INVARIANT MANIFOLDS FOR WEAK SOLUTIONS TO STOCHASTIC EQUATIONS

DAMIR FILIPOVIĆ

DEPARTMENT OF MATHEMATICS, ETH, RÄMISTRASSE 101
CH-8092 ZÜRICH, SWITZERLAND
E-MAIL: FILIPO@MATH.ETHZ.CH

ABSTRACT. Viability and invariance problems related to a stochastic equation in a Hilbert space H are studied. Finite dimensional invariant C^2 submanifolds of H are characterized. We derive Nagumo type conditions and prove a regularity result: Any weak solution, which is viable in a finite dimensional C^2 submanifold, is a strong solution.

These results are related to finding finite dimensional realizations for stochastic equations. There has recently been increased interest in connection with a model for the stochastic evolution of forward rate curves.

1. INTRODUCTION

Consider a stochastic equation

$$\begin{cases} dX_t = (AX_t + F(t, X_t))\,dt + B(t, X_t)\,dW_t \\ X_0 = x_0 \end{cases} \tag{1}$$

on a separable Hilbert space H. Here W denotes a Q-Wiener process on some separable Hilbert space G. The operator A is the infinitesimal generator of a strongly continuous semigroup in H. The random mappings $F = F(t, \omega, x)$ and $B = B(t, \omega, x)$ satisfy appropriate measurability conditions.

This paper studies the *stochastic viability* and *invariance problem* related to equation (1) for finite dimensional regular manifolds in H. A subset $\mathcal{M}$ of H is called *(locally) invariant* for (1), if for any space-time initial point $(t_0, x_0) \in \mathbb{R}_+ \times \mathcal{M}$ any solution $X^{(t_0, x_0)}$ to (1) is *(locally) viable* in $\mathcal{M}$, i.e. stays (locally) on $\mathcal{M}$ almost surely.

Our purpose is to characterize invariant *finite dimensional C^2 submanifolds* $\mathcal{M}$ of H in terms of A, F and B. Under mild conditions on F and B it is shown that $\mathcal{M}$ lies necessarily in the domain $D(A)$ of A. The Nagumo type *consistency conditions*

$$\mu(t, \omega, x) := Ax + F(t, \omega, x) - \frac{1}{2}\sum_j DB^j(t, \omega, x)B^j(t, \omega, x) \in T_x\mathcal{M} \tag{2}$$

$$B^j(t, \omega, x) \in T_x\mathcal{M}, \quad \forall j, \tag{3}$$

equivalent to local invariance of $\mathcal{M}$, are derived. Here $B^j := \sqrt{\lambda_j}Be_j$, where $\{\lambda_j, e_j\}$ is an eigensequence defined by Q. If moreover $\mathcal{M}$ is closed we can prove that local invariance implies invariance.

The vector fields $\mu(t, \omega, \,.\,)$ and $B^j(t, \omega, \,.\,)$ are shown to be continuous on $\mathcal{M}$. It turns out that the stochastic invariance problem related to (1) is equivalent to the set of deterministic invariance problems related to $\mu(t, \omega, \,.\,)$ and $B^j(t, \omega, \,.\,)$.

By a solution to (1) we mean a *weak* solution, which in contrast to a *strong* solution is the natural concept for stochastic equations in Hilbert spaces (see below for the exact definitions). It is well known that under some Lipschitz and linear growth conditions on F and B, a classical fixed point argument ensures the existence of a unique weak solution X to (1), see [6]. But there is nothing comparable for the existence of a strong solution, due to the discontinuity of A. However we derive the following *regularity result*: If $\dim H < \infty$, any linear operator on H is continuous and the concepts of a weak and strong solution to (1) coincide (at least if $\dim G < \infty$). We show that this is also true in the general case provided that X is viable in a finite dimensional C^2 submanifold of H.

The main idea behind the proofs is the fact that locally any finite dimensional submanifold $\mathcal{M}$ of H can be projected diffeomorphically onto a finite dimensional linear subspace of $D(A^\star)$, where $A^\star$ is the adjoint of A. Thus any weak solution viable in $\mathcal{M}$ is locally mapped onto a finite dimensional semimartingale, which can be "pulled back" by Itô's formula (this is why we assumed $\mathcal{M}$ to be C^2). This way equation (1) is locally transformed into a finite dimensional stochastic equation, which in our setting has a unique strong solution. We use implicitly that $A^\star$ restricted to a finite dimensional subset of its domain $D(A^\star)$ is bounded. Therefore this idea cannot be directly extended to infinite dimensional invariant submanifolds.

Many phenomena, say in physics or economics, are described by stochastic equations of the form (1). The Hilbert space H is typically a space of functions. Suppose (1) represents a stochastic model. Let $\mathcal{G} := \{G(\cdot, z) \mid z \in \mathcal{Z}\}$ be a parametrized family of functions in H, for some parameter set $\mathcal{Z} \subset \mathbb{R}^m$. Assume that $\mathcal{G}$ is used for statistical estimation of the model coefficients A, F and B in (1). That is, one observes the $\mathcal{Z}$-valued process Z which is related to X by

$$G(\cdot, Z_t) = X_t. \tag{4}$$

The following problems naturally arise:

  i) What are the sufficient and necessary conditions on $\mathcal{G}$ such that there exists a non-trivial $\mathcal{Z}$-valued process Z satisfying (4)? We then say that the model is consistent with $\mathcal{G}$.

 ii) If so, what kind of dynamics does Z inherit from X by (4)? (Notice that X is a priori a weak solution to (1), hence not a semimartingale in H.)

If G is regular enough, then $\mathcal{G}$ is an m-dimensional C^2 submanifold in H. Accordingly our results apply: Problem i) is solved by the consistency conditions (2) and (3). These can be expressed in local coordinates involving the first and second order derivatives of G, making them feasible for applications. By our regularity result, Z is an Itô process. This responds to Problem ii).

Similar invariance problems have been studied in [3], [4] and [17]. In [3] and [4] the consistency problem for HJM models is solved. However they work with strong solutions in a Hilbert space, which in general is a rather cutting restriction. In [17] the invariance question for finite dimensional linear subspaces with respect to Ornstein–Uhlenbeck processes is completely solved and applied to various interest rate models. But the methods used in [17] have not yet been established for general equations (1). Therefore our results cannot straightly been obtained that way.

Further recent results can be found in [2], [12] for finite dimensional systems, and in [11], [16] for infinite dimensional ones. See also the references therein. Compared to these studies we put very mild assumptions on the coefficients of (1), which is due to the strong structure of $\mathcal{M}$.

The paper is organized as follows. In Section 2 we give the exact setting for equation (1) and define (local) weak and strong solutions. Some classical results on stochastic equations are presented. Section 3 contains the main results on stochastic viability and invariance for finite dimensional C^2 submanifolds. The proofs are postponed to Sections 4 and 5, where we also recall Itô's formula for our setting. We sketch an application of the obtained results for HJM models in Section 6. The appendix includes a summary on finite dimensional submanifolds in a Hilbert space. Their crucial properties are deduced.

2. Preliminaries on stochastic equations

For the stochastic background and notations we refer to [6]. We are given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ together with a normal filtration $(\mathcal{F}_t)_{0 \leq t < \infty}$. Let H and G be separable Hilbert spaces and $Q \in L(G)$ a self-adjoint strictly positive operator. Set $G_0 := Q^{1/2}(G)$, equipped with the scalar product $\langle u, v \rangle_{G_0} := \langle Q^{-1/2}u, Q^{-1/2}v \rangle_G$. We assume W in (1) to be a Q-Wiener process on G and $\operatorname{Tr} Q < \infty$. Otherwise there always exists a separable Hilbert space $G_1 \supset G$ on which W has a realization as a finite trace class Wiener process, see [6, Chapt. 4.3].

Denote by $L_2^0 := L_2(G_0; H)$ the space of Hilbert–Schmidt operators from G_0 into H, which itself is a separable Hilbert space. We will focus on the stochastic equation (1) under the following set of (standard) assumptions.

- The operator A is the infinitesimal generator of a strongly continuous semigroup in H.
- The mappings F and B are measurable from $(\mathbb{R}_+ \times \Omega \times H, \mathcal{P} \otimes \mathcal{B}(H))$ into $(H, \mathcal{B}(H))$, resp. $(L_2^0, \mathcal{B}(L_2^0))$.
- The initial value is a non random point $x_0 \in H$.

Two processes Y and Z are *indistinguishable* if $\mathbb{P}[Y_t = Z_t, \ \forall t < \infty] = 1$. We will not distinguish between them.

There is an equivalent way of looking at equation (1) which will be used here. Let $\{e_j\}$ be an orthonormal basis of eigenvectors of Q, such that $Qe_j = \lambda_j e_j$ for a bounded sequence of strictly positive real numbers λ_j. Then $\{\sqrt{\lambda_j}e_j\}$ is an orthonormal basis of G_0 and W has the expansion

$$W = \sum_j \sqrt{\lambda_j}\beta^j e_j, \tag{5}$$

where $\{\beta^j\}$ is a sequence of real independent $(\mathcal{F}_t)$-Brownian motions. This series is convergent in the space of G-valued continuous square integrable martingales $\mathcal{M}_T^2(G)$, for all $T < \infty$.

