On the Existence of Finite Dimensional Realizations for Nonlinear Forward Rate Models *

Tomas Björk †
Department of Finance
Stockholm School of Economics
Box 6501
S-113 83 Stockholm, SWEDEN
e-mail: fintb@hhs.se

Lars Svensson
Department of Mathematics
Royal Insitute of Technology
S-100 44 Stockholm, SWEDEN
e-mail: larss@math.kth.se

October 12, 1999

Abstract

We consider interest rate models of Heath-Jarrow-Morton type, where the forward rates are driven by a multidimensional Wiener process, and where he volatility is allowed to be an arbitrary smooth functional of the present forward rate curve. Using ideas from differential geometry as well as from systems and control theory, we investigate when the forward rate process can be realized by a finite dimensional Markovian state space model, and we give general necessary and sufficient conditions, in terms of the volatility structure, for the existence of a finite dimensional realization. A number of concrete applications are given, and most previously known realization results for time homogenous Wiener driven models are included and extended. As a special case we give a general and easily applicable necessary and sufficient condition for when the induced short rate is a Markov process. In particular we show that the only forward rate models, with short rate dependent volatility structures, which generically give rise to a Markovian short rate are the affine ones. These models are thus the only generic short rate models from a forward rate point of view.

JEL: E43, G13.
Key Words: Forward rate curves, interest rate models, factor models, state space models, Markovian realizations.

*We are very grateful to J.Zabzcyk for valuable discussions and to J.Lund for important references.
†Support from ITM is gratefully acknowledged.

Contents

1 Introduction **3**

2 Mathematical Background and the Frobenius Theorem **4**

3 Realizations and Invariant Manifolds **9**

3.1 Homogenous Systems 9

3.2 Interpreting the State Variables 13

3.3 Time Varying Systems 14

4 Interest Rate Models **16**

4.1 Basics 16

4.2 The Space 17

4.3 The General Model 19

4.4 Main Results 21

5 Deterministic Volatility **23**

5.1 Finite Dimensional Realizations 23

5.2 Invariant Manifolds 25

6 Deterministic Direction Volatility **27**

6.1 A Scalar Driving Wiener Process 28

6.2 The Time Varying Case 29

6.3 Several Driving Wiener Processes 30

6.4 The Time Varying Case 32

7 Short Rate Realizations **33**

7.1 General Theory 33

7.2 Deterministic Volatility 37

7.3 Constant Direction Volatility 39

7.4 "All Generic Short Rate Models are Affine" 42

1 Introduction

The purpose of the present paper is to investigate when and how a given, a priori infinite dimensional, forward rate model can be realized by a finite dimensional state space model in the form of a nonlinear diffusion. In particular we are interested in minimal realizations, i.e. realizations having a minimal dimension.

For the case of a deterministic volatility structure (which gives rise to linear forward rate dynamics) the realization problem has been completely solved in [3]. The nonlinear realization problem has so far only been studied for special cases and the typical results are in terms of sufficient conditions for the existence of a finite dimensional realization (henceforth denoted FDR).

Most papers consider some version of a multiplicative volatility structure. Such a volatility structure, which we below refer to as a "deterministic direction volatility", has been studied in [1], [23], and [17]. The typical assumption is that volatilitites only depends on the short rate. In the recent preprint [9], which is a massive extension of [1], there are several new results on Markovian subsystems for an interesting class of time varying forward rate models where the volatility structure is allowed to depend upon a finite number of benchmark forward rates.

A classic paper on Markovian short rates is [8], where a deterministic volatility is considered. See also [13] for a similar case with a driving Levy process. The deepest study of short rate realizations so far is [19].

The present paper, in contrast to previous work in the field, is a study of the general nonlinear realization problem. More precisely:

- The forward rate volatilities are allowed to general smooth functionals of the present forward rate curve.
- In this general setting we obtain necessary and sufficient conditions for the existence of an FDR.

We are thus able to give a complete solution to the general FDR problem and, apart from [9] where the setup is lightly different, all earlier results are included and extended.

The methodology is roughly as follows: Our problem turns out to be closely connected to certain geometric properties of the forward rate process. In a previous paper [2], it is investigated when the forward rate curve for a given (nonlinear) forward rate model actually evolves on a given finite dimensional submanifold in the forward curve space. Such a submanifold is thus invariant under the action of the forward rate process, and in [2] it is shown that there exists a FDR if and only if there exists an invariant finite dimensional submanifold passing through each point in the space of forward rate curves. In the present work we use results from [2], together with ideas from differential geometry as well as from

systems and control theory to give general necessary and sufficient conditions, in terms of the volatility structure, for the existence of an FDR.

The structure of the paper is as follows. In Section 2 we give the necessary mathematical background. In particular we give a self-contained proof of an infinite dimensional version of the Frobenius Theorem on involutive distributions, which is the key result for our entire theory. In Section 3 we solve the realization problem for a general SDE evolving in Hilbert space, and we also provide an explicit characterization of the minimal invariant manifold generated by such an SDE.

The applications to finance starts in Section 4, where we construct a suitable Hilbert space of forward rate curves to work in, and apply our general realization results to the forward rate equation. The main result of the paper is Theorem 4.1, which says that the forward rate model has a finite dimensional realization if and only if the Lie algebra generated by the drift and the volatility of the forward rate equation is finite dimensional.

The rest of the paper is devoted to applications of the main theorem, and in Section 5 we study the case of a deterministic volatility structure. The more interesting case of a multiplicative volatility is studied, in some detail, in Section 6 where we include and extend almost all previously known results in the field. We end the paper with a study of short rate Markovian realizations in Section 7. The most important result here is Theorem 7.1 which provides a very simple necessary and sufficient condition for when the short rate is Markovian. As an application of this result we rederive, in Theorem 7.2, a remarkable result from [19] which shows that the only forward rate models, with short rate dependent volatility structures, which generically give rise to a Markovian short rate are the affine ones. These models are thus the only generic short rate models from a forward rate point of view.

The geometric ideas presented below are intimately connected to controllability problems in systems theory, where they have been used extensively (see [18]). They have also been used in filtering theory, where the problem is to find a finite dimensional realization of the unnormalized conditional density process, the evolution of which is given by the Zakai equation. A very good account of these ideas can be found in [7]. For general background on interest rate theory see [4].

2 Mathematical Background and the Frobenius Theorem

In this section we give the necessary (and sufficient) background on infinite dimensional differential geometry. In the context of interest rate theory below, we will only be working in a Hilbert space setting, but the results are just as easy to prove in Banach space.

Let us thus consider a real Banach space X. By an n-dimensional **distribution**, we mean a mapping F, which to each x in an open subset V of X associates an n-dimensional subspace $F(x) \subseteq X$. A mapping (vector field) $f : U \to X$, where U is an open subset of X, is said to **lie in** F (on U), if $U \subseteq V$ and $f(x) \in F(x)$ for every $x \in U$. A collection $f_1, \ldots f_n$ of vector fields lying in F on U **generate** (or span) F on U if $span\{f_1(x), \ldots f_n(x)\} = F(x)$ for every x in U, where *span* denotes the linear hull over the real field. The distribution is **smooth** if, to every x in V, there exist an open set U such that $x \in U \bigcap V$, and smooth vector fields $f_1, \ldots f_n$ spanning F on U. If F and G are distributions and $G(x) \subseteq F(x)$ for all x we say that F **contains** G, and we write $G \subseteq F$. The **dimension** of a distribution F is defined pointwise as $dimF(x)$.

Remark 2.1 The prefix **smooth** above is interpreted as C^∞. Note, however, that in many situations below we only need C^k for some k which depends upon the particular context.

Let f and g be smooth vector fields on U. Their **Lie bracket** is the vector field

$$[f, g](x) = f'(x)g(x) - g'(x)f(x),$$

where $f'(x)$ denotes the Frechet derivative of f at x, and similarly for g. For clarity we will sometimes write $f'(x)[g(x)]$ instead of $f'(x)g(x)$ A distribution F is called **involutive** if, for every smooth f and g lying in F on U, their Lie bracket also lies in F, i.e.

$$[f, g](x) \in F(x), \quad \forall x \in U.$$

Let the distribution F be as above, and let $\varphi : V \to W$ be a diffeomorphism between the open subsets V and W of X. Then we can define a new distribution $\varphi_\star F$ on W by

$$(\varphi_\star F)(\varphi(x)) = \varphi'(x)F(x).$$

Similarly, for any smooth vector field $f \in C^\infty(U, X)$, we define the field $\varphi_\star f$ by $\varphi_\star f = \left(\varphi' \circ \varphi^{-1}\right)\left(f \circ \varphi^{-1}\right)$. By a straightforward calculation one verifies easily that

$$\varphi_\star [f, g] = [\varphi_\star f, \varphi_\star g].$$

From this we conclude that if F is generated by $f_1, \ldots, f_n$, then $\varphi_\star F$ is generated by $\varphi_\star f_1, \ldots, \varphi_\star f_n$, and we see that F is involutive if and only if $\varphi_\star F$ is involutive.

The main result of this section is the following infinite dimensional extension of the standard Frobenius Theorem. For our purposes we need a particular formulation of the theorem and, since this was not immediately available in the literature (see e.g. [20]), we also provide a self contained proof.

Theorem 2.1 (Frobenius) *Let F be a smooth distribution on the open set V in the Banach space X. Let furthermore x be an arbitrary point in V. Then there exists a diffeomorphism $\varphi : U \to X$ on some neighborhood $U \subseteq V$ of x, such that $\varphi_\star F$ is constant on $\varphi(U)$ if and only if F is involutive.*

Proof. The only if part is obvious from the remarks above. Let us thus assume that F is involutive. We proceed by induction on n. For $n = 1$ the result follows basically from the existence theorem for ordinary differential equations, and a sketch of this part of the proof is as follows.

Without loss of generality (WLOG) we may assume that the origin is in V, i.e. that $0 \in V$. Define the vector e by $e = f(0)$, and write X as the direct sum $X = Re + Y$. Note that, since Re is finite dimensional, it is complemented, so the space Y always exists. Now consider the equation

$$\psi(te + y) = y + \int_0^t f(\psi(se + y))\, ds, \quad t \in R,\ y \in Y.$$

Using a standard contraction argument, it can be shown that, for some neighborhood U of $0 \in X$, there exists a unique continuous solution ψ of this equation with $\psi : U \to U$. To prove smoothness (i.e. C^∞) of Ψ is harder, but it can be done, see e.g. [20]. From the equation above we have $\psi'(te+y)e = f(\psi(te + y))$. We also note that $\psi(y) = y$ and that $\psi'(0)(te + y) = tf(0) + y = te + y$ where the last equality follows from the formula

$$\begin{aligned} \psi(te + y) &= y + \int_0^t f(\psi(0) + \mathcal{O}(se + y))\, ds \\ &= y + tf(0) + o(te + y) = y + te + o(te + y). \end{aligned}$$

Thus $\psi'(0)$ is the identity mapping, and the inverse function theorem provides us with a local inverse $\varphi = \psi^{-1}$. Now let $x = \psi(te+y)$, i.e. $\varphi(x) = te+y$. Then we have $(\varphi_\star f)(te + y) = \varphi'(x)f(x)$. But $\psi'(te + y)e = f(\psi(te + y)) = f(x)$, so $(\varphi_\star f)(te + y) = \varphi'(\psi(te + y))\psi'(te + y)e = (\varphi \circ \psi)'(te + y)e = e$. Thus $\varphi_\star f = e$ is a constant vector field.

For the induction step, consider an n with $n > 1$ and suppose that the theorem holds for every m-dimensional distribution whenever $m < n$. We may again WLOG assume that the origin is in V, i.e. that $0 \in V$. Assume furthermore that $f_1, \ldots f_n$ generate F in V. Put $e_i = f_i(0);\ i = 1, \ldots, n$ and choose a closed subspace Z of X such that

$$X = Re_1 + \ldots + Re_n + Z,$$

and

$$(Re_1 + \ldots + Re_n) \cap Z = 0.$$

Such a subspace always exists, since $Re_1 + \ldots + Re_n$ is finite dimensional, and every finite dimensional subspace in a Banach space can be complemented. We now define the subspace Y by

$$Y = Re_2 + \ldots + Re_n + Z.$$

Since the theorem holds for $n = 1$ we may (modulo a diffeomorphism) assume that $f_1 = e_1$. By applying Gauss elimination we also see that we may WLOG assume that

$$f_j = e_j + g_j, \quad j = 2, \ldots, n,$$

where $g_j \in C^\infty(V, Z)$.

