ON THE GEOMETRY OF THE TERM STRUCTURE OF INTEREST RATES

DAMIR FILIPOVIĆ AND JOSEF TEICHMANN

Abstract. We present recently developed geometric methods for the analysis of finite dimensional term structure models of the interest rates. This includes an extension of the Frobenius theorem for Fréchet spaces in particular. This approach puts new light on many of the classical models, such as the Hull-White extended Vasicek and Cox-Ingersoll-Ross short rate models. The notion of a finite dimensional realization (FDR) is central for this analysis: we motivate it, classify all generic FDRs and provide some new results for the corresponding factor processes, such as hypoellipticity of its generators and the existence of smooth densities. Furthermore we include finite dimensional external factors, thus admitting a stochastic volatility structure.

1. Introduction

¿From the point of view of mathematics the present article treats a stochastic invariance problem which has been motivated by mathematical finance. In general we consider a stochastic equation

$$\begin{cases} dh_t = (Ah_t + \Theta(h_t))\,dt + \sum_{j=1}^{d} \Sigma_j(h_t)\,dW_t^j \\ h_0 \in \mathcal{U} \end{cases} \tag{1.1}$$

on some convex open subset $\mathcal{U}$ in a separable Hilbert space $\mathcal{H}$, in the spirit of Da Prato and Zabczyk [6]. The operator $A : D(A) \subset \mathcal{H} \to \mathcal{H}$ generates a strongly continuous semigroup $(T_t)_{t\geq 0}$ on $\mathcal{H}$. Here $d \in \mathbb{N}$, and $W = (W^1, \ldots, W^d)$ denotes a standard d-dimensional Brownian motion defined on a fixed reference probability space $(\Omega, \mathcal{F}, \mathbb{P})$ (see [6]). The mappings

$$\Theta : \mathcal{U} \subset \mathcal{H} \to \mathcal{H} \quad \text{and} \quad \Sigma = (\Sigma_1, \ldots, \Sigma_d) : \mathcal{U} \subset \mathcal{H} \to \mathcal{H}^d$$

satisfy appropriate regularity conditions (e.g. are smooth, which means C^∞) on $\mathcal{U}$. We distinguish, in decreasing order of generality, between (local) mild, weak and strong solutions of equation (1.1). The reader is referred to [6] or [9] for the precise definitions. We provide necessary and sufficient conditions for the existence of finite dimensional invariant manifolds for (1.1). Since A is an unbounded operator in general, this requires an extension of the classical Frobenius theorem for Fréchet spaces.

Date: March 19, 2002 (first draft); April 2, 2003 (this draft).

We thank T. Björk, A. Kriegl, P. Malliavin, P. Michor, W. Schachermayer, L. Svensson and J. Zabczyk for interesting and helpful discussions during the work on these topics. The second author is grateful for the distinguished hospitality of Ceremade, Université Paris Dauphine, where the work on this article began, in particular Ivar Ekeland, Elyes Jouini and Huyen Pham are mentioned.

¿From the point of view of financial mathematics we provide the characterization of all finite dimensional Heath–Jarrow–Morton (henceforth HJM) [17] interest rate models which admit arbitrary initial yield curves. This is an extension and completion of a series of results obtained by Björk et al. [1, 2, 4, 3] and [9, 11, 12, 35]. It is well known that affine term structure models with time-dependent coefficients (such as the Hull–White extension of the Vasicek short rate model [21]) perfectly fit any initial term structure. Under reasonable assumptions on the volatility structure, we find that such affine models are in fact the only finite-factor term structure models with the aforementioned property. We also show that there is usually an invariant singular set of initial yield curves where the affine term structure model becomes time-homogeneous. This is again well known for the classical Vasicek [34] and Cox–Ingersoll–Ross (CIR) [5] short rate models, where the set of consistent inital curves is given explicitely by the model parameters.

Below we shall say a bit more about the motivation for finite dimensional realizations and give the general setup for the corresponding stochastic invariance problem.

1.1. **Finite Dimensional Term Structure Models.** An HJM model for the forward curve, $x \mapsto r_t(x)$, is determined by the volatility structure and the market price of risk (see [17]). Here $r_t(x)$ denotes the forward rate at time t for date $t+x$ (this is the Musiela [28] parameterization). That is, the price at time t of a zero-coupon bond maturing at date $T \geq t$ is given by

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

It is shown in [9] that essentially every (classical) HJM model can be realized as a stochastic equation

$$\begin{cases} dr_t = \left(\dfrac{d}{dx} r_t + \alpha_{HJM}(r_t) \right) dt + \displaystyle\sum_{j=1}^{d} \sigma_j(r_t)\, dW_t^j \\ r_0 \in \mathcal{U} \end{cases} \tag{1.2}$$

on some open convex subset $\mathcal{U}$ in a Hilbert space H of forward curves ($\mathcal{U}$ is for example the halfspace $\{r \in H \mid r(0) > 0\}$). We will enlarge the model (1.2) by adding an external m-dimensional factor process ($m \in \mathbb{N}$), admitting a stochastic volatility structure. That is, we let $b, c_1, \dots, c_d : \mathcal{U} \times \mathbb{R}^m \to \mathbb{R}^m$ be smooth vector fields and

$$\begin{cases} dr_t = \left(\dfrac{d}{dx} r_t + \alpha_{HJM}(r_t, Y_t) \right) dt + \displaystyle\sum_{j=1}^{d} \sigma_j(r_t, Y_t)\, dW_t^j \\ dY_t = b(r_t, Y_t) dt + \displaystyle\sum_{j=1}^{d} c_j(r_t, Y_t)\, dW_t^j \\ r_0 \in \mathcal{U}, \quad Y_0 \in \mathbb{R}^m. \end{cases} \tag{1.3}$$

This extension has recently been introduced and studied by Björk et al. [3] (there however Y was Markov, that is, b and c_j were only Y-dependent).

Equation (1.3) is obviously of the form (1.1) with $\mathcal{H} = H \times \mathbb{R}^m$ and $\mathcal{U}$ replaced by $\mathcal{U} \times \mathbb{R}^m$ (the precise setup is given below in Subsection 1.3). The process Y can stand for an abstract latent factor but also for an observable quantity such as (the logarithm of) an index or stock price, or a combination of such. In any case we require that (1.3) is an arbitrage-free model and that $\mathbb{P}$ already is the risk neutral measure. This means that the HJM drift condition holds

$$\alpha_{HJM}(r,Y,x)=\sum_{j=1}^{d}\sigma_j(r,Y,x)\int_0^x\sigma_j(r,Y,\eta)\,d\eta,$$

and if Y^i is the log-price process of a tradable asset for some i then necessarily (since the return of S has to the risk-free rate)

$$b_i(r,Y)=r(0)-\frac{1}{2}\sum_{j=1}^{d}c_{ji}^2(r,Y). \tag{1.4}$$

In general, the solution process r of (1.3) cannot be realized by a finite dimensional Markov state process. An HJM is said to admit a *finite dimensional realization (FDR)* at the initial forward curve r_0 if, roughly speaking, there exists an n-dimensional diffusion state process Z and a map $\phi:\mathbb{R}^n\to H$ such that $r_t=\phi(Z_t)$. Notice that n, Z and ϕ may depend on r_0 and Y_0. We will investigate those HJM models that admit a generic FDR, that is, an FDR of the same dimension at every initial state (r_0,Y_0) in an open set of $\mathcal{H}$.

Practitioners and academics alike have a vital interest in finite-factor term structure models, and the distinction of time-homogenous and inhomogeneous ones. According to [18] there are two groups of practitioners in the fixed income market.

Fund managers trade on the yield curve (buy and sell swaps at different maturities), trying to make money out of it. They do not believe that all the interest rate market quotes are "correct". Instead, they in general use a time-homogeneous two- or three-factor model, estimate the model parameters from long time series data, and then update the state variables (factors) each day to fit the current term structure. Hence the term structure is considered as a derivative based on more fundamental state variables (factors), such as in an equilibrium model. The discrepancies between the fitted term structure and the market prices are perceived as potential trading opportunities. For example, if the fitted curve is above the two year and ten year swap rates, but is below the five year swap rate. Then one does a butterfly trade: receiving the five year rate (as one thinks it is high) and delivering the two year and ten year rates (as one thinks they are low compared to the five year rate). After this trade, one usually needs to wait for six months or longer for the rates to "reverse" (as predicted by the model) so that one can make money. Since this is a long term game, the model parameters must not change every day. Parameters have to be constant. If a parameter is time-varying, it is a factor and one needs to specify its dynamics so that one can make corresponding adjustments for the hedging. A state variable (factor) is time-varying, but since one has a stochastic model for its evolution, one can check on a daily basis whether its realized value lies within a statistical confidence interval or not.

Interest rate option traders, on the other hand, often take the quoted yield curve data, with minimal or no smoothing, as model input. To fit the observed yield curve perfectly, they allow some of the model parameters to be time-inhomogeneous. They intend to hedge away instantly all the risks on the yield curve and only worry about the risk in the implied volatility structure. Yet, low-dimensionality of the model is desirable, since the number of factors usually equals the number of instruments one needs to hedge in the model. And the daily adjustment of a large number of instruments becomes infeasible in practice due to transaction costs. Of course, the model factors have to represent tradable values. This can usually be achieved by a coordinate transformation.

Given the above considerations there are three main points which speak for a deeper analysis of finite dimensional structures.

Consistency: A curve-fitting procedure should be consistent with an arbitrage-free stochastic model, that is, the model output curves should be of the curve-fitting type. Only such models can give a reasonable framework for the statistical comparison of the curve-fitting data over time.

Model calibration: Finite dimensional models with identifiable factors are inevitable for model calibration. Hence, given an arbitrary initial curve, the possible finite factor models evolving from this curve should be completely understood.

Analytical and computational tractability: For the purpose of calculating derivatives prices, the stochastic characteristics of the factor processes have to be known.

1.2. **Stochastic Invariance Problems ...** We go back to the general setup (1.1) and recall

Definition 1.1. *A subset K of $\mathcal{U}$ is called* locally invariant *for (1.1) if, for every initial point $h_0 \in K$, there exists a continuous local weak solution h to (1.1) with strictly positive lifetime $\tau > 0$ such that $h_{t\wedge\tau} \in K$, for all $t \geq 0$.*

We are going to address the problem: given A, Θ and Σ, do there exist locally invariant submanifolds with boundary $\mathcal{M}$ of $\mathcal{U}$ for arbitrary initial values $h_0 \in \mathcal{U}$? For the notion of a (smooth) finite dimensional submanifold with boundary $\mathcal{M}$ of a Hilbert space (Fréchet space) and its tangent spaces $T_h\mathcal{M}$, $h \in \mathcal{M}$, we refer to [14] or to Section 3. Submanifolds *with boundary* appear naturally since the vector field Ξ in (1.5) below only generates a *semi*flow (see the Frobenius Theorem 3.14 below), so one direction is distinguished.

Finite dimensional locally invariant submanifolds (without boundary) have been characterized in [11], see also [9, 12]. These results carry over to submanifolds with boundary. For example [9, Theorem 6.2.3]:

Theorem 1.2. *Suppose that Θ is locally Lipschitz continuous and locally bounded, and Σ is C^1. Let $\mathcal{M}$ be an n-dimensional submanifold with boundary of $\mathcal{U}$. Then the following conditions are equivalent:*

i) *$\mathcal{M}$ is locally invariant for (1.1)*
ii) *$\mathcal{M} \subset D(A)$ and*

$$\Xi(h) := Ah + \Theta(h) - \frac{1}{2}\sum_{j=1}^{d} D\Sigma_j(h)\Sigma_j(h) \in T_h\mathcal{M} \tag{1.5}$$

$$\Sigma_j(h) \in T_h\mathcal{M}, \quad j = 1, \dots, d, \tag{1.6}$$

for all $h \in \mathcal{M}$, where $\Xi(h)$ is inward pointing and the $\Sigma_j(h)$ are parallel to the boundary for $h \in \partial\mathcal{M}$.

The proof is essentially the same as for [9, Theorem 6.2.3]. Indeed, the only geometric difference is that now the local coordinates of $\mathcal{M}$ vary in open sets of the half-space

$$\mathbb{R}^n_{\geq 0} := \{z \in \mathbb{R}^n \mid z_n \geq 0\}.$$

The coordinate process of h is thus a diffusion

$$dZ_t = \beta(Z_t)\,dt + \sum_{j=1}^{d} \rho_j(Z_t)\,dW_t^j$$

in some open $V \subset \mathbb{R}^n_{\geq 0}$ (see [9, p. 109]). A stochastic viability result in [26] yields that $\beta(z) \in \mathbb{R}^n_{\geq 0}$ (inward pointing) and $\rho_{j,n}(z) = 0$ (parallel) for all $z \in \partial V$. Whence the last statement in Theorem 1.2.

An FDR is essentially equivalent to a finite dimensional invariant submanifold with boundary in the following sense. If $\phi : \mathbb{R}^n \to \mathcal{U}$ is an FDR for (1.1) at some $h_0 \in \mathcal{U}$, then there exists an open neighborhood V_0 of $\phi^{-1}(h_0)$ in $\mathbb{R}^n_{\geq 0}$ such that $\phi(V_0)$ is an n-dimensional submanifold with boundary of $\mathcal{U}$, which is locally invariant for (1.1). The converse is given by the following result, which is a straightforward modification of [9, Theorem 6.4.1].

Theorem 1.3. *Let Θ, Σ and $\mathcal{M}$ be as in Theorem 1.2. Suppose $\mathcal{M}$ is locally invariant for (1.1). Then, for any $h_0 \in \mathcal{M}$, there exists an FDR $\phi : V \subset \mathbb{R}^n_{\geq 0} \to \mathcal{H}$ for (1.1) at h_0 such that $\phi(V) = \mathcal{V} \cap \mathcal{M}$, where $\mathcal{V}$ is an open set in $\mathcal{H}$.*

Hence the stochastic (local) invariance problem for (1.1) is equivalent to the deterministic invariance problems related to the vector fields $\Xi, \Sigma_1, \dots, \Sigma_d$.

Theorem 1.2 provides conditions for the local invariance of a given submanifold with boundary $\mathcal{M}$. However, it does not say anything about the *existence* of an FDR for (1.1). Conclusions on the existence can be drawn by Frobenius type theorems, which allow to answer the question whether there is a submanifold tangent to a given set of vector fields in a space. We have to face the problem that the vector field Ξ is neither continuous nor everywhere defined on $\mathcal{H}$. This fundamental problem has to be taken fully into account (compare [4]) to solve this problem completely. This means that we have to find a better space of definition for the geometric problem without loosing solutions of it. The Fréchet space

$$D(A^\infty) := \cap_{n\in\mathbb{N}} D(A^n)$$

fulfills these requirements. It is explained in Section 2 why all FDRs can be found in $D(A^\infty)$ and in Section 3 why geometric questions can be solved thereon.

1.3. ... Applied to HJM Models.

In view of the preceding subsection it is now clear how we tackle the issue of finite dimensional HJM interest rate models.

The rigorous setup for (1.3), extending [9], is given by the following structure. The Hilbert space $\mathcal{H}$ is axiomatically characterized by the properties

- **(H0):** $\mathcal{H} = H \times \mathbb{R}^m$, where m is the dimension of the external factor influencing the interest rate market.
- **(H1):** H is a separable Hilbert space continuously embedded in $C(\mathbb{R}_{\geq 0}; \mathbb{R})$ (that is, for every $x \in \mathbb{R}_{\geq 0}$, the pointwise evaluation $\mathrm{ev}_x : r \mapsto r(x)$ is a continuous linear functional on H), and $1 \in H$ (the constant function 1).
- **(H2):** The family of right-shifts $S_t r = r(t + \cdot)$, for $t \in \mathbb{R}_{\geq 0}$, forms a strongly continuous semigroup S on H with generator denoted by d/dx.