Recall the following classical result on stochastic integration, see [6, Chapt. 4].

Proposition 2.1. *Let E be a separable Hilbert space and σ_t^j be E-valued predictable processes, such that $\sum_j \|\sigma_t^j(\omega)\|_E^2 < \infty$ for all $(t, \omega) \in \mathbb{R}_+ \times \Omega$ and*

$$\mathbb{P}\Big[\int_0^t \sum_j \|\sigma_s^j\|_E^2 \, ds < \infty \Big] = 1, \quad \forall t < \infty.$$

Then the series of stochastic integrals

$$M_t := \sum_j \int_0^t \sigma_s^j \, d\beta_s^j$$

converges uniformly on compact time intervals in probability. Moreover M_t is an E-valued continuous local martingale and for any bounded stopping time τ

$$M_{t\wedge\tau} = \sum_j \int_0^{t\wedge\tau} \sigma_s^j \, d\beta_s^j = \sum_j \int_0^t \sigma_s^j 1_{[0,\tau]}(s) \, d\beta_s^j.$$

Set $B^j(t,\omega,x) := \sqrt{\lambda_j} B(t,\omega,x) e_j \in H$ and take into account the identity

$$\|B(t,\omega,x)\|_{L_2^0}^2 = \sum_j \|B^j(t,\omega,x)\|_H^2 < \infty. \tag{6}$$

Then equation (1) may be equivalently rewritten in the form

$$\begin{cases} dX_t = (AX_t + F(t,X_t)) \, dt + \sum_j B^j(t,X_t) \, d\beta_t^j \\ \\ X_0 = x_0. \end{cases} \tag{1'}$$

Definition 2.2. *An H-valued predictable process X is a* local weak solution *to (1), resp. (1'), if there exists a stopping time $\tau > 0$, called the* lifetime *of X, such that*

$$\mathbb{P}\Big[\int_0^{t\wedge\tau} \Big(\|X_s\|_H + \|F(s,X_s)\|_H + \|B(s,X_s)\|_{L_2^0}^2\Big) ds < \infty\Big] = 1, \quad \forall t < \infty$$

and for any $t < \infty$ and $\zeta \in D(A^\star)$, $\mathbb{P}$-a.s.

$$\begin{aligned} \langle \zeta, X_{t\wedge\tau}\rangle &= \langle \zeta, x_0\rangle + \int_0^{t\wedge\tau} \Big(\langle A^\star\zeta, X_s\rangle + \langle\zeta, F(s,X_s)\rangle\Big) ds \\ &+ \int_0^{t\wedge\tau} \langle\zeta, B(s,X_s)\,dW_s\rangle, \end{aligned} \tag{7}$$

where

$$\int_0^{t\wedge\tau} \langle\zeta, B(s,X_s)\,dW_s\rangle = \sum_j \int_0^{t\wedge\tau} \langle\zeta, B^j(s,X_s)\rangle \, d\beta_s^j.$$

A local weak solution X is a local strong solution *to (1), resp. (1'), if in addition*

$$X_{t\wedge\tau} \in D(A), \quad dt \otimes d\mathbb{P}\text{-a.s.}$$

$$\mathbb{P}\Big[\int_0^{t\wedge\tau} \|AX_s\|_H \, ds < \infty\Big] = 1, \quad \forall t < \infty$$

and for any $t < \infty$, $\mathbb{P}$-a.s.

$$X_{t\wedge\tau} = x_0 + \int_0^{t\wedge\tau} \Big(AX_s + F(s,X_s)\Big) ds + \int_0^{t\wedge\tau} B(s,X_s)\,dW_s,$$

where

$$\int_0^{t\wedge\tau} B(s,X_s)\,dW_s = \sum_j \int_0^{t\wedge\tau} B^j(s,X_s)\,d\beta_s^j.$$

If $\tau = \infty$ we just refer to X as weak, *resp.* strong solution.

This definition may be straightly extended to random $\mathcal{F}_0$-measurable initial values x_0. Notice that the lifetime τ is by no means maximal.

Lemma 2.3. *Let X be a weak solution to (1) and τ be a bounded stopping time. Then $X_{\tau+t}$ is a weak solution to*

$$\begin{cases} dY_t = (AY_t + F(\tau + t, Y_t)) \, dt + B(\tau + t, Y_t) \, d\tilde{W}_t \\ Y_0 = X_\tau \end{cases} \tag{8}$$

with respect to the filtration $\tilde{\mathcal{F}}_t := \mathcal{F}_{\tau+t}$, where $\tilde{W}_t := W_{\tau+t} - W_\tau$ is a Q-Wiener process with respect to $\tilde{\mathcal{F}}_t$. Moreover (5) induces the following expansion

$$\tilde{W} = \sum_j \sqrt{\lambda_j} \tilde{\beta}^j e_j, \tag{9}$$

where $\tilde{\beta}_t^j := \beta_{\tau+t}^j - \beta_\tau^j$ is a sequence of real independent $(\tilde{\mathcal{F}}_t)$-Brownian motions.

Proof. First we show that $\tilde{W}$ is a Q-Wiener process with respect to $\tilde{\mathcal{F}}_t$. Obviously $\tilde{W}_0 = 0$ and $\tilde{W}$ is continuous and $\tilde{\mathcal{F}}_t$-adapted. Now we proceed as in the proof of [13, Theorem (3.6), Chapter IV]. Let $h \in G$. Define the function $f \in C^\infty(\mathbb{R}_+ \times G; \mathbb{C})$ as follows

$$f(t, x) := \exp\left(i\langle h, x\rangle_G + \frac{t}{2}\langle Qh, h\rangle_G\right).$$

Identify $\mathbb{C} \cong \mathbb{R}^2$, then compute

$$f_t(t, x) = \frac{1}{2}f(t, x)\langle Qh, h\rangle_G \in L(\mathbb{R}; \mathbb{C})$$
$$f_x(t, x) = if(t, x)\langle h, \,\cdot\,\rangle_G \in L(H; \mathbb{C})$$
$$f_{xx}(t, x) = -f(t, x)\langle h, \,\cdot\,\rangle_G\langle h, \,\cdot\,\rangle_G \in L(H \times H; \mathbb{C})$$

and Itô's formula, see [6, Chapt 4.5], gives

$$f(t, W_t) = 1 + i \int_0^t f(s, W_s) \, d\langle h, W_s\rangle_G.$$

Since $f(t, W_t)$ is uniformly bounded on compact time intervals, $M_t := f(t, W_t)$ is a complex martingale (its real and imaginary parts are martingales). By the optional stopping theorem $M_{\tau+t}$ is a nowhere vanishing complex $\tilde{\mathcal{F}}_t$-martingale. For $0 \le s < t < \infty$

$$\mathbb{E}\left[\frac{M_{\tau+t}}{M_{\tau+s}} \mid \tilde{\mathcal{F}}_s\right] = 1.$$

Whence for $A \in \tilde{\mathcal{F}}_s$ we get

$$\mathbb{E}[1_A \exp(i\langle h, \tilde{W}_t - \tilde{W}_s\rangle_G)] = \mathbb{P}[A] \exp\left(-\frac{1}{2}(t-s)\langle Qh, h\rangle_G\right).$$

Since this is true for any $h \in G$, the increment $\tilde{W}_t - \tilde{W}_s$ is independent of $\tilde{\mathcal{F}}_s$ and has a Gaussian distribution with covariance operator $(t-s)Q$. Hence $\tilde{W}$ is a Q-Wiener process with respect to $\tilde{\mathcal{F}}_t$.

Next we claim that

$$\int_\tau^{\tau+t} \Phi_s \, dW_s = \int_0^t \Phi_{\tau+s} \, d\tilde{W}_s, \tag{10}$$

for any predictable L_2^0-valued process Φ with

$$\mathbb{P}\Big[\int_0^t \|\Phi_s\|_{L_2^0}^2\, ds < \infty\Big] = 1, \quad \forall t < \infty.$$

If Φ is elementary and τ a simple stopping time, then (10) holds by inspection. The general case follows by a well known localization procedure, see [6, Lemma 4.9].

Since X is a weak solution to (1), for $\zeta \in D(A^\star)$ we thus have

$$\langle \zeta, X_{\tau+t} \rangle = \langle \zeta, X_\tau \rangle + \int_\tau^{\tau+t} \Big(\langle A^\star \zeta, X_s \rangle + \langle \zeta, F(s, X_s) \rangle \Big)\, ds$$

$$+ \int_\tau^{\tau+t} \langle \zeta, B(s, X_s)\, dW_s \rangle$$

$$= \langle \zeta, X_\tau \rangle + \int_0^t \Big(\langle A^\star \zeta, X_{\tau+s} \rangle + \langle \zeta, F(\tau+s, X_{\tau+s}) \rangle \Big)\, ds$$

$$+ \int_0^t \langle \zeta, B(\tau+s, X_{\tau+s})\, d\tilde{W}_s \rangle.$$

Hence it follows that $X_{\tau+t}$ is a weak solution to (8). □

In view of Lemma 2.3 the following definition makes sense.

