a constant vector field.

**If $\mathbf{n > 1}$**

For the induction step, we consider an $n$-dimensional distribution, and suppose that the theorem holds for every $m$-dimensional distribution, with $m < n$.

As done before, we can assume the the origin belongs to $V$. Therefore, we suppose that $\xi_1, \dots, \xi_n$ are vector fields generating $S$ on $V$. We denote $v_i = f_i(0)$, $i = 1, \dots n$, and we decompose $\mathcal{H}$ as follows:

$$\mathcal{H} = \langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle \oplus Z,$$

supposing that $\langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle \cap Z = \langle 0 \rangle$. Similarly as before, we can say that such a space $Z$ always exists, since $\langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle$ is finite-dimensional.

We now introduce the following subspace:

$$\mathcal{H}_1 = \langle v_2 \rangle \oplus \cdots \oplus \langle v_n \rangle \oplus Z.$$

For the inductive step, we can suppose that (modulo diffeomorphism) $\xi_1 = v_1$. Hence, we can apply Gauss elimination, in order to rewrite our generating base in the following way:

$$\xi_i = v_i + g_i, \quad j = 2, \dots, n,$$

where $g_j \in \mathcal{C}^\infty(V, Z)$.

By assumption, $F$ is assumed to be involutive, then there exists a family of scalar fields $a_{jk} \in \mathcal{C}^\infty(V, \mathbb{R})$ such that: $[\xi_1, \xi_j] = \sum_{k=1}^n a_{jk}\xi_k$, for $j = 2, \dots, n$. Hence:

$$\begin{aligned}[\xi_1, \xi_j] &= \xi_1'\xi_j - \xi_j'\xi_1 = 0 - g_j'v_1 \\ &= a_{j1}v_1 + a_{j2}(v_2 + g_2)\cdots + a_{jn}(v_n + g_n).\end{aligned}$$

Note that $g_j'v_1 \in Z$. As a consequence, it follows that: $j = 2, \dots, n$, $k = 1, \dots, n$: $a_{jk} = 0$ and therefore $g_j'v_1 = 0$ for $j = 2, \dots, n$.

Then for $j = 2, \dots, n$, we have:

$$\frac{\partial}{\partial t_1} g_j(t_1v_1 + h) = g_j'(t_1v_1 + h)v_1 = 0, \quad \forall t_1 \in \mathbb{R}, \quad \forall h \in \mathcal{H}_1.$$

Thus $g_j$ does not depend on $t_1$, then:

$$g_j(t_1v_1 + h) = g_j(h), \ \forall h \in \mathcal{H}_1, \ \forall t_1 \in \mathbb{R}.$$

If we consider the restriction of $\xi_1, \dots, \xi_n$ to $\mathcal{H}_1$, they generate an $(n-1)$-dimensional distribution $F_{\mathcal{H}_1}$, indeed: $\xi_j(t_1 v_1 + h) = v_j + g_j(t_1 v_1 + h) = v_j + g_j(h)$ is well defined on $\mathcal{H}_1$.
Clearly $F_{\mathcal{H}_1}$ is smooth and involutive.
Therefore, from induction hypothesis there exists $U \subset \mathcal{H}$ and a diffeomorphism $\Phi_{\mathcal{H}_1} : U \cap \mathcal{H}_1 \to U \cap \mathcal{H}_1$ such that $\Phi_{Y*}F$ is constant near $\Phi(0)$.
Finally, we define the map $\Phi : U \to U$, in the following way:

$$\Phi(t_1 v_1 + h) = t_1 v_1 + \Phi_Y(h),$$

we get a diffeomorphism around $0 \in \mathcal{H}$ such that $\Phi_* F$ is constant near $\Phi(0)$. □

## B.3 Tangential manifolds for involutive distributions $S$

In this section, we introduce the concept of tangential manifold for a given distribution $F$. Such a manifold is defined as follows:

**Definition B.3.1.** *Let $F$ be a smooth distribution and let $x_0$ be a fixed point in $\mathcal{X}$, an $\mathcal{H}$-manifold.*
*A submanifold $\mathcal{G} \subset \mathcal{X}$, with $x_0 \in \mathcal{G}$, is called* tangential manifold *through $x_0$ for $F$, if $F(x) \leq T_x\mathcal{G}$, $\forall x \in U$, where $U$ is an open neighborhood of $x_0 \in \mathcal{G}$.*

