Diss. ETH No. 13603

Consistency Problems for HJM Interest Rate Models

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
ZURICH

for the degree of
Dr. sc. math.

presented by
DAMIR FILIPOVIĆ
Dipl. Math. ETH
born March 26, 1970
citizen of Schmiedrued AG

accepted on the recommendation of
Prof. Dr. F. Delbaen, examiner
Prof. Dr. P. Embrechts, co-examiner

2000

Contents

Abstract iii

Introduction 1

Part 1. The HJM Methodology 7

Chapter 1. Stochastic Equations in Infinite Dimensions 9
- 1.1. Infinite Dimensional Brownian Motion 9
- 1.2. The Stochastic Integral 11
- 1.3. Fundamental Tools 13
- 1.4. Stochastic Equations 16

Chapter 2. The HJM Methodology Revisited 19
- 2.1. The Bond Market 19
- 2.2. The Musiela Parametrization 20
- 2.3. Arbitrage Free Term Structure Movements 24
- 2.4. Contingent Claim Valuation 28
- A Summary of Conditions 32
- 2.5. What Is a Model? 32

Chapter 3. The Forward Curve Spaces H_w 35
- 3.1. Definition of H_w 35
- 3.2. Volatility Specification 40
- 3.3. The Yield Curve 41
- 3.4. Local State Dependent Volatility 42
- 3.5. Functional Dependent Volatility 45
- 3.6. The BGM Model 46

Chapter 4. Consistent HJM Models 51
- 4.1. Consistency Problems 51
- 4.2. A Simple Criterion for Regularity of G 53
- 4.3. Regular Exponential-Polynomial Families 54
- 4.4. Affine Term Structure 58

Part 2. Publications 63

Chapter 5. A Note on the Nelson–Siegel Family 65
- 5.1. Introduction 65
- 5.2. The Interest Rate Model 66
- 5.3. Consistent State Space Processes 67
- 5.4. The Class of Consistent Itô Processes 69

5.5. E-Consistent Itô Processes 73

Chapter 6. Exponential-Polynomial Families and the Term Structure of Interest Rates 75
6.1. Introduction 75
6.2. Consistent Itô Processes 77
6.3. Exponential-Polynomial Families 78
6.4. Auxiliary Results 81
6.5. The Case $BEP(1, n)$ 84
6.6. The General Case $BEP(K, n)$ 86
6.7. E-Consistent Itô Processes 92
6.8. The Diffusion Case 93
6.9. Applications 95
6.10. Conclusions 97

Chapter 7. Invariant Manifolds for Weak Solutions to Stochastic Equations 99
7.1. Introduction 99
7.2. Preliminaries on Stochastic Equations 101
7.3. Invariant Manifolds 105
7.4. Proof of Theorem 7.6 106
7.5. Proof of Theorems 7.7, 7.9 and 7.10 110
7.6. Consistency Conditions in Local Coordinates 113

Appendix: Finite Dimensional Submanifolds in Banach Spaces 115

Bibliography 121

Abstract

This thesis brings together estimation methods and stochastic factor models for the term structure of interest rates within the HJM framework. It is based on the complex of consistency problems introduced by Björk and Christensen [**7**].

There exist commonly used methods for fitting the current term structure. Any curve fitting method can be represented as a parametrized family of smooth curves $\mathcal{G} = \{G(\cdot, z) \mid z \in \mathcal{Z}\}$ with finite dimensional parameter set $\mathcal{Z}$. A lot of cross-sectional data z is available, and one may ask for a suitable stochastic model for z which provides accurate bond option prices. However, this requires the absence of arbitrage. We characterize all consistent $\mathcal{Z}$-valued state space Itô processes Z which, by definition, provide an arbitrage-free model when representing the parameter z. Obviously, $G(T-t, Z_t)$ has to satisfy the HJM drift condition. It turns out that selected common curve fitting methods do not go well with the HJM framework.

We then consider the preceding consistency problem from a geometric point of view, as proposed by Björk and Christensen [**7**]. We extend the HJM framework to incorporate an infinite dimensional driving Brownian motion. We then perform the change of parametrization due to Musiela [**37**] and arrive at a stochastic equation in a Hilbert space H, describing the arbitrage-free evolution of the forward curve. The family $\mathcal{G}$ can be treated as a subset of H and the above consistency considerations yield a stochastic invariance problem for the previously derived stochastic equation. Under the assumption that $\mathcal{G}$ is a regular submanifold of H, we derive sufficient and necessary conditions for its invariance. Expressed in local coordinates they turn out to equal the HJM drift condition.

Classical models, such as the Vasicek [**48**] and CIR [**17**] short rate model, and the popular BGM [**12**] LIBOR rate model, are shown to fit well into that framework. By their very definition, affine HJM models are consistent with finite dimensional linear submanifolds of H. A straight application of our results yields a complete characterization of them, as obtained by Duffie and Kan [**21**].

In conclusion, we provide a general tool for exploiting the interplay between curve fitting methods and HJM factor models.

Introduction

Bond markets differ in one fundamental aspect from standard stock markets. While the latter are built up by a finite number of traded assets, the underlying basis of a bond market is the entire term structure of interest rates: an infinite dimensional variable which is not directly observable. On the empirical side this necessitates curve fitting methods for the daily estimation of the term structure. Pricing models on the other hand, are usually built upon stochastic factors representing the term structure in a finite dimensional state space.

The aim of this thesis is threefold: to bring together estimation methods and factor models for interest rates, to provide appropriate consistency conditions and to explore some important examples.

By a *bond* with maturity T we mean a default-free zero coupon bond with nominal value 1. Its price at time t is denoted by $P(t,T)$. There is a one to one relation between the time t term structure of bond prices and the time t term structure of interest rates or *forward curve* $\{f(t,T) \mid T \geq t\}$ given by

$$P(t,T) = \exp\left(-\int_t^T f(t,s)\,ds\right).$$

Accordingly, $f(t,T)$ is the continuously compounded instantaneous *forward rate* for date T prevailing at time t. The forward curve contains all the necessary information for pricing bonds, swaps and forward rate agreements of all maturities. Furthermore, it is a vital indicator for central banks in setting monetary policy.

Several algorithms for constructing the current term structure of interest rates from market data are applied in practice. A recent document from the Bank for International Settlements (BIS [**6**]) provides an overview of all estimation methods used by some of the most important central banks, see Table 1 on page 76. Prominent among them are the curve fitting procedures by Nelson and Siegel [**39**] and Svensson [**44**] and smoothing splines. Any common method can be represented as a parametrized family $\mathcal{G} = \{G(\cdot, z) \mid z \in \mathcal{Z}\}$ of smooth curves where $\mathcal{Z} \subset \mathbb{R}^m$ is a finite dimensional parameter set ($m \in \mathbb{N}$). By appropriate choice of the parameter $z \in \mathcal{Z}$, an optimal fit of the forward curve $x \mapsto G(x,z)$ to the observed data is achieved. Here $x \geq 0$ denotes *time to maturity*. In that sense z represents the current state of the economy, taking values in the *state space* $\mathcal{Z}$.

According to [**6**] a lot of cross-sectional data, i.e. daily estimations of z, is available for selected estimation methods $\mathcal{G}$. In addition there exists an extensive literature on statistical inference for diffusion processes based on discrete time observations, see e.g. [**5**]. It therefore seems promising to explore the stochastic evolution of the parameter z continuously over time.

Henceforth we are given a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ satisfying the usual conditions. Let $Z = (Z_t)_{t\in\mathbb{R}_+}$ be a diffusion model for the above parameter z. One would think that now $f(t,T) = G(T-t, Z_t)$ provides an accurate factor model for the interest rates. However, there are constraints on the dynamics of Z, such as the requirement of absence of arbitrage. By this we mean the existence of a probability measure $\mathbb{Q} \sim \mathbb{P}$ under which discounted bond price processes follow local martingales ($\mathbb{Q}$ is called an *equivalent local martingale measure* or *risk neutral measure*). Necessary conditions can be formulated in terms of the generator of Z applied to G. These conditions turn out to be very restrictive. For G fixed we arrive at a type of *inverse problem* for the generator of Z. State space processes Z which comply with these constraints, are called *consistent* with $\mathcal{G}$. It can be shown for example that the Nelson–Siegel family admits no non-trivial consistent diffusion Z. We achieve these negative results even for generic state space Itô processes Z (thereby the inverse problem applies to the characteristics of Z instead of the generator as for diffusions).

An application of Itô's formula shows that the preceding approach is part of the framework introduced by Heath, Jarrow and Morton (henceforth HJM) [**28**] which basically unifies all continuous interest rate models. For arbitrary but fixed $T \in \mathbb{R}_+$, HJM let $f(\cdot, T)$ evolve as an Itô process

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T) + \int_0^t \sigma(s,T)\, dW_s, \quad t \in [0,T]. \tag{1}$$

Originally in [**28**], W is a finite dimensional Brownian motion. We will however extend this approach at once by considering an *infinite dimensional* Brownian motion W. This generalization is straightforward but one has to take care of the definition of the stochastic integral.

The no-arbitrage requirement cumulates in the well-known *HJM drift condition*: existence of an equivalent local martingale measure $\mathbb{Q}$ yields a representation for the $\mathbb{Q}$-drift of $f(\cdot,T)$ in terms of σ. Thus, pricing formula for interest rate sensitive contingent claims do only depend on σ. The above mentioned inverse problem for the characteristics of Z is simply this drift condition.

Note that (1) provides a generic description of the arbitrage-free evolution rather than a model for the interest rates. By HJM's result it is convenient and standard to describe the evolution of $f(\cdot,T)$ under the risk neutral measure $\mathbb{Q}$, involving only σ and the Girsanov transform $\tilde{W}$ of W.

Musiela [**37**] proposed the reparametrization

$$r_t(x) = f(t, x+t), \tag{2}$$

taking better into account the nature of the forward curve $x \mapsto r_t(x)$ as an (infinite dimensional) state variable, where $x \geq 0$ denotes time to maturity. We ask: what is the implied stochastic dynamics for $(r_t(x))_{t\in\mathbb{R}_+}$?

The semigroup theory for stochastic equations in a Hilbert space provided by Da Prato and Zabczyk [**18**] is tailor-made for that approach. We give an axiomatic scheme for the choice of an admissible Hilbert space H of functions $h : \mathbb{R}_+ \to \mathbb{R}$ in which $(r_t)_{t\in\mathbb{R}_+}$ can be realized as a solution to the stochastic equation

$$\begin{cases} dr_t = (Ar_t + F_{HJM}(t, r_t))\, dt + \sigma(t, r_t)\, d\tilde{W}_t \\ r_0 = h_0. \end{cases} \tag{3}$$

Here A is the generator of the semigroup of right-shifts $S(t)h(x) = h(x+t)$. The coefficients[1] $\sigma = \sigma(t,\omega,h)$ and $F_{HJM} = F_{HJM}(t,\omega,h)$ are random mappings. The latter is fully specified by σ according to the HJM drift condition under the measure $\mathbb{Q}$ (under $\mathbb{P}$ the market price of risk is also involved). Equation (3) with $h_0 = f(0,\cdot)$ represents the system of infinitely many Itô processes (1) as *one* dynamical system in infinite dimensions, now allowing for arbitrary initial curves $r_0 \in H$.

The axiomatic exposure of an admissible Hilbert space of forward curves is new. We propose a particular class of admissible Hilbert spaces, which are both economically reasonable and appropriate for providing existence and uniqueness results for (3) in full generality. Any Lipschitz continuous and bounded volatility coefficient σ provides an HJM model therein. Classical models are shown to fit well into that framework. These include the short rate models of Vasicek [**48**] and Cox, Ingersoll and Ross (henceforth CIR) [**17**], and also the popular LIBOR model by Brace, Gątarek and Musiela (henceforth BGM) [**12**].

Equation (3) was introduced and analyzed first in [**37**] and [**13**], but only for deterministic volatility structure. In that Gaussian framework Musiela also characterized the set of attainable forward curves by the model. This is an issue which is strongly related to the present discussion. Further aspects of equation (3) have been exploited by several authors. Among them are Björk et al. [**7**], [**8**], [**10**], Zabczyk [**50**], Vargiolu [**47**], [**46**] and Milian [**34**]. Research is ongoing. A rigorous exposure of the no-arbitrage property for equation (3) as provided by this thesis therefore seems to be needed.

In view of the previously mentioned negative consistency results for common estimation methods, one can proceed in two possible ways. Either one considers more general non-continuous (Markovian) state space processes Z, or one stays within the continuous (that is HJM) regime and looks for appropriate forward curve families $\mathcal{G}$. The first approach is not a topic of this thesis and will be exploited elsewhere, see [**26**]. The second approach leads naturally to the complex of *consistency problems for HJM models* introduced by Björk and Christensen [**7**]. They propose the following geometric point of view.

Considering $\mathcal{G}$ as a subset of H, we can study the *stochastic invariance problem* related to (3) for $\mathcal{G}$. The set $\mathcal{G}$ is called *(locally) invariant* for (3) if for any space-time initial point $(t_0,h_0) \in \mathbb{R}_+ \times \mathcal{G}$, the solution $r = r^{(t_0,h_0)}$ of (3) stays (locally) in $\mathcal{G}$ almost surely. Accordingly, we call the HJM model provided by (3) *consistent* with $\mathcal{G}$.

By a solution to (3) we mean a *mild* or *weak* solution rather than a *strong* solution. These three concepts will be defined below, see Definition 1.19. It is essential, however, to notice that a mild or a weak solution is *not* an H-valued semimartingale.

The existence of a state space diffusion process Z which is (locally) consistent with $\mathcal{G}$ for any initial point $(t_0,z_0) \in \mathbb{R}_+ \times \mathcal{Z}$ obviously implies consistency of the induced HJM model $G(\cdot,Z)$ with $\mathcal{G}$. The converse is not true in general: if the model (3) is consistent with $\mathcal{G}$ it's not clear what the $\mathcal{Z}$-valued *coordinate process* Z determined by

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

[1] By a slight abuse of notation we use the same letter σ in (3) as well as (1), although they represent different objects.

looks like. We cannot expect it to be an Itô process.

We provide a general regularity result for this problem. Suppose $\mathcal{G}$ is a finite dimensional regular submanifold of H and locally invariant for a stochastic equation like (3). Then the coordinate process Z in (4) is a $\mathcal{Z}$-valued diffusion. Furthermore, we find Nagumo-type *consistency conditions* in terms of σ and F_{HJM} and the tangent bundle of $\mathcal{G}$. Expressing these consistency conditions in local coordinates we arrive at exactly the above inverse problem for the characteristics of the coordinate (i.e. state space) process Z. That way the consistency problem for $\mathcal{G}$ is completely solved.

These results are then applied to the estimation methods provided by [**6**]. Under some technical restrictions, we show that *exponential-polynomial families*, such as those of Nelson–Siegel or Svensson, are regular submanifolds of H. Thus, the previously derived negative consistency results can be restated within the present setup.

A further application yields the characterization of all HJM models which are consistent with linear submanifolds. These are just the *affine* HJM models. We derive the well-known results of Duffie and Kan [**21**] from our general point of view.

The present consistency results have been basically established by Björk and Christensen [**7**]. However, they proceeded under much stronger assumptions which are not necessarily convenient for the HJM framework, see Remark 2.16 below. Moreover, they did not address the difference between a regular and an immersed submanifold, see Figure 1 on page 52. By exploiting the strong structure of the regular submanifold $\mathcal{G}$ we derive *global* invariance results in contrast to [**7**] where only local results are provided. New is also the representation of the consistency conditions in local coordinates, making them feasible for applications, which work mainly in one direction: given a curve fitting method $\mathcal{G}$, we can determine the consistent HJM models. The converse problem of the existence of an invariant finite dimensional submanifold, given a particular HJM model, is treated in Björk and Svensson [**10**], by using the Frobenius theorem. The non-consistency of exponential-polynomial families has been proved in [**7**] for the particular case of a deterministic volatility structure.

Similar stochastic invariance problems have been studied by Zabczyk [**50**], Jachimiak [**30**] and for finite dimensional systems by Milian [**33**] (see also the references therein). Björk et al. [**8**], [**10**] discuss further the existence of (minimal) finite dimensional realizations for (3). In [**50**] the invariance question for finite dimensional linear subspaces with respect to Ornstein–Uhlenbeck processes is resolved. Applications to HJM models, the BGM model and to a second order term structure model by Cont [**16**] are exploited. But the support theorem methods in [**50**] have not yet been established for the general equation (3). Jachimiak [**30**] exploits invariance for closed sets $K \subset H$. However, his methods cannot be applied directly to (3).

Outline. The thesis is divided into two parts. Part 2 contains a series of three articles ([**25**], [**23**], [**24**]) developed during my Ph.D. research over the last three years. All have been accepted for publication by renowned journals. Part 1 represents a synthesis of the three articles cumulating in the central Chapter 4.

In Chapter 1 we repeat the relevant material from [**18**] about stochastic analysis in infinite dimensions. We define an infinite dimensional Brownian motion W and show how it can be realized as Hilbert space valued Wiener process. The construction of the stochastic integral with respect to W is then sketched and we provide the main tools from stochastic analysis: Itô's formula, the stochastic Fubini theorem and Girsanov's theorem. We introduce stochastic equations in a Hilbert space, define the concepts for a solution and state an existence and uniqueness result.

Chapter 2 recaptures the HJM [**28**] methodology and extends it to incorporate an infinite dimensional driving Brownian motion W. We perform the change of parametrization (2) and arrive at the stochastic equation (3) in a Hilbert space H which is axiomatically scheduled. The no-arbitrage (or HJM drift) condition is rigorously exposed. We give a sketch for contingent claim valuation in the preceding bond market and discuss in particular the forward measure, forward LIBOR rates and caplets. This is introductory for the BGM model to be recaptured in the subsequent chapter. Finally, we define what we mean by an HJM model.

In Chapter 3 we introduce a class of Hilbert spaces H_w which are both economically reasonable and convenient for analyzing the previously derived stochastic equation (3). We provide a general existence and uniqueness result for Lipschitz continuous and bounded volatility coefficients σ. Classical models are shown to fit well into that framework. So is the classical HJM model which corresponds to locally state dependent volatility coefficients σ. We also recapture the popular BGM model and provide further promising examples.

Chapter 4 is central and can be considered as the main chapter. It unifies the results obtained in [**25**], [**23**] and [**24**] within the previously constructed framework. We state the main result for regular families $\mathcal{G}$ and thus solve completely the consistency problem. We discuss some pitfalls which arise from the subtle but important difference between a regular and an immersed submanifold. A simple criterion for differentiability in H_w is presented and applied to the class of exponential-polynomial families. We restate the (negative) consistency results for the common Nelson–Siegel and Svensson families. Finally, we apply our methods for identifying the class of affine term structure models which are characterized by possessing linear invariant submanifolds. We provide some of the well-known results from [**21**] and recapture the Vasicek and CIR short rate models.

Chapter 5 is [**25**] and shows that there exists no non-trivial state space Itô process which is consistent with the Nelson–Siegel family.

Chapter 6 is [**23**]. Here we extend the methods used in Chapter 5 for characterizing the class of state space Itô processes which are consistent with general exponential-polynomial families. We obtain remarkably restrictive results. In particular we identify the only non-trivial HJM model which is consistent with the Svensson family: an extended Vasicek short rate model. Chapters 5 and 6 were motivated by the examples given in [**7**], see Remarks 5.3 and 7.1 therein.

Chapter 7 is an extended version (with an additional section and complete proofs) of [**24**]. It is not directly related to interest rate models and provides general results on stochastic invariance for finite dimensional submanifolds in a Hilbert space. The main results of [**7**] are considerably extended.

The appendix includes a comprehensive discussion on finite dimensional submanifolds in Banach spaces. Their crucial properties are deduced.

Remark on Notation. Basically, we follow the notation of [**40**] and [**18**].

Let G and H be Banach spaces. The norm of H is denoted by $\|\cdot\|_H$ and we write $\mathcal{B}(H)$ for the σ-algebra generated by its topology. If H is a Hilbert space we use the notation $\langle\cdot,\cdot\rangle_H$ for its scalar product. The closure of a subset $M \subset H$ is denoted by $\overline{M}$. We define the open ball of radius $R > 0$

$$B_R(H) := \{h \in H \mid \|h\|_H < R\}.$$

The space of continuous linear operators from G into H is denoted $L(G; H)$ and $L(H) = L(H; H)$ for short. We write $D(A)$ for the domain of an (unbounded) linear operator $A : G \to H$ and A^* for its adjoint.

We write $C(G; H)$ and $C^k(G; H)$ for the space of continuous and k times continuously differentiable mappings $\phi : G \to H$, respectively. The derivatives are denoted by $D\phi$ and $D^k\phi$.

For $U \subset \mathbb{R}^m$, $m \in \mathbb{N}$, the space of continuous, resp. k times continuously differentiable functions $f : U \to \mathbb{R}$ is written shortly $C(U)$, resp. $C^k(U)$, and $C^k_c(U)$ consists of those elements of $C^k(U)$ which have compact support in U.

If $(\Omega, \mathcal{F}, \mathbb{P})$ is a measure space and $p \geq 1$, we write $L^p(\Omega, \mathcal{F}, \mathbb{P})$ for the Banach space of integrable functions[2] $X : \Omega \to \mathbb{R}$ equipped with norm

$$\|X\|^p_{L^p(\Omega,\mathcal{F},\mathbb{P})} = \int_\Omega |X(\omega)|^p \, d\mathbb{P}(\omega).$$

If Ω is a Borel subset of $\mathbb{R}$ and $\mathbb{P}$ the Lebesgue measure, we just write $L^p(\Omega)$. The space of locally integrable functions[3] $h : \Omega \to \mathbb{R}$ is denoted by $L^1_{\mathrm{loc}}(\Omega)$.

Letters such as C, $\tilde{C}$, K, ... represent positive real constants which, if not otherwise stated, can vary from line to line.

All remaining notation is either standard and self-explanatory or it is introduced and explained in the text.

Acknowledgements. It is a great pleasure to thank my adviser F. Delbaen for his guidance throughout my Ph.D. studies. He has given me his time and his insights in the course of many stimulating conversations. I am very grateful to him for his generosity and for giving me the opportunity of several educational excursions abroad.

I have profited from fruitful discussions with many people at several meetings and seminars. Very special thanks go to T. Björk and J. Zabczyk for their helpful and motivating suggestions, and their hospitality, which have been essential for the development of this thesis. I would like to thank also B. J. Christensen and G. Da Prato for their interest and hospitality. Furthermore, I thank P. Embrechts for being co-examiner.

Thanks go to Axel Schulze–Halberg, Paul Harpes, Freweini Tewelde, Uwe Schmock and all my colleagues from ETH for their friendship and support, and especially to Manuela for her love over the last years.

Financial support from Credit Suisse is gratefully acknowledged.

[2]In fact, equivalence classes of functions where $X \sim Y$ if $X = Y$ $\mathbb{P}$-a.s.

[3]That is, equivalence classes ($h \sim g$ if $h = g$ a.s.) of measurable functions h satisfying

$$\int_{B_R(\mathbb{R})\cap\Omega} |h(x)| \, dx < \infty \quad \forall R \in \mathbb{R}_+.$$

Part 1

The HJM Methodology

CHAPTER 1

Stochastic Equations in Infinite Dimensions

In this chapter we provide some introduction to infinite dimensional stochastic analysis and sketch the relevant material from Da Prato and Zabczyk [**18**] without giving detailed proofs. For terminology and notation we refer to [**18**] and [**40**].

Here and subsequently, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ stands for a complete filtered probability space satisfying the usual conditions[1]. The predictable σ-field is denoted by $\mathcal{P}$. We will suppress the notational dependence of stochastic variables on ω when no confusion can arise.

Throughout, H denotes a separable Hilbert space.

Let $I \subset \mathbb{R}_+$. A stochastic process $(t,\omega) \mapsto X_t(\omega) : I \times \Omega \to H$ will be written indifferently[2] $(X_t)_{t\in I}$, resp. X or (X_t) if there is no ambiguity about the index set I. Whereas $X_t : \Omega \to H$ is a random variable for each $t \in I$.

Two processes $(X_t)_{t\in I}$ and $(Y_t)_{t\in I}$ are *indistinguishable* if

$$\mathbb{P}[X_t = Y_t,\, \forall t \in I] = 1.$$

We do not distinguish between them and write $X = Y$.

1.1. Infinite Dimensional Brownian Motion

What is an H-valued "standard Brownian motion" W? A natural requirement would be that $\langle W, h\rangle_H$ were a real Brownian motion with

$$\mathbb{E}[\langle W_t, h\rangle_H \langle W_t, h'\rangle_H] = t\langle h, h'\rangle_H, \quad \forall h, h' \in H.$$

If $\dim H < \infty$ this certainly applies. But in general

$$\mathbb{E}\left[\|W_t\|_H^2\right] = \sum_{j\in\mathbb{N}} \mathbb{E}\left[|\langle W_t, h_j\rangle_H|^2\right] = \sum_{j\in\mathbb{N}} t = \infty$$

where $\{h_j \mid j \in \mathbb{N}\}$ is an orthonormal basis in H. Whence W_t does not exist in H, see [**49**]. Usually one relaxes the condition on W_t belonging to H. Yor [**49**] gives a comprehensive overview of the different concepts for defining W.

However, we will use a more direct approach here. For us, an *infinite dimensional Brownian motion* is a sequence

$$W = (\beta^j)_{j\in\mathbb{N}} \tag{1.1}$$

of independent real-valued standard $(\mathcal{F}_t)$-Brownian motions. For being consistent with [**18**] we show first how W can be realized as a Hilbert space valued Wiener process.

[1] $\mathcal{F}$ is $\mathbb{P}$-complete, $(\mathcal{F}_t)$ is increasing and right continuous, $\mathcal{F}_0$ contains all $\mathbb{P}$-nullsets of $\mathcal{F}$

[2] Differing notations like $(B(t))_{t\in\mathbb{R}_+}$ or $(P(t,T))_{t\in[0,T]}$ are self-explanatory.

For that purpose we introduce the usual Hilbert sequence space

$$\ell^2 := \left\{ v = (v_j)_{j\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \;\middle|\; \|v\|_{\ell^2}^2 := \sum_{j\in\mathbb{N}} |v_j|^2 < \infty \right\}$$

and denote by $\{g_j \mid j \in \mathbb{N}\}$ the standard orthonormal basis[3] in ℓ^2. As mentioned above, $W = (\beta^j)_{j\in\mathbb{N}} \equiv \sum_{j\in\mathbb{N}} \beta^j g_j$ does not exist in ℓ^2. In the notation of [**18**, Section 4.3.1], W is a *cylindrical Wiener process* in ℓ^2. It can be realized in a larger space as follows.

Let $\lambda = (\lambda_j)_{j\in\mathbb{N}}$ be a sequence of strictly positive numbers with $\sum_j \lambda_j < \infty$. We define the weighted sequence Hilbert space

$$\ell^2_\lambda := \left\{ v = (v_j)_{j\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \;\middle|\; \|v\|_{\ell^2_\lambda}^2 := \sum_{j\in\mathbb{N}} \lambda_j |v_j|^2 < \infty \right\}.$$

Clearly, $\ell^2 \subset \ell^2_\lambda$ with Hilbert–Schmidt embedding, and $e_j := (\lambda_j)^{-\frac{1}{2}} g_j$ form an orthonormal basis in ℓ^2_λ. Define $Q \in L(\ell^2_\lambda)$ by $Qe_j := \lambda_j e_j$. Obviously, the operator Q is strictly positive, self-adjoint and nuclear. Moreover we have $Q^{\frac{1}{2}}(\ell^2_\lambda) = \ell^2$ and $Q^{-\frac{1}{2}} : \ell^2 \to \ell^2_\lambda$ is an isometry since for $u, v \in \ell^2$

$$\langle u, v\rangle_{\ell^2} = \sum_{j\in\mathbb{N}} u_j v_j = \sum_{j\in\mathbb{N}} \lambda_j \frac{u_j}{\sqrt{\lambda_j}} \frac{v_j}{\sqrt{\lambda_j}} = \langle Q^{-\frac{1}{2}} u, Q^{-\frac{1}{2}} v\rangle_{\ell^2_\lambda}. \tag{1.2}$$

Fix $T \in \mathbb{R}_+$.

Definition 1.1. *We denote by $\mathcal{H}^2_T(H)$ the space of H-valued* continuous martingales[4] *M on $[0,T]$ such that*

$$\|M\|^2_{\mathcal{H}^2_T(H)} := \mathbb{E}\left[\|M_T\|_H^2\right] < \infty. \tag{1.3}$$

Write $M_T^* := \sup_{t\in[0,T]} \|M_t\|_H$. Doob's inequality [**18**, Theorem 3.8] says

$$\mathbb{E}\left[|M_T^*|^2\right] \le 4\|M\|^2_{\mathcal{H}^2_T(H)}, \quad \text{for } M \in \mathcal{H}^2_T. \tag{1.4}$$

Using (1.4) it can be shown as in [**18**, Proposition 3.9] that $\mathcal{H}^2_T(H)$ is a Hilbert space for the norm (1.3). Denote by $\mathcal{H}^{0,2}_T(H)$ the closed subspace consisting of $M \in \mathcal{H}^2_T(H)$ with $M_0 = 0$.

We adopt the terminology of [**18**, Section 4.1].

Definition 1.2. *An ℓ^2_λ-valued continuous martingale X is called a Q-*Wiener process *if*

i) $X_0 = 0$
ii) *X has independent increments*
iii) *$X_t - X_s$ is centered Gaussian*[5] *with covariance operator $(t-s)\,Q$, for all $0 \le s \le t$.*

Proposition 1.3. *The series*

$$W = \sum_{j\in\mathbb{N}} \beta^j g_j$$

converges in $\mathcal{H}^{0,2}_T(\ell^2_\lambda)$ for all $T \in \mathbb{R}_+$ and defines a Q-Wiener process in ℓ^2_λ.

---

[3] That is, $g_1 = (1, 0, \dots)$, $g_2 = (0, 1, 0, \dots)$, ...

[4] see [**18**, Section 3.4]

[5] see [**18**, Section 2.3.2]

Observe that this realization of W as a Hilbert space valued Wiener process depends on the choice of λ. Indeed, the same procedure applies for any given sequence λ sharing the above properties. Yet the class of integrable processes does not depend on λ as we shall see.

PROOF. Fix $T \in \mathbb{R}_+$. We define $W^n \in \mathcal{H}^{0,2}_T(\ell^2_\lambda)$ by

$$W^n = \sum_{j=1}^{n} \beta^j g_j, \quad n \in \mathbb{N}. \tag{1.5}$$

Obviously, $\mathbb{E}\left[\|W^n_T - W^m_T\|^2_{\ell^2_\lambda}\right] = T\sum_{j=m+1}^n \lambda_j \to 0$ for $m, n \to 0$. Hence $(W^n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathcal{H}^{0,2}_T(\ell^2_\lambda)$, converging to W. Since this holds for any $T \in \mathbb{R}_+$, we get that W is an ℓ^2_λ-valued continuous martingale.

For any $v \in \ell^2_\lambda$ the $\mathbb{R}$-valued random variables $\langle W^n_t - W^n_s, v\rangle_{\ell^2_\lambda}$ are centered Gaussian, converging in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ to $\langle W_t - W_s, v\rangle_{\ell^2_\lambda}$ which is therefore centered Gaussian too. By the same reasoning

$$\mathbb{E}\left[\langle W_t - W_s, e_i\rangle_{\ell^2_\lambda} \langle W_t - W_s, e_j\rangle_{\ell^2_\lambda}\right] = (t-s)\langle Qe_i, e_j\rangle_{\ell^2_\lambda}.$$

Whence Condition iii) of Definition 1.2 holds. Similarly one shows ii) and clearly $W_0 = 0$. □

REMARK 1.4. *We notice that any Hilbert space valued (cylindrical) Wiener process in the terminology of Da Prato and Zabczyk* [**18**] *can be considered in the preceding $\mathbb{R}^{\mathbb{N}}$-framework (1.1), see* [**18**, Propositions 4.1 and 4.11].

1.2. The Stochastic Integral

For the definition of the stochastic integral with respect to W, Da Prato and Zabczyk [**18**] introduce the space $L^0_2(H)$ of Hilbert Schmidt operators from ℓ^2 into H.[6] It consists of linear mappings $\Phi : \ell^2 \to H$ with

$$\|\Phi\|^2_{L^0_2(H)} := \sum_{j\in\mathbb{N}} \|\Phi^j\|^2_H < \infty, \tag{1.6}$$

where $\Phi^j := \Phi(g_j)$. In the sequel we shall identify Φ with $(\Phi^j)_{j\in\mathbb{N}}$. On the other hand, observe that any sequence $(\Phi^j)_{j\in\mathbb{N}}$ of H-valued predictable processes satisfying (1.6) provides[7] an $L^0_2(H)$-valued predictable process Φ by $\Phi(g_j) := \Phi^j$.

Notice the particular case $L^0_2(\mathbb{R}) = \ell^2$.

We summarize the construction of the stochastic integral, based on the theory from [**18**, Sections 4.1-4.4]. Fix $T \in \mathbb{R}_+$.

DEFINITION 1.5. *We call $\mathcal{L}^2_T(H)$ the Hilbert space of equivalence classes of $L^0_2(H)$-valued predictable processes Φ with norm*

$$\|\Phi\|^2_{\mathcal{L}^2_T(H)} := \mathbb{E}\left[\int_0^T \|\Phi_t\|^2_{L^0_2(H)}\, dt\right]. \tag{1.7}$$

---

[6]Remember that $\ell^2 = Q^{\frac{1}{2}}(\ell^2_\lambda)$ endowed with the implied scalar product, see (1.2) and compare with [**18**, (4.7)].

[7]Notice that $\mathcal{B}(L^0_2(H)) = \bigotimes_{\mathbb{N}} \mathcal{B}(H)$.

Then there exists an isometry $\Phi \mapsto \Phi \cdot W$ between $\mathcal{L}^2_T(H)$ and $\mathcal{H}^{0,2}_T(H)$, see [**18**, Proposition 4.5], defined on the dense[8] subset of elementary $L(\ell^2_\lambda; H)$-valued[9] processes

$$\Phi = \sum_{m=0}^{k-1} \Phi_m 1_{(t_m, t_{m+1}]}, \quad \Phi_m \text{ is } \mathcal{F}_{t_m}\text{-measurable,} \tag{1.8}$$

by

$$(\Phi \cdot W)_t := \sum_{m=0}^{k-1} \Phi_m (W_{t_{m+1}\wedge t} - W_{t_m \wedge t}), \quad t \in [0,T].$$

Definition 1.6. *The H-valued continuous martingale $(\Phi \cdot W)_{t\in[0,T]}$ is called the* stochastic integral *of Φ with respect to W and is also denoted by*

$$(\Phi \cdot W)_t = \int_0^t \Phi_s \, dW_s.$$

It is clear, how stochastic integration with respect to the d-dimensional Brownian motion $W^d = (\beta^1, \dots, \beta^d)$, $d \in \mathbb{N}$, is included in this concept: any H^d-valued predictable process $\Phi = (\Phi^1, \dots, \Phi^d)$ can be considered as $L^0_2(H)$-valued predictable process $\tilde{\Phi}$ by setting $\tilde{\Phi}^j = \Phi^j$ if $j \le d$ and $\tilde{\Phi}^j \equiv 0$ for $j > d$. This way one gets easily

$$(\Phi \cdot W^d)_t = (\tilde{\Phi} \cdot W)_t = \sum_{j=1}^d \int_0^t \Phi^j_s \, d\beta^j_s. \tag{1.9}$$

Here are some elementary properties of the stochastic integral. Denote by E a separable Hilbert space.

Proposition 1.7. *Let $\Phi \in \mathcal{L}^2_T(H)$.*

i) *For any stopping time $\tau \le T$*

$$(\Phi \cdot W)_{t\wedge\tau} = ((\Phi 1_{[0,\tau]}) \cdot W)_t. \tag{1.10}$$

ii) *The following series converges uniformly on $[0,T]$ in probability*

$$(\Phi \cdot W)_t = \sum_{j\in\mathbb{N}} \int_0^t \Phi^j_s \, d\beta^j_s. \tag{1.11}$$

iii) *Let $A \in L(H; E)$. Then $A \circ \Phi \in \mathcal{L}^2_T(E)$ and*

$$A(\Phi \cdot W) = (A \circ \Phi) \cdot W. \tag{1.12}$$

Proof. Property i) is [**18**, Lemma 4.9]. Property ii) follows by considering (1.9) and $(\Phi^1, \dots, \Phi^d, 0, \dots)$ which converges in $\mathcal{L}^2_T(H)$ towards Φ for $d \to \infty$. Now apply Doob's inequality (1.4). Finally, iii) can be shown in the same way as [**18**, Proposition 4.15].

Notice that all equalities hold in $\mathcal{H}^{0,2}_T(H)$, hence up to indistinguishability. □

There is a larger class of integrable processes than $\mathcal{L}^2_T(H)$.

[8] see [**18**, Proposition 4.7]

[9] for an arbitrary choice of $\lambda = (\lambda_j)_{j\in\mathbb{N}}$ as in Section 1.1

DEFINITION 1.8. *We call $\mathcal{L}_T^{loc}(H)$ the space of all $L_2^0(H)$-valued predictable processes Φ such that*

$$\mathbb{P}\left[\int_0^T \|\Phi_t\|^2_{L_2^0(H)} < \infty\right] = 1.$$

Observe that $\Phi \in \mathcal{L}_T^{\mathrm{loc}}(H)$ if and only if there exists an increasing sequence of stopping times $\tau_n \uparrow T$ such that $\Phi 1_{[0,\tau_n]}$ is in $\mathcal{L}_T^2(H)$ for all $n \in \mathbb{N}$. This allows us to extend the notion of the stochastic integral, see [**18**, p. 94–96].

PROPOSITION 1.9. *Let $\Phi \in \mathcal{L}_T^{loc}(H)$. Then there exists a unique H-valued continuous local martingale $(M_t)_{t\in[0,T]}$ characterized by*

$$M_{t\wedge\tau} = ((\Phi 1_{[0,\tau]}) \cdot W)_t$$

whenever $\Phi 1_{[0,\tau]} \in \mathcal{L}_T^2(H)$. Again, M is called the stochastic integral *of Φ with respect to W and it is written*

$$M_t =: (\Phi \cdot W)_t = \int_0^t \Phi_s \, dW_s.$$

Many properties of the stochastic integral for $\mathcal{L}_T^2(H)$-integrands carry over for $\mathcal{L}_T^{\mathrm{loc}}(H)$-integrands by *localization*. Here is one feature of this method.

LEMMA 1.10. *Let $(M^n)_{n\in\mathbb{N}}$ be a sequence of H-valued continuous local martingales. Assume that for all $\delta > 0$ there exists a stopping time τ with $\mathbb{P}[\tau < T] \leq \delta$ such that $(M^n_{t\wedge\tau}) \to 0$ in $\mathcal{H}_T^2(H)$. Then $(M^n)^*_T \to 0$ in probability.*

PROOF. Let $\delta, \epsilon > 0$ and τ be as in the lemma. Then

$$\mathbb{P}\left[(M^n)^*_T > \epsilon\right] \leq \delta + \mathbb{P}\left[\sup_{t\in[0,T]} \|M^n_{t\wedge\tau}\|_H > \epsilon\right]$$

and the last term goes to zero as $n \to \infty$ by Doob's inequality (1.4). $\square$

Using Lemma 1.10 the following can be shown.