Definition 2.4. *For $(t_0, x_0) \in \mathbb{R}_+ \times H$ we denote by $X^{(t_0, x_0)}$ any (local) weak solution to*

$$\begin{cases} dX_t = (AX_t + F(t_0 + t, X_t))\, dt + B(t_0 + t, X_t)\, dW_t^{(t_0)} \\ X_0 = x_0 \end{cases} \tag{11}$$

with respect to the filtration $\mathcal{F}_t^{(t_0)} := \mathcal{F}_{t_0+t}$. Here $W_t^{(t_0)} := W_{t_0+t} - W_{t_0}$ is a Q-Wiener process with respect to $\mathcal{F}_t^{(t_0)}$.

All results in this paper for local weak solutions $X = X^{(0, x_0)}$ to (1) will straightly apply for local weak solutions $X^{(t_0, x_0)}$ to (11). Hence they are stated only for the case $t_0 = 0$.

In the sequel we assume X to be a (local) weak solution to (1). The main results of the next section follow under some regularity assumptions, which for clarity are presented here in a collected form.

(A1): (*continuity*) There exists a continuous version of X, still denoted by X.

(A2): (*differentiability*) $B(t, \omega, .) \in C^1(H; L_2^0)$ for all $(t, \omega) \in \mathbb{R}_+ \times \Omega$.
    This implies $B^j(t, \omega, .) \in C^1(H; H)$, and since

$$DB^j(t, \omega, x)B^j(t, \omega, x) = (DB(t, \omega, x)B^j(t, \omega, x))\sqrt{\lambda_j} e_j$$

hence

$$\sum_j \|DB^j(t, \omega, x)B^j(t, \omega, x)\|_H^2 \le \|DB(t, \omega, x)\|_{L(H; L_2^0)}^2 \|B(t, \omega, x)\|_{L_2^0}^2 < \infty$$

and $\sum_j DB^j(t, \omega, .)B^j(t, \omega, .) \in C^0(H; H)$, for all $(t, \omega) \in \mathbb{R}_+ \times \Omega$.

(A3): (*integrability*) For any $t < \infty$ we have

$$\mathbb{E}\Big[\int_0^t \Big(\|X_s\|_H + \|F(s, X_s)\|_H + \|B(s, X_s)\|_{L_2^0}^2 \Big) ds \Big] < \infty.$$

(A4): (*local Lipschitz*) For all $T, R < \infty$ there exists a real constant $C = C(T, R)$ such that

$$\|F(t, \omega, x) - F(t, \omega, y)\|_H + \|B(t, \omega, x) - B(t, \omega, y)\|_{L_2^0} \le C\|x - y\|_H$$

for all $(t, \omega) \in [0, T] \times \Omega$ and $x, y \in H$ with $\|x\|_H \le R$ and $\|y\|_H \le R$.

(A5): (*right continuity*) The mappings $F(t, \omega, x)$ and $B(t, \omega, x)$ are right continuous in t, for all $x \in H$ and $\omega \in \Omega$.

Finally we present a classical existence and uniqueness result for local weak solutions to equation (1).

Lemma 2.5. *Assume* **(A4)**. *Then for any x_0 there exists a local weak solution to* *(1) satisfying* **(A1)**. *Moreover any weak solution to (1) satisfying* **(A1)** *is unique.*

Proof. Let $x_0 \in H$. Set $l := 2\|x_0\|_H$ and define

$$\tilde{F}(t, \omega, x) := \begin{cases} F(t, \omega, x), & \text{if } \|x\|_H \le l \\ F(t, \omega, \frac{l}{\|x\|_H} x), & \text{if } \|x\|_H > l \end{cases}$$

and similarly $\tilde{B}(t, \omega, x)$. Then $\tilde{F}$ and $\tilde{B}$ are bounded and globally Lipschitz in x on $[0, T] \times \Omega \times H$, for all $T < \infty$. Hence by [6, Theorems 7.4 and 6.5] there exists a unique continuous weak solution $\tilde{X}$ to (1), when replacing F by $\tilde{F}$ and B by $\tilde{B}$. Define the stopping time $\tau := \inf\{t \ge 0 \mid \|\tilde{X}\|_H \ge l\}$. Then $\tau > 0$ and $X_t := \tilde{X}_{t \wedge \tau}$ is a local weak solution to (1) with lifetime τ and satisfies **(A1)**.

If X is a continuous weak solution to (1) then by the above arguments it is unique on $[0, \tau_n]$ for $n \ge 2$, where $\tau_n := \inf\{t \ge 0 \mid \|X\|_H \ge n\|x_0\|_H\}$. Now use that $\tau_n \uparrow \infty$. $\square$

3. INVARIANT MANIFOLDS

Let $\mathcal{M}$ denote an m-dimensional C^2 submanifold of H, $m \ge 1$. See the appendix for the definition and properties of a finite dimensional submanifold in a Hilbert space. The concept of a submanifold is used in more than one sense in the literature, see [5]. We therefore point out that $\mathcal{M}$ represents what is also called a *regular* or *embedded* submanifold, i.e. the topology and differentiable structure on $\mathcal{M}$ is the one induced by H. This allows for deriving global regularity and invariance results, see Theorems 1 and 2 below.

Essential for our needs is the following observation. Since H is separable, there exists a countable open covering $(U_k)_{k \in \mathbb{N}}$ of $\mathcal{M}$ and for each k a parametrization $\phi_k : V_k \subset \mathbb{R}^m \to U_k \cap \mathcal{M}$, where $\phi_k \in C_b^2(\mathbb{R}^m; H)$, see Remark A.3. Using the fact that $D(A^\star)$ is dense in H, by Proposition A.5 we can assume that for each k there exists a linearly independent set $\{\zeta_{k,1}, \ldots, \zeta_{k,m}\}$ in $D(A^\star)$ such that

$$\phi_k(\langle \zeta_{k,1}, x \rangle, \ldots, \langle \zeta_{k,m}, x \rangle) = x, \quad \forall x \in U_k \cap \mathcal{M}. \tag{12}$$

To simplify notation, we write $\langle \zeta_k, x \rangle$ instead of $(\langle \zeta_{k,1}, x \rangle, \ldots, \langle \zeta_{k,m}, x \rangle)$ and skip the index k whenever there is no ambiguity.

For the sake of readability all proofs are postponed to the following sections. First we state a (global) regularity result.

Theorem 1. *Let X be a local weak solution to (1) with lifetime τ, assume* **(A1)** *and*

$$X_{t \wedge \tau} \in \mathcal{M}, \quad \forall t < \infty. \tag{13}$$

Then there exists a stopping time $0 < \tau' \le \tau$, such that X is a local strong solution to (1) with lifetime τ'.

If in addition $\tau = \infty$ and **(A3)** *holds, then X is a strong solution to (1).*

We can give sufficient conditions for a weak solution X to be viable in $\mathcal{M}$. Notice that $\mu(t, \omega, x)$ in (2) is well defined by assuming **(A2)**.

Theorem 2. *Let X be a weak solution to (1) with $X_0 \in \mathcal{M}$. Assume* **(A1)**, **(A2)** *and* **(A4)**. *If $\mathcal{M}$ is closed, lies in $D(A)$ and satisfies the consistency conditions (2) and (3) for $dt \otimes d\mathbb{P}$-a.e. $(t, \omega) \in \mathbb{R}_+ \times \Omega$, for all $x \in \mathcal{M}$. Then (13) holds for $\tau = \infty$.*

Moreover A is continuous on $\mathcal{M}$. Consequently (2) and (3) hold for all $x \in \mathcal{M}$, for $dt \otimes d\mathbb{P}$-a.e. $(t, \omega) \in \mathbb{R}_+ \times \Omega$.

It turns out that the viability condition (13) is too weak to imply (2) and (3). Neither does it imply that $\mathcal{M}$ lies in $D(A)$. As a link between Theorems 1 and 2 we shall formulate equivalent conditions for (2) and (3). Rather than just asking for the assumptions of Theorem 1 we have to require that (13) holds for any space-time starting point (t_0, x_0).

Definition 3.1. *The submanifold $\mathcal{M}$ is called* locally invariant *for (1), if for all $(t_0, x_0) \in \mathbb{R}_+ \times \mathcal{M}$ there exists a continuous local weak solution $X^{(t_0, x_0)}$ to (11) with lifetime $\tau = \tau(t_0, x_0)$, such that*

$$X^{(t_0, x_0)}_{t \wedge \tau} \in \mathcal{M}, \quad \forall t < \infty.$$

Observe that this definition involves local existence.

Theorem 3. *Assume* **(A2)**, **(A4)** *and* **(A5)**. *Then the following conditions are equivalent:*

  i) *$\mathcal{M}$ is locally invariant for (1)*
 ii) *$\mathcal{M} \subset D(A)$ and (2), (3) hold for $dt \otimes d\mathbb{P}$-a.e. $(t, \omega) \in \mathbb{R}_+ \times \Omega$, for all $x \in \mathcal{M}$.*
iii) *$\mathcal{M} \subset D(A)$ and (2), (3) hold for all $(t, x) \in \mathbb{R}_+ \times \mathcal{M}$, for $\mathbb{P}$-a.e. $\omega \in \Omega$.*

We finally mention the important case where $\mathcal{M}$ is a linear submanifold. Then the above results are true under much weaker assumptions.