Since F was assumed to be involutive, there exist scalar fields $\alpha_{jk} \in C^\infty(V, R)$ such that $[f_1, f_j] = \sum_{k=1}^n \alpha_{jk} f_k$ for $j = 2, \ldots, n$. Hence

$$\begin{aligned}[f_1, f_j] &= f_1' f_j - f_j' f_1 = 0 - g_j' e_1 \\ &= \alpha_{j1} e_1 + \ldots + \alpha_{jn} e_n + \alpha_{j2} g_2 + \ldots + \alpha_{jn} g_n.\end{aligned}$$

Now, $g_j' : V \to L(X, Z)$ (where $L(X, Z)$ denotes the space of bounded linear maps), so $g_j' e_1 \in Z$, from which it follows that for $j = 2, \ldots, n$ we have $\alpha_{j1} = \ldots = \alpha_{jn} = 0$. Hence $g_j' e_1 = 0$, so for $j = 2, \cdots, n$ we have $g_j(t_1 e_1 + y) = g_j(y)$ for every $t_1 \in R$ and $y \in Y$. Thus the restrictions of $f_2, \ldots, f_n$ to Y generates an $(n-1)$-dimensional distribution F_Y which obviously also is smooth and involutive. From our induction hypethesis there will thus exist a diffeomorphism φ_Y, defined near $0 \in Y$, such that $\varphi_{Y\star} F_Y$ is constant. Let φ be the extension of φ_Y, defined by

$$\varphi(t_1 e_1 + y) = t_1 e_1 + \varphi_Y(y).$$

Clearly φ is a diffeomorphism around $0 \in X$ such that $\varphi_\star F$ is constant near $\varphi(0)$. ∎

We will use the Frobenius theorem in order to prove existence of so called tangential manifolds.

Definition 2.1 *Let F be a smooth distribution, and let x_0 be a fixed point in X. A submanifold $\mathcal{G} \subseteq \mathcal{X}$ with $x_0 \in \mathcal{G}$ is called a* **tangential manifold** *through x for F, if $F(x) \subseteq T_{\mathcal{G}}(x)$ for each x in a neighborhood of x_0 in $\mathcal{G}$. Here $T_{\mathcal{G}}(x)$ denotes the tangent space to $\mathcal{G}$ at x.*

Remark 2.2 We note that the definition of a tangential manifold (the concept seems to be new) is similar to the definition of an integral manifold, but that for an integral manifold we have the inclusion $T_{\mathcal{G}}(x) \subseteq F(x)$. Just as one typically is looking for maximal integral manifolds, we will be looking for minimal tangential manifolds.

We now have the following result which we will use below.

Theorem 2.2 *Let F be an n-dimensional distribution, and let x_0 be a fixed point in X. There exists an n-dimensional tangential manifold through x for all x in a neighborhood of x_0 if and only if F is involutive.*

Proof. Assume that F is involutive. Then the Frobenius Theorem provides us with smooth vector fields $f_1, \ldots, f_n$ spanning F, and a local diffeomorphism such that the fields $\varphi_\star f_1, \ldots \varphi_\star f_n$ are constant. Denoting $\varphi_\star f_i$ by e_i we see that for

every x near x_0, the plane $P_x = \varphi(x) + Re_1 + \ldots + Re_n$ is an tangential manifold for the distribution $\varphi_\star F$, passing through $\varphi(x)$. The pull-back $\varphi^{-1}(P_x)$ of this plane is then an tangential manifold for F passing through x. The proof of the other implication is easy and the finite dimensional argument goes through (see e.g. [24]). ∎

Suppose now that we are given an n-dimensional involutive distribution F and a point $x_0 \in X$. From the result above we know that there exists a tangential manifold for F passing through x_0, and a natural question is if it possible to parameterize this manifold in some constructive way. This is in fact possible but we need some new notation.

Definition 2.2 *Let f be a smooth vector field on X, and let x be a fixed point in X. Consider the ODE*

$$\begin{cases} \dfrac{dx_t}{dt} &= f(x_t), \\ x_0 &= x. \end{cases}$$

We denote the solution x_t as $x_t = e^{ft}x$.

We have thus defined a group of operators $\{e^{ft} : t \in R\}$, and we note that the set $\{e^{ft}x : t \in R\} \subseteq X$ is nothing else than the integral curve of the vector field f, with initial point x, i.e. the curve obtained by starting at x and then "gluing together infinitesimally small pieces" of the vector field f.

Proposition 2.1 *Take as given an n-dimensional involutive distribution spanned by $f_1, \ldots, f_n$, and a point $x_0 \in X$. Denote the tangential manifold through x_0 by $\mathcal{G}$. Define the mapping $G : R^n \to X$ by*

$$G(z_1, \ldots z_n) = e^{f_n z_n} \ldots e^{f_1 z_1} x_0.$$

Then G is a local parameterization of $\mathcal{G}$ in the sense that there exist open neighborhoods $U \subseteq R^n$ and $V \subseteq \mathcal{G}$, of 0 and x_0 respectively, such that $V = G(U)$. Furthermore, the inverse of G restricted to V is a local coordinate system for $\mathcal{G}$ at x_0.

Proof. It follows directly from the definition of an tangential manifold that $G(z) \in \mathcal{G}$ for all z near 0 in R^n. Furthermore it is easy to see that $G'(0)h = \sum_{i=1}^n h_i f_i(x_0)$ and , since $f_1, \ldots, f_n$ are independent, $G'(0)$ is injective. The inverse function theorem does the rest. ∎

3 Realizations and Invariant Manifolds

3.1 Homogenous Systems

Take as given an m-dimensional standard Wiener process W with components $W^1, \ldots, W^m$, and a Hilbert space $\mathcal{H}$, where a generic point will be denoted by r. Let furthermore $\mu, \sigma_1, \ldots, \sigma_m$ be smooth vector fields on $\mathcal{H}$. For a given initial point $r^o \in \mathcal{H}$ we can then consider the following Stratonovich SDE on $\mathcal{H}$.

$$\begin{cases} dr_t &= \mu(r_t)dt + \sigma(r_t) \circ dW_t, \\ r_0 &= r^o, \end{cases} \tag{1}$$

where $\sigma(r_t) \circ dW_t = \sum_1^m \sigma_i(r_t) \circ dW_t^i$ and where $\circ$ denotes the Stratonovich integral. We note that, because of the assumed smoothness, the SDE (1) will locally (up to a positive stopping time), always have a unique strong solution. Let us emphasize that in the sequel we are dealing exclusively with local strong (as opposed to weak or mild) solutions to all SDEs. For information on SDEs in Hilbert space see [11], which is the standard text on the subject.

The reason for denoting the points in $\mathcal{H}$ by r is that in our applications below, the space $\mathcal{H}$ will be a space of forward rate curves.

Remark 3.1 Note that we are here only considering time invariant systems. For the time varying case, see next section.

The process r above is inherently an infinite dimensional process, but we will investigate under what conditions it can be realized by means of a finite dimensional SDE. We will to some extent use results from [2], where the reader also can find a more detailed conceptual discussion.

Definition 3.1 *We say that the SDE (1) has a (local) d-dimensional* **realization** *if there exists a point $z_0 \in R^d$, smooth vector fields $a, b_1, \ldots, b_m$ on some open subset $\mathcal{Z}$ of R^d and a smooth (submanifold) map $G : \mathcal{Z} \to \mathcal{H}$, such that r has the local representation*

$$r_t = G(Z_t), \tag{2}$$

where Z is the solution of the d-dimensional Stratonovich SDE

$$\begin{cases} dZ_t &= a(Z_t)dt + b(Z_t) \circ dW_t, \\ Z_0 &= z_0. \end{cases} \tag{3}$$

The prefix "local" above means that the representation is assumed to hold for all t with $0 \leq t < \tau(r^o)$, P-a.s. where, for each $r^o \in \mathcal{H}$, $\tau(r^o)$ is a a strictly positive stopping time.

We want to give conditions for the existence of a finite realization in terms of the local characteristics μ, σ and in this context a local realization is the best one can hope for. See [2] for a discussion on this point. We will thus often suppress the prefix "local", so in the sequel the word realization should always be interpreted as local realization.

The realization concept is closely connected to the concept of an invariant submanifold.

Definition 3.2 *A submanifold $\mathcal{G}$ in $\mathcal{H}$ is said to be (locally)* **invariant** *under the action of the SDE (1), if for every choice of $r^0 \in \mathcal{G}$ we have $r_t \in \mathcal{G}$ for $0 \leq t < \tau(r^o)$, where τ is a strictly positive stopping time.*

The first step to a solution of the realization problem lies in the following result.

Proposition 3.1 *There exists a local d-dimensional realization to (1) if and only if there exists an invariant submanifold $\mathcal{G}$ with $r^o \in \mathcal{G}$.*

Proof. For a formal proof see the proof of Proposition 4.2 in [2], which goes through in our present setting. A simple intuitive argument runs as follows. Suppose that there exists a finite dimensional invariant manifold $\mathcal{G}$ with $r^o \in \mathcal{G}$. Then $\mathcal{G}$ has a local coordinate system, and we may define the Z process as the local coordinate process for the r-process. On the other hand it is clear that if r has a finite dimensional realization as in (2)-(3), then all points r produced by the realization are of the form $r = G(z)$ so there obviously exists a finite dimensional invariant submanifold $\mathcal{G}$ containing the initial point r^o, namely $\mathcal{G} = Im[G] = \{G(z);\ z \in \mathcal{Z}\}$. ∎

The problem of finding a realization is thus reduced to the problem of finding a finite dimensional invariant submanifold and, using another result from [2], we can now connect this problem to the Frobenius theory in the preceeding section.

Theorem 3.1 *A submanifold $\mathcal{G}$ is invariant under the action of the SDE (1) if and only if, for every point $r \in \mathcal{G}$, the vectors $\mu(r), \sigma_1(r), \ldots, \sigma_m(r)$ belong to the tangent space of $\mathcal{G}$ at r. Thus $\mathcal{G}$ is invariant if and only if it is an tangential manifold for the distribution generated by $\mu, \sigma_1, \ldots, \sigma_m$.*

Proof. See [2], Theorem 4.1. ∎

Putting these results together we immediately have the following.

Proposition 3.2 *The SDE (1) has a finite dimensional realization if and only if there exists a finite dimensional tangential manifold for $\mu, \sigma_1, \ldots, \sigma_m$, containing the initial point r^o. The dimension of a minimal (w.r.t. the dimension d) realization coincides with the dimension of the minimal tangential manifold.*

Before going to our main result we need a new concept.

Definition 3.3 *Let F be a smooth distribution on $\mathcal{H}$. The* **Lie algebra** *generated by F, denoted by $\{F\}_{LA}$, is defined as the minimal (under inclusion) involutive distribution containing F.*

If, for example, the distribution F is spanned by the vector fields $f_1, \ldots, f_n$ then, to construct the Lie algebra $\{f_1, \ldots, f_n\}_{LA}$, you simply form all possible brackets, and brackets of brackets, etc. of the fields $f_1, \ldots, f_n$, and adjoin these to the original distribution until the dimension of the distribution is no longer increased.

We can now formulate and prove our main result on the existence of finite dimensional realizations.

Theorem 3.2 *Take as given the vector fields μ, σ and a point $\hat{r} \in \mathcal{H}$. The following statements are equivalent.*

- *For each choice of initial point r^o near $\hat{r} \in \mathcal{H}$, there exists a d-dimensional realization of the infinite dimensional SDE (1).*
- *The Lie algebra $\{\mu, \sigma_1, \ldots, \sigma_m\}_{LA}$ has dimension d near $\hat{r}$.*

Proof. From Frobenius it follows that $\mathcal{G}$ is an tangential manifold for the distribution spanned by $\{\mu, \sigma_1, \ldots, \sigma_m\}$, if and only if is tangential for $\{\mu, \sigma_1, \ldots, \sigma_m\}_{LA}$. The rest follows from Proposition 3.2. ∎

Remark 3.2 Note that, as far as finite realizations are concerned, we are only considering pure existence results. We do not treat the problem of how to construct a concrete realization. "In principle" this is simple: You "just" fix a coordinate system on the invariant manifold (obtained through Proposition 2.1) and write down the coordinate dynamics of the r-process, but this is easier said than done. Note also that there is no such thing as a unique realization, since any diffeomorphic transformation of the state space will give a new equivalent realization. The problem of finding canonical realizations is a subject of ongoing research.

In applications, when one is trying to construct the Lie algebra generated by a concrete choice of $\{\mu, \sigma_1, \ldots, \sigma_m\}$, the following observations are often quite useful

Lemma 3.1 *Take the vector fields $f_1, \ldots, f_k$ as given. The Lie algebra $\{f_1, \ldots, f_k\}_{LA}$ remains unchanged under the following operations.*

- *The vector field f_i may be replaced by αf_i, where α is any smooth nonzero scalar field.*
- *The vector field f_i may be replaced by*

$$f_i + \sum_{j \neq i} \alpha_j f_j,$$

where $\alpha_1, \ldots, \alpha_k$ are any smooth scalar field.

Proof. The first point is geometrically obvious, since multiplication by a scalar field will only change the length of the vector field f_i, and not its direction, and thus not the tangential manifold. Formally it follows from the "Leibnitz rule" $[f, \alpha \cdot g] = \alpha \cdot [f, g] - (\alpha' f) \cdot g$. The second point follows from the bilinear property of the Lie bracket together with the fact that $[f, f] = 0$. ∎

We end this section with some remarks on the structural stability of realizations. Consider therefore a given dynamical structure, specified by μ, σ. It is important to note that it may well happen that the SDE generated by $\{\mu, \sigma\}$ has a finite dimensional realization for a particular choice of initial point r^o, whereas no finite dimensional realization exists for points close to r^o. We say that such a system has a **non-generic** or "accidental" finite dimensional realization at r^o. If, on the other hand, the system has a finite dimensional realization for **all** points in a neighborhood of r^o we say that the system has a **generic** finite dimensional realization at r^o. The existence of a non-generic realization is of course of very limited value, since the situation is structurally unstable. We note (with satisfaction) that our Lie algebraic result above guarantees the existence of a **generic** finite dimensional realization.