We use the Frobenius Theorem in order to prove the following result:

**Theorem B.3.2.** *Let $F$ be an $n$-dimensional distribution and let $x_0$ be a fixed point on an $\mathcal{H}$-manifold $\mathcal{X}$. Then, there exists an $n$-dimensional tangential manifold through $x$ for each $x$ in a neighborhood of $x_0$, if and only if $F$ is involutive.*

*Proof.* **Part 1**
($\Leftarrow$) If $F$ is involutive, using the Frobenius theorem, we get $n$-smooth vector fields $\xi_1, \dots, \xi_n$ and a local diffeomorphism $\Phi : U \to U$, defined on $U$ open neighborhood of $x_0$ on $\mathcal{X}$, such that $\Phi_*\xi_1, \dots, \Phi_*\xi_n$ are constant. Denoting $\Phi_*\xi_i = w_i$, we see that for each $x \in U$, the hyperplane:

$$\pi_x = \Phi(x) + \langle w_1, \dots, w_n \rangle$$

is a tangential manifold for the distribution $\Phi_* F$, passing through $\Phi(x)$. Pulling back this plane with $\Phi$, we get $\Phi^{-1}(\pi_x)$, which is a tangential manifold for $F$, passing through $x$, denoted with $\mathcal{G}_{|U}$.

**Part 2**
($\Rightarrow$) If there exists an $n$-dimensional tangential manifold $\mathcal{G}$ through $x$, for each

$x \in U$, where $U \subset \mathcal{X}$ neighborhood of $x_0$, then: $F(x) = T_{\mathcal{G}}x$, for each $x \in U$ (we can restrict $\mathcal{G}$ in order to have the equivalence).
If $\xi_1, \xi_2$ are vector fields spanning $F$, then $\xi_1(x), \xi_2(x) \in T_{\mathcal{G}}x$, for each $x \in U$, then $\xi_1, \xi_2$ are vector fields on the manifold $U \cap \mathcal{G}$. But we have seen that also $[\xi_1, \xi_2]$ is a vector field on $U \cap \mathcal{G}$ and this means that: $[\xi_1, \xi_2](x) \in T_{\mathcal{G}}x = F(x)$. In conclusion, we obtain that $F$ is involutive. □

Recalling that, given a smooth vector field $\xi$ on an $\mathcal{H}$-manifold $\mathcal{X}$, we denote the integral curve passing through a point $x$ with $\sigma_x(t) = e^{\xi t}x$, with the following result we can describe how to build the tangential manifold.

**Proposition B.3.3.** *Consider an $n$-dimensional involutive distribution spanned by $\xi_1, \dots, \xi_n$ and a point $x_0 \in X$. We have seen that a tangential manifold through $x_0$ exists and let us denote with $\mathcal{G}$.*
*Defining a mapping $G : \mathbb{R}^n \to \mathcal{X}$ by:*

$$G(z^1, \dots, z^n) = e^{\xi_n z^n} \cdots e^{\xi_1 z^1} x_0,$$

*then $G$ is a local parametrization of $\mathcal{G}$ in the sense that: there exists $U \subset \mathbb{R}^n$ open, containing $0$ and $V \subset \mathcal{G}$ open, containing $x_0$ such that $V = G(U)$.*
*Furthermore, the inverse of $G_{|V}$ is a local coordinate system for $\mathcal{G}$ at $x_0$.*

*Proof.* From the definition of tangential manifold we have that $G(z) \in \mathcal{G}$, for $z \in U \subset \mathbb{R}^n$, open subset containing 0 (we can suppose without loss of generality, that $0 \in U$).