Theorem 4. *If $\mathcal{M}$ is an m-dimensional linear submanifold, then Theorems 2 and 3 remain valid* without *assuming* **(A2)** *and with $\mu(t, \omega, x)$ in (2) replaced by*

$$\nu(t, \omega, x) := Ax + F(t, \omega, x).$$

4. Proof of Theorem 1

We recall Itô's formula for our setting.

Lemma 4.1 (Itô's formula). *Assume that we are given the continuous $\mathbb{R}^m$-valued adapted process*

$$Y_t = Y_0 + \int_0^t b_s \, ds + \sum_j \int_0^t \sigma_s^j \, d\beta_s^j,$$

where b and σ^j are $\mathbb{R}^m$-valued predictable processes such that

$$\sum_j \|\sigma_t^j(\omega)\|_{\mathbb{R}^m}^2 < \infty, \quad \forall(t,\omega) \in \mathbb{R}_+ \times \Omega \tag{14}$$

$$\mathbb{P}\Big[\int_0^t \Big(\|b_s\|_{\mathbb{R}^m} + \sum_j \|\sigma_s^j\|_{\mathbb{R}^m}^2\Big) ds < \infty\Big] = 1, \quad \forall t < \infty \tag{15}$$

and Y_0 is $\mathcal{F}_0$-measurable. Let $\phi \in C_b^2(\mathbb{R}^m; H)$. Then $J_t := \frac{1}{2}\sum_j D^2\phi(Y_t)(\sigma_t^j, \sigma_t^j)$ is an H-valued predictable process,

$$\mathbb{P}\Big[\int_0^t \Big(\|D\phi(Y_s)b_s\|_H + \frac{1}{2}\sum_j \|D^2\phi(Y_s)(\sigma_s^j, \sigma_s^j)\|_H$$
$$+ \sum_j \|D\phi(Y_s)\sigma_s^j\|_H^2\Big) ds < \infty\Big] = 1, \quad \forall t < \infty, \tag{16}$$

and $\phi \circ Y$ is a continuous H-valued adapted process given by

$$(\phi \circ Y)_t = (\phi \circ Y)_0 + \int_0^t \Big(D\phi(Y_s)b_s + J_s\Big) ds + \sum_j \int_0^t D\phi(Y_s)\sigma_s^j \, d\beta_s^j. \tag{17}$$

Proof. Point-wise convergence of the series defining J_t follows from (14). Hence J_t is predictable. Integrability (16) is a direct consequence of (15). Denote the right hand side of (17) with I_t. Clearly by (16) it is a well defined continuous adapted process in H, see Proposition 2.1. Choose an orthonormal basis $\{e_i\}$ in H. It is enough to show equality (17) for each component with respect to $\{e_i\}$. Define $\phi_i(y) := \langle e_i, \phi(y)\rangle$. Then $\phi_i \in C_b^2(\mathbb{R}^m; \mathbb{R})$ and Itô's formula applies, see [6, Theorem 4.17]. It follows

$$\langle e_i, (\phi \circ Y)_t\rangle = (\phi_i \circ Y)_0 + \int_0^t \Big(D\phi_i(Y_s)b_s + \frac{1}{2}\sum_j D^2\phi_i(Y_s)(\sigma_s^j, \sigma_s^j)\Big) ds$$
$$+ \sum_j \int_0^t D\phi_i(Y_s)\sigma_s^j \, d\beta_s^j$$
$$= \langle e_i, I_t\rangle.$$

$\square$

Proof of Theorem 1. Assume first that X is a weak solution, that is $\tau = \infty$, and that **(A3)** holds. Fix $T > 0$. Following [7, Lemma (3.5)] there exists an increasing sequence of predictable stopping times $0 = T_0 \leq T_1 \leq T_2 \leq \cdots \leq T$ with $\sup_n T_n = T$, such that on each of the intervals

$$[T_n, T_{n+1}] \cap \{T_{n+1} > T_n\}$$

the process X takes values in some $U_{\alpha(n)}$ (here $[0,T] \times \{T_{n+1} > T_n\}$ is identified with $\{T_{n+1} > T_n\}$). This holds due to **(A1)**.

Now let $n \in \mathbb{N}_0$ and $k = \alpha(n)$. We will analyse X locally on U_k. For simplicity we shall skip the index k for ϕ_k, U_k, V_k and ζ_k in what follows. Define the $\mathbb{R}^m$-valued predictable processes b and σ^j as

$$b_t := \langle A^\star\zeta, X_t\rangle + \langle \zeta, F(t, X_t)\rangle$$
$$\sigma_t^j := \langle \zeta, B^j(t, X_t)\rangle,$$

where logically $\langle A^\star \zeta, X_t \rangle$ stands for $(\langle A^\star \zeta_1, X_t \rangle, \ldots, \langle A^\star \zeta_m, X_t \rangle)$. Then b and σ^j satisfy (14) and (15), see (6), and by the very definition of X

$$Y_t := \langle \zeta, X_t \rangle = \langle \zeta, x_0 \rangle + \int_0^t b_s \, ds + \sum_j \int_0^t \sigma_s^j \, d\beta_s^j.$$

Hence Itô's formula (Lemma 4.1) for $\phi \circ Y$ applies. Since moreover $Y_{(T_n+t) \wedge T_{n+1}}$ takes values in V on $\{T_{n+1} > T_n\}$

$$X_{(T_n+t) \wedge T_{n+1}} = (\phi \circ Y)_{(T_n+t) \wedge T_{n+1}}, \quad \text{on } \{T_{n+1} > T_n\}. \tag{18}$$

Now using (7), (17) and (18) we have for $\xi \in D(A^\star)$

$$\begin{aligned}
0 = \int_{T_n}^{(T_n+t) \wedge T_{n+1}} &\Big(\langle A^\star \xi, X_s \rangle + \langle \xi, F(s, X_s) \rangle \\
&\quad - \langle \xi, D\phi(Y_s) b_s + \frac{1}{2} \sum_j D^2 \phi(Y_s)(\sigma_s^j, \sigma_s^j) \rangle \Big) ds \\
+ \sum_j \int_{T_n}^{(T_n+t) \wedge T_{n+1}} &\Big(\langle \xi, B^j(s, X_s) \rangle - \langle \xi, D\phi(Y_s) \sigma_s^j \rangle \Big) d\beta_s^j.
\end{aligned} \tag{19}$$

Notice that the series in the first integral converges point-wise and defines an H-valued predictable process, see Lemma 4.1. By Proposition 2.1 the last term in (19) is a continuous local martingale with respect to the filtration $(\mathcal{F}_{T_n+t})$. Therefore

$$\langle A^\star \xi, X_t \rangle = \langle \xi, -F(t, X_t) + D\phi(Y_t) b_t + \frac{1}{2} \sum_j D^2 \phi(Y_t)(\sigma_t^j, \sigma_t^j) \rangle \tag{20}$$

on $[T_n, T_{n+1}] \cap \{T_{n+1} > T_n\}$, $dt \otimes d\mathbb{P}$-a.s. (see the proof of [9, Proposition 3.2]). By separability of H there exists a countable dense subset $\mathcal{D} \subset D(A^\star)$ such that (20) holds simultaneously for all $\xi \in \mathcal{D}$ and for all $(t, \omega) \in [T_n, T_{n+1}] \cap \{T_{n+1} > T_n\}$ outside an exceptional $dt \otimes d\mathbb{P}$-nullset. Let (t, ω) be such a point of validity. Then $\langle A^\star \xi, X_t(\omega) \rangle$ is a linear functional on $D(A^\star)$, which is bounded on $\mathcal{D}$ by (20). Hence it is bounded and therefore continuous on $D(A^\star)$. Since $A^{\star\star} = A$, see [14, Theorem 13.12], we conclude that $X_t(\omega) \in D(A)$ and – using full notation again –

$$\begin{aligned}
AX_t + F(t, X_t) = D\phi(Y_t) &\Big(\langle A^\star \zeta, X_t \rangle + \langle \zeta, F(t, X_t) \rangle \Big) \\
&+ \frac{1}{2} \sum_j D^2 \phi(Y_t)(\langle \zeta, B^j(t, X_t) \rangle, \langle \zeta, B^j(t, X_t) \rangle)
\end{aligned} \tag{21}$$

on $[T_n, T_{n+1}] \cap \{T_{n+1} > T_n\}$, $dt \otimes d\mathbb{P}$-a.s.

Similarly one shows that (compute the quadratic covariation of (19) with β^j)

$$B^j(t, X_t) = D\phi(Y_t) \langle \zeta, B^j(t, X_t) \rangle, \quad \forall j, \tag{22}$$

on $[T_n, T_{n+1}] \cap \{T_{n+1} > T_n\}$, $dt \otimes d\mathbb{P}$-a.s.