PROPOSITION 1.11. *Equality (1.9) and Proposition 1.7 remain valid if $\mathcal{L}_T^2(H)$ is replaced by $\mathcal{L}_T^{loc}(H)$.*

Up to now, the stochastic integral was defined only on the finite time interval $[0,T]$. But we can consider integrands in $\mathcal{L}^2(H) := \mathcal{L}^2_\infty(H)$ (with T in (1.7) replaced by ∞), respectively $\mathcal{L}^{\mathrm{loc}}(H) := \bigcap_{T\in\mathbb{R}_+} \mathcal{L}_T^{\mathrm{loc}}(H)$.

Let $\Phi \in \mathcal{L}^2(H)$. Write $M^n := \Phi 1_{[0,n]} \cdot W$ for the stochastic integral on $[0,n]$. By (1.10), M^{n+1} coincides with M^n on $[0,n]$ for all $n \in \mathbb{N}$. Therefore we can define unambiguously a process $\Phi\cdot W$ – the *stochastic integral* of Φ with respect to W – by stipulating that it is equal to M^n on $[0,n]$. This process is obviously an H-valued continuous martingale.

By localization the same procedure applies for $\Phi \in \mathcal{L}^{\mathrm{loc}}(H)$, and $\Phi \cdot W$ is an H-valued continuous local martingale.

1.3. Fundamental Tools

Three fundamental theorems of stochastic analysis are *Itô's formula*, the *stochastic Fubini theorem* and *Girsanov's theorem*. The first is given by Lemma 7.11. This section provides the other two.

1.3.1. The Stochastic Fubini Theorem. Let $(E,\mathcal{E})$ be a measurable space and let $\Phi : (t,\omega,x) \mapsto \Phi_t(\omega,x)$ be a measurable mapping from $(\mathbb{R}_+ \times \Omega \times E, \mathcal{P} \otimes \mathcal{E})$ into $(L_2^0(H), \mathcal{B}(L_2^0(H)))$. In addition let μ denote a finite positive measure on $(E,\mathcal{E})$.

Fix $T > 0$. By localization we get the following version of [**18**, Theorem 4.18].

THEOREM 1.12 (Stochastic Fubini theorem). *Assume that*

$$\int_0^T \int_E \|\Phi_t(x)\|^2_{L_2^0(H)} \, \mu(dx) \, dt < \infty, \quad \mathbb{P}\text{-}a.s.$$

Then there exists an $\mathcal{F}_T \otimes \mathcal{E}$-measurable version $\xi(\omega,x)$ of the stochastic integral $\int_0^T \Phi_t(x) \, dW_t$ which is μ-integrable $\mathbb{P}$-a.s. such that

$$\int_E \xi(x) \, \mu(dx) = \int_0^T \left(\int_E \Phi_t(x) \mu(dx) \right) dW_t, \quad \mathbb{P}\text{-}a.s.$$

1.3.2. Girsanov's Theorem. Let $\gamma = (\gamma^j)_{j\in\mathbb{N}} \in \mathcal{L}^{\mathrm{loc}}(\mathbb{R})$[10]. We define the *stochastic exponential* $\mathcal{E}(\gamma \cdot W)$ of $\gamma \cdot W$ by

$$\mathcal{E}\left(\gamma \cdot W\right)_t := \exp\left((\gamma \cdot W)_t - \frac{1}{2} \int_0^t \|\gamma_s\|^2_{\ell^2} \, ds \right). \tag{1.13}$$

By [**18**, Lemma 10.15] there exists a real-valued standard $(\mathcal{F}_t)$-Brownian motion β^0 such that $(\gamma \cdot W)_t = \int_0^t \|\gamma_s\|_{\ell^2} \, d\beta^0_s$. This way we are lead to the real-valued theory.

LEMMA 1.13. *The stochastic exponential $\mathcal{E}(\gamma \cdot W)$ is a non-negative local martingale. It is a martingale if and only if $\mathbb{E}\left[\mathcal{E}(\gamma \cdot W)_t\right] = 1$ for all $t \in \mathbb{R}_+$.*

PROOF. By Itô's formula

$$\mathcal{E}(\gamma \cdot W)_t = 1 + \int_0^t \left(\mathcal{E}(\gamma \cdot W)_s \|\gamma_s\|_{\ell^2}\right) d\beta^0_s.$$

Whence the first assertion. For the second statement we refer the reader to [**40**, Remark 3, p. 141]. □

We recall [**40**, Proposition (1.15), Chapter VIII].

LEMMA 1.14 (Novikov's criterion). *If*

$$\mathbb{E}\left[\exp\left(\frac{1}{2} \int_{\mathbb{R}_+} \|\gamma_t\|^2_{\ell^2} \, dt \right) \right] < \infty, \tag{1.14}$$

then $\mathcal{E}(\gamma \cdot W)$ is a uniformly integrable martingale. Accordingly, $\mathcal{E}(\gamma \cdot W)_t \to \mathcal{E}(\gamma \cdot W)_\infty$ in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ for $t \to \infty$ and $\mathcal{E}(\gamma \cdot W)_\infty > 0$, $\mathbb{P}$-a.s.

The last statement is a direct consequence of the fact that $\mathcal{E}(\gamma \cdot W)_\infty > 0$, $\mathbb{P}$-a.s. on the set $\{\int_{\mathbb{R}_+} \|\gamma_t\|^2_{\ell^2} \, dt < \infty\}$, see [**40**, Proposition (1.26), Chapter IV].

Let $T \in \mathbb{R}_+$. The next theorem will assure the existence of an equivalent local martingale measure.

THEOREM 1.15 (Girsanov's Theorem). *Let $\gamma = (\gamma^j)_{j\in\mathbb{N}} \in \mathcal{L}^{loc}_T(\mathbb{R})$ satisfy*

$$\mathbb{E}\left[\mathcal{E}(\gamma \cdot W)_T\right] = 1.$$

[10] Recall that $L_2^0(\mathbb{R}) = \ell^2$.

Then

$$\tilde{\beta}_t^j := \beta_t^j - \int_0^t \gamma_s^j \, ds, \quad t \in [0,T], \quad j \in \mathbb{N}$$

are independent real-valued standard $(\mathcal{F}_t)_{t\in[0,T]}$*-Brownian motions with respect to the measure* $\tilde{\mathbb{P}}_T \sim \mathbb{P}$ *on* $\mathcal{F}_T$ *given by*

$$\frac{d\tilde{\mathbb{P}}_T}{d\mathbb{P}} = \mathcal{E}(\gamma \cdot W)_T.$$

PROOF. See [**18**, Theorem 10.14]. □

We can now define, as above, the spaces $\tilde{\mathcal{H}}_T^2(H)$, $\tilde{\mathcal{H}}_T^{0,2}(H)$, $\tilde{\mathcal{L}}_T^2(H)$ and $\tilde{\mathcal{L}}_T^{\mathrm{loc}}(H)$ on the filtered probability space $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, \tilde{\mathbb{P}}_T)$. Obviously we have $\tilde{\mathcal{L}}_T^{\mathrm{loc}}(H) = \mathcal{L}_T^{\mathrm{loc}}(H)$. Notice that a sequence of $\mathcal{F}_T$-measurable random variables converging in probability $\mathbb{P}$ also convergences in probability $\tilde{\mathbb{P}}_T$ and vice versa.

PROPOSITION 1.16. *The series*

$$\tilde{W} = \sum_{j\in\mathbb{N}} \tilde{\beta}^j g_j$$

converges in $\tilde{\mathcal{H}}_T^{0,2}(\ell_\lambda^2)$ *and defines a* Q*-Wiener process in* ℓ_λ^2 *with respect to* $\tilde{\mathbb{P}}_T$.

Moreover, for any $\Phi \in \tilde{\mathcal{L}}_T^{loc}(H)$,

$$(\Phi \cdot \tilde{W})_t = (\Phi \cdot W)_t - \int_0^t \Phi_s(\gamma_s) \, ds, \quad \forall t \in [0,T], \quad \mathbb{P}\text{-}a.s. \tag{1.15}$$

PROOF. For $n \in \mathbb{N}$ we define $\tilde{W}^n := \sum_{j=1}^n \tilde{\beta}^j g_j$. By Proposition 1.3, the sequence $(\tilde{W}^n)_{n\in\mathbb{N}}$ converges to $\tilde{W}$ in $\tilde{\mathcal{H}}_T^{0,2}(H)$, and $\tilde{W}$ is a Q-Wiener process in ℓ_λ^2 for the measure $\tilde{\mathbb{P}}_T$. In view of (1.5) and Theorem 1.15

$$\tilde{W}_t^n = W_t^n - \sum_{j=1}^n \int_0^t \gamma_s^j g_j \, ds, \quad \forall n \in \mathbb{N}.$$

By Doob's inequality (1.4) and Proposition 1.3, the right hand side converges to $W_t - \sum_{j\in\mathbb{N}} \int_0^t \gamma_s^j g_j \, ds$ uniformly on $[0,T]$ in probability, which is therefore indistinguishable from $\tilde{W}$ on $[0,T]$.

Equality (1.15) is certainly true for elementary $L(\ell_\gamma^2; H)$-valued processes, recall (1.8). Now let $\Phi \in \tilde{\mathcal{L}}_T^{\mathrm{loc}}(H) = \mathcal{L}_T^{\mathrm{loc}}(H)$. We define the stopping times

$$\tau_n := T \wedge \inf\left\{ t \in \mathbb{R}_+ \;\middle|\; \int_0^t \|\Phi_s\|_{L_2^0(H)}^2 \, ds \geq n \right\}, \quad n \in \mathbb{N}.$$

Then $\tau_n \uparrow T$ and $\Phi^n := \Phi 1_{[0,\tau_n]} \in \tilde{\mathcal{L}}_T^2(H) \cap \mathcal{L}_T^2(H)$. By Lemma 1.17 below, (1.15) holds for each Φ^n. By (1.10) therefore also for Φ. □

LEMMA 1.17. *Let* $\Phi \in \tilde{\mathcal{L}}_T^2(H) \cap \mathcal{L}_T^2(H)$*. Then there exists a sequence* $(\Phi^n)_{n\in\mathbb{N}}$ *of elementary* $L(\ell_\gamma^2; H)$*-valued processes converging to* Φ *in* $\tilde{\mathcal{L}}_T^2(H)$ *and in* $\mathcal{L}_T^2(H)$ *simultaneously.*

PROOF. Standard measure theory. □

1.4. Stochastic Equations

This section provides the basic concepts and results for stochastic equations in infinite dimensions. Let $W=(\beta^j)_{j\in\mathbb{N}}$ be an infinite dimensional Brownian motion as introduced in Section 1.1. Our standing set of ingredients is the following:

- A strongly continuous semigroup $\{S(t)\mid t\in\mathbb{R}_+\}$ on H with infinitesimal generator A.
- Two measurable mappings $F(t,\omega,h)$ and $\sigma(t,\omega,h)=(\sigma^j(t,\omega,h))_{j\in\mathbb{N}}$ from $(\mathbb{R}_+\times\Omega\times H,\mathcal{P}\otimes\mathcal{B}(H))$ into $(H,\mathcal{B}(H))$, resp. $(L_2^0(H),\mathcal{B}(L_2^0(H)))$.
- A non random initial value $h_0\in H$.

1.4.1. Mild, Weak and Strong Solutions. We shall assign a meaning to the *stochastic equation* in H

$$\begin{cases} dX_t=(AX_t+F(t,X_t))\,dt+\sigma(t,X_t)\,dW_t\\ X_0=h_0.\end{cases}\tag{1.16}$$

In view of (1.11) this can be rewritten equivalently

$$\begin{cases} dX_t=(AX_t+F(t,X_t))\,dt+\sum_{j\in\mathbb{N}}\sigma^j(t,X_t)\,d\beta_t^j\\[2ex] X_0=h_0.\end{cases}\tag{1.16$'$}$$

Remark 1.18. *Notice that F and σ are random mappings and may depend* explicitly *on ω. In accordance with the notation of* [**18**]*, however, we abbreviate $F(t,\omega,X_t(\omega))$ and $\sigma(t,\omega,X_t(\omega))$ to $F(t,X_t)$ and $\sigma(t,X_t)$, respectively.*

Here are three concepts of a solution.

Definition 1.19. *Suppose X is an H-valued predictable process satisfying*

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

for a stopping time $\tau>0$, the lifetime *of X. We call X*

i) *a* local mild solution *to (1.16), if the stopped variation of constants formula holds*[11]

$$\begin{aligned} X_t=S(t\wedge\tau)h_0&+\int_0^{t\wedge\tau}S((t\wedge\tau)-s)F(s,X_s)\,ds\\ &+\sum_{j\in\mathbb{N}}\int_0^{t\wedge\tau}S((t\wedge\tau)-s)\sigma^j(s,X_s)\,d\beta_s^j,\quad\mathbb{P}\text{-a.s.}\quad\forall t\in\mathbb{R}_+.\end{aligned}$$

ii) *a* local weak solution *to (1.16), if for arbitrary $\zeta\in D(A^\star)$*

$$\begin{aligned}\langle\zeta,X_t\rangle_H=\langle\zeta,h_0\rangle_H&+\int_0^{t\wedge\tau}\Big(\langle A^\star\zeta,X_s\rangle_H+\langle\zeta,F(s,X_s)\rangle_H\Big)\,ds\\ &+\sum_{j\in\mathbb{N}}\int_0^{t\wedge\tau}\langle\zeta,\sigma^j(s,X_s)\rangle_H\,d\beta_s^j,\quad\mathbb{P}\text{-a.s.}\quad\forall t\in\mathbb{R}_+.\end{aligned}$$

[11] The stochastic convolution always has a predictable version, see [**18**, Proposition 6.2].

iii) *a* local strong solution *to (1.16), if $X \in D(A)$, $dt \otimes d\mathbb{P}$-a.s.*

$$\mathbb{P}\left[\int_0^{t\wedge\tau} \|AX_s\|_H \, ds < \infty\right] = 1, \quad \forall t \in \mathbb{R}_+ \tag{1.17}$$

and the integral version of (1.16) holds

$$X_t = h_0 + \int_0^{t\wedge\tau} \Big(AX_s + F(s, X_s)\Big) ds + \int_0^{t\wedge\tau} \sigma(s, X_s) \, dW_s, \quad \mathbb{P}\text{-a.s.} \quad \forall t \in \mathbb{R}_+.$$

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

This definition may be straightly extended to random $\mathcal{F}_0$-measurable initial values h_0. Notice that the lifetime τ is by no means maximal.

Remark 1.20. *For stochastic equations in finite dimensions one usually distinguishes between strong and weak solutions, see e.g.* [**32**, Chapter 5]. *There, however,* weak *has a different meaning than in the present context. Here we follow the terminology of* [**18**].

According to [**18**, Theorem 6.5] the concepts in Definition 1.19 are related by

$$iii) \Rightarrow ii) \Rightarrow i),$$

and $i) \Rightarrow ii)$ if $\sigma(\cdot, X) \in \mathcal{L}^2(H)$.

Since the coefficients F and σ are not homogeneous in time, we will frequently consider the time shifted version of (1.16). The following result is classic.

Lemma 1.21. *Let τ_0 denote a bounded stopping time. Then*

$$\beta_t^{(\tau_0),j} := \beta^j_{\tau_0+t} - \beta^j_{\tau_0}, \quad t \in \mathbb{R}_+, \quad j \in \mathbb{N}$$

are independent standard Brownian motions for the filtration $(\mathcal{F}_t^{(\tau_0)}) := (\mathcal{F}_{\tau_0+t})$.

Proof. See Lemma 7.3. □

We write $W^{(\tau_0)} = (\beta^{(\tau_0),j})_{j\in\mathbb{N}}$. It is obvious how $W^{(\tau_0)}$ can be realized as Hilbert space valued $(\mathcal{F}_t^{(\tau_0)})$-Wiener process and how to define $\mathcal{L}^{\mathrm{loc}}(H, (\mathcal{F}_t^{(\tau_0)}))$ etc. In fact, $(\Phi_t) \in \mathcal{L}^{\mathrm{loc}}(H)$ if and only if $(\Phi_{\tau_0+t}) \in \mathcal{L}^{\mathrm{loc}}(H, (\mathcal{F}_t^{(\tau_0)}))$.

The *time τ_0-shifted version* of (1.16) now reads

$$\begin{cases} dX_t = (AX_t + F(\tau_0 + t, X_t)) \, dt + \sigma(\tau_0 + t, X_t) \, dW_t^{(\tau_0)} \\ X_0 = h_0 \end{cases} \tag{1.18}$$

and similarly the expanded form (1.16′). A local mild (weak, strong) solution to (1.18) will be denoted by $X^{(\tau_0,h_0)}$. Suppose that $X = X^{(0,h_0)}$ is a local mild (weak, strong) solution to (1.16) with lifetime $\tau > \tau_0$. Then (X_{τ_0+t}) is local mild (weak, strong) solution to (1.18) with h_0 replaced by X_{τ_0} and lifetime $\tau - \tau_0$. This can be shown using the identity

$$\int_{\tau_0}^{\tau_0+t} \Phi_s \, dW_s = \int_0^t \Phi_{\tau_0+s} \, dW_s^{(\tau_0)}, \quad \Phi \in \mathcal{L}^{\mathrm{loc}}(H)$$

which is derived in the proof of Lemma 7.3.

In the sequel we shall mostly consider deterministic times $\tau_0 \equiv t_0 \in \mathbb{R}_+$.

1.4.2. Existence and Uniqueness. Lipschitz continuity plays an important role in the theory of stochastic equations. Assume that D and E are Banach spaces.

DEFINITION 1.22. *A mapping $\Phi = \Phi(t, \omega, h) : \mathbb{R}_+ \times \Omega \times D \to E$ is called* locally Lipschitz continuous in h*, if for all $R \in \mathbb{R}_+$ there exists a number $C = C(R)$, the* Lipschitz constant *for Φ, such that for all $(t, \omega) \in \mathbb{R}_+ \times \Omega$*

$$\|\Phi(t, \omega, h_1) - \Phi(t, \omega, h_2)\|_E \leq C\|h_1 - h_2\|_D, \quad \forall h_1, h_2 \in B_R(D). \tag{1.19}$$

If C does not depend on R,[12] *we call Φ simply* Lipschitz continuous in h*.*

We can now formulate the main result on stochastic equations which is a combination of Theorems 6.5 and 7.4 in [**18**]. It is based on a classical fixed point argument.

THEOREM 1.23. *Suppose $F(t, \omega, h)$ and $\sigma(t, \omega, h)$ are Lipschitz continuous in h and satisfy the linear growth condition*

$$\|F(t, \omega, h)\|_H^2 + \|\sigma(t, \omega, h)\|_{L_2^0(H)}^2 \leq C(1 + \|h\|_H^2), \quad \forall (t, \omega, h) \in \mathbb{R}_+ \times \Omega \times H,$$

where C does not depend on (t, ω, h).

Then for any space time initial point $(t_0, h_0) \in \mathbb{R}_+ \times H$ there exists a unique continuous weak solution $X = X^{(t_0,h_0)}$ to the time t_0-shifted equation (1.18). Moreover, for any $T \in \mathbb{R}_+$ there exists a constant $C = C(T)$ such that

$$\mathbb{E}\left[|X_T^*|^2\right] \leq C(1 + \|h_0\|_H^2).$$

Here is the local analogue.

COROLLARY 1.24. *Suppose $F(t, \omega, h)$ and $\sigma(t, \omega, h)$ are locally Lipschitz continuous in h.*

Then for any space time initial point $(t_0, h_0) \in \mathbb{R}_+ \times H$ there exists a unique continuous local weak solution $X = X^{(t_0,h_0)}$ to the time t_0-shifted equation (1.18).

PROOF. See Lemma 7.5. □

REMARK 1.25. *Theorem 1.23 and its corollary indicate that a (local)* weak *solution is the proper concept for equation (1.16). Indeed, (local) strong solutions exist very rarely in general. Even if $F, \sigma^j \in D(A)$ and $D(A)$ is invariant for $S(t)$, this does not imply (1.17) yet. See e.g.* [**15**, Proposition 3.26] *for the case $F \equiv 0$.*

[12] Actually C may depend on T and then (1.19) holds for all $(t, \omega) \in [0, T] \times \Omega$. The subsequent existence results remain valid. For ease of notation we abandon this generalization.

CHAPTER 2

The HJM Methodology Revisited

2.1. The Bond Market

This chapter is devoted to a detailed discussion of the seminal paper by HJM [**28**].

We consider a continuous trading economy with an infinite trading interval. The case of a finite time interval for the term structure is not issue of this thesis but will be exploited elsewhere.

The basic contract is a (default-free) *zero coupon bond* with maturity date T, which pays the holder with certainty one unit of cash at time T. Let $P(t,T)$ denote its time t price for $t \leq T$. Clearly $P(t,t) = 1$, and we require that the following term structure representation holds

$$P(t,T) = \exp\left(-\int_t^T f(t,s)\,ds\right), \quad \forall T \geq t, \tag{2.1}$$

for some locally integrable function $f(t,\cdot) : [t,\infty) \to \mathbb{R}$, the time t *forward curve*. The number $f(t,T)$ is the instantaneous *forward rate* prevailing at time t for a risk-less loan that begins at date $T \geq t$ and is returned an instant later. In the particular case $T = t$ we call $f(t,t)$ the *short rate*.

HJM [**28**] represent each particular forward rate process $(f(t,T))_{t\in[0,T]}$ as an Itô process based on a fixed finite dimensional Brownian motion. We follow their terminology at the beginning, but will then allow for an infinite dimensional underlying Brownian motion.

We adopt the space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in\mathbb{R}_+},\mathbb{P})$ from Chapter 1 and consider first the standard d-dimensional $(\mathcal{F}_t)$-Brownian motion $(\beta^1,\dots,\beta^d)$, $d \in \mathbb{N}$. Define the triangle subset of $\mathbb{R}^2$

$$\Delta^2 := \{(t,T) \in \mathbb{R}^2 \mid 0 \leq t \leq T\}.$$

We recapture the HJM setup [**28**, Condition C.1] for the term structure movements. Let $\alpha(t,T,\omega)$ and $\sigma^j(t,T,\omega)$, $1 \leq j \leq d$, be measurable mappings from $(\Delta^2 \times \Omega, \mathcal{B}(\Delta^2)\otimes\mathcal{F})$ into $(\mathbb{R},\mathcal{B}(\mathbb{R}))$, such that for each $T \in \mathbb{R}_+$ the processes $(\alpha(t,T))_{t\in[0,T]}$ and $(\sigma^j(t,T))_{t\in[0,T]}$ are progressively measurable. Moreover

$$\int_0^T \left(|\alpha(t,T)| + \|\sigma(t,T)\|^2_{\mathbb{R}^d}\right) dt < \infty, \quad \mathbb{P}\text{-a.s.} \quad \forall T \in \mathbb{R}_+.$$

Then for fixed, but arbitrary $T \in \mathbb{R}_+$, the forward rate for date T evolves as the Itô process

$$f(t,T) = f(0,T) + \int_0^t \alpha(s,T)\,ds + \sum_{j=1}^d \int_0^t \sigma^j(s,T)\,d\beta^j_s, \quad t \in [0,T]. \tag{2.2}$$

Here $f(0,\cdot)$ is a non-random initial forward curve which is locally integrable on $\mathbb{R}_+$.

We emphasize that (2.2) represents a system of infinitely many Itô processes indexed by $T \in \mathbb{R}_+$. In what follows we tend to transform this into *one* infinite dimensional process.

Under some additional assumptions – which we do not further specify here, since we will take this point up in the next section – the *savings account*

$$B(t) := \exp\left(\int_0^t f(s,s)\,ds\right), \quad t \in \mathbb{R}_+,$$

can be shown to follow an Itô process. So does $(P(t,T))_{t\in[0,T]}$, see [**28**, Conditions C.2 and C.3].

2.2. The Musiela Parametrization

Musiela [**37**] proposed to change the parametrization of the term structure according to $f(t, x+t)$, where $x \geq 0$ denotes *time to maturity*. There are indeed several reasons for doing so.

Non-varying state space: The basic underlying state variable is the time t forward curve $\{f(t,T) \mid T \geq t\}$. In the HJM time framework, however, this means that the state space is a space of functions on an interval varying with t. The above modification eliminates this difficulty.

Non-local state dependence: Models within the classical HJM framework have state dependent volatility coefficients of the form $\sigma(t,T,\omega,f(t,\omega,T))$, see [**28**], [**36**] and [**35**]. For short rate models the dependence is on $f(t,t)$, respectively. In any case, the dependence on the state variable is *local*. It would be natural, however, if the volatility depended on the whole forward curve. Hence if also non-local (like LIBOR rate) dependence were allowed, incorporating all LIBOR or simple rate models in one framework. This aim is provided by considering $f(t, \cdot + t)$ as state variable in a function space, evolving according to functional sensitive dynamics.

Consistency problems: One major issue of this thesis is to investigate when the system of infinitely many processes (2.2) does allow for a finite dimensional realization. This question on one hand is of interest by its own. On the other hand it occurs in connection with statistical purposes, as explained in the introduction.

Suppose $\mathcal{G} = \{G(\cdot, z) \mid z \in \mathcal{Z}\}$ represents a forward curve fitting method. It is natural to ask whether there exists any HJM model $(f(t,T))_{(t,T)\in\Delta^2}$ which is consistent with $\mathcal{G}$. That is, which produces forward curves which lie within the class $\mathcal{G}$.

Let's consider for the moment the Nelson–Siegel family, see Chapter 5,

$$G(x,z) = z_1 + (z_2 + z_3 x)e^{-z_4 x}, \quad \mathcal{Z} \subset \mathbb{R}^4.$$

Formally speaking, we search a family of forward rate processes given by (2.2) and a $\mathcal{Z}$-valued process Z such that

$$G(T-t, Z_t) = f(t,T), \quad \forall (t,T) \in \Delta^2. \tag{2.3}$$

A naive approach is to fix a tenor $0 < T_1 < \cdots < T_5$, define the mapping

$$\tilde{G}(t,z) := (G(T_1 - t, z), \ldots, G(T_5 - t, z)) : [0,T_1] \times \mathcal{Z} \to \mathbb{R}^5$$

and try to (locally) invert $\tilde{G}$, which yields

$$(t, Z_t) = \tilde{G}^{-1}(f(t,T_1), \ldots, f(t,T_5)), \quad t \in (0, T_1).$$

The inverse mapping theorem requires non-singularity of the 5×5-matrix

$$\begin{bmatrix} -\partial_x G(T_1 - t, z) & \partial_{z_1} G(T_1 - t, z) & \dots & \partial_{z_4} G(T_1 - t, z) \\ \vdots & \vdots & & \vdots \\ -\partial_x G(T_5 - t, z) & \partial_{z_1} G(T_5 - t, z) & \dots & \partial_{z_4} G(T_5 - t, z) \end{bmatrix}.$$

But this is impossible since

$$\begin{aligned} \partial_x G(x, z) &= (-z_2 z_4 + z_3 - z_3 z_4 x) e^{-z_4 x} \\ \partial_{z_1} G(x, z) &= 1 \\ \partial_{z_2} G(x, z) &= e^{-z_4 x} \\ \partial_{z_3} G(x, z) &= x e^{-z_4 x} \\ \partial_{z_4} G(x, z) &= (-z_2 x - z_3 x^2) e^{-z_4 x} \end{aligned}$$

are linearly dependent functions in x!

The problem originates from the appearance of $\partial_x G(x, z)$ in the above matrix, which is due to the explicit t-dependence of G in (2.3). A way out is given by the new parametrization $f(t, x + t)$ since then (2.3) reads

$$G(x, Z_t) = f(t, x + t), \quad \forall x, t \in \mathbb{R}_+. \tag{2.4}$$

This however causes a new difficulty: $(f(t, x+t))_{t \in \mathbb{R}_+}$ is not an Itô process in general. So even if (2.4) can be inverted for all $x \in \mathbb{R}_+$, what kind of process is $(Z_t) = \big(G(x, \cdot)^{-1}(f(t, x + t))\big)$? We will address these issues in Chapter 4.

We now shall see how the HJM dynamics can be expressed in the language of stochastic equations in infinite dimensions (Section 1.4), allowing for the Musiela parametrization.

Let $\{S(t) \mid t \in \mathbb{R}_+\}$ denote the semigroup of right shifts which is defined by $S(t)f(x) = f(x + t)$, for any function $f : \mathbb{R}_+ \to \mathbb{R}$.

Fix $t, x \in \mathbb{R}_+$. Then (2.2) can be simply rewritten, for T replaced by $x + t$,

$$\begin{aligned} f(t, x + t) = S(t) f(0, x) &+ \int_0^t S(t - s) \alpha(s, x + s)\, ds \\ &+ \sum_{j=1}^d \int_0^t S(t - s) \sigma^j(s, x + s)\, d\beta_s^j \end{aligned} \tag{2.5}$$

where $S(t)$ acts on $f(0, x)$, $\alpha(s, x + s)$ and $\sigma^j(s, x + s)$ as functions in x. We shall work out in an axiomatic way the minimal requirements on a Hilbert space H such that (2.5) can be given a meaning when

$$r_t := f(t, \cdot + t), \quad t \in \mathbb{R}_+$$

is considered as an H-valued process. Also the HJM no arbitrage condition [**28**, Condition C.4] has to be recaptured. In particular, this means that $(P(t, T))_{t \in [0,T]}$ and $(B(t))_{t \in \mathbb{R}_+}$ have to follow Itô processes.

Clearly $H \subset L^1_{\mathrm{loc}}(\mathbb{R}_+)$, by (2.1). Observe that (2.5) is a pointwise equality for all $x \in \mathbb{R}_+$. Hence pointwise evaluation has to be well specified for $h \in H$. Moreover, we have to interchange integration and pointwise evaluation since x appears under the integral sign. Accordingly, we assume

(H1): The functions $h \in H$ are continuous[1] and the pointwise evaluation $\mathcal{J}_x(h) := h(x)$ is a continuous linear functional[2] on H, for all $x \in \mathbb{R}_+$.

Indeed, it is not very restrictive to assume that the forward curves $f(t,T)$ given by (2.2) are continuous in $T \geq t$, see [**28**, Lemma 2].

Here is an immediate consequence of assumption **(H1)**.

LEMMA 2.1. *For any $u \in \mathbb{R}_+$ there exists a number $k(u)$ such that*

$$\|\mathcal{J}_x\|_H \leq k(u), \quad \forall x \in [0,u].$$

Furthermore, $(x,h) \mapsto \mathcal{J}_x(h) : \mathbb{R}_+ \times H \to \mathbb{R}$ is jointly continuous.

PROOF. Let $u \in \mathbb{R}_+$. By **(H1)**, $\sup_{x\in[0,u]} |\mathcal{J}_x(h)| < \infty$ for all $h \in H$. Now the Banach–Steinhaus theorem [**41**, Theorem 2.6] yields the existence of $k(u)$.

Clearly, $x \mapsto \mathcal{J}_x : \mathbb{R}_+ \to H$ is weakly continuous[3]. The second assertion of the lemma follows by considering

$$|\mathcal{J}_{x_n}(h_n) - \mathcal{J}_x(h)| \leq \|\mathcal{J}_{x_n}\|_H \|h_n - h\|_H + |\mathcal{J}_{x_n}(h) - \mathcal{J}_x(h)|$$

and using the above estimate. □

In view of Section 1.4 we have to require furthermore

(H2): The semigroup $\{S(t) \mid t \in \mathbb{R}_+\}$ is strongly continuous in H with infinitesimal generator denoted by A.

For any $h \in D(A)$ we have $\frac{d}{dt}S(t)h = S(t)Ah$ in H for all $t \in \mathbb{R}_+$, see e.g. [**41**, Theorem 13.35]. By **(H1)** therefore $\frac{d}{dt}h(t) = Ah(t)$ for all $t \in \mathbb{R}_+$. Hence $Ah = h' \in H$. But this implies $h \in C^1(\mathbb{R}_+)$. We have thus shown

LEMMA 2.2. *Under* **(H1)**–**(H2)** *we have*

$$D(A) \subset \{h \in H \cap C^1(\mathbb{R}_+) \mid h' \in H\}.$$

For the coefficients in (2.5) we formulate the analogon to [**28**, Condition C.1]. At this point we generalize the hypothesis on the noise and allow from now on for an infinite dimensional underlying $(\mathcal{F}_t)$-Brownian motion $W = (\beta^j)_{j\in\mathbb{N}}$, see Section 1.1.

Write $\alpha_t(\omega) := \alpha(t,\omega,\cdot+t)$ and $\sigma = (\sigma^j)_{j\in\mathbb{N}}$ where $\sigma^j_t(\omega) := \sigma^j(t,\omega,\cdot+t)$.

(C1): The initial forward curve $r_0 = f(0,\cdot)$ lies in H.

(C2): The processes α and σ are H-, resp. $L^0_2(H)$-valued predictable and

$$\mathbb{P}\left[\int_0^t \left(\|\alpha_s\|_H + \|\sigma_s\|^2_{L^0_2(H)}\right) ds < \infty\right] = 1, \quad \forall t \in \mathbb{R}_+.$$

In the classical HJM framework $\sigma^j \equiv 0$ for $j > d$.

[1] That is, for any $h \in H$ there exists a continuous representative. We shall systematically replace the equivalence class h by its continuous representative which we still denote by h.

[2] Hence $\mathcal{J}_x$ is element of H which we identify with its dual here and subsequently.

[3] That is, continuous with respect to the weak topology on H.

Now (2.5) reads, for any $t, x \in \mathbb{R}_+$,

$$\begin{aligned}\mathcal{J}_x(r_t) &= \mathcal{J}_x(S(t)r_0) + \int_0^t \mathcal{J}_x(S(t-s)\alpha_s)\,ds + \sum_{j\in\mathbb{N}} \int_0^t \mathcal{J}_x(S(t-s)\sigma_s^j)\,d\beta_s^j \\ &= \mathcal{J}_x\left(S(t)r_0 + \int_0^t S(t-s)\alpha_s\,ds + \sum_{j\in\mathbb{N}} \int_0^t S(t-s)\sigma_s^j\,d\beta_s^j\right), \quad \mathbb{P}\text{-a.s.}\end{aligned} \tag{2.6}$$

see (1.12). For t fixed, the exceptional $\mathbb{P}$-nullset in equality (2.6) depends on x. But according to **(H1)** both sides are continuous in x. A standard argument yields equality $\mathbb{P}$-a.s. simoultaneously for all $x \in \mathbb{R}_+$. Pointwise equality of two functions yields identity of the functions, hence

$$r_t = S(t)r_0 + \int_0^t S(t-s)\alpha_s\,ds + \sum_{j\in\mathbb{N}} \int_0^t S(t-s)\sigma_s^j\,d\beta_s^j, \quad \mathbb{P}\text{-a.s.} \tag{2.7}$$

By [**18**, Proposition 6.2] the right hand side of (2.7) has a predictable version which we still denote by r_t. Consequently, r is a mild solution to the stochastic equation

$$\begin{cases} dr_t = (Ar_t + \alpha_t)\,dt + \sum_{j\in\mathbb{N}} \sigma_t^j\,d\beta_t^j \\ r_0 = f(0,\cdot). \end{cases} \tag{2.8}$$

We make an additional assumption.

(C3): There exists a continuous modification of r, still denoted by r.

Condition **(C3)** is satisfied if $S(t)$ forms a contraction semigroup in H, see [**18**, Theorem 6.10], or if we require stronger integrability of σ, see [**18**, Proposition 7.3]. Notice, however, that r is *not* an H-valued semimartingale in general.

Remark 2.3. *Combining* **(C2)**, **(C3)** *and Lemma 2.1 we conclude that the random fields* $r_t(\omega, x)$, $\alpha_t(\omega, x)$ *and* $\sigma_t^j(\omega, x)$ *are* $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}_+)$*-measurable and* $r_t(\omega, x)$ *is jointly continuous in* (t, x) *for* $\mathbb{P}$*-a.e.* ω.

So far we reformulated the classical HJM setup in the spirit of Musiela [**37**] by putting additional assumptions **(C1)** and **(C2)** on the coefficients (reflecting essentially [**28**, Conditions C.1-C.3]). Moreover, we allow for an infinite dimensional driving Brownian motion. The process r, *defined* as the continuous mild solution to (2.8), is related to $f(t, T)$ from (2.2) by

$$r_t(x) = f(t, x+t) \quad \forall x \in \mathbb{R}_+, \quad \mathbb{P}\text{-a.s.} \quad \forall t \in \mathbb{R}_+.$$

From now on, the bond prices and the savings account are given by

$$\begin{aligned} P(t,T) &= \exp\left(-\int_0^{T-t} r_t(x)\,dx\right), \quad (t,T) \in \Delta^2 \\ B(t) &= \exp\left(\int_0^t r_s(0)\,ds\right), \quad t \in \mathbb{R}_+. \end{aligned}$$

2.3. Arbitrage Free Term Structure Movements

We derive the dynamics of the discounted bond prices. Since r is not an H-valued semimartingale, we cannot use Itô formula directly but we rather have to deal with the random field $(r_t(x))_{t,x\in\mathbb{R}_+}$ and apply the stochastic Fubini theorem. We essentially adapt the technique from [**28**]. Sufficient conditions for the existence of an equivalent martingale measure, which ensures the absence of arbitrage, are provided.

First we introduce the linear functional $\mathcal{I}_u$ on H defined by

$$\mathcal{I}_u(h) := \int_0^u h(x)\,dx, \quad u \in \mathbb{R}_+.$$

There is an immediate consequence of assumption **(H1)**.

Lemma 2.4. *The linear functional $\mathcal{I}_u$ is continuous on H. In particular,*

$$\|\mathcal{I}_u\|_H \le uk(u), \quad \forall u \in \mathbb{R}_+$$

where $k(u)$ is the constant from Lemma 2.1.

Moreover, the mapping $(u,h) \mapsto \mathcal{I}_u(h) : \mathbb{R}_+ \times H \to \mathbb{R}$ is jointly continuous.

Proof. The estimate is trivial. The second assertion follows as in the proof of Lemma 2.1 by the weak continuity of $u \mapsto \mathcal{I}_u : \mathbb{R}_+ \to H$. □

It follows from **(C3)** and Lemma 2.4 that $P(t,T) = \exp(-\mathcal{I}_{T-t}(r_t))$ is $\mathcal{F}_t$-measurable and jointly continuous in $(t,T) \in \Delta^2$.