(H3): There exists a closed subspace H_0 of H such that

$$M(r_1, r_2)(x) := r_1(x) \int_0^x r_2(\eta)\, d\eta,$$

defines a continuous bilinear mapping $M : H_0 \times H_0 \to H$.

Equation (1.3) is obviously of the form (1.1) with coefficients

$$A = \begin{pmatrix} \frac{d}{dx} & 0 \\ 0 & 0 \end{pmatrix}, \quad \Theta(r, Y) = \begin{pmatrix} \alpha_{HJM}(r, Y) \\ b(r, Y) \end{pmatrix}, \quad \Sigma_j(r, Y) = \begin{pmatrix} \sigma_j(r, Y) \\ c_j(r, Y) \end{pmatrix}, \tag{1.7}$$

for $j = 1, \dots, d$, on the separable Hilbert space $\mathcal{H}$. Assuming (H0)–(H3) we see that A generates a strongly continuous semigroup $t \mapsto (S_t r, Y)$ on $\mathcal{H}$, and $D(A) = D(d/dx) \times \mathbb{R}^m$. If we further assume that the volatility coefficients Σ_j map $\mathcal{H}$ into $\mathcal{H}_0 := H_0 \times \mathbb{R}^m$ then the HJM drift coefficient

$$\Theta = \begin{pmatrix} \alpha_{HJM} \\ b \end{pmatrix} := \begin{pmatrix} \sum_{j=1}^d M(\sigma_j, \sigma_j) \\ b \end{pmatrix} : \mathcal{H} \to \mathcal{H} \tag{1.8}$$

is a well-defined map. Hence any HJM model is uniquely determined by the specification of its volatility structure $\Sigma = (\Sigma_1, \dots, \Sigma_d)$ and the vector field b.

As an illustration we shall always have the following example in mind (see [9, Section 5]).

Example 1.4. *Let $w : \mathbb{R}_{\geq 0} \to [1, \infty)$ be a non-decreasing C^1-function such that $w^{-1/3} \in L^1(\mathbb{R}_{\geq 0})$. We may think of $w(x) = e^{\alpha x}$ or $w(x) = (1 + x)^\alpha$, for $\alpha > 0$ or $\alpha > 3$, respectively. The space H consisting of absolutely continuous functions h on $\mathbb{R}_{\geq 0}$ and equipped with the norm*

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

is a Hilbert space satisfying (H1)–(H2). Property (H3) is satisfied for $H_0 := \{h \in H \mid \lim_{x \to \infty} h(x) = 0\}$.

It is easy to see that $D(d/dx) \subset \{h \in H \cap C^1(\mathbb{R}_{\geq 0}; \mathbb{R}) \mid \partial_x h \in H\}$ and $(d/dx)h = \partial_x h$ (differentiation). Without much loss of generality we shall in fact assume

(H4): $D(d/dx) = \{h \in H \cap C^1(\mathbb{R}_{\geq 0}; \mathbb{R}) \mid \partial_x h \in H\}$.

Also (H4) is satisfied for the spaces H from Example 1.4.

Denote by $A_0 : D(A_0) \subset \mathcal{H}_0 \to \mathcal{H}_0$ the restriction of A to $\mathcal{H}_0$. That is, $D(A_0) = \{h \in D(A) \cap \mathcal{H}_0 \mid Ah \in \mathcal{H}_0\}$. This induces a Fréchet subspace

$$D(A_0^\infty) := \cap_{n \in \mathbb{N}} D(A_0^n)$$

of $\mathcal{H}_0$. As a consequence of (H0)-(H4) we have that

$$(r_1, Y_1, r_2, Y_2) \mapsto \begin{pmatrix} M(r_1, r_2) \\ 0 \end{pmatrix} : D(A_0^\infty) \times D(A_0^\infty) \to D(A^\infty)$$

is a continuous bilinear mapping.

The preceding specifications for Σ are still too general for concrete implementations on $D(A^\infty)$. We actually have the idea of $\Sigma(r, Y)$ being sensitive with respect to Y and functionals of the forward curve r. That is, $\Sigma_j(r, Y) = \phi_j(\ell_1(r), \dots, \ell_p(r), Y)$, for some $p \geq 1$, where $\phi_j : \mathbb{R}^{p+m} \to D(A_0^\infty)$ is a smooth map and $\ell_1, \dots, \ell_p$ denote continuous linear functionals on H. We may think of $\ell_i(r) = (1/x_i) \int_0^{x_i} r(\eta)\, d\eta$

(benchmark yields), or $\ell_i(r) = r(x_i)$ (benchmark forward rates). This idea is (generalized and) expressed in terms of the following regularity and non-degeneracy assumptions:

(A1): We have

$$\Sigma_i(r, Y) = \phi_i(\ell(r), Y), \quad 1 \leq i \leq d,$$

where $\ell \in L(H, \mathbb{R}^p)$, for some $p \in \mathbb{N}$, and $\phi_1, \dots, \phi_d : \mathbb{R}^{p+m} \to D(A_0^\infty)$ are smooth and pointwise linearly independent maps. Moreover

$$b(r, Y) = \phi_0(\ell(r), Y),$$

where $\phi_0 : \mathbb{R}^{p+m} \to \mathbb{R}^m$ is smooth (in view of (1.4) we usually have to assume $\ell_1(r) = r(0)$). Hence

$$\Sigma_j : \mathcal{H} \to D(A_0^\infty) \quad \text{and} \quad \Theta : \mathcal{H} \to D(A^\infty)$$

are Banach maps (see Section 3).

(A2): For every $q \geq 0$, the map

$$(\ell, \ell \circ (d/dx), \dots, \ell \circ (d/dx)^q) : D\left((d/dx)^\infty\right) \to \mathbb{R}^{p(q+1)}$$

is open.

(A3): A is unbounded; that is, $D(A)$ is a strict subset of $\mathcal{H}$. Equivalently, $A : D(A^\infty) \to D(A^\infty)$ is not a Banach map (see Section 3).

Anticipating the results of Section 2 the problem of finding FDRs for (1.3) is now reduced to the following question (recall the definitions (1.7)): do there exist at most n-dimensional submanifolds with boundary $\mathcal{M}$ of $D(A^\infty) \cap \mathcal{U}$ such that (1.5)– (1.6) hold for all $h \in \mathcal{M}$, where $\Xi(h)$ is inward pointing and the $\Sigma_j(h)$ are parallel to the boundary for $h \in \partial\mathcal{M}$?

Remark 1.5. *Replacing the generator $\frac{d}{dx}$ by other generators in (H1)–(H4) is in principle no problem and one can easily formulate the appropriately adapted conditions (A1)–(A3). Then the conclusions of Section 4 hold, too. This is partly worked out in* [14].

2. Finite dimensional Realizations

Identifying finite dimensional realizations with submanifolds with boundary $\mathcal{M}$ of the respective Hilbert space $\mathcal{H}$ has lead to deterministic consistency problems as outlined in Thereom 1.2. To solve these consistency problems we are going to apply Frobenius type Theorems. We are therefore forced to look for a better adapted space of definition for the vector fields in question, which – in our setting – will be given by the Fréchet space $D(A^\infty)$.

Nevertheless we have to face the problem, that there might exist locally invariant submanifolds with boundary outside $D(A^\infty)$. That this is impossible will be the first part of this section. This is essentially a review of [12]. In the second part we introduce the notion of *generic* finite dimensional realizations in contrast to *accidental* ones.

Let $k \geq 1$ be given. We consider a Banach space X and a *continuous local semiflow Fl of C^k-maps* on it, i.e.

i) There is $\varepsilon > 0$ and $V \subset X$ open with $Fl : [0, \varepsilon[\times V \to X$ a continuous map.
ii) $Fl(0, x) = x$ and $Fl(s, Fl(t, x)) = Fl(s + t, x)$ for $s, t, s + t \in [0, \varepsilon[$ and $x, Fl(t, x) \in V$.

iii) The map $Fl_t : V \to X$ is C^k for $t \in [0,\varepsilon[$.

Continuous local semiflows of C^k-maps appear naturally as mild solutions of non-linear evolution equations. The continuous local semiflow Fl is called *local C^k-semiflow* or simply C^k if $Fl : [0,\varepsilon[\times V \to X$ is C^k.

Let X be a Banach space, S a strongly continuous semigroup on X with infinitesimal generator A, and $P : \mathbb{R}_{\geq 0} \times X \to X$ a continuous map. The basic existence, uniqueness and regularity results for the evolution equation

$$\frac{d}{dt}x(t) = Ax(t) + P(t, x(t)). \tag{2.1}$$

is the following (see [30, Theorem 1.2, Chapter 6]).

We say that $P : \mathbb{R}_{\geq 0} \times X \to X$ is *locally Lipschitz continuous on* X if for every $T \geq 0$ and $K \geq 0$ there exists $C = C(T,K)$ such that

$$\|P(t,x) - P(t,y)\| \leq C\|x - y\|$$

for all $t \in [0,T]$, and $x, y \in X$ with $\|x\| \leq K$ and $\|y\| \leq K$.

Theorem 2.1. *Suppose $P : \mathbb{R}_{\geq 0} \times X \to X$ is locally Lipschitz continuous on X. Let $x_0 \in X$. Then there exist a neighborhood U of x_0 and $T > 0$ such that, for every $x \in U$, equation (2.1) has a unique mild solution $x(t)$, $t \in [0,T]$, with $x(0) = x$. If $x(t)$ and $y(t)$ are two mild solutions of* (2.1) *with $x(0) = x \in U$ and $y(0) = y \in U$ then*

$$\sup_{t \in [0,T]} \|x(t) - y(t)\| \leq M e^{MCT} \|x - y\|, \tag{2.2}$$

holds, where

$$M := \sup_{t \in [0,T]} \|S_t\| \tag{2.3}$$

with some $C = C(T,U)$.

Here is the announced regularity result.

Theorem 2.2. *Let $k \geq 1$. Suppose $P : \mathbb{R}_{\geq 0} \times X \to X$ is C^k in x, $D_x^k P$ is locally Lipschitz continuous on X and $D_x^r P$ is continuous on $\mathbb{R}_{\geq 0} \times X$, for all $r \leq k$. Let $x_0 \in X$. Then there exists an open neighborhood U of x_0 and $T > 0$, and a map $F \in C([0,T] \times U, H)$ such that, for every $x \in U$, $F(\cdot, x)$ is the unique mild solution of (2.1) with $F(0,x) = x$. Moreover $F(t,\cdot) \in C^k(U,X)$ for all $t \in [0,T]$.*

This regularity result together with the following fundamental generalization of results from [27] constitutes the final result, for the proofs see [12]:

Theorem 2.3. *Let $k \geq 1$ be given and let $Fl : [0,\varepsilon[\times U \to \mathcal{M}$ be a local semiflow on a finite-dimensional C^k-manifold $\mathcal{M}$ with boundary, which satisfies the following conditions:*

i) *The semiflow $Fl : [0,\varepsilon[\times U \to \mathcal{M}$ is continuous with $U \subset \mathcal{M}$ open.*
ii) *The mapping $Fl(t,.)$ is C^k.*
iii) *For fixed $x \in U$ there exists $\varepsilon_x > 0$ such that $T_x Fl(t,.)$ is invertible for $0 \leq t \leq \varepsilon_x$.*

Then Fl is C^k and for any $x \in U \setminus \partial\mathcal{M}$ there is a local C^k-flow $\tilde{Fl} :]-\delta,\delta[\times V \to \mathcal{M}$ with $V \subset U \setminus \partial\mathcal{M}$ open around x and $\delta \leq \varepsilon$ such that $Fl(y,t) = \tilde{Fl}(y,t)$ for $y \in V$ and $0 \leq t \leq \delta$.

Now we trace back to our original problem. We assume (A1)–(A3). Given $\mathcal{M} \subset \mathcal{H}$ such that the vector fields $\Xi, \Sigma_1, \dots, \Sigma_d$ are tangent to $\mathcal{M}$. By Theorem 2.2 and 2.3 we see that Ξ restricts to a smooth vector field along $\mathcal{M}$, in particular by Theorem 3.1 of [12] we obtain $\mathcal{M} \subset D(A^\infty)$.

Now we assume additionally that we are given $P_1, \dots, P_N$ Banach map vector fields (see Section 3.3) such that pointwise the tangent space of $\mathcal{M}$ is spanned by $\Xi, P_1, \dots, P_N$. In particular $\Sigma_1, \dots, \Sigma_d$ can be represented as linear combinations of P_i along $\mathcal{M}$. Under the assumption that the dimension of $\mathcal{M}$ equals the dimension of D_{LA} at a point $h \in \mathcal{M}$ (see section 4), we can construct the vector fields P_i by iterated Lie brackets of $\Xi, \Sigma_1, \dots, \Sigma_d$, which are defined on $D(A^\infty)$. These Lie brackets are tangent to $\mathcal{M}$ in the differentiable structure as submanifold of $\mathcal{H}$ by representing them as derivatives of smooth pullbacks with respect to the local flows $Fl^{\Sigma_1}, \dots, Fl^{\Sigma_d}$, which leave $\mathcal{M}$ locally invariant by assumption, i.e.

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

for $X, Y : U \subset D(A^\infty) \to D(A^\infty)$ vector fields, where X admits a smooth local flow, and $h \in U$. The same formula holds for vector fields on $\mathcal{M}$. Hence the Lie brackets restrict to smooth vector fields on $\mathcal{M}$ in the differentiable structure as submanifold of $\mathcal{H}$ and the procedure can be iterated. Finally the Lie brackets span the tangent space at every point.

The Banach map principle ([16, Theorem 5.6.3] or Theorem 3.18) yields that each Banach map vector field P_i generates a local flow Fl^{P_i} on $D(A^\infty)$. Given $h_0 \in \mathcal{M} \subset D(A^\infty)$, the composition

$$(u_0, \dots, u_N) \mapsto Fl^{\Xi}_{u_0} \circ Fl^{P_1}_{u_1} \circ \dots \circ Fl^{P_N}_{u_N}(h_0)$$

then yields a chart for $\mathcal{M}$ since the map is smooth to $D(A^\infty) \subset \mathcal{H}$, so a smooth map to $\mathcal{H}$, and as a smooth map to $\mathcal{H}$ it is an immersion locally onto $\mathcal{M}$, so in particular locally one-to-one.

Theorem 2.4. *Let $\mathcal{M} \subset \mathcal{U}$ be a $(N+1)$-dimensional C^∞-submanifold with boundary of $\mathcal{H}$. If $\mathcal{M}$ is locally invariant for Fl^{Ξ}, $Fl^{P_1}, \dots, Fl^{P_N}$, then $\mathcal{M}$ is a C^∞-submanifold with boundary of $D(A^\infty)$.*

In Section 1 we have argued that finite dimensional realizations can essentially be seen as locally invariant submanifolds with boundary of $D(A^\infty)$. We now make this more precise. Let $r_0^* \in \mathcal{U} \cap D(A^\infty)$ and $n \in \mathbb{N}$.