A concrete, and very simplified picture of what is going on, is the following two-dimensional example. In the (x, y)-plane we consider the distribution spanned by the following vector fields

$$\begin{aligned} f(x, y) &= (y, -x), \\ g(x, y) &= \big((x^2 + y^2)y, -x\big) \end{aligned}$$

It is clear that these two fields are independent at every point not lying on the unit circle, whereas they are identical on the unit circle. For any point off the unit circle, the minimal tangential manifold will thus be two-dimensional and this property is generic. On the contrary, for any given point on the unit circle, the tangential manifold will be one-dimensional (it will be the circle itself), but this is a non-generic case, since the tangential manifold will be of higher dimension at points arbitrarily close to the starting point.

3.2 Interpreting the State Variables

As noted above, there is no such thing as a unique realization of a given system, since any diffeomorphic transformation of $\mathcal{Z}$ will provide an equivalent realization. Nevertheless, in many applications it is of interest to find realizations where the state variables have a natural (for example economic) interpretation. For that purpose we will now prove a fairly general result concerning coordinate systems, and for the rest of this subsection we make the following standing assumption.

Assumption 3.1 *We assume that the Hilbert space $\mathcal{H}$ is a space of functions of the form $r : I \to R$, where $x \longmapsto r(x)$. Here I is assumed to be some interval on the real line.*

In the applications below $\mathcal{H}$ will be a space of (analytic) forward rate curves.

Theorem 3.3 *Suppose that there exists a d-dimensional realization for a given system. Then there exists real numbers $x_1, \ldots, x_d$ in the interval I, and a realization Z such that $Z_t = (r_t(x_1), \ldots, r_t(x_d))$. If the space $\mathcal{H}$ is a space of real analytic functions, then the numbers $x_1, \ldots, x_d$ can, apart from a discrete set of "forbidden values", be choose completely freely.*

Proof. Since there exists a d-dimensional realization, there will exist a d-dimensional invariant manifold, $\mathcal{G}$, in our Hilbert space $\mathcal{H}$. It is enough to show that we can choose $x = (x_1, \ldots, x_d)$ in R^d such that the map $\hat{x} : \mathcal{G} \to R^d$, defined by $\hat{x}(r) = (r(x_1), \ldots, r(x_d))$ is a local coordinate mapping. By the inverse mapping theorem it is enough to show that the Frechet derivative of $\hat{x}$ is bijective. Let therefore r be a fixed point on $\mathcal{G}$ and denote by $T\mathcal{G}_r$ the tangent space of $\mathcal{G}$ at r. Since $\mathcal{G} \subseteq \mathcal{H}$ we can identify $T\mathcal{G}_r$ with a d-dimensional subspace of $\mathcal{H}$. Furthermore, since $\hat{x}$ as a mapping form $\mathcal{H}$ is linear and continuous, the derivative $\hat{x}_{\star r}$ of $\hat{x}$ at r is, for $q \in T\mathcal{G}_r$, given by $\hat{x}_{\star r}(q) = (q(x_1), \ldots, q(x_d))$. To prove that this mapping is bijective for some choice of x we let $e_1, \ldots, e_d$ be a basis for $T\mathcal{G}_r$. Defining, for $k \leq d$, the matrix

$$E_k(x) = \begin{bmatrix} e_1(x_1) & \ldots & e_1(x_k) \\ \vdots & & \vdots \\ e_k(x_1) & \ldots & e_k(x_k) \end{bmatrix},$$

it is thus enough to show that $\mathbf{det}\,[E_d(x)] \neq 0$ for some x in R^d. We proceed by induction on k.

Since e_1 is not the null function, there is some x_1 such that $\mathbf{det}\,[E_1(x)] = e_1(x_1) \neq 0$. Now assume that $x_1, \ldots, x_k$ have been chosen such that $\mathbf{det}\,[E_k(x)] \neq$

0 and consider

$$E_{k+1}(x) = \begin{bmatrix} e_1(x_1) & \dots & e_1(x_k) & e_1(x_{k+1} \\ \vdots & & \vdots & \vdots \\ e_k(x_1) & \dots & e_k(x_k) & e_k(x_{k+1}) \\ e_{k+1}(x_1) & \dots & e_{k+1}(x_k) & e_{k+1}(x_{k+1}) \end{bmatrix}.$$

Suppose that $\mathbf{det}\,[E_{k+1}(x)] = 0$ for any choice of x_{k+1}. Then, since $\mathbf{det}\,[E_k(x)] \neq 0$, the last row is a linear combination of the other rows. If the j:th row is denoted by r_j we would then have

$$r_{k+1} = \lambda_1 r_1 + \dots \lambda_k r_k.$$

However, since $\mathbf{det}\,[E_k(x)] \neq 0$, the λ:s only depend on $x_1, \dots, x_k$. Thus we would have

$$e_{k+1}(x_{k+1}) = \lambda_1 e_1(x_{k+1}), + \dots \lambda_k e_k(x_{k+1})$$

for every choice of x_{k+1}. This, however, is a contradiction since $e_1, \dots, e_k$ were supposed to be linearly independent. If $\mathcal{H}$ is a space of analytic functions, we use the fact that a nonzero analytic function only has isolated zeroes. ∎

3.3 Time Varying Systems

We now extend the analysis to cover time varying systems of the form

$$\begin{cases} dr_t &= \mu(r_t, t)dt + \sigma(r_t, t) \circ dW_t, \\ r_s &= r^o. \end{cases} \tag{4}$$

Note that we no longer initiate the system at $t = 0$ but at an arbitrary point s in time. The relevant definitions of realizations and invariant manifolds are the natural ones with the obvious extension of the prefix "local".

Definition 3.4 *We say that the SDE (4) has a (local) d-dimensional* **realization** *at (s, r^o), if there exists a point $z_s \in R^d$, smooth vector fields $a, b_1, \dots, b_m$ on some open subset $\mathcal{Z}$ of R^d and a smooth (submanifold) map $G : \mathcal{Z} \to \mathcal{H}$, such that r has the local representation*

$$r_t = G(Z_t), \quad t \geq s \tag{5}$$

where Z is the solution of the d-dimensional Stratonovich SDE

$$\begin{cases} dZ_t &= a(Z_t)dt + b(Z_t) \circ dW_t, \\ Z_s &= z_s. \end{cases} \tag{6}$$

Definition 3.5 *A submanifold $\mathcal{G}$ in $\mathcal{H}$ is said to be (locally)* **invariant** *under the action of the SDE (4), if for every choice of $(s, r^0) \in R_+ \times \mathcal{G}$ we have $r_t \in \mathcal{G}$ locally for $s \leq t$.*

As before we are investigating the existence of finite dimensional realizations and invariant manifolds. We handle this new situation by enlarging the state space and introducing running time as a new state. Defining, for each fixed s, the Wiener process W^s as $W_t^s = W_t - W_s$, we can write (4) as

$$\begin{cases} dr_u &= \mu(r_u, t)du + \sigma(r_u, t) \circ dW_u^s, \\ dt &= 1 \cdot du + 0 \cdot dW_u^s, \\ r_0 &= r^o, \\ t_0 &= s. \end{cases}$$

We are led to the following natural notation.

Definition 3.6 *Define the following extended objects.*

$$\hat{\mathcal{H}} = \mathcal{H} \times R, \tag{7}$$

$$\hat{r} = \begin{bmatrix} r \\ t \end{bmatrix}, \tag{8}$$

$$\hat{\mu}(\hat{r}) = \begin{bmatrix} \mu(r, t) \\ 1 \end{bmatrix}, \tag{9}$$

$$\hat{\sigma}(\hat{r}) = \begin{bmatrix} \sigma(r, t) \\ 0 \end{bmatrix}, \tag{10}$$

We can now write (4) as the following SDE on $\hat{\mathcal{H}}$ where, in order to save space, we suppress the superindex in W^s. and

$$\begin{cases} d\hat{r}_u &= \hat{\mu}(\hat{r}_u)du + \hat{\sigma}(\hat{r_u}) \circ dW_u, \\ \hat{r}_0 &= \hat{r}^o. \end{cases} \tag{11}$$

Here we have used the notation

$$\hat{r}^o = \begin{bmatrix} r^o \\ s \end{bmatrix}.$$

It is now easy to see that a manifold $\mathcal{G}$ is invariant in $\mathcal{H}$ for the time varying system (4) if and only if $\mathcal{G} \times R$ is invariant for the homogenous system (11). Furthermore it is clear that finding a realization for (4) is equivalent to that of finding a realization for (11), and we have our main result, which follows immediately from Theorem 3.2.

Theorem 3.4 *The time varying system (4) has a finite dimensional realization if and only if*

$$dim\{\hat{\mu}, \hat{\sigma}\}_{LA} < \infty$$

4 Interest Rate Models

4.1 Basics

As in Heath, Jarrow, and Morton (1992), we consider a default free bond market and let $p(t,T)$ denote the price at time $t \geq 0$ of a zero coupon bond maturing at $T \geq t$. We assume frictionless markets and perfectly divisible bonds. The forward rates $f(t,T)$ are as usual defined by

$$f(t,T) = -\frac{\partial \log p(t,T)}{\partial T}, \tag{12}$$

and the short rate R is defined by

$$R(t) = f(t,t).$$

Consider now as given a concrete term structure model, defined on a filtered probability space $(\Omega, \mathcal{F}, Q, \{\mathcal{F}_t\}_{t\geq 0})$. We assume that the model is free of arbitrage in the sense that the probability measure Q is a martingale measure for the model, i.e. for each fixed T, the process $p(t,T)/B_t$ is a Q-martingale, where the money account B is defined as usual by $B_t = \exp\left\{\int_0^t R(s)ds\right\}$. The physical probability measure P will play no role below, so all calculations are carried out under Q.

More specifically, we assume that under the martingale measure Q, the dynamics of the forward rates are of the form

$$df(t,T) = \alpha(t,T)dt + \sigma_0(t,T)dW_t, \tag{13}$$

where W is an m-dimensional Q-Wiener process.

As we will see below, it is convenient to use the Musiela ([6], [21]) parameterization of forward rates

$$r(t,x) = f(t,t+x), \tag{14}$$

where the symbol x denotes time **to** maturity as opposed to T, which denotes time **of** maturity. We denote the induced dynamics for the r-process by

$$dr(t,x) = \beta(t,x)dt + \sigma(t,x)dW_t, \tag{15}$$

and it is easy to see that there is a one-to-one correspondence between the formulations (13) and (15), namely

$$\beta(t,x) = \frac{\partial}{\partial x}r(t,x) + \alpha(t,t+x), \tag{16}$$

$$\sigma(t,x) = \sigma_0(t,t+x). \tag{17}$$

The HJM no arbitrage drift condition can now be transferred to the new parameterization. Writing * for transpose, the result is as follows ([21]).

Proposition 4.1 (Forward Rate Equation) *Under the martingale measure Q, the r-dynamics are*

$$dr(t,x) = \left\{ \frac{\partial}{\partial x} r(t,x) + \sigma(t,x) \int_0^x \sigma(t,u)^\star du \right\} dt + \sigma(t,x) dW_t. \qquad (18)$$

4.2 The Space

The big problem with the forward rate equation (18), from our point of view, is the operator $\frac{\partial}{\partial x}$ in the drift term. In most "naturally" chosen spaces, like a weighted Sobolev space, the operator is unbounded, which means that we have no existence results concerning strong solutions. Thus, in order to study the realization problem using our previously developed framework, we need a very regular space to work in (see Remark 4.3).

Definition 4.1 *Consider fixed real numbers $\beta > 1$, and $\gamma > 0$. The space $\mathcal{H}_{\beta,\gamma}$ is defined as the space of all infinitely differentiable functions*

$$r : R_+ \to R$$

satisfying the norm condition $\|r\|_{\beta,\gamma} < \infty$. Here the norm is defined as

$$\|r\|^2_{\beta,\gamma} = \sum_{n=0}^{\infty} \beta^{-n} \int_0^\infty \left(\frac{d^n r}{dx^n}(x) \right)^2 e^{-\gamma x} dx.$$

Note that $\mathcal{H}_{\beta,\gamma}$ is not a space of distributions, but a space of functions. The reason for the exponential weighting is to include all constant functions in the space. Most of the results below are uniform w.r.t. (β,γ), so in the sequel we will often suppress the subindices. With the obvious inner product $\mathcal{H}$ is a pre-Hilbert space, and we have the following result.

Proposition 4.2 *The space $\mathcal{H}$ is a Hilbert space (i.e. complete), and if $f_n \to f$ in $\mathcal{H}$ then $f_n^{(m)} \to f^{(m)}$ uniformly on compacts for every $m \geq 0$, where $f^{(m)} = d^m f / dx^m$. Thus in particular, for any fixed $x \in R_+$, the point evaluation mapping $r \longmapsto r(x)$ is a bounded linear functional. Furthermore, every function in $\mathcal{H}$ is in fact real analytic, and can thus be uniquely extended to a holomorphic function in the entire complex plane.*

Proof. Let us introduce measures ν and μ on N (the natural numbers) and R_+ respectively (with obvious sigma algebras), by

$$\begin{aligned} \nu(E) &= \sum_{n \in E} \beta^{-n}, \quad E \subseteq N, \\ \mu(F) &= \int_F e^{-\gamma x} dx. \quad F \subseteq R_+. \end{aligned}$$

Let furthermore $\nu \times \mu$ denote the product measure.