Fix $(t,T) \in \Delta^2$. Then we have by (2.7), Lemma 2.4 and (1.12)

$$\begin{aligned}
I &:= -\log P(t,T) \\
&= \mathcal{I}_{T-t}(S(t)r_0) + \int_0^t \mathcal{I}_{T-t}(S(t-s)\alpha_s)\,ds + \sum_{j\in\mathbb{N}} \int_0^t \mathcal{I}_{T-t}(S(t-s)\sigma_s^j)\,d\beta_s^j \\
&= \mathcal{I}_T(r_0) - \mathcal{I}_t(r_0) + \int_0^t \left(\mathcal{I}_{T-s}(\alpha_s) - \mathcal{I}_{t-s}(\alpha_s)\right)\,ds \\
&\quad + \sum_{j\in\mathbb{N}} \int_0^t \left(\mathcal{I}_{T-s}(\sigma_s^j) - \mathcal{I}_{t-s}(\sigma_s^j)\right)\,d\beta_s^j, \qquad \mathbb{P}\text{-a.s.}
\end{aligned}$$

where we took into account the obvious relation $\mathcal{I}_u \circ S(t) = \mathcal{I}_{t+u} - \mathcal{I}_t$. The processes $(\mathcal{I}_{T-s}(\sigma_s^j))_{s\in[0,T]}$ are predictable and

$$\sum_{j\in\mathbb{N}} \left|\mathcal{I}_{T-s}(\sigma_s^j(\omega))\right|^2 \le (Tk(T))^2 \sum_{j\in\mathbb{N}} \|\sigma_s^j(\omega)\|_H^2 < \infty, \quad \forall (s,\omega) \in [0,T]\times\Omega.$$

Hence $(\mathcal{I}_{T-s}\circ\sigma_s)_{s\in[0,T]}$ is ℓ^2-valued predictable. Similarly

$$\mathbb{P}\left[\int_0^T \sum_{j\in\mathbb{N}} |\mathcal{I}_{T-s}(\sigma_s^j)|^2\,ds < \infty\right] = 1,$$

thus $(\mathcal{I}_{T-s}\circ\sigma_s) \in \mathcal{L}_T^{\mathrm{loc}}(\mathbb{R})$. Replacing T by t, the same holds for $(\mathcal{I}_{t-s}\circ\sigma_s)_{s\in[0,t]}$. By linearity we can thus split up the integrals and write

$$I = I_1 - I_2, \quad \mathbb{P}\text{-a.s.}$$

where

$$I_1 := \mathcal{I}_T(r_0) + \int_0^t \mathcal{I}_{T-s}(\alpha_s)\,ds + \sum_{j\in\mathbb{N}} \int_0^t \mathcal{I}_{T-s}(\sigma_s^j)\,d\beta_s^j$$

$$I_2 := \mathcal{I}_t(r_0) + \int_0^t \mathcal{I}_{t-s}(\alpha_s)\,ds + \sum_{j\in\mathbb{N}} \int_0^t \mathcal{I}_{t-s}(\sigma_s^j)\,d\beta_s^j.$$

The integrals in I_2 need to be transformed. We introduce the mappings

$$\tilde{\sigma}_s^j(\omega,u) := \begin{cases} \sigma_s^j(\omega,u-s), & \text{if } s \le u \\ 0, & \text{otherwise.} \end{cases}$$

Substituting $u = x+s$, we have

$$\mathcal{I}_{t-s}(\sigma_s^j) = \int_0^{t-s} \sigma_s^j(x)\,dx = \int_s^t \sigma_s^j(u-s)\,du = \int_0^t \tilde{\sigma}_s^j(u)\,du.$$

In view of Remark 2.3, $\tilde{\sigma} = (\tilde{\sigma}^j)_{j\in\mathbb{N}}$ is measurable from $(\mathbb{R}_+\times\Omega\times\mathbb{R}_+, \mathcal{P}\otimes\mathcal{B}(\mathbb{R}_+))$ into $(\ell^2, \mathcal{B}(\ell^2))$. Moreover, by Lemma 2.1 and **(C2)**

$$\begin{aligned}\int_0^t \int_0^t \|\tilde{\sigma}_s(u)\|_{\ell^2}^2\,du\,ds &= \int_0^t \int_0^t \sum_{j\in\mathbb{N}} |\tilde{\sigma}_s^j(u)|^2\,du\,ds \\ &= \int_0^t \sum_{j\in\mathbb{N}} \left(\int_0^{t-s} |\sigma_s^j(x)|^2\,dx\right)\,ds \\ &\le t(k(t))^2 \int_0^t \sum_{j\in\mathbb{N}} \|\sigma_s^j\|_H^2\,ds < \infty, \quad \mathbb{P}\text{-a.s.}\end{aligned}$$

Consequently, the stochastic Fubini theorem 1.12 applies and, by (1.10)–(1.11),

$$\begin{aligned}\sum_{j\in\mathbb{N}} \int_0^t \mathcal{I}_{t-s}(\sigma_s^j)\,d\beta_s^j &= \sum_{j\in\mathbb{N}} \int_0^t \left(\int_0^t \tilde{\sigma}_s^j(u)\,du\right)\,d\beta_s^j = \int_0^t \left(\sum_{j\in\mathbb{N}} \int_0^t \tilde{\sigma}_s^j(u)\,d\beta_s^j\right)\,du \\ &= \int_0^t \left(\sum_{j\in\mathbb{N}} \int_0^u \sigma_s^j(u-s)\,d\beta_s^j\right)\,du, \quad \mathbb{P}\text{-a.s.}\end{aligned} \tag{2.9}$$

Using the ordinary Fubini theorem we derive similarly

$$\int_0^t \mathcal{I}_{t-s}(\alpha_s)\,ds = \int_0^t \left(\int_0^u \alpha_s(u-s)\,ds\right)\,du, \quad \mathbb{P}\text{-a.s.} \tag{2.10}$$

Combining (2.9) and (2.10) we can write, by (1.12),

$$\begin{aligned}I_2 &= \int_0^t \mathcal{J}_0\left(S(u)r_0 + \int_0^u S(u-s)\alpha_s\,ds + \sum_{j\in\mathbb{N}} \int_0^u S(u-s)\sigma_s^j\,d\beta_s^j\right)\,du \\ &= \int_0^t r_u(0)\,du, \quad \mathbb{P}\text{-a.s.}\end{aligned}$$

Taking into account $\mathcal{I}_T(r_0) = -\log P(0,T)$ we arrive at the following representation

$$\log P(t,T) = \log P(0,T) + \int_0^t \left(r_s(0) - \mathcal{I}_{T-s}(\alpha_s)\right) ds + \sum_{j\in\mathbb{N}} \int_0^t (-\mathcal{I}_{T-s}(\sigma_s^j))\, d\beta_s^j, \quad \mathbb{P}\text{-a.s.} \quad \forall (t,T) \in \Delta^2.$$

Since $P(t,T)$ is jointly continuous in $(t,T) \in \Delta^2$, the $\mathbb{P}$-nullset can be chosen for each $T \in \mathbb{R}_+$ independently of $t \in [0,T]$. Whence $(\log P(t,T))_{t\in[0,T]}$ is a continuous real-valued semimartingale. Accordingly, so is $(\log P(t,T) - \log B(t))_{t\in[0,T]}$. Applying Itô's formula to e^x gives for the bond price process

$$\begin{aligned} P(t,T) = P(0,T) &+ \int_0^t P(s,T)\left(r_s(0) - \mathcal{I}_{T-s}(\alpha_s) + \frac{1}{2}\sum_{j\in\mathbb{N}} (\mathcal{I}_{T-s}(\sigma_s^j))^2\right) ds \\ &+ \sum_{j\in\mathbb{N}} \int_0^t P(s,T)(-\mathcal{I}_{T-s}(\sigma_s^j))\, d\beta_s^j, \qquad t \in [0,T] \end{aligned} \tag{2.11}$$

and for the discounted bond price process $Z(t,T) := \frac{P(t,T)}{B(t)}$

$$\begin{aligned} Z(t,T) = P(0,T) &+ \int_0^t Z(s,T)\left(-\mathcal{I}_{T-s}(\alpha_s) + \frac{1}{2}\sum_{j\in\mathbb{N}} (\mathcal{I}_{T-s}(\sigma_s^j))^2\right) ds \\ &+ \sum_{j\in\mathbb{N}} \int_0^t Z(s,T)(-\mathcal{I}_{T-s}(\sigma_s^j))\, d\beta_s^j, \qquad t \in [0,T]. \end{aligned} \tag{2.12}$$

The following mapping will play an important role. For any continuous function f on $\mathbb{R}_+$ we define the continuous function $\mathcal{S}f : \mathbb{R}_+ \to \mathbb{R}$ by

$$(\mathcal{S}f)(x) := f(x) \int_0^x f(\eta)\, d\eta, \quad x \in \mathbb{R}_+.$$

Here is an elementary result on $\mathcal{S}$.

LEMMA 2.5. *Let $\Phi = (\Phi^j)_{j\in\mathbb{N}} \in L_2^0(H)$. Then $f(x) := \frac{1}{2}\sum_{j\in\mathbb{N}} (\mathcal{I}_x(\Phi^j))^2$ is a continuously differentiable function in $x \in \mathbb{R}_+$ with*

$$f'(x) = \sum_{j\in\mathbb{N}} (\mathcal{S}\Phi^j)(x).$$

Both series converge uniformly in x on compacts.

PROOF. Set $f_n(x) := \frac{1}{2}\sum_{j=1}^n (\mathcal{I}_x(\Phi^j))^2$ which is obviously continuously differentiable in $x \in \mathbb{R}_+$ with

$$f_n'(x) = \sum_{j=1}^n (\mathcal{S}\Phi^j)(x).$$

Moreover, for $m < n \in \mathbb{N}$

$$|f_n(x) - f_m(x)| + |f_n'(x) - f_m'(x)| \le \left(\|\mathcal{I}_x\|_H^2 + \|\mathcal{J}_x\|_H \|\mathcal{I}_x\|_H\right) \sum_{j=m+1}^n \|\Phi^j\|_H^2$$

which tends to zero for $m, n \to \infty$ uniformly in x on compacts, by Lemmas 2.1 and 2.4. Consequently, the function $f(x) = \lim_{n\to\infty} f_n(x)$ exists and is continuously differentiable in $x \in \mathbb{R}_+$ with $f'(x) = \lim_{n\to\infty} f_n'(x)$. $\square$

Below we will consider a particular Girsanov transformation from $\mathbb{P}$ to some risk neutral measure $\mathbb{Q} \sim \mathbb{P}$ on $\mathcal{F}_\infty$ which only affects a change in the drift α for the $\mathbb{Q}$-dynamics of r. So let's pretend for the moment that $\mathbb{P}$ itself is a risk neutral measure. Then the HJM drift condition [**28**, (18)] reads as follows.

LEMMA 2.6. *The discounted bond price process* $(Z(t,T))_{t\in[0,T]}$ *follows a local martingale for all* $T \in \mathbb{R}_+$ *if and only if*

$$\alpha = \sum_{j\in\mathbb{N}} \mathcal{S}\sigma^j, \quad dt \otimes d\mathbb{P}\text{-}a.s. \tag{2.13}$$

PROOF. Let $T \in \mathbb{R}_+$. By [**40**, Proposition (1.2)], $(Z(t,T))_{t\in[0,T]}$ is a local martingale if and only if the integral with respect to ds in (2.12) is indistinguishable from zero. Since $Z(s,T) > 0$, this is equivalent to

$$\mathcal{I}_{T-t}(\alpha_t) = \frac{1}{2}\sum_{j\in\mathbb{N}} (\mathcal{I}_{T-t}(\sigma_t^j))^2, \quad dt \otimes d\mathbb{P}\text{-a.s.} \tag{2.14}$$

see the proof of (5.6). The $dt \otimes d\mathbb{P}$-nullset depends on T. But by Lemma 2.5 the right hand side of (2.14) is continuously differentiable in $T \geq t$ and so is obviously the left hand side. By a standard argument we find a $dt \otimes d\mathbb{P}$-nullset $\mathcal{N} \in \mathcal{P}$ with the property that

$$\mathcal{I}_x(\alpha_t(\omega)) = \frac{1}{2}\sum_{j\in\mathbb{N}} (\mathcal{I}_x(\sigma_t^j(\omega)))^2, \quad \forall (t,\omega,x) \in \mathcal{N}^c \times \mathbb{R}_+.$$

Differentiation yields the assertion. $\square$

Observe that Lemma 2.6 imposes through (2.13) implicitly a measurability and integrability condition on $\sum_{j\in\mathbb{N}} \mathcal{S}\sigma^j$. Their validity is guaranteed by the following pair of assumptions.

(C4): The processes $\mathcal{S}\sigma^j$ are H-valued predictable.

(H3): There exists a constant K such that

$$\|\mathcal{S}h\|_H \leq K\|h\|_H^2$$

for all $h \in H$ with $\mathcal{S}h \in H$.

In general, $\mathbb{P}$ is considered as physical measure. Accordingly, the conditions of Lemma 2.6 are not satisfied under $\mathbb{P}$.

DEFINITION 2.7. *We call a measure* $\mathbb{Q}$ *on* $\mathcal{F}_\infty$ *an* equivalent local martingale measure (ELMM) *if*

i) $\mathbb{Q} \sim \mathbb{P}$ *on* $\mathcal{F}_\infty$
ii) $(Z(t,T))_{t\in[0,T]}$ *is a* $\mathbb{Q}$*-local martingale for all* $T \in \mathbb{R}_+$.

The next condition assures the existence of an ELMM, see [**28**, Condition C.4].

(C5): There exists an ℓ^2-valued predictable process $\gamma = (\gamma^j)_{j\in\mathbb{N}}$ satisfying Novikov's condition (1.14) and

$$\sum_{j\in\mathbb{N}} \gamma^j\sigma^j = \sum_{j\in\mathbb{N}} \mathcal{S}\sigma^j - \alpha, \quad dt \otimes d\mathbb{P}\text{-a.s.} \tag{2.15}$$

By Hölder's inequality and Lemma 2.1, the series

$$\mathcal{J}_x\left(\sum_{j\in\mathbb{N}}\gamma_t^j\sigma_t^j\right)=\sum_{j\in\mathbb{N}}\gamma_t^j\mathcal{J}_x(\sigma_t^j)$$

converges uniformly in x on compacts. Taking into account Lemma 2.5 and integrating, we see that (2.15) is equivalent to

$$\sum_{j\in\mathbb{N}}\gamma_t^j\,\mathcal{I}_{T-t}(\sigma_t^j)=\frac{1}{2}\sum_{j\in\mathbb{N}}\left(\mathcal{I}_{T-t}(\sigma_t^j)\right)^2-\mathcal{I}_{T-t}(\alpha_t),\quad \forall T\geq t,\quad dt\otimes d\mathbb{P}\text{-a.s.} \tag{2.16}$$

There is an intuitive interpretation of $-\gamma$ as *market price of risk*, by considering the bond price dynamics (2.11): after localization, the right hand side of (2.16) is, roughly speaking, the difference between the "average return $\mathbb{E}\left[\frac{dP(t,T)}{P(t,T)}\right]$" of the bond and the "risk-less rate" $r_t(0)$. Whence the interpretation of $-\gamma_t^j$ as market price of risk in units of the volatility, $-\mathcal{I}_{T-t}(\sigma_t^j)$, for the j-th noise factor. Whence also the expression *risk neutral measure* for $\mathbb{Q}$.

Condition **(C5)** allows for Lemma 1.14. Hence we can define $\mathbb{Q}\sim\mathbb{P}$ on $\mathcal{F}_\infty$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}}=\mathcal{E}(\gamma\cdot W)_\infty,$$

recall the definition of the stochastic exponential (1.13). Girsanov's theorem, see Proposition 1.16, yields that

$$\tilde{W}_t=W_t-\int_0^t\gamma_s\,ds,\quad t\in\mathbb{R}_+$$

is a Wiener process for the measure $\mathbb{Q}$. The dynamics of r change accordingly.

PROPOSITION 2.8. *Under the above assumptions, r satisfies*

$$r_t=S(t)r_0+\int_0^t\left(S(t-s)\sum_{j\in\mathbb{N}}\mathcal{S}\sigma_s^j\right)ds+\sum_{j\in\mathbb{N}}\int_0^t S(t-s)\sigma_s^j\,d\tilde{\beta}_s^j \tag{2.17}$$

with respect to $\mathbb{Q}$. Consequently, $\mathbb{Q}$ is an ELMM.

PROOF. The statement on (r_t) follows by combining (1.15), (2.7) and (2.15). Lemma 2.6 yields the assertion on $\mathbb{Q}$. □

2.4. Contingent Claim Valuation

This section sketches how to value contingent claims in the preceding bond market. Yet we do not address here the question of completeness, which essentially is equivalent to uniqueness of the ELMM $\mathbb{Q}$. Since the underlying noise is an infinite dimensional Brownian motion, the notion of infinite dimensional admissible trading strategies were convenient. This, however, is beyond the scope of this thesis and will be exploited elsewhere. For a treatment of the finite dimensional case we refer the reader to [**38**, Section 10.1]. Measure-valued trading strategies are introduced in [**9**].

We content ourself in postulating that $\mathbb{Q}$ is *the* risk neutral measure[4] and all prices are computed as expectations under $\mathbb{Q}$. This is justified at least if all discounted bond price processes $(Z(t,T))_{t\in[0,T]}$ follow true $\mathbb{Q}$-martingales. Since

[4]Equivalently, $-\gamma$ is *the* market price of risk.

then any attainable claim[5] will admit a unique arbitrage price. It is determined by the replicating strategy.

The subsequent introduction of the forward measure, the forward LIBOR rates and caplets is preliminary for the discussion of the popular BGM model in Section 3.6.

2.4.1. When Is the Discounted Bond Price Process a $\mathbb{Q}$-Martingale? We can give a sufficient condition for the discounted bond price process to follow a true $\mathbb{Q}$-martingale. The dynamics (2.12) read under $\mathbb{Q}$

$$Z(t,T) = P(0,T) + \sum_{j\in\mathbb{N}} \int_0^t Z(s,T)(-\mathcal{I}_{T-s}(\sigma_s^j))\, d\tilde{\beta}_s^j, \qquad t \in [0,T],$$

see (2.16). Itô's formula yields the representation as the stochastic exponential

$$Z(t,T) = P(0,T)\,\mathcal{E}\left(\sum_{j\in\mathbb{N}} \int_0^{\cdot} (-\mathcal{I}_{T-s}(\sigma_s^j))\, d\tilde{\beta}_s^j\right)_t . \tag{2.18}$$

But now we can use Novikov's criterion, see Lemma 1.14, to derive

LEMMA 2.9. *If*

$$\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\frac{1}{2}\int_0^T \|\mathcal{I}_{T-t}\circ\sigma_t\|_{\ell^2}^2\, dt\right)\right] < \infty$$

then $(Z(t,T))_{t\in[0,T]}$ *is a* $\mathbb{Q}$*-martingale.*

Another sufficient criterion is positivity of the forward rates. Indeed, if $r_t \geq 0$ for all $t \in [0,T]$ then $0 \leq Z(t,T) \leq 1$ for all $t \in [0,T]$. Moreover, a bounded local martingale is a uniformly integrable martingale. Positivity of r is related to stochastic invariance of the positive cone of functions $h \geq 0$ in H under the dynamics (2.8). We do not address this problem here and refer the reader to [**34**].

2.4.2. The Forward Measure. For pricing purposes it is convenient to use a technique called "change of numéraire", see [**20**] for a thorough treatment. We give an application for pricing bond options.

Assume that $(Z(t,T))_{t\in[0,T]}$ is a $\mathbb{Q}$-martingale. We can define the measure $\mathbb{Q}^T \sim \mathbb{Q}$ on $(\Omega,\mathcal{F}_T)$ by $\frac{d\mathbb{Q}^T}{d\mathbb{Q}} = \frac{1}{P(0,T)B(T)}$. By (2.18) we then have

$$\mathbb{E}_{\mathbb{Q}}\left[\frac{d\mathbb{Q}^T}{d\mathbb{Q}} \mid \mathcal{F}_t\right] = \frac{Z(t,T)}{P(0,T)} = \mathcal{E}\left(\sum_{j\in\mathbb{N}} \int_0^{\cdot} (-\mathcal{I}_{T-s}(\sigma_s^j))\, d\tilde{\beta}_s^j\right)_t, \quad t \in [0,T]. \tag{2.19}$$

DEFINITION 2.10. *The measure* $\mathbb{Q}^T$ *is called the* T-forward measure*.*

A claim X due at time T is in $L^1(\Omega,\mathcal{F}_T,\mathbb{Q}^T)$ if and only if $\frac{X}{B(T)} \in L^1(\Omega,\mathcal{F}_T,\mathbb{Q})$. Bayes' rule, see e.g. [**32**, Lemma 5.3, Chapter 3], yields the equality

$$P(t,T)\,\mathbb{E}_{\mathbb{Q}^T}[X \mid \mathcal{F}_t] = B(t)\,\mathbb{E}_{\mathbb{Q}}\left[\frac{X}{B(T)} \mid \mathcal{F}_t\right], \quad \mathbb{Q}\text{-a.s.} \quad \forall t \in [0,T]. \tag{2.20}$$

On the right hand side stands the time t price of the claim X. This is a very useful relation, since the left hand side allows for closed form expressions if X has a nice distribution under $\mathbb{Q}^T$. We will later use it for pricing a cap.

[5] if ever defined in a reasonable way

2.4.3. Forward LIBOR Rates. Let's introduce an important $\mathbb{Q}^T$-local martingale. We will use the notation $\mathcal{I}_u^v := \mathcal{I}_v - \mathcal{I}_u$ for $(u,v) \in \Delta$. Fix $\delta > 0$.

DEFINITION 2.11. *The* forward δ-period[6] LIBOR rate *for the future date T prevailing at time t is*

$$L(t,T) := \frac{1}{\delta}\left(e^{\mathcal{I}_{T-t}^{T+\delta-t}(r_t)} - 1\right), \quad (t,T) \in \Delta^2.$$

We can re-express $L(t,T)$ directly in term of bond prices

$$1 + \delta L(t,T) = \frac{P(t,T)}{P(t,T+\delta)}.$$

Whence the meaning of $L(t,T)$ as forward average simple rate for a loan over the future time period $[T, T+\delta]$.

Let $T \in \mathbb{R}_+$ and write $\ell(t,T) := \mathcal{I}_{T-t}^{T+\delta-t}(r_t)$. Combining Lemma 2.4 and (2.17) we conclude that $(\ell(t,T))_{t\in[0,T]}$ is a real-valued continuous semimartingale and

$$\begin{aligned}
\ell(t,T) = \ell(0,T) &+ \int_0^t \mathcal{I}_{T-t}^{T+\delta-t}\left(S(t-s)\sum_{j\in\mathbb{N}} \mathcal{S}\sigma_s^j\right) ds \\
&+ \sum_{j\in\mathbb{N}} \int_0^t \mathcal{I}_{T-t}^{T+\delta-t}\left(S(t-s)\sigma_s^j\right) d\tilde{\beta}_s^j \\
&= \ell(0,T) + \frac{1}{2}\int_0^t \sum_{j\in\mathbb{N}} \left((\mathcal{I}_{T+\delta-s}(\sigma_s^j))^2 - (\mathcal{I}_{T-s}(\sigma_s^j))^2\right) ds \\
&+ \sum_{j\in\mathbb{N}} \int_0^t \mathcal{I}_{T-s}^{T+\delta-s}(\sigma_s^j)\, d\tilde{\beta}_s^j, \qquad t \in [0,T].
\end{aligned}$$

We have used the relation $\mathcal{I}_{T-t}^{T+\delta-t} \circ S(t-s) = \mathcal{I}_{T-s}^{T+\delta-s}$ and Lemma 2.5. Itô's formula yields

$$\begin{aligned}
L(t,T) &= L(0,T) \\
&\quad + \frac{1}{2\delta}\int_0^t e^{\ell(s,T)} \sum_{j\in\mathbb{N}} \left((\mathcal{I}_{T+\delta-s}(\sigma_s^j))^2 - (\mathcal{I}_{T-s}(\sigma_s^j))^2 + (\mathcal{I}_{T-s}^{T+\delta-s}(\sigma_s^j))^2\right) ds \\
&\quad + \sum_{j\in\mathbb{N}} \frac{1}{\delta}\int_0^t e^{\ell(s,T)} \mathcal{I}_{T-s}^{T+\delta-s}(\sigma_s^j)\, d\tilde{\beta}_s^j \\
&= L(0,T) + \frac{1}{\delta}\int_0^t e^{\ell(s,T)} \sum_{j\in\mathbb{N}} \left(\mathcal{I}_{T-s}^{T+\delta-s}(\sigma_s^j)\, \mathcal{I}_{T+\delta-s}(\sigma_s^j)\right) ds \\
&\quad + \sum_{j\in\mathbb{N}} \frac{1}{\delta}\int_0^t e^{\ell(s,T)} \mathcal{I}_{T-s}^{T+\delta-s}(\sigma_s^j)\, d\tilde{\beta}_s^j, \qquad t \in [0,T].
\end{aligned}$$

Now assume that $(Z(t,T+\delta))_{t\in[0,T+\delta]}$ is a $\mathbb{Q}$-martingale. Using (2.19) and Girsanov's theorem 1.15 we conclude that

$$\beta_t^{(T+\delta),j} = \tilde{\beta}_t^j + \int_0^t \mathcal{I}_{T+\delta-s}(\sigma_s^j)\, ds, \quad t \in [0, T+\delta], \quad j \in \mathbb{N}$$

[6] Typically, δ stands for 3 or 6 months.

are independent $(\mathcal{F}_t)_{t\in[0,T+\delta]}$-Brownian motions under $\mathbb{Q}^{T+\delta}$. Consequently, the process $(L(t,T))_{t\in[0,T]}$ is a $\mathbb{Q}^{T+\delta}$-local martingale with representation

$$L(t,T) = L(0,T) + \sum_{j\in\mathbb{N}} \frac{1}{\delta} \int_0^t e^{\ell(s,T)} \mathcal{I}_{T-s}^{T+\delta-s}(\sigma_s^j)\, d\beta_s^{(T+\delta),j}, \quad t \in [0,T]. \tag{2.21}$$

2.4.4. Caps. For a thorough introduction to interest rate derivatives, such as swaps, caps, floors and swaptions, we refer to [**38**, Chapter 16]. We restrict our consideration to the following basic instrument.

A *caplet* with reset date T and settlement date $T+\delta$ pays the holder the difference between the LIBOR $L(T,T)$ and the strike rate κ. It's cash-flow at time $T+\delta$ is

$$\delta(L(T,T)-\kappa)^+.$$

A *cap* is a strip of caplets. If we can price caplets, we can price caps. The arbitrage price of the caplet at time $t \le T$ equals, see [**38**, p. 391],

$$\begin{aligned} Cpl(t) &= B(t)\mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B(T+\delta)}\delta(L(T,T)-\kappa)^+ \mid \mathcal{F}_t\right] \\ &= P(t,T+\delta)\mathbb{E}_{\mathbb{Q}^{T+\delta}}\left[\delta(L(T,T)-\kappa)^+ \mid \mathcal{F}_t\right], \end{aligned}$$

where we have used (2.20).

Suppose there exists a deterministic ℓ^2-valued measurable function $\pi(\cdot,T) = (\pi^j(\cdot,T))_{j\in\mathbb{N}}$ with $\int_0^T \|\pi(t,T)\|_{\ell^2}^2\, dt < \infty$ and such that

$$\mathcal{I}_{T-t}^{T+\delta-t}(\sigma_t^j) = (1-e^{-\ell(t,T)})\pi^j(t,T), \quad dt\otimes d\mathbb{Q}\text{-a.s. on } [0,T]\times\Omega. \tag{2.22}$$

Then (2.21) reads

$$L(t,T) = L(0,T) + \sum_{j\in\mathbb{N}} \int_0^t L(s,T)\pi^j(s,T)\, d\beta_s^{(T+\delta),j}, \quad t\in[0,T].$$

This equation is uniquely solved by the stochastic exponential

$$L(t,T) = L(0,T)\mathcal{E}\left(\sum_{j\in\mathbb{N}}\int_0^{\cdot} \pi^j(s,T)\, d\beta_s^{(T+\delta),j}\right)_t.$$

By [**18**, Lemma 10.15] there exists a real-valued standard $(\mathcal{F}_t)_{t\in[0,T]}$-Brownian motion $(\beta_t^0)_{t\in[0,T]}$ under $\mathbb{Q}^{T+\delta}$ such that

$$\sum_{j\in\mathbb{N}}\int_0^t \pi^j(s,T)\, d\beta_s^{(T+\delta),j} = \int_0^t \|\pi(s,T)\|_{\ell^2}^2\, d\beta_s^0, \quad t\in[0,T].$$

Accordingly, $L(T,T)$ is lognormally distributed under $\mathbb{Q}^{T+\delta}$, and the following pricing formula holds, see [**38**, Lemma 16.3.1].

LEMMA 2.12. *Under the above assumptions, the time t price of the caplet equals*

$$Cpl(t) = \delta P(t,T+\delta)\left(L(t,T)N(e_1(t,T)) - \kappa N(e_2(t,T))\right) \tag{2.23}$$

where

$$e_{1,2} := \frac{\log\left(\frac{L(t,T)}{\kappa}\right) \pm \frac{1}{2}v_0^2(t,T)}{v_0(t,T)}$$

$$v_0^2(t,T) := \int_t^T \|\pi(s,T)\|_{\ell^2}^2 \, ds.$$

and N stands for the standard Gaussian cumulative distribution function.

BGM [**12**] present a method for modelling the forward LIBOR rates such that (2.22) holds for all $T \in \mathbb{R}_+$. We will recapture their approach in the present setting in Section 3.6.

A Summary of Conditions

For the reader's convenience we summarize the conditions made so far. Recall the definition of $\mathcal{S}$

$$(\mathcal{S}f)(x) := f(x) \int_0^x f(\eta)\, d\eta, \quad x \in \mathbb{R}_+$$

for any continuous function $f : \mathbb{R}_+ \to \mathbb{R}$.

(H1): The functions $h \in H$ are continuous, and the pointwise evaluation $\mathcal{J}_x(h) := h(x)$ is a continuous linear functional on H, for all $x \in \mathbb{R}_+$.

(H2): The semigroup $\{S(t) \mid t \in \mathbb{R}_+\}$ is strongly continuous in H with infinitesimal generator denoted by A.

(H3): There exists a constant K such that

$$\|\mathcal{S}h\|_H \le K \|h\|_H^2$$

for all $h \in H$ with $\mathcal{S}h \in H$.

(C1): The initial forward curve $r_0 = f(0, \cdot)$ lies in H.

(C2): The processes α and σ are H-, resp. $L_2^0(H)$-valued predictable and

$$\mathbb{P}\left[\int_0^t \left(\|\alpha_s\|_H + \|\sigma_s\|_{L_2^0(H)}^2\right) ds < \infty\right] = 1, \quad \forall t \in \mathbb{R}_+.$$

(C3): There exists a continuous modification of r, still denoted by r.

(C4): The processes $\mathcal{S}\sigma^j$ are H-valued predictable.

(C5): There exists an ℓ^2-valued predictable process $\gamma = (\gamma^j)_{j\in\mathbb{N}}$ satisfying Novikov's condition (1.14) and

$$\sum_{j\in\mathbb{N}} \gamma^j \sigma^j = \sum_{j\in\mathbb{N}} \mathcal{S}\sigma^j - \alpha, \quad dt \otimes d\mathbb{P}\text{-a.s.} \tag{2.24}$$

2.5. What Is a Model?

Up to now we analyzed the stochastic behavior of a single generic forward curve process r within the HJM framework. For modelling purposes we rather want the stochastic evolution being described by a dynamical system like (2.8) with state dependent coefficients $\alpha_t(\omega) = \alpha(t, \omega, r_t(\omega))$ and $\sigma_t(\omega) = \sigma(t, \omega, r_t(\omega))$, which allows for arbitrary initial curves r_0.

From the last section it seems plausible to specify as model ingredients the volatility structure σ and a market price of risk vector $-\gamma$. We then let the drift α being *defined* by (2.24). Proposition 2.8, accordingly, assures the existence of an ELMM $\mathbb{Q}$.

As it can be seen from (2.17), γ does not appear in the $\mathbb{Q}$-dynamics of r. Thus, instead of solving (2.8) directly we rather start with the Wiener process $\tilde{W}$ under $\mathbb{Q}$, solve (2.17) and transform $\tilde{W} \rightsquigarrow W$, by Girsanov's theorem, to have γ back in the drift for the $\mathbb{P}$-dynamics of r. Compare with [**28**, Lemma 1]. That way, however, the measure $\mathbb{P}$ and the Wiener process W will depend on r implicitly by $\gamma_t(\omega) = \gamma(t, \omega, r_t(\omega))$. Yet W will be adapted to $(\mathcal{F}_t)$. Which is convenient since we never needed the filtration $(\mathcal{F}_t)$ being generated by W.

Henceforth we are given a complete filtered probability space

$$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{Q})$$

satisfying the usual assumptions, the measure $\mathbb{Q}$ being considered as the risk neutral measure. Let $\tilde{W} = (\tilde{\beta}^j)_{j\in\mathbb{N}}$ be an infinite dimensional $(\mathcal{F}_t)$-Brownian motion under $\mathbb{Q}$. The random mappings $\sigma(t, \omega, h)$ (volatility structure) and $-\gamma(t, \omega, h)$ (market price of risk) are specified in accordance with the following assumptions.

- **(D1):** The mappings σ and γ are measurable from $(\mathbb{R}_+ \times \Omega \times H, \mathcal{P} \otimes \mathcal{B}(H))$ into $(L_2^0(H), \mathcal{B}(L_2^0(H)))$, resp. $(\ell^2, \mathcal{B}(\ell^2))$.
- **(D2):** The mappings $\mathcal{S}\sigma^j$ are measurable from $(\mathbb{R}_+ \times \Omega \times H, \mathcal{P} \otimes \mathcal{B}(H))$ into $(H, \mathcal{B}(H))$.
- **(D3):** There exists a function $\Gamma \in L^2(\mathbb{R}_+)$ such that

$$\|\gamma(t, \omega, h)\|_{\ell^2} \le \Gamma(t), \quad \forall (t, \omega, h) \in \mathbb{R}_+ \times \Omega \times H.$$

Observe that assumptions **(D1)**–**(D2)** together with **(H3)** imply that $\sum_{j\in\mathbb{N}} \mathcal{S}\sigma^j$ is a measurable mapping from $(\mathbb{R}_+ \times \Omega \times H, \mathcal{P} \otimes \mathcal{B}(H))$ into $(H, \mathcal{B}(H))$.

Definition 2.13. *By the* HJMM equation *(HJM equation in the Musiela parametrization) we mean the following stochastic equation*[7] *in* H

$$\begin{cases} dr_t = (Ar_t + F_{HJM}(t, r_t))\, dt + \sigma(t, r_t)\, d\tilde{W}_t \\ r_0 = h_0 \end{cases} \tag{2.25}$$

where $F_{HJM}(t, \omega, h) := \sum_{j\in\mathbb{N}} \mathcal{S}\sigma^j(t, \omega, h)$.

Definition 2.14. *Let* σ, γ *satisfy* **(D1)**–**(D3)** *and let* $I \subset H$ *be a set of initial forward curves. We call* (σ, γ, I) *a* (local) HJM model *in* H *if for every space time initial point* $(t_0, h_0) \in \mathbb{R}_+ \times I$ *there exists a unique continuous (local) mild solution* $r = r^{(t_0,h_0)}$ *to the time* t_0*-shifted*[8] *HJMM equation (2.25).*

This definition is justified by the following result.

Theorem 2.15. *Let* (σ, γ, I) *be an HJM model. Then any space time initial point* $(t_0, h_0) \in \mathbb{R}_+ \times I$ *specifies a measure* $\mathbb{P} \sim \mathbb{Q}$ *on* $\mathcal{F}_\infty$ *and a Wiener process* W *with respect to* $(\Omega, \mathcal{F}_\infty, (\mathcal{F}_t^{(t_0)})_{t\in\mathbb{R}_+}, \mathbb{P})$ *such that* **(C1)**–**(C5)** *hold for*

$$\alpha_t(\omega) = F_{HJM}(t_0 + t, \omega, r_t(\omega)) - \sum_{j\in\mathbb{N}} \gamma^j(t_0 + t, \omega, r_t(\omega))\sigma^j(t_0 + t, \omega, r_t(\omega))$$

$$\sigma_t(\omega) = \sigma(t_0 + t, \omega, r_t(\omega))$$

$$\gamma_t(\omega) = \gamma(t_0 + t, \omega, r_t(\omega))$$

where $r = r^{(t_0,h_0)}$. *Accordingly,* $\mathbb{Q}$ *is an ELMM for this setup.*

---

[7]See Remark 1.18 for notation.

[8]recall (1.18)

PROOF. Let $r = r^{(t_0,h_0)}$ be the continuous mild solution to the time t_0-shifted version of (2.25). From **(D3)** we deduce that

$$\int_{\mathbb{R}_+} \| - \gamma(t_0 + s, \omega, r_s(\omega))\|_{\ell^2}^2 \, ds \le \|\Gamma\|_{L^2(\mathbb{R}_+)}^2 < \infty, \quad \forall \omega \in \Omega. \tag{2.26}$$

Hence Lemma 1.14 applies and we can define the measure $\mathbb{P} \sim \mathbb{Q}$ on $\mathcal{F}_\infty = \mathcal{F}_\infty^{(t_0)}$ by

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = \mathcal{E}(-\gamma \cdot \tilde{W}^{(t_0)})_\infty. \tag{2.27}$$

According to Girsanov's theorem, see Proposition 1.16,

$$W_t = \tilde{W}_t^{(t_0)} + \int_0^t \gamma_s \, ds$$

is a Wiener process under $\mathbb{P}$.

By **(H3)** and Hölder's inequality

$$\|\alpha_t\|_H \le K \sum_{j \in \mathbb{N}} \|\sigma_t^j\|_H^2 + \|\gamma_t\|_{\ell^2} \|\sigma_t\|_{L_2^0(H)} \le (K + \frac{1}{2}) \|\sigma_t\|_{L_2^0(H)}^2 + \frac{1}{2} \|\gamma_t\|_{\ell^2}^2.$$

Now (2.26) yields **(C2)**, since $(\sigma_t)_{t \in \mathbb{R}_+} \in \mathcal{L}^{\mathrm{loc}}(H)$ by the very definition of a mild solution. Using (1.15), the process r is seen to be the continuous mild solution of (2.8) with $f(0, \cdot)$ replaced by h_0, whence **(C3)**. Conditions **(C4)**–**(C5)** are implied by **(D1)**–**(D3)** and, clearly, $\mathbb{Q}$ is an ELMM for this setup. ☐

REMARK 2.16. *Theorem 2.15 points out that a* mild *solution r to the HJMM equation (2.25) is the right concept for the HJM framework. This is also in accordance with Remark 1.25. It is therefore an essential request to derive the results from* [**7**] *without assuming that r is a (local) strong solution, see* [**7**, Assumption 2.1]

REMARK 2.17. *Local HJM models are useless for pricing financial instruments. Yet in some cases we only can assert local existence and uniqueness for the HJMM equation (2.25), which is still adapted for the consistency analysis in Chapter 4. Therefore we also included the local case in Definition 2.14.*

The specification of the initial forward curve set I is convenient and allows to incorporate also the following classical examples in our framework. Let $-\gamma$ be any market price of risk satisfying **(D1)** and **(D3)**. See Sections 4.4.1–4.4.2 for terminology and results.