Definition 2.5. *We say that (1.1) admits a* generic n-dimensional realization around r_0^* *if there exists an open neighborhood $\mathcal{V}$ of r_0^* in $\mathcal{U} \cap D(A^\infty)$, an open set U in $\mathbb{R}^n_{\geq 0}$, and a C^∞-map $\alpha : U \times \mathcal{V} \to \mathcal{U} \cap D(A^\infty)$ such that*

i) *$r \in \alpha(U, r)$ for all $r \in \mathcal{V}$,*
ii) *$D_z\alpha(z, r) : \mathbb{R}^n \to D(A^\infty)$ is injective for every $(z, r) \in U \times \mathcal{V}$,*
iii) *$\alpha(z_1, r_1) = \alpha(z_2, r_2)$ implies $D_z\alpha(z_1, r_1)(\mathbb{R}^n) = D_z\alpha(z_2, r_2)(\mathbb{R}^n)$ for all $(z_i, r_i) \in U \times \mathcal{V}$,*
iv) *for every $r^* \in \mathcal{V}$ there exists a U-valued diffusion process Z and a stopping time $\tau > 0$ such that*

$$r_{t \wedge \tau} = \alpha(Z_{t \wedge \tau}, r^*) \tag{2.4}$$

is the (unique) local solution of (1.1) with $r_0 = r^*$.

Definition 2.5 states that a generic n-dimensional realization around r_0^* implies the existence of an FDR at every point r^* in a neighborhood of r_0^*, and these FDRs have a smooth dependence on r^*. In fact, by i) and ii), each $\alpha(\cdot, r^*) : U \to \mathcal{U} \cap D(A^\infty)$ is (after a localization) the parametrization of an n-dimensional submanifold with boundary, say $\mathcal{M}_{r^*}$, of $\mathcal{U} \cap D(A^\infty)$, and (2.4) says that

$$r_{t\wedge\tau} \in \mathcal{M}_{r^*} \text{ for all } \quad t \geq 0.$$

Condition iii) implies that two such leafs $\mathcal{M}_{r_1}$ and $\mathcal{M}_{r_2}$ can only intersect at points where their tangent spaces coincide. According to [14], the family $\{\mathcal{M}_r\}_{r\in\mathcal{V}}$ is called an n-dimensional *weak foliation* on $\mathcal{V}$.

In contrast to a generic n-dimensional realizations we call one single submanifold with boundary $\mathcal{M} \subset \mathcal{U}$, locally invariant with respect to (1.1), an *accidental finite dimensional realization* if there exists an $r_0^* \in \mathcal{M}$ that does not admit a generic FDR around r_0^*.

3. Geometric and Analytic Methods

We are treating the problem of existence of finite dimensional realizations for equations of the type (1.1). In Section 2 we explained that it is sufficient to solve the deterministic consistency problem on the Fréchet space $D(A^\infty)$. Besides the precise analytical formulation of equations of type (1.1) the methods of this section are crucial for the analysis. We shall sketch several ideas to provide a feeling for this analysis, the details can be found in [14]. First we give a guided tour through the proof:

Analysis: Analysis on Fréchet spaces is a subtle subject since – given a map $f : E \to F$ on Fréchet spaces with derivative $Df : U \to L(E, F)$ – it is not clear how to define the second derivative consistently due to the fact that $L(E, F)$ is no more a Fréchet space in general. Therefore some new concepts enter the scenery, which are even for classical analysis a considerable simplification.

Geometry: Given vector fields $\Xi, \Sigma_1, \ldots, \Sigma_d$ on an open subset $\mathcal{U}$ with associated flows Fl, the map

$$(u_0, \ldots, u_d) \mapsto Fl^{\Xi}_{u_0} \circ \cdots \circ Fl^{\Sigma^d}_{u_d}(r^*)$$

is the obvious candidate for a parametrization of a submanifold with boundary at $r^* \in \mathcal{U}$ tangent to $\Xi, \Sigma_1, \ldots, \Sigma_d$, if we expect the dimension to be $d + 1$. To formulate the algebraic obstructions for this assertion, namely that the Lie brackets of the involved vector fields lie in the span of the vector fields, e.g.

$$[\Xi, \Sigma](r) = \lambda_0(r)\Xi(r) + \sum_{i=0}^{d} \lambda_i(r)\Sigma_i(r)$$

for some smooth, real valued functions λ_i, we need an applicable analysis at hand.

Synthesis: To be able to apply the geometric results reasonably to our problem we have to reinvestigate the ingredients of equation (1.1), namely the class of involved vector fields, to obtain finally a fairly general classification result. We shall see that the vector fields Ξ and Σ_j have different analytic properties and are rarely linearly dependent under conditions (A1)–(A3).

3.1. **Analysis.** For the purposes of analysis on open subsets of Fréchet spaces we shall follow two equivalent approaches. The classical Gateaux-approach as outlined in [16] and so called "convenient analysis" as in [23]. On Fréchet spaces these two notions of smoothness coincide and convenient calculus is the appropriate extension of analysis to more general locally convex spaces. The combinations of these methods allow simple and elegant calculations. The main advantage of convenient calculus is however, that one can give a precise analytic meaning (in simple terms) to geometric objects on Fréchet spaces such as vector fields or differential forms (see [23]), which do not lie in Fréchet spaces generically. First we recall the definitions of Gateaux-C^n-calculus.

Definition 3.1. *Let E, F be Fréchet spaces and $U \subset E$ an open subset. A map $P : U \to F$ is called* Gateaux-C^1 *if*

$$DP(f)h := \lim_{t \to 0} \frac{P(f+th) - P(f)}{t}$$

exists for all $f \in U$ and $h \in E$ and $DP : U \times E \to F$ is a continuous map.

For the definition of Gateaux-C^2-maps the ambiguities of calculus on Fréchet spaces already appear. Since there is no Fréchet space topology on the vector space of continuous linear mappings $L(E, F)$ one has to work by point evaluations:

Definition 3.2. *Let E, F be Fréchet spaces and $U \subset E$ an open subset. A map $P : U \to F$ is called* Gateaux-C^2 *if*

$$D^2P(f)(h_1, h_2) := \lim_{t \to 0} \frac{DP(f+th_2)h_1 - DP(f)h_1}{t}$$

exists for all $f \in U$ and $h_1, h_2 \in E$ and $D^2P : U \times E \times E \to F$ is a continuous map. Higher derivatives are defined in a similar way. A map is called Gateaux-smooth *or* Gateaux-C^∞ *if it is Gateaux-C^n for all $n \geq 0$.*

For the construction of differential calculus on locally convex spaces we need the concept of smooth curves into locally convex spaces and the concept of smooth maps on open subsets of locally convex spaces. We remark that already on Fréchet spaces the situation concerning analysis was complicated and unclear until convenient calculus was invented (see [23], pp. 73–77, for extensive historical remarks). The reason for inconsistencies can be found in the fundamental difference between bounded and open subsets on locally convex vector spaces.

We denote the set of continuous linear functionals on a locally convex space E by E'_c. A subset $B \subset E$ is called *bounded* if $l(B)$ is a bounded subset of $\mathbb{R}$ for all $l \in E'_c$. A multilinear map $m : E_1 \times ... \times E_n \to F$ is called *bounded* if bounded sets $B_1 \times ... \times B_n$ are mapped onto bounded subsets of F. Continuous linear functionals are clearly bounded. The locally convex vector space of bounded linear operators with uniform convergence on bounded sets is denoted by $L(E, F)$, the dual space formed by bounded linear functionals by E'. These spaces are locally convex vector spaces, which we shall need for analysis (see [23], 3.17).

Definition 3.3. *Let E be a locally convex space, then $c : \mathbb{R} \to E$ is called* smooth *if all derivatives exist as limits of difference quotients. The set of smooth curves is denoted by $C^\infty(\mathbb{R}, E)$.*

A subset $U \subset E$ is called c^∞-open *if $c^{-1}(U)$ is open in $\mathbb{R}$ for all $c \in C^\infty(\mathbb{R}, E)$. The generated topology on E is called* c^∞-topology *and E equiped with this topology is denoted by $c^\infty E$.*

If U is c^∞-open, a map $f : U \subset E \to \mathbb{R}$ is called smooth *if $f \circ c \in C^\infty(\mathbb{R},\mathbb{R})$ for all $c \in C^\infty(\mathbb{R}, E)$.*

These definitions work for any locally convex vector space, but for the following theorem we need a weak completeness assumption. A locally convex vector space E is called *convenient* if the following property holds: a curve $c : \mathbb{R} \to E$ is smooth if and only if it is weakly smooth, i.e. $l \circ c \in C^\infty(\mathbb{R},\mathbb{R})$ for all $l \in E'$. This is equivalent to the assertion that any smooth curve $c : \mathbb{R} \to E$ can be (Riemann-) integrated in E on compact intervals (see [23], 2.14). The spaces $L(E, F)$ and E' are convenient vector spaces (see [23], 3.17), if E and F are convenient.

Theorem 3.4. *Let E, G, H be convenient vector spaces, $U \subset E$, $V \subset G$ c^∞-open subsets:*

i) *Smooth maps are continuous with respect to the c^∞-topology.*
ii) *Multilinear maps are smooth if and only if they are bounded.*
iii) *If $P : U \to G$ is smooth, then $DP : U \to L(E, G)$ is smooth and bounded linear in the second component, where*
$$DP(f)h := \frac{d}{dt}|_{t=0}P(f + th).$$
iv) *The chain rule holds.*
v) *Let $[f, f + h] := \{f + sh \text{ for } s \in [0, 1]\} \subset U$, then Taylor's formula is true at $f \in U$, where higher derivatives are defined as usual (see iii.),*
$$P(f + h) = \sum_{i=0}^{n} \frac{1}{i!}D^iP(f)h^{(i)} + \int_0^1 \frac{(1-t)^n}{n!} D^{n+1}P(f + th)\,(h^{(n+1)})dt$$
for all $n \in \mathbb{N}$.
vi) *There are natural convenient locally convex structures on $C^\infty(U, F)$ and we have* cartesian closedness
$$C^\infty(U \times V, H) \simeq C^\infty(U, C^\infty(V, H)).$$
via the natural map $f \mapsto \check{f} : U \to C^\infty(V, H)$ for $f \in C^\infty(U \times V, H)$. This natural map is well defined and a smooth linear isomorphism.
vii) *The evaluation and the composition*
$$\begin{aligned} &\mathrm{ev} : C^\infty(U, F) \times U \to F, \quad (P, f) \mapsto P(f) \\ &.\circ. : C^\infty(F, G) \times C^\infty(U, F) \to C^\infty(U, G), \quad (Q, R) \mapsto Q \circ R \end{aligned}$$
are smooth maps.
viii) *A map $P : U \subset E \to L(G, H)$ is smooth if and only if $(\mathrm{ev}_g \circ P)$ is smooth for all $g \in G$.*

Proof. For the proofs see [23] in Subsections 3.12, 3.13, 3.18, 5.11, 5.12, 5.18. □

Convenient Calculus is an extension of the Gateaux-Calculus to locally convex spaces, where all necessary tools for analysis are preserved. Since typically vector spaces like $C^\infty(U, F)$ or $L(E, F)$ are not Fréchet spaces, this extension is very useful for the analysis of the geometric objects in Section 3.

Theorem 3.5. *Let E, F be Fréchet spaces and $U \subset E$ a c^∞-open subset, then U is open and $P : U \subset E \to F$ is Gateaux-smooth if and only if P is smooth (in the convenient sense).*

Concerning differential equations, there are possible counterexamples on non-normable Fréchet spaces in all directions, which causes some problems in the foundations of differential geometry (see [23] and the review article [25]).

If not otherwise stated, E and F denote Fréchet spaces and B a Banach space in what follows. A *vector field* P on an open subset $U \subset E$ is a smooth map $P : U \to E$. We denote by $\mathfrak{X}(U)$ the convenient space of all vector fields on an open subset of a Fréchet space E. Given $P : U \subset E \to E$ a vector field on U. We are looking for solutions of the ordinary differential equation with initial value $g \in U$

$$\begin{aligned} f :\,]-\varepsilon,\varepsilon[&\to U \text{ smooth} \\ \frac{d}{dt} f(t) &= P(f(t)) \\ f(0) &= g \in U. \end{aligned}$$

If for any initial value g in a small neighborhood V of $f_0 \in U$ there is a unique smooth solution $t \mapsto f_g(t)$ for $t \in]-\varepsilon,\varepsilon[$ depending smoothly on the initial value g, then $Fl(t,g) := f_g(t)$ defines a *local flow*, i.e. a smooth map

$$\begin{aligned} Fl :\,]-\varepsilon,\varepsilon[\times V &\to E \\ Fl(0,g) &= g \\ Fl(t, Fl(s,g)) &= Fl(s+t, g) \end{aligned}$$

if $s, t, s+t \in]-\varepsilon,\varepsilon[$ and $Fl(s,g) \in V$. If there is a local flow around $f_0 \in U$ (this shall mean once and for all: "in an open, convex neighborhood of f_0"), the differential equation is uniquely solvable around $f_0 \in U$ and the dependence on initial values is smooth (see Lemma 3.6 for the proof). Notice at this point that it is irrelevant if we define "smooth dependence" on initial values via smooth maps $V \to C^\infty(]-\varepsilon,\varepsilon[, E)$ or $V \times]-\varepsilon,\varepsilon[\to E$ by cartesian closedness. We shall denote $f_g(t) = Fl_t(g) = Fl(t,g)$.

We can replace in the above definition of a local flow the interval $]-\varepsilon,\varepsilon[$ by $[0,\varepsilon[$ to obtain *local semiflows*. The initial value problem

$$\begin{aligned} f : [0,\varepsilon[&\to U \text{ smooth} \\ \frac{d}{dt} f(t) &= P(f(t)) \\ f(0) &= g \in U. \end{aligned}$$

admits unique solutions around an initial value depending smoothly on the initial values if and only if a local semiflow exists. Here we need convenient analysis of non-open domains, see [14] or [23]. The notion of a local semiflow is redundant on Banach spaces.

Lemma 3.6. *Let Fl be a local semiflow on $[0,\varepsilon[\times U \to E$, then the map $P(f) := \frac{d}{dt}|_{t=0} Fl(t,f)$ is a well defined smooth vector field. We obtain*

$$DFl_t(f)P(f) = P(Fl_t(f))$$

and the initial value problem has unique solutions for small times which coincide with the given semiflow.

3.2. **Geometry.** We are interested in the geometry generated by a finite number of vector fields given on an open subset of a Fréchet space E. Therefore we need the notions of finite-dimensional submanifolds (with boundary) of a Fréchet space (see [23] for all details and more). Here and subsequent E denotes a Fréchet space. From the classical definition we can conclude the following Lemma (see [14]):

Lemma 3.7 (Submanifolds by Parametrization)**.** *Let E be a Fréchet space and $\phi : U \subset \mathbb{R}^n_{\geq 0} \to E$ a smooth immersion, i.e. for $u \in U$ the map $D\phi(u)$ is injective, then for any $u_0 \in U$ there is a small open neighborhood V of u_0 such that $\phi(V)$ is a submanifold with boundary of E, $\phi|_V$ is a called a* parametrization.

Definition 3.8 (Lie bracket)**.** *The* Lie bracket *of two vector fields $X, Y \in \mathfrak{X}(U)$ is defined by the following formula:*

$$[X,Y](f) = DX(f) \cdot Y(f) - DY(f) \cdot X(f)$$

and is a bounded, skew-symmetric bilinear map from $\mathfrak{X}(U) \times \mathfrak{X}(U)$ into $\mathfrak{X}(U)$.