In view of **(A3)**

$$\mathbb{E}\Big[\int_0^T \Big(\|D\phi(Y_s)b_s\|_H + \frac{1}{2}\sum_j \|D^2\phi(Y_s)(\sigma_s^j,\sigma_s^j)\|_H + \sum_j \|D\phi(Y_s)\sigma_s^j\|_H^2\Big)ds\Big]$$

$$\leq C_1\mathbb{E}\Big[\int_0^T \Big(\|b_s\|_{\mathbb{R}^m} + \sum_j \|\sigma_s^j\|_{\mathbb{R}^m}^2\Big)ds\Big] \tag{23}$$

$$\leq C_2\mathbb{E}\Big[\int_0^T \Big(\|X_s\|_H + \|F(s,X_s)\|_H + \|B(s,X_s)\|_{L_2^0}^2\Big)ds\Big] < \infty,$$

where C_1 and C_2 depend only on n.

Summing up over $0 \leq n \leq N-1$ we get from (21) and (23)

$$\mathbb{E}\Big[\int_0^{T_N} \|AX_s\|_H\, ds\Big] < \infty, \quad \forall N \in \mathbb{N}.$$

Since $\lim_N \uparrow T_N = T$

$$\mathbb{P}\Big[\int_0^t \|AX_s\|_H\, ds < \infty\Big] = 1, \quad \forall t < T,$$

and by the identities (21) and (22)

$$X_t = x_0 + \int_0^t (AX_s + F(s,X_s))\, ds + \sum_j \int_0^t B^j(s,X_s)\, d\beta_s^j.$$

Since T was arbitrary the second part of the theorem is proved.

Now assume X to be a local weak solution with lifetime τ. Let $\phi : V \to U \cap \mathcal{M}$ be a parametrization in X_0 satisfying (12). By **(A1)** there exists a stopping time $T_1 > 0$, such that $X_{t\wedge\tau}$ takes values in U on $[0,T_1]$. Set $\tau' := \tau \wedge T_1$.

The rest of the proof runs as before, only estimation (23) has to be replaced. We give a sketch. Define

$$b_t := \Big(\langle A^\star\zeta, X_{t\wedge\tau'}\rangle + \langle \zeta, F(t,X_{t\wedge\tau'})\rangle\Big)\mathbf{1}_{[0,\tau']}(t)$$

$$\sigma_t^j := \langle \zeta, B^j(t,X_{t\wedge\tau'})\rangle\mathbf{1}_{[0,\tau']}(t).$$

Then

$$X_{t\wedge\tau'} = (\phi \circ Y)_t, \quad \forall t < \infty, \tag{24}$$

where

$$Y_t := \langle \zeta, X_{t\wedge\tau'}\rangle = \langle \zeta, x_0\rangle + \int_0^t b_s\, ds + \sum_j \int_0^t \sigma_s^j\, d\beta_s^j.$$

Instead of (23) we now have (15) by the very definition of X. Using (24) and Lemma 4.1 we get the relations (21) and (22) for X on $[0,\tau']$, $dt \otimes d\mathbb{P}$-a.s. In particular $X_{t\wedge\tau'} \in D(A)$, $dt \otimes d\mathbb{P}$-a.s. Hence from (16) and (21)

$$\mathbb{P}\Big[\int_0^{t\wedge\tau'} \|AX_s\|_H\, ds < \infty\Big] = 1, \quad \forall t < \infty$$

and

$$X_{t\wedge\tau'} = x_0 + \int_0^{t\wedge\tau'} \Big(AX_s + F(s,X_s)\Big)ds + \sum_j \int_0^{t\wedge\tau'} B^j(s,X_s)\, d\beta_s^j.$$

$\square$

5. Proof of Theorems 2, 3 and 4

A key step in proving Theorems 2, 3 and 4 consists in the following property.

Lemma 5.1. *Assume* **(A2)** *and* **(A4)**. *Let* $\phi : V \to U \cap \mathcal{M}$ *be a parametrization satisfying* (12). *Suppose that* $U \cap \mathcal{M} \subset D(A)$. *Then* (2) *and* (3) *hold for* $dt \otimes d\mathbb{P}$-a.e. $(t, \omega) \in \mathbb{R}_+ \times \Omega$, *for all* $x \in U \cap \mathcal{M}$ *if and only if*

$$Ax + F(t, \omega, x) = D\phi(y)\Big(\langle A^\star \zeta, x\rangle + \langle \zeta, F(t, \omega, x)\rangle\Big)$$
$$+ \frac{1}{2} \sum_j D^2\phi(y)(\langle \zeta, B^j(t, \omega, x)\rangle, \langle \zeta, B^j(t, \omega, x)\rangle) \qquad (25)$$

$$B^j(t, \omega, x) = D\phi(y)\langle \zeta, B^j(t, \omega, x)\rangle, \quad \forall j, \qquad (26)$$

where $y = \langle \zeta, x\rangle$, *for all* $x \in U \cap \mathcal{M}$, *for* $dt \otimes d\mathbb{P}$-a.e. $(t, \omega) \in \mathbb{R}_+ \times \Omega$.
Consequently A *is contiuous on* $U \cap \mathcal{M}$.

Proof. Point-wise equivalence of (3) and (26), and of (2) and the relation

$$Ax + F(t, \omega, x) - \frac{1}{2} \sum_j DB^j(t, \omega, x) B^j(t, \omega, x)$$
$$= D\phi(y)\langle \zeta, Ax + F(t, \omega, x) - \frac{1}{2} \sum_j DB^j(t, \omega, x) B^j(t, \omega, x)\rangle \quad (27)$$

follows from Lemma A.6. Assume (2) and (3) – and hence (26) and (27) – hold for $dt \otimes d\mathbb{P}$-a.e. $(t, \omega) \in \mathbb{R}_+ \times \Omega$, for all $x \in U \cap \mathcal{M}$. But both sides of (26) are continuous in x by **(A2)**. Hence there exists a $dt \otimes d\mathbb{P}$-nullset $\mathcal{N} \subset \mathbb{R}_+ \times \Omega$, such that (26) is true simoultaneously for all $x \in U \cap \mathcal{M}$ and $(t, \omega) \in \mathcal{N}^c$.

Fix $(t, \omega) \in \mathcal{N}^c$. In view of **(A2)** Proposition A.7 applies and the last summand in (25) equals

$$\frac{1}{2} \sum_j DB^j(t, \omega, x) B^j(t, \omega, x) - \frac{1}{2} D\phi(y)\langle \zeta, \sum_j DB^j(t, \omega, x) B^j(t, \omega, x)\rangle.$$

Hence (25) holds for $(t, \omega, x) \in \mathcal{N}^c \times (U \cap \mathcal{M})$ if and only if (27) does so.

It remains to show validity of (25) for all $x \in U \cap \mathcal{M}$, for $dt \otimes d\mathbb{P}$-a.e. (t, ω). We abbreviate the right hand side of (25) to $R(t, \omega, x)$. Then it reads

$$Ax = -F(t, \omega, x) + R(t, \omega, x) \qquad (28)$$

for $dt \otimes d\mathbb{P}$-a.e. (t, ω), for all $x \in U \cap \mathcal{M}$. But due to **(A4)** the right hand side of (28) is continuous in x, see (6). Hence for any (countable) sequence $x_n \to x$ in $U \cap \mathcal{M}$ we have $Ax_n \to Ax$, by the closedness of A. In other words A restricted to $U \cap \mathcal{M}$ is continuous. This establishes the lemma. $\square$

Proof of Theorem 2. Define the stopping time $\tau_0 := \inf\{t \geq 0 \mid X_t \notin \mathcal{M}\}$. By closedness of $\mathcal{M}$ and **(A1)** we have $X_{\tau_0} \in \mathcal{M}$ on $\{\tau_0 < \infty\}$. We claim $\tau_0 = \infty$.

Assume $\mathbb{P}[\tau_0 < K] > 0$ for a number $K \in \mathbb{N}$. By countability of the covering (U_k) there exists a parametrization $\phi : V \to U \cap \mathcal{M}$, satisfying (12) and $\mathbb{P}[X_{\tau_0} \in U \text{ and } \tau_0 < K] > 0$. Define the bounded stopping time $\tau_1 := \tau_0 \wedge K$ and set $Y_0 := \langle \zeta, X_{\tau_1}\rangle$. Since $\phi \in C_b^2(\mathbb{R}^m; H)$, the $\mathbb{R}^m$-valued, resp. $L_2(G_0; \mathbb{R}^m)$-valued

mappings

$$b(t, \omega, y) := \langle A^\star \zeta, \phi(y) \rangle + \langle \zeta, F(\tau_1(\omega) + t, \omega, \phi(y)) \rangle$$
$$\sigma(t, \omega, y) := \langle \zeta, B(\tau_1(\omega) + t, \omega, \phi(y))(\,.\,) \rangle$$
$$\sigma^j(t, \omega, y) := \sqrt{\lambda_j} \sigma(t, \omega, y) e_j = \langle \zeta, B^j(\tau_1(\omega) + t, \omega, \phi(y)) \rangle$$

are bounded and globally Lipschitz in y on $[0, T] \times \Omega \times \mathbb{R}^m$ for all $T < \infty$, which is due to **(A4)**. By Lemma 2.3 the process $\tilde{W}_t := W_{\tau_1 + t} - W_{\tau_1}$ is a Q-Wiener process expanded by $\{\tilde{\beta}_t^j\}$ given by (9). Hence the stochastic differential equation in $\mathbb{R}^m$

$$Y_t = Y_0 + \int_0^t b(s, Y_s)\, ds + \sum_j \int_0^t \sigma^j(s, Y_s)\, d\tilde{\beta}_s^j \tag{29}$$

has a unique continuous strong solution Y, see [6, Theorem 7.4].