EXAMPLE 2.18 (CIR Model). *Here $I = \{g_0 + y g_1 \mid y > 0\}$ where g_0, g_1 are given functions, $\sigma^1 = \sigma^1(h) = \sqrt{\alpha \mathcal{J}_0(h)} g_1$ for all $h \in I$ and $\sigma^j \equiv 0$ for $j \ge 2$. Then (σ, γ, I) is a local HJM model.*

EXAMPLE 2.19 (Vasicek Model). *The volatility is constant, $\sigma^1 \equiv \sqrt{a} g_1$, where $g_1(x) = e^{\beta x}$ and $\sigma^j \equiv 0$ for $j \ge 2$. Moreover, $I = \{g_0 + y g_1 \mid y \in \mathbb{R}\}$ and g_0 is a given function. Then (σ, γ, I) is an HJM model.*

It is desirable to find easy verifiable criterions for σ to provide a (local) HJM model. We will approach this by the classical Lipschitz condition, see Section 1.4.2. But if $\sigma(t, \omega, h)$ is (locally) Lipschitz continuous in h, what about $F_{HJM}(t, \omega, h)$? In the next chapter we present a Hilbert space of forward curves which is both economically reasonable and convenient for deriving (local) existence and uniqueness results for the HJMM equation (2.25).

CHAPTER 3

The Forward Curve Spaces H_w

3.1. Definition of H_w

We introduce a class of Hilbert spaces satisfying **(H1)**–**(H3)**[1] and which are coherent with economical reasoning about the forward curve $x \mapsto r_t(x)$.

Since in practice the forward curve is obtained by smoothing data points using smooth fitting methods, see Chapters 5 and 6, it is reasonable to assume

$$\int_{\mathbb{R}_+} |r_t'(x)|^2 \, dx < \infty.$$

Moreover, the curve flattens for large time to maturity x. There is no reason to believe that the forward rate for an instantaneous loan that begins in 10 years differs much from one which begins one day later. We take this into account by penalizing irregularities of $r_t(x)$ for large x by some increasing weighting function $w(x) \geq 1$, that is,

$$\int_{\mathbb{R}_+} |r_t'(x)|^2 \, w(x) dx < \infty.$$

However, this does not define a norm yet since constant functions are not distinguished. So we add the square of the short rate $|r_t(0)|^2$.

We shall give a mathematically correct approach. Recall the fact that if $h \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$ possesses a weak derivative[2] $h' \in L^1_{\mathrm{loc}}(\mathbb{R}_+)$, then there exists an absolutely continuous representative of h, still denoted by h, such that

$$h(x) - h(y) = \int_y^x h'(u) \, du, \quad \forall x, y \in \mathbb{R}_+. \tag{3.1}$$

This can be found e.g. in [**14**, Section VIII.2]. Accordingly, the following definition makes sense.

Definition 3.1. *Let $w : \mathbb{R}_+ \to [1, \infty)$ be a non-decreasing C^1-function such that*

$$w^{-\frac{1}{3}} \in L^1(\mathbb{R}_+). \tag{3.2}$$

We write

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

and define

$$H_w := \left\{ h \in L^1_{loc}(\mathbb{R}_+) \mid \exists h' \in L^1_{loc}(\mathbb{R}_+) \text{ and } \|h\|_w < \infty \right\}.$$

The choice of H_w is established by the next theorem.

[1]see page 32

[2]uniquely specified by $\int_{\mathbb{R}_+} h(x)\varphi'(x) \, dx = -\int_{\mathbb{R}_+} h'(x)\varphi(x) \, dx$ for all $\varphi \in C^1_c((0, \infty))$

THEOREM 3.2. *The set H_w equipped with $\|\cdot\|_w$ forms a separable Hilbert space meeting* **(H1)**–**(H3)**.

Before proving the theorem we give two guiding examples of admissible weighting functions w which satisfy condition (3.2).

EXAMPLE 3.3. $w(x) = e^{\alpha x}$, *for* $\alpha > 0$.

EXAMPLE 3.4. $w(x) = (1+x)^{\alpha}$, *for* $\alpha > 3$.

PROOF. It is clear that $\|\cdot\|_w$ is a norm. Consider the separable Hilbert space $\mathbb{R}\times L^2(\mathbb{R}_+)$ equipped with the norm $\left(|\cdot|^2 + \|\cdot\|^2_{L^2(\mathbb{R}_+)}\right)^{\frac{1}{2}}$. Then the linear operator $T : H_w \to \mathbb{R}\times L^2(\mathbb{R}_+)$ given by

$$Th = \left(h(0), h' w^{\frac{1}{2}}\right), \quad h \in H_w$$

is an isometry with inverse

$$(T^{-1}(u,f))(x) = u + \int_0^x f(\eta) w^{-\frac{1}{2}}(\eta)\, d\eta, \quad (u,f) \in \mathbb{R}\times L^2(\mathbb{R}_+).$$

Hence H_w is a separable Hilbert space.

We claim that

$$\mathcal{D}_0 := \{f \in C^2(\mathbb{R}_+) \mid f' \in C^1_c(\mathbb{R}_+)\}$$

is dense in H_w. Indeed, $C^1_c(\mathbb{R}_+)$ is dense in $L^2(\mathbb{R}_+)$, see [**14**, Corollaire IV.23]. Fix $h \in H_w$ and let (f_n) be a approximating sequence of $h' w^{\frac{1}{2}}$ in $L^2(\mathbb{R}_+)$ with $f_n \in C^1_c(\mathbb{R}_+)$ for all $n \in \mathbb{N}$. Write $h_n := T^{-1}(h(0), f_n)$. Obviously, $h_n \in \mathcal{D}_0$ and $h_n \to h$ in H_w. Whence the claim.

Since w^{-1} is bounded, (3.2) implies in particular $w^{-1} \in L^1(\mathbb{R}_+)$. The subsequent Sobolev-type inequalities (3.4)–(3.8) are crucial. Throughout, h denotes a function in H_w. The constants C_1–C_4 are universal and depend only on w.

First we prove

$$\|h'\|_{L^1(\mathbb{R}_+)} \le C_1 \|h\|_w. \tag{3.3}$$

This is established by Hölder's inequality

$$\begin{aligned}\int_{\mathbb{R}_+} |h'(x)|\, dx &= \int_{\mathbb{R}_+} |h'(x)| w^{\frac{1}{2}}(x) w^{-\frac{1}{2}}(x)\, dx \\ &\le \left(\int_{\mathbb{R}_+} |h'(x)|^2\, w(x) dx\right)^{\frac{1}{2}} \left(\int_{\mathbb{R}_+} w^{-1}(x)\, dx\right)^{\frac{1}{2}} \le \|h\|_w\, \|w^{-1}\|^{\frac{1}{2}}_{L^1(\mathbb{R}_+)}.\end{aligned}$$

An important consequence of (3.3) is that $h(x)$ converges to a limit $h(\infty) \in \mathbb{R}$ for $x \to \infty$, see (3.1). This is of course a desirable property for forward curves.

For $C_2 = 1 + C_1$ it follows easily from (3.1) and (3.3) that

$$\|h\|_{L^\infty(\mathbb{R}_+)} \le C_2 \|h\|_w. \tag{3.4}$$

Write $\Pi(x) := \int_x^\infty w^{-1}(\eta)\, d\eta$. Then we have, by monotonicity of w,

$$\Pi(x) \le w^{-\frac{2}{3}}(x) \int_x^\infty w^{-\frac{1}{3}}(\eta)\, d\eta \le w^{-\frac{2}{3}}(x) \|w^{-\frac{1}{3}}\|_{L^1(\mathbb{R}_+)}, \quad x \in \mathbb{R}_+$$

and hence, by (3.2),

$$\|\Pi^{\frac{3}{2}} w\|_{L^\infty(\mathbb{R}_+)} \leq \|w^{-\frac{1}{3}}\|_{L^1(\mathbb{R}_+)} < \infty. \tag{3.5}$$

Moreover,

$$\|\Pi^{\frac{1}{2}}\|_{L^1(\mathbb{R}_+)} \leq \|w^{-\frac{1}{3}}\|^2_{L^1(\mathbb{R}_+)} < \infty. \tag{3.6}$$

By the same method as above we derive

$$\begin{aligned}\int_{\mathbb{R}_+} |h(x) - h(\infty)|\, dx &= \int_{\mathbb{R}_+} \Big| \int_x^\infty h'(\eta) w^{\frac{1}{2}}(\eta) w^{-\frac{1}{2}}(\eta)\, d\eta \Big|\, dx \\ &\leq \int_{\mathbb{R}_+} \left(\int_x^\infty |h'(\eta)|^2\, w(\eta) d\eta \right)^{\frac{1}{2}} \Pi^{\frac{1}{2}}(x)\, dx \\ &\leq \|h\|_w\, \|\Pi^{\frac{1}{2}}\|_{L^1(\mathbb{R}_+)}.\end{aligned}$$

Hence, by (3.6),

$$\|h - h(\infty)\|_{L^1(\mathbb{R}_+)} \leq C_3 \|h\|_w. \tag{3.7}$$

Finally we compute, using Hölder's inequality again,

$$\begin{aligned}\int_{\mathbb{R}_+} |h(x) - h(\infty)|^4\, w(x) dx &= \int_{\mathbb{R}_+} \Big| \int_x^\infty h'(\eta) w^{\frac{1}{2}}(\eta) w^{-\frac{1}{2}}(\eta)\, d\eta \Big|^4 w(x) dx \\ &\leq \int_{\mathbb{R}_+} \left(\int_{\mathbb{R}_+} |h'(\eta)|^2\, w(\eta) d\eta \right)^2 \Pi^2(x)\, w(x) dx \\ &\leq \|h\|_w^4\, \|\Pi^{\frac{3}{2}} w\|_{L^\infty(\mathbb{R}_+)}\, \|\Pi^{\frac{1}{2}}\|_{L^1(\mathbb{R}_+)}.\end{aligned}$$

Therefore, by (3.5)–(3.6),

$$\|(h - h(\infty))^4 w\|_{L^1(\mathbb{R}_+)} \leq C_4 \|h\|_w^4. \tag{3.8}$$

Now we can prove **(H1)**–**(H3)**. Estimate (3.4) yields continuity of the linear functional $\mathcal{J}_x$ for all $x \in \mathbb{R}_+$. Since by the preceding remarks, see (3.1), H_w consists of continuous functions, **(H1)** is established.

Let $h \in H_w$ and $\varphi \in C_c^1((0,\infty))$. Then the shift semigroup $S(t)$ satisfies

$$\begin{aligned}\int_{\mathbb{R}_+} S(t)h(x)\varphi'(x)\, dx &= \int_{\mathbb{R}_+} h(x)\varphi'(x-t)\, dx = -\int_{\mathbb{R}_+} h'(x)\varphi(x-t)\, dx \\ &= -\int_{\mathbb{R}_+} h'(x+t)\varphi(x)\, dx\end{aligned} \tag{3.9}$$

since $\varphi(\cdot - t) \in C_c^1((t,\infty))$. Consequently, $(S(t)h)'$ exists and equals $S(t)h'$. In view of (3.4) it follows

$$\|S(t)h\|_w^2 = |h(t)|^2 + \int_{\mathbb{R}_+} |h'(x+t)|^2\, w(x) dx \leq (C_2^2 + 1)\|h\|_w^2. \tag{3.10}$$

We have used the monotonicity of w. Hence $S(t)h \in H_w$ and $S(t)$ is bounded for all $t \in \mathbb{R}_+$.

We have to show strong continuity of $S(t)$. We start with the following observation. By (3.1)

$$h(x+t) - h(x) = t \int_0^1 h'(x+st)\, ds, \quad h \in H_w. \tag{3.11}$$

Let $g \in \mathcal{D}_0$. Since $g' \in H_w$, (3.10) and (3.11) yield

$$\begin{aligned}\|S(t)g - g\|_w^2 &= |g(t) - g(0)|^2 + \int_{\mathbb{R}_+} |g'(x+t) - g'(x)|^2\, w(x)dx \\ &\le |g(t) - g(0)|^2 + t^2 \int_0^1 \int_{\mathbb{R}_+} |g''(x+st)|^2 w(x)\, dx\, ds \\ &\le |g(t) - g(0)|^2 + t^2 \int_0^1 \|S(st)g'\|_w^2\, ds \to 0 \quad \text{for } t \to 0.\end{aligned}$$

Hence $S(t)$ is strongly continuous on $\mathcal{D}_0$. But for any $h \in H_w$ and $\epsilon > 0$ there exists $g \in \mathcal{D}_0$ with $\|h - g\|_w \le \frac{\epsilon}{4(C_2+1)}$. Combining this with (3.10) yields

$$\begin{aligned}\|S(t)h - h\|_w &\le \|S(t)(h-g)\|_w + \|S(t)g - g\|_w + \|g - h\|_w \\ &\le (C_2+1)\frac{\epsilon}{4(C_2+1)} + \|S(t)g - g\|_w + \frac{\epsilon}{4(C_2+1)} \le \epsilon\end{aligned}$$

for t small enough. We conclude that $S(t)$ is strongly continuous on H_w, whence the validity of **(H2)**.

It remains to prove the estimate in **(H3)**. Let $h \in H_w$. Similarly to (3.9) one shows that $(\mathcal{S}h)'$ exists with

$$(\mathcal{S}h)'(x) = h'(x) \int_0^x h(\eta)\, d\eta + h^2(x).$$

We define

$$H_w^0 := \{h \in H_w \mid h(\infty) = 0\}. \tag{3.12}$$

By (3.4), H_w^0 is a closed subspace of H_w. We claim that

$$\mathcal{S}h \in H_w \iff h \in H_w^0. \tag{3.13}$$

Assume that $h(\infty) > 0$. Then there exists $x_0 \in \mathbb{R}_+$ such that $h(x) \ge \frac{h(\infty)}{2}$ for all $x \ge x_0$. But then

$$\mathcal{S}h(x) \ge h(x) \int_0^{x_0} h(\eta)\, d\eta + \left(\frac{h(\infty)}{2}\right)^2 (x - x_0), \quad \forall x \ge x_0$$

which explodes for $x \to \infty$. The same reasoning applies if $h(\infty) < 0$. Whence the necessity in (3.13).

Conversely, assume that $h(\infty) = 0$. Using (3.7) and (3.8) we estimate

$$\begin{aligned}\|\mathcal{S}h\|_w^2 &= \int_{\mathbb{R}_+} \left| h'(x) \int_0^x h(\eta)\, d\eta + h^2(x) \right|^2 w(x)dx \\ &\le 2 \int_{\mathbb{R}_+} \left\{ |h'(x)|^2 \left(\int_0^x |h(\eta)|\, d\eta \right)^2 + h^4(x) \right\} w(x)dx \\ &\le 2(\|h\|_{L^1(\mathbb{R}_+)}^2 \|h\|_w^2 + C_4 \|h\|_w^4) \le 2(C_3^2 + C_4)\|h\|_w^4.\end{aligned}$$

Hence $\mathcal{S}h \in H_w$ and (3.13) is proved. Moreover, we conclude that **(H3)** holds with $K = \sqrt{2(C_3^2 + C_4)}$. □

REMARK 3.5. *It is not difficult to see that* **(H1)** *and* **(H2)** *hold without assumption (3.2). Yet it makes the proof more handy.*

The preceding proof allows for identifying the full generator A of $S(t)$ in H_w.

COROLLARY 3.6. *We have* $D(A) = \{h \in H_w \mid h' \in H_w\}$ *and* $Ah = h'$.

PROOF. Write $\mathcal{D} := \{h \in H_w \mid h' \in H_w\}$. By Lemma 2.2, $D(A) \subset \mathcal{D}$. Conversely, let $h \in \mathcal{D}$. Applying (3.11) we get

$$\frac{h'(x+t) - h'(x)}{t} - h''(x) = \int_0^1 (h''(x+st) - h''(x))\, ds.$$

Consequently,

$$\begin{aligned} n(t) &:= \int_{\mathbb{R}_+} \Big| \frac{h'(x+t) - h'(x)}{t} - h''(x) \Big|^2 w(x)dx \\ &\leq \int_0^1 \int_{\mathbb{R}_+} |h''(x+st) - h''(x)|^2\, w(x)dx\, ds \\ &\leq \int_0^1 \|S(st)h' - h'\|_w^2\, ds \to 0 \quad \text{for } t \to 0 \end{aligned}$$

by dominated convergence, see (3.10). Since h' is continuous on $\mathbb{R}_+$, see (3.1), we conclude

$$\Big\| \frac{S(t)h - h}{t} - h' \Big\|_w^2 = \Big| \frac{h(t) - h(0)}{t} - h'(0) \Big|^2 + n(t) \to 0 \quad \text{for } t \to 0,$$

whence $h \in D(A)$. □

Most important for the HJMM equation is the following result.

COROLLARY 3.7. *The mapping* $\mathcal{S} : H_w^0 \to H_w^0$ *is locally Lipschitz continuous, that is*

$$\|\mathcal{S}g - \mathcal{S}h\|_w \leq C_5(\|g\|_w + \|h\|_w)\, \|g - h\|_w, \quad \forall g, h \in H_w^0$$

where the constant C_5 *only depends on* w.

PROOF. In view of (3.13) and (3.7) it's immediate that $\mathcal{S}h \in H_w^0$ for all $h \in H_w^0$. Let $g, h \in H_w^0$. Using Hölder's inequality and (3.7)–(3.8) we calculate

$$\begin{aligned} \|\mathcal{S}g - \mathcal{S}h\|_w^2 &= \int_{\mathbb{R}_+} \Big| g'(x) \int_0^x g(\eta)\, d\eta - h'(x) \int_0^x h(\eta)\, d\eta + g^2(x) - h^2(x) \Big|^2 w(x)dx \\ &\leq 3 \int_{\mathbb{R}_+} |g'(x)|^2 \Big| \int_0^x (g(\eta) - h(\eta))\, d\eta \Big|^2 w(x)dx \\ &\quad + 3 \int_{\mathbb{R}_+} \Big| \int_0^x h(\eta)\, d\eta \Big|^2 |g'(x) - h'(x)|^2\, w(x)dx \\ &\quad + 3 \int_{\mathbb{R}_+} (g(x) + h(x))^2 w^{\frac{1}{2}}(x) (g(x) - h(x))^2 w^{\frac{1}{2}}(x)\, dx \\ &\leq 3 \left(\|g\|_w^2\, \|g - h\|_{L^1(\mathbb{R}_+)}^2 + \|h\|_{L^1(\mathbb{R}_+)}^2\, \|g - h\|_w^2 \right) \\ &\quad + 3\, \|(g+h)^4 w\|_{L^1(\mathbb{R}_+)}^{\frac{1}{2}}\, \|(g-h)^4 w\|_{L^1(\mathbb{R}_+)}^{\frac{1}{2}} \\ &\leq 3(C_3^2 + 2C_4)(\|g\|_w^2 + \|h\|_w^2)\, \|g - h\|_w^2 \end{aligned}$$

and we have used the universal inequality

$$|x_1 + \cdots + x_k|^2 \leq k \left(|x_1|^2 + \cdots + |x_k|^2\right), \quad k \in \mathbb{N}. \tag{3.14}$$

Setting $C_5 = \sqrt{3(C_3^2 + 2C_4)}$ yields the result. $\square$

3.2. Volatility Specification

Working in the space H_w we can give simple criteria for $\sigma(t, \omega, h)$ to provide a (local) HJM model.

Condition **(D2)** and (3.13) clearly force $\sigma^j \in H_w^0$, recall (3.12). This is of course a desirable property since we don't want the forward rates $r_t(x)$ being too volatile for large x, see (2.6). In fact, Corollary 3.7 yields

LEMMA 3.8. *Let σ meet* **(D1)** *with $H = H_w$ and $L_2^0(H)$ replaced by $L_2^0(H_w^0)$. Then* **(D2)** *is satisfied as well.*

From now on we require that σ satisfies the assumptions of Lemma 3.8. In view of Corollary 1.24 it's enough to require local Lipschitz continuity of $\sigma(t, \omega, h)$ and $F_{HJM}(t, \omega, h)$ in h for asserting local existence and uniqueness for the HJMM equation (2.25). The choice of H_w gains in interest if we consider the following property.

Let $\sigma(t, \omega, h)$ be locally Lipschitz continuous in h with Lipschitz constant $C = C(R)$ and suppose

$$\|\sigma(t, \omega, 0)\|_{L_2^0(H_w)} \leq \tilde{C}, \quad \forall (t, \omega) \in \mathbb{R}_+ \times \Omega \tag{3.15}$$

for a constant $\tilde{C}$. This clearly gives

$$\|\sigma(t, \omega, h)\|_{L_2^0(H_w)} \leq \tilde{C} + C(R)R, \quad \forall h \in B_R(H_w), \quad \forall R \in \mathbb{R}_+. \tag{3.16}$$

Whence σ is locally bounded.

LEMMA 3.9. i) *Let σ satisfy the above assumptions. Then $F_{HJM}(t, \omega, h)$ is locally Lipschitz continuous in h.*

ii) *Suppose in addition that $\sigma(t, \omega, h)$ is Lipschitz continuous in h and uniformly bounded:*

$$\|\sigma(t, \omega, h)\|_{L_2^0(H_w)} \leq C, \quad \forall (t, \omega, h) \in \mathbb{R}_+ \times \Omega \times H_w, \tag{3.17}$$

for some constant C. Then the same holds true for $F_{HJM}(t, \omega, h)$.

PROOF. For simplicity of notation we abandon temporarily the (t, ω)-dependence of σ and F_{HJM}.

Let $h_1, h_2 \in B_R(H_w)$ and write $\Delta := \|F_{HJM}(h_1) - F_{HJM}(h_2)\|_w$. By Corollary 3.7 and Hölder's inequality

$$\begin{aligned}
\Delta &\leq \sum_{j \in \mathbb{N}} \|\mathcal{S}\sigma^j(h_1) - \mathcal{S}\sigma^j(h_2)\|_w \\
&\leq C_5 \sum_{j \in \mathbb{N}} \left(\|\sigma^j(h_1)\|_w + \|\sigma^j(h_2)\|_w\right) \|\sigma^j(h_1) - \sigma^j(h_2)\|_w \\
&\leq C_5 \left(\|\sigma(h_1)\|_{L_2^0(H_w)} + \|\sigma(h_2)\|_{L_2^0(H_w)}\right) \|\sigma(h_1) - \sigma(h_2)\|_{L_2^0(H_w)} \\
&\leq C_5 C(R) \left(\|\sigma(h_1)\|_{L_2^0(H_w)} + \|\sigma(h_2)\|_{L_2^0(H_w)}\right) \|h_1 - h_2\|_w.
\end{aligned}$$

Combining this with (3.16) proves i) and the first part of ii). The boundedness assertion on F_{HJM} is a consequence of **(H3)**. $\square$

Combining Lemmas 3.8–3.9 with Theorem 1.23 and Corollary 1.24 we get our main existence and uniqueness result for the HJMM equation. Denote by $-\gamma$ a market price of risk vector meeting **(D1)** and **(D3)**.

THEOREM 3.10. i) *Let σ be as in Lemma 3.9.i). Then (σ, γ, H_w) is a local HJM model in H_w.*

ii) *Let σ be as in Lemma 3.9.ii). Then (σ, γ, H_w) is an HJM model in H_w and all discounted bond price processes $(Z(t,T))_{t\in[0,T]}$ are true $\mathbb{Q}$-martingales. Moreover, any mild solution $r^{(t_0,h_0)}$ of the time t_0-shifted HJMM equation (2.25) is also a weak solution, $(t_0, h_0) \in \mathbb{R}_+ \times H_w$.*

PROOF. Only the statement on $(Z(t,T))_{t\in[0,T]}$ remains to be proved. But this is a straightforward consequence of Lemma 2.9 and the estimate

$$\|\mathcal{I}_{T-t} \circ \sigma(t, r_t)\|_{\ell^2}^2 \leq (Tk(T))^2 \|\sigma(t, r_t)\|_{L_2^0(H_w)}^2 \leq (Tk(T))^2 C^2,$$

see Lemma 2.4. □

We shall examine three types of possible volatility specifications. The first one corresponds to the classical HJM framework where the volatility depends locally on the state variable. It turns out that in this case one merely gets local HJM models. For example the one which is presented in [**28**, Section 7].

The second type is much more convenient. Here the volatility depends on a functional value of the state variable. It is easy to find sufficient conditions for providing an HJM model in this case.

Lastly, we recapture the BGM [**12**] model which we already mentioned in Section 2.4.4. However, since also this type of volatility is of local nature, we merely get a local HJM model.

Subsequently, there follows an intermediary section where we introduce an important financial variable.

3.3. The Yield Curve

We already know by Riesz representation theorem that the functionals $\mathcal{J}_x$ and $\mathcal{I}_x$ can be identified with elements in H_w, see Theorem 3.2 and Lemma 2.4. Let's just mention how they look like as functions.

LEMMA 3.11. *We have*

$$\mathcal{J}_x(\eta) = 1 + \int_0^{x\wedge\eta} \frac{1}{w(u)}\, du, \quad \eta \in \mathbb{R}_+$$
$$\mathcal{I}_x(\eta) = x + \int_0^{\eta} \frac{x - u \wedge x}{w(u)}\, du, \quad \eta \in \mathbb{R}_+$$

as elements in H_w.

PROOF. We only proof the statement on $\mathcal{I}_x$. Let $h \in H_w$. Integration by parts yields

$$\begin{aligned}\langle \mathcal{I}_x, h\rangle_w &= xh(0) + \int_{\mathbb{R}_+} (x - \eta \wedge x) h'(\eta)\, d\eta \\ &= xh(0) + \int_0^x (x-\eta) h'(\eta)\, d\eta \\ &= xh(0) - xh(0) + \int_0^x h(\eta)\, d\eta = \int_0^x h(\eta)\, d\eta.\end{aligned}$$

□

The next result will prove very useful. Recall that A^* denotes the adjoint of A.

LEMMA 3.12. *We have $\mathcal{I}_x \in D(A^*)$ and $A^*\mathcal{I}_x = \mathcal{J}_x - \mathcal{J}_0$ for all $x \in \mathbb{R}_+$.*

PROOF. From Corollary 3.6 we know that

$$\langle \mathcal{I}_x, Ah\rangle_w = \int_0^x h'(\eta)\,d\eta = h(x) - h(0) = \langle \mathcal{J}_x - \mathcal{J}_0, h\rangle_w, \quad \forall h \in D(A).$$

□

Besides the forward curve and the term structure of bond prices there is another measurement of the bond market in between.

DEFINITION 3.13. *Given the continuous time t forward curve $x \mapsto r_t(x)$, the time t* yield curve *$x \mapsto y_t(x)$ is defined by*

$$y_t(x) := \frac{1}{x}\int_0^x r_t(\eta)\,d\eta, \quad x \in \mathbb{R}_+.$$

Notice that $y_t(0) = r_t(0)$ is well defined.

In contrast to forward rates, yields of zero coupon bonds can be directly observed on the market. We will exploit this fact in Example 3.18 below.

For completeness we shall describe the implied dynamics for y. It turns out that the yield curve process enjoys nice regularity properties.

Let σ satisfy the assumptions of Lemma 3.9.ii) and let r be a weak solution of the HJMM equation (2.25) in H_w. For a fixed $x > 0$ we have by Lemma 3.12 that $(xy_t(x)) = (\mathcal{I}_x(r_t))$ is a real-valued semimartingale. Taking into account Lemma 2.5 we derive the decomposition

$$\begin{aligned} xy_t(x) &= \mathcal{I}_x(r_0) + \int_0^t \Big((\mathcal{J}_x - \mathcal{J}_0)(r_s) + \frac{1}{2}\sum_{j\in\mathbb{N}}(\mathcal{I}_x(\sigma^j(s,r_s)))^2\Big)\,ds \\ &\quad + \sum_{j\in\mathbb{N}}\int_0^t \mathcal{I}_x(\sigma^j(s,r_s))\,d\tilde{\beta}^j_s. \end{aligned}$$

Hence the random field $(Y(t,x)) := (xy_t(x))$ is a classical solution (that is, (t,x)-pointwise) of the stochastic partial differential equation (SPDE)

$$\begin{cases} dY(t,x) = \Big(\frac{\partial}{\partial x}Y(t,x) - r_t(0) + \frac{1}{2}\sum_{j\in\mathbb{N}}(\mathcal{I}_x(\sigma^j(t,r_t)))^2\Big)\,dt + \sum_{j\in\mathbb{N}}\mathcal{I}_x(\sigma^j(t,r_t))\,d\tilde{\beta}^j_t \\ Y(0,x) = \mathcal{I}_x(r_0). \end{cases}$$

3.4. Local State Dependent Volatility

HJM present in [**28**, Section 7] a class of models with local state dependent volatility. In the present setting this reads

$$\sigma^j(t,\omega,h)(x) = \Phi^j(t,\omega,x,h(x)), \quad j \in \mathbb{N} \tag{3.18}$$

for some real valued functions $\Phi^j(t,\omega,x,s)$. We shall formulate sufficient conditions on $(\Phi^j)_{j\in\mathbb{N}}$ such that $\sigma = (\sigma^j)_{j\in\mathbb{N}}$ given by (3.18) satisfies the hypotheses of

Lemma 3.9.i). Writing $f(t,T) = r_t(T-t)$ for a mild solution $r = r^{(0,h_0)}$ to the HJMM equation (2.25) (if it exists) we get

$$f(t,T) = h_0(T) + \int_0^t \sum_{j\in\mathbb{N}} \left(\Phi^j(s, T-s, f(s,T)) \int_s^T \Phi^j(s, T-s, f(s,u))\, du \right) ds$$
$$+ \sum_{j\in\mathbb{N}} \int_0^t \Phi^j(s, T-s, f(s,T))\, d\tilde{\beta}_s^j \tag{3.19}$$

which is [**28**, Equation (42)].

Proposition 3.14. *Let $\Phi^j(t,\omega,x,s) : \mathbb{R}_+ \times \Omega \times \mathbb{R}_+ \times \mathbb{R} =: D \to \mathbb{R}$, $j \in \mathbb{N}$, satisfy the following assumptions*

i) *$\Phi^j(t,\omega,x,s)$ is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R})$-measurable*
ii) *$\Phi^j(t,\omega,x,s)$ is continuously differentiable in (x,s) for all $(t,\omega,x,s) \in D$*
iii) *$\lim_{x\to\infty} \Phi^j(t,\omega,x,s) = 0$ for all $(t,\omega,s) \in \mathbb{R}_+ \times \Omega \times \mathbb{R}$.*

Suppose, moreover, there exist a continuous mapping $c = (c^j)_{j\in\mathbb{N}} : \mathbb{R} \to \ell^2$ with c^j convex and $c^j(-s) = c^j(s)$, and a measurable mapping $\Psi = (\Psi^j)_{j\in\mathbb{N}} : \mathbb{R}_+ \to \ell^2$ with

$$\int_{\mathbb{R}_+} \|\Psi(x)\|_{\ell^2}^2\, w(x) dx < \infty$$

meeting

iv) $|\Phi^j(t,\omega,0,s)| \le c^j(s)$
v) $|\partial_x \Phi^j(t,\omega,x,s)| \le \Psi^j(x) c^j(s)$
vi) $|\partial_s \Phi^j(t,\omega,x,s)| \le c^j(s)$
vii) $|\Phi^j(t,\omega,0,s_1) - \Phi^j(t,\omega,0,s_2)| \le (c^j(s_1) + c^j(s_2))|s_1 - s_2|$
viii) $|\partial_x \Phi^j(t,\omega,x,s_1) - \partial_x \Phi^j(t,\omega,x,s_2)| \le \Psi^j(x)(c^j(s_1) + c^j(s_2))|s_1 - s_2|$
ix) $|\partial_s \Phi^j(t,\omega,x,s_1) - \partial_s \Phi^j(t,\omega,x,s_2)| \le (c^j(s_1) + c^j(s_2))|s_1 - s_2|$

for all $(t,\omega,x,s) \in D$, $s_1, s_2 \in \mathbb{R}$ and $j \in \mathbb{N}$. Then $\sigma = (\sigma^j)_{j\in\mathbb{N}}$ defined by (3.18) fulfills the hypotheses of Lemma 3.9.i).

Proof. Whenever there is no ambiguity we shall skip the notional dependence on (t,ω). The proof is divided into four steps.

Step 1: Let's recall the chain rule for the weak derivative. For $h \in H_w$ we have

$$\frac{d}{dx}\Phi^j(x,h(x)) = \partial_x \Phi^j(x,h(x)) + \partial_s \Phi^j(x,h(x)) h'(x).$$

This can be shown using ii) as in the proof of [**14**, Corollaire VIII.10].

Step 2: Now we show that $\sigma(h) \in L_2^0(H_w^0)$ for all $h \in H_w$.

Let $h \in H_w$. Clearly $\lim_{x\to\infty} \sigma^j(h)(x) = 0$, by ii) and iii). Assumptions iv)– vi) yield

$$\begin{aligned}\|\sigma^j(h)\|_w^2 &\le |\Phi^j(0,h(0))|^2 \\ &\quad + 2\int_{\mathbb{R}_+} \left(|\partial_x \Phi^j(x,h(x))|^2 + |\partial_s \Phi^j(x,h(x))|^2 |h'(x)|^2 \right) w(x) dx \\ &\le |c^j(\|h\|_w)|^2 \left(1 + 2\int_{\mathbb{R}_+} \left(|\Psi^j(x)|^2 + |h'(x)|^2 \right) w(x) dx \right)\end{aligned}$$

which implies the assertion.

Step 3: Next we check measurability of $(t,\omega,h) \mapsto \sigma^j(t,\omega,h)$. It's enough to show measurability of $(t,\omega,h) \mapsto \langle \sigma^j(t,\omega,h), g\rangle_w$ for all $g \in H_w$, see [**18**, Proposition 1.3].

Fix $g \in H_w$. By i), ii) and Lemma 2.1 the mapping

$$(t,\omega,x,h) \mapsto \partial_x \Phi^j(t,\omega,x,h(x))$$

is $\mathcal{P}\otimes\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{B}(H_w)$-measurable. The same applies for $\partial_s\Phi^j(t,\omega,x,h(x))$. Hence for any $f \in H_w$ the function

$$(t,\omega,h) \mapsto I(t,\omega,h)(f) := \int_{\mathbb{R}_+} \partial_s\Phi^j(t,\omega,x,h(x)) f'(x) g'(x)\, w(x)dx$$

is $\mathcal{P}\otimes\mathcal{B}(H_w)$-measurable. Here we used vi) and a monotone class argument, see [**4**, Lemma 23.2]. Since $f \in H_w$ was arbitrary this implies that $(t,\omega,h) \mapsto I(t,\omega,h)$ is a $\mathcal{P}\otimes\mathcal{B}(H_w)$-measurable application into the dual space of H_w which we identify with H_w as usual (Notice that $I(t,\omega,h)$ *is* a bounded linear functional on H_w, by vi)). Similar reasoning, using v), shows that

$$(t,\omega,h) \mapsto J(t,\omega,h) := \int_{\mathbb{R}_+} \partial_x\Phi^j(t,\omega,x,h(x)) g'(x)\, w(x)dx$$

is $\mathcal{P}\otimes\mathcal{B}(H_w)$-measurable into $\mathbb{R}$.

But now we can write

$$\langle \sigma^j(t,\omega,h), g\rangle_w = \Phi^j(t,\omega,0,\mathcal{J}_0(h))g(0) + J(t,\omega,h) + I(t,\omega,h)(h)$$

and deduce the desired measurability assertion.

Step 4: It remains to check the boundedness and Lipschitz properties. Observe that (3.15) is a direct consequence of iv)–v). Let $h_1, h_2 \in H_w$. In view of (3.14) we have

$$\begin{aligned}
\|\sigma^j(h_1) - \sigma^j(h_2)\|_w^2 \le{} & |\Phi^j(0,h_1(0)) - \Phi^j(0,h_2(0))|^2 \\
& + 3\int_{\mathbb{R}_+} |\partial_x\Phi^j(x,h_1(x)) - \partial_x\Phi^j(x,h_2(x))|^2\, w(x)dx \\
& + 3\int_{\mathbb{R}_+} |\partial_s\Phi^j(x,h_1(x))\,(h_1'(x) - h_2'(x))\,|^2\, w(x)dx \\
& + 3\int_{\mathbb{R}_+} |h_2'(x)\left(\partial_s\Phi^j(x,h_1(x)) - \partial_s\Phi^j(x,h_2(x))\right)|^2\, w(x)dx \\
\le{} & C(h_2)\left(c^j(\|h_1\|_w) + c^j(\|h_2\|_w)\right)^2 \|h_1 - h_2\|_w^2
\end{aligned}$$

where

$$C(h_2) = 1 + 3\int_{\mathbb{R}_+} \|\Psi(x)\|_{\ell^2}^2\, w(x)dx + 3 + 3\|h_2\|_w^2$$

and we have used vi)–ix). Therefore $\sigma(t,\omega,h)$ is locally Lipschitz continuous in h, and the proposition follows.

□

The following two examples are presented in [**28**, Section 7].

EXAMPLE 3.15. *Let $\phi \in H_w^0$. Define $\Phi^1(t,\omega,x,s) \equiv \Phi^1(x,s) := \phi(x)\,s$ and $\Phi^j \equiv 0$ for $j > 1$. The assumptions of Proposition 3.14 are easily seen to be fulfilled with $c^1(s) = (1+\|\phi\|_w)(1+|s|)$ and $\Psi^1(x) = |\phi'(x)|$, and $c^j = \Psi^j \equiv 0$ for $j > 1$. According to Theorem 3.10, the so specified volatility mapping σ provides a local*

*HJM model. However, it is shown in Morton's thesis [**36**] that every mild solution r to the HJMM equation (2.25) explodes with probability one in finite time.*

We define

$$\chi(s) := \frac{s}{|s|}\left(|s| \wedge \lambda\right), \quad s \in \mathbb{R} \tag{3.20}$$

where λ stands for some positive number.

EXAMPLE 3.16. *Like Example 3.15, but this time $\Phi^1(x,s) := \phi(x)\chi(s)$. A slight modification of the preceding proof shows that also here σ fulfills the hypotheses of Lemma 3.9.i). It is proved in* [**28**, Proposition 4] *that (3.19) has a jointly continuous solution $(f(t,T))_{(t,T)\in\Delta^2}$. Yet we still cannot assert global existence for the HJMM equation (2.25) since the norm $\|r_t\|_w$ is not a priori controllable.*

In view of Example 3.16 one may wish to consider the HJMM equation (2.25) in a larger space like $L^2_\alpha(\mathbb{R}_+)$ equipped with norm

$$\|f\|^2_{L^2_\alpha(\mathbb{R}_+)} := \int_{\mathbb{R}_+} |f(x)|^2\, e^{-\alpha x} dx, \quad \alpha > 0.$$

This approach is used by Goldys and Musiela [**27**]. Indeed, Example 3.16 provides a global mild solution to equation (2.25) if interpreted in $L^2_\alpha(\mathbb{R}_+)$. However, $L^2_\alpha(\mathbb{R}_+)$ does not meet **(H1)** and **(H3)** and it is not clear how the analysis of Chapter 2 could be performed in this space.

3.5. Functional Dependent Volatility

Very convenient for the present framework is a functional dependent volatility

$$\sigma^j(t,\omega,h)(x) = \Phi^j(t,\omega,x,\zeta^j(h)), \quad j \in \mathbb{N} \tag{3.21}$$

where ζ^j are continuous linear functionals on H_w and $\Phi^j(t,\omega,x,s)$ are real-valued functions. The following proposition contains appropriate conditions on Φ^j to make sure that $\sigma = (\sigma^j)_{j\in\mathbb{N}}$ in (3.21) provides an HJM model.