Definition 3.9. *Let $\phi : U \subset \mathbb{R}^n_{\geq 0} \to E$ be a smooth parametrization of a submanifold with boundary $\mathcal{M} \subset E$, i.e. $\phi(U) = \mathcal{M}$. A vector field $X : U \to \mathbb{R}^n$ is called ϕ-related to $Y : \mathcal{V} \subset E \to E$ if $T_u\phi(X_u) = Y_{\phi(u)}$ for all $u \in U$. This is denoted by $X \sim_\phi Y$.*

Proposition 3.10. *Let $\mathcal{U}$ be an open set in E, and $\mathcal{M} \subset \mathcal{U}$ be a submanifold with boundary. If two vector fields $X_1, X_2 \in \mathfrak{X}(\mathcal{U})$ are tangent to $\mathcal{M}$, then $[X,Y](h) \in T_h\mathcal{M}$ for $h \in \mathcal{U}$.*

Proof. Take a parametrization $\phi : U \subset \mathbb{R}^n_{\geq 0} \to E$ of $\mathcal{M}$ around $h_0 \in \mathcal{M}$. By [23] we obtain, that if two vector fields are ϕ-related, then their Lie bracket is ϕ-related, too. Given two vector fields $X_1, X_2 : \mathcal{U} \to E$ such that for all $h \in \mathcal{M}$ we have $X_i(h) \in T_h\mathcal{M}$ for $i = 1, 2$. Then we can define vector fields Y_1, Y_2 on U by restriction and pulling back to U such that

$$T_u\phi(Y_i(u)) = X_i(\phi(u))$$

for $i = 1, 2$ and $u \in U$. So X_i is ϕ-related to Y_i for $i = 1, 2$ and therefore their Lie bracket as well. Consequently all Lie brackets take along $\mathcal{M}$ values in its tangent space, since we can choose a parametrization around any point $h_0 \in \mathcal{M}$. $\square$

¿From this observation Frobenius theory can be built up. We denote by $\langle \ldots \rangle$ the generated vector space over the reals $\mathbb{R}$.

Definition 3.11. *Let E be a Fréchet space, U an open subset. A* distribution *on U is a collection of vector subspaces $D = \{D_f\}_{f \in U}$ of E. A vector field $X \in \mathfrak{X}(U)$ is said to take values in D if $X(f) \in D(f)$ for $f \in U$. A distribution D on U is said to be* involutive *if for any two locally given vector fields X, Y with values in D the Lie bracket $[X,Y]$ takes values in D.*

A distribution is said to have constant rank *on U if $\dim_{\mathbb{R}} D_f$ is locally constant for $f \in U$. A distribution is called* smooth *if there is a set S of locally defined vector fields on U such that*

$$D_f = \langle \{X(f) | (X : U_X \to E) \in S \text{ and } f \in U_X\} \rangle.$$

We say that the distribution admits local frames *on U if for any $f \in U$ there is an open neighborhood $f \in V \subset U$ and n smooth, pointwise linearly independent vector* *fields* $X_1, ..., X_n$ *on* V *with*

$$\langle X_1(g), ..., X_n(g) \rangle = D_g$$

for $g \in V$.

Remark 3.12. *Given a distribution* D *on* U *generated by a set of local vector fields* S, *such that the dimensions of* D_f *are bounded by a fixed constant* N. *Let* $f \in U$ *be a point with maximal dimension* $n_f = \dim_{\mathbb{R}} D_f$, *then there are* n_f *smooth local vector fields* $X_1, ..., X_{n_f} \in S$ *with common domain of definition* U' *such that*

$$\langle X_1(f), ..., X_n(f) \rangle = D_f.$$

Choosing n_f *continuous linear functionals* $l_1, ..., l_{n_f} \in E'$ *with* $l_i(X_j(f)) = \delta_{ij}$, *then the continuous mapping* $M : U' \to L(\mathbb{R}^{n_f}), \quad g \mapsto (l_i(X_j(g)))$ *has range in the invertible matrices in a small neighborhood of* f. *Consequently in this neighborhood the dimension of* D_g *is at least* n_f. *It follows by maximality of* n_f *that it is exactly* n_f. *In particular the distribution admits a local frame at* f.

The concept of weak foliations will be perfectly adapted to the FDR-problem:

Definition 3.13. *A* weak foliation $\mathcal{F}$ *of dimension* n *on an open subset* U *of a Fréchet space* E *is a collection of submanifolds with boundary* $\{M_r\}_{r \in U}$ *such that*

i) *For all* $r \in U$ *we have* $r \in M_r$ *and the dimension of* M_r *is* n.
ii) *The distribution*

$$D(\mathcal{F})(f) := \langle T_f M_r \text{ for all } r \in U \text{ with } f \in M_r \rangle$$

has dimension n *for all* $f \in U$, *i.e. given* $f \in U$ *the tangent spaces* $T_f M_r$ *agree for all* $M_r \ni f$. *This distribution is called the* tangent distribution *of* $\mathcal{F}$.

Given any distribution D *we say that* D *is* tangent *to* $\mathcal{F}$ *if* $D(f) \subset D(\mathcal{F})(f)$ *for all* $f \in U$.

Theorem 3.14. *Let* D *be an smooth distribution of constant rank* n *on an open subset* U *of a Fréchet space* E. *Assume that for any point* f_0 *the distribution admits a local frame of vector fields* $X_1, ..., X_n$, *where* $X_1, ..., X_{n-1}$ *admit local flows* $Fl_t^{X_i}$ *and* X_n *admits a local semiflow* $Fl_t^{X_n}$. *Then* D *is involutive if and only if it is tangent to an* n*-dimensional weak foliation.*

For details on Frobenius theorems in the classical setting see [22]. The phenomenon that there is no Frobenius chart is due to the fact that there is one vector field among the vector fields X_1,...,X_n (generating the distribution D) admitting only a local semiflow. If all of them admitted flows, there would exist a Frobenius chart, which can be given by a construction outlined in [33]. The non-existence of a Frobenius-chart means that the leafs cannot be parallelized, since they follow a semiflow, which means in turn that "gaps" between two leafs can occur and leafs can touch. This is an infinite dimensional phenomenon, which does not appear in finite dimensions.

3.3. **Synthesis.** From the geometric considerations we can conclude the existence of FDRs if

1: the distribution D_{LA} generated by $\Xi, \Sigma_1, \ldots, \Sigma_d$ and all mutually generated Lie brackets has locally constant dimension $N_{LA} \geq 1$.

2: there is a frame of vector fields $X_1, \ldots, X_{N_{LA}}$ for D_{LA} around any point, such that $X_1, \ldots, X_{N_{LA}-1}$ admit local flows and $X_{N_{LA}}$ admits a local semiflow.

Applying Theorem 3.14 then yields the existence of N_{LA}-dimensional locally invariant submanifolds with boundary. For an interesting application one has to investigate the meaning of 1. closely and guarantee 2. in as many situations as possible. In this subsection the conditions (A1)–(A3) are applied such that only the algebraic condition remains and one can nicely distinguish between the vector fields admitting flows and the one only admitting a semiflow.

Definition 3.15. *Given a Fréchet space E, a smooth map $P : U \subset E \to E$ is called a* Banach map *if there are smooth (not necessarily linear) maps $R : U \subset E \to B$ and $Q : V \subset B \to E$ such that $P = Q \circ R$, where B is a Banach space and $V \subset B$ is an open set.*

We denote by $\mathcal{B}(U)$ the set of Banach map vector fields on an open subset of a Fréchet space E.

Theorem 3.16. *$\mathcal{B}(U)$ is a $C^\infty(U,\mathbb{R})$-submodule of $\mathfrak{X}(U)$.*

Proof. We have to show that for $\psi, \eta \in C^\infty(U,\mathbb{R})$ and $P_1, P_2 \in \mathcal{B}(U)$ the linear combination $\psi P_1 + \eta P_2 \in \mathcal{B}(U)$. Given $P_i = Q_i \circ R_i$ for $i = 1,2$ with intermediate Banach spaces B_i, then $\psi P_1 + \eta P_2 = Q \circ R$ with $Q : \mathbb{R}^2 \times V_1 \times V_2 \subset \mathbb{R}^2 \times B_1 \times B_2 \to E$ and $R : U \to \mathbb{R}^2 \times B_1 \times B_2$ such that

$$Q(r, s, v_1, v_2) = rQ_1(v_1) + sQ_2(v_2)$$
$$R(f) = (\psi(f), \eta(f), R_1(f), R_2(f))$$

So the sum $\psi P_1 + \eta P_2$ is a Banach map and therefore the set of all Banach map vector fields carries the asserted submodule structure. □

Lemma 3.17. *Let U be an open set in a Fréchet space E, then $\mathcal{B}(U)$ is a subalgebra with respect to the Lie bracket. Let A be a bounded linear operator on E, then $[A, \mathcal{B}(U)] \subset \mathcal{B}(U)$. Consequently the Lie algebra $L(E)$ acts on $\mathcal{B}(U)$ by the Lie bracket.*

Proof. Given two Banach maps P_1 and P_2, $DP_1(f) \cdot P_2(f) = DQ_1(R_1(f)) \cdot DR_1(f) \cdot P_2(f)$ holds, which can be written as composition of $DQ_1(v) \cdot w$ for $v, w \in B$ and $(R_1(f), DR_1(f) \cdot P_2(f))$ for $f \in U$. So the Lie bracket lies in $\mathcal{B}(U)$. Given $A \in L(E)$, we see that $AP_1(f) - DP_1(f) \cdot Af$ is a Banach map by an obvious decomposition. □

Banach map vector fields admit solutions of initial value problems.

Theorem 3.18 (Banach map principle). *Let $P : U \subset E \to E$ be a Banach map, then P admits a local flow around any point $g \in U$.*

Proof. For the proof see [16], Theorem 5.6.3. □

We are in particular interested in special types of differential equations on Fréchet spaces E, namely Banach map perturbed bounded linear equations. Given a bounded linear operator $A : E \to E$, the abstract Cauchy problem associated to A is given by the initial value problem associated to A. We assume that there is a smooth semigroup of bounded linear operators $S : \mathbb{R}_{\geq 0} \to L(E,E)$ such that

$$\lim_{t \downarrow 0} \frac{S_t - id}{t} = A$$

which is a global semiflow for the linear vector field $f \mapsto Af$. Notice that the theory of bounded linear operators on Fréchet spaces contains as a special case Hille-Yosida-Theory of unbounded operators on Banach spaces (see for example [32]).

Given a strongly continuous semigroup S_t for $t \geq 0$ of bounded linear operators on a Banach space B, then $D(A^n)$ with the respective operator norms $p_n(f) := \sum_{i=0}^n ||A^i f||$ for $n \geq 0$ and $f \in D(A^n)$ is a Banach space, where the semigroup restricts to a strongly continuous semigroup $S^{(n)}$. Consequently the semigroup restricts to the Fréchet space $D(A^\infty)$. This semigroup is now smooth.

Given a Banach map $P : U \subset E \to E$, we want to investigate the solutions of the initial value problem

$$\frac{d}{dt} f(t) = Af(t) + P(f(t)), \quad f(0) = f_0.$$

Theorem 3.19. *Let E be a Fréchet space and A be the generator of a smooth semigroup $S : \mathbb{R} \to L(E)$ of bounded linear operators on E. Let $P : U \subset E \to E$ be a Banach map. For any $f_0 \in U$ there is $\varepsilon > 0$ and an open set V containing f_0 and a local semiflow $Fl : [0, \varepsilon[\times V \to U$ satisfying*

$$\begin{aligned} \frac{d}{dt} Fl(t,f) &= AFl(t,f) + P(Fl(t,f)) \\ Fl(0,f) &= f \end{aligned}$$

for all $(t,f) \in [0, \varepsilon[\times V$.

Proof. For the proof see [14]. □

For the purposes of classification we shall need the following result, see [14].

Lemma 3.20. *Let A be the generator of a strongly continuous semigroup S on a Banach space B, then the operator $A : D(A^\infty) \to D(A^\infty)$ is a Banach map if and only if $A : B \to B$ is bounded.*

Now we can formulate the following conclusions from (A1)–(A3), which makes the geometric conditions applicable:

1: The vector fields Σ_i and Θ are Banach map vector fields for $i = 1, \ldots, n$.
2: The Lie brackets $[\Xi, \Sigma_i]$ and $[\Sigma_i, \Sigma_j]$ are Banach maps for $i, j = 1, \ldots, n$. Any further Lie bracket with a Banach map vector field yields a Banach map vector field. This is due to Lemma 3.17.
3: The vector field Ξ is not a Banach map due to (A3) and Lemma 3.20, but generates a local semiflow due to Theorem 3.19.

4. Results on the Existence of Finite dimensional Realizations

Now we can derive several results with the developed geometric tools. We consider the setup from Section 1.3 and shall always assume (H0)–(H4) and (A1)–(A3). This is an extension of what we derived in [14], therefore all proofs are given here.

Lemma 4.1. *Let $X_1, \ldots, X_k$ be pointwise linearly independent Banach maps on an open set $\mathcal{V}$ in $D(A^\infty)$, for some $k \in \mathbb{N}$. Then the set*

$$\mathcal{N} = \{h \in \mathcal{V} \mid \Xi(h) \in \langle X_1(h), \ldots, X_k(h) \rangle\}$$

is closed and nowhere dense in $\mathcal{V}$.

Proof. Clearly, $\mathcal{N}$ is closed by continuity of Ξ and $X_1, \dots, X_k$. Now suppose there exists a set $\mathcal{W} \subset \mathcal{N}$ which is open in $D(A^\infty)$, such that for every $h \in \mathcal{W}$ there exist unique numbers $c_1(h), \dots, c_k(h)$ such that

$$\Xi(h) = \sum_{j=1}^{k} c_j(h) X_j(h). \tag{4.1}$$

We can choose linear functionals $\xi_1, \dots, \xi_k$ on $D(A^\infty)$ such that the $k \times k$-matrix $M_{ij}(h) := \xi_i(X_j(h))$ is smooth and invertible on $\mathcal{W}$ (otherwise we choose a smaller open subset $\mathcal{W}$). Hence

$$c_i(h) = \sum_{j=1}^{k} M_{ij}^{-1}(h) \xi_j(\Xi(h))$$

are smooth functions on $\mathcal{W}$. Then (4.1) implies that A is a Banach map on $\mathcal{W}$. But this contradicts (A3), whence the claim. □

The vector fields $\Xi, \Sigma_1, \dots, \Sigma_d$ induce two distributions on $D(A^\infty)$: their linear span $D = \langle \Xi, \Sigma_1, \dots, \Sigma_d \rangle$, and the distribution D_{LA} generated by all multiple Lie brackets of these vector fields. As a consequence of (A1) and Lemma 4.1 there exists a closed and nowhere dense set $\mathcal{N}$ in $D(A^\infty)$ such that

$$\dim D_{LA}(h) \geq \dim D(h) = d + 1 \quad \text{for } h \in D(A^\infty) \setminus \mathcal{N}. \tag{4.2}$$

Remark 4.2. *The preceding observation proves a conjecture in* [4], *namely that every nontrivial generic short rate model is of dimension 2 (see* [4, Remark 7.1]*).*

The following is a modification of the necessary condition in Theorem 3.14.

Proposition 4.3. *Let $\mathcal{V}$ be an open set in $D(A^\infty)$, and $\mathcal{M} \subset \mathcal{V}$ be a submanifold with boundary. If D is tangent to $\mathcal{M}$, then $D_{LA}(h) \subset T_h\mathcal{M}$ for $h \in \mathcal{V}$.*

Proof. See Proposition 3.10. □

Let $\mathcal{V}$ denote an open connected set in $D(A^\infty)$ in what follows. Proposition 4.3 tells us that boundedness of $\dim D_{LA}$ on $\mathcal{V}$ is a necessary condition for the existence of a finite-dimensional weak foliation on $\mathcal{V}$. To avoid the difficulties in the analysis of degenerate situations where $\dim D_{LA}$ is not constant on $\mathcal{V}$, we only consider the non-degenerate case. This is our appropriate Frobenius condition

(F): D_{LA} has constant finite dimension N_{LA} on $\mathcal{V}$.

Here and subsequently, we let (F) be in force. In view of (4.2) we have $N_{LA} \geq d + 1$ and the following proposition holds.