Now $\mathbb{P}[Y_0 \in V \mid \tau_0 < K] \cdot \mathbb{P}[\tau_0 < K] = \mathbb{P}[Y_0 \in V \text{ and } \tau_0 < K] > 0$ and the distribution of Y_0 under $\mathbb{P}[\,.\mid \tau_0 < K]$ is regular. Hence by normality of $\mathbb{R}^m$ there exists two open sets V_0, V_1 such that $\overline{V}_0 \subset V_1 \subset \overline{V}_1 \subset V$ and $\mathbb{P}[Y_0 \in \overline{V}_0 \text{ and } \tau_0 < K] > 0$. Define the $(\tilde{\mathcal{F}}_t)$-stopping time

$$\tau_2 := \begin{cases} \inf\{t \geq 0 \mid Y_t \notin V_1\}, & \text{if } Y_0 \in \overline{V}_0 \text{ and } \tau_0 < K \\ 0, & \text{otherwise.} \end{cases} \tag{30}$$

Then by continuity of Y we have $\mathbb{P}[\tau_2 > 0] = \mathbb{P}[Y_0 \in \overline{V}_0 \text{ and } \tau_0 < K] > 0$ and

$$Y \in V \quad \text{on } [0, \tau_2] \cap \{\tau_2 > 0\}. \tag{31}$$

From the boundedness of b and σ we derive

$$\mathbb{E}\Big[\int_0^t \Big(\|b(s, Y_s)\|_{\mathbb{R}^m} + \sum_j \|\sigma^j(s, Y_s)\|_{\mathbb{R}^m}^2 \Big) ds \Big] < \infty, \quad \forall t < \infty. \tag{32}$$

Set $\tilde{X} := \phi \circ Y$. The assumptions of Lemma 4.1 are satisfied, hence

$$\tilde{X}_t = X_{\tau_1} + \int_0^t \Big(D\phi(Y_s) b(s, Y_s) + \frac{1}{2} \sum_j D^2\phi(Y_s)(\sigma^j(s, Y_s), \sigma^j(s, Y_s)) \Big) ds$$
$$+ \sum_j \int_0^t D\phi(Y_s)\sigma^j(s, Y_s)\, d\tilde{\beta}_s^j, \quad \text{on } [0, \tau_2] \cap \{\tau_2 > 0\}, \tag{33}$$

where the series in the first integral converges point-wise and defines an H-valued predictable process. Moreover by (31) we have

$$\tilde{X} \in (U \cap \mathcal{M}) \subset D(A) \quad \text{on } [0, \tau_2] \cap \{\tau_2 > 0\}. \tag{34}$$

Hence due to **(A2)** and **(A4)** Lemma 5.1 applies and

$$A\tilde{X}_t + F(\tau_1 + t, \tilde{X}_t) = D\phi(Y_t) b(t, Y_t) + \frac{1}{2} \sum_j D^2\phi(Y_t)(\sigma^j(t, Y_t), \sigma^j(t, Y_t))$$
$$B^j(\tau_1 + t, \tilde{X}_t) = D\phi(Y_t)\sigma^j(t, Y_t), \quad \forall j,$$

on $[0, \tau_2] \cap \{\tau_2 > 0\}$, $dt \otimes d\mathbb{P}$-a.s. This together with (32) and (33) implies

$$\mathbb{P}\Big[\int_0^{t \wedge \tau_2} \|A\tilde{X}_s\|_H \, ds < \infty \Big] = 1, \quad \forall t < \infty$$

$$\tilde{X}_{t \wedge \tau_2} = X_{\tau_1} + \int_0^{t \wedge \tau_2} \Big(A\tilde{X}_s + F(\tau_1 + s, \tilde{X}_s) \Big) ds + \sum_j \int_0^{t \wedge \tau_2} B^j(\tau_1 + s, \tilde{X}_s) \, d\tilde{\beta}_s^j,$$

on $\{\tau_2 > 0\}$. On the other hand we know by Lemmas 2.3 and 2.5 that $X_{\tau_1 + t}$ is the unique continuous weak solution to

$$\begin{cases} dZ_t = (AZ_t + F(\tau_1 + t, Z_t)) \, dt + \sum_j B^j(\tau_1 + t, Z_t) \, d\tilde{\beta}^j{}_t \\[2mm] Z_0 = X_{\tau_1} \end{cases}$$

Whence $\tilde{X}_t = X_{\tau_1 + t}$ on $[0, \tau_2] \cap \{\tau_2 > 0\}$. Because of (34) in particular $X_{\tau_1 + \tau_2} \in \mathcal{M}$. But by construction $\{\tau_0 < K\} = \{\tau_0 < K\} \cap \{\tau_0 = \tau_1\}$ and

$$0 < \mathbb{P}[\tau_2 > 0] \leq \mathbb{P}[\{\tau_1 + \tau_2 > \tau_1\} \cap \{\tau_0 < K\}] \leq \mathbb{P}[\tau_1 + \tau_2 > \tau_0],$$

which is absurd. Hence $\tau_0 = \infty$.

The last statement of the theorem follows from Lemma 5.1 and **(A2)**. $\qquad\square$

Proof of Theorem 3. i)$\Rightarrow$ii): Fix $(t_0, x_0) \in \mathbb{R}_+ \times \mathcal{M}$ and denote by $X = X^{(t_0, x_0)}$ a continuous local weak solution to (11) with lifetime τ. We proceed as in the proof of Theorem 1 and adapt the notation. Define the $\mathcal{F}_t^{(t_0)}$-stopping time $T_1 > 0$ with the property that $X_{t \wedge \tau'}$ takes values in $U_{\alpha(0)}$ for $\tau' := \tau \wedge T_1$. Analogously to (20) we derive the equality

$$\langle A^\star \xi, X_t \rangle = \Big\langle \xi, -F(t_0 + t, X_t) + D\phi(Y_t)\Big(\langle A^\star \zeta, X_t \rangle + \langle \zeta, F(t_0 + t, X_t) \rangle \Big) \tag{35}$$
$$+ \frac{1}{2} \sum_j D^2\phi(Y_t)(\langle \zeta, B^j(t_0 + t, X_t) \rangle, \langle \zeta, B^j(t_0 + t, X_t) \rangle) \Big\rangle,$$

for all $\xi \in D(A^\star)$ on $[0, \tau']$, $dt \otimes d\mathbb{P}$-a.s. Due to **(A4)** and **(A5)** both sides of (35) are right continuous in t. Hence for $\mathbb{P}$-a.e. ω the limit $t \downarrow 0$ exists. By (6) we can interchange the limiting $t \downarrow 0$ with the summation over j. Arguing as for (21) we conclude that $x_0 \in D(A)$ and

$$Ax_0 + F(t_0, \omega, x_0) = D\phi(y_0)\Big(\langle A^\star \zeta, x_0 \rangle + \langle \zeta, F(t_0, \omega, x_0) \rangle \Big)$$
$$+ \frac{1}{2} \sum_j D^2\phi(y_0)(\langle \zeta, B^j(t_0, \omega, x_0) \rangle, \langle \zeta, B^j(t_0, \omega, x_0) \rangle), \quad \mathbb{P}\text{-a.s.}$$

Similarly

$$B^j(t_0, \omega, x_0) = D\phi(y_0)\langle \zeta, B^j(t_0, \omega, x_0) \rangle, \quad \forall j, \quad \mathbb{P}\text{-a.s.}$$

Since (t_0, x_0) was arbitrary, Lemma 5.1 yields condition ii).

ii)$\Rightarrow$i): Let $(t_0, x_0) \in \mathbb{R}_+ \times \mathcal{M}$ and let $\phi : V \to U \cap \mathcal{M}$ be a parametrization in x_0 satisfying (12). As in the proof of Theorem 2 (setting $\tau_0 = \tau_1 = 0$) we get a

unique continuous strong solution Y to

$$Y_t = \langle \zeta, x_0 \rangle + \int_0^t \left(\langle A^\star \zeta, \phi(Y_s) \rangle + \langle \zeta, F(t_0 + s, \phi(Y_s)) \rangle \right) ds$$
$$+ \sum_j \int_0^t \langle \zeta, B^j(t_0 + s, \phi(Y_s)) \rangle \, d\beta_s^{(t_0),j}.$$

Here $\beta_t^{(t_0),j} := \beta_{t_0+t}^j - \beta_{t_0}^j$ is the sequence of $(\mathcal{F}_t^{(t_0)})$-Brownian motions related to $W_t^{(t_0)}$ by (9). The stopping time τ_2 given by (30) is strictly positive and $Y_{t \wedge \tau_2} \in V$. Analysis similar to that in the proof of Theorem 2 shows that $X_t^{(t_0,x_0)} := (\phi \circ Y)_{t \wedge \tau_2}$ is a continuous local strong solution to (11) with lifetime τ_2 and $X_t^{(t_0,x_0)} \in U \cap \mathcal{M}$, for all $t < \infty$.

ii)$\Leftrightarrow$iii): This is a direct consequence of Lemma 5.1 and **(A5)**. $\qquad\square$

Proof of Theorem 4. In view of Proposition A.5 we can choose any parametrization ϕ_k in (12) such that $D^2\phi_k \equiv 0$ on V_k. By straightforward inspection of the proofs we see that Lemma 5.1 and Theorems 2 and 3 remain valid under the conditions stated in the theorem. $\qquad\square$

6. Consistent HJM models

We shall briefly sketch how the theorems in Section 3 can be applied to the HJM models. See [3], [17] and [10] for more details.