Moreover, if we denote the local flow of the vector field $\xi_i$ with $\Theta^{\xi_i}$ we have that:

$$G(z^1, \dots, z^n) = \Theta^{\xi_n}(z^n, \Theta^{\xi_{n-1}}(z^{n-1}, \dots (z^2, \Theta^{\xi_1}(z^1, x_0)) \dots)),$$

and the differential of $G$ at the arbitrary point $(z^1, \dots, z^n)$ is given by:

$$dG_{(z^1,\dots,z^n)} = \Big(\frac{\partial}{\partial z^1} G, \dots, \frac{\partial}{\partial z^1} G\Big)\Big|_{(z^1,\dots,z^n)},$$

In particular, for each $h \in \mathbb{R}^n$ it holds that: $dG_{(z^1,\dots,z^n)}(h) = \sum_{j=1}^n \frac{\partial}{\partial z^j} G_{(z^1,\dots,z^n)} h^j$.
Recalling Example B.1.11, we aim to prove that $dG$ is injective around 0. In theorem B.3.2, we have seen that $\mathcal{G}|_V = \Phi^{-1}(\pi_x)$, where $V$ is an open neighborhood of $x_0$ and $\Phi : V \to \Phi(V) \subset \pi_x$. For what we have told at the beginning of the proof, $G(z) \in \mathcal{G}$, for each $z \in U$, then we get $\Phi(G(z)) \in \pi_x$, where $\pi_x$ is the plane introduced in Theorem B.3.2.
Then, by the Remark B.1.21: $\Phi_*[\xi_i, \xi_j] = [\Phi_*\xi_i, \Phi_*\xi_j] = 0$, because the transformed vector fields are constants.

Therefore, if we consider the submanifold $\Phi(G(U))$, it is generated by $\Phi_*\xi_i := \eta_i \ \ \forall\, i = 1, \ldots, n$. In particular:

$$\Phi(G(z)) = e^{\eta_n z^n} \cdots e^{\eta_1 z^1} x_0.$$

It can be proved that if $[\eta_j, \eta_i] = 0$, then the local flows of $\eta_i, \eta_j$ commute (see [16], chapter $V\S I$, theorem 1.5). This fact allows us to permute the integral curves which define $\Phi(G(U))$, in order to compute:

$$\begin{aligned}\frac{\partial}{\partial z^i}\Phi(G(z))\Big|_{z=0} =& \frac{\partial}{\partial z^i} e^{\eta_n z^n} \cdots e^{\eta_1 z^1}\Phi(x_0)\Big|_{z=0} = \eta_i(\Phi(x_0))\\ =&(\Phi_*\xi_i)(\Phi(x_0)) = \Phi'(x_0)(\xi_i(x_0)),\end{aligned}$$

On the other hand, by definition: $\frac{\partial}{\partial z^i}\Phi(G(z))|_{z=0} = \Phi'(G(0))\frac{\partial}{\partial z^i}G'(z)|_{z=0}$.
In conclusion, due to the previous equivalences and remembering that $G(0) = x_0$ and that $\Phi'(x_0)$ is invertible:

$$\frac{\partial}{\partial z^i}(G(z))\Big|_{z=0} = \xi_i(x_0).$$

Since $dG_{(z=0)}(h) = \sum_{j=1}^n \frac{\partial}{\partial z^j} G_{(z^1,\ldots,z^n)}|_{(z=0)}h^j = \sum_{j=1}^n \xi(x_0)h^j$ and $\xi_1, \ldots, \xi_n$ form a base, then they are independent, we obtain that $dG_{z=0}$ is injective.
By the theorem of inverse function, we can find an open neighborhood of $0 \in \mathbb{R}^n$, $U$ and an open neighborhood of $x_0 = G(0)$, open subset $V$ , in which $G$ is invertible. This fact means that we can provide a local coordinate in $\mathbb{R}^n$ for the tangential submanifold $\mathcal{G}|_U$, indeed:

$$G^{-1}\big|_V : \mathcal{G}\big|_V \xrightarrow{\sim} U \subset \mathbb{R}^n.$$

□

We end this chapter with a concept which will be crucial in the search of finite dimensional realizations for a forward rate model described through $HJM$-approach.

**Definition B.3.4.** *Let $F$ be a smooth distribution on $U \subset \mathcal{H}$ open. The* Lie algebra *generated by $F$, denoted by $\{F\}_{LA}$, is defined as the minimal (under inclusion) involutive distribution containing $F$.*

We prove now a result, which can be useful when we have to determine the Lie Algebra generated by a set of smooth vector fields. It is based on Lemma 4.1 of [5].