Proposition 4.4. *We have*

$$\Xi(h) \notin \langle \Sigma_1(h), \dots, \Sigma_d(h) \rangle, \text{ for all } \quad h \in \mathcal{V}. \tag{4.3}$$

Moreover, for any $h_0 \in \mathcal{V}$ there exists an open neighborhood $\mathcal{W}$ and Banach maps $X_{d+1}, \dots, X_{N_{LA}-1}$ on $\mathcal{W}$ such that

$$D_{LA} = \langle \Xi, \Sigma_1, \dots, \Sigma_d, X_{d+1}, \dots, X_{N_{LA}-1} \rangle \quad \text{on } \mathcal{W}.$$

In particular, D_{LA} is tangent to an N_{LA}-dimensional weak foliation $\mathcal{F}$ on $\mathcal{V}$.

Proof. Suppose $\Xi(h_0) \in \langle \Sigma_1(h_0), \dots, \Sigma_d(h_0)\rangle$, for some $h_0 \in \mathcal{V}$. By the definition of D_{LA} and Lemma 3.17 there exist $N_{LA} - d$ Banach maps $X_{d+1}, \dots, X_{N_{LA}}$ on $\mathcal{V}$ such that

$$D_{LA}(h) = \langle \Sigma_1(h), \dots, \Sigma_d(h), X_{d+1}(h), \dots, X_{N_{LA}}(h)\rangle,$$

for $h = h_0$, and hence for all h in a neighborhood of h_0, by continuity. But this implies that $\Xi(h)$ lies in the span of Banach maps, for all h in an open set. This contradicts Lemma 4.1, whence (4.3). The rest of the proposition follows by Remark 3.12 and Theorem 3.14. $\square$

The following theorem provides a strong necessary condition for the structure of D_{LA}, which leads to a full classification of $\mathcal{F}$ in the case where $m = 0$ (see Section 5).

Theorem 4.5. *Under the above assumption (F) there exist pointwise linearly independent vector fields $\lambda_1, \dots, \lambda_{N_{LA}-1} \in C^\infty(V, D(A_0^\infty))$ such that*

$$D_{LA}(r, Y) = \langle \Xi(r, Y), \lambda_1(Y), \dots, \lambda_{N_{LA}-1}(Y)\rangle$$

and

$$\Sigma_j(r, Y) \in \langle \lambda_1(Y), \dots, \lambda_{N_{LA}-1}(Y)\rangle, \quad 1 \le j \le d, \tag{4.4}$$

for all $(r, Y) \in \mathcal{V}$.

Thus, if $m = 0$ then the range of Σ_j is the constant finite-dimensional subspace $\langle \lambda_1, \dots, \lambda_{N_{LA}-1}\rangle$ in $D(A_0^\infty)$.

Proof. Let $1 \le i, j \le d$, recall (A1)–(A3) and calculate

$$D\Sigma_i(r, Y) \cdot \begin{pmatrix} \overline{r} \\ \overline{Y} \end{pmatrix} = D\phi_i(\ell(r), Y) \cdot \begin{pmatrix} \ell(\overline{r}) \\ \overline{Y} \end{pmatrix} \tag{4.5}$$

(here the linearity of ℓ is essential!), hence

$$D\Sigma_i(r, Y) \cdot \Sigma_j(r, Y) = D\phi_i(\ell(r), Y) \cdot \begin{pmatrix} \ell(\phi_j^{(1)}(\ell(r), Y)) \\ \phi_j^{(2)}(\ell(r), Y) \end{pmatrix}, \tag{4.6}$$

where we use the notation

$$h = (h^{(1)}, h^{(2)}) \in H \times \mathbb{R}^m = \mathcal{H}.$$

In view of (1.5), (1.8) and (4.6)we define the smooth map

$$\Gamma := \sum_{j=0}^{d} \Gamma_j : \mathbb{R}^{p+m} \to D(A^\infty)$$

by $\Gamma_0(y, Y) := \begin{pmatrix} 0 \\ \phi_0(y, Y) \end{pmatrix}$ and, for $1 \le j \le d$,

$$\Gamma_j(y, Y) := \begin{pmatrix} M(\phi_j^{(1)}(r, Y), \phi_j^{(1)}(r, Y)) \\ 0 \end{pmatrix} - \frac{1}{2} D\phi_i(y, Y) \cdot \begin{pmatrix} \ell(\phi_j^{(1)}(y, Y)) \\ \phi_j^{(2)}(y, Y) \end{pmatrix},$$

such that we can write

$$\Xi(r, Y) = \begin{pmatrix} \frac{d}{dx} r \\ 0 \end{pmatrix} + \Gamma(\ell(r), Y). \tag{4.7}$$

We already know from Lemma 3.17 that $[\Sigma_i, \Sigma_j]$ and $[\Xi, \Sigma_j]$ are Banach maps. In fact, combining (4.5)–(4.7) yields the decompositions

$$\begin{aligned}[\Sigma_i, \Sigma_j](r, Y) &= \Upsilon_{ij}(\ell(r), Y),\\ [\Xi, \Sigma_j](r, Y) &= \Phi_j(\ell(r), \ell((d/dx)r), Y),\end{aligned}$$

for some smooth maps $\Upsilon_{ij} : \mathbb{R}^{p+m} \to D(A^\infty)$ and $\Phi_j : \mathbb{R}^{2p+m} \to D(A^\infty)$ and $(r, Y) \in \mathcal{V}$.

Now fix $(r^*, Y^*) \in \mathcal{V}$. By induction of the preceding argument and Proposition 4.4 there exists an open neighborhood of the product form (this we can assume without loss of generality)

$$\mathcal{V}^* = U^* \times V^* \in D((d/dx)^\infty) \times \mathbb{R}^m$$

of (r^*, Y^*), an integer $q \geq -1$, and Banach maps $X_1, \dots, X_{N_{LA}-1}$ with decomposition

$$X_i(r, Y) = \Psi_i\left(\ell(r), \ell((d/dx)r), \dots, \ell((d/dx)^q r), Y\right), \tag{4.8}$$

for pointwise linearly independent smooth maps $\Psi_i : \mathbb{R}^{p(q+1)} \times \mathbb{R}^m \to D(A^\infty)$ such that

$$D_{LA} = \langle \Xi, X_1, \dots, X_{N_{LA}-1}\rangle \quad \text{on } \mathcal{V}^*. \tag{4.9}$$

Notice that the case $q = -1$ is included in a consistent way: it simply means that Ψ_i in (4.8), and hence X_i, does only depend on Y.

There exists a minimal integer, still denoted by q, with the above properties. We shall show that $q = -1$.

We argue by contradiction and suppose that $q \geq 0$. We claim that then there exists smooth maps $\tilde{\Psi}_i : \mathbb{R}^{pq} \times \mathbb{R}^m \to D(A^\infty)$ such that we can replace X_i in (4.9) with $\tilde{X}_i = \tilde{\Psi}_i \circ (\ell, \dots, \ell \circ (d/dx)^{q-1}, id_{\mathbb{R}^m})$. Indeed, since $[\Xi, X_i]$ is a Banach map on $\mathcal{V}^*$ (see Lemma 3.17), for every $(r, Y) \in \mathcal{V}^*$ there exists numbers $c_{ij}(r, Y)$ such that

$$[\Xi, X_i](r, Y) = \sum_{j=1}^{N_{LA}-1} c_{ij}(r, Y) X_j(r, Y), \quad 1 \leq i \leq N_{LA} - 1. \tag{4.10}$$

By explicit calculation we obtain

$$[\Xi, X_i] = \Delta_i \circ (\ell, \dots, \ell \circ (d/dx)^{q+1}, id_{\mathbb{R}^m}), \tag{4.11}$$

where

$$\begin{aligned}\Delta_i(y, z, Y) = A\Psi_i(y, Y) &+ D\Gamma(y_0, Y) \cdot \begin{pmatrix} \ell(\Psi_i^{(1)}(y, Y)) \\ \Psi_i^{(2)}(y, Y) \end{pmatrix}\\ &- D\Psi_i(y, Y) \cdot \left(\begin{pmatrix} y_1 \\ \vdots \\ y_q \\ z \\ 0 \end{pmatrix} + \begin{pmatrix} \ell(\Gamma^{(1)}(y_0, Y)) \\ \vdots \\ \ell \circ (d/dx)^{q-1}(\Gamma^{(1)}(y_0, Y)) \\ \ell \circ (d/dx)^q(\Gamma^{(1)}(y_0, Y)) \\ \Gamma^{(2)}(y_0, Y) \end{pmatrix} \right),\end{aligned} \tag{4.12}$$

for $(y, z, Y) = (y_0, \dots, y_q, z, Y) \in \mathbb{R}^{p(q+1)} \times \mathbb{R}^p \times \mathbb{R}^m$. As in the proof of Lemma 4.1 we find linear functionals $\xi_1, \dots, \xi_{N_{LA}-1}$ on $D(A^\infty)$ such that the $(N_{LA} - 1) \times (N_{LA} - 1)$-matrix $M_{ij}(y, Y) := \xi_i(\Psi_j(y, Y))$ is smooth and invertible on

$$(\ell, \dots, \ell \circ (d/dx)^q)(U^*) \times V^* =: W \times V^*,$$

which is an open set in $\mathbb{R}^{p(q+1)} \times \mathbb{R}^m$ by (A2). Equating (4.10) and (4.11), applying the functionals ξ_k and inverting gives that

$$(y,z,Y) \mapsto \gamma_{ij}(y,z,Y) := \sum_{k=1}^{N_{LA}-1} M_{jk}^{-1}(y,Y)\xi_k(\Delta_i(y,z,Y))$$

are smooth functions from

$$(\ell,\dots,\ell\circ(d/dx)^{q+1})(U^*) \times V^* =: W' \times V^* \subset \mathbb{R}^{p(q+2)} \times \mathbb{R}^m$$

into $\mathbb{R}$, and they satisfy

$$c_{ij}(r,Y) = \gamma_{ij} \circ (\ell(r),\dots,\ell((d/dx)^{q+1}r),Y)$$

on $\mathcal{V}^*$, hence

$$\Delta_i(y,z,Y) = \sum_{j=1}^{N_{LA}-1} \gamma_{ij}(y,z,Y)\Psi_j(y,Y), \quad \forall (y,z,Y) \in W' \times V^*. \tag{4.13}$$

Differentiating (4.13) with respect to z (which makes sense since W' is open by (A2)) yields, see (4.12),

$$D_{y_q}\Psi_i(y,Y) = \sum_{j=1}^{N_{LA}-1} \Psi_j(y,Y)D_z\gamma_{ij}(y,z,Y), \quad \forall (y,z,Y) \in W' \times V^*.$$

Arguing again by linear independence of $\Psi_1,\dots,\Psi_{N_{LA}-1}$ we see that the maps

$$D_z\gamma_{ij}(y,z,Y) \equiv: \beta_{ij}(y,Y)$$

depend only on (y,Y). We may assume that $W = W_0 \times W_1$ where $W_0 \subset \mathbb{R}^{pq}$ and $W_1 \subset \mathbb{R}^p$ are open such that $(y_0^*,\dots,y_q^*) := (\ell,\ell\circ(d/dx),\dots,\ell\circ(d/dx)^q)(r^*) \in W_0 \times W_1$, and W_1 is star-shaped with respect to y_q^* (otherwise replace U^* accordingly). Now let $(y,Y) \in W_0 \times W_1 \times V^*$ and define

$$\psi_i(t) := \Psi_i(y_0,\dots,y_{q-1},y_q^* + t(y_q - y_q^*),Y).$$

Then there exists an open interval I containing $[0,1]$ such that

$$\frac{d}{dt}\psi_i(t) = \sum_{j=1}^{N_{LA}-1} \left(\beta_{ij}(y_0,\dots,y_{q-1},y_q^* + t(y_q - y_q^*),Y)\cdot(y_q - y_q^*)\right)\psi_j(t)$$
$$\psi_i(0) = \Psi_i(y_0,\dots,y_{q-1},y_q^*,Y), \quad i = 1,\dots,N_{LA}-1,$$

for $t \in I$. This system of differential equations has a unique solution, which is of the form

$$\psi_i(t) = \sum_{j=1}^{N_{LA}-1} \alpha_{ij}(t)\psi_j(0),$$

for some smooth curves $\alpha_{ij} : I \to \mathbb{R}$. In particular, for $t = 1$,

$$\Psi_i(y_0,\dots,y_q,Y) = \psi_i(1) = \sum_{j=1}^{N_{LA}-1} \alpha_{ij}(1)\Psi_j(y_0,\dots,y_{q-1},y_q^*,Y).$$

This way we find a smooth real-valued matrix-valued map, again denoted by (α_{ij}), on $W_0 \times W_1 \times V^*$ such that

$$\Psi_i(y,Y) = \sum_{j=1}^{N_{LA}-1} \alpha_{ij}(y,Y)\Psi_j(y_0,\dots,y_{q-1},y_q^*,Y), \quad \forall (y,Y) \in W_0 \times W_1 \times V^*.$$

But this implies that Ξ and the Banach maps $\Psi_j(\cdot, y_q^*, \cdot) \circ (\ell, \dots, \ell \circ (d/dx)^{q-1}, id_{\mathbb{R}^m})$ span the Lie algebra D_{LA} on $\mathcal{V}^*$. Whence the claim.

But q was supposed to be minimal – a contradiction. Hence $q = -1$; that is, $X_1, \dots, X_{N_{LA}-1}$ in (4.9) can be chosen to depend only on Y in some neighborhood of (r^*, Y^*). Since $(r^*, Y^*) \in \mathcal{V}$ was arbitrary and $\mathcal{V}$ is connected, the theorem now follows by a continuity argument. □

5. Properties of the Factor processes

5.1. **Finite dimensional HJM-models.** Throughout this section we let $m = 0$. That is, $\mathcal{H} = H$ and $\mathcal{U}$ is a convex open set in $\mathcal{H}$ where equation (1.2) is defined. Moreover, (1.7) now reads

$$A = \frac{d}{dx}, \quad \Theta(r) = \alpha_{HJM}(r) = \sum_{j=1}^{d} M(\Sigma_j(r), \Sigma_j(r)), \quad \Sigma_j(r) = \sigma_j(r), \tag{5.1}$$

for $j = 1, \dots, d$. At this point we also recall (1.5)

$$\begin{aligned} \Xi(r) &= Ar + \Theta(r) - \frac{1}{2}\sum_{j=1}^{d} D\Sigma_j(r)\Sigma_j(r) \\ &= \frac{d}{dx}r + \alpha_{HJM}(r) - \frac{1}{2}\sum_{j=1}^{d} D\sigma_j(r)\sigma_j(r) \end{aligned} \tag{5.2}$$

and introduce

$$\Pi(r) := Ar + \Theta(r). \tag{5.3}$$

for $r \in D(A)$. We shall provide a nice representation for generic finite dimensional realizations and the associated factor processes Z and prove a support theorem for the solution of equation (1.2).

Theorem 4.5 tells us that, under the assumption (F), there exist $\lambda_1, \dots, \lambda_{N_{LA}} \in D(A_0^\infty)$ such that

$$D_{LA}(r) = \langle \Xi(r), \lambda_1, \dots, \lambda_{N_{LA}} \rangle$$

and

$$\Sigma_j(r) \in \langle \lambda_1, \dots, \lambda_{N_{LA}} \rangle \tag{5.4}$$

for all $r \in \mathcal{V}$. Theorem 4.5 is a global result in so far as it holds for every open connected set $\mathcal{V} \subset \mathcal{U} \cap D(A^\infty)$ where (F) is satisfied. We now are interested in the question whether there exist a priori structural restrictions on the choice of $\mathcal{V}$. In view of (F) and Theorem 4.5 it is clear that $\mathcal{V}$ must not intersect with the singular set

$$\mathfrak{S} := \{h \in \mathcal{U} \cap D(A^\infty) \mid \Xi(h) \in \langle \lambda_1, \dots, \lambda_{N_{LA}-1} \rangle\}. \tag{5.5}$$

By Lemma 4.1, $\mathfrak{S}$ is closed and nowhere dense in $D(A^\infty)$.