PROPOSITION 3.17. *Let $\Phi^j(t,\omega,x,s) : \mathbb{R}_+ \times \Omega \times \mathbb{R}_+ \times \mathbb{R} =: D \to \mathbb{R}$, $j \in \mathbb{N}$, satisfy the following assumptions*

i) $\Phi^j(t,\omega,x,s)$ *is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{B}(\mathbb{R})$-measurable*
ii) $\Phi^j(t,\omega,x,s)$ *is continuously differentiable in x*
iii) $\lim_{x\to\infty} \Phi^j(t,\omega,x,s) = 0$ *for all $(t,\omega,s) \in \mathbb{R}_+ \times \Omega \times \mathbb{R}$*
iv) $|\partial_x \Phi^j(t,\omega,x,s)| \le \Psi^j(x)$
v) $|\partial_x \Phi^j(t,\omega,x,s_1) - \partial_x \Phi^j(t,\omega,x,s_2)| \le \Psi^j(x)|s_1 - s_2|$

for all $(t,\omega,x,s) \in D$, $s_1, s_2 \in \mathbb{R}$, and $\Psi = (\Psi^j)_{j\in\mathbb{N}}$ as in Proposition 3.14. Then $\sigma = (\sigma^j)_{j\in\mathbb{N}}$ defined by (3.21) fulfills the hypotheses of Lemma 3.9.ii).

PROOF. We proceed as in the proof of Proposition 3.14. Again we skip the notional dependence on (t,ω) whenever there is no ambiguity.

Assumptions ii), iii) and iv) yield $\Phi^j(\cdot,s) \in H_w^0$ and, using Hölder's inequality,

$$\begin{aligned}\|\Phi^j(\cdot,s)\|_w^2 &\le |\Phi^j(0,s)|^2 + \int_{\mathbb{R}_+} |\Psi^j(x)|^2\, w(x)dx \\ &\le \left(\int_{\mathbb{R}_+} |\partial_x \Phi^j(x,s)|\, dx\right)^2 + \int_{\mathbb{R}_+} |\Psi^j(x)|^2\, w(x)dx \\ &\le \left(\int_{\mathbb{R}_+} |\Psi^j(x)|^2\, w(x)dx\right)\left(\|w^{-1}\|_{L^1(\mathbb{R}_+)} + 1\right)\end{aligned}$$

for all $s \in \mathbb{R}$. From this we deduce $\sigma(h) \in L_2^0(H_w^0)$ for all $h \in H_w$. Moreover, (3.17) holds.

By a similar argument as in Step 3 in the proof of Proposition 3.14, using i) and ii), we derive the measurability of

$$\begin{aligned}(t,\omega,h) \mapsto \langle \sigma^j(t,\omega,h), g\rangle_w &= \Phi^j(t,\omega,0,\zeta^j(h))g(0) \\ &\quad + \int_{\mathbb{R}_+} \partial_x \Phi^j(t,\omega,x,\zeta^j(h))g'(x)\, w(x)dx\end{aligned}$$

for arbitrary $g \in H_w$.

It remains to show that $\sigma(t,\omega,h)$ is Lipschitz continuous in h. But this is a direct consequence of v) and iii) which implies the equality

$$\Phi^j(t,\omega,0,s) = -\int_{\mathbb{R}_+} \partial_x \Phi^j(t,\omega,x,s)\, dx.$$

□

The following example looks promising.

EXAMPLE 3.18. *Fix $m \in \mathbb{N}$ and m* benchmark yields*, see Definition 3.13,*

$$\frac{1}{\delta_1}\mathcal{I}_{\delta_1}(r_t), \ldots, \frac{1}{\delta_m}\mathcal{I}_{\delta_m}(r_t), \quad 0 < \delta_1 < \cdots < \delta_m,$$

which are directly observable *on the market. Let $\phi^1, \ldots, \phi^m \in H_w^0$ and set*

$$\sigma^j(t,\omega,h) \equiv \sigma^j(h) := \begin{cases} \phi^j\, \chi\left(\frac{1}{\delta_j}\mathcal{I}_{\delta_j}(h)\right), & j \le m \\ 0, & j > m \end{cases}$$

see (3.20). Then Proposition 3.17 applies with

$$\Phi^j(t,\omega,x,s) \equiv \Phi^j(x,s) = \begin{cases} \phi^j(x)\chi(s), & j \le m \\ 0, & j > m \end{cases}$$

and $\Psi^j = \lambda\left|\left(\phi^j\right)'\right|$ for $j \le m$, resp. $\Psi^j = 0$ for $j > m$.

The functions $\phi^1, \ldots, \phi^m$ can be inferred by statistical methods.

3.6. The BGM Model

A prominent representative within the HJM framework is the BGM [**12**] model which was already motivated in Section 2.4.4. It is desirable to have it included in the present setup. In this section we construct therefore a volatility structure σ satisfying the assumptions of Lemma 3.9.i) and which meets (2.22).

Recall the notation from Section 2.4. Assume that we want to price any caplet with settlement date $T+\delta \le T^*$ according to formula (2.23), for some $T^* \in \mathbb{R}_+$. Without loss of generality $T^* = K^*\delta$ for $K^* \in \mathbb{N}$. Motivated by (2.22) we are looking for a family of deterministic functions $(\theta^j(t,x))_{j\in\mathbb{N}}$ such that for any $(t,h) \in \mathbb{R}_+ \times H_w$

$$\mathcal{I}_x^{x+\delta}(\sigma^j(t,h)) = \left(1 - e^{-\mathcal{I}_x^{x+\delta}(h)}\right)\theta^j(t,x), \quad \forall x \in [0,(K^*-1)\delta]. \tag{3.22}$$

We suppress temporarily the notional dependence on t and j. Suppose for the moment that $\theta(x)$ is arbitrarily smooth in x. Taking derivative in (3.22) we get

$$\sigma(h)(x+\delta) - \sigma(h)(x) = \phi(h,x), \quad \forall x \in [0,(K^*-1)\delta), \tag{3.23}$$

where

$$\phi(h,x) := (h(x+\delta)-h(x))e^{-\mathcal{I}_x^{x+\delta}(h)}\theta(x) + \left(1 - e^{-\mathcal{I}_x^{x+\delta}(h)}\right)\partial_x\theta(x).$$

Recurrence relationship (3.23) defines a volatility structure $\sigma(h)(x)$ for $x \in [\delta, K^*\delta)$ provided $\sigma(h)(x)$ is defined on $[0,\delta)$. Notice that (3.23) implies equality (3.22) only up to a integration constant C which is specified by setting $x=0$

$$\mathcal{I}_\delta(\sigma(h)) = (1 - e^{-\mathcal{I}_\delta(h)})\,\theta(0) + C.$$

We will later set $\theta(0)=0$, see (3.24). Thus $C=0$ if and only if $\mathcal{I}_\delta(\sigma(h))=0$ for all $h \in H_w$. Following BGM [**12**] we set $\sigma(h)(x)=0$ on $[0,\delta)$. Solving (3.23) we derive

$$\sigma(h)(x) = \sum_{i=1}^{k}\phi(h,x-i\delta) \quad \text{for } x \in [k\delta,(k+1)\delta) \text{ and } 1 \le k \le K^*-1.$$

This implies $\lim_{x\uparrow(k+1)\delta}\sigma(h)(x) = \sum_{i=1}^k \phi(h,i\delta)$ and we see that $\sigma(h)(x)$ is continuous in $x \in [0,K^*\delta)$ for all $h \in H_w$ if and only if $\phi(h,0)=0$ for all $h \in H_w$. That is, if and only if

$$\theta(0) = \partial_x\theta(0) = 0. \tag{3.24}$$

We have to extend the definition of $\sigma(h)(x)$ to the whole $\mathbb{R}_+$. Set

$$\sigma(h)(K^*\delta) := \lim_{x\uparrow K^*\delta}\sigma(h)(x) = \sum_{i=1}^{K^*-1}\phi(h,i\delta)$$

and extrapolate linearly to 0

$$\sigma(h)(x) := \sigma(h)(K^*\delta)\left(1 - \frac{(x-K^*\delta)}{\delta}\right)^+, \quad x \ge K^*\delta.$$

This construction implies that $\sigma(h)'$ exists in $L^1_{\mathrm{loc}}(\mathbb{R}_+)$ and

$$\sigma(h)'(x) = \begin{cases} 0, & x \in [0,\delta) \\ \sum_{i=1}^k \partial_x\phi(h,x-i\delta), & x \in (k\delta,(k+1)\delta), \quad 1 \le k \le K^*-1 \\ -\frac{1}{\delta}\sigma(h)(K^*\delta), & x \in (K^*\delta,(K^*+1)\delta) \\ 0, & x > (K^*+1)\delta. \end{cases} \tag{3.25}$$

We have to check whether $\sigma(h) \in H_w^0$ and local Lipschitz continuity. In view of (3.25) we compute

$$\begin{aligned}\partial_x \phi(h,x) &= (h'(x+\delta) - h'(x)) e^{-\mathcal{I}_x^{x+\delta}(h)} \theta(x) \\ &\quad - (h(x+\delta) - h(x))^2 e^{-\mathcal{I}_x^{x+\delta}(h)} \theta(x) \\ &\quad + 2(h(x+\delta) - h(x)) e^{-\mathcal{I}_x^{x+\delta}(h)} \partial_x \theta(x) \\ &\quad + \left(1 - e^{-\mathcal{I}_x^{x+\delta}(h)}\right) \partial_x^2 \theta(x).\end{aligned}$$

The following estimates are basic. From (3.14) we obtain

$$\begin{aligned}\frac{|\partial_x \phi(h,x)|^2}{4} \le \left(1 + e^{2\delta\|h\|_w}\right) \Bigg(& 2(|h'(x+\delta)|^2 + |h'(x)|^2) \, |\theta(x)|^2 \\ & + \left(\int_x^{x+\delta} |h'(\eta)| \, d\eta \right)^4 |\theta(x)|^2 \\ & + 4 \left(\int_x^{x+\delta} |h'(\eta)| \, d\eta \right)^2 |\partial_x \theta(x)|^2 \\ & + 2 |\partial_x^2 \theta(x)|^2 \Bigg)\end{aligned}$$

and thus by (3.3)

$$\begin{aligned}\frac{|\partial_x \phi(h,x)|^2}{4} \le \left(1 + e^{2\delta\|h\|_w}\right) \Bigg(& 2(|h'(x+\delta)|^2 + |h'(x)|^2) \, |\theta(x)|^2 \\ & + C_1^4 \|h\|_w^4 \, |\theta(x)|^2 \\ & + 4 C_1^2 \|h\|_w^2 \, |\partial_x \theta(x)|^2 \\ & + 2 |\partial_x^2 \theta(x)|^2 \Bigg).\end{aligned} \tag{3.26}$$

Let $h_1, h_2 \in H_w$. We decompose $\Delta := \partial_x \phi(h_1, x) - \partial_x \phi(h_2, x)$ according to

$$\begin{aligned}\Delta &= (h_1'(x+\delta) - h_1'(x)) \left(e^{-\mathcal{I}_x^{x+\delta}(h_1)} - e^{-\mathcal{I}_x^{x+\delta}(h_2)}\right) \theta(x) \\ &\quad + e^{-\mathcal{I}_x^{x+\delta}(h_2)} (h_1'(x+\delta) - h_2'(x+\delta) + h_1'(x) - h_2'(x)) \theta(x) \\ &\quad - \left(\int_x^{x+\delta} h_1'(\eta) \, d\eta \right)^2 \left(e^{-\mathcal{I}_x^{x+\delta}(h_1)} - e^{-\mathcal{I}_x^{x+\delta}(h_2)}\right) \theta(x) \\ &\quad - e^{-\mathcal{I}_x^{x+\delta}(h_2)} \left(\int_x^{x+\delta} (h_1'(\eta) + h_2'(\eta)) \, d\eta \right) \left(\int_x^{x+\delta} (h_1'(\eta) - h_2'(\eta)) \, d\eta \right) \theta(x) \\ &\quad + 2 \left(\int_x^{x+\delta} h_1'(\eta) \, d\eta \right) \left(e^{-\mathcal{I}_x^{x+\delta}(h_1)} - e^{-\mathcal{I}_x^{x+\delta}(h_2)}\right) \partial_x \theta(x) \\ &\quad + 2 e^{-\mathcal{I}_x^{x+\delta}(h_2)} \left(\int_x^{x+\delta} (h_1'(\eta) - h_2'(\eta)) \, d\eta \right) \partial_x \theta(x) \\ &\quad - \left(e^{-\mathcal{I}_x^{x+\delta}(h_1)} - e^{-\mathcal{I}_x^{x+\delta}(h_2)}\right) \partial_x^2 \theta(x).\end{aligned}$$

Using

$$\left|e^{-\mathcal{I}_x^{x+\delta}(h_1)} - e^{-\mathcal{I}_x^{x+\delta}(h_2)}\right| \le \delta\|h_1 - h_2\|_w e^{\delta(\|h_1\|_w + \|h_2\|_w)},$$

(3.3) and (3.14) we get

$$\begin{aligned}\frac{|\Delta|^2}{7} \le e^{2\delta(\|h_1\|_w+\|h_2\|_w)}\Big(&2(|h_1'(x+\delta)|^2 + |h_1'(x)|^2)\delta^2\|h_1-h_2\|_w^2\,|\theta(x)|^2\\ &+ 2(|h_1'(x+\delta) - h_2'(x+\delta)|^2 + |h_1'(x) - h_2'(x)|^2)|\theta(x)|^2\\ &+ C_1^4\|h_1\|_w^4\delta^2\|h_1-h_2\|_w^2\,|\theta(x)|^2\\ &+ 2C_1^4(\|h_1\|_w^2 + \|h_2\|_w^2)\|h_1-h_2\|_w^2\,|\theta(x)|^2\\ &+ 4C_1^2\|h_1\|_w^2\delta^2\|h_1-h_2\|_w^2\,|\partial_x\theta(x)|^2\\ &+ 4C_1^2\|h_1-h_2\|_w^2\,|\partial_x\theta(x)|^2\\ &+ \delta^2\|h_1-h_2\|_w^2\,|\partial_x^2\theta(x)|^2\Big).\end{aligned} \tag{3.27}$$

In the sequel, C denotes a real constant which only depends on w, δ and K^* but represents different values from line to line.

In view of (3.25)

$$\begin{aligned}\int_{\mathbb{R}_+} |\sigma(h)'(x)|^2\,w(x)dx = &\int_0^{K^*\delta} |\sigma(h)'(x)|^2\,w(x)dx\\ &+ \frac{|\sigma(h)(K^*\delta)|^2}{\delta^2}\int_{K^*\delta}^{(K^*+1)\delta} w(x)dx.\end{aligned}$$

But Hölder's inequality yields

$$|\sigma(h)(K^*\delta)|^2 \le \left(\int_0^{K^*\delta} |\sigma(h)'(x)|\,dx\right)^2 \le C\left(\int_0^{K^*\delta} |\sigma(h)'(x)|^2\,w(x)dx\right),$$

and so, by (3.14) and (3.25)

$$\begin{aligned}\int_{\mathbb{R}_+} |\sigma(h)'(x)|^2\,w(x)dx \le C &\int_0^{K^*\delta} |\sigma(h)'(x)|^2\,w(x)dx\\ &= C\sum_{k=1}^{K^*-1}\int_{k\delta}^{(k+1)\delta} |\sum_{i=1}^{k}\partial_x\phi(h,x-i\delta)|^2\,w(x)dx\\ &\le C\sum_{k=1}^{K^*-1}\sum_{i=1}^{k}\int_{k\delta}^{(k+1)\delta} |\partial_x\phi(h,x-i\delta)|^2\,w(x)dx\\ &\le Cw(K^*\delta)\int_I |\partial_x\phi(h,x)|^2\,dx.\end{aligned} \tag{3.28}$$

where $I := [0,(K^*-1)\delta]$. In the same manner we can see that

$$\int_{\mathbb{R}_+} |\sigma(h_1)'(x) - \sigma(h_2)'(x)|^2\,w(x)dx \le Cw(K^*\delta)\int_I |\partial_x\phi(h_1,x) - \partial_x\phi(h_2,x)|^2\,dx. \tag{3.29}$$

From (3.26) and (3.28) it follows now easily

$$\begin{aligned}\|\sigma(h)\|_w^2 &\le C\left(1+e^{2\delta\|h\|_w}\right)\left(\|\theta\|_{L^\infty(I)}^2+\|\theta\|_{L^2(I)}^2+\|\partial_x\theta\|_{L^2(I)}^2+\|\partial_x^2\theta\|_{L^2(I)}^2\right)\\ &\le C\left(1+e^{2\delta\|h\|_w}\right)\|\partial_x^2\theta\|_{L^2(I)}^2\end{aligned} \tag{3.30}$$

and we have used (3.24). Combining (3.27) and (3.29) gives similarly

$$\|\sigma(h_1)-\sigma(h_2)\|_w^2 \le Ce^{2\delta(\|h_1\|_w+\|h_2\|_w)}\|h_1-h_2\|_w^2\,\|\partial_x^2\theta\|_{L^2(I)}^2. \tag{3.31}$$

We are thus led to the following result.

THEOREM 3.19. *Let $(\theta^j(t,x))_{j\in\mathbb{N}}$ satisfy*

i) $(t,x)\mapsto\theta^j(t,x):\mathbb{R}_+\times I\to\mathbb{R}$ *is jointly measurable*
ii) $\theta^j(t,\cdot)\in C^2(I),\quad \forall t\in\mathbb{R}_+$
iii) $\theta^j(t,0)=\partial_x\theta^j(t,0)=0,\quad \forall t\in\mathbb{R}_+$
iv) $\sup_{t\in\mathbb{R}_+}\sum_{j\in\mathbb{N}}\|\partial_x^2\theta^j(t,\cdot)\|_{L^2(I)}^2<\infty$.

Then the above construction yields a measurable mapping

$$\sigma=(\sigma^j)_{j\in\mathbb{N}}:(\mathbb{R}_+\times H_w,\mathcal{B}(\mathbb{R}_+)\otimes\mathcal{B}(H_w))\to(L_2^0(H_w^0),\mathcal{B}(L_2^0(H_w^0)))$$

which meets (3.22) and the assumptions of Lemma 3.9.i).

PROOF. The structure and arguments of the proof are similar to those of the proof of Proposition 3.14. We shall give a sketch.

Condition iii) is (3.24). Estimate (3.30) and iv) clearly yield $\sigma(t,h)\in L_2^0(H_w^0)$ and (3.15). Let $g\in H_w$. Measurability of $(t,h)\mapsto\langle\sigma^j(t,h),g\rangle_w$ follows, using i) and ii), by similar arguments as in Step 3 in the proof of Proposition 3.14.

Finally, iv) and (3.31) imply local Lipschitz continuity of $\sigma(t,h)$ in h. □

The so found local HJM model is a slight modification of the BGM [**12**] model. The difference is that (2.22) holds for $T\le T^*-\delta$ with $\pi^j(t,T)=\theta^j(t,T-t)$. Consequently, only caplets with settlement date $T+\delta\le T^*$ can be priced according to formula (2.23). From a practical point of view this is no restriction at all, since T^* can be chosen arbitrarily large. From the mathematical point of view this restriction is essential: $T^*=K^*\delta$ has to be finite in estimates (3.28) and (3.29).

Unfortunately, we cannot prove existence of an HJM model which has property (3.22). In contrast to the preceding result, BGM [**12**] can assert global existence for equation (2.25) if considered as an SPDE, see [**12**, Corollary 2.1]. Yet their SPDE methods give no information about $\|r_t\|_w$. Moreover, they do not provide a general framework for (local) HJM models as in the present thesis.

In summary, we have shown that the BGM [**12**] model is (essentially) recaptured by our setting. In particular, it is available for the consistency results in Chapter 4.

CHAPTER 4

Consistent HJM Models

4.1. Consistency Problems

In practice, the forward curve $x \mapsto r_t(x)$ is often estimated by a parametrized family $\mathcal{G} = \{G(\cdot, z) \mid z \in \mathcal{Z}\}$ of functions in H_w with parameter set $\mathcal{Z} \subset \mathbb{R}^m$, $m \in \mathbb{N}$. By slight abuse of notation, we use the same letter G for the mapping $z \mapsto G(\cdot, z) : \mathcal{Z} \to H_w$. The parameter $z \in \mathcal{Z}$ is daily chosen in such a way that the forward curve $x \mapsto G(x, z)$ optimally fits the data available on the bond and swap market. This means that one observes the time t evolution $(r_t)_{t\in\mathbb{R}_+}$ of the forward curve by its $\mathcal{Z}$-valued trace process $(Z_t)_{t\in\mathbb{R}_+}$ given by

$$G(Z_t) = r_t. \tag{4.1}$$

Relation (4.1) implies $r_t \in \mathcal{G}$. If we assume that r is given by an HJM model (σ, γ, I) in H_w then this requires a certain consistency between (σ, γ, I) and the family $\mathcal{G}$. This problem was first formulated by Björk and Christensen [**7**]. Two questions naturally arise:

i) What are the conditions on (σ, γ, I) such that for any $(t_0, h_0) \in \mathbb{R}_+ \times (I \cap \mathcal{G})$ there exists a non-trivial $\mathcal{Z}$-valued process Z and (4.1) is (locally) satisfied for $r = r^{(t_0,h_0)}$?
ii) If so, what kind of process is Z?

Both questions are related to invertibility of (4.1) and have been studied in the subsequent articles for exponential-polynomial families (Chapters 5–6) and for the regular but general case where $\mathcal{G}$ is a submanifold in H_w (Chapter 7). Accordingly, we considered different concepts of consistency.

Definition 7.8 straightly extends to the set $\mathcal{M} = \mathcal{G}$.

Definition 4.1. *A local HJM model (σ, γ, I) in H_w is* consistent *with $\mathcal{G}$ if*

i) $\mathcal{G} \subset I$
ii) *$\mathcal{G}$ is locally invariant for the HJMM equation (2.25) equipped with σ.*

This definition replies to question i).

Remark 4.2. *It is evident that consistency is a property of σ and I only. This fact is also reflected by the invariance of the Nagumo type consistency conditions (7.2)–(7.3) with respect to the Girsanov transformation (2.27).*

Here is a stronger version of the previous definition, which corresponds to question ii).

Definition 4.3. *A local HJM model (σ, γ, I) in H_w is* r-consistent[1] *with $\mathcal{G}$ if*

i) $\mathcal{G} \subset I$

[1] This terminology is in accordance with Definition 4.1 in [**7**].

ii) *For any* $(t_0, h_0) \in \mathbb{R}_+ \times \mathcal{G}$ *there exists a* $\mathcal{Z}$*-valued Itô process*[2]

$$Z_t = Z_t^{(t_0,h_0)} = Z_0 + \int_0^t b_s\, ds + \int_0^t \rho_s\, d\tilde{W}_s^{(t_0)}$$

such that $G(Z)$ *is a local mild solution to the time* t_0*-shifted HJMM equation (2.25) equipped with* σ *and initial condition* $G(Z_0) = h_0$*.*

Any such state space process Z *in turn is called* consistent *with* $\mathcal{G}$ *if* $G(Z)$ *is a global mild solution.*[3]

By uniqueness of $r = r^{(t_0,h_0)}$, r-consistency of (σ, γ, I) with $\mathcal{G}$ implies (4.1) for r and $Z = Z^{(t_0,h_0)}$ and therefore implies consistency. The converse is not true in general. In Chapter 7 we give criterions for the latter to hold. A restatement of Corollary 3.7 and Lemma 3.9 yields their applicability. Conditions **(A1)**–**(A5)** are formulated at the end of Section 7.2.

Lemma 4.4. *Suppose that* σ *satisfies the assumptions of Lemma 3.9.i). Then* **(A4)** *is satisfied. If in addition* $\sigma(t, \omega, h)$ *is right continuous in* $t \in \mathbb{R}_+$ *for all* $h \in H_w$ *and* $\omega \in \Omega$*, then* **(A5)** *holds too.*

We can now rephrase Theorem 7.13 as follows.

Theorem 4.5. *Let* σ *satisfy* **(A2)** *and the hypotheses of Lemma 4.4. Suppose that* $\mathcal{G}$ *is an* m*-dimensional* C^2 *submanifold of* H_w*. Then r-consistency of* (σ, γ, I) *with* $\mathcal{G}$ *is equivalent to consistency.*

If $\mathcal{G}$ *is linear, then the same conclusion can be drawn without assumption* **(A2)***.*

There is a subtle but important difference between $\mathcal{G}$ being a *regular* or merely an *immersed* submanifold.

Assume that G is a C^2 immersion at some $z_0 \in \mathcal{Z}$. From Proposition A.1 we deduce that there exists an open neighborhood V of z_0 in $\mathcal{Z}$ such that $G(V)$ is an m-dimensional C^2 submanifold in H_w and $G : V \to G(V)$ is a parametrization. But $\mathcal{G}$ can have self-intersections or can accumulate on $G(V)$ as shown in Figure 1.

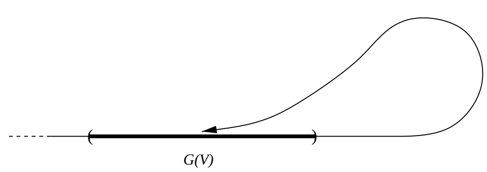

Figure 1

Thus, in principle, $G(V)$ may very well fail to be consistent with (σ, γ, I) but still (σ, γ, I) is consistent with $\mathcal{G}$. The same reasoning applies for seeing that consistency does not imply r-consistency: even if $\mathcal{G}$ is an immersed submanifold, the trace process Z may have a jump at z_0 if $G(z_0)$ is an accumulation point of $\mathcal{G}$.

[2]for some $\mathbb{R}^m$-, resp. $L_2^0(\mathbb{R}^m)$-valued $(\mathcal{F}_t^{(t_0)})$-predictable processes b and $\rho = (\rho^j)_{j\in\mathbb{N}}$ with

$$\int_0^t \left(\|b_s\|_{\mathbb{R}^m} + \|\rho_s\|^2_{L_2^0(\mathbb{R}^m)} \right) ds < \infty, \quad \mathbb{Q}\text{-a.s.} \quad \forall t \in \mathbb{R}_+$$

[3]Compare with Definition 5.1, resp. Definition 6.1.

On the other hand it is true that r-consistency with $\mathcal{G}$ implies r-consistency with $G(V)$. If furthermore σ satisfies the hypotheses of Theorem 4.5, then Theorem 7.13 applies. We shall express the consistency conditions (7.2)–(7.3) in local coordinates, see (4.4) below. But first we have to check differentiability of G.

4.2. A Simple Criterion for Regularity of G

Consider $G \in C^2(\mathbb{R}_+ \times \mathcal{Z})$ with $G(z) \equiv$ [4] $G(\cdot, z) \in H_w$ for all $z \in \mathcal{Z}$. We can give a simple criterion for differentiability of G in H_w.

Let $\xi_1, \dots, \xi_m$ denote the standard basis in $\mathbb{R}^m$ and let $z_0 \in \mathcal{Z}$. The *partial derivative* $D_i G(z_0)$ is the derivative of the mapping $s \mapsto G(z_0 + s\xi_i) : \mathbb{R} \to H_w$ in $s = 0$. For the differential calculus in Banach spaces we refer to [**1**, Chapter 2]. By **(H1)** we have

$$\mathcal{J}_x \left(\frac{d}{ds} G(z_0 + s\xi_i)|_{s=0} \right) = \frac{d}{ds} \mathcal{J}_x \left(G(z_0 + s\xi_i) \right) |_{s=0} = \partial_{z_i} G(x, z_0), \quad x \in \mathbb{R}_+.$$

Therefore $D_i G(z_0) = \partial_{z_i} G(\cdot, z_0)$ if it exists.

LEMMA 4.6. *Let $\tilde{w} : \mathbb{R}_+ \to [1, \infty)$ be a non-decreasing function such that $w\tilde{w}^{-1} \in L^1$ and let V be an open subset of $\mathcal{Z}$. Suppose*

$$\sup_{(x,z) \in \mathbb{R}_+ \times V} |\partial_{z_i} \partial_x G(x, z)|^2 \tilde{w}(x) < \infty.$$

Then the partial derivatives $D_i G(z)$ exists for all $z \in V$ and $z \mapsto D_i G(z)$ is continuous from V into H_w, for all $1 \le i \le m$. Consequently, $G \in C^1(V; H_w)$ and

$$DG(z) = (\partial_{z_1} G(\cdot, z), \dots, \partial_{z_m} G(\cdot, z)) \in L(\mathbb{R}^m; H_w). \tag{4.2}$$

PROOF. Let $z \in V$. Write $\epsilon_i = \epsilon \xi_i$, where $\epsilon > 0$. We have to show that

$$T(\epsilon) := \left\| \frac{G(z + \epsilon_i) - G(z)}{\epsilon} - \partial_{z_i} G(\cdot, z) \right\|_w^2 \to 0 \quad \text{for } \epsilon \to 0.$$

We decompose $T(\epsilon) = T_1(\epsilon) + T_2(\epsilon)$, where

$$T_1(\epsilon) := \left| \frac{G(0, z + \epsilon_i) - G(0, z)}{\epsilon} - \partial_{z_i} G(0, z) \right|^2 \to 0 \quad \text{for } \epsilon \to 0$$

and

$$\begin{aligned} T_2(\epsilon) &:= \int_{\mathbb{R}_+} \left| \frac{\partial_x G(x, z + \epsilon_i) - \partial_x G(x, z)}{\epsilon} - \partial_{z_i} \partial_x G(x, z) \right|^2 w(x) dx \\ &\le \int_{\mathbb{R}_+} \int_0^1 \left| \partial_{z_i} \partial_x G(x, z + s\epsilon_i) - \partial_{z_i} \partial_x G(x, z) \right|^2 w(x) \, ds \, dx. \end{aligned}$$

For ϵ small enough we have $z + s\epsilon_i \in V$ for all $s \in [0, 1]$ and by assumption

$$\left| \partial_{z_i} \partial_x G(x, z + s\epsilon_i) - \partial_{z_i} \partial_x G(x, z) \right|^2 w(x) \le 2C w(x) \tilde{w}^{-1}(x) \in L^1([0, 1] \times \mathbb{R}_+)$$

where C is independent of x and ϵ. Hence $T_2(\epsilon) \to 0$ for $\epsilon \to 0$, by dominated convergence. We conclude that $D_i G(z)$ exists. Continuity of $z \mapsto D_i G(z) : V \to H_w$ follows similarly. This proves the first part of the lemma.

For the second part we refer to [**1**, Proposition 2.4.12]. □

[4] Terminology introduced at the beginning of Section 4.1.

By [**1**, Proposition 2.4.12], $G \in C^2(\mathcal{Z}; H_w)$ if and only if $D_i G \in C^1(\mathcal{Z}; H_w)$ for $1 \leq i \leq m$, which again can be checked with Lemma 4.6. Moreover

$$D^2 G(z) = \big(\partial_{z_i} \partial_{z_j} G(\cdot, z)\big)_{1 \leq i,j \leq m} \tag{4.3}$$

which is a bilinear mapping from $\mathbb{R}^m \times \mathbb{R}^m$ into H_w.

Assume that $G \in C^2(\mathcal{Z}; H_w)$ and that $\partial_{z_1} G(\cdot, z_0), \ldots, \partial_{z_m} G(\cdot, z_0)$ are linearly independent functions, for some $z_0 \in \mathcal{Z}$. Then G is a C^2 immersion at z_0 and we can go on with the reasoning at the end of the preceding section. Thus let σ meet the hypotheses of Theorem 4.5 and let (σ, γ, I) be r-consistent with $\mathcal{G}$.

By Theorem 7.9 necessarily $G(V) \in D(A)$. Combining (4.2)–(4.3) with (7.41)–(7.42) and Theorem 7.13 yields the consistency condition in local coordinates

$$\begin{aligned} \partial_x G(x,z) = &\sum_{i=1}^m b^i(t,\omega,z) \partial_{z_i} G(x,z) + \frac{1}{2} \sum_{k,l=1}^m a^{k,l}(t,\omega,z) \partial_{z_k} \partial_{z_l} G(x,z) \\ &- \sum_{k,l=1}^m a^{k,l}(t,\omega,z) \partial_{z_k} G(x,z) \int_0^x \partial_{z_l} G(\eta, z)\, d\eta \end{aligned} \tag{4.4}$$

for all $(t, \omega, z) \in \mathbb{R}_+ \times \Omega \times V$ (after removing a nullset from Ω). Here $a^{k,l} := \sum_{j \in \mathbb{N}} \rho^{j,k} \rho^{j,l}$ denotes the diffusion matrix and b the drift vector of the coordinate process.

Compare condition (4.4) with Propositions 5.2 and 6.2. It is just the stronger (since pointwise) version of (5.4) and (6.3), respectively, which is due to the different meanings of "consistency" in Chapters 5–6 and Chapter 7, see Definition 4.3.

If $\mathcal{G} = G(\mathcal{Z})$ is a regular submanifold, then condition (4.4) can be used to check whether *any* local HJM model is consistent with $\mathcal{G}$. In some cases this leads to the complete characterization of all consistent local HJM models. So for regular exponential-polynomial families.

4.3. Regular Exponential-Polynomial Families

We restate the analysis of Chapters 5–6 from the preceding point of view. As mentioned at the end of the previous section, this needs a slight modification.

Throughout this section we choose as weighting function $w(x) = (1+x)^\alpha$ for some $\alpha > 3$, see Example 3.4.

We adapt the notion of Section 6.3 and write

$$G(x,z) = \sum_{i=1}^K p_i(x,z) e^{-z_{i,n_i+1} x} \tag{4.5}$$

where $p_i(x,z) = \sum_{j=0}^{n_i} z_{i,j} x^j$, for some vector $n = (n_1, \ldots, n_K) \in \mathbb{N}_0^K$ and $K \in \mathbb{N}$. Since H_w contains only bounded functions, we restrict to the bounded exponential-polynomial families $BEP(K,n)$, see Definition 6.3. But $BEP(K,n)$ is not a submanifold in H_w. We have to extract the singularities. An easy computation yields

$$\partial_{z_{i,j}} G(x,z) = \begin{cases} x^j e^{-z_{i,n_i+1} x}, & 0 \leq j \leq n_i \\ -x p_i(x,z) e^{-z_{i,n_i+1} x}, & j = n_i + 1. \end{cases}$$

Hence $\{\partial_{z_{i,j}} G(\cdot, z) \mid 0 \leq j \leq n_i + 1,\ 1 \leq i \leq K\}$ are linearly independent functions if and only if

$$z_{i,n_i+1} \neq z_{j,n_j+1} \quad \text{and} \quad z_{i,n_i} \neq 0, \quad \forall 1 \leq i \neq j \leq K. \tag{4.6}$$

Under these constraints $G(\cdot, z) \in H_w$ if and only if $z_{i,n_i+1} > 0$ or both $z_{i,n_i+1} = 0$ and $n_i = 0$, compare with (6.8). But the parameter set $\mathcal{Z}$ has to be open. To include the latter case, we will therefore assume that

$$\begin{gathered} n_1 = 0 \\ z_{1,1} \equiv 0 \quad \text{and} \quad z_{i,n_i+1} > 0 \quad \text{for } 2 \leq i \leq K. \end{gathered} \tag{4.7}$$

Consequently $G(\infty, z) = z_{1,0}$. For regularity reasons, see Remark 4.9 below, we have to sharpen condition (4.6) and let the distance between different exponents be bounded away from zero by some fixed $\epsilon > 0$. This leads to the following parameter set

$$\mathcal{Z}_\epsilon := \mathbb{R} \times \left\{ (z_{2,0}, \ldots, z_{K,n_K+1}) \;\middle|\; \begin{array}{ll} z_{i,n_i} \neq 0, & \\ z_{i,n_i+1} > 0, & \forall 2 \leq i \neq j \leq K \\ |z_{i,n_i+1} - z_{j,n_j+1}| > \epsilon, & \end{array} \right\}$$

in $\mathbb{R}^m$ where $m = 1 + (n_2 + 2) + \cdots + (n_K + 2)$. To avoid self-intersections of G we further require that

$$n_i \neq n_j, \quad 2 \leq i \neq j \leq K. \tag{4.8}$$

Now the representation (4.5) of $G(\cdot, z)$ with $z \in \mathcal{Z}_\epsilon$ is unique: $G(\cdot, z) = G(\cdot, \tilde{z})$ with $z, \tilde{z} \in \mathcal{Z}_\epsilon$ implies $z = \tilde{z}$.

Definition 4.7. *For $K \in \mathbb{N}$, $n \in \mathbb{N}_0^K$ satisfying (4.7)–(4.8), $\epsilon > 0$ and $\mathcal{Z}_\epsilon$ as above we define*

$$REP(K, n, \epsilon) := G(\mathcal{Z}_\epsilon)$$

the regular exponential-polynomial family*.*

Clearly, we have $REP(K, n, \epsilon) \subset BEP(K, n)$.

Setting $\tilde{w}(x) = (1 + x)^{\alpha+2}$, we conclude by Lemma 4.6 that $G \in C^2(\mathcal{Z}_\epsilon, H_w)$. Up to now, $G : \mathcal{Z}_\epsilon \to H_w$ is an injective immersion and $REP(K, n, \epsilon)$ is an immersed submanifold in H_w.

Proposition 4.8. *The mapping $G : \mathcal{Z}_\epsilon \to H_w$ is an embedding. Its image $REP(K, n, \epsilon)$ is consequently an m-dimensional C^2 submanifold in H_w, with m given above.*

Proof. We have to show that $G : \mathcal{Z}_\epsilon \to REP(K, n, \epsilon)$ is a homeomorphism.

Let (z^k) be a sequence in $\mathcal{Z}_\epsilon$ and $z \in \mathcal{Z}_\epsilon$ such that $G(z^k) \to G(z)$ in H_w. Inequality (3.4) yields $G(z^k) \to G(z)$ in $L^\infty(\mathbb{R}_+)$. The proof is completed by showing that $z^k \to z$.

In view of (3.3) we conclude $z^k_{1,0} = G(\infty, z^k) \to G(\infty, z) = z_{1,0}$. Thus $(G(\infty, z^k) - z^k_{1,0}) \to (G(\infty, z) - z_{1,0})$ in H_w and we may assume in the sequel that $z^k_{1,0} \equiv z_{1,0} = 0$.

Write $\overline{\theta}_i := \limsup_{k\to\infty} z^k_{i,n_i+1} \leq \infty$, $i = 2, \ldots, K$. Then there exists a subsequence (z^{k_μ}) of (z^k) such that $\lim_{\mu\to\infty} z^{k_\mu}_{i,n_i+1} = \overline{\theta}_i$ for all i.