**Lemma B.3.5.** *Let us consider $n$ smooth vector fields $\xi_1, \ldots, \xi_n$ defined on an $\mathcal{X}$-manifold. Then the following operations does not modify the Lie algebra $\mathcal{L} = \{\xi_1, \ldots, \xi_n\}_{LA}$:*

1. *The vector field* $\xi_i$ *can be replaced by* $\alpha\xi_i$ *where* $\alpha$ *is a smooth non-zero scalar field defined on* $\mathcal{X}$*;*

2. *The vector field* $\xi_i$ *can be replaced by:*

$$\xi_i + \sum_{j \neq i} \alpha_j \xi_j,$$

*where* $\alpha_j$ *si a smooth scalar vector field, for each* $j$*.*

*Proof.* If we substitute any vector $v$ with $\lambda v$ with $\lambda \in \mathbb{R} \setminus \{0\}$, the vector space generated by $v$ does not change. This fact proves point *1.*.

Point *2.* follows directly from the bilinearity of Lie Bracket, point *1.* and the fact that $[\xi, \xi] = 0$. □

Given a forward rate model, which is described by a distribution $F$, we exploit the result of the Frobenius theorem in order to provide a set of vector fields which span $\{F\}_{LA}$. Doing that, we obtain, by Proposition B.3.3, a local set of coordinates for a tangential submanifold, associated with the minimal extension of $F$, which is involutive. To do that we base on [5],[21].

# Conclusions

The aim of this thesis was to generalize the geometric approach developed by Björk in order to face the problems of consistency and existence of finite-dimensional realizations in a post-crisis interest-rate market. As regards the problem of consistency, we understood that it was no longer possible to consider an identical approach to the one developed in [5] for the pre-crisis context. Indeed, although the theoretical conditions can be easily generalized from the pre-crisis environment, the presence of the spreads between interest rates associated with different tenors had led to a more complex structure to manage in concrete examples. Therefore, we first tried to understand if it was possible to circumvent the presence of spreads by adding them to the finite-dimensional process $Z_t$ determined by the consistency conditions. As we described in Remark 2.2.9, this result can not be achieved without requesting additional hypotheses. Hence, we concluded that it was necessary to provide conditions on a parameterized family $\mathcal{G}$ for the components associated with the spreads too, in order to guarantee the consistency between a given model $\mathcal{M}$ and $\mathcal{G}$. As a consequence of this fact, we studied concrete examples of forward rate models $\mathcal{M}$, as the Ho-Lee model and, especially, the Hull-White model, in comparison to widely used parameterized families, the Svensson family and the Nelson-Siegel family. In particular, we considered the generalizations introduced in [2] of the above mentioned families, in order to guarantee the consistency with each forward rate components of the analysed models. The main problem was related to the presence of the spreads. In the analysed examples, we exploited the independence between the coordinates of the volatility term $\hat{\sigma}$ from the entire structure of the solution of system (1.33) in order to construct a procedure which allows to satisfy the consistency conditions with a very simple functions for the components associated with the spreads, by adding a suitable number of real parameters. Vice versa, in some cases we were able to determine the relations on the coordinates of the vector $\hat{\sigma}$, which guaranteed the consistency between the model $\mathcal{M}$ determined by $\hat{\sigma}$ and a suitable parameterized family $\mathcal{G}$, introduced without adding other parameters. We proved these results in the case of models driven by a 1-dimensional Brownian motion and, for the Hull-White model and the forward parameterized family determined by the function (2.74), we provided those results in the general case of a $d$-dimensional Brownian motion.

For the problem of the existence of finite-dimensional realizations (FDR), we exploited the analogy between the interest rate market in the post-crisis framework modeled by (1.33) and a system of SDEs which described the multi-currency interest rate market, for which the problem of FDR was analysed in [21] by Slinko. In this article, the problem of the existence of FDR was faced for a model $\mathcal{M}$ given by 2 different currencies. We generalized those results to the case of a general tenor structure composed by $m$ tenors and for models driven by a $d$-dimensional Brownian motion. In particular, we proved that in the case of a constant volatility term $\hat{\sigma}$ the existence of FDR is equivalent to request that the coordinates of $\hat{\sigma}$ are given by quasi-exponential (QE) functions. Moreover, for a model $\mathcal{M}$ given by a constant direction volatility term $\hat{\sigma}$ as (3.31), we proved that if $\hat{\sigma}$ is determined by QE functions, then finite-dimensional realizations exist. Finally, we constructed a simplified constant direction volatility model, for which, under suitable technical conditions on the volatility term $\hat{\sigma}$, requesting that $\lambda_i^j$ is QE for each $j = 0, \dots, m$ and $i = 1, \dots, d$ is equivalent to the existence of FDR.