Lemma 5.1. *If (5.4) holds on the open connected set $\mathcal{V} \subset \mathcal{U} \cap D(A^\infty)$, then*

$$\mathfrak{S} \cap \overline{\mathcal{V}} \qquad (\overline{\mathcal{V}}\text{=closure of } \mathcal{V})$$

lies in a finite-dimensional linear subspace $\mathcal{O}$ in $D(A^\infty)$ with $N_{LA} \le \dim \mathcal{O} \le N_{LA} + (N_{LA} - 1)^2$.

Proof. Since Σ is continuous, (5.4) holds on $\overline{\mathcal{V}}$. Assumption (A1) yields

$$\langle \Sigma_1(h), \dots, \Sigma_d(h)\rangle \subset D(A_0^\infty), \text{ for all } \quad h \in \mathcal{H}.$$

Hence there exists $d \le d^* \le N_{LA} - 1$ such that (after a change of coordinates if necessary) $\lambda_1, \dots, \lambda_{d^*} \in D(A_0^\infty)$, and

$$\Sigma_i(h) = \sum_{j=1}^{d^*} \beta_{ij}(h)\lambda_j, \quad 1 \le i \le d, \text{ for all } \quad h \in \overline{\mathcal{V}}, \tag{5.6}$$

for smooth functions $\beta_{ij} : \mathcal{H} \to \mathbb{R}$. Moreover, $D\Sigma_i(h)\Sigma_i(h) \in \langle \lambda_1, \dots, \lambda_{N_{LA}-1}\rangle$, for all $h \in \overline{\mathcal{V}}$. By (5.2)–(5.3) hence

$$\mathfrak{S} \cap \overline{\mathcal{V}} = \{h \in D(A^\infty) \mid \Pi(h) \in \langle \lambda_1, \dots, \lambda_{N_{LA}-1}\rangle\} \cap \overline{\mathcal{V}}. \tag{5.7}$$

Since $\Lambda_{ij} := M(\lambda_i, \lambda_j)$ is a well-defined element in $D(A^\infty)$, for all $1 \le i, j \le d^*$, we obtain

$$\Pi(h) = Ah + \sum_{i,j=1}^{d^*} a_{ij}(h)\Lambda_{ij}, \text{ for all } \quad h \in \mathcal{V}, \tag{5.8}$$

where $a_{ij}(h) := \sum_{k=1}^{d} \beta_{ki}(h)\beta_{kj}(h)$, see (5.1). Hence $h \in \mathfrak{S} \cap \overline{\mathcal{V}}$ if and only if there exist real numbers $c_1(h), \dots, c_{N_{LA}-1}(h)$ such that

$$Ah + \sum_{i,j=1}^{d^*} a_{ij}(h)\Lambda_{ij} = \sum_{i=1}^{N_{LA}-1} c_i(h)\lambda_i. \tag{5.9}$$

Let $\mathcal{R}$ be the subspace spanned by $\lambda_1, \dots, \lambda_{N_{LA}-1}$ and $\Lambda_{11}, \Lambda_{12} \dots, \Lambda_{d^*d^*}$, and let I be a set of indices (i,j) such that $\{\lambda_1, \dots, \lambda_{N_{LA}-1}, \Lambda_{ij} \mid (i,j) \in I\}$ is linear independent and spans $\mathcal{R}$. In view of (5.9) it is clear that $\mathfrak{S} \cap \overline{\mathcal{V}}$ lies in $\mathcal{O} := A^{-1}(\mathcal{R})$. Since the kernel of $A = d/dx$ is spanned by 1 (see (H1)), the dimension of $\mathcal{O}$ is $1 + \dim \mathcal{R} = N_{LA} + |I|$. □

The maximal possible choice of $\mathcal{V}$ is $\mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$. In this case we can say more about $\mathfrak{S}$.

Lemma 5.2. *Suppose that $\mathcal{V} = \mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$. Then $h \in \mathfrak{S}$ implies*

$$\mathcal{U} \cap (h + \langle \lambda_1, \dots, \lambda_{N_{LA}-1}\rangle) \subset \mathfrak{S}.$$

Proof. By Theorem 4.5 and since $[\Xi, \lambda_i]$ is a Banach map on $\mathcal{V}$ (see Lemma 3.17), we have

$$[\Xi, \lambda_i](h) = D\Xi(h)\lambda_i \in \langle \lambda_1, \dots, \lambda_{N_{LA}-1}\rangle, \tag{5.10}$$

for all $h \in \mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$, and hence for all $h \in \mathcal{U} \cap D(A^\infty)$, by smoothness of Ξ. Now let $h \in \mathfrak{S}$ and $u \in \mathbb{R}^{N_{LA}-1}$. Using Taylor's formula we calculate for $h + \sum_{i=1}^{N_{LA}-1} u_i\lambda_i \in \mathcal{U}$

$$\Xi\left(h + \sum_{i=1}^{N_{LA}-1} u_i\lambda_i\right) = \Xi(h) + \sum_{i=1}^{N_{LA}-1} u_i \int_0^1 D\Xi\left(h + t\sum_{i=1}^{N_{LA}-1} u_i\lambda_i\right)\lambda_i\, dt, \tag{5.11}$$

which lies in $\langle \lambda_1, \dots, \lambda_{N_{LA}-1}\rangle$ by (5.10), and the lemma follows. □

We now can give the classification of the corresponding finite dimensional realizations as well.

Theorem 5.3. *Suppose (F) holds on $\mathcal{V} = \mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$, where $\mathfrak{S}$ is given by (5.5). Then, for every $h_0 \in \mathcal{U} \cap D(A^\infty)$, there exists an $\mathbb{R}^{N_{LA}-1}$-valued diffusion process Z with $Z_0 = 0$ such that*

$$r_t = Fl_t^\Xi(h_0) + \sum_{i=1}^{N_{LA}-1} Z_t^i \lambda_i \tag{5.12}$$

is the unique continuous local solution to (1.1) with $r_0 = h_0$. If $h_0 \in \mathfrak{S} \cap \mathcal{U}$ we can even choose Z such that

$$r_t = h_0 + \sum_{i=1}^{N_{LA}-1} Z_t^i \lambda_i. \tag{5.13}$$

In particular, $\mathfrak{S}$ is locally invariant for (1.1).

The coordinate process Z will be explicitely constructed in the proof below (see (5.18)).

Remark 5.4. *There is a straightforward modification of the process Z such that the semiflow Fl^Ξ generated by Ξ in (5.12) can be replaced by the semiflow Fl^Π generated by Π. This follows since $\Xi(r) - \Pi(r) \in \langle \lambda_1, \dots, \lambda_d \rangle$ for all $r \in \mathcal{U} \cap D(A^\infty)$.*

Remark 5.5. *HJM models that satisfy (5.12), or (5.13), are known in the finance literature as* affine term structure models. *Hence Theorem 5.3 can be roughly reformulated in the following way:* HJM models that admit an FDR at every inital point $h_0 \in \mathcal{U} \cap D(A^\infty)$ are necessarily affine term structure models.

Affine term structure models have been extensively studied in [8], [9], [7] *(see also references therein).*

Proof. By smoothness of Σ and Ξ, (5.4) and (5.10) hold on $\mathcal{U}$ and $\mathcal{U} \cap D(A^\infty)$, respectively. Let $h_0 \in \mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$ and $\mathcal{M}_{h_0}$ a leaf of the weak foliation $\mathcal{F}$ through h_0 (see Proposition 4.4). As in the proof of Theorem 3.14 (see [14, Theorem 3.9]) we obtain a parametrization of $\mathcal{M}_{h_0}$ at h_0 by

$$\alpha(u, h_0) = Fl_{u_0}^\Xi(h_0) + \sum_{i=1}^{N_{LA}-1} u_i \lambda_i, \quad u = (u_0, \dots, u_{N_{LA}-1}) \in [0, \varepsilon) \times V, \tag{5.14}$$

for some $\varepsilon > 0$ and some open neighborhood V of 0 in $\mathbb{R}^{N_{LA}-1}$, where Fl^Ξ is the local semiflow induced by Ξ. (Strictly speaking, $\alpha(\cdot, h_0)$ is a parametrization of a submanifold with boundary of $\mathcal{M}_{h_0}$.) Now we proceed as in [9, Section 6.4] to find the appropriate coordinate process Z. Using Taylor's formula we obtain as in (5.11)

$$\begin{aligned} \Xi(\alpha(u, h_0)) &= \Xi\left(Fl_{u_0}^\Xi(h_0)\right) + \sum_{i=1}^{N_{LA}-1} \tilde{b}_i(u, h_0)\lambda_i \\ &= D\alpha(u, h_0) \cdot (1, \tilde{b}_1(u, h_0), \dots, \tilde{b}_{N_{LA}-1}(u, h_0)), \end{aligned} \tag{5.15}$$

where $\tilde{b}_i(\cdot, h_0) : [0, \varepsilon) \times V \to \mathbb{R}$ are smooth maps well specified by

$$\sum_{i=1}^{N_{LA}-1} \tilde{b}_i(u, h_0)\lambda_i := \sum_{i=1}^{N_{LA}-1} u_i \int_0^1 D\Xi\left(Fl_{u_0}^\Xi(h_0) + t \sum_{i=1}^{N_{LA}-1} u_i \lambda_i\right) \lambda_i \, dt.$$

On the other hand, we have

$$\Sigma_i(\alpha(u, h_0)) = D\alpha(u, h_0) \cdot (0, \rho_i(u, h_0), 0, \dots, 0), \quad 1 \le i \le d, \tag{5.16}$$

where $\rho_i(\cdot, h_0) = (\rho_{i1}(\cdot, h_0), \ldots, \rho_{id^*}(\cdot, h_0)) : [0, \varepsilon) \times V \to \mathbb{R}^{d^*}$ are smooth maps given by

$$\rho_{ij}(u, h_0) := \beta_{ij}(\alpha(u, h_0)),$$

see (5.6). Define the smooth map $b_i(\cdot, h_0) : [0, \varepsilon) \times V \to \mathbb{R}$ by

$$b_i(u, h_0) := \begin{cases} \tilde{b}_i(u, h_0) + \frac{1}{2} \sum_{j=1}^d D\rho_{ji}(u, h_0) \cdot \rho_j(u, h_0), & 1 \le i \le d^*, \\ \tilde{b}_i(u, h_0), & d^* < i \le N_{LA} - 1. \end{cases} \tag{5.17}$$

Then the stochastic differential equation

$$\begin{cases} dZ_t^i = b_i((t, Z_t), h_0)\, dt + \displaystyle\sum_{j=1}^d \rho_{ji}((t, Z_t), h_0)\, dW_t^j, & 1 \le i \le d^*, \\ dZ_t^i = b_i((t, Z_t), h_0)\, dt, & d^* < i \le N_{LA} - 1, \\ Z_0 = 0, \end{cases} \tag{5.18}$$

has a unique V-valued continuous local solution. By Itô's formula it follows that $r_t = \alpha((t, Z_t), h_0)$ is the unique continuous local solution to (1.1), see [9, Section 6.4], whence the theorem is proved for $h_0 \in \mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$.

Now let $h_0 \in \mathfrak{S} \cap \mathcal{U}$. By Lemma 5.2, the $(N_{LA}-1)$-dimensional affine submanifold $\mathcal{N}_{h_0} := \mathcal{U} \cap (h_0 + \langle \lambda_1, \ldots, \lambda_{N_{LA}-1} \rangle)$ lies in $\mathfrak{S}$. Since (1.5) and (1.6) are clearly satisfied for all $h \in \mathcal{M} = \mathcal{N}_{h_0}$, Theorem 1.2 gives that $\mathcal{N}_{h_0}$ is locally invariant for (1.1). Replace α in (5.14) by

$$\tilde{\alpha}(u, h_0) := h_0 + \sum_{i=1}^{N_{LA}-1} u_i \lambda_i, \quad u = (u_1, \ldots, u_{N_{LA}-1}) \in \mathbb{R}^{N_{LA}-1},$$

which is a parametrization of $\mathcal{N}_{h_0}$. A similar procedure as above yields an $\mathbb{R}^{N_{LA}-1}$-valued diffusion process Z such that $r_t = \tilde{\alpha}(Z_t, h_0)$ is the unique continuous local solution to (1.1), whence (5.13). (Notice that, by construction, Z is time-homogeneous.) Since $Fl_t^{\Xi}(h_0) \in \mathcal{N}_{h_0}$, for all $t \ge 0$ where it is defined, it is easy to modify Z such that (5.12) is satisfied too. $\square$

We remark that the form of the FDRs, (5.12) and (5.18), has already been derived in [2] and [4] under the *assumption* of (5.4) and $D_{LA} = \langle \Xi, \lambda_1, \ldots, \lambda_{N_{LA}-1} \rangle$. Above we have provided the sufficiency and *necessity* of these conditions and its consequences in a more general (and appropriate) functional-analytic setup.

5.2. **Distributional Properties.** In the sequel we shall argue why under the hypotheses of Theorem 5.3 the stopped factor process $Z_{t \wedge \tau}$ has nice distributional properties. We could argue by stochastic methods such as Malliavin Calculus (outlined in [15]), but here we argue directly by arguments on hypoelliptic differential operators as outlined in [19].

Given $U \subset \mathbb{R}^n$ open and $d+1$ smooth vector fields $V_0, \ldots, V_d : U \to \mathbb{R}^n$ and a smooth function $c : U \to \mathbb{R}$, then the second order differential operator

$$L(f) := \frac{1}{2} \sum_{i=1}^d V_i^2(f) + V_0(f) + cf$$

for smooth $f : U \to \mathbb{R}$ is said to be of *sum of the squares* type. The action of a vector field $V = (V^1, \dots, V^n)^\intercal$ on f is defined via

$$V(f)(x) = \sum_{i=1}^{n} V^i(x) \frac{\partial}{\partial x_i} f(x)$$

for $x \in U$ and $f : U \to \mathbb{R}$ smooth. The formal adjoint of L, denoted by L^*, is also of this type

$$L^* = \frac{1}{2} \sum_{i=1}^{d} V_i^2 - V_0 + \sum_{i=1}^{d} \phi_i V_i + \psi,$$

with smooth functions $\phi_1, \dots, \phi_d, \psi : U \to \mathbb{R}$, only drift and potential are changed. This is due to the formula $V^* = -V - \sum_{i=1}^n \frac{\partial}{\partial x_i} V^i(x)$, where the second term acts as multiplication operator on smooth functions.