Fix $\alpha > 3$. In [10] we introduce the appropriate model Hilbert space H_α, consisting of weakly differentiable functions $h : \mathbb{R}_+ \to \mathbb{R}$ and equipped with norm

$$\|h\|_\alpha^2 := |h(0)|^2 + \int_{\mathbb{R}_+} |h'(\eta)|^2 (1 + \eta)^\alpha \, d\eta.$$

Let A be the generator of the shift semigroup $S(t)h(\eta) = h(\eta + t)$ in H_α. Then the stochastic evolution of the forward rate curve $\{X_t(\eta) \mid \eta \in \mathbb{R}_+\}$ can be modeled by equation (1), where F and B satisfy the HJM condition

$$F(t, \omega, x)(\cdot) = \sum_j B^j(t, \omega, x)(\cdot) \int_0^{\cdot} B^j(t, \omega, x)(\eta) \, d\eta. \qquad (36)$$

Hence any good choice of B yields an HJM model.

In practice, daily observations of the forward rate curve are provided by parametrized families $\mathcal{G} = \{G(\cdot, z) \mid z \in \mathcal{Z}\}$ of smooth curves in H_α, for a parameter set $\mathcal{Z} \subset \mathbb{R}^m$. See [3] for concrete examples. If G is regular enough, we can check whether any HJM model is consistent with $\mathcal{G}$: combining the consistency conditions (2) and (3) in local coordinates with the HJM condition (36) yields an algebraic relation between G and B^j, which can be quite restrictive.

Widely used fitting procedures are the exponential-polynomial families. Their consistency with HJM models has been discussed for constant B^j in [3] and for general B^j in [9] and [8]. In the latter case however we started with a generic $\mathcal{Z}$-valued Itô process Z, which we inserted in (4). Problem ii) in Section 1, asserting the generality of that approach, left open and is now clarified by the present results. It turns out that – despite their popularity – exponential-polynomial families do not go very well with HJM models.

Appendix A. Finite dimensional submanifolds in Hilbert spaces

In this section we discuss and prove some crucial properties of finite dimensional submanifolds in Hilbert spaces. Let E be a Hilbert space. For a direct sum decomposition $E = E_1 \oplus E_2$ we denote by $\Pi_{(E_2, E_1)}$ the induced projection onto E_1. Notice that in general $\Pi_{(E_2, E_1)}$ is by no means orthogonal.

First we summarize some well known facts about finite dimensional submanifolds in E, see [1] for the general theory. For a detailed presentation we refer the reader to [10]. Let $k, m \in \mathbb{N}$. The following definition straightly extends the concept of a regular surface in $\mathbb{R}^3$.

Definition A.1. *A subset $\mathcal{M} \subset E$ is an m-dimensional C^k submanifold of E, if for all $x \in \mathcal{M}$ there is a neighborhood U in E, an open set $V \subset \mathbb{R}^m$ and a C^k map $\phi : V \to E$, such that*

  i) *$\phi : V \to U \cap \mathcal{M}$ is a homeomorphism*
  ii) *$D\phi(y)$ is one to one, for all $y \in V$.*
The map ϕ is called a parametrization *in x.*

$\mathcal{M}$ is a linear submanifold, *if for all $x \in \mathcal{M}$ there exists a linear parametrization of the form $\phi(y) = x + \sum_{i=1}^{m} y_i e_i$ in x.*

In what follows, $\mathcal{M}$ denotes an m-dimensional C^k submanifold of E. By the Inverse Mapping Theorem, see [1, Theorem 2.5.12], one shows that $\mathcal{M}$ shares in fact the characterizing property of a C^k manifold. That is, the change of parameters is a C^k diffeomorphism. Therefore the following concept is well defined.

Definition A.2. *For $x \in \mathcal{M}$ the* tangent space *to $\mathcal{M}$ at x is the subspace*

$$T_x \mathcal{M} := D\phi(y)\mathbb{R}^m, \quad y = \phi^{-1}(x),$$

where $\phi : V \subset \mathbb{R}^m \to \mathcal{M}$ is a parametrization in x.

It will be useful to extend a parametrization to the whole of $\mathbb{R}^m$. Let $x \in \mathcal{M}$ and $\phi : V \to U \cap \mathcal{M}$ be a parametrization in x. Set $y = \phi^{-1}(x)$. Since V is a neighborhood of y there exists $\epsilon > 0$ such that the open ball

$$B_{2\epsilon}(y) := \{ v \in \mathbb{R}^m \mid |y - v| < 2\epsilon \}$$

is contained in V. On $B_{2\epsilon}(y)$ one can define a function $\psi \in C^\infty(\mathbb{R}^m; [0,1])$ satisfying $\psi \equiv 1$ on $\overline{B_\epsilon(y)}$ and $\mathrm{supp}(\psi) \subset B_{2\epsilon}(y)$, see [5, Theorem (5.1), Chapt. II]. Since ϕ is a homeomorphism there exists a neighborhood U' of x in E with $\phi(B_\epsilon(y)) = U' \cap \mathcal{M}$. Set $\tilde{\phi} := \psi\phi$. Then $\tilde{\phi} \in C_b^k(\mathbb{R}^m; E)$ and $\tilde{\phi}_{|B_\epsilon(y)} = \phi_{|B_\epsilon(y)} : B_\epsilon(y) \to U' \cap \mathcal{M}$ is a parametrization in x. We have thus shown

Remark A.3. *We may and will assume that any parametrization $\phi : V \to U \cap \mathcal{M}$ extends to $\phi \in C_b^k(\mathbb{R}^m; E)$.*

Let $x \in \mathcal{M}$. Then $T_x \mathcal{M}$ is spanned by some orthonormal set $\{e_1, \ldots, e_m\}$. The following lemma is well known. Geometrically speaking, it says that locally in x, the orthogonal projection of $\mathcal{M}$ onto $T_x \mathcal{M}$ is a diffeomorphism. For a proof we refer to [10].

Lemma A.4. *There exists a parametrization $\phi : V \to U \cap \mathcal{M}$ in x such that*

$$\phi(\langle e_1, z \rangle, \ldots, \langle e_m, z \rangle) = z, \quad \forall z \in U \cap \mathcal{M}.$$

The following result is crucial for our discussion on weak solutions to stochastic equations which are viable in $\mathcal{M}$.

Proposition A.5. *Let $D \subset E$ be a dense subset. Then for any $x \in \mathcal{M}$ there exist elements $\zeta_1, \ldots, \zeta_m$ in D and a parametrization $\phi : V \to U \cap \mathcal{M}$ in x such that*

$$\phi(\langle \zeta_1, z \rangle, \ldots, \langle \zeta_m, z \rangle) = z, \quad \forall z \in U \cap \mathcal{M}.$$

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

Proof. The idea is to find a decomposition $E = E_1 \oplus E_2$, such that E_1 is "not too far" from $T_x \mathcal{M}$ and such that $E_1 = \mathrm{span}\{\zeta_1 \ldots, \zeta_m\}$ with $\zeta_1 \ldots, \zeta_m \in D$. The expression "not too far" means that $\Pi_{(E_2, E_1)|T_x \mathcal{M}} : T_x \mathcal{M} \to E_1$ is an isomorphism.