In conclusion, we analyzed an open problem concerning interest rate market models, adopting a geometric approach described by some strong results of functional analysis and differential geometry, in a stochastic framework.

# Bibliography


[1] Ametrano, F. M. & Bianchetti, M. (2013): *Everything You Always Wanted to Know About Multiple Interest Rate Curve Bootstrapping but Were Afraid to Ask*, SSRN.

[2] Björk, T. & Christensen, B. J. (1999): *Interest Rate Dynamics and Consistent Forward Rate Curves*, Mathematical Finance, Vol. 9, No. 4, 323-348.

[3] Björk, T. (2009-3rd edition): *Arbitrage Theory in Continuous Time*, Oxford University Press.

[4] Björk, T. & Svensson, L. (2001): *On the Existence of Finite-Dimensional Realizations for Nonlinear Forward Rate Models*, Mathematical Finance, Vol. 11, No. 2, 205-243.

[5] Björk T. (2004): *On the geometry of interest rate models*, In: Carmona R.A. *et al.* (Eds), Paris-Princeton Lectures on Mathematical Finance 2003, Lecture Notes in Mathematics, Springer, Berlin, Heidelberg, Vol. 1847, pp 133-215.

[6] Brezis, H. (2011): *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Spinger.

[7] Coddington, E.A. & Carlson, R. (1997): *Linear Ordinary Differential Equations*, SIAM.

[8] Cuchiero, C., Fontana C. & Gnoatto, A. (2016): *A general HJM framework for multiple yield curve modelling.* Finance and Stochastics, Vol. 20, No. 2, 267-320.

[9] Da Prato, G. & Zabczyk, J. (1992-1st edition): *Stochastic Equations in Infinite Dimensions* Encyclopedia of Mathematics and its Applications, Cambridge University Press.

[10] Gawarecki, L. & Mandrekar, V. (2011): *Stochastic Differential Equations in Infinite Dimensions with Applications to Stochastic Partial Differential Equations*, Springer.

[11] Grbac, Z. & Runggaldier, W.J. (2016): *Interest Rate Modeling: Post-Crisis Challenges and Approaches*, Springer.

[12] Fontana, C., Grbac, Z., Gumbel, S. & Schmidt, T. (2018): *Term-structure modeling for multiple curves with stochastic discontinuities*, preprint.

[13] Filipović D. (1999): *A note on the Nelson-Siegel family*, Mathematical Finance, Vol. 9, No. 4, 349-359.

[14] Filipović D. (2009): *Term-Structure Models, A Graduate Course*, Springer.

[15] Ho, T.S.Y. & Lee, S.B. (1986): *Term Structure Movements and Pricing Interest Rate Contingent Claims*, The Journal of Finance, Vol. 41, No. 5, 1011-1029.

[16] Lang, S. (1999): *Fundamentals of Differential Geometry*, Springer.

[17] Nelson, C.R. & Siegel, A.F. (1987):*Parsimonious modeling of yield curves*, The Journal of Business, Vol. 60, No. 4. pp 473-489.

[18] Rudin, W. (1991): *Functional Analysis*, McGraw-Hill.

[19] Sharef, E. & Filipović, D. (2003): *Conditions for consistent exponential-polynomial forward rate processes with multiple non trivial factors*, International Journal of Theoretical and Applied Finance, Vol. 7, No. 6, 685-700.

[20] Slinko, I. (2006): *Essays in Option Pricing and Interest Rate Models*, Stockholm School Of Economics.

[21] Slinko, I. (2010): *On finite dimensional realizations of two country interest rate models*, Mathematical Finance, Vol. 20, No. 1, 117-143.