For every $f \in C^\infty(R_+)$ we define $Tf : N \times R_+ \to R$ by $Tf(n,x) = f^{(n)}(x)$, so $\mathcal{H} = \left\{ f \in C^\infty(R_+) : \ Tf \in L^2(\nu \times \mu) \right\}$.

Now let $\{f_n\}$ be a Cauchy sequence in $\mathcal{H}$ and put $f_{n,k} = f_n - f_k$. Then, as $n, k \to \infty$, $Tf_{n,k} \to 0$ in $L^2(\nu \times \mu)$ Thus there exists some $g \in L^2(\nu \times \mu)$ such that $Tf_n \to g$ in $L^2(\nu \times \mu)$, and we can choose a subsequence $\{f_{n'}\}$ from $\{f_n\}$ such that

$$Tf_{n'} \to g, \quad (\nu \times \mu) - a.e.$$

But then, for every $m \geq 0$ we have that

$$f_n^{(m)} \to g_m, \quad in \ L^2(\mu),$$

and

$$f_{n'}^{(m)} \to g_m, \quad (\mu) - a.e.,$$

where $g_m(x) = g(m,x)$. Let, for $T > 0$, $C[0,T]$ denote the space of continuous functions $[0,T] \to R$ with supremum norm. We now want to prove that, for every m, $\left\{ f_n^{(m)} \right\}$ is a Cauchy sequence in $C[0,T]$, and for this we need the following lemma.

Lemma 4.1 *The sequence* $\left\{ f_n^{(m)}(0) \right\}_n$ *is Cauchy in* R.

Proof of Lemma: To get a contradiction, assume that there exists some $\epsilon > 0$ such that, for every N, $f_{n,k}^{(m)}(0) > \epsilon$ for some $n, k > N$. Now consider

$$f_{n,k}^{(m)}(x) = f_{n,k}^{(m)}(0) + \int_0^x f_{n,k}^{(m+1)}(y)dy.$$

We have, however,

$$\begin{aligned} & \left| \int_0^x f_{n,k}^{(m+1)}(y)dy \right|^2 = \left| \int_0^\infty e^{(\gamma y)/2} I_{[0,x]}(y) e^{-\gamma y/2} f_{n,k}^{(m+1)}(y)dy \right|^2 \\ \leq & \left(\int_0^x e^{\gamma y} dy \right) \left(\int_0^\infty \left\{ f_{n,k}^{(m+1)}(y) \right\}^2 e^{-\gamma y} dy \right) = \left(\frac{e^{\gamma x} - 1}{\gamma} \right) \left\| f_{n,k}^{(m+1)} \right\|_{L^2(\mu)}^2 \end{aligned}$$

Thus

$$\left\| f_{n,k}^{(m)} - f_{n,k}^{(m)}(0) \right\|_{C[0,T]} \leq K \left\| f_{n,k}^{(m+1)} \right\|_{L^2(\mu)}$$

for some K depending on γ and T. Now choose N such that

$$K \left\| f_{n,k}^{(m+1)} \right\|_{L^2(\mu)} < \epsilon/2$$

if $n, k > N$. Then obviously $f_{n,k}^{(m)}(x) > \epsilon/2$ for all $x \in [0,T]$ and for arbitrarily large n, k. This, however, contradicts the assumption that $\| f_{n,k}^{(m)} \|_{L^2(\mu)} \to 0$

when $n, k \to \infty$. Thus $f_{n,k}^{(m)}(0) \to 0$ as $n, k \to \infty$.
End of proof of Lemma.

From the above proof we also have

$$\|f_{n,k}^{(m)}\|_{C[0,T]} \leq \left|f_{n,k}^{(m)}(0)\right| + K\|f_{n,k}^{(m+1)}\|_{L^2(\mu)} \to 0$$

as $n, k \to \infty$, so $\left\{f_n^{(m)}\right\}$ is Cauchy in $C[0,T]$. Thus $f_n^{(m)} \to h_m$ in $C[0,T]$ for some h_m in $C[0,T]$. Since $f_{n'}^{(m)} \to g_m,\ \mu - a.e.$ we conclude that $g_m = h_m,\ \mu - a.e.$ on $[0,T]$.

From

$$f_n^{(m)}(x) = f_n^{(m)}(0) + \int_0^x f_n^{(m+1)}(y)dy,$$

we get

$$h_m(x) = h_m(0) + \int_0^x h_{m+1}(y)dy,$$

from which it follows that $h_m' = h_{m+1}$, and consequently that $Th_0 = g$, $h_0 \in \mathcal{H}$ and $f_n \to h_0$ in $\mathcal{H}$. Thus we have proved completeness and uniform convergence of derivatives on compacts.

To prove analyticity we make a Taylor expansion with exact rest term. Thus

$$f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(0)x^k + \frac{1}{n!}\int_0^x t^n f^{(n+1)}(x-t)dt,$$

and it is enough to prove that this series converges for every x, i.e, to prove that the remainder term tends to zero as $n \to \infty$. Using obvious estimates and Cauchy-Schwartz we obtain

$$\begin{aligned} & \frac{1}{n!}\left|\int_0^x t^n f^{(n+1)}(x-t)dt\right| \leq \frac{x^n}{n!}\int_0^x \left|f^{(n+1)}(t)\right| dt \\ \leq\ & \frac{x^n \beta^{\frac{n}{2}}}{n!}\sqrt{\frac{e^{\gamma x}-1}{\gamma}}\left(\beta^{-n}\int_0^\infty e^{-\gamma t}\left|f^{(n+1)}(t)\right|^2\right)^{1/2}. \end{aligned}$$

In this product the first term tends to zero because of the factorial, the second is constant, and the third tend to zero because $\|f\|_{\beta,\gamma} < \infty$. ∎

4.3 The General Model

We take as given a volatility σ of the form

$$\sigma : \mathcal{H} \times R_+ \to R^m,$$

i.e. each component of $\sigma(r,x) = [\sigma_1(r,x), \ldots, \sigma_m(r,x)]$ is a functional of the infinite dimensional r-variable, and a function of the real variable x. We will

also view σ_i as a mapping from $\mathcal{H}$ to a space of functions, and we will in fact assume that each component σ_i is a smooth vector field on $\mathcal{H}$.

Remark 4.1 Note that we are primarily interested in homogenous systems, i.e. we consider forward rate volatilities without an explicit time-dependence. This is partly due to the fact that it (to us) seems unnatural to assume time-varying volatilities, but mainly in order to save space. In our experience, the introduction of a time dependent volatility seems to add very little of interest, and time varying models are easily handled by the methods in Section 3.3. We will give examples below.

We need to introduce some more compact notation.

Definition 4.2 *In the sequel, we will denote integration w.r.t. time to maturity by* $\mathbf{H}$*. Thus, we will write*

$$\mathbf{H}\sigma(r,x) = \int_0^x \sigma(r,s)ds. \tag{19}$$

Suppressing the x-variable, the Itô dynamics (18) of the forward rates under the martingale measure Q can thus be written as

$$dr_t = \left\{ \frac{\partial}{\partial x} r_t + \sigma(r_t)\mathbf{H}\sigma(r_t)^\star \right\} dt + \sigma(r_t)dW_t. \tag{20}$$

Remark 4.2 Note that $\sigma(r_t)\mathbf{H}\sigma(r_t)^\star$ is shorthand for $\sum_1^m \sigma_i(r_t)\mathbf{H}\sigma_i(r_t)$, where the product $\sigma_i(r_t)\mathbf{H}\sigma_i(r_t)$ is, for each fixed r, the pointwise product of two real valued functions. This notational convention will be kept in the sequel.

It is not hard to show (see [2]) that the forward rate model on Stratonovich form is given by

$$dr_t = \mu(r_t)dt + \sigma(r_t) \circ dW_t \tag{21}$$

where

$$\mu(r) = \frac{\partial}{\partial x} r + \sigma(r)\mathbf{H}\sigma(r)^\star - \frac{1}{2}\sigma_r'(r)[\sigma(r)], \tag{22}$$

where $\sigma_r'(r)[\sigma(r)]$ denotes the Frechet derivative $\sigma'(r)$ operating on $\sigma(r)$. We need some regularity assumptions.

Assumption 4.1 *From now on we assume that* σ *has the following properties.*

- *The mappings* $\sigma_1, \ldots, \sigma_m$ *are smooth vector fields on* $\mathcal{H}$*.*
- *The mapping*

$$r \longmapsto \sigma(r)\mathbf{H}\sigma(r)^\star$$

is a smooth vector field on $\mathcal{H}$*.*

Remark 4.3 The reason for our choice of $\mathcal{H}$ as the underlying space, is that the linear operator $\mathbf{F} = \frac{\partial}{\partial x}$ is bounded in this space. Together with the assumptions above, this implies that both μ and σ are smooth vector fields on $\mathcal{H}$, thus ensuring the local existence of a strong solution to the forward rate equation for every initial point $r^o \in \mathcal{H}$.

4.4 Main Results

We now go on to apply our general results from Sections 2 and 3 to the class of interest rate models presented in the Section 4.3. Assumption 4.1 is in force throughout the entire section.

Theorem 4.1 (Main Theorem) *Take as given the volatility mapping $\sigma = (\sigma_1, \ldots, \sigma_m)$ as well as an initial forward rate curve $r^o \in \mathcal{H}$. Then the forward rate model generated by σ generically admits a finite dimensional realization at r^o, if and only if*

$$dim \{\mu, \sigma_1, \ldots, \sigma_m\}_{LA} < \infty \tag{23}$$

in a neighborhood of r^o, where μ is given by

$$\mu(r) = \frac{\partial}{\partial x} r + \sigma(r)\mathbf{H}\sigma(r)^\star - \frac{1}{2}\sigma_r'(r)[\sigma(r)], \tag{24}$$

and $\mathbf{H}\sigma$ is defined by

$$\mathbf{H}\sigma(r, x) = \int_0^x \sigma(r, s)ds. \tag{25}$$

Proof. Follows at once from Theorem 3.2 and (22). ∎

In the sequel we will, in order to save space, denote the Lie algebra above by $\{\mu, \sigma\}_{LA}$, by $\mathcal{L}(\mu, \sigma)$ or by $\mathcal{L}$.

If the assumptions of Theorem 4.1 are in force, we can furthermore use Proposition 2.1 to give a give a description of the minimal invariant (tangential) manifold, containing r^o, which is generated by the interest rate dynamics. In other words, we can describe the structure of all possible forward curves that can be produced by the model.

Theorem 4.2 *Assume that the Lie algebra $\{\mu, \sigma\}_{LA}$ is spanned by the smooth vector fields $f_1, \ldots, f_d$. Then, for the initial point r^o, all forward rate curves produced by the model will belong to the induced tangential manifold $\mathcal{G}$, which can be parameterized as $\mathcal{G} = Im[G]$, where*

$$G(z_1, \ldots, z_d) = e^{f_d z_d} \ldots e^{f_1 z_1} r^o, \tag{26}$$

and where the operator $e^{f_i z_i}$ is given in Definition 2.2.

With this machinery we can also very easily solve a related question, which was left open in [2]. Consider a fixed interest rate model, specified by the volatility σ and also a fixed family of forward rate curves parameterized by the mapping $G_0 : R^k \to \mathcal{H}$. Now, if $\mathcal{G}_0 = Im[G_0]$ is invariant, then the interest rate model will, given any initial point r^o in $\mathcal{G}_0$, only produce forward rate curves belonging to $\mathcal{G}_0$, in which case we say that the model and the family $\mathcal{G}_0$ are **consistent**. If the family is not consistent, then an initial forward rate curve in $\mathcal{G}_0$ may produce future forward rate curves outside $\mathcal{G}_0$, and the question arises how to construct the smallest possible family of forward rate curves which contains the initial family $\mathcal{G}_0$, and is consistent (i.e. invariant) w.r.t the interest rate model. As a concrete example, one may want to find the minimal extension of the Nelson-Siegel family of forward rate curves (see [22], [2]) which is consistent with the Hull-White (extended Vasiček) model. In particular one would like to know under what conditions this minimal extension of $\mathcal{G}_0$ is finite dimensional.

In geometrical terms we thus want to construct the minimal manifold containing $\mathcal{G}_0$, which is tangential w.r.t. the vector fields $\mu, \sigma_1, \ldots, \sigma_m$. The solution is obvious: For very point on $\mathcal{G}_0$ we construct the minimal tangential manifold through that point, and then we define the extension $\mathcal{G}$ as the union of all these fibers. Thus we have the following result, the proof of which is obvious. Concrete applications will be given below.