Let $\{e_1, \ldots, e_m\}$ be an orthonormal basis for $T_x \mathcal{M}$ and let $\psi : V_1 \to U_1 \cap \mathcal{M}$ the parametrization in x given by Lemma A.4. Set $y = \psi^{-1}(x)$. Since D is dense in E, there exist elements $\zeta_1, \ldots, \zeta_m$ in D with $\|\zeta_i - e_i\|_E < 2^{-m}$. Hence

$$|\langle \zeta_i, e_j \rangle| \le |\langle e_i, e_j \rangle| + |\langle \zeta_i - e_i, e_j \rangle| < 2^{-m}, \quad \text{if } i \ne j$$

and

$$|\langle \zeta_i, e_i \rangle| \ge |\langle e_i, e_i \rangle| - |\langle \zeta_i - e_i, e_i \rangle| > 1 - 2^{-m}.$$

The matrix $M := (\langle \zeta_i, e_j \rangle)_{1 \le i,j \le m}$ is therefore invertible, which follows from the theorem of Gerschgorin, see [15]. Consequently the family $\{\zeta_1, \ldots, \zeta_m\}$ is linearly independent in E. Define $E_1 := \mathrm{span}\{\zeta_1, \ldots, \zeta_m\}$ and let E_2 be its orthogonal complement. Since

$$\Pi_{(E_2, E_1)} e_j = \sum_{i=1}^{m} M_{ij} f_i, \quad \text{for } 1 \le j \le m,$$

we see that $\Pi_{(E_2, E_1)|T_x \mathcal{M}} : T_x \mathcal{M} \to E_1$ is an isomorphism. Let $I_{E_1} : E_1 \to \mathbb{R}^m$ denote the canonical isomorphism $I_{E_1} := (\langle \zeta_1, \cdot \rangle, \ldots, \langle \zeta_m, \cdot \rangle)$. Then the map $h := I_{E_1} \circ \Pi_{(E_2, E_1)} \circ \psi : V_1 \to \mathbb{R}^m$ is C^k and $Dh(y) = I_{E_1} \circ \Pi_{(E_2, E_1)} \circ D\psi(y)$ is one to one. By the Inverse Mapping Theorem there exist neighborhoods $W \subset V_1$ of y and $V \subset \mathbb{R}^m$ of $h(y)$ such that $h : W \to V$ is a C^k diffeomorphism. Since ψ is a homeomorphism there is a neighborhood $U \subset E$ of x with $\psi(W) = U \cap \mathcal{M}$. The map $\phi := \psi \circ h^{-1} : V \to U \cap \mathcal{M}$ is then the desired parametrization in x, that is $\phi^{-1} = I_{E_1} \circ \Pi_{(E_2, E_1)}$ on $U \cap \mathcal{M}$.

If $\mathcal{M}$ is linear, we choose $\psi(u) = x + \sum_{i=1}^{m} u_i e_i$. An easy computation shows that $\phi(v) = (x - (\Pi_{(E_2, E_1)|T_x \mathcal{M}})^{-1} \circ \Pi_{(E_2, E_1)} x) + \sum_{i=1}^{m}(\sum_{j=1}^{m} M_{ij}^{-1} v_j) e_i$. Setting $N := M^{-1}$ and $e_0 := x - (\Pi_{(E_2, E_1)|T_x \mathcal{M}})^{-1} \circ \Pi_{(E_2, E_1)} x$ completes the proof. $\square$

The proof of Proposition A.5 asserts the existence of a particular parametrization $\phi : V \to U \cap \mathcal{M}$ in $x \in \mathcal{M}$ related to the decomposition $E = E_1 \oplus E_2$. Actually E_2 is complemented in E simultaneously by all tangent spaces $T_z \mathcal{M}$, $z \in U \cap \mathcal{M}$.

Lemma A.6. *Let $\phi : V \to U \cap \mathcal{M}$ be a parametrization. If there exist elements $\zeta_1, \ldots, \zeta_m$ in E with the property that*

$$\phi(\langle \zeta_1, x \rangle, \ldots, \langle \zeta_m, x \rangle) = x, \quad \forall x \in U \cap \mathcal{M},$$

then $\zeta_1, \ldots, \zeta_m$ are linearly independent in E and

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

where E_2 is the orthogonal complement of $\mathrm{span}\{\zeta_1, \ldots, \zeta_m\}$. Moreover the induced projections are given by

$$\Pi_{(E_2, T_x \mathcal{M})} = D\phi(y)(\langle \zeta_1, \cdot \rangle, \ldots, \langle \zeta_m, \cdot \rangle), \quad y = \phi^{-1}(x), \quad \forall x \in U \cap \mathcal{M}. \qquad (37)$$

Proof. Define $\Lambda := (\langle\zeta_1,\cdot\rangle,\ldots,\langle\zeta_m,\cdot\rangle) \in L(E;\mathbb{R}^m)$. Fix $x \in U \cap \mathcal{M}$ and let $y = \phi^{-1}(x)$. By assumption $\Lambda \circ \phi(v) = v$ on V, therefore $\Lambda \circ D\phi(y) = Id_{\mathbb{R}^m}$. It follows that $\Lambda_{|T_x\mathcal{M}} : T_x\mathcal{M} \to \mathbb{R}^m$ is an isomorphism. Consequently $\zeta_1,\ldots,\zeta_m$ are linearly independent in E.

Another consequence is that $\Lambda : E \to \mathbb{R}^m$ is onto. Now the map $\Pi := D\phi(y) \circ \Lambda$ satisfies (i) $\Pi \in L(E)$, (ii) $\mathcal{N}(\Pi) = \mathcal{N}(\Lambda) = E_2$, (iii) $\mathcal{R}(\Pi) = T_x\mathcal{M}$ and (iv) $\Pi^2 = \Pi$, hence $E = T_x\mathcal{M} \oplus E_2$ and Π is the corresponding projection. Since $x \in U \cap \mathcal{M}$ was arbitrary, this completes the proof. □

Finally we derive a result which will prove crucial for Itô calculus on submanifolds. In the remainder of this section we assume $\mathcal{M}$ to be a C^2 submanifold. Let $B \in C^1(E)$ have the property that

$$B(x) \in T_x\mathcal{M}, \quad \forall x \in \mathcal{M}.$$

Fix $x \in \mathcal{M}$ and let $\phi : V \to U \cap \mathcal{M}$ be a parametrization in x. Set $y = \phi^{-1}(x)$. For $\delta > 0$ small enough, the curve $c : (-\delta,\delta) \to U \cap \mathcal{M}$

$$c(t) := \phi(y + tD\phi(y)^{-1}B(x)), \quad t \in (-\delta,\delta),$$

is well defined and satisfies $c \in C^1((-\delta,\delta);E)$, $c(0) = x$ and $c'(0) = B(x)$. Hence

$$\frac{d}{dt}B(c(t))_{|t=0} = DB(x)B(x). \tag{38}$$

Suppose now that ϕ satisfies the assumptions of Lemma A.6. To shorten notation, we write $\langle\zeta,\cdot\rangle$ instead of $(\langle\zeta_1,\cdot\rangle,\ldots,\langle\zeta_m,\cdot\rangle)$. By (37)

$$B(c(t)) = D\phi(\langle\zeta,c(t)\rangle)\langle\zeta,B(c(t))\rangle, \quad \forall t \in (-\delta,\delta).$$

Differentiation with respect to t gives

$$\frac{d}{dt}B(c(t))_{|t=0} = D^2\phi(y)(\langle\zeta,B(x)\rangle,\langle\zeta,B(x)\rangle) + D\phi(y)\langle\zeta,DB(x)B(x)\rangle. \tag{39}$$

Combining (37), (38) and (39) we get the following proposition.

Proposition A.7. *Let $\mathcal{M}$, B and $\phi : V \to U \cap \mathcal{M}$ be as above. Then*

$$DB(x)B(x) = D\phi(y)\langle\zeta,DB(x)B(x)\rangle + D^2\phi(y)(\langle\zeta,B(x)\rangle,\langle\zeta,B(x)\rangle),$$

$y = \phi^{-1}(x)$, *is the decomposition according to $E = T_x\mathcal{M} \oplus E_2$, for all $x \in U \cap \mathcal{M}$.*

REFERENCES

[1] R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, tensor analysis, and applications*, Springer-Verlag, Berlin-Heidelberg-New York, 1988.

[2] M. Bardi and P. Goatin, *Invariant sets for controlled degenerate diffusions: a viscosity solutions approach*, Stochastic analysis, control, optimization and applications, Birkhäuser Boston, Boston, MA, 1999, pp. 191–208.

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

[4] T. Björk and L. Svensson, *On the existence of finite dimensional realizations for nonlinear forward rate models*, Working paper, Stockholm School of Economics, submitted, 1997.

[5] W. M. Boothby, *An introduction to differentiable manifolds and Riemannian geometry*, second ed., Academic Press, 1986.

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

[7] M. Emery, *Stochastic calculus in manifolds*, Springer-Verlag, Berlin-Heidelberg-New York, 1989.

[8] D. Filipović, *Exponential-polynomial families and the term structure of interest rates*, to appear in BERNOULLI.

[9] ———, *A note on the Nelson–Siegel family*, Math. Finance **9** (1999), 349–359.

[10] ———, *Consistency problems for HJM interest rate models*, Ph.D. thesis, ETH Zürich, 2000.

[11] W. Jachimiak, *Stochastic invariance in infinite dimension*, Working paper, Polish Academy of Sciences, 1998.

[12] A. Milian, *Invariance for stochastic equations with regular coefficients*, Stochastic Anal. Appl. **15** (1997), no. 1, 91–101.

[13] D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Grundlehren der mathematischen Wissenschaften, vol. 293, Springer-Verlag, Berlin-Heidelberg-New York, 1994.

[14] W. Rudin, *Functional analysis*, second ed., McGraw-Hill, 1991.

[15] J. Stoer and R. Bulirsch, *Einführung in die Numerische Mathematik*, second ed., vol. II, Springer-Verlag, Berlin-Heidelberg-New York, 1978.

[16] G. Tessitore and J. Zabczyk, *Comments on transition semigroups and stochastic invariance*, Preprint, 1998.

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