The second order differential operator is said to be *hypoelliptic* if for all $u \in \mathcal{D}'(U)$ with $Lu \in C^\infty(U)$ the conclusion $u \in C^\infty(U)$ holds. We now state Lars Hörmander's famous Theorem on sum of the squares operators (see [19], also for the notions on distributions):

Theorem 5.6. *Given smooth vector fields $V_0, \dots, V_d : U \to \mathbb{R}^n$ and assume that the involutive distribution generated by $V_0, \dots, V_d, [V_i, V_j], \dots$ for $i, j = 0, \dots, d$ has full rank at all points of U, then the operator $L = \frac{1}{2} \sum_{i=1}^d V_i^2 + V_0 + c$ is hypoelliptic for any smooth function $c : U \to \mathbb{R}$.*

Remark 5.7. *Notice that on could replace V_0 by $-V_0 + \sum_{i=1}^d \phi_i V_i$ without changing the generated distribution. This means that – under the above assumptions – also L^* is hypoelliptic.*

Given smooth vector fields $X_0, \dots, X_d : U \to \mathbb{R}^n$. The generator of a $\mathbb{R}^n$-valued Ito diffusion Z_t on U

$$dZ_t = X_0(t, Z_t) + \sum_{i=1}^{d} X_i(t, Z_t) \circ dW_t^i,$$

written in Stratonovich form, is a sum of the squares operator namely

$$L = \frac{1}{2} \sum_{i=1}^{d} X_i^2 + X_0.$$

Theorem 5.8. *Given the assumptions of Theorem 5.3, $h_0 \in \mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$ and in particular the parametrization (5.14). For $K \subset V$ a compact neighborhood of $0 \in V$, we define the stopping time $\tau := \inf\{t \geq 0,\ Z_t \notin K\} \wedge \varepsilon$, then the solution of equation (1.1) with initial value h_0 satisfies*

$$r_t = Fl_t^{\Xi}(h_0) + \sum_{i=1}^{N_{LA}-1} Z_t^i \lambda_i$$

for $0 \leq t \leq \tau$. The distribution of the factor process $Z_{t\wedge\tau}$ can be decomposed according to

$$((Z_{t\wedge\tau})_* P)(A) := P(Z_{t\wedge\tau} \in A) = \int_A \lambda_t(x) dx + \mu_{\partial K, t}(A),$$

for $0 < t < \varepsilon$ and $A \subset K$. Here $\lambda \in C^\infty(]0, \varepsilon[\times K^\circ, \mathbb{R})$ is a non-negative function and $\mu_{\partial K, t}$ is a positive measure on K with support in ∂K for $0 \leq t < \varepsilon$. In *particular* $\mu_{\partial K,t}(A) = P(\{Z_{t\wedge\tau} \in A\} \cap \{\tau \leq t\})$ *and* $P(\{(Z_{t\wedge\tau}) \in A\} \cap \{\tau > t\}) = \int_A \lambda_t(x)dx$.

Proof. We only have to prove that $P(\{(Z_{t\wedge\tau}) \in A\} \cap \{\tau > t\})$ admits the described representation with $\lambda \in C^\infty(]0,\varepsilon[\times K^\circ, \mathbb{R})$. We first define a time dependent distribution on K°, namely

$$u(t,f) := E(f(Z_{t\wedge\tau})|\ \tau > t)P(\tau > t)$$

for any test functions $f \in \mathcal{D}(K^\circ)$ and $t \in [0,\varepsilon[$. For the notations of distributions see [19]. We observe that $u(0) = \delta_0$, the Dirac distribution at $0 \in K^\circ \subset V$. From classical stochastic analysis we know that

$$\frac{\partial}{\partial t}E(f(Z_{\tau\wedge t})) = E((L_t f)(Z_{\tau\wedge t})),$$

where L_t denotes the (time-dependent) generator of the process $Z_{\tau\wedge t}$ on K° and $f \in \mathcal{D}(K^\circ)$. Translated to distributions we obtain

$$\frac{\partial}{\partial t}u(t) = L_t^* u_t$$

for $t \in [0,\varepsilon[$, since the support of f lies in K°. Now we can identify the operator $-\frac{\partial}{\partial t} + L_t^* = (\frac{\partial}{\partial t} + L_t)^*$ on $]0,\varepsilon[\times K^\circ$ with the adjoint of the $\tilde{L} := (\frac{\partial}{\partial t} + L_t)$. But $\tilde{L}$ is the generator of the process $[0,\varepsilon[\times V$-valued diffusion $\tilde{Z}_t$ defined by

$$\tilde{Z}_t^0 = t,\ \tilde{Z}_t^i = Z_t^i \text{ for } i = 1,\dots,N_{LA}.$$

This is seen by directly calculating the generator $\tilde{L}$ of $\tilde{Z}$, compare also the proof of Theorem 5.3. Hence we have to prove that $\tilde{L}$ is hypoelliptic. The involved vector fields $\tilde{X}_0,\dots,\tilde{X}_d$ are α-related to the vector fields $\mu,\sigma_1,\dots,\sigma_d$, where α denotes the parametrization from (5.14).

$$d\tilde{Z}_t = \tilde{X}_0(\tilde{Z}_t)dt + \sum_{i=1}^d \tilde{X}_i(\tilde{Z}_t)dW_t^i$$
$$\tilde{X}_0 \sim_\alpha \Xi, \dots \tilde{X}_d \sim_\alpha \Sigma_d$$

By the assumptions (F) the operator $\tilde{L}$ and $\tilde{L}^*$ are seen to satisfy the assumptions of Theorem 5.6 (see also Theorem 22.2.1 of [19]), since by the proof of Proposition 3.10 the Lie brackets of α-related vector fields are α-related. Therefore by $\tilde{L}^* u = 0$ and hypoellipticity of $\tilde{L}^*$ we obtain $u(t,f) = \int_K \lambda(t,x)f(x)dx$ for all $f \in \mathcal{D}(K^\circ)$ and $0 < t < \varepsilon$ with stated regularity for λ. □

5.3. **Classification of the manifolds.** We finally show that $\lambda_1,\dots,\lambda_{N_{LA}-1}$ have to satisfy a functional relation which depends on β_{ij} (see (5.6)). Let the assumptions of Theorem 5.3 be in force. As shown in the proof of Lemma 5.1 we obtain $D_{LA} = \langle \Pi, \lambda_1,\dots,\lambda_{N_{LA}-1}\rangle$ on $\mathcal{U} \cap D(A^\infty)$. Hence, as in (5.10), there exist smooth functions c_{ij} on $\mathcal{U} \cap D(A^\infty)$ such that, for all $h \in \mathcal{U} \cap D(A^\infty)$,

$$D\Pi(h)\lambda_i = A\lambda_i + \sum_{k,l=1}^{d^*} (Da_{kl}(h)\lambda_i)\,\Lambda_{kl} = \sum_{j=1}^{N_{LA}-1} c_{ij}(h)\lambda_j. \tag{5.19}$$

Here we have used the notation from the proof of Lemma 5.1, see (5.8). Now fix $h \in \mathcal{U} \cap D(A^\infty)$. Expressed as a point-wise equality for functions, (5.19) reads

$$\partial_x \left(\lambda_i(x) + \frac{1}{2} \sum_{k,l=1}^{d^*} (Da_{kl}(h)\lambda_i)\, \Lambda_k(x)\Lambda_l(x) \right) = \sum_{j=1}^{N_{LA}-1} c_{ij}(h)\lambda_j(x), \quad \forall x \in \mathbb{R}_{\geq 0},$$

where $\Lambda_i(x) := \int_0^x \lambda_i(\eta)\, d\eta$. Integration with respect to x yields

$$\partial_x \Lambda_i(x) = -\frac{1}{2} \sum_{k,l=1}^{d^*} (Da_{kl}(h)\lambda_i)\, \Lambda_k(x)\Lambda_l(x) + \sum_{j=1}^{N_{LA}-1} c_{ij}(h)\Lambda_j(x) + \lambda_i(0),$$

for all $x \in \mathbb{R}_{\geq 0}$. Thus every $h \in D(A^\infty)$ implies a system of ODEs (Riccati equations) for the functions $\Lambda_1, \ldots, \Lambda_{N_{LA}-1}$, which have to hold simultaneously for all $h \in D(A^\infty)$.

Remark 5.9. *As a simple remark we can state that the long rates $r_t^L := \lim_{x\to\infty} r_t(x)$ are deterministic in the case of a finite dimensional realization, which is due to the fact that $\lim_{x\to\infty} \lambda_i(x) = 0$, compare with* [20].

6. Applications

In the seminal papers [4] and [2] finite-dimensional realizations, in particular the Hull-White extensions of the Vasicek and CIR-model, are considered for the first time from the geometric point of view. In addition to their excellent treatment (compare Section 5 of [2] or Section 7 of [4]), we prove that the Hull-White extensions of the Vasicek and CIR model are the only 2-dimensional local HJM models and we demonstrate the importance of the corresponding singular sets. The same type of analysis can also be performed in higher dimensional cases, which will be done elsewhere. At the end of this section we provide an example of how to embed the Svensson family as a leaf of a weak foliation associated to a functional dependent volatility structure.

As in Section 5 we let $m = 0$ and here also $d = 1$ (dimension of the Brownian motion), hence $N_{LA} = 2$. We let (F) be in force on $\mathcal{V} = \mathcal{U} \cap D(A^\infty)$. Hence (5.4) tells us that

$$\Sigma(r) = \Phi(r)\lambda, \quad r \in \mathcal{U},$$

for some $\lambda \in D(A_0^\infty) \setminus \{0\}$ and a smooth map $\Phi : \mathcal{U} \to \mathbb{R}$ (which is of the form $\Phi = \phi \circ \ell$ by (A1)). Without loss of generality we can assume that $\Phi > 0$, since (A1) requires "linear independence" of Φ which here simply means $\Phi \neq 0$. We want to specify under which conditions this volatility structure admits 2-dimensional realizations and how they look like. We shall show that it has to be either of the Vasicek or CIR type. This is already done in Section 7.3 of [4], however, their special setting does not allow to treat the CIR-case.

Writing $\psi(r) := \Phi(r)(D\Phi(r) \cdot \lambda)$, we obtain for $r \in \mathcal{U} \cap D(A^\infty)$

$$\begin{aligned}
D\Sigma(r) \cdot h &= (D\Phi(r) \cdot h)\lambda \\
D\Sigma(r) \cdot \Sigma(r) &= \Phi(r)(D\Phi(r) \cdot \lambda)\lambda = \psi(r)\lambda \\
\Xi(r) &= \frac{d}{dx} r + \Phi(r)^2 \lambda \int \lambda - \frac{1}{2}\psi(r)\lambda \\
D\Xi(r) \cdot h &= \frac{d}{dx} h + 2\Phi(r)(D\Phi(r) \cdot h)\lambda \int \lambda - \frac{1}{2}(D\psi(r) \cdot h)\lambda.
\end{aligned}$$

Consequently we can calculate the Lie bracket

$$[\Xi,\Sigma](r)=\Phi(r)\frac{d}{dx}\lambda+2\Phi(r)\psi(r)\lambda\int\lambda-\frac{1}{2}\Phi(r)(D\psi(r)\cdot\lambda)\lambda-$$
$$-(D\Phi(r)\cdot\frac{d}{dx}r)\lambda-\Phi(r)^2(D\Phi(r)\cdot\lambda\int\lambda)\lambda+\frac{1}{2}\psi(r)(D\Phi(r)\cdot\lambda)\lambda.$$

As in (5.10) we have $[\Xi,\Sigma](r)\in\langle\lambda\rangle$ on $\mathcal{U}\cap D(A^\infty)$. We can divide by $\Phi(r)$ and obtain an equation

$$\frac{d}{dx}\lambda+2\psi(r)\lambda\int\lambda-\theta(r)\lambda=0$$

with some smooth function $\theta:\mathcal{U}\cap D(A^\infty)\to\mathbb{R}$. There are consequently two cases:

i) If λ and $\lambda\int\lambda$ are linearly independent in $D(A^\infty)$, then by derivation with respect to r we obtain that ψ and θ are constant, say $2\psi(r)=a$ and $\theta(r)=b$ with real numbers a and b. Defining $\Lambda:=\int\lambda$ we obtain finally a Riccati equation for Λ, which yields the CIR-type if $a\neq 0$ or the Vasicek-type if $a=0$:
$$\frac{d}{dx}\Lambda+\frac{a}{2}\Lambda^2+b\Lambda=\lambda(0),\quad \Lambda(0)=0. \tag{6.1}$$
The Ho-Lee model is considered as particular case of the Vasicek model for $b=0$.

ii) If λ and $\lambda\int\lambda$ are linearly dependent in $D(A^\infty)$, then we necessarily obtain an equation of the type
$$\frac{d}{dx}\lambda+b\lambda=0,$$
which yields that λ is vanishes identically, since otherwise λ and $\lambda\int\lambda$ are linearly independent. This case was excluded at the beginning.

Notice that by (6.1), $\lambda(0)=0$ if and only if $\lambda=0$, which is not possible. Hence a fortiori we have $\lambda(0)\neq 0$, such that by rescaling we always can assume that $\lambda(0)=1$. This observation slightly improves the discussion in Section 7.3 in [4].

By the definition of ψ we have $D\Phi^2(r)\cdot\lambda=a$, hence we obtain the following representation for Φ. We split $D(A^\infty)$ into $\mathbb{R}\lambda+E$, where $E:=\ker ev_0$. We denote by $pr:D(A^\infty)\to E$ the corresponding projection. Then

$$\Phi(h)=\sqrt{a(\mathrm{ev}_0(h))+\eta(pr(h))}, \tag{6.2}$$

where $\eta:pr(\mathcal{U}\cap D(A^\infty))\subset E\to\mathbb{R}$ is a smooth function (compare with Proposition 7.3 of [4]).

In view of (5.7) we have

$$\mathfrak{S}=\left\{h\in\mathcal{U}\cap D(A^\infty)\mid \Pi(h)=Ah+\Phi(h)^2\lambda\int\lambda\in\langle\lambda\rangle\right\},$$

see (5.5). Thus, if λ and $\lambda\int\lambda$ are linearly independent in $D(A^\infty)$ then any $h\in\mathfrak{S}$ is necessarily of the form

$$h=a_1+a_2\Lambda^2+a_3\Lambda$$

in all cases for some real numbers a_i. By the particular representation of Φ we obtain that $a_2=aa_1+g(a_3)$, where g is some smooth real function derived from

$$aa_1+\eta(a_2\Lambda^2+a_3\Lambda)=a_2.$$

By Fl^X we denote the local (semi-)flow of a vector field X on $\mathcal{U} \cap D(A^\infty)$. The leaves through r^* of the weak foliation are given by the local parametrization

$$(u_0, u_1) \mapsto Fl^{\Pi}_{u_0}(r^*) + u_1 \frac{d}{dx}\Lambda$$

if r^* does not lie in the singular set $\mathfrak{S}$. If $r^* \in \mathfrak{S}$, then the leaf is a one dimensional immersed submanifold of $\langle 1, \Lambda, \Lambda^2 \rangle$. Notice that the stochastic evolution of the factor process takes place in the u_1-component, see Theorem 5.3 and Remark 5.4.

We summarize the preceding results in the following theorem.

Theorem 6.1. *Let $\mathfrak{S}$ and $\mathcal{U}$ be as above. Assume that Σ admits a 2-dimensional realization around any initial curve $r^* \in \mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$. Then there exists $\lambda \in D(A_0^\infty)$ and a function $\Phi : \mathcal{U} \to \mathbb{R}_{>0}$ of the form (6.2) such that $\Sigma(h) = \Phi(h)\lambda$. The singular set $\mathfrak{S}$ is a (possibly empty) subset of $\langle 1, \Lambda, \Lambda^2 \rangle$, where $\Lambda = \int \lambda$ satisfies the Riccati equation (6.1). The local HJM model is an affine short rate model. That is, for every initial curve $r^* \in \mathcal{U} \cap D(A^\infty)$ there exist functions $b : \mathbb{R}_{\geq 0} \times \mathbb{R} \to \mathbb{R}$, $\theta : \mathbb{R}_{\geq 0} \to \mathbb{R}$ and a stopping time $\tau > 0$ such that*

$$r_{t\wedge\tau} = Fl^{\Pi}_{t\wedge\tau}(r^*) + R_{t\wedge\tau}\,\lambda \tag{6.3}$$

is the unique $\mathcal{U}$-valued local solution to (1.2) and the short rates $R_t = r_t(0)$ follow, locally for $t \leq \tau$, a time-inhomogeneous diffusion process

$$dR_t = b(t, R_t)\,dt + \sqrt{aR_t + \theta(t)}\,dW_t.$$

This process becomes time-homogeneous if and only if $r^* \in \mathfrak{S}$, and then $r_{t\wedge\tau} \in \mathfrak{S}$ for all $t \geq 0$.