Proposition 4.3 *Consider a fixed volatility mapping σ, and let $\mathcal{G}_0$ be a k-dimensional submanifold parameterized by $G_0 : R^k \to \mathcal{H}$. Then $\mathcal{G}_0$ can be extended to a finite dimensional invariant submanifold $\mathcal{G}$, if and only if*

$$dim\left\{\mu, \sigma_1, \ldots, \sigma_m\right\}_{LA} < \infty.$$

Moreover, if $\mathcal{G}_0$ is transversal to $\{\mu, \sigma\}_{LA}$ and if the Lie algebra is spanned by the independent vector fields $f_1, \ldots, f_d$, then $dim\mathcal{G} = k + d$ and a parameterization of $\mathcal{G}$ is given by the map $G : R^{k+d} \to \mathcal{H}$, defined by

$$G(z_1, \ldots, z_k, y_1, \ldots, y_d) = e^{f_d y_d} \ldots e^{f_1 y_1} G(z_1, \ldots, z_k). \tag{27}$$

Remark 4.4 The term "transversal" above means that no vector in the Lie algebra $\mathcal{L}(\mu, \sigma)$ is contained the tangent space of $\mathcal{G}_0$ at any point of $\mathcal{G}_0$. This prohibits an integral curve of $\mathcal{L}$ to be contained in $\mathcal{G}_0$, which otherwise would lead to an extension with lower dimension than $d + k$. In such a case the parameterization above would amount to an over parameterization in the sense that G would not be injective.

We end this section by showing that it is always possible to choose the state variables as a set of benchmark forward rates.

Proposition 4.4 *Suppose that $dim\{\mu, \sigma_1, \ldots, \sigma_m\}_{LA} = d$. Then, for almost every choice of distinct benchmark maturities $x_1, \ldots, x_d$, the realization can be*

chosen such that the state process Z_t is given by $Z_t = (r_t(x_1), \dots r_t(x_d))$. The expression "almost every choice" above means that, apart from a discrete set of forbidden values, $x_1, \dots, x_d$ can be chosen freely.

Proof. Follows immediately from Theorem 3.3 and the fact that $\mathcal{H}$ is a space of analytic functions. ∎

5 Deterministic Volatility

We now move on to present some applications of the theory developed above. We start with the simplest case, which is when the volatility $\sigma(r, x)$ does not depend on r, i.e. when we can write $\sigma(r, x) = \sigma(x)$. This case was first studied in [21], and more or less completely analyzed in [3]. For simplicity we also assume (for the moment) that we have only one driving Wiener process, so the volatility is a constant vector field in $\mathcal{H}$.

5.1 Finite Dimensional Realizations

In order to use Theorem 4.1 we now compute the Lie algebra $\mathcal{L}(\mu, \sigma)$. Since the vector field σ is constant we have $\sigma' = 0$ so, from (24) we have

$$\mu(r) = \mathbf{F}r + D,$$

where $D(x) = \sigma(x) \int_0^x \sigma(u)du$. The Frechet derivatives are trivial in this case. Since $\mathbf{F}$ is linear (and bounded in our space), we obtain

$$\begin{aligned} \mu' &= \mathbf{F}, \\ \sigma' &= 0. \end{aligned}$$

Thus the Lie bracket $[\mu, \sigma]$ is given by

$$[\mu, \sigma] = \mathbf{F}\sigma,$$

and in the same way we have

$$[\mu, [\mu, \sigma]] = \mathbf{F}^2\sigma.$$

Continuing in the same way it is easily seen that the relevant Lie algebra $\mathcal{L}$ is given by

$$\mathcal{L} = \{\mu, \sigma\}_{LA} = span\left\{\mu, \sigma, \mathbf{F}\sigma, \mathbf{F}^2\sigma, \dots\right\}.$$

It is thus clear that $\mathcal{L}$ is finite dimensional (at each point r) if and only if the function space

$$span\left\{\mathbf{F}^n\sigma \; ; n = 0, 1, 2, \dots\right\}$$

is finite dimensional. This argument easily carries over to the case of several driving Wiener processes, and we have the following result, which originally was proved in [3]. See also [21] for a closely related result.

Proposition 5.1 *Assume that the volatility components $\sigma_1 \dots, \sigma_m$ are deterministic. Then the model possesses a finite dimensional realization if and only if the function space*

$$span\left[\mathbf{F}^k \sigma_i \ ; \ i = 1, \cdots, m. \ k = 0, 1, \cdots\right]$$

is finite dimensional. If the dimension of the function space above equals d, then the dimension of a minimal realization is $d + 1$.

We thus have a finite dimensional realization if and only if the components of σ solve, as functions of x, a multidimensional linear ODE with constant coefficients. Using standard results from ODE theory, a more concrete characterization is given as follows.

Corollary 5.1 *Given the assumptions above there exists a finite dimensional realization if and only if the volatility $\sigma(x) = [\sigma_1(x), \dots, \sigma_m(x)]$ can be written*

$$\sigma(x) = ce^{Ax}B, \tag{28}$$

where c is a row vector, whereas A and B are matrices. In particular, if every component of σ is of the form

$$\sigma_i(x) = p_i(x)e^{-\lambda_i x}, \tag{29}$$

for some polynomial p_i and some real positive number λ_i, then there exists a finite dimensional realization.

Remark 5.1 Let us call a function of the form $ce^{Ax}b$, where c is a row vector, A is a square matrix and b is a column vector, a **quasi-exponential** (or QE) function. All functions of the form (29) are of course quasi exponential, and the general form of a quasi-exponential function f is given by

$$f(x) = \sum_i e^{\lambda_i x} + \sum_j e^{\alpha_i x}\left[p_j(x)\cos(\omega_j x) + q_j(x)\sin(\omega_j x)\right], \tag{30}$$

where $\lambda_i, \alpha_1, \omega_j$ are real numbers, whereas p_j and q_j are real polynomials.

QE functions will turn up again, so we list some simple properties.

Lemma 5.1 *The following hold for the quasi-exponential functions.*

- *A function is QE if and only if it is a component of the solution of a vector valued linear ODE with constant coefficients.*

- *A function is QE if and only if it can be written as* $f(x) = ce^{Ax}b$.
- *If* f *is QE, then* f' *is QE.*
- *If* f *is QE, then its primitive function is QE.*
- *If* f *and* g *are QE, then* fg *is QE.*

5.2 Invariant Manifolds

We now turn to the construction of invariant manifolds, and to this end we assume that the Lie algebra above is finite dimensional. Thus it is spanned by a finite number of vector fields as

$$\{\mu, \sigma\}_{LA} = span\left\{\mu, \sigma_i^{(k)};\ i = 1, \ldots, m;\ k = 0, 1, \ldots, n_i\right\},$$

where

$$\sigma_i^{(k)}(x) = \frac{\partial^k \sigma_i}{\partial x^k}(x).$$

In order to apply Theorem 4.2 and Proposition 4.3, we have to compute the operators $\exp[\mu t]$ and $\exp\left[\sigma_i^{(k)}\right]$, i.e. we have to solve $\mathcal{H}$-valued ODEs. We recall that

$$\mu(r) = \mathbf{F}r + D,$$

where the constant field D is given

$$D(x) = \sum_i^m \sigma_i(x) \int_0^x \sigma_i(s)ds,$$

which can be written as

$$D(x) = \frac{1}{2}\frac{\partial}{\partial x}\|S(x)\|^2,$$

where, as before, $S(x) = \int_0^x \sigma(s)ds$. Thus $e^{\mu t}$ is obtained by solving

$$\frac{dr}{dt} = \mathbf{F}r + D.$$

This is a linear equation with solution

$$r_t = e^{\mathbf{F}t}r_0 + \int_0^t e^{\mathbf{F}(t-s)}Dds$$

so

$$\left(e^{\mu t}r_0\right)(x) = r_0(x+t) + \frac{1}{2}\left(\|S(x+t)\|^2 - \|S(x)\|^2\right).$$

The vector fields $\sigma_i^{(k)}$ are constant, so the corresponding ODEs are trivial. We have

$$e^{\sigma_i^{(k)}t}r_0 = r_0 + \sigma_i^{(k)}t.$$

We thus have the following results on the parameterization of invariant manifolds. For a given mapping $G : R^n \to \mathcal{H}$, we write $G(z)(x)$ or $G(z,x)$ to denote the function $G(z) \in \mathcal{H}$ evaluated at $x \in R_+$.

Proposition 5.2 *The invariant manifold generated by the initial forward rate curve r_0 is parameterized as*

$$\begin{aligned} & G(z_0, z_i^k;\ i = 1, \ldots, m;\ k = 0, \ldots, n_i)(x) \\ = \quad & r_0(x + z_0) + \frac{1}{2}\left(\|S(x+t)\|^2 - \|S(x)\|^2\right) + \sum_{i=1}^{m}\sum_{k=0}^{n_i} \sigma_i^{(k)}(x) z_i^k. \end{aligned}$$

If the k-dimensional manifold $\mathcal{G}_0$ is transversal to $\mathcal{L}\{\mu, \sigma\}$ and parameterized by $G_0(y_1, \ldots, y_k)$, then the minimal consistent (i.e. invariant) extension is parameterized as

$$\begin{aligned} & G(y_1, \ldots, y_k, z_0, z_i^k;\ i = 1, \ldots, m;\ k = 0, \ldots, n_i)(x) \\ = \quad & G_0(y_1, \ldots, y_k)(x + z_0) + \frac{1}{2}\left(\|S(x+z_0)\|^2 - \|S(x)\|^2\right) + \sum_{i=1}^{m}\sum_{k=0}^{n_i} \sigma_i^{(k)}(x) z_i^k. \end{aligned}$$

Note that if $\mathcal{G}_0$ is invariant under shift in the x-variable (this is in fact the typical case), then a simpler parameterization of $\mathcal{G}$ is given by

$$\begin{aligned} & G(y_1, \ldots, y_k, z_0, z_i^k;\ i = 1, \ldots, m;\ k = 0, \ldots, n_i)(x) \\ = \quad & G_0(y_1, \ldots, y_k)(x) + \frac{1}{2}\left(\|S(x+z_0)\|^2 - \|S(x)\|^2\right) + \sum_{i=1}^{m}\sum_{k=0}^{n_i} \sigma_i^{(k)}(x) z_i^k. \end{aligned}$$

As a concrete application let us consider the simple case when $m = 1$ and

$$\sigma(x) = \sigma e^{-ax},$$

where, with a slight abuse of notation, a and σ denote positive constants. As is well known, this is the HJM formulation of the Hull-White extension of the Vasiček model [16],[25]. In this case we have

$$S(x) = \frac{\sigma}{a}\left[1 - e^{-ax}\right].$$

The relevant function space

$$\{\mathbf{F}^n\sigma;\ n \geq 0\} = \left\{\frac{\partial^n}{\partial x^n} e^{-ax};\ n \geq 0\right\}$$

is obviously one-dimensional and spanned by the single function e^{-ax}, so the Lie algebra is two-dimensional.

As the given manifold $\mathcal{G}_0$ we take the Nelson-Siegel ([22]) family of forward rate curves, parameterized as

$$G_0(y_1, \ldots, y_4)(x) = y_1 + y_2 e^{-y_4 x} + y_3 x e^{-y_4 x}.$$

This family is obviously invariant under shift in x, so we have the following result.

Proposition 5.3 *For a given initial forward rate curve r_0, the invariant manifold generated by the Hull-White extended Vasiček model is parameterized by*

$$G(z_0, z_1)(x) = r_0(x+z_0) + e^{-ax}\frac{\sigma^2}{a^2}\left[1 - e^{-az_0}\right] - e^{-2ax}\frac{\sigma^2}{2a^2}\left[1 - e^{-2az_0}\right] + z_1 e^{-ax}.$$

The minimal extension of the NS family consistent with the Hull-White extended Vasiček model is parameterized by

$$\begin{aligned} G(z_0, z_1, y_1, \ldots, y_4)(x) &= y_1 + y_2 e^{-y_4 x} + y_3 x e^{-y_4 x} \\ &+ e^{-ax}\frac{\sigma^2}{a^2}\left[1 - e^{-az_0}\right] - e^{-2ax}\frac{\sigma^2}{2a^2}\left[1 - e^{-2az_0}\right] + z_1 e^{-ax}. \end{aligned}$$

6 Deterministic Direction Volatility

We go on to study the most natural extension of the deterministic volatility case, namely the case when the volatility is of the form

$$\sigma_i(r, x) = \varphi_i(r)\lambda_i(x), \quad i = 1, \ldots, m. \tag{31}$$

In this case the individual vector field σ_i has constant direction $\lambda_i \in \mathcal{H}$, but is of varying length, determined by φ_i. For ease of exposition we mainly restrict ourselves to the time invariant case, but the results can easily be extended to the time varying case (see below).

In order to avoid trivialities we make the following assumption.

Assumption 6.1 *We assume that $\varphi_i(r) \neq 0$ for all $r \in \mathcal{H}$ and for all $i = 1, \ldots, m$.*

Note that φ_i is allowed to be any smooth functional of the entire forward rate curve. The simpler special case when φ_i is a point evaluation of the short rate, i.e. of the form $\varphi(r) = h(r(0))$ has been studied in [1], [17] and [23]. A more general case, treated in [9], occurs when φ is a finite point evaluation, i.e. when $\varphi(r) = h(r(x_1), \ldots r(x_k))$ for fixed benchmark maturities $x_1, \ldots, x_k$. All these setups are special cases of our present framework and, apart from [9] which also consider non-multiplicative volatilities, all results are included and extended below.

The special case of only one driving Wiener process turns out to have some features which will be used later, so we make a separate investigation of that particular case.