Suppose first $\overline{\theta}_i < \infty$ for all i. By definition of $\mathcal{Z}_\epsilon$ it follows that

$$|\overline{\theta}_i - \overline{\theta}_j| \geq \epsilon, \quad \forall 2 \leq i \neq j \leq K. \tag{4.9}$$

We write $N := m - K$ and define the $\mathbb{R}^N$-vectors

$$\begin{aligned}
v^\mu(x) &:= \Big(e^{-z^{k_\mu}_{2,n_2+1}x}, xe^{-z^{k_\mu}_{2,n_2+1}x}, \dots, x^{n_2}e^{-z^{k_\mu}_{2,n_2+1}x}, \dots, x^{n_K}e^{-z^{k_\mu}_{K,n_K+1}x}\Big)\\
v(x) &:= \Big(e^{-\overline{\theta}_2 x}, xe^{-\overline{\theta}_2 x}, \dots, x^{n_2}e^{-\overline{\theta}_2 x}, \dots, x^{n_K}e^{-\overline{\theta}_K x}\Big)\\
y^\mu &:= \Big(z^{k_\mu}_{2,0}, z^{k_\mu}_{2,1}, \dots, z^{k_\mu}_{2,n_2}, \dots, z^{k_\mu}_{K,n_K}\Big).
\end{aligned}$$

Then

$$\langle v^\mu(x), y^\mu\rangle_{\mathbb{R}^N} = G(x, z^{k_\mu}) \to G(x,z), \quad \forall x \in \mathbb{R}_+. \tag{4.10}$$

Fix a sequence $x_1 < \cdots < x_N$ in $\mathbb{R}_+$. The vectors $v^\mu(x_1), \dots, v^\mu(x_N)$ are linearly independent for any μ and so are $v(x_1), \dots, v(x_N)$, by (4.9). Hence the $N \times N$-matrices

$$M^\mu = M^\mu(x_1, \dots, x_N) := \begin{pmatrix} v^\mu(x_1)\\ \vdots \\ v^\mu(x_N)\end{pmatrix}, \quad M = M(x_1,\dots,x_N) := \begin{pmatrix} v(x_1)\\ \vdots \\ v(x_N)\end{pmatrix}$$

are regular and $M^\mu \to M$ for $\mu \to \infty$. Consequently, $(M^\mu)^{-1} \to M^{-1}$ and

$$y^\mu \to M^{-1}\begin{pmatrix} G(x_1,z)\\ \vdots \\ G(x_N,z)\end{pmatrix}.$$

Hence $y_{i,j} := \lim_{\mu\to\infty} z^{k_\mu}_{i,j}$ exists in $\mathbb{R}$ for $0 \le j \le n_i$, $2 \le i \le K$. We conclude

$$\lim_{\mu\to\infty} G(x, z^{k_\mu}) = \sum_{i=2}^{K}\Big(\sum_{j=1}^{n_i} y_{i,j}x^j\Big)e^{-\overline{\theta}_i x} = G(x,z), \quad \forall x \in \mathbb{R}_+. \tag{4.11}$$

Combining (4.8) with (4.11) we get $\overline{\theta}_i = z_{i,n_i+1}$ and $y_{i,j} = z_{i,j}$. That way one concludes that any subsequence of (z^k) contains a convergent subsequence with limit z. Hence $\lim_{k\to\infty} z^k = z$, and the proposition is proved in case $\overline{\theta}_i < \infty$ for all $i = 2, \dots, K$.

Now assume that $I := \{2 \le i \le K \mid \overline{\theta}_i = \infty\}$ is not empty. Let $i \in I$. Then there exists $x_i^* \in \mathbb{R}_+$ such that

$$|p_i(x, z^{k_\mu})|e^{-z^{k_\mu}_{i,n_i+1}x} \le \sum_{j=0}^{n_i} |z^{k_\mu}_{i,j}|x^j\Big(e^{-z^{k_\mu}_{i,n_i+1}}\Big)^x \to 0 \quad \text{for } \mu\to\infty, \quad \forall x \ge x_i^*.$$

We write $\tilde{N} := N - \sum_{i\in I}(n_i+1)$ and define the $\mathbb{R}^{\tilde{N}}$-vectors $\tilde{v}^\mu(x)$, $\tilde{v}(x)$ and $\tilde{y}^\mu$ by skipping all components of $v^\mu(x)$, $v(x)$, resp. y^μ which include $z^{k_\mu}_{i,j}$ or $\overline{\theta}_i$ for $i \in I$. Set $x^* := \max_{i\in I} x_i^*$. Instead of (4.10) we now have

$$\langle \tilde{v}^\mu(x), \tilde{y}^\mu\rangle_{\mathbb{R}^{\tilde{N}}} = G(x, z^{k_\mu}) - \sum_{i\in I} p_i(x, z^{k_\mu})e^{-z^{k_\mu}_{i,n_i+1}x} \to G(x,z) \quad \forall x \ge x^*.$$

Choose a sequence $x^* \le x_1 < \cdots < x_{\tilde{N}}$. Inequality (4.9) still holds for $i, j \notin I$. Accordingly, the vectors $\tilde{v}(x_1), \dots, \tilde{v}(x_{\tilde{N}})$ are linearly independent and so are $\tilde{v}^\mu(x_1), \dots, \tilde{v}^\mu(x_{\tilde{N}})$ for all μ. A passage to the limit similar to the above implies

$$\lim_{\mu\to\infty} G(x, z^{k_\mu}) = \sum_{i\notin I}\Big(\sum_{j=1}^{n_i} y_{i,j}x^j\Big)e^{-\overline{\theta}_i x} = G(x,z), \quad \forall x \ge x^*,$$

which is impossible since $[x^*, \infty)$ is a set of uniqueness for analytic functions. Hence I is empty, and the proof is complete. $\square$

REMARK 4.9. *Estimate (4.9) is essential for the preceding proof. Although Proposition 4.8 seems to be true also for $\epsilon = 0$, we do not see how to prove it for that case.*

We can now summarize the main results from Chapter 6 as follows.

THEOREM 4.10. *Let σ satisfy the assumptions*[5] *of Lemma 4.4, and suppose the local HJM model (σ, γ, I) is consistent with $REP(K, n, \epsilon)$. Then*

$$\sigma(t, \omega, \cdot) \equiv 0 \quad on \quad \overline{REP(K, n, \epsilon)}.$$

PROOF. The proof is divided into two steps.

Step 1: Write

$$\mathcal{B} := \left\{ \beta = (0, \beta_2, \ldots, \beta_K) \in \mathbb{R}^K \mid \beta_i > 0,\ |\beta_i - \beta_j| > \epsilon,\ \forall 2 \le i \ne j \le K \right\}.$$

For $\beta \in \mathcal{B}$ we define

$$E_\beta := REP(K, n, \epsilon) \cap \operatorname{span}\left\{ x^j e^{-\beta_i x} \mid 0 \le j \le n_i,\ 1 \le i \le K \right\}.$$

We claim that (σ, γ, I) is consistent with E_β. Indeed, let $(t_0, h_0) \in \mathbb{R}_+ \times E_\beta$. That is, $h_0 = G(z^*)$ with $z^* \in \mathcal{Z}_\epsilon$ and $z^*_{i,n_i+1} = \beta_i$ for $1 \le i \le K$. By Theorem 7.6 there exists a $\mathcal{Z}_\epsilon$-valued Itô process[6]

$$Z_t^{i,j} = z^*_{i,j} + \int_0^t b_s^{i,j}\, ds + \sum_{k \in \mathbb{N}} \int_0^t \rho_s^{i,j;k}\, d\tilde{\beta}_s^k$$

and a stopping time $\tau > 0$ such that $r^{(t_0,h_0)}_{t \wedge \tau} = G(Z_{t\wedge\tau})$ for all $t \in \mathbb{R}_+$. Accordingly, $r^{(t_0,h_0)}$ is a local strong solution to the time t_0-shifted HJMM equation (2.25). This in turn implies relation (6.3) for $G(x, Z_{\cdot\wedge\tau})$. By Theorem 6.4 we conclude that there exists a stopping time $0 < \tau' \le \tau$ such that

$$b^{i,n_i+1} = \sum_{k \in \mathbb{N}} \left(\rho^{i,n_i+1;k} \right)^2 = 0, \quad dt \otimes d\mathbb{Q}\text{-a.s. on } [0, \tau'].$$

Therefore $Z^{i,n_i+1}_{t\wedge\tau'} \equiv \beta_i$ and the claim follows.

Step 2: Write

$$\tilde{\mathcal{B}} := \{ \beta \in \mathcal{B} \mid 2\beta_i = \beta_j \text{ for some } i \ne j \}.$$

The set $\tilde{\mathcal{B}}$ is contained in a finite union of hyperplanes in $\mathbb{R}^K$. Therefore $\overline{(\mathcal{B} \setminus \tilde{\mathcal{B}})} \cap \mathcal{B} = \mathcal{B}$.

Since E_β is a linear submanifold, Theorem 4.5 yields that (σ, γ, I) is r-consistent with E_β. From Theorem 7.13 we derive the consistency condition in local coordinates (4.4) which in turn implies (6.3). Hence Theorem 6.15 applies and $\sigma(t, \omega, h) \equiv 0$ for $h \in E_\beta$, for all $\beta \in \mathcal{B} \setminus \tilde{\mathcal{B}}$. By continuity of $h \mapsto \sigma(t, \omega, h)$ the assertion follows.

$\square$

[5] We don't need **(A2)** here.

[6] In fact, $Z = (G^{-1} \circ \phi)(Y)$ where ϕ is as in the proof of Theorem 7.6 and Y is given by (7.27).

Theorem 4.10 contains a stricter statement than Theorems 6.14–6.15. It is because "consistency" in Chapter 6 is only related to one particular initial point h_0 in $REP(K, n, \epsilon)$, see Definition 6.1 and Proposition 6.2 and compare with Definitions 4.1–4.3.

Let's recall the two leading examples of Chapters 5–6.

4.3.1. The Nelson–Siegel Family.

$$G_{NS}(x, z) = z_{1,0} + (z_{2,0} + z_{2,1}x)e^{-z_{2,2}x}$$

In the preceding context we have $G_{NS}(\mathcal{Z}_\epsilon) = REP(2, (0, 1), \epsilon)$, where actually ϵ is redundant.

4.3.2. The Regular Svensson Family.

$$G_S(x, z) = z_{1,0} + (z_{2,0} + z_{2,1}x)e^{-z_{2,2}x} + z_{3,1}xe^{-z_{3,2}x}$$

Obviously, condition (4.8) is not satisfied. So the preceding discussion does not apply immediately. Yet we can "regularize" also the Svensson family. Set $\epsilon > 0$ and

$$\mathcal{Z}_{S,\epsilon} := \mathbb{R} \times \{(z_{2,0}, z_{2,1}, z_{2,2}, z_{3,1}, z_{3,2}) \mid |z_{2,0}| > \epsilon,\ z_{2,2}, z_{3,2} > 0,\ |z_{2,2} - z_{3,2}| > \epsilon\}.$$

Similar as in the proof of Proposition 4.8 one shows that $G_S : \mathcal{Z}_{S,\epsilon} \to H_w$ is an embedding. As before we conclude $\sigma(t, \omega, h) \equiv 0$ for $h \in \overline{G_S(\mathcal{Z}_{S,\epsilon})}$, for every consistent local HJM model (σ, γ, I).

In Section 6.9.2 we identify the coordinate process Z for the non-trivial HJM model[7] which is r-consistent with the full (bounded) Svensson family. It is amazing that the support of the non-deterministic component $Z_t^{2,0}$ contains 0 for arbitrary small t (it is Gaussian distributed). Thus $\mathbb{P}[Z_t \notin \mathcal{Z}_{S,\epsilon}] > 0$ for every $t > 0$. However, as far as I can see, there is no deeper connection between failure of (4.8) and the existence of a non-trivial consistent state space process for the Svensson family.

Remark 4.11. *Consistency of HJM models with exponential-polynomial families has been checked first by Björk and Christensen* [**7**] *for deterministic σ. Their result was extended by the two articles in Chapter 5 and 6. There, however, we merely checked r-consistency for $EP(K, n)$, resp. $BEP(K, n)$. Question ii) in Section 4.1, asserting the generality of that approach, was left open and it is now clarified by Theorem 4.10.*

4.4. Affine Term Structure

In this section we apply Theorem 7.13 to identify the particular class of *affine local HJM models*. That is, those local HJM models which are consistent with finite dimensional linear submanifolds.

It is required throughout that σ meets the assumptions of Lemma 4.4. Suppose that $\mathcal{M}$ is an m-dimensional linear submanifold of H_w which is locally invariant for the HJMM equation (2.25). Without loss of generality, $\mathcal{M}$ is connected. According to Theorem 7.10 there exists $V \subset \mathbb{R}^m$ open and connected and a global parametrization $\phi : V \to \mathcal{M}$ with $\phi(y) = g_0 + \sum_{i=1}^m y_i g_i$ where $g_0, \dots, g_m \in D(A)$ are linearly independent. Clearly $D\phi(y) \equiv (g_1, \dots, g_m)$.

[7] It's a generalized Vasicek model.

Applying Theorem 7.13 yields, see (4.4),

$$g_0' + \sum_{i=1}^m y_i g_i' = \sum_{i=1}^m b^i(t,\omega,y) g_i - \frac{1}{2}\left(\sum_{i,j=1}^m a^{i,j}(t,\omega,y) G_i G_j\right)' \tag{4.12}$$

for all $(t,\omega,y) \in \mathbb{R}_+ \times \Omega \times V$ (after removing a $\mathbb{P}$-nullset from Ω), where $G_i(x) := \int_0^x g_i(\eta)\,d\eta$. To shorten notation, we let $'$ stand for ∂_x.

Write $\gamma_i = g_i(0)$. Integrating (4.12) yields

$$g_0 - \gamma_0 + \sum_{i=1}^m y_i (g_i - \gamma_i) = \sum_{i=1}^m b^i(t,\omega,y) G_i - \frac{1}{2} \sum_{i,j=1}^m a^{i,j}(t,\omega,y) G_i G_j. \tag{4.13}$$

Now if G_i and $G_i G_j$, $1 \le i,j \le m$, are linearly independent functions, we can invert the linear equation (4.13) for $b^i(t,\omega,y)$ and $a^{i,j}(t,\omega,y)$. Since the left hand side of (4.13) is affine in y, we obtain that also they are affine functions in y:

$$\begin{aligned} b^i(t,\omega,y) &= b^i(t,\omega) + \sum_{j=1}^m \beta^{i,j}(t,\omega) y_j, \quad 1 \le i \le m \\ a^{i,j}(t,\omega,y) &= a^{i,j}(t,\omega) + \sum_{k=1}^m \alpha^{i,j,k}(t,\omega) y_k, \quad 1 \le i,j \le m \end{aligned}$$

for some predictable functions $b^i(t,\omega)$, $\beta^{i,j}(t,\omega)$, $a^{i,j}(t,\omega)$ and $\alpha^{i,j,k}(t,\omega)$ being right continuous in t.

Since affine functions are analytic and V is open, equation (4.13) splits into the following system of Riccati equations for the G_i's

$$\begin{aligned} G_0' &= \gamma_0 + \sum_{k=1}^m b^k(t,\omega) G_k - \frac{1}{2} \sum_{k,l=1}^m a^{k,l}(t,\omega) G_k G_l \\ G_i' &= \gamma_i + \sum_{k=1}^m \beta^{k,i}(t,\omega) G_k - \frac{1}{2} \sum_{k,l=1}^m \alpha^{k,l,i}(t,\omega) G_k G_l, \quad 1 \le i \le m \end{aligned} \tag{4.14}$$

with initial conditions $G_0(0) = \cdots = G_m(0) = 0$.

Hence we derived some of the results in [**21**] as a special case of our general point of view. Clearly, Theorem 7.13 imposes constraints on the choice of V in order that b and ρ are locally Lipschitz, see (7.44), and that (7.43) has a local solution in V. Moreover, there is a tradeoff between the functions g_i and the coefficients $a^{i,j}$, $\alpha^{i,j,k}$, b^i, $\beta^{i,j}$ and γ_i by the Riccati equations (4.14). We shall discuss this problem for the case $m = 1$. For a fuller treatment we refer to [**21**].

For simplicity we assume that there is only a one dimensional driving Brownian motion W and that all coefficients are autonomous, that is

$$\begin{aligned} b(t,\omega,y) &\equiv b^1(y) = b + \beta y \\ a(t,\omega,y) &\equiv a^{1,1}(y) = a + \alpha y. \end{aligned}$$

Accordingly

$$\rho(t,\omega,y) \equiv \rho^{1,1}(y) = \sqrt{a + \alpha y}.$$

We rewrite the Riccati equations (4.14) for G_0 and G_1:

$$\begin{aligned} G_0' &= \gamma_0 + bG_1 - \frac{1}{2} a G_1^2, \quad G_0(0) = 0 \\ G_1' &= \gamma_1 + \beta G_1 - \frac{1}{2} \alpha G_1^2, \quad G_1(0) = 0. \end{aligned} \tag{4.15}$$

Equation (4.15) admits a non-exploding solution $G_1 : \mathbb{R}_+ \to \mathbb{R}$ if and only if $\alpha\gamma_1 \geq 0$ or both $\alpha\gamma_1 < 0$ and $\beta \leq -\sqrt{2|\alpha\gamma_1|}$.

Suppose $\alpha > 0$. Theorem 7.13 tells us that $y \mapsto \sqrt{a + \alpha y}$ has to be locally Lipschitz in V, see (7.44). Consequently, $a + \alpha y$ has to be bounded away from zero. Thus $V \subset [\frac{\epsilon - a}{\alpha}, \infty)$ for some $\epsilon > 0$. The dynamics of the coordinate process, see (7.43), is given by the stochastic differential equation

$$dY_t = (b + \beta Y_t)\, dt + \sqrt{a + \alpha Y_t}\, dW_t, \quad Y_0 \in V. \tag{4.16}$$

It can be shown[8], see [**29**, p. 221], that (4.16) admits a (global) unique continuous strong solution on $[\frac{-a}{\alpha}, \infty)$ if $b \geq \frac{a}{\alpha}\beta$.

If $\alpha < 0$, then $V \subset (-\infty, \frac{\epsilon - a}{\alpha}]$ for some $\epsilon > 0$ and (4.16) remains valid.

4.4.1. The Cox–Ingersoll–Ross (CIR) Model. For $\alpha > 0$, $\gamma_1 = 1$ and $a = \gamma_0 = 0$ (and hence $b \geq 0$) we get the CIR [**17**] short rate model. There exists a closed form solution to (4.15), see [**17**, Formula (23)],

$$G_1(x) = \frac{2\left(e^{\left(\sqrt{\beta^2 + 2\alpha}\right)x} - 1\right)}{\left(\sqrt{\beta^2 + 2\alpha} - \beta\right)\left(e^{\left(\sqrt{\beta^2 + 2\alpha}\right)x} - 1\right) + 2\sqrt{\beta^2 + 2\alpha}}$$

and $G_0(x) = b \int_0^x G_1(\eta)\, d\eta$. It holds by inspection[9] that $g_0 = G_0' = bG_1 \in H_w$ and $g_1 = G_1' \in H_w^0$ for any polynomial weighting function w as in Example 3.4. Hence the CIR model is contained in our setting.

From (7.41) and the identity $y = g_0(0) + y g_1(0) = \mathcal{J}_0(\phi(y))$ we obtain

$$\sigma(t, \omega, h) \equiv \sigma^1(h) = \sqrt{\alpha \mathcal{J}_0(h)}\, g_1, \quad h \in \phi((0, \infty)).$$

Denote by $\tilde{\sigma} : H_w \to H_w^0$ a measurable continuation of σ. Let Y be the continuous solution of (4.16) with $Y_0 > \epsilon$ and write $\tau_\epsilon = \inf\{t \mid Y_t \leq \epsilon\}$. It is straightforward that $(r_{t \wedge \tau_\epsilon}) = (\phi(Y_{t \wedge \tau_\epsilon}))$ with $r_0 = \phi(Y_0)$ is the unique continuous local strong solution to the HJMM equation (2.25) equipped with $\tilde{\sigma}$, see Theorem 7.13

However, $\sigma(h)$ – and hence $\tilde{\sigma}(h)$ – cannot be extended Lipschitz continuously to $h = \phi(0)$ (naturally $\sigma(\phi(0)) = 0$). Thus, although $\phi(Y)$ is a strong solution to the HJMM equation (2.25) with $\tilde{\sigma}(\phi(0)) = 0$, we cannot assert its global uniqueness. Uniqueness, however, holds for Y in $\mathbb{R}_+$, by the preceding discussion.

[8] Substitute $X := Y + \frac{a}{\alpha}$

[9] More general it can be shown that the critical point $y_0 = \frac{\beta + \sqrt{\beta^2 + 2\alpha}}{\alpha} \in \mathbb{R}_+$ for (4.15) is globally exponentially stable, see [**2**, Chapter IV]. A fact which also applies for the multi-dimensional system (4.14). We will investigate this further in connection with general affine models.

4.4.2. The Vasicek Model. The case $\alpha = 0$ leads to a Gaussian coordinate process Y. There are no constraints on V, but a has to be non-negative.

For $\gamma_0 = 0$, $\gamma_1 = 1$, $b \geq 0$ and $\beta < 0$ this is the Vasicek [**48**] short rate model. The solution for (4.15) is given by

$$G_1(x) = \frac{1}{\beta}\left(e^{\beta x} - 1\right)$$

and $G_0(x) = b\int_0^x G_1(\eta)\,d\eta - \frac{1}{2}a\int_0^x G_1^2(\eta)\,d\eta$. Obviously, $g_0 = bG_1 - \frac{1}{2}aG_1^2 \in H_w$ and $g_1(x) = e^{\beta x} \in H_w^0$. Hence also this classical model is contained in our setting.

From (7.41) we deduce that $\sigma(t,\omega,h) \equiv \sigma^1(h) \equiv \sqrt{a}g_1$ for $h \in \phi(\mathbb{R})$.

Part 2

Publications

CHAPTER 5

A Note on the Nelson–Siegel Family

Abstract. We study a problem posed in Björk and Christensen [**7**]: does there exist any nontrivial interest rate model which is consistent with the Nelson–Siegel family? They show that within the HJM framework with deterministic volatility structure the answer is no.

In this paper we give a generalized version of this result including stochastic volatility structure. For that purpose we introduce the class of *consistent state space processes*, which have the property to provide an arbitrage-free interest rate model when representing the parameters of the Nelson–Siegel family. We characterize the consistent state space Itô processes in terms of their drift and diffusion coefficients. By solving an inverse problem we find their explicit form. It turns out that there exists no nontrivial interest rate model driven by a consistent state space Itô process.

5.1. Introduction

Björk and Christensen [**7**] introduce the following concept: let $\mathcal{M}$ be an interest rate model and $\mathcal{G}$ a parameterized family of forward curves. $\mathcal{M}$ and $\mathcal{G}$ are called consistent, if all forward rate curves which may be produced by $\mathcal{M}$ are contained within $\mathcal{G}$, provided that the initial forward rate curve lies in $\mathcal{G}$. Under the assumption of a deterministic volatility structure and working under a martingale measure, they show that within the Heath–Jarrow–Morton (henceforth HJM) framework there exists no nontrivial forward rate model, consistent with the Nelson–Siegel family $\{G(\,.\,,z)\}$. The curve shape of $G(\,.\,,z)$ is given by the well known expression

$$G(x,z) = z_1 + z_2 e^{-z_4 x} + z_3 x e^{-z_4 x}, \tag{5.1}$$

introduced by Nelson and Siegel [**39**].

For an optimal todays choice of the parameter $z \in \mathbb{R}^4$, expression (5.1) represents the current term structure of interest rates, i.e. $x \geq 0$ denotes time to maturity. This method of fitting the forward curve is widely used among central banks, see the BIS [**6**] documentation.

From an economic point of view it seems reasonable to restrict z to the state space $\mathcal{Z} := \{z = (z_1, \dots, z_4) \in \mathbb{R}^4 \mid z_4 > 0\}$.

The corresponding term structure of the bond prices is given by

$$\Pi(x,z) := \exp\left(-\int_0^x G(\eta, z)\, d\eta\right).$$

Then $\Pi \in C^\infty([0,\infty) \times \mathcal{Z})$.

In order to imply a stochastic evolution of the forward rates, we introduce in Section 5.2 some state space process $Z = (Z_t)_{0 \leq t < \infty}$ with values in $\mathcal{Z}$ and ask whether $G(\,.\,,Z)$ provides an arbitrage-free interest rate model. We call Z consistent, if the corresponding discounted bond prices are martingales, see Section 5.3.

Solving an inverse problem we characterize in Section 5.4 the class of consistent state space Itô processes. Since a diffusion is a special Itô process, the very important class of consistent state space diffusion processes is characterized as well. Still we are able to derive a more general result. It turns out that all consistent Itô processes have essentially deterministic dynamics. The corresponding interest rate models are in turn trivial.

Consistent state space Itô processes are, by definition, specified under a martingale measure. This seems to be a restriction at first and one may ask whether there exists any Itô process Z under some objective probability measure inducing a nontrivial arbitrage free interest rate model $G(\,.\,,Z)$. However if the underlying filtration is not too large we show in Section 5.5 that our (negative) result holds for Itô processes modeled under any probability measure, provided that there exists an equivalent martingale measure. Hence under the requirement of absence of arbitrage there exists no nontrivial interest rate model driven by Itô processes and consistent with the Nelson–Siegel family.

Using the same ideas, still larger classes of consistent processes like Itô processes with jumps could be characterized.

5.2. The Interest Rate Model

For the stochastic background and notations we refer to [**40**] and [**31**].

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t<\infty}, \mathbb{P})$ be a filtered complete probability space, satisfying the usual conditions, and let $W = (W_t^1, \dots, W_t^d)_{0\le t<\infty}$ denote a standard d-dimensional $(\mathcal{F}_t)$-Brownian motion, $d \ge 1$.

We assume as given a $\mathcal{Z}$-valued Itô process $Z = (Z^1, \dots, Z^4)$ of the form

$$Z_t^i = Z_0^i + \int_0^t b_s^i\, ds + \sum_{j=1}^d \int_0^t \sigma_s^{ij}\, dW_s^j, \quad i = 1, \dots, 4, \quad 0 \le t < \infty, \tag{5.2}$$

where Z_0 is $\mathcal{F}_0$-measurable, and b, σ are progressively measurable processes with values in $\mathbb{R}^4$, resp. $\mathbb{R}^{4\times d}$, such that

$$\int_0^t \left(|b_s| + |\sigma_s|^2\right) ds < \infty, \quad \mathbb{P}\text{-a.s.,} \quad \text{for all finite } t. \tag{5.3}$$

Z could be for instance the (weak) solution of a stochastic differential equation, but in general Z is not Markov.

Write $r(t,x) := G(x, Z_t)$ for the *instantaneous forward rate* prevailing at time t for date $t + x$.

It is shown in [**19**, Section 7] that traded assets have to follow semimartingales. Hence it is of importance for us to observe that the price at time t for a *zero coupon bond* with maturity T

$$P(t,T) := \Pi(T - t, Z_t), \quad 0 \le t \le T < \infty,$$

and the *short rates*

$$r(t,0) = \lim_{x\to 0} r(t,x) = G(0, Z_t) = -\frac{\partial}{\partial x}\Pi(0, Z_t), \quad 0 \le t < \infty,$$

form continuous semimartingales, by the smoothness of G and Π. Therefore the same holds for the process of the *savings account*

$$B(t) := \exp\left(\int_0^t r(s,0)\, ds\right), \quad 0 \le t < \infty.$$

5.3. Consistent State Space Processes

We are going to define consistency in our context, which slightly differs from that in [**7**]. We focus on the state space process Z, which follows an Itô process or may follow some more general process.

DEFINITION 5.1. *Z is called* consistent *with the Nelson–Siegel family, if*

$$\left(\frac{P(t,T)}{B(t)}\right)_{0\le t\le T}$$

is a $\mathbb{P}$-martingale, for all $T<\infty$.

The next proposition is folklore in case that Z follows a diffusion process, i.e. if $b_t(\omega)=b(t,Z_t(\omega))$ and $\sigma_t(\omega)=\sigma(t,Z_t(\omega))$ for Borel mappings b and σ from $[0,\infty)\times\mathcal{Z}$ into the corresponding spaces. That case usually leads to a PDE including the generator of Z. The standard procedure is then to find a solution u (the term structure of bond prices) to this PDE on $(0,\infty)\times\mathcal{Z}$ with initial condition $u(0,\,.\,)=1$. It is well known in the financial literature that Z is necessarily consistent with the corresponding forward rate curve family. In contrast we ask the other way round and are more general what concerns Z. Our aim is, *given* G, to derive *conditions on b and σ* being *necessary* for consistency of Z with $\{G(\,.\,,z)\}_{z\in\mathcal{Z}}$. But the coefficients b and σ are progressively measurable processes. Hence Z given by (5.2) is not Markov, i.e. there is no infinitesimal generator. By the nature of b and σ such conditions can therefore only be stated $dt\otimes d\mathbb{P}$-a.s. (note that equation (5.4) below is not a PDE). On the other hand the argument mentioned in the diffusion case works just in one direction: consistency of Z with a forward curve family $\mathcal{G}=\{v(\,.\,,z)\}_{z\in\mathcal{Z}}$ does not imply validity of the PDE condition for $u(x,z)=\exp(-\int_0^x v(\eta,z)\,d\eta)$ in general.

Actually one could re-parameterize $f(t,T):=r(t,T-t)$, for $0\le t\le T<\infty$, and work within the HJM framework. Equation (5.4) below corresponds to the well known HJM drift condition for $(f(t,T))_{0\le t\le T}$. But as soon as Z is not an Itô process anymore, this connection fails and one has to proceed like in the following proof (see Remark 5.3), which therefore is given in its full form.

Set $a:=\sigma\sigma^*$, where σ^* denotes the transpose of σ, i.e. $a_t^{ij}=\sum_{k=1}^d \sigma_t^{ik}\sigma_t^{jk}$, for $1\le i,j\le 4$ and $0\le t<\infty$. Then a is a progressively measurable process with values in the symmetric nonnegative definite 4×4-matrices.

PROPOSITION 5.2. *Z is consistent with the Nelson–Siegel family only if*

$$\begin{aligned}\frac{\partial}{\partial x}G(x,Z) &= \sum_{i=1}^4 b^i \frac{\partial}{\partial z_i}G(x,Z)\\ &\quad+\frac12\sum_{i,j=1}^4 a^{ij}\Bigg(\frac{\partial^2}{\partial z_i\partial z_j}G(x,Z)-\frac{\partial}{\partial z_i}G(x,Z)\int_0^x\frac{\partial}{\partial z_j}G(\eta,Z)\,d\eta\\ &\qquad\qquad -\frac{\partial}{\partial z_j}G(x,Z)\int_0^x\frac{\partial}{\partial z_i}G(\eta,Z)\,d\eta\Bigg),\end{aligned} \tag{5.4}$$

for all $x\ge 0$, $dt\otimes d\mathbb{P}$-a.s.

PROOF. For $f \in C^2(\mathcal{Z})$ we set

$$\mathcal{A}_t(\omega) f(z) := \sum_{i=1}^{4} b_t^i(\omega) \frac{\partial f}{\partial z_i}(z) + \frac{1}{2} \sum_{i,j=1}^{4} a_t^{ij}(\omega) \frac{\partial^2 f}{\partial z_i \partial z_j}(z), \quad 0 \le t < \infty, \quad z \in \mathcal{Z}.$$

Using Itô's formula we get for $T < \infty$

$$P(t,T) = P(0,T) + \int_0^t \left(\mathcal{A}_s \Pi(T-s, Z_s) - \frac{\partial}{\partial x} \Pi(T-s, Z_s) \right) ds$$
$$+ \int_0^t \sigma_s^* \nabla_z \Pi(T-s, Z_s)\, dW_s, \quad 0 \le t \le T, \quad \mathbb{P}\text{-a.s.}$$

where ∇_z denotes the gradient with respect to (z_1, z_2, z_3, z_4), and

$$\frac{1}{B(t)} = 1 + \int_0^t \frac{1}{B(s)} \frac{\partial}{\partial x} \Pi(0, Z_s)\, ds, \quad 0 \le t < \infty, \quad \mathbb{P}\text{-a.s.}$$

For $0 \le t \le T$ define

$$H(t,T) := \frac{1}{B(t)} \Big(\mathcal{A}_t \Pi(T-t, Z_t) - \frac{\partial}{\partial x} \Pi(T-t, Z_t) + \frac{\partial}{\partial x} \Pi(0, Z_t) \Pi(T-t, Z_t) \Big)$$

and the local martingale

$$M(t,T) := \int_0^t \frac{1}{B(s)} \sigma_s^* \nabla_z \Pi(T-s, Z_s)\, dW_s.$$

Integration by parts then yields

$$\frac{1}{B(t)} P(t,T) = P(0,T) + \int_0^t H(s,T)\, ds + M(t,T), \quad 0 \le t \le T, \quad \mathbb{P}\text{-a.s.}$$

Let's suppose now that Z is consistent. Then necessarily for $T < \infty$

$$\int_0^t H(s,T)\, ds = 0, \quad \forall t \in [0,T], \quad \mathbb{P}\text{-a.s.} \tag{5.5}$$

Since b and σ are progressive and Π is smooth, $H(\,.\,,T)$ is progressively measurable on $[0,T] \times \Omega$. We claim that (5.5) yields

$$H(\,.\,,T) = 0, \quad \text{on } [0,T] \times \Omega, \quad dt \otimes d\mathbb{P}\text{-a.s.} \tag{5.6}$$

Proof of (5.6). Define $N := \{(t,\omega) \in [0,T] \times \Omega \mid H(t,T)(\omega) > 0\}$. Then N is a $\mathcal{B} \otimes \mathcal{F}$- measurable set. Since $H(\,.\,,T)$ is positive on N we can use Tonelli's theorem

$$\int_N H(t,T)(\omega)\, dt \otimes d\mathbb{P} = \int_\Omega \left(\int_{N_\omega} H(t,T)(\omega)\, dt \right) d\mathbb{P}(\omega) = 0,$$

where $N_\omega := \{t \mid (t,\omega) \in N\} \in \mathcal{B}$ and we have used (5.5) and the inner regularity of the measure dt. We therefore conclude that N has $dt \otimes d\mathbb{P}$-measure zero. By using a similar argument for $-H(\,.\,,T)$ we have proved (5.6).

Note that (5.6) holds for all $T < \infty$, where the $dt \otimes d\mathbb{P}$-nullset depends on T. But since $H(t,T)$ is continuous in T, a standard argument yields

$$H(t, t+x)(\omega) = 0, \quad \forall x \ge 0, \quad \text{for } dt \otimes d\mathbb{P}\text{-a.e. } (t,\omega).$$

Multiplying this equation with $B(t)$ and using again the full form this reads

$$\mathcal{A}\, \Pi(x,Z) - \frac{\partial}{\partial x} \Pi(x,Z) + \frac{\partial}{\partial x} \Pi(0,Z) \Pi(x,Z) = 0, \quad \forall x \ge 0, \quad dt \otimes d\mathbb{P}\text{-a.s.} \tag{5.7}$$

Differentiation gives

$$\int_0^x \mathcal{A}\, G(\eta, Z)\, d\eta - \frac{1}{2}\sum_{i,j=1}^{4} a^{ij} \left(\int_0^x \frac{\partial}{\partial z_i} G(\eta, Z)\, d\eta\right)\left(\int_0^x \frac{\partial}{\partial z_j} G(\eta, Z)\, d\eta\right)$$
$$- G(x, Z) + G(0, Z) = 0, \quad \forall x \geq 0, \quad dt \otimes d\mathbb{P}\text{-a.s.}$$

where we have divided by $\Pi(x, Z)$, since $\Pi > 0$ on $[0, \infty) \times \mathcal{Z}$. Differentiating with respect to x finally yields (5.4). □

Remark 5.3. *Definition 5.1 can be extended in a natural way to a wider class of state space processes Z. We mention here just two possible directions:*

a) Z *a* time homogeneous Markov process with infinitesimal generator $\mathcal{L}$. *The corresponding version of Proposition 5.2 can be formulated in terms of equation (5.7), where $\mathcal{A}$ has to be replaced by $\mathcal{L}$. The difficulty here consists of checking whether Π goes well together with Z, i.e. $x \mapsto \Pi(x, \,.\,)$ has to be a nice mapping from $[0, \infty)$ into the domain of $\mathcal{L}$.*

b) Z *an* Itô process with jumps. *Again one could reformulate Proposition 5.2 on the basis of equation (5.7) by adding the corresponding stochastic integral with respect to the compensator of the jump measure, which we assume to be absolutely continuous with respect to dt. The jump measure could be implied for example by a homogeneous Poisson process.*

5.4. The Class of Consistent Itô Processes

Equation (5.4) characterizes b and a, resp. σ, just up to a $dt \otimes d\mathbb{P}$-nullset. But the stochastic integral in (5.2) is (up to indistinguishability) defined on the equivalence classes with respect to the $dt \otimes d\mathbb{P}$-nullsets. Hence this is enough to determine the process Z, given Z_0, uniquely (up to indistinguishability). On the other hand Z cannot be represented in the form (5.2) with integrands that differ from b and σ on a set with $dt \otimes d\mathbb{P}$-measure strictly greater than zero (the characteristics of Z are unique up to indistinguishability). Therefore it makes sense to pose the following *inverse problem* on the basis of equation (5.4): For which choices of coefficients b and σ do we get a consistent state space Itô process Z starting in Z_0?

The answer is rather remarkable:

Theorem 5.4. *Let Z be a consistent Itô process. Then Z is of the form*

$$\begin{aligned}
Z_t^1 &= Z_0^1 \\
Z_t^2 &= Z_0^2 e^{-Z_0^4 t} + Z_0^3 t e^{-Z_0^4 t} \\
Z_t^3 &= Z_0^3 e^{-Z_0^4 t} \\
Z_t^4 &= Z_0^4 + \left(\int_0^t b_s^4\, ds + \sum_{j=1}^{d} \int_0^t \sigma_s^{4j}\, dW_s^j\right) 1_{\Omega_0}
\end{aligned}$$

for all $0 \leq t < \infty$, where $\Omega_0 := \{Z_0^2 = Z_0^3 = 0\}$.

Remark 5.5. *On Ω_0 the processes Z^2 and Z^3 are zero. Hence $G(x, Z_t) = Z_0^1$ on Ω_0, i.e. the process Z^4 has no influence on $G(\,.\,, Z_t)$ on Ω_0. So it holds that the* *corresponding interest rate model is of the form*

$$\begin{aligned}
r(t,x) &= G(x, Z_t) \\
&= Z_0^1 + \left(Z_0^2 e^{-Z_0^4 t} + Z_0^3 t e^{-Z_0^4 t} \right) e^{-Z_0^4 x} + Z_0^3 e^{-Z_0^4 t} x e^{-Z_0^4 x} \\
&= Z_0^1 + Z_0^2 e^{-Z_0^4 (t+x)} + Z_0^3 (t+x) e^{-Z_0^4 (t+x)} \\
&= r(0, t+x), \quad \forall t, x \geq 0,
\end{aligned}$$

and is therefore quasi deterministic, i.e. all randomness remains $\mathcal{F}_0$-measurable.

REMARK 5.6. *One can show a similar (negative) result even for the wider class of state space Itô processes with jumps on a finite or countable mark space, see Remark 5.3. However allowing for more general exponential-polynomial families $\{G(\,.\,, z)\}$ there exist (although very restricted) consistent Itô processes providing a nontrivial interest rate model, see* [**23**].