Proof. We know that $\lambda(0) \neq 0$. Hence (6.3) follows from (5.12) and Remark 5.4. The rest of the theorem is a consequence of Theorem 5.3 and the preceding discussion. □

6.1. **The Hull-White extension of the Vasicek model.** We consider the volatility structure of the Vasicek model: $\Sigma(r)(x) = \rho\exp(-\beta x) = \rho\lambda$ with $\rho > 0$ and $\beta > 0$, for $r \in \mathcal{U} \cap D(A^\infty) = D(A^\infty)$ and $x \geq 0$. Then by the above formulas

$$[\Xi, \Sigma] = -\beta\rho\lambda.$$

The singular set $\mathfrak{S}$ is characterized by

$$\frac{d}{dx}h + \frac{\rho^2}{\beta}\exp(-\beta x)(1 - \exp(-\beta x)) = c\exp(-\beta x)$$

for some real c. Therefore a_2 is some fixed value, namely $a_2 = \frac{\rho^2}{2}$ and a_1, a_3 are arbitrary. Consequently the singular $\mathfrak{S}$ set is an affine subspace for the fixed values ρ, β:

$$h = a_1 - \frac{\rho^2}{2}\Lambda^2 + a_3\Lambda.$$

Going back to traditional notations for the Vasicek model we write

$$\Lambda(x) = \frac{1}{\beta}(1 - \exp(-\beta x))$$

$$B_V(x) = \Lambda'(x) = e^{-\beta x}$$

$$A_V(x) = b\Lambda(x) - \frac{\rho^2}{2}\Lambda(x)^2,$$

then h lies in the singular set $\mathfrak{S}$ if and only if

$$h \in A_V + \langle B_V \rangle$$

for some value b (which becomes an additional parameter in the short rate equation). The solution for r^* in the singular set reads as follows

$$\begin{aligned} r_t &= A_V + B_V R_t \\ dR_t &= (b - \beta R_t)\, dt + \rho\, dW_t, \end{aligned}$$

where $R_t = ev_0(r_t)$ denotes the short rate, which is the Vasicek short rate model.

Outside the singular set $\mathfrak{S}$ we have a 2-dimensional realization. First we calculate the deterministic part of the dynamics

$$\begin{aligned} Fl^{\Pi}_{u_0}(r^*)(x) &= S_{u_0} r^*(x) + \int_0^{u_0} S_{u_0 - s}\Big(\frac{\rho^2}{\beta}\exp(-\beta x)(1 - \exp(-\beta x)\Big)\, ds \\ &= S_{u_0} r^*(x) + \frac{\rho^2}{2}\int_0^{u_0} \frac{d}{dx}(\Lambda)^2(x + u_0 - s)\, ds \\ &= r^*(x + u_0) + \frac{\rho^2}{2}\Lambda(x + u_0)^2 - \frac{\rho^2}{2}\Lambda(x)^2. \end{aligned}$$

If we identify u_0 with the time variable t, which is possible since the stochastics only occurs in direction of B_V (see Remark 5.4), we obtain by direct calculations for (5.12)

$$\begin{aligned} r_t(x) &= r^*(x + t) + \frac{\rho^2}{2}\Lambda(x + t)^2 - \frac{\rho^2}{2}\Lambda(x)^2 + \Lambda'(x) Z_t \\ dZ_t &= -\beta Z_t\, dt + \rho\, dW_t. \end{aligned}$$

A parameter transformation yields the customary form, namely

$$R_t = e^{-\beta t} r^*(0) + \int_0^t e^{-\beta(t-s)} b(s)\, ds + Z_t.$$

This yields the following expressions:

$$\begin{aligned} A_{HWV}(t, x) &= r^*(x + t) + \frac{\rho^2}{2}\Lambda(x + t)^2 - \frac{\rho^2}{2}\Lambda(x)^2 - \\ &\quad - \Lambda'(x)\Lambda'(t) r^*(0) - \Lambda'(x)\int_0^t e^{-\beta(t-s)} b(s)\, ds \\ B_{HWV}(x) &= B_V(x) = \Lambda'(x) \\ dR_t &= (b(t) - \beta R_t)\, dt + \rho\, dW_t \\ r_t &= A_{HWV}(t) + B_{HWV} R_t \\ b(t) &= \frac{d}{dt} r^*(t) + \beta r^*(t) + \frac{\rho^2}{2\beta}(1 - \exp(-2\beta t)). \end{aligned}$$

The functions A_{HWV} and B_{HWV} are solutions of time-dependent Riccati equations constructed by geometric methods. The equation for b follows from the fact that $A_{HWV}(t, 0) = 0$.

6.2. **The Hull-White extension of the CIR model.** We proceed in the same spirit: $\Sigma(r) := \rho\sqrt{ev_0(r)}\lambda$ for $\rho > 0$. The volatility structure is defined on the convex open set $\mathcal{U} = \{ev_0(r) > \varepsilon\}$ for some $\varepsilon > 0$. The function $\Lambda := \int \lambda$ satisfies (in certain normalization) a Riccati equation, namely

$$\frac{d}{dx}\Lambda + \frac{\rho^2}{2}\Lambda^2 + \beta\Lambda = 1, \quad \Lambda(0) = 0.$$

We obtain the solution (see e.g. [9, Section 7.4.1])

$$\Lambda(x) = \frac{2(\exp(x\sqrt{\beta^2 + 2\rho^2}) - 1)}{(\sqrt{\beta^2 + 2\rho^2} - \beta)(\exp(x\sqrt{\beta^2 + 2\rho^2}) - 1) + 2\sqrt{\beta^2 + 2\rho^2}}.$$

Under this assumption we can proceed as above: the singular set $\mathfrak{S}$ is determined by the equation

$$\Pi(h) = \frac{d}{dx}h + \rho^2 ev_0(h)\Lambda\Lambda' \in \langle\lambda\rangle\,,$$

hence

$$h = a_1 + \frac{\rho^2}{2}a_1\Lambda^2 + a_3\Lambda.$$

Again a_1 and a_3 can be chosen freely, which completely determines $\mathfrak{S}$. Traditionally one writes the singular set in the following form:

$$A_{CIR} = b\Lambda$$
$$B_{CIR} = 1 - \beta\Lambda - \frac{\rho^2}{2}\Lambda^2 = \Lambda'$$

with some additional parameter b and we obtain equally that h lies in $\mathfrak{S}$ if and only if

$$h \in A_{CIR} + \langle B_{CIR}\rangle\,.$$

The short rate dynamics follows the known pattern:

$$r_t = A_{CIR} + B_{CIR}R_t$$
$$dR_t = (b - \beta R_t)\,dt + \rho\sqrt{R_t}\,dW_t$$

for $r^* \in \mathfrak{S}$. Outside the singular set we have a 2-dimensional realization. First we calculate the deterministic part, by the variation of constants formula,

$$Fl^{\Pi}_{u_0}(r^*)(x) = S_{u_0}r^*(x) + \rho^2\int_0^{u_0} Fl^{\Pi}_s(r^*)(0)(S_{u_0-s}(\Lambda'\Lambda))(x)\,ds.$$

Identifying u_0 with the time parameter yields the following formula 2-dimensional realization, which is derived by direct calculations,

$$r_t = Fl^{\Pi}_t(r^*) + \Lambda' Z_t$$
$$dZ_t = -\beta Z_t\,dt + \rho\sqrt{c(t) + Z_t}\,dW_t,$$

where $c(t) = Fl^{\Pi}_t(r^*)(0)$. The short rate is given through $R_t = c(t) + Z_t$ and

$$\begin{aligned} dR_t &= (c'(t) - \beta Z_t)\,dt + \rho\sqrt{R_t}\,dW_t \\ &= (b(t) - \beta R_t)\,dt + \rho\sqrt{R_t}\,dW_t. \end{aligned}$$

Notice that $\lambda(0) = \Lambda'(0) = 1$ by the Riccati equation and $b(t) = c'(t) + \beta c(t)$.

This formula closes the circle with the classical Hull-White extension of the CIR-model:

$$
\begin{aligned}
A_{HWCIR}(t,x) &= Fl_t^{\Pi}(r^*)(x) - c(t)\Lambda'(x) \\
B_{HWCIR} &= B_{CIR} = \Lambda' \\
dR_t &= (b(t) - \beta R_t)\,dt + \rho\sqrt{R_t}\,dW_t \\
r_t &= A_{HWCIR}(t) + B_{HWCIR}R_t \\
b(t) &= \beta c(t) + \frac{d}{dt}c(t) \\
c(t) &= r^*(t) + \rho^2\int_0^t c(s)(\Lambda\Lambda')(t-s)\,ds.
\end{aligned}
$$

Again this is a geometrical construction of solutions of time-dependent Riccati equations.

6.3. **Fitting procedures as leaves of foliations.** A popular forward curve-fitting method is the Svensson [31] family

$$G_S(x,z) = z_1 + z_2e^{-z_5x} + z_3xe^{-z_5x} + z_4xe^{-z_6x}.$$

It is shown in [10] that the only non-trivial interest rate model that is consistent with the Svensson family is of the form

$$r_t = Z_t^1 g_1 + \cdots + Z_t^4 g_4, \tag{6.4}$$

where

$$g_1(x) \equiv 1, \quad g_2(x) = e^{-\alpha x}, \quad g_3(x) = xe^{-\alpha x}, \quad g_4(x) = xe^{-2\alpha x},$$

for some fixed $\alpha > 0$. Moreover,

$$Z_t^1 \equiv Z_0^1, \quad Z_t^3 = Z_0^3 e^{-\alpha t}, \quad Z_t^4 = Z_0^4 e^{-2\alpha t} \quad (Z_0^4 \geq 0)$$

and Z^2 satisfies

$$dZ_t^2 = \left(Z_t^3 + Z_t^4 - \alpha Z_t^2\right)\,dt + \sqrt{\alpha Z_t^4}\,dW_t. \tag{6.5}$$

As above, W is a real-valued Brownian motion.

We now shall find a 2-dimensional local HJM model that is of the form (6.4) whenever $r_0 = \sum_{j=1}^4 z_j g_j$ with $z_4 \geq 0$. In view of (6.5), a candidate for Σ is given on the half space $\mathcal{U} := \{\ell > 0\}$ by

$$\Sigma(h) = \sqrt{\alpha\ell(h)}g_2,$$

where ℓ is some continuous linear functional on H with $\ell(g_1) = \ell(g_2) = \ell(g_3) = 0$ and $\ell(g_4) = 1$. Straightforward calculations show, for $h \in \mathcal{U} \cap D(A^\infty)$,

$$
\begin{aligned}
\Xi(h) &= \frac{d}{dx}h + \ell(h)g_2 - \ell(h)g_2^2 \\
[\Xi,\Sigma](h) &= -\alpha\sqrt{\alpha\ell(h)}g_2 - \frac{\ell(\Xi(h))}{2\sqrt{\alpha\ell(h)}}g_2.
\end{aligned}
$$

(the clue is that $\ell \circ \Sigma \equiv 0$). Hence indeed the Lie algebra generated by Σ and Ξ has dimension 2 on $\mathcal{U} \cap D(A^\infty) \setminus \mathfrak{S}$.

References

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

[2] T. Björk and C. Landén, *On the construction of finite-dimensional realizations for nonlinear forward rate models*, forthcoming in Finance Stochast.

[3] T. Björk, C. Landén, and L. Svensson, *Finite dimensional Markovian realizations for stochastic volatility forward rate models*, working paper, 2002.

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

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

[6] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, 1992.

[7] D. Duffie, D. Filipović, and W. Schachermayer, *Affine processes and applications in finance*, Annals of applied probability, to appear (2003).

[8] D. Duffie and R. Kan, *A yield-factor model of interest rates*, Math. Finance **6** (1996), no. 4, 379–406.

[9] D. Filipović, *Consistency problems for Heath–Jarrow–Morton interest rate models*, Springer-Verlag, Berlin, 2001.

[10] ———, *Exponential-polynomial families and the term structure of interest rates*, Bernoulli **6** (2000), no. 6, 1–27.

[11] ———, *Invariant manifolds for weak solutions to stochastic equations*, Probab. Theory Relat. Fields **118** (2000), no. 3, 323–341.

[12] D. Filipović, and J. Teichmann, *A note on the regularity of finite-dimensional realizations for evolution equations*, Journal of Functional Analysis, to appear(2002).

[13] ———, *On finite-dimensional term structure models*, working paper (2002).

[14] ———, *Existence of finite dimensional realizations for stochastic equations*, Journal of Functional Analysis, to appear (2002).

[15] Patrick Florchinger, *Malliavin Calculus with time dependent coefficients and application to nonlinear filtering*, Probab. Theory Relat. Fields **86** (1990), 203–223 (1990).

[16] Richard S. Hamilton, *The inverse function theorem of Nash and Moser*, Bull. Am. Math. Soc. **7** (1982), 65–222.

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

[18] M. Heidari and L. Wu, *Are interest rate derivatives spanned by the term structure of interest rates?*, Working paper, Fordham University, 2001.

[19] L. Hörmander, *The Analysis of Linear Partial Differential Operators I–III*, Grundlehren der mathematischen Wissenschaften **256**,**257**,**274**, Springer Verlag Berlin, 1983, 1983, 1985.

[20] Friedrich Hubalek, Irene Klein and Josef Teichmann, *A general proof of the Dybvig-Ingersoll-Ross-Theorem: Long forward rates can never fall*, Mathematical Finance, to appear (2002).

[21] J. Hull and A. White, *Pricing interest-rate derivative securities*, Rev. Finan. Stud. **3** (1990), 573–592.

[22] I. Kolář, Peter W. Michor, and J. Slovák, *Natural operations in differential geometry*, (1993).

[23] Andreas Kriegl and Peter W. Michor, *The convenient setting for global analysis*, 'Surveys and Monographs 53', AMS, Providence, 1997.

[24] Serge Lang, *Fundamentals of differential geometry*, Graduate Texts in Mathematics 191, Springer, 1999.

[25] S. G. Lobanov and O. G. Smolyanov, *Ordinary differential equations in locally convex spaces*, Russian Mathematical Surveys (1993), 97–175.

[26] A. Milian, *Stochastic viability and a comparison theorem*, Colloquium Mathematicum, 68 (1995), 297–316.

[27] D. Montgomery and L. Zippin, *Topological transformation groups*, Interscience, 1957.

[28] M. Musiela, *Stochastic PDEs and term structure models*, Journées Internationales de Finance, IGR-AFFI, La Baule, 1993.

[29] R. Nagel(ed.), *One-parameter semigroups*, Springer-Verlag, Berlin-New York-Tokyo, 2000.

[30] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, Appl. Math. Sci., vol. 44, Springer-Verlag, 1983.

[31] L. E. O. Svensson, *Estimating and interpreting forward interest rates: Sweden 1992-1994*, IMF Working Paper No. 114, September 1994.

[32] Josef Teichmann, *Convenient Hille-Yosida theory*, forthcoming in Revista Mathematica Complutense.

[33] ———, *A Frobenius theorem on convenient manifolds*, forthcoming in Monatshefte für Mathematik.

[34] O. Vasicek, *An equilibrium characterization of the term structure*, J. Finan. Econom. **5** (1977), 177–188.

[35] J. Zabczyk, *Stochastic invariance and consistency of financial models*, Preprint, Polish Academy of Sciences, Warsaw, 1999.

Damir Filipović, Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544, USA. Josef Teichmann, Institute of financial and actuarial mathematics, TU Vienna, Wiedner Hauptstrasse 8-10, A-1040 Vienna, Austria

E-mail address: `dfilipov@princeton.edu, josef.teichmann@fam.tuwien.ac.at`