6.1 A Scalar Driving Wiener Process

In this case we have $\sigma(r,x) = \varphi(r)\lambda(x)$, and after a simple calculation the drift vector μ turns out to be

$$\mu(r) = \mathbf{F}r + \varphi^2(r)D - \frac{1}{2}\varphi'(r)[\lambda]\varphi(r)\lambda, \tag{32}$$

where $\varphi'(r)[\lambda]$ denotes the Frechet derivative $\varphi'(r)$ acting on the vector λ, and where the constant vector $D \in \mathcal{H}$ is given by

$$D(x) = \lambda(x)\int_0^x \lambda(s)ds.$$

We now want to know under what conditions on φ and λ we have a finite dimensional realization, i.e. when the Lie algebra generated by

$$\begin{aligned}\mu(r) &= \mathbf{F}r + \varphi^2(r)D - \frac{1}{2}\varphi'(r)[\lambda]\varphi(r)\lambda,\\ \sigma(r) &= \varphi(r)\lambda,\end{aligned}$$

is finite dimensional. Under Assumption 6.1 we can use Lemma 3.1, to see that the Lie algebra is in fact generated by the simpler system of vector fields

$$\begin{aligned}f_0(r) &= \mathbf{F}r + \Phi(r)D,\\ f_1(r) &= \lambda,\end{aligned}$$

where we have used the notation

$$\Phi(r) = \varphi^2(r).$$

Since the field f_1 is constant, it has zero Frechet derivative. Thus the first Lie bracket is easily computed as

$$[f_0, f_1](r) = \mathbf{F}\lambda + \Phi'(r)[\lambda]D.$$

The next bracket to compute is $[[f_0, f_1], f_1]$ which is given by

$$[[f_0, f_1], f_1] = \Phi''(r)[\lambda;\lambda]D.$$

Note that $\Phi''(r)[\lambda;\lambda]$ is the second order Frechet derivative of Φ operating on the vector pair $[\lambda;\lambda]$. This pair is to be distinguished from (notice the semicolon) the Lie bracket $[\lambda,\lambda]$ (with a comma), which if course would be equal to zero. We now make a further assumption.

Assumption 6.2 *We assume that* $\Phi''(r)[\lambda;\lambda] \neq 0$ *for all* $r \in \mathcal{H}$.

Given this assumption we may again use Lemma 3.1 to see that the Lie algebra is generated by the following vector fields

$$\begin{aligned} f_0(r) &= \mathbf{F}r, \\ f_1(r) &= \lambda, \\ f_3(r) &= \mathbf{F}\lambda, \\ f_4(r) &= D. \end{aligned}$$

Of these vector fields, all but f_0 are constant, so all brackets are easy. After elementary calculations we see that in fact

$$\{\mu, \sigma\}_{LA} = span\left\{\mathbf{F}r, \mathbf{F}^n\lambda,\ \mathbf{F}^n D;\ n = 0, 1, \ldots\right\}.$$

From this expression it follows immediately that a necessary condition for the Lie algebra to be finite dimensional is that the vector space spanned by $\{\mathbf{F}^n\lambda;\ n \geq 0\}$ is finite dimensional. This occurs if and only if λ is quasi-exponential (see Remark 5.1). If, on the other hand, λ is quasi-exponential, then we know from Lemma 5.1, that also D is quasi-exponential, since it is the integral of the QE function λ multiplied by the QE function λ. Thus the space $\{\mathbf{F}^n D;\ n = 0, 1, \ldots\}$ is also finite dimensional, and we have proved the following result.

Proposition 6.1 *Under Assumptions 6.1 and 6.2, the interest rate model with volatility given by $\sigma(r, x) = \varphi(r)\lambda(x)$ has a finite dimensional realization if and only if λ is a quasi-exponential function. The scalar field φ is allowed to be any smooth field.*

Remark 6.1 The case when assumption 6.2 is not in force is interesting when one is looking for a two-dimensional realization, and we will come back to this case in Section 7.

6.2 The Time Varying Case

For the time varying case of the model above we may apply Theorem 3.4 and, using the using the same technique as above, the following result can easily be derived. The proof is very similar to the one above, and therefore omitted.

Proposition 6.2 *Consider a volatility of the form*

$$\sigma(r, t, x) = \varphi(r, t)\lambda(t, x),$$

and define D, Φ by $D(t, x) = \lambda(t, x)\int_0^x \lambda(t, s)ds$ and $\Phi(r, t) = \varphi^2(r, t)$. Assume that $\Phi''_r(r, t)[\lambda, \lambda] \neq 0$, where Φ''_r denotes the second Frechet derivative w.r.t. r. Then the model has a finite dimensional realization if and only if, for each t, the function space

$$\left\{\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)^n \lambda,\ \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)^n D;\ n \geq 0\right\}$$

has finite dimension.

In the theorem above we view, for each t, the function space above as a space of functions of the single variable x. Thus there exists a finite dimensional realization if and only if there exist scalar functions $\{\alpha_i(t); i = 1, \ldots, N\}$ and $\{\beta_j(t); j = 1, \ldots, M\}$ with $\alpha_N = \beta_M = 1$, such that the PDE:s

$$\sum_{i=0}^{N} \alpha_i(t) \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x} \right)^i \lambda(t, x) = 0,$$
$$\sum_{j=0}^{M} \beta_j(t) \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x} \right)^j D(t, x) = 0.$$

are satisfied for all t and x.

Sometimes the models above are defined in terms of time **of** maturity T, instead of time **to** maturity x, and it may be convenient to have the result above in terms this (HJM-type) parameterization. By definition we have $x = T - t$, so we may therefore define the objects λ_0 and D_0 by

$$\begin{aligned} \lambda_0(t, T) &= \lambda(t, T - t), \\ D_0(t, T) &= D(t, T - t). \end{aligned}$$

It is then easily seen that there exists a finite dimensional realization if and only if there exist scalar functions $\{\alpha_i(t); i = 1, \ldots, N\}$ and $\{\beta_j(t); j = 1, \ldots, M\}$ with $\alpha_N = \beta_M = 1$ such that the ODE:s

$$\sum_{i=0}^{N} \alpha_i(t) \frac{\partial^i}{\partial t^i} \lambda_0(t, T) = 0,$$
$$\sum_{j=0}^{M} \beta_j(t) \frac{\partial^i}{\partial t^i} D_0(t, T) = 0.$$

are satisfied for all t and T.

6.3 Several Driving Wiener Processes

We now go back to the multidimensional model with m driving Wiener processes and volatilities of the form (31). With notation parallel to that of the scalar case, the drift field of the forward rate equation is given by

$$\mu(r) = \mathbf{F}r + \sum_{i=1}^{m} \varphi_i^2(r) D_i - \frac{1}{2} \sum_{i=1}^{m} \varphi_i'(r)[\lambda_i] \varphi_i(r) \lambda_i, \tag{33}$$

where

$$D_i(x) = \lambda_i(x) \int_0^x \lambda_i(s) ds, \quad i = 1, \ldots, m.$$

We now have to study the dimension of the Lie algebra generated by

$$\begin{aligned}
\mu(r) &= \mathbf{F}r + \sum_{i=1}^{m} \varphi_i^2(r) D_i - \frac{1}{2} \sum_{i=1}^{m} \varphi_i'(r)[\lambda_i]\varphi_i(r)\lambda_i, \\
\sigma_1(r) &= \varphi_1(r)\lambda_1, \\
&\vdots \\
\sigma_m(r) &= \varphi_m(r)\lambda_m.
\end{aligned}$$

Under Assumption 6.1 we can again use Lemma 3.1, to see that the Lie algebra is in fact generated by the much simpler system of vector fields

$$\begin{aligned}
f_0(r) &= \mathbf{F}r + \sum_{i=1}^{m} \Phi_i(r) D_i, \\
f_1(r) &= \lambda_1, \\
&\vdots \\
f_m(r) &= \lambda_m,
\end{aligned}$$

where we have used the notation

$$\Phi_i(r) = \varphi_i^2(r).$$

In this case, the structure of the Lie algebra is no longer as simple as in the case of a scalar Wiener process, so we will restrict ourselves to giving sufficient conditions for finite dimensionality. Such conditions are, on the other hand, very easy to provide. From the set of generating vector fields given above, it is obvious that the Lie algebra $\{\mu, \sigma\}_{LA}$ is included in the algebra $\{\mathbf{F}r, D_1, \ldots, D_m, \lambda_1, \ldots, \lambda_m\}_{LA}$. For this Lie algebra we have, however, the following trivial relation

$$\begin{aligned}
& \{\mathbf{F}r, D_1, \ldots, D_m, \lambda_1, \ldots, \lambda_m\}_{LA} \\
= \;& span\left\{\mathbf{F}r, \mathbf{F}^n D_i, \mathbf{F}^n \lambda_i;\ i = 1, \ldots, m;\ n \geq 0\right\}.
\end{aligned}$$

Arguing as in the proof of Proposition 6.1 we have the following sufficient conditions.

Proposition 6.3 *Under Assumption 6.1, a sufficient condition for the volatility structure $\sigma_i(r, x) = \varphi_i(r)\lambda_i(x)$ to have a finite dimensional realization, is that all the functions $\lambda_1, \ldots, \lambda_m$ are quasi-exponential. The functions $\varphi_1, \ldots, \varphi_m$ are allowed to be arbitrary smooth scalar fields on $\mathcal{H}$.*

Proposition 6.3 can quite easily be extended to the case when each volatility component σ_i is a finite linear combination of QE functions.

Proposition 6.4 *Assume that the volatility has the structure*

$$\sigma_i(r,x) = \sum_{j=1}^{N} \varphi_{ij}(r)\lambda_j(x), \quad i = 1, \dots m.$$

A sufficient condition for the existence of a finite dimensional realization, is that all the functions λ_j are quasi-exponential. The functions φ_{ij} are allowed to be arbitrary smooth scalar fields on $\mathcal{H}$.

Proof. Arguing as in the proof of the Proposition 6.3, we see that $\{\mu, \sigma\}_{LA}$ is included in

$$span\left\{\mathbf{F}r, \mathbf{F}^n\lambda_j, \mathbf{F}^n D_{jk};\ j,k = 1, \dots N;\ n \geq 0\right\},$$

where

$$D_{jk}(x) = \lambda_j(x)\int_0^x \lambda_k(s)ds.$$

Using Lemma 3.1 we see that this function space is finite dimensional. ∎

6.4 The Time Varying Case

As in Section 6.2 we can easily handle the time varying case.

Proposition 6.5 *Consider volatilities of the form*

$$\sigma_i(r,t,x) = \varphi_i(r,t)\lambda_i(t,x),\ \ i = 1, \dots, m.$$

and define D_i, Φ_i by $D_i(t,x) = \lambda_i(t,x)\int_0^x \lambda(t,s)ds$ and $\Phi_i(r,t) = \varphi_i^2(r,t)$. A sufficient condition for the existence of a finite dimensional realization is that, for each t, the function space

$$\left\{\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)^n \lambda_i,\ \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)^n D_i;\ n \geq 0,\ i = 1, \dots, m.\right\}$$

has finite dimension.

This result can obviously be reformulated in terms of T instead of in terms of x along the lines of Section 6.2.

7 Short Rate Realizations

One of the classical problems concerning the HJM approach to interest rate modeling is that of determining when a given forward rate model is realized by a short rate model, i.e. when the short rate induced by the forward rate model is a one-dimensional Markov process. There are several results in this area providing partial answers (see references in the introduction). Note that we, as usual, mainly restrict ourselves to the case of time homogenous forward rate volatilities.

7.1 General Theory

Using our previous results, we immediately have the following general necessary condition.

Proposition 7.1 *The forward rate model generated by $\sigma(r,x)$ is a generic short rate model, i.e the short rate is generically a Markov process, only if*

$$dim\left\{\mu,\sigma\right\}_{LA} \leq 2 \tag{34}$$

Proof. Suppose that the model is really a short rate model and assume for notational simplicity that we have a scalar driving Wiener process, so that the short rate R satisfies a scalar SDE of the form

$$dR = a(t,R_t)dt + b(R_t)dW_t.$$

Note that, since our r-model is time invariant, the volatility b must also be time invariant. Then bond prices are given as $p(t,T) = F^T(t,R_t)$ where F^T solves the term structure PDE

$$\begin{aligned} \frac{\partial F^T}{\partial t}(t,R) + a(t,R)\frac{\partial F^T}{\partial R}(t,R) + \frac{1}{2}b^2(R)\frac{\partial^2 F^T}{\partial R^2}(t,R) - RF^T(t,R) &= 0, \\ F^T(T,R) &= 1 \end{aligned}$$

Thus we have

$$r(t,x) = G(t,R_t,x),$$

where

$$G(t,R,x) = -\frac{\partial}{\partial x}\ln F^{t+x}(t,R).$$

Hence we see that the forward rate process is generated by a two dimensional state space model with time t and the short rate R as the states. ∎

Remark 7.1 The natural case is of course $dim\{\mu,\sigma\}_{LA} = 2$, since we are basically looking for a short rate realization where the coefficients are time dependent (in order to make it possible to fit any initial forward rate curve). It is presently an open question whether there exists a nontrivial generic short rate model with $dim\{\mu,\sigma\}_{LA} = 1$.

Let us start by observing that, in this context, the case of a scalar Wiener process is the only non pathological case.