PROOF. Let Z be a consistent Itô process given by equation (5.2). The proof of the theorem relies on expanding equation (5.4).

First of all we subtract $\frac{\partial}{\partial x} G(x, Z)$ from both sides of (5.4) to obtain a null equation. Fix then a point (t, ω) in $[0, \infty) \times \Omega$. For simplicity we write (z_1, z_2, z_3, z_4) for $Z_t(\omega)$, a_{ij} for $a_t^{ij}(\omega)$ and b_i for $b_t^i(\omega)$. Notice that $Z_t(\omega) \in \mathcal{Z}$, i.e. $z_4 > 0$. Observe then that our null equation is in fact of the form

$$p_1(x) + p_2(x) e^{-z_4 x} + p_3(x) e^{-2 z_4 x} = 0, \tag{5.8}$$

which has to hold simultaneously for all $x \geq 0$. The expressions p_1, p_2 and p_3 denote some polynomials in x, which depend on the z_i's, b_i's and a_{ij}'s. Since the functions $\{1, e^{-z_4 x}, e^{-2 z_4 x}\}$ are independent over the ring of polynomials, (5.8) can only be satisfied if each of the p_i's is 0. This again yields that all coefficients of the p_i's have to be zero. To proceed in our analysis we list all terms appearing in (5.4):

$$\begin{aligned}
\frac{\partial}{\partial x} G(x, z) &= (-z_2 z_4 + z_3 - z_3 z_4 x) e^{-z_4 x}, \\
\nabla_z G(x, z) &= \left(1, e^{-z_4 x}, x e^{-z_4 x}, (-z_2 x - z_3 x^2) e^{-z_4 x}\right), \\
\frac{\partial^2}{\partial z_i \partial z_j} G(x, z) &= 0, \quad \text{for } 1 \leq i, j \leq 3, \\
\frac{\partial}{\partial z_4} \nabla_z G(x, z) &= \left(0, -x e^{-z_4 x}, -x^2 e^{-z_4 x}, (z_2 x^2 + z_3 x^3) e^{-z_4 x}\right).
\end{aligned}$$

Finally we need the relation

$$\int_0^x \eta^m e^{-z_4 \eta} \, d\eta = -q_m(x) e^{-z_4 x} + \frac{m!}{z_4^{m+1}}, \quad m = 0, 1, 2, \ldots,$$

where $q_m(x) = \sum_{k=0}^m \frac{m!}{(m-k)!} \frac{x^{m-k}}{z_4^{k+1}}$ is a polynomial in x of order m.

First we shall analyze p_1. The terms that contribute to p_1 are those containing $\frac{\partial}{\partial z_1} G(x, z)$ and $\frac{\partial}{\partial z_1} G(x, z) \int_0^x \frac{\partial}{\partial z_j} G(\eta, z) \, d\eta$, for $1 \leq j \leq 4$. Actually p_1 is of the form

$$p_1(x) = a_{11} x + \cdots + b_1,$$

where $\ldots$ stands for terms of zero order in x containing the factors $a_{1j} = a_{j1}$, for $1 \leq j \leq 4$. It follows that $a_{11} = 0$. But the matrix (a_{ij}) has to be nonnegative definite, so necessarily

$$a_{1j} = a_{j1} = 0, \quad \text{for all } 1 \le j \le 4,$$

and therefore also $b_1 = 0$. Thus p_1 is done.

The contributing terms to p_3 are those containing $\frac{\partial}{\partial z_i} G(x,z) \int_0^x \frac{\partial}{\partial z_j} G(\eta, z)\, d\eta$, for $2 \le i, j \le 4$. But observe that the degree of p_3 and p_2 depends on whether z_2 or z_3 are equal to zero or not. Hence we have to distinguish between the four cases

i) $z_2 \neq 0,\ z_3 \neq 0$ **iii)** $z_2 = 0,\ z_3 \neq 0$
ii) $z_2 \neq 0,\ z_3 = 0$ **iv)** $z_2 = z_3 = 0$.

case i): The degree of p_3 is 4. The fourth order coefficient contains a_{44}, i.e.

$$p_3(x) = a_{44} \frac{z_3^2}{z_4} x^4 + \dots,$$

where $\dots$ stands for terms of lower order in x. Hence

$$a_{4j} = a_{j4} = 0, \quad \text{for } 1 \le j \le 4.$$

The degree of p_3 reduces to 2. The second order coefficient is $\frac{a_{33}}{z_4}$. Hence $a_{3j} = a_{j3} = 0$, for $1 \le j \le 4$. It remains $p_3(x) = \frac{a_{22}}{z_4}$. Thus the diffusion matrix (a_{ij}) is zero. This implies that (σ_{ij}) is zero, independent of the choice of d (the number of Brownian motions in (5.2)). Now we can write down p_2:

$$p_2(x) = -b_4 z_3 x^2 + (b_3 - b_4 z_2 + z_3 z_4) x + b_2 + z_2 z_4 - z_3.$$

It follows that $b_4 = 0$ and

$$b_2 = z_3 - z_2 z_4,$$
$$b_3 = -z_3 z_4.$$

For the other three cases we will need the following lemma, which is a direct consequence of the occupation times formula, see [**40**, Corollary (1.6), Chapter VI].

Lemma 5.7. *Using the same notation as in (5.2), it holds for* $1 \le i \le 4$ *that*

$$a^{ii} 1_{\{Z^i = 0\}} = b^i 1_{\{Z^i = 0\}} = 0, \quad dt \otimes d\mathbb{P}\text{-a.s.}$$

As a consequence, since we are characterizing a and b up to a $dt \otimes d\mathbb{P}$-nullset, we may and will assume that $z_i = 0$ implies $a_{ij} = a_{ji} = b_i = 0$, for $1 \le j \le 4$ and $i = 2, 3$.

case ii): We have $\nabla_z G(x,z) = (1, e^{-z_4 x}, x e^{-z_4 x}, -z_2 x e^{-z_4 x})$. Hence the degree of p_3 is 2. Since $a_{3j} = a_{j3} = b_3 = 0$, for $1 \le j \le 4$, the second order coefficient comes from

$$-a_{44} \frac{\partial}{\partial z_4} G(x,z) \int_0^x \frac{\partial}{\partial z_4} G(\eta, z)\, d\eta = a_{44} \frac{z_2^2}{z_4} x^2 e^{-2 z_4 x} + \dots,$$

where $\dots$ denotes terms of lower order in x. Hence $a_{4j} = a_{j4} = 0$, for $1 \le j \le 4$. The polynomial p_3 reduces to $p_3(x) = \frac{a_{22}}{z_4}$. It follows that also in this case the diffusion matrix (a_{ij}) is zero. From case i) we derive immediately that now

$$p_2(x) = -b_4 z_2 x + b_2 + z_2 z_4,$$

hence $b_2 = -z_2 z_4$ and $b_4 = 0$.

case iii): Since $a_{2j} = a_{j2} = b_2 = 0$, for $1 \leq j \leq 4$, the zero order coefficient of p_2 reduces to $-z_3$. We conclude that $z_2 = 0$ implies $z_3 = 0$, so this case doesn't enter $dt \otimes d\mathbb{P}$-a.s.

case iv): In this case $a_{ij} = b_k = 0$, for all $(i,j) \neq (4,4)$ and $k \neq 4$. Also $\frac{\partial}{\partial z_4} G(x,z) = \frac{\partial}{\partial x} G(x,z) = 0$. Hence $p_2(x) = p_3(x) = 0$, independently of the choice of b_4 and a_{44}.

Summarizing the four cases we conclude that equation (5.4) implies

$$\begin{aligned} b_1 &= 0 & b_3 &= -z_3 z_4 \\ b_2 &= z_3 - z_2 z_4 & a_{ij} &= 0, \quad \text{for } (i,j) \neq (4,4). \end{aligned}$$

Whereas b_4 and a_{44} are arbitrary real, resp. nonnegative real, numbers whenever $z_2 = z_3 = 0$. Otherwise $b_4 = a_{44} = 0$.

This has to hold for $dt \otimes d\mathbb{P}$-a.e. (t,ω). So Z is uniquely (up to indistinguishability) determined and satisfies

$$\begin{aligned} Z_t^1 &= Z_0^1 \\ Z_t^2 &= Z_0^2 + \int_0^t \left(Z_s^3 - Z_s^2 Z_s^4 \right) ds \\ Z_t^3 &= Z_0^3 - \int_0^t Z_s^3 Z_s^4 \, ds \\ Z_t^4 &= Z_0^4 + \int_0^t b_s^4 \, ds + \sum_{j=1}^d \int_0^t \sigma_s^{4j} \, dW_s^j \end{aligned}$$

for some progressively measurable processes b^4 and σ^{4j}, $j = 1, \dots, d$, being compatible with (5.3) and vanishing outside the set $\{(t,\omega) \mid Z_t^2(\omega) = Z_t^3(\omega) = 0\}$.

Note that Z^2 and Z^3 satisfy (path-wise) a system of linear ODE's with continuous coefficients. Hence they are indistinguishable from zero on Ω_0. So the statement of the theorem is proved on Ω_0. It remains to prove it on $\Omega_1 := \Omega \setminus \Omega_0$.

Introduce the stopping time $\tau := \inf\{s > 0 \mid Z_s^2 = Z_s^3 = 0\}$. We have just argued that $\Omega_0 \subset \{\tau = 0\}$. By continuity of Z also the reverse inclusion holds, hence $\Omega_0 = \{\tau = 0\}$. The stopped process $Z^\tau =: Y$ satisfies (path-wise) the following system of linear stochastic integral equations

$$\begin{aligned} Y_t^1 &= Z_0^1 \\ Y_t^2 &= Z_0^2 + \int_0^{t\wedge\tau} \left(Y_s^3 - Y_s^2 Z_0^4 \right) ds \\ Y_t^3 &= Z_0^3 + \int_0^{t\wedge\tau} Y_s^3 Z_0^4 \, ds \\ Y_t^4 &= Z_0^4, \quad \text{for } 0 \leq t < \infty. \end{aligned} \tag{5.9}$$

We have used the fact that $b^4 = b^4 1_{[\tau,\infty]}$ and $\sigma^4 = \sigma^4 1_{[\tau,\infty]}$. Then the last equation follows from an elementary property of the stopped stochastic integral:

$$\sum_{j=1}^d \int_0^{t\wedge\tau} \sigma_s^{4j} \, dW_s^j = \sum_{j=1}^d \int_0^t \left(\sigma_s^{4j} 1_{[\tau,\infty]} \right) 1_{[0,\tau]} \, dW_s^j = \sum_{j=1}^d \int_0^t \sigma_s^{4j} 1_{[\tau]} \, dW_s^j = 0,$$

by continuity of W.

The system (5.9) has the unique solution for $0 \le t < \infty$

$$\begin{aligned} Y_t^1 &= Z_0^1 \\ Y_t^2 &= Z_0^2 e^{-Z_0^4(t\wedge\tau)} + Z_0^3(t\wedge\tau)e^{-Z_0^4(t\wedge\tau)} \\ Y_t^3 &= Z_0^3 e^{-Z_0^4(t\wedge\tau)} \\ Y_t^4 &= Z_0^4. \end{aligned}$$

Since $Z = Y$ on the stochastic interval $[0,\tau]$ and since $Y_t \neq 0$, $\forall t < \infty$, $\mathbb{P}$-a.s. on Ω_1, it follows by the continuity of Z, that $\Omega_1 = \{\tau > 0\} = \{\tau = \infty\}$. Inserting this in the above solution, the theorem is proved also on Ω_1. □

5.5. E-Consistent Itô Processes

Note that by definition Z is consistent if and only if $\mathbb{P}$ is a martingale measure for the discounted bond price processes. We could generalize this definition and call a state space process Z *e-consistent* if there exists an equivalent martingale measure $\mathbb{Q}$. Then obviously consistency implies e-consistency, and e-consistency implies the absence of arbitrage opportunities, as it is well known.

In case where the filtration is generated by the Brownian motion W, i.e. $(\mathcal{F}_t) = (\mathcal{F}_t^W)$, we can give the following stronger result:

Proposition 5.8. *If $(\mathcal{F}_t) = (\mathcal{F}_t^W)$, then any e-consistent Itô process Z is of the form as stated in Theorem 5.4. In particular the corresponding interest rate model is purely deterministic.*

Proof. Let Z be an e-consistent Itô process under $\mathbb{P}$, and let $\mathbb{Q}$ be an equivalent martingale measure. Since $(\mathcal{F}_t) = (\mathcal{F}_t^W)$, we know that all $\mathbb{P}$-martingales have the representation property relative to W. By Girsanov's theorem it follows therefore that Z remains an Itô process under $\mathbb{Q}$. In particular Z is a consistent state space Itô process with respect to the stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t^W)_{0\le t<\infty}, \mathbb{Q})$. Hence Z is of the form as stated in Theorem 5.4.

Since $\mathcal{F}_0^W$ consists of sets of measure 0 or 1, the processes Z^1, Z^2, Z^3 are (after changing the Z_0^i's on a set of measure 0) purely deterministic and therefore also $(r(t, .))_{0\le t<\infty}$, see Remark 5.5. □

CHAPTER 6

Exponential-Polynomial Families and the Term Structure of Interest Rates

ABSTRACT. Exponential-polynomial families like the Nelson–Siegel or Svensson family are widely used to estimate the current forward rate curve. We investigate whether these methods go well with inter-temporal modelling. We characterize the consistent Itô processes which have the property to provide an arbitrage free interest rate model when representing the parameters of some bounded exponential-polynomial type function. This includes in particular diffusion processes. We show that there is a strong limitation on their choice. Bounded exponential-polynomial families should rather not be used for modelling the term structure of interest rates.

6.1. Introduction

The current term structure of interest rates contains all the necessary information for pricing bonds, swaps and forward rate agreements of all maturities. It is used furthermore by the central banks as indicator for their monetary policy.

There are several algorithms for constructing the current forward rate curve from the (finitely many) prices of bonds and swaps observed in the market. Widely used are splines and parameterized families of smooth curves $\{G(\,.\,,z)\}_{z\in\mathcal{Z}}$, where $\mathcal{Z}\subset\mathbb{R}^N$, $N\geq 1$, denotes some finite dimensional parameter set. By an optimal choice of the parameter z in $\mathcal{Z}$ an optimal fit of the forward curve $x\mapsto G(x,z)$ to the observed data is attained. Here $x\geq 0$ denotes *time to maturity*. In that sense z represents the current state of the economy taking values in the state space $\mathcal{Z}$.

Examples are the Nelson–Siegel [**39**] family with curve shape

$$G_{NS}(x,z)=z_1+(z_2+z_3x)e^{-z_4x}$$

and the Svensson [**44**] family, an extension of Nelson–Siegel,

$$G_S(x,z)=z_1+(z_2+z_3x)e^{-z_5x}+z_4xe^{-z_6x}.$$

Table 1 gives an overview of the fitting procedures used by some selected central banks. It is taken from the documentation of the Bank for International Settlements [**6**].

Despite the flexibility and low number of parameters of G_{NS} and G_S, their choice is somewhat arbitrary. We shall discuss them from an inter-temporal point of view: A lot of cross-sectional data, i.e. daily estimations of z, is available. Therefore it would be natural to ask for the stochastic evolution of the parameter z over time. But then there exist economic constraints based on no arbitrage considerations.

Table 1. Forward rate curve fitting procedures

central bank	**curve fitting procedure**
Belgium	Nelson–Siegel, Svensson
Canada	Svensson
Finland	Nelson–Siegel
France	Nelson–Siegel, Svensson
Germany	Svensson
Italy	Nelson–Siegel
Japan	smoothing splines
Norway	Svensson
Spain	Nelson–Siegel (before 1995), Svensson
Sweden	Svensson
UK	Svensson
USA	smoothing splines

Following [**7**], instead of G_{NS} and G_S we consider general exponential-polynomial families containing curves of the form

$$G(x,z) = \sum_{i=1}^{K} \Big(\sum_{\mu=0}^{n_i} z_{i,\mu} x^{\mu} \Big) e^{-z_{i,n_i+1}x}.$$

Hence linear combinations of exponential functions $\exp(-z_{i,n_i+1}x)$ over some polynomials of degree $n_i \in \mathbb{N}_0$. Obviously G_{NS} and G_S are of this type. We replace then z by an Itô process $Z = (Z_t)_{t \geq 0}$ taking values in $\mathcal{Z}$. The following questions arise:

- Does $G(\,.\,,Z)$ provide an arbitrage free interest rate model?
- And what are the conditions on Z for it?

Working in the Heath–Jarrow–Morton [**28**] – henceforth HJM – framework with deterministic volatility structure, Björk and Christensen [**7**] showed that the exponential-polynomial families are in a certain sense to large to carry an interest rate model. This result has been generalized for the Nelson–Siegel family in [**25**], including stochastic volatility structure. Expanding the methods used in there, we give in this paper the general result for bounded exponential-polynomial families.

The paper is organized as follows. In Section 6.2 we introduce the class of Itô processes consistent with a given parameterized family of forward rate curves. Consistent Itô processes provide an arbitrage free interest rate model when driving the parameterized family. They are characterized in terms of their drift and diffusion coefficients by the HJM drift condition.

By solving an inverse problem we get the main result for consistent Itô processes, stated in Section 6.3. It is shown that they are remarkably limited. The proof is divided into several steps, given in Sections 6.4, 6.5 and 6.6.

In Section 6.7 we extend the notion of consistency to e-consistency when $\mathbb{P}$ is not a martingale measure.

The main result reads much clearer when restricting to diffusion processes, as shown in Section 6.8. It turns out that e-consistent diffusion processes driving bounded exponential-polynomial families like Nelson–Siegel or Svensson are very limited: Most of the factors are either constant or deterministic. It is shown in

Section 6.9, that there is no non-trivial diffusion process which is e-consistent with the Nelson–Siegel family. Furthermore we identify the diffusion process which is e-consistent with the Svensson family. It contains just one non deterministic component. The corresponding short rate model is shown to be the generalized Vasicek model.

We conclude that bounded exponential-polynomial families, in particular G_{NS} and G_S, should rather not be used for modelling the term structure of interest rates.

6.2. Consistent Itô Processes

For the stochastic background and notations we refer the reader to [**40**] and [**31**]. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t<\infty}, \mathbb{P})$ be a filtered complete probability space, satisfying the usual conditions, and let $W = (W_t^1, \dots, W_t^d)_{0\le t<\infty}$ denote a standard d-dimensional $(\mathcal{F}_t)$-Brownian motion, $d \ge 1$.

Let $Z = (Z^1, \dots, Z^N)$ denote an $\mathbb{R}^N$–valued Itô process, $N \ge 1$, of the form

$$Z_t^i = Z_0^i + \int_0^t b_s^i \, ds + \sum_{j=1}^d \int_0^t \sigma_s^{i,j} \, dW_s^j, \quad i = 1, \dots, N, \quad 0 \le t < \infty,$$

where Z_0 is $\mathcal{F}_0$-measurable, and b, σ are progressively measurable processes with values in $\mathbb{R}^N$, resp. $\mathbb{R}^{N\times d}$, such that

$$\int_0^t \left(|b_s| + |\sigma_s|^2\right) ds < \infty, \quad \mathbb{P}\text{-a.s.}, \quad \text{for all finite } t.$$

Let $G(x,z)$ be a function in $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^N)$, i.e. G and the partial derivatives $(\partial G/\partial x)$, $(\partial G/\partial z_i)$, $(\partial^2 G/\partial z_i \partial z_j)$, which exist for $1 \le i, j \le N$, are continuous functions on $\mathbb{R}_+ \times \mathbb{R}^N$. Interpreting Z_t as the state of the economy at time t, we let $x \mapsto G(x, Z_t)$ stand for the corresponding term structure of interest rates. Meaning that $G(x, Z_t)$ denotes the *instantaneous forward rate* at time t for date $t + x$.

Notice that

$$\Pi(x,z) := \exp\left(-\int_0^x G(\eta, z)\, d\eta\right)$$

is in $C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^N)$ too. Therefore the price processes for *zero coupon T-bonds*

$$P(t,T) := \Pi(T - t, Z_t), \quad 0 \le t \le T < \infty, \tag{6.1}$$

and the process of the *savings account*

$$B(t) := \exp\left(-\int_0^t \frac{\partial}{\partial x}\Pi(0, Z_s)\, ds\right), \quad 0 \le t < \infty,$$

form continuous semimartingales.

Let $\mathcal{Z}$ denote an arbitrary subset of $\mathbb{R}^N$. The function G generates in a canonical way a parameterized set of forward curves $\{G(\,.\,,z)\}_{z\in\mathcal{Z}}$. We shall refer to $\mathcal{Z}$ as the state space of the economy.

Definition 6.1. *Z is called* consistent *with $\{G(\,.\,,z)\}_{z\in\mathcal{Z}}$, if the support of Z is contained in $\mathcal{Z}$ and*

$$\left(\frac{P(t,T)}{B(t)}\right)_{0\le t\le T} \tag{6.2}$$

is a $\mathbb{P}$-martingale, for all $T < \infty$.

Set $a := \sigma\sigma^*$, where σ^* denotes the transpose of σ, i.e. $a_t^{i,j} = \sum_{k=1}^d \sigma_t^{i,k}\sigma_t^{j,k}$, for $1 \le i, j \le N$ and $0 \le t < \infty$. Then a is a progressively measurable process with values in the symmetric nonnegative definite $N \times N$-matrices.

Using Itô's formula, the dynamics of (6.2) can be decomposed into finite variation and local martingale part. Requiring consistency the former has to vanish. This is the well known HJM drift condition and is stated explicitly in the following proposition.

Proposition 6.2. *If Z is consistent with $\{G(\,.\,, z)\}_{z\in\mathcal{Z}}$ then*

$$
\begin{aligned}
\frac{\partial}{\partial x}G(x,Z) = &\sum_{i=1}^N b^i \frac{\partial}{\partial z_i}G(x,Z) \\
&+ \frac{1}{2}\sum_{i,j=1}^N a^{i,j}\left(\frac{\partial^2}{\partial z_i \partial z_j}G(x,Z) - \frac{\partial}{\partial z_i}G(x,Z)\int_0^x \frac{\partial}{\partial z_j}G(\eta,Z)\,d\eta \right.\\
&\qquad\left. -\frac{\partial}{\partial z_j}G(x,Z)\int_0^x \frac{\partial}{\partial z_i}G(\eta,Z)\,d\eta\right),
\end{aligned}
\tag{6.3}
$$

for all $x \ge 0$, $dt \otimes d\mathbb{P}$-a.s.

Proof. Analogous to the proof of Proposition 5.2. □

6.3. Exponential-Polynomial Families

In this section we introduce a particular class of functions G. As the main result we characterize the corresponding consistent Itô processes.

Let K denote a positive integer and let $n = (n_1, \dots, n_K)$ be a vector with components $n_i \in \mathbb{N}_0$, for $1 \le i \le K$. Write $|n| := n_1 + \cdots + n_K$. For a point

$$z = (z_{1,0}, \dots, z_{1,n_1+1}, z_{2,0}, \dots, z_{2,n_2+1}, \dots, z_{K,0}, \dots, z_{K,n_K+1}) \in \mathbb{R}^{|n|+2K} \tag{6.4}$$

define the polynomials $p_i(z)$ as

$$p_i(z) = p_i(x,z) := \sum_{\mu=0}^{n_i} z_{i,\mu}x^\mu, \quad 1 \le i \le K.$$

The function G is now defined as

$$G(x,z) := \sum_{i=1}^K p_i(x,z)e^{-z_{i,n_i+1}x}. \tag{6.5}$$

Obviously $G \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R}^{|n|+2K})$. Hence the preceding section applies with $N = |n| + 2K$.

From an economic point of view it seems reasonable to restrict to bounded forward rate curves. Let therefore $\mathcal{Z}$ denote the set of all $z \in \mathbb{R}^N$ such that $\sup_{x\in\mathbb{R}_+}|G(x,z)| < \infty$.

Definition 6.3. *The* exponential-polynomial family $EP(K,n)$ *is defined as the set of forward curves $\{G(\,.\,,z)\}_{z\in\mathbb{R}^N}$.*

The bounded exponential-polynomial family $BEP(K,n) \subset EP(K,n)$ *is defined as the set of forward curves $\{G(\,.\,,z)\}_{z\in\mathcal{Z}}$.*

Clearly $G_{NS}(x,z) \in BEP(2,(0,1))$ and $G_S(x,z) \in BEP(3,(0,1,1))$, if at each case the parameter z is chosen such that the curve is bounded. From now on, the Nelson–Siegel and Svensson families are considered as subsets of $BEP(2,(0,1))$ and $BEP(3,(0,1,1))$ respectively.

If two exponents z_{i,n_i+1} and z_{j,n_j+1} coincide, the sum (6.5) defining G reduces to a linear combination of $K-1$ exponential functions. Thus for $z \in \mathbb{R}^N$ we introduce the equivalence relation

$$i \sim_z j :\Longleftrightarrow z_{i,n_i+1} = z_{j,n_j+1} \tag{6.6}$$

on the set $\{1,\dots,K\}$ and denote by $[i] = [i]_z$ the equivalence class of i. We will use the notation

$$\begin{aligned} n_{[i]} = n_{[i]}(z) &:= \max\{n_j \mid j \in [i]_z\} \\ \mathcal{I}_{[i],\mu} = \mathcal{I}_{[i],\mu}(z) &:= \{j \in [i]_z \mid n_j \geq \mu\}, \quad 0 \leq \mu \leq n_{[i]}(z) \\ z_{[i],\mu} = z_{[i],\mu}(z) &:= \sum_{j \in \mathcal{I}_{[i],\mu}(z)} z_{j,\mu}, \quad 0 \leq \mu \leq n_{[i]}(z) \\ p_{[i]}(z) &:= \sum_{j \in [i]_z} p_j(z). \end{aligned} \tag{6.7}$$

In particular $p_{[i]}(z) = \sum_{\mu=0}^{n_{[i]}} z_{[i],\mu} x^\mu$ and (6.5) reads now

$$G(x,z) = \sum_{[i] \in \{1,\dots,K\}/\sim} p_{[i]}(z) e^{-z_{i,n_i+1}x}.$$

Observe that for $z \in \mathcal{Z}$ we have

$$\begin{aligned} z_{i,n_i+1} = 0, &\quad \text{only if} \quad p_{[i]}(z) = z_{[i],0} \\ z_{i,n_i+1} < 0, &\quad \text{only if} \quad p_{[i]}(z) = 0. \end{aligned} \tag{6.8}$$

We shall write the $\mathbb{R}^N$-valued Itô process Z with the same indices which we use for a point $z \in \mathbb{R}^N$, see (6.4),

$$Z_t^{i,\mu} = Z_0^{i,\mu} + \int_0^t b_s^{i,\mu}\, ds + \sum_{\lambda=1}^d \int_0^t \sigma_s^{i,\mu;\lambda}\, dW_s^\lambda, \quad 0 \leq \mu \leq n_i + 1, \quad 1 \leq i \leq K. \tag{6.9}$$

It's diffusion matrix a consists of the components

$$a^{i,\mu;j,\nu} = \sum_{\lambda=1}^d \sigma^{i,\mu;\lambda} \sigma^{j,\nu;\lambda}, \quad 0 \leq \mu \leq n_i + 1, \quad 0 \leq \nu \leq n_j + 1, \quad 1 \leq i, j \leq K.$$

Notice that for $1 \leq i \leq K$

$$\begin{aligned} \{z \mid p_{[i]}(z) = 0\} = \bigcup_{\substack{J \subset \{1,\dots,K\} \\ J \ni i}} \Big(&\{z \mid z_{j,n_j+1} = z_{i,n_i+1} \text{ for all } j \in J\} \\ &\cap \bigcap_{\mu=0}^{\max\{n_j \mid j \in J\}} \{z \mid \sum_{\substack{j \in J \\ n_j \geq \mu}} z_{j,\mu} = 0\} \\ &\setminus \bigcup_{l \in J^c} \{z \mid z_{l,n_l+1} = z_{i,n_i+1}\} \Big) \end{aligned} \tag{6.10}$$

is not closed in general but nevertheless a Borel set in $\mathbb{R}^N$. We introduce the following, thus optional, random sets of singular points (t, ω)

$$\mathcal{A}_i := \left\{p_i(Z) = 0 \text{ or } p_{[i]}(Z) = 0\right\}, \quad 1 \leq i \leq K$$
$$\mathcal{B} := \bigcup_{\substack{i,j=1\\ i\neq j}}^{K} \left\{Z^{i,n_i+1} = Z^{j,n_j+1}\right\}$$
$$\mathcal{C} := \bigcup_{\substack{i,j=1\\ i\neq j}}^{K} \left\{2Z^{i,n_i+1} = Z^{j,n_j+1}\right\}$$

and the optional random sets of regular points (t, ω)

$$\mathcal{D} := (\mathbb{R}_+ \times \Omega) \setminus (\bigcup_{i=1}^{K} \mathcal{A}_i \cup \mathcal{B} \cup \mathcal{C})$$
$$\mathcal{D}' := (\mathbb{R}_+ \times \Omega) \setminus (\mathcal{B} \cup \mathcal{C}).$$

Let us recall that for S and T stopping times, a stochastic interval like $[S, T]$ is a subset of $\mathbb{R}_+ \times \Omega$. Hence $[S] = [S, S]$ is the restriction of the graph of the mapping $S : \Omega \to [0, \infty]$ to the set $\mathbb{R}_+ \times \Omega$.

For any stopping time τ with $[\tau] \in (\mathbb{R}_+ \times \Omega) \setminus \mathcal{A}_i$ we define

$$\tau'(\omega) := \inf\{t > \tau(\omega) \mid (t, \omega) \in \mathcal{A}_i\},$$

the debut of the optional set $[\tau, \infty[\cap \mathcal{A}_i$. Observe that in general it is not true that $\tau' > \tau$ on $\{\tau < \infty\}$. This can be seen from the following example: For

$$G(x, z) = z_{1,0}e^{-z_{1,1}x} + z_{2,0}e^{-z_{2,1}x} + z_{3,0}e^{-z_{3,1}x} \in BEP(3, (0, 0, 0))$$

let $Z_t^{1,0} = Z_t^{3,0} = 1$, $Z_t^{2,0} = -1$, $Z_t^{3,1} = 1+t$ and $Z_t^{1,1} = Z_t^{2,1} = 1$ for $t \in [0, 1]$. Then $p_1(Z_0) = p_{[1]}(Z_0) = 1$ and $p_{[1]}(Z_t) = 0$ for all $t \in (0, 1]$. Hence $[0] \in (\mathbb{R}_+ \times \Omega) \setminus \mathcal{A}_1$, but $\tau' = 0$. However, by continuity of Z we always have

$$\tau < \tau' \quad \mathbb{P}\text{-a.s. on } \{\omega \mid (\tau(\omega), \omega) \in \mathcal{D}'\}. \tag{6.11}$$

Recall the fact that there is a one to one correspondence between the Itô processes Z starting in Z_0 (up to indistinguishability) and the equivalence classes of b and σ with respect to the $dt \otimes d\mathbb{P}$-nullsets in $\mathcal{R}_+ \otimes \mathcal{F}$. Hence we may state the following *inverse problem* to equation (6.3): Given a family of forward curves. For which choices of coefficients b and σ do we get a consistent Itô process Z starting in Z_0?

The main result is the following characterization of all consistent Itô processes, which is remarkably restrictive. The proof of the theorem will be given in Sections 6.5 and 6.6.

THEOREM 6.4. *Let $K \in \mathbb{N}$, $n = (n_1, \dots, n_K) \in \mathbb{N}_0^K$ and Z as above. If Z is consistent with $BEP(K, n)$, then necessarily for $1 \leq i \leq K$*

$$a^{i,n_i+1;i,n_i+1} = 0, \quad \text{on } \{p_i(Z) \neq 0\}, \quad dt \otimes d\mathbb{P}\text{-a.s.} \tag{6.12}$$
$$b^{i,n_i+1} = 0, \quad \text{on } \{p_i(Z) \neq 0\} \cap \{p_{[i]}(Z) \neq 0\}, \quad dt \otimes d\mathbb{P}\text{-a.s.} \tag{6.13}$$

Consequently, Z^{i,n_i+1} is constant on intervals where $p_i(Z) \neq 0$ and $p_{[i]}(Z) \neq 0$. That is, for $\mathbb{P}$-a.e. ω

$$Z_t^{i,n_i+1}(\omega) = Z_u^{i,n_i+1}(\omega) \quad \text{for } t \in [u, v],$$

if $p_i(Z_t(\omega)) \neq 0$ *and* $p_{[i]}(Z_t(\omega)) \neq 0$ *for* $t \in (u, v)$.

For a stopping time τ *with* $[\tau] \subset \mathcal{D}'$ *let* $\tau'(\omega) := \inf\{t \geq \tau(\omega) \mid (t, \omega) \notin \mathcal{D}'\}$ *denote the debut of the optional random set* $(\mathcal{B} \cup \mathcal{C}) \cap [\tau, \infty[$. *Then it holds furthermore that* $\tau < \tau'$ *on* $\{\tau < \infty\}$ *and*

$$Z^{i,\mu}_{\tau+t} = Z^{i,\mu}_{\tau} e^{-Z^{i,n_i+1}_{\tau} t} + Z^{i,\mu+1}_{\tau} t e^{-Z^{i,n_i+1}_{\tau} t}$$
$$Z^{i,n_i}_{\tau+t} = Z^{i,n_i}_{\tau} e^{-Z^{i,n_i+1}_{\tau} t}$$

on $[0, \tau' - \tau[$, *for* $0 \leq \mu \leq n_i - 1$ *and* $1 \leq i \leq K$, *up to evanescence.*

If $\mathcal{D}'$ *above is replaced by* $\mathcal{D}$ *and* τ' *is the debut of* $(\cup_{i=1}^K \mathcal{A}_i \cup \mathcal{B} \cup \mathcal{C}) \cap [\tau, \infty[$, *then* $\tau' = \infty$ *and in addition*

$$Z^{i,n_i+1}_{\tau+t} = Z^{i,n_i+1}_{\tau}$$

for $1 \leq i \leq K$, $\mathbb{P}$*-a.s. on* $\{\tau < \infty\}$.

Remark 6.5. *It will be made clear in the proof of the theorem that it is actually sufficient to assume* Z *to be consistent with* $EP(K, n)$ *for (6.12) to hold.*

As an immediate consequence we may state the following corollaries. The notation is the same as in the theorem.

Corollary 6.6. *If* Z *is consistent with* $BEP(K, n)$ *and if the optional random sets* $\{p_i(Z) = 0\}$ *and* $\{p_{[i]}(Z) = 0\}$ *have* $dt \otimes d\mathbb{P}$*-measure zero, then the exponent* Z^{i,n_i+1} *is indistinguishable from* Z^{i,n_i+1}_0, $1 \leq i \leq K$.

Proof. If $\{p_i(Z) = 0\}$ and $\{p_{[i]}(Z) = 0\}$ have $dt \otimes d\mathbb{P}$-measure zero, then $\{p_i(Z) \neq 0\} \cap \{p_{[i]}(Z) \neq 0\} = \mathbb{R}_+ \times \Omega$ up to a $dt \otimes d\mathbb{P}$-nullset. The claim follows using (6.12) and (6.13). ☐

Corollary 6.7. *If* Z *is consistent with* $BEP(K, n)$ *and if the following three points are* $\mathbb{P}$*-a.s. satisfied*

i) $p_i(Z_0) \neq 0$, *for all* $1 \leq i \leq K$,
ii) *there exists no pair of indices* $i \neq j$ *with* $Z^{i,n_i+1}_0 = Z^{j,n_j+1}_0$,
iii) *there exists no pair of indices* $i \neq j$ *with* $2Z^{i,n_i+1}_0 = Z^{j,n_j+1}_0$,

then Z *and hence the interest rate model* $G(x, Z)$ *is quasi deterministic, i.e. all randomness remains* $\mathcal{F}_0$*-measurable. In particular the exponents* Z^{i,n_i+1} *are indistinguishable from* Z^{i,n_i+1}_0, *for* $1 \leq i \leq K$.

Proof. If i), ii) and iii) hold $\mathbb{P}$-a.s. then $[0] \subset \mathcal{D}$. The claim follows from the second part of the theorem setting $\tau = 0$. ☐

6.4. Auxiliary Results

For the proof of the main result we need three auxiliary lemmas, presented in this section. First there is a result on the identification of the coefficients of Itô processes.

Lemma 6.8. *Let*

$$X_t = X_0 + \int_0^t \beta^X_s \, ds + \sum_{j=1}^d \int_0^t \gamma^{X,j}_s \, dW^j_s$$
$$Y_t = Y_0 + \int_0^t \beta^Y_s \, ds + \sum_{j=1}^d \int_0^t \gamma^{Y,j}_s \, dW^j_s$$

be two Itô processes. Then $dt \otimes d\mathbb{P}$-a.s.

$$1_{\{X=Y\}} \sum_{j=1}^{d} (\gamma^{X,j})^2 = 1_{\{X=Y\}} \sum_{j=1}^{d} \gamma^{X,j}\gamma^{Y,j} = 1_{\{X=Y\}} \sum_{j=1}^{d} (\gamma^{Y,j})^2$$
$$1_{\{X=Y\}}\beta^X = 1_{\{X=Y\}}\beta^Y.$$

PROOF. We write $\langle\, . \,,\, . \,\rangle$ for the scalar product in $\mathbb{R}^d$. Then

$$|\langle \gamma^X, \gamma^X\rangle - \langle \gamma^X, \gamma^Y\rangle| = |\langle \gamma^X, \gamma^X - \gamma^Y\rangle| \le \sqrt{\langle \gamma^X, \gamma^X\rangle}\sqrt{\langle \gamma^X - \gamma^Y, \gamma^X - \gamma^Y\rangle}$$

By the occupation times formula, see [**40**, Corollary (1.6), Chapter VI],

$$\int_0^t 1_{\{X_s=Y_s\}} \langle \gamma^X_s - \gamma^Y_s, \gamma^X_s - \gamma^Y_s\rangle\, ds = 0, \quad \text{for all } t < \infty, \mathbb{P}\text{-a.s.}$$

Hence by Hölder inequality

$$\begin{aligned}
&\int_0^t 1_{\{X_s=Y_s\}} |\langle \gamma^X_s, \gamma^X_s\rangle - \langle \gamma^X_s, \gamma^Y_s\rangle|\, ds \\
&\quad \le \int_0^t 1_{\{X_s=Y_s\}} \sqrt{\langle \gamma^X_s, \gamma^X_s\rangle}\sqrt{\langle \gamma^X_s - \gamma^Y_s, \gamma^X_s - \gamma^Y_s\rangle}\, ds \\
&\quad \le \Big(\int_0^t 1_{\{X_s=Y_s\}} \langle \gamma^X_s, \gamma^X_s\rangle\, ds\Big)^{\frac12} \Big(\int_0^t 1_{\{X_s=Y_s\}} \langle \gamma^X_s - \gamma^Y_s, \gamma^X_s - \gamma^Y_s\rangle\, ds\Big)^{\frac12} \\
&\quad = 0, \quad \text{for all } t < \infty, \mathbb{P}\text{-a.s.}
\end{aligned}$$

Thus by symmetry

$$1_{\{X=Y\}}\langle \gamma^X, \gamma^X\rangle = 1_{\{X=Y\}}\langle \gamma^X, \gamma^Y\rangle = 1_{\{X=Y\}}\langle \gamma^Y, \gamma^Y\rangle, \quad dt \otimes d\mathbb{P}\text{-a.s.}$$

By continuity of the processes X and Y there are sequences of stopping times (S_n) and (T_n), $S_n \le T_n$, with $[S_m, T_m] \cap [S_n, T_n] = \emptyset$ for all $m \ne n$ and

$$\{X = Y\} = \bigcup_{n\in\mathbb{N}} [S_n, T_n], \quad \text{up to evanescence.}$$

To see this, let $n \in \mathbb{N}$ and let $S(n,1) := \inf\{t > 0 \mid |X_t - Y_t| = 0\}$. Define $T(n,p) := \inf\{t > S(n,p) \mid |X_t - Y_t| > 0\}$ and inductively

$$S(n,p+1) := \inf\{t > S(n,p) \mid |X_t - Y_t| = 0 \text{ and } \sup_{S(n,p)\le s\le t} |X_s - Y_s| > 2^{-n}\}.$$

Then by continuity we have $\lim_{p\to\infty} S(n,p) = \infty$ for all $n \in \mathbb{N}$ and it follows that $\{X = Y\} = \bigcup_{n,p\in\mathbb{N}} [S(n,p), T(n,p)]$. Now proceed as in [**31**, Lemma I.1.31] to find the sequences (S_n) and (T_n) with the desired properties.