Lemma 7.1 *Assume that there are m driving Wiener processes, and that the volatilities are of the general form $\sigma_1(r,x),\ldots,\sigma_m(r,x)$. Then, a necessary condition for the Lie algebra $\{\mu,\sigma\}_{LA}$ to be two dimensional is that there exists a vector field $\sigma_0(r,x)$ and scalar fields $\varphi_i(r),\ i=1,\ldots,m$ such that*

$$\sigma_i(r,x) = \varphi_i(r)\sigma_0(r,x)$$

Proof. If the Lie algebra is two-dimensional, then the fields $\sigma_1,\ldots,\sigma_m$ must be parallel at each point r, otherwise $\{\mu(r),\sigma_1(r),\ldots,\sigma_m(r)\}$ will span a subspace of $\mathcal{H}$ of dimension greater than two. ∎

To avoid degenerate cases like this we make a standing assumption.

Assumption 7.1 *We assume that we have only one scalar driving Wiener process, i.e. that $m = 1$.*

Note that condition (34) is only a sufficient condition for the existence of a short rate realization. It guarantees that there exists a two-dimensional realization, but the question remains whether the realization can chosen in such a way that the short rate and running time are the state variables. This question is completely resolved by the following central result.

Theorem 7.1 *Assume that the model is not deterministic, and take as given a time invariant volatility $\sigma(r,x)$. Then there exists a short rate realization if and only if the vector fields $[\mu,\sigma]$ and σ are parallel, i.e. if and only if there exists a scalar field $\alpha(r)$ such that the following relation holds (locally) for all r.*

$$[\mu,\sigma](r) = \alpha(r)\sigma(r). \tag{35}$$

This can also be written

$$[\mathbf{F}r + D,\sigma](r) - \frac{1}{2}\sigma''(r)[\sigma(r);\sigma(r)] = \alpha(r)\sigma(r). \tag{36}$$

where

$$D(r,x) = \sigma(r,x)\int_0^x \sigma(r,s)ds.$$

If the volatility is time varying of the form $\sigma(r,t,x)$, then (35) is replaced by

$$\mu_r(r,t)[\sigma(r,t)] - \sigma_r(r,t)[\mu(r,t)] - \sigma_t(r,t) = \alpha(r,t)\sigma(r,t), \tag{37}$$

where subindex denotes the Frechet derivative w.r.t. the indicated variable.

Remark 7.2 We note that the necessary and sufficient condition (35) is very explicit, in the sense that it can easily be checked for any concrete specification of the volatility structure. The only thing required is that we calculate some derivatives.

Proof. We start by noting that (35) obviously implies that the Lie algebra is two dimensional. The hard part to prove is that (35) is necessary and sufficient. The other results then follow by direct calculations and an application of Theorem 3.4.

The proof consists of two parts. We begin by showing that (35) is necessary and sufficient for a two dimensional realization where running time is one of the state variables. Then we show that (35) is also sufficient for us to be able to choose the short rate as the other state variable.

We begin the first part by showing necessity. Assume thus that there exists a two dimensional realization where running time is one of the state variables. Then it must have the form

$$\begin{aligned} dz_1 &= 1 \cdot dt + 0 \circ dW, \\ dz_2 &= a(z)dt + b(z) \circ dW, \\ r &= G(z). \end{aligned}$$

In vector notation this reads

$$dz = A(z)dt + B(z) \circ dW,$$

where the vector fields A and B are given by

$$A(z) = \begin{bmatrix} 1 \\ a(z) \end{bmatrix}, \quad B(z) = \begin{bmatrix} 0 \\ b(z) \end{bmatrix}.$$

The Frechet derivatives (Jacobians) are easily obtained as follows, where subscript denotes partial derivatives.

$$A'(z) = \begin{bmatrix} 0 & 0 \\ a_1(z) & a_2(z) \end{bmatrix}, \quad B'(z) = \begin{bmatrix} 0 & 0 \\ b_1(z) & b_2(z) \end{bmatrix}.$$

Thus the Lie bracket is given by

$$[A,B] = A'B - B'A = \begin{bmatrix} 0 \\ a_2 b_1 - b_2 a_1 - b_1 \end{bmatrix}.$$

It is now clear by inspection that we have $[A,B]//B$, where // denotes parallel. On the other hand, because of the relation $r = G(z)$, we also have $\mu = G_\star A$ and $\sigma = G_\star B$. Using the fact that the bracket is preserved under smooth mappings we obtain

$$[\mu,\sigma] = [G_\star A, G_\star B] = G_\star [A,B] .$$

Since $[A,B]//B$ we thus have $G_\star [A,B]//G_\star B$, but $G_\star B = \sigma$ and we are finished.

We now go on to prove sufficiency, and thus assume that (35) actually holds. The Lie algebra is then two-dimensional so there will exist a two dimensional realization of the form

$$\begin{aligned} dz_1 &= a^1(z)dt + b^1(z) \circ dW, \\ dz_2 &= a^2(z)dt + b^2(z) \circ dW, \\ r &= G(z). \end{aligned}$$

Furthermore, it follows from the first part (the case $n = 1$) of the proof of the Frobenius Theorem that, after a diffeomorphism, the realization can be written on the form

$$\begin{aligned} dz_1 &= a^1(z)dt + 0 \circ dW, \\ dz_2 &= a^2(z)dt + 1 \circ dW. \end{aligned}$$

We now have to show that we can find a diffeomorphism $H : R^2 \to R^2$ such that

$$H' \begin{bmatrix} a^1 \\ a^2 \end{bmatrix} = \begin{bmatrix} 1 \\ c \end{bmatrix}, \tag{38}$$

and

$$H' \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ d \end{bmatrix}, \tag{39}$$

for some functions c and d. Writing

$$H = \begin{bmatrix} H^1 \\ H^2 \end{bmatrix},$$

we have, with subscripts denoting partial derivatives,

$$H' = \begin{bmatrix} H^1_1 & H^1_2 \\ H^2_1 & H^2_2 \end{bmatrix},$$

so (38)-(39) become

$$\begin{bmatrix} H^1_2 \\ H^2_2 \end{bmatrix} = \begin{bmatrix} 0 \\ c \end{bmatrix}, \quad \begin{bmatrix} H^1_1 a^1 + H^1_2 a^2 \\ H^1_2 a^1 + H^2_2 a^2 \end{bmatrix} = \begin{bmatrix} 1 \\ d \end{bmatrix}$$

which can be written

$$\begin{bmatrix} H^1_2 \\ H^2_2 \end{bmatrix} = \begin{bmatrix} 0 \\ c \end{bmatrix}, \quad \begin{bmatrix} H^1_1 a^1 \\ H^1_2 a^1 + H^2_2 a^2 \end{bmatrix} = \begin{bmatrix} 1 \\ d \end{bmatrix}$$

The problem is thus reduced to that of finding a function $h = H^1(z_1, z_2)$ such that

$$\frac{\partial h}{\partial z_2}(z) = 0, \quad \frac{\partial h}{\partial z_1}(z)a^1(z) = 1. \tag{40}$$

For an arbitrarily chosen function a^1 this is of course not possible. We can, however, now use our assumption that $[\mu, \sigma]//\sigma$. Using the functorial property of the Lie bracket w.r.t. diffeomorphisms, the parallel condition can be transferred to the local coordinate system (i.e. to our realization above) so that, with obvious notation, $[A, B]//B$. Written out in detail this means that

$$\begin{bmatrix} a_2^1 \\ a_2^2 \end{bmatrix} \; // \; \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

so we see that a^1 is a function of z_1 only, and (40) becomes

$$\frac{\partial h}{\partial z_2}(z_1, z_2) = 0, \quad \frac{\partial h}{\partial z_1}(z_1, z_2)a^1(z_1) = 1.$$

We can thus define our local coordinate transformation H by choosing $H^1(z_1, z_2) = h(z_1)$ as any primitive function to $1/a^1(z_1)$, and choosing H^2 arbitrarily. We note in passing that $a^1 \neq 0$ since otherwise the entire system is deterministic.

We have thus shown that if condition (35) is satisfied then there exists a realization of the form

$$\begin{aligned} dy &= a(y,t)dt + b(y,t) \circ dW \\ r_t(x) &= G\left(y_t, t, x\right), \end{aligned}$$

where y is a scalar process (corresponding to z_2 above), and it remains to show that we can choose y as the short rate R. In order to see this we use the realization to obtain $R_t = G(y_t, t, 0)$. This gives us the R-dynamics as

$$dR = \left(\frac{\partial G}{\partial t} + a\frac{\partial G}{\partial y}\right)dt + b\frac{\partial G}{\partial y} \circ dW,$$

from which we see that $\frac{\partial G}{\partial y}(y, t, 0) \neq 0$, since otherwise we would have a deterministic model. We can thus, for each t, locally invert the mapping $y \longmapsto G(y, t, 0)$. Denoting the inverse by $H(R, t)$ we have $y_t = H(R_t, t)$, and with this change of variable we have thus obtained a realization with R as the state variable. ∎

7.2 Deterministic Volatility

We now turn to applications of Theorem 7.1, and in this section we study the simplest non-trivial example, which is the case studied by Carverhill in [8]. We have m driving Wiener processes, and in our notation the volatilities are

assumed to be deterministic, i.e. (with a slight misuse of notation) they have the form

$$\sigma_i(t, r, x) = \sigma_i(t, x).$$

Adapting Lemma 7.1 to our present setting, we see that there must exist functions $g(t)$ and $\sigma_0(t, x)$ such that

$$\sigma_i(t, x) = g_i(t)\sigma_0(t, x), \ i = 1, \ldots, m.$$

This however, implies that we can replace the m driving Wiener processes with just one Wiener process so, as was expected, we may as well assume that $m = 1$. Denoting σ_1 by σ, we obtain

$$\mu(r, t, x) = \frac{\partial}{\partial x} r(x) + \sigma(t, x) \int_0^x \sigma(t, s) ds.$$

and we see that $\mu_r = \mathbf{F}r$ and $\sigma_r = 0$. Since we have a time varying system, we use the second part of Theorem 7.1 which says that there exists a short rate realization if and only if here exists a scalar field $\alpha(r, t)$, such that for all r, t, x we have

$$\sigma_x(t, x) - \sigma_t(t, x) = \alpha(r, t)\sigma(t, x).$$

From this it is clear that $\alpha(r, t) = \alpha(t)$. Dividing by σ and defining g by $g(t, x) = \ln \sigma(t, x)$ we have the equation

$$g_x(t, x) - g_t(t, x) = \alpha(t).$$

Taking x-derivatives, and defining h by $h = g_x$ we get

$$h_x(t, x) - h_t(t, x) = 0.$$

This is the simplest possible example of the wave equation, with solution $h(t, x) = \lambda(t + x)$ for some arbitrary function λ. Thus

$$g(t, x) = \int_0^x \lambda(t + s) ds + \beta(t),$$

for some function β and, going back all the steps to σ, we have the following result, which was proved in [8].

Proposition 7.2 *The deterministic volatility structure $\sigma(t, x)$ induces a Markovian short rate if and only if it can be written as*

$$\sigma(t, x) = c(t) \exp\left\{ \int_t^{t+x} \lambda(s) ds \right\}.$$

for some functions c and λ, where $c > 0$.

7.3 Constant Direction Volatility

We now turn to models with a constant direction volatility, and for simplicity we will carry out the discussion only for the case when there is no explicit time dependence. We thus assume that

$$\sigma(r, x) = \varphi(r)\lambda(x). \tag{41}$$

We recall from Section 6.1 that for this model μ is given by

$$\mu(r) = \mathbf{F}r + \varphi^2(r)D - \frac{1}{2}\varphi'(r)[\lambda]\varphi(r)\lambda.$$

Using Theorem 7.1, Lemma 3.1, and calculations already made in Section 6.1 we conclude that there exists a short rate realization if and only if there exists a scalar field $c(r)$ such that

$$\mathbf{F}\lambda + \Phi'(r)[\lambda]D + c(r)\lambda = 0.$$

for all r. This is a vector equation, and if we write it in component form it reads

$$\frac{d\lambda}{dx}(x) + \Phi'(r)[\lambda]D(x) + c(r)\lambda(x) = 0, \quad \forall (x, r) \in R_+ \times \mathcal{H}. \tag{42}$$

Setting $x = 0$, and recalling that $D(x) = 0$, we obtain

$$\frac{d\lambda}{dx}(0) + c(r)\lambda(0) = 0, \quad \forall (x, r) \in R_+ \times \mathcal{H}, \quad \forall r \in \mathcal{H}.$$

Assuming that $\lambda(0) \neq 0$, this implies that the scalar field c is in fact constant. Denoting this constant by c, (42) transforms into

$$\frac{d\lambda}{dx}(x) + \Phi'(r)[\lambda]D(x) + c\lambda(x) = 0, \quad \forall (x, r) \in R_+ \times \mathcal{H}.$$

From this equation it is clear that we must have

$$\Phi'(r)[\lambda] = d, \tag{43}$$

for some constant d. Plugging this into (42) we get

$$\frac{d\lambda}{dx}(x) + d \cdot D(x) + c\lambda(x) = 0, \quad \forall x \in R. \tag{44}$$

This is an ODE for λ, and we obviously should treat the cases $d = 0$ and $d \neq 0$ separately.

If $d = 0$, then (44) is the simple equation

$$\frac{d\lambda}{dx}(x) + c\lambda(x) = 0,$$

with solution

$$\lambda(x) = \lambda(0)e^{-cx}.$$

From (43), the condition $d = 0$ says that $\Phi'(r)[\lambda] = 0$ for all r. This in turns implies that $\Phi(r)$ is constant on all lines in $\mathcal{H}$, which are parallel to λ. Writing $\mathcal{H} = \{\lambda\} + \mathcal{Y}$, where $\mathcal{Y} = \{\lambda\}^{\perp}$, we thus see that Φ must be of the form

$$\Phi(y + s\lambda) = \Psi(y), s \in R,\ y \in \mathcal{Y},$$

for some smooth mapping $\Psi : \mathcal{Y} \to R$.

We now turn to the case $d \neq 0$. In this case we rewrite (44) by noting that

$$D(x) = \frac{1}{2}\frac{d}{dx}L^2(x),$$

where

$$L(x) = \int_0^x \lambda(s)ds. \tag{45}$$

Thus $\lambda = \frac{d}{dx}L$ and (44) becomes,

$$\frac{d^2L}{dx^2} + d\frac{d}{dx}L^2 + c\frac{dL}{dx} = 0.$$

Thus

$$\frac{dL}{dx}(x) + dL^2(x) + cL(x) + e = 0,$$

which is a Riccati equation for L.

To obtain information from condition (43), we may write it as

$$\Gamma'(r)[\lambda] = 0,$$

where $\Gamma(r) = \Phi(r) - (v, r)$, for any fixed vector $v \in \mathcal{H}$ such that $(v, \lambda) = d$. Here $(\cdot, \cdot)$ denotes then inner product in $\mathcal{H}$. Thus Γ is constant on all lines parallel to λ, and we can summarize our findings.

Proposition 7.3 *Assume that the volatility is of the form $\sigma(r, x) = \varphi(r)\lambda(x)$, and that $\lambda(0) \neq 0$. Then there exists a short rate realization if and only if one of the following two cases occur.*

1. *The vector λ is of the form*

$$\lambda(x) = \lambda_0 e^{-cx}, \tag{46}$$

for some constant c. Furthermore, the scalar field φ is of the form

$$\varphi(r, t) = \sqrt{\Psi(Pr)}, \tag{47}$$

where Ψ is an arbitrary nonlinear smooth map $\Psi : \{\lambda\}^{\perp} \to R$, and P is the orthogonal projection in $\mathcal{H}$ onto $\lambda^{\perp}$.

2. *Defining L as in (45), L satisfies a Riccati equation of the form*

$$\frac{dL}{dx}(x) + dL^2(x) + cL(x) + e = 0. \tag{48}$$

Furthermore φ has the form

$$\varphi(r) = \sqrt{(v, r) + \Psi(Pr)} \tag{49}$$

where Ψ is as above, and where v is any vector with $(v, \lambda) = d$.

Remark 7.3 The assumption $\lambda(0) \neq 0$ is very innocent, since it says that the dynamics of the short rate has nonzero diffusion term. The reader will probably notice the structural similarity with the results above and the class of affine short rate models. Also note that all our results are local, otherwise we would of course have problems with expressions like $\sqrt{(v, r) + \Psi(Pr)}$ where the linear functional can take arbitrarily large negative values.

The result above has an easy an interesting corollary, which will be used later.

Corollary 7.1 *Assume that the scalar field φ above is a point evaluation of the short rate, i.e. that (with a slight abuse of notation)*

$$\varphi(r) = \varphi(R).$$

Then the function Ψ in Proposition 7.3 must be constant w.r.t. r. Thus (47) and (49) are, for some choice of constants α and β, replaced by

$$\varphi(R) = \beta, \tag{50}$$

and

$$\varphi(R) = \sqrt{\alpha R + \beta}, \tag{51}$$

respectively.

Proof. In this case, with $\Phi(r) = \varphi^2(R)$, the condition (43) becomes

$$\frac{d\Phi(R)}{dR}\lambda(0) = d,$$

which, given the standing assumption $\lambda(0) \neq 0$, immediately give us the result. ∎

Remark 7.4 From Corollary 7.1 we recognize, by inspection of φ and λ, three well known short rate models.

- $\varphi(R) = \beta$ and $c = 0$. This is the Ho-Lee model.
- $\varphi(R) = \beta$ and $c \neq 0$. This is the Hull-White extension of the Vasiček model.
- $\varphi(R) = \sqrt{\alpha R + \beta}$. This is the Hull-White extension of the Cox-Ingersoll-Ross model.

These models are henceforth referred to as the **affine** short rate models. We say that a forward rate model which is generated (realized) by an affine short rate model is an **affine forward rate model**.

7.4 "All Generic Short Rate Models are Affine"

We now consider the perhaps most interesting special case, which is when the forward rate volatility $\sigma(r)$ only depends on the short rate R. We will use the techniques above in order to rederive a rather astonishing result from [19].

Assumption 7.2 *The volatility is of the form*

$$\sigma = \sigma(R, x),$$

where, with a slight abuse of notation, σ is a smooth function of the two real variables R and x.

The question is now when a volatility specification like this will possess a short rate realization and, as we will see, the requirement $[\mu, \sigma] \,//\, \sigma$ in Theorem 7.1 turns out to be very severe. We have in fact the following strong result.

Proposition 7.4 *Under Assumption 7.2, the there exists a short rate realization if and only if σ can be factored as*

$$\sigma(R, x) = \varphi(R)\lambda(x).$$

Proof. Writing out all variables, the drift vector field μ is easily calculated as

$$\mu(R, x) = \mathbf{F}r(x) + \sigma(R, x) \int_0^x \sigma(R, s)ds - \frac{1}{2}\sigma'_R(R, x)g(R),$$

where $g(R) = \sigma(R, 0)$ and where $\sigma'_R = \frac{\partial \sigma}{\partial R}$. We now have to compute the Lie bracket $[\mu, \sigma]$, and we obtain

$$\sigma'(r)[\mu(r)](x) = \sigma'_R(R, x)\mathbf{F}r(0) - \frac{1}{2}\sigma'_R(R, x)g'(R)g(R)$$

and

$$\begin{aligned}\mu'(r)[\sigma(r)](x) &= \sigma'_x(R,x) - \frac{1}{2}\left\{\sigma''_R(R,x)g^2(R) - \sigma'_R(R,x)g'(R)g(R)\right\} \\ &+ g(R)\left\{\sigma'_R(R,x)\int_0^x \sigma(R,s)ds + \sigma_R(R,x)\int_0^x \sigma'_R(R,s)ds\right\}.\end{aligned}$$

Thus

$$\begin{aligned}[\sigma,\mu](x) &= \sigma'_x(R,x) + g(R)\frac{\partial}{\partial R}\left(\sigma(R,x)\int_0^x \sigma(R,s)ds\right) \\ &- \frac{1}{2}\sigma''_R(R,x)g^2(R) - \sigma'_R(R,x)\mathbf{F}r(0)\end{aligned}$$

Applying Theorem 7.1, a necessary and sufficient for the existence of a short rate realization is that σ and $[\mu,\sigma]$ are parallel. Denoting $\mathbf{F}r(0)$ by z, there must thus exist a function $\alpha(z,R)$ such that the equation

$$\begin{aligned}&\sigma'_x(R,x) + g(R)\frac{\partial}{\partial R}\left(\sigma(R,x)\int_0^x \sigma(R,s)ds\right) \\ -\frac{1}{2}\sigma''_R(R,x)g^2(R) - \sigma'_R(R,x)z + \alpha(z,R)\sigma(R,x) &= 0\end{aligned}$$

holds for all R and z. Since this equation holds identically in z, we can take the z-derivative, to obtain the equation

$$\sigma'_R(R,x) = \alpha'_z(z,R)\sigma(R,x),$$

from which it follows that in fact $\alpha'_z(z,R) = f(R)$. We thus obtain

$$\sigma'_R(R,x) = f(R)\sigma(R,x),$$

which is an ODE for $\sigma(R,x)$ with the solution

$$\sigma(R,x) = \exp\left\{\int_0^R f(s)ds\right\}\sigma(0,x),$$

and we have the desired factorization. ∎

We may now state the main result of the section (see Remark 7.4). It was first stated and proved (using completely different techniques) in the remarkable paper [19].

Theorem 7.2 *Assume that the forward rate volatilities are of the form*

$$\sigma(r,x) = \sigma(R,x).$$

Then the model has a short rate realization if and only if it is affine.

Proof. This follows immediately from Proposition 7.4, Corollary 7.1, and the comments above. ∎

We have thus completely characterized the class of forward rate models with a volatility structure of the form $\sigma(R,x)$ which are generated by short rate models. Another way of formulating this somewhat surprising result is to say that "the only generic short rate models are the affine ones". To understand the geometric picture one can think of the following program.

1. Choose an arbitrary short rate model, say of the form

$$dR_t = a(t,R_t)dt + b(R_t)dW_t$$

   with a fixed initial point R_0.

2. Solve the associated PDE in order to compute bond prices. This will also produce:
   - An initial forward rate curve $\hat{r}^o(x)$.
   - Forward rate volatilities of the form $\sigma(R,x)$.
3. Forget about the underlying short rate model, and take the forward rate volatility structure $\sigma(R,x)$ as given in the forward rate equation (20).
4. Initiate the forward rate equation with an arbitrary initial forward rate curve $r^o(x)$

The question is now whether the thus constructed forward rate model will produce a Markovian short rate process. Obviously, if you choose the initial forward rate curve r^o as $r^o = \hat{r}^o$, then you are back where you started, and everything is OK. If, however, you choose another initial forward rate curve $\hat{r}^o$, say the observed forward rate curve of today, then it is no longer clear that the short rate will be Markovian. What the theorem above says, is that only the models listed above will produce a Markovian short rate model for all initial points in a neighborhood of r^o. If you take another model (like, say, the Dothan model) then a generic choice of the initial forward rate curve will produce a short rate process which is not Markovian.

References

[1] Bhar, R. & Chiarella, C. (1997) Transformation of Heath-Jarrow-Morton Models to Markovian systems. *European Journal of Finance,* **3**, No. 1, 1-26.

[2] Björk, T. & Christensen, B.J. (1997) Interest Rate Dynamics and Consistent Forward Rate Curves. To appear in *Mathematical Finance.*

[3] Björk, T. & Gombani A. (1999) Minimal Realization of Forward Rates. *Finance and Stochastics*, **3**, No. 4, 413-432.

[4] Björk, T. (1997). Interest Rate Theory. In W. Runggaldier (ed.), *Financial Mathematics.* Springer Lecture Notes in Mathematics, Vol. 1656. Springer Verlag, Berlin.

[5] Björk, T. (1998) *Arbitrage Theory in Continuous Time.* Oxford University Press.

[6] Brace, A. & Musiela M. (1994) A Multi-Factor Gauss Markov Implementation of Heath Jarrow and Morton. *Mathematical Finance* **4**, Vol.3, 563-576.

[7] Brockett, R.W. (1981) Nonlinear Systems and Nonlinear Estimation Theory. In *Stochastic systems: The Mathematics of Filtering and Identification and Applications* (eds. Hazewinkel, M & Willems, J.C.) Reidel.

[8] Carverhill, A. (1994) When is the Spot Rate Markovian? *Mathematical Finance*, , 305-312.

[9] Chiarella, C & Kwon, K. (1998) Forward Rate Dependent Markovian Transformations of the Heath-Jarrow-Morton Term Structure Model. Working paper. School of Finance and Economics, University of Technology, Sydney.

[10] Cox, J., Ingersoll, J. & Ross, S. (1985b). A Theory of the Term Structure of Interest Rates. *Econometrica* **53**, 385–408.

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

[12] Duffie, D. & Kan, R (1996) A yield Factor Model of Interest Rates. *Mathematical Finance*, **6**, 379-406.

[13] Eberlein, E. & Raible, S. (1999) Term Structure Models Driven by General Levy Processes. *Mathematical Finance,* **9**, No 1, 31-53.

[14] Heath, D. & Jarrow, R. & Morton, A. (1992) Bond Pricing and the Term Structure of Interest Rates. *Econometrica* **60** No.1, 77-106.

[15] Ho, T. & Lee, S. (1986) Term Structure Movements and Pricing Interest Rate Contingent Claims. *Journal of Finance* **41**, 1011-1029.

[16] Hull, J & White, A. (1990) Pricing Interest-Rate-Derivative Securities. *The Review of Financial Studies* **3**, 573-592.

[17] Inui, K. & Kijima, M. (1998) A Markovian Framework in Multi-Factor Heath-Jarrow-Morton Models. *JFQA* **33** no. 3, 423-440.

[18] Isidori, A. (1989) *Nonlinear Control Systems.* Springer-Verlag, Berlin.

[19] Jeffrey, A. (1995) Single Factor Heath-Jarrow-Morton Term Structure Models Based on Markovian Spot Interest Rates. *JFQA* **30** no.4, 619-642.

[20] Lang, S. (1995) *Differentiable Manifolds and Riemannian Geometry.* Springer-Verlag, Berlin.

[21] Musiela M. (1993) Stochastic PDE:s and Term Structure Models. Preprint.

[22] Nelson, C. & Siegel, A. (1987): Parsimonious Modeling of Yield Curves. *Journal of Business*, **60**, 473-489.

[23] Ritchken, P. & Sankarasubramanian, L. (1995) Volatility Structures of Forward Rates and the Dynamics of the Term Structure. *Mathematical Finance*, **5**, no. 1, 55-72.

[24] Warner, F.W. (1979) *Foundations of Differentiable Manifolds and Lie Groups.* Scott, Foresman, Hill.

[25] Vasiček, O. (1977) An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics* **5**, 177-188.