From above we have $1_{\{X=Y\}}(\gamma^X - \gamma^Y)^2 = 0$, $dt \otimes d\mathbb{P}$-a.s. For any $0 \le t < \infty$ therefore $\int_{S_n\wedge t}^{T_n\wedge t} (\gamma^X_s - \gamma^Y_s)\, dW_s = 0$, $\mathbb{P}$-a.s. Hence

$$0 = (X - Y)_{T_n\wedge t} - (X - Y)_{S_n\wedge t} = \int_{S_n\wedge t}^{T_n\wedge t} (\beta^X_s - \beta^Y_s)\, ds, \quad \mathbb{P}\text{-a.s.}$$

We conclude

$$\int_0^t 1_{\{X_s=Y_s\}} (\beta^X_s - \beta^Y_s)\, ds = \sum_{n\in\mathbb{N}} \int_{S_n\wedge t}^{T_n\wedge t} (\beta^X_s - \beta^Y_s)\, ds = 0, \quad \text{for } 0 \le t < \infty, \mathbb{P}\text{-a.s.}$$

Using the same arguments as in the proof of [**25**, Proposition 3.2], we derive the desired result. □

Secondly there are listed two results in matrix algebra.

LEMMA 6.9. *Let* $\gamma = (\gamma_{i,j})$ *be a* $N \times d$*-matrix and define the symmetric non-negative definite* $N \times N$*-matrix* $\alpha := \gamma\gamma^*$*, i.e.* $\alpha_{i,j} = \alpha_{j,i} = \sum_{\lambda=1}^{d} \gamma_{i,\lambda}\gamma_{j,\lambda}$*. Let* I *and* J *denote two arbitrary subsets of* $\{1, \dots, N\}$*. Define*

$$\alpha_{I,J} = \alpha_{J,I} := \sum_{j\in J}\sum_{i\in I} \alpha_{i,j}.$$

Then it holds that $\alpha_{I,I} \geq 0$ *and* $|\alpha_{I,J}| \leq \sqrt{\alpha_{I,I}}\sqrt{\alpha_{J,J}}$*.*

PROOF. For $1 \leq \lambda \leq d$ define $\gamma_{I,\lambda} := \sum_{i\in I} \gamma_{i,\lambda}$. Then by definition

$$\alpha_{I,J} = \sum_{j\in J}\sum_{i\in I}\sum_{\lambda=1}^{d} \gamma_{i,\lambda}\gamma_{j,\lambda} = \sum_{\lambda=1}^{d}\Big(\sum_{i\in I}\gamma_{i,\lambda}\Big)\Big(\sum_{j\in J}\gamma_{j,\lambda}\Big) = \sum_{\lambda=1}^{d}\gamma_{I,\lambda}\gamma_{J,\lambda}.$$

Hence

$$\alpha_{I,I} = \sum_{\lambda=1}^{d} (\gamma_{I,\lambda})^2 \geq 0$$

and by Schwarz inequality

$$|\alpha_{I,J}| \leq \sqrt{\sum_{\lambda=1}^{d} (\gamma_{I,\lambda})^2}\sqrt{\sum_{\lambda=1}^{d} (\gamma_{J,\lambda})^2} = \sqrt{\alpha_{I,I}}\sqrt{\alpha_{J,J}}.$$

□

LEMMA 6.10. *Let* $\alpha = (\alpha_{i,j})$ *be a* $n \times n$*-matrix,* $n \in \mathbb{N}$*, which is diagonally dominant from the right, i.e.*

$$|\alpha_{i,i}| \geq \sum_{\substack{j=1\\ j\neq i}}^{n} |\alpha_{i,j}|$$

$$|\alpha_{i,i}| > \sum_{j=i+1}^{n} |\alpha_{i,j}|, \quad \Big(set \sum_{j=n+1}^{n} \cdots := 0\Big),$$

for all $1 \leq i \leq n$*. Then* α *is regular.*

PROOF. The proof is a slight modification of an argument given in [**42**, Theorem 1.5].

Gaussian elimination: by assumption $|\alpha_{1,1}| > \sum_{j=2}^{n} |\alpha_{1,j}| \geq 0$, in particular $\alpha_{1,1} \neq 0$. If $n = 1$ we are done. If $n > 1$, the elimination step

$$\alpha_{i,j}^{(1)} := \alpha_{i,j} - \frac{\alpha_{i,1}}{\alpha_{1,1}}\alpha_{1,j}, \quad 2 \leq i, j \leq n,$$

leads to the $(n-1) \times (n-1)$-matrix $\alpha^{(1)} = (\alpha_{i,j}^{(1)})_{2\leq i,j\leq n}$. We show that $\alpha^{(1)}$ is diagonal dominant from the right. If $\alpha_{i,1} = 0$, there is nothing to prove for the i-th row. Let $\alpha_{i,1} \neq 0$, for some $2 \leq i \leq n$. We have

$$|\alpha_{i,j}^{(1)}| \geq |\alpha_{i,j}| - \left|\frac{\alpha_{i,1}}{\alpha_{1,1}}\right| |\alpha_{1,j}|, \quad 2 \leq j \leq n.$$

Therefore

$$\begin{aligned}\sum_{j=i+1}^{n}|\alpha_{i,j}^{(1)}| \le \sum_{\substack{j=2\\ j\neq i}}^{n}|\alpha_{i,j}^{(1)}| = \sum_{\substack{j=2\\ j\neq i}}^{n}\left|\alpha_{i,j}-\frac{\alpha_{i,1}}{\alpha_{1,1}}\alpha_{1,j}\right| &\le \sum_{\substack{j=2\\ j\neq i}}^{n}|\alpha_{i,j}| + \left|\frac{\alpha_{i,1}}{\alpha_{1,1}}\right|\sum_{\substack{j=2\\ j\neq i}}^{n}|\alpha_{1,j}| \\ &= \sum_{\substack{j=1\\ j\neq i}}^{n}|\alpha_{i,j}| - |\alpha_{i,1}| + \left|\frac{\alpha_{i,1}}{\alpha_{1,1}}\right|\Big(\sum_{j=2}^{n}|\alpha_{1,j}| - |\alpha_{1,i}|\Big) \\ &< |\alpha_{i,i}| - |\alpha_{i,1}| + \left|\frac{\alpha_{i,1}}{\alpha_{1,1}}\right|\Big(|\alpha_{1,1}| - |\alpha_{1,i}|\Big) \\ &= |\alpha_{i,i}| - \left|\frac{\alpha_{i,1}}{\alpha_{1,1}}\right||\alpha_{1,i}| \le |\alpha_{i,i}^{(1)}|.\end{aligned}$$

Proceed inductively to $\alpha^{(2)}, \dots, \alpha^{(n-1)}$. □

6.5. The Case $BEP(1, n)$

We will treat the case $K = 1$ separately, since it represents a key step in the proof of the general $BEP(K, n)$ case. For simplicity we shall skip the index $i = 1$ and write $n = n_1 \in \mathbb{N}_0$, $p = p_1$, $b^j = b^{1,j}$, $a^{i,j} = a^{1,i;1,j}$, etc. In particular we use the notation of Section 6.2 with $N = n + 2$.

LEMMA 6.11. *Let $n \in \mathbb{N}_0$ and Z be as above. If Z is consistent with $BEP(1, n)$, then necessarily*

$$\begin{aligned} Z_t^i &= Z_0^i e^{-Z_0^{n+1}t} + Z_0^{i+1} t e^{-Z_0^{n+1}t} \\ Z_t^n &= Z_0^n e^{-Z_0^{n+1}t} \\ Z_t^{n+1} &= Z_0^{n+1} + \left(\int_0^t b_s^{n+1}\, ds + \sum_{j=1}^{d}\int_0^t \sigma_s^{n+1,j}\, dW_s^j\right) 1_{\Omega_0}, \end{aligned}$$

for $0 \le i \le n-1$ and $0 \le t < \infty$, $\mathbb{P}$-a.s., where $\Omega_0 := \{p(Z_0) = 0\}$.

Consequently, if Z is consistent with $BEP(1, n)$, then $\{p(Z) = 0\} = \mathbb{R}_+ \times \Omega_0$. Hence $\{Z^{n+1} \neq Z_0^{n+1}\} \subset \{p(Z) = 0\}$. Therefore we may state

COROLLARY 6.12. *If Z is consistent with $BEP(1, n)$, then Z is as in the lemma and*

$$G(x, Z) = p(x, Z) e^{-Z_0^{n+1}x}.$$

Hence the corresponding interest rate model is quasi deterministic, i.e. all randomness remains $\mathcal{F}_0$-measurable.

PROOF OF LEMMA 6.11. Let $n \in \mathbb{N}_0$ and let Z be an Itô process, consistent with $BEP(1, n)$. Fix a point (t, ω) in $\mathbb{R}_+ \times \Omega$. For simplicity we write z_i for $Z_t^i(\omega)$, $a_{i,j}$ for $a_t^{i,j}(\omega)$ and b_i for $b_t^i(\omega)$. The proof relies on expanding equation (6.3) in the point $z=(z_0,\dots,z_{n+1})$. The involved terms are

$$\frac{\partial}{\partial x}G(x,z)=\left(\frac{\partial}{\partial x}p(x,z)-z_{n+1}p(x,z)\right)e^{-z_{n+1}x} \tag{6.14}$$

$$\frac{\partial}{\partial z_i}G(x,z)=\begin{cases}x^ie^{-z_{n+1}x}, & 0\le i\le n\\ -xp(x,z)e^{-z_{n+1}x}, & i=n+1\end{cases} \tag{6.15}$$

$$\frac{\partial^2 G(x,z)}{\partial z_i\partial z_j}=\frac{\partial^2 G(x,z)}{\partial z_j\partial z_i}=\begin{cases}0, & 0\le i,j\le n\\ -x^{i+1}e^{-z_{n+1}x}, & 0\le i\le n, j=n+1\\ x^2p(x,z)e^{-z_{n+1}x}, & i=j=n+1.\end{cases} \tag{6.16}$$

Finally it's useful to know the following relation for $m\in\mathbb{N}_0$

$$\int_0^x \eta^m e^{-z_{n+1}\eta}\,d\eta=\begin{cases}-q_m(x)e^{-z_{n+1}x}+\frac{m!}{z_{n+1}^{m+1}}, & z_{n+1}\neq 0\\ \frac{x^{m+1}}{m+1}, & z_{n+1}=0,\end{cases} \tag{6.17}$$

where $q_m(x)=\sum_{k=0}^m\frac{m!}{(m-k)!}\frac{x^{m-k}}{z_{n+1}^{k+1}}$ is a polynomial in x of order m.

Let's suppose first that $z_{n+1}\neq 0$. Thus, subtracting $\frac{\partial}{\partial x}G(x,Z)$ from both sides of (6.3) we get a null equation of the form

$$q_1(x)e^{-z_{n+1}x}+q_2(x)e^{-2z_{n+1}x}=0, \tag{6.18}$$

which has to hold simultaneously for all $x\ge 0$. The polynomials q_1 and q_2 depend on the z_i's, b_i's and $a_{i,j}$'s. Equality (6.18) implies $q_1=q_2=0$. This again yields that all coefficients of the q_i's have to be zero.

To proceed we have to distinguish the two cases $p(z)\neq 0$ and $p(z)=0$. Let's suppose first the former is true. Then there exists an index $i\in\{0,\dots,n\}$ such that $z_i\neq 0$. Set $m:=\max\{i\le n\mid z_i\neq 0\}$. With regard to (6.15)–(6.17) it follows that $\deg q_2=2m+2$. In particular

$$q_2(x)=a_{n+1,n+1}\frac{z_m^2}{z_{n+1}}x^{2m+2}+\dots,$$

where $\dots$ denotes terms of lower order in x. Hence $a_{n+1,n+1}=0$. But the matrix a has to be nonnegative definite, so necessarily

$$a_{n+1,j}=a_{j,n+1}=0,\quad \text{for all } 1\le j\le n+1.$$

In view of Lemma 6.8 (setting $Y=0$), since we are characterizing a and b up to $dt\otimes d\mathbb{P}$-nullsets, we may assume $a_{i,j}=a_{j,i}=0$, for $0\le j\le n+1$, for all $i\ge m+1$. Thus the degree of q_2 reduces to $2m$. Explicitly

$$q_2(x)=\frac{a_{m,m}}{z_{n+1}}x^{2m}+\dots.$$

Hence $a_{m,m}=0$ and so $a_{m,j}=a_{j,m}=0$, for $0\le j\le n+1$. Proceeding inductively for $i=m-1,m-2,\dots,0$ we finally get that the diffusion matrix a is equal to zero and hence $q_2=0$ is fulfilled.

Now we determine the drift b. By Lemma 6.8, we may assume $b_i=0$ for $m+1\le i\le n$. With regard to (6.14) and (6.15), q_1 reduces therefore to

$$q_1(x)=-b_{n+1}z_mx^{m+1}+\dots.$$

It follows $b_{n+1} = 0$ and it remains

$$
\begin{aligned}
q_1(x) &= (b_m + z_{n+1}z_m)x^m + \sum_{i=0}^{m-1}(b_i - z_{i+1} + z_{n+1}z_i)x^i \\
&= (b_n + z_{n+1}z_n)x^n + \sum_{i=0}^{n-1}(b_i - z_{i+1} + z_{n+1}z_i)x^i.
\end{aligned}
$$

We now turn to the singular cases. If $p(z) = 0$, that is $z_0 = \cdots = z_n = 0$, we may assume $a_{i,j} = a_{j,i} = b_i = 0$, $0 \le j \le n+1$, for all $i \le n$. But this means that $q_1 = q_2 = 0$, independently of the choice of b_{n+1} and $a_{n+1,n+1}$.

For the case where $z_{n+1} = 0$ we need the boundedness assumption $z \in \mathcal{Z}$. By (6.8) it follows that $z_1 = \cdots = z_n = 0$. So by Lemma 6.8 again $a_{i,j} = a_{j,i} = b_i = 0$, $0 \le j \le n+1$, for all $i \ge 1$. Thus in this case equation (6.3) reduces to

$$0 = b_0 - a_{0,0}x$$

and therefore $b_0 = a_{0,0} = 0$.

Summarizing all cases we conclude that necessarily

$$
\begin{aligned}
b_i &= -z_{n+1}z_i + z_{i+1}, \quad 0 \le i \le n-1 \\
b_n &= -z_{n+1}z_n \\
a_{i,j} &= 0, \quad \text{for } (i,j) \ne (n+1, n+1).
\end{aligned}
$$

Whereas b_{n+1} and $a_{n+1,n+1}$ are arbitrary real, resp. nonnegative real, numbers whenever $p(z) = 0$. Otherwise $b_{n+1} = a_{n+1,n+1} = 0$.

The rest of the proof is analogous to the proof of [**25**, Proposition 4.1]. $\square$

6.6. The General Case $BEP(K, n)$

Using again the notation of Section 6.3 we give the proof of the main result for the case $K \ge 2$. The exposure is somewhat messy, which is due to the multi-dimensionality of the problem. The idea however is simple: For a fixed point $(t, \omega) \in \mathbb{R}_+ \times \Omega$ we expand equation (6.3), which turns out to be a linear combination of linearly independent exponential functions over the ring of polynomials, equaling zero. Consequently many of the coefficients have to vanish, which leads to our assertion.

The difficulty is that some exponents may coincide. This causes a considerable number of singular cases which require a separate discussion.

Proof of Theorem 6.4. Let $K \ge 2$, $n = (n_1, \dots, n_K) \in \mathbb{N}_0^K$, and let Z be consistent with $BEP(K, n)$. As in the proof of Lemma 6.11 we fix a point (t, ω) in $\mathbb{R}_+ \times \Omega$ and use the shorthand notation $z_{i,\mu}$ for $Z_t^{i,\mu}(\omega)$, $a_{i,\mu;j,\nu}$ for $a_t^{i,\mu;j,\nu}(\omega)$ and $b_{i,\mu}$ for $b_t^{i,\mu}(\omega)$, etc. Since we are characterizing a and b up to a $dt \otimes d\mathbb{P}$-nullset, we assume that (t, ω) is chosen outside of an exceptional $dt \otimes d\mathbb{P}$-nullset. In particular the lemmas from Section 6.4 shall apply each time we use them.

The strategy is the same as for the case $K = 1$. Thus we expand equation (6.3) in the point $z = (z_{1,0}, \dots, z_{K,n_K+1})$ to get a linear combination of (ideally) linearly independent exponential functions over the ring of polynomials

$$\sum_{i=1}^{K} q_i(x)e^{-z_{i,n_i+1}x} + \sum_{1\le i\le j\le K} q_{i,j}(x)e^{-(z_{i,n_i+1}+z_{j,n_j+1})x} = 0. \tag{6.19}$$

Consequently, all polynomials q_i and $q_{i,j}$ have to be zero. The main difference to the case $K=1$ is that representation (6.19) may not be unique due to the possibly multiple occurrence of the following singular cases

i) $z_{i,n_i+1} = z_{j,n_j+1}$, for $i \neq j$,
ii) $2z_{i,n_i+1} = z_{j,n_j+1} + z_{k,n_k+1}$,
iii) $2z_{i,n_i+1} = z_{j,n_j+1}$,
iv) $z_{i,n_i+1} = z_{j,n_j+1} + z_{k,n_k+1}$,

for some indices $1 \le i,j,k \le K$. However, the lemmas in Section 6.4 and the boundedness assumption $z \in \mathcal{Z}$ are good enough to settle these four cases.

Let's suppose first that $p_i(z) \neq 0$, for all $i \in \{1,\dots,K\}$. To settle Case i), let $\sim$ denote the equivalence relation defined in (6.6). After re-parametrization if necessary we may assume that

$$\{1,\dots,K\}/_{\sim} = \{[1],\dots,[\tilde{K}]\}$$

and $z_{1,n_1+1} < \cdots < z_{\tilde{K},n_{\tilde{K}}+1}$ for some integer $\tilde{K} \le K$. Write $I := \{1,\dots,\tilde{K}\}$. In view of Lemma 6.8 we may assume

$$a_{j,n_j+1;j,n_j+1} = a_{i,n_i+1;i,n_i+1} \text{ and } b_{j,n_j+1} = b_{i,n_i+1} \text{ for all } j \in [i],\ i \in I. \tag{6.20}$$

The proof of (6.12) and (6.13) is divided into four claims.

Claim 1. $a_{i,n_i+1;i,n_i+1} = 0$, *for all* $i \in I$.

Expression (6.19) takes the form

$$\sum_{i\in I} \tilde{q}_i(x)e^{-z_{i,n_i+1}x} + \sum_{\substack{i,j\in I\\ i\le j}} \tilde{q}_{i,j}(x)e^{-(z_{i,n_i+1}+z_{j,n_j+1})x} = 0, \tag{6.21}$$

for some polynomials $\tilde{q}_i$ and $\tilde{q}_{i,j}$. Taking into account Cases ii)–iv), this representation may still not be unique. However if for an index $i \in I$ there exist no $j,k \in I$ such that $2z_{i,n_i+1} = z_{j,n_j+1} + z_{k,n_k+1}$ or $2z_{i,n_i+1} = z_{j,n_j+1}$ (in particular $z_{i,n_i+1} \neq 0$) then we have

$$\tilde{q}_{i,i}(x) = a_{i,n_i+1;i,n_i+1} \frac{\sum_{j\in\mathcal{I}_{[i],\mu_m}} z_{j,\mu_m}^2}{z_{i,n_i+1}} x^{2\mu_m+2} + \dots,$$

where $\mu_m := \max\{\nu \mid \nu \le n_j \text{ and } z_{j,\nu} \neq 0 \text{ for some } j \in [i]\} \in \mathbb{N}_0$. Hence $a_{i,n_i+1;i,n_i+1} = 0$ and Claim 1 is proved for the regular case.

For the singular cases observe first that $z_{i,n_i+1} = 0$ implies $a_{i,n_i+1;i,n_i+1} = 0$, which follows from Lemma 6.8. Now we split I into two disjoint subsets I_1 and I_2, where

$$\begin{aligned} I_1 &:= \{i \in I \mid z_{i,n_i+1} \neq 0 \text{ and there exist } j,k \in I, \text{ such that} \\ &\qquad 2z_{i,n_i+1} = z_{j,n_j+1} + z_{k,n_k+1} \text{ or } 2z_{i,n_i+1} = z_{j,n_j+1}\} \\ I_2 &:= I \setminus I_1. \end{aligned}$$

Observe that $z_{\tilde{K},n_{\tilde{K}}+1} > 0$ implies $\tilde{K} \in I_2$ and $z_{1,n_1+1} < 0$ implies $1 \in I_2$. Since at least one of these events has to happen, the set I_2 is not empty. We have shown above that $a_{i,n_i+1;i,n_i+1} = 0$, for $i \in I_2$. If I_1 is not empty, we will show that for each $i \in I_1$, the parameter z_{i,n_i+1} can be written as a linear combination of z_{j,n_j+1}'s with $j \in I_2$. From this it follows by Lemmas 6.8 and 6.9 that $a_{i,n_i+1;i,n_i+1} = 0$ for all $i \in I_1$ and Claim 1 is completely proved. We proceed as follows. Write $I_1 = \{i_1, \dots, i_r\}$ with $z_{i_1,n_{i_1}+1} < \dots < z_{i_r,n_{i_r}+1}$. For each $i_k \in I_1$ there exists one linear equation of the form

$$(*, \dots, *, 2, *, \dots, *) \begin{pmatrix} z_{i_1,n_{i_1}+1} \\ \vdots \\ z_{i_k,n_{i_k}+1} \\ \vdots \\ z_{i_r,n_{i_r}+1} \end{pmatrix} = \alpha_k,$$

where $*$ stands for 0 or -1, but at most one -1 on each side of 2. The α_k on the right hand side is either 0 or z_{i,n_i+1} or $z_{i,n_i+1} + z_{j,n_j+1}$ for some indices $i, j \in I_2$. Hence we get the system of linear equations

$$\begin{pmatrix} 2 & * & \dots & * \\ * & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ * & \dots & * & 2 \end{pmatrix} \begin{pmatrix} z_{i_1,n_{i_1}+1} \\ \vdots \\ \vdots \\ z_{i_r,n_{i_r}+1} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \vdots \\ \alpha_r \end{pmatrix}.$$

By Lemma 6.10, the matrix on the left hand side is invertible, from which follows our assertion.

Claim 2. *$a_{j,n_j+1;k,\nu} = a_{k,\nu;j,n_j+1} = 0$, for $0 \le \nu \le n_k$, for all $1 \le j, k \le K$.*

In view of (6.20), Claim 2 follows immediately from Claim 1 and Lemma 6.9.

Analogous to the notation introduced in (6.7) we set

$$b_{[i],\mu} := \sum_{j \in \mathcal{I}_{[i],\mu}} b_{j,\mu}$$

$$\sigma_{[i],\mu;\lambda} := \sum_{j \in \mathcal{I}_{[i],\mu}} \sigma_{j,\mu;\lambda}$$

$$a_{[i],\mu;k,\nu} := \sum_{j \in \mathcal{I}_{[i],\mu}} a_{j,\mu;k,\nu},$$

for $0 \le \mu \le n_{[i]}$, $0 \le \nu \le n_k$, $1 \le k \le K$, $1 \le \lambda \le d$, $i \in I$, and

$$a_{[i],\mu;[k],\nu} := \sum_{l \in \mathcal{I}_{[k],\nu}} \sum_{j \in \mathcal{I}_{[i],\mu}} a_{j,\mu;l,\nu},$$

for $0 \le \mu \le n_{[i]}$, $0 \le \nu \le n_{[k]}$, $i, k \in I$.

Claim 3. *If $z_{[i],\mu} = 0$, for $i \in I$ and $\mu \in \{0, \dots, n_{[i]}\}$, then*

$$b_{[i],\mu} = a_{[i],\mu;[i],\mu} = a_{[i],\mu;k,\nu} = a_{k,\nu;[i],\mu} = 0,$$

for all $0 \le \nu \le n_k$, $1 \le k \le K$.

Notice that $a_{[i],\mu;[i],\mu} = \sum_{\lambda=1}^{d} \sigma_{[i],\mu;\lambda}^2$. Hence Claim 3 follows by Lemma 6.8 and Lemma 6.9.

Claim 4. *$b_{i,n_i+1} = 0$, for all $i \in I$ such that $p_{[i]}(z) \neq 0$.*

Suppose first that $z_{i,n_i+1} \neq 0$, for all $i \in I$. Let $i \in I$ such that $p_{[i]}(z) \neq 0$, and let's assume there exist no $j, k \in I$ with $z_{i,n_i+1} = z_{j,n_j+1} + z_{k,n_k+1}$. How does the polynomial $\tilde{q}_i$ in (6.21) look like? With regard to (6.20), Claim 2, Lemma 6.8 and equalities (6.14)–(6.17) the contributing terms are

$$\begin{aligned} \frac{\partial}{\partial x} p_j(x, z) e^{-z_{j,n_j+1}x} &= \Big(\Big(\sum_{\mu=1}^{\mu_m \wedge n_j} z_{j,\mu} x^{\mu-1}\Big) - z_{i,n_i+1}\Big(\sum_{\mu=0}^{\mu_m \wedge n_j} z_{j,\mu} x^{\mu}\Big)\Big) e^{-z_{i,n_i+1}x}, \\ \sum_{\mu=0}^{n_j+1} b_{j,\mu} \frac{\partial}{\partial z_{j,\mu}} G(x, z) &= \Big(\Big(\sum_{\mu=0}^{\mu_m \wedge n_j} b_{j,\mu} x^{\mu}\Big) - b_{i,n_i+1}\Big(\sum_{\mu=0}^{\mu_m \wedge n_j} z_{j,\mu} x^{\mu+1}\Big)\Big) e^{-z_{i,n_i+1}x} \end{aligned} \tag{6.22}$$

and

$$\begin{aligned} &- 2\frac{1}{2}\Big(\sum_{\mu=0}^{n_j} a_{j,\mu;k,\nu} \frac{\partial}{\partial z_{j,\mu}} G(x, z) \int_0^x \frac{\partial}{\partial z_{k,\nu}} G(\eta, z)\, d\eta\Big) \\ &\quad = -\Big(\sum_{\mu=0}^{\mu_m \wedge n_j} a_{j,\mu;k,\nu} \frac{n_k!}{z_{k,n_k+1}^{n_k+1}} x^{\mu}\Big) e^{-z_{i,n_i+1}x} - \binom{\text{polynomial}}{\text{in } x} e^{-(z_{i,n_i+1}+z_{k,n_k+1})x}, \end{aligned} \tag{6.23}$$

for $0 \leq \nu \leq n_k$, for all $1 \leq k \leq K$ and $j \in [i]$. We have used the integer

$$\mu_m := \max\{\lambda \mid \lambda \leq n_l \text{ and } z_{l,\lambda} \neq 0 \text{ for some } l \in [i]\}.$$

Define $\tilde{\mu}_m := \max\{\lambda \mid \lambda \leq n_{[i]} \text{ and } z_{[i],\lambda} \neq 0\} \in \mathbb{N}_0$. Obviously $\tilde{\mu}_m \leq \mu_m$. By Claim 3 we have $a_{[i],\mu;k,\nu} = 0$, for all $\tilde{\mu}_m < \mu \leq n_{[i]}$. Thus summing up the above expressions over $j \in [i]$ we get

$$\tilde{q}_i(x) = -b_{i,n_i+1} z_{[i],\tilde{\mu}_m} x^{\tilde{\mu}_m+1} + \dots. \tag{6.24}$$

Consequently $b_{i,n_i+1} = 0$ in the regular case.

For the singular cases the boundedness assumption $z \in \mathcal{Z}$ is essential. We split I into two disjoint subsets J_1 and J_2, where

$$\begin{aligned} J_1 := \{i \in I \mid &\text{ there exist } j, k \in I, \text{ such that } z_{i,n_i+1} = z_{j,n_j+1} + z_{k,n_k+1} \\ &\text{ and } z_{j,n_j+1} > 0 \text{ and } z_{k,n_k+1} > 0\} \\ J_2 := I \setminus J_1&. \end{aligned}$$

Notice that in any case $1 \in J_2$. We have shown above that for each $i \in J_2$ such that z_{i,n_i+1} is not the sum of two other z_{j,n_j+1}'s it follows that $b_{i,n_i+1} = 0$. We will now show that $b_{i,n_i+1} = 0$ for all $i \in J_2$. Let $i \in J_2$ and assume there exist $j, k \in I$ with $z_{i,n_i+1} = z_{j,n_j+1} + z_{k,n_k+1}$. Then necessarily one of the summands is strictly less than zero. Without loss of generality $z_{j,n_j+1} < 0$. Since $z \in \mathcal{Z}$, we have $p_{[j]}(z) = 0$, see (6.8). Thus $a_{[j],\mu;[j],\mu} = 0$ by Claim 3 and therefore $a_{[j],\mu;k,\nu} = 0$, for all $0 \leq \mu \leq n_{[j]}$, $0 \leq \nu \leq n_k$, $1 \leq k \leq K$. The contributing terms to the polynomial in front of $e^{-z_{i,n_i+1}x}$, i.e. $\tilde{q}_i + \tilde{q}_{j,k} + \dots$, are those in (6.22) and (6.23) and also

$$\begin{aligned} &- 2\frac{1}{2} a_{l,\mu;m,\nu}\Big(\frac{\partial G(x,z)}{\partial z_{l,\mu}} \int_0^x \frac{\partial G(\eta,z)}{\partial z_{m,\nu}}\, d\eta + \frac{\partial G(x,z)}{\partial z_{m,\nu}} \int_0^x \frac{\partial G(\eta,z)}{\partial z_{l,\mu}}\, d\eta\Big) \\ &= -a_{l,\mu;m,\nu}\Big(x^{\mu} e^{-z_{j,n_j+1}x} \int_0^x \eta^{\nu} e^{-z_{k,n_k+1}\eta} d\eta + x^{\nu} e^{-z_{k,n_k+1}x} \int_0^x \eta^{\mu} e^{-z_{j,n_j+1}\eta} d\eta\Big), \end{aligned} \tag{6.25}$$

for $0 \leq \mu \leq n_l$, $0 \leq \nu \leq n_m$, $l \in [j]$, $m \in [k]$. However, summing up – for fixed μ, m and ν – the right hand side of (6.25) over $l \in \mathcal{I}_{[j],\mu}$ gives zero. Hence the terms in (6.25) actually don't contribute to the meant polynomial. The same conclusion can be drawn for all $j, k \in I$ with the property that $z_{i,n_i+1} = z_{j,n_j+1} + z_{k,n_k+1}$. It finally follows as in the regular case that $b_{i,n_i+1} = 0$ for all $i \in J_2$.

If J_1 is not empty, we show that for each $i \in J_1$, the parameter z_{i,n_i+1} can be written as a linear combination of z_{j,n_j+1}'s with $j \in J_2$. From this it follows by Lemma 6.8 that $b_{i,n_i+1} = 0$ for all $i \in J_1$. We proceed as follows. Write $J_1 = \{i_1, \ldots, i_{r'}\}$ with $z_{i_1,n_{i_1}+1} < \cdots < z_{i_{r'},n_{i_{r'}}+1}$. For each $i_k \in J_1$ there exists one linear equation of the form

$$\big(*, \ldots, *, 1, 0, \ldots, 0\big) \begin{pmatrix} z_{i_1,n_{i_1}+1} \\ \vdots \\ z_{i_k,n_{i_k}+1} \\ \vdots \\ z_{i_{r'},n_{i_{r'}}+1} \end{pmatrix} = \alpha'_k,$$

where $*$ stands for 0 or -1, but at most two of them are -1. The α'_k on the right hand side is either 0 or z_{i,n_i+1} or $z_{i,n_i+1} + z_{j,n_j+1}$ for some indices $i, j \in J_2$ with $z_{i,n_i+1} > 0$ and $z_{j,n_j+1} > 0$. Obviously α'_1 is of the latter form. Hence we get the system of linear equations

$$\begin{pmatrix} 1 & 0 & \ldots & 0 \\ * & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \ldots & * & 1 \end{pmatrix} \begin{pmatrix} z_{i_1,n_{i_1}+1} \\ \vdots \\ \vdots \\ z_{i_{r'},n_{i_{r'}}+1} \end{pmatrix} = \begin{pmatrix} z_{i,n_i+1} + z_{j,n_j+1} \\ \alpha'_2 \\ \vdots \\ \alpha'_{r'} \end{pmatrix},$$

for some $i, j \in J_2$. On the left hand side stands a lower-triangular matrix, which is therefore invertible. Hence Claim 4 is proved in the case where $z_{i,n_i+1} \neq 0$ for all $i \in I$.

Assume now that there exists $i \in I$ with $z_{i,n_i+1} = 0$. Then $i \in J_2$. We have to make sure that also in this case $b_{j,n_j+1} = 0$, for all $j \in J_2$. Clearly b_{i,n_i+1} is zero by Lemma 6.8. The problem is that $z_{j,n_j+1} = z_{i,n_i+1} + z_{j,n_j+1}$ for all $j \in J_2$. But following the lines above it is enough to show $a_{[i],\mu;[i],\mu} = 0$, for all $0 \leq \mu \leq n_{[i]}$. From the boundedness assumption $z \in \mathcal{Z}$ we know that $p_{[i]}(z) = z_{[i],0}$, see (6.8). Hence $a_{[i],\mu;[i],\mu} = 0$, for $1 \leq \mu \leq n_{[i]}$. Suppose there is no pair of indices $j, k \in I \backslash \{i\}$ with $z_{j,n_j+1} + z_{k,n_k+1} = 0$. Summing up the contributing terms in (6.22) and (6.23) over $j \in [i]$ we get the polynomial in front of e^0, i.e.

$$\tilde{q}_i(x) + \tilde{q}_{i,i}(x) = -a_{[i],0;[i],0}x + \ldots, \tag{6.26}$$

hence $a_{[i],0;[i],0} = 0$. If there exist a pair of indices $j, k \in I \setminus \{i\}$ with $z_{j,n_j+1} + z_{k,n_k+1} = 0$, then one of these summands is strictly less than zero. Arguing as before, the polynomial in front of e^0 remains of the form (6.26) and again $a_{[i],0;[i],0} = 0$. Thus Claim 4 is completely proved.

Up to now we have established (6.12) and (6.13) under the hypothesis that $p_i(z) \neq 0$, for all $i \in \{1, \ldots, K\}$. Suppose now, there is an index $i \in \{1, \ldots, K\}$ with $p_i(z) = 0$. By Lemma 6.8, we may assume $a_{i,\mu;i,\mu} = b_{i,\mu} = 0$, for all $0 \leq \mu \leq n_i$. But then Lemma 6.9 tells us that none of the terms including the index i appears in (6.19). In particular $a_{i,n_i+1;i,n_i+1}$ and b_{i,n_i+1} can be chosen arbitrarily without affecting equation (6.19). This means that we may skip i and proceed, after a re-parametrization if necessary, with the remaining index set $\{1,\dots,K-1\}$ to establish Claims 1–4 as above.

This all has to hold for $dt\otimes d\mathbb{P}$-a.e. (t,ω). Hence (6.12) and (6.13) are fully proved. A closer look to the proof of (6.12), i.e. Claim 1, shows that the boundedness assumption $z\in\mathcal{Z}$ was not explicitly used there. Whence Remark 6.5.

Next we prove that the exponents Z^{i,n_i+1} are locally constant on intervals where $p_i(Z)$ and $p_{[i]}(Z)$ do not vanish. Let $v\geq 0$ be a rational number and let $T_v := \inf\{t>v \mid p_i(Z_t)=0 \text{ or } p_{[i]}(Z_t)=0\}$ denote the debut of the optional set $[v,\infty[\cap\mathcal{A}_i$. By (6.12) and (6.13) and the continuity of Z we have that Z^{i,n_i+1} is $\mathbb{P}$-a.s. constant on $[v,T_v]$, hence $\mathbb{P}$-a.s. constant on every such interval $[v,T_v]$. Since every open interval where $p_i(Z_t)\neq 0$ or $p_{[i]}(Z_t)\neq 0$ is covered by a countable union of intervals $[v,T_v]$ and by continuity of Z the assertion follows and the first part of the theorem is proved.

For establishing the second part of the theorem let τ be a stopping time with $[\tau]\in\mathcal{D}'$ and $\mathbb{P}(\tau<\infty)>0$. Define the stopping time $\tau'(\omega):=\inf\{t\geq\tau(\omega)\mid (t,\omega)\notin\mathcal{D}'\}$. By continuity of Z, we conclude that $\tau<\tau'$ on $\{\tau<\infty\}$. Choose a point (t,ω) in $[\tau,\tau'[$. We use shorthand notation as above.

By definition of $\mathcal{D}'$ we can exclude the singular cases $z_{i,n_i+1}=z_{j,n_j+1}$ or $2z_{i,n_i+1}=z_{j,n_j+1}$, for $i\neq j$. In particular $\tilde{K}=K$, hence $I=\{1,\dots,K\}$. First we show that the diffusion matrix for the coefficients of the polynomials $p_i(z)$ vanishes.

Claim 5. *$a_{i,\mu;j,\nu}=a_{j,\nu;i,\mu}=0$, for $0\leq\mu\leq n_i$, for $0\leq\nu\leq n_j$, for all $i,j\in I$.*

By Lemma 6.8 it's enough to prove that the diagonal $a_{i,\mu;i,\mu}$ vanishes for $0\leq\mu\leq n_i$ and $i\in I$. If there is an index $i\in I$ with $p_i(z)=0$ then argued as above $a_{i,\mu;i,\mu}=b_{i,\mu}=0$, for all $0\leq\mu\leq n_i$, and we may skip the index i. Hence we assume now that there is a $K'\leq K$ such that $p_i(z)\neq 0$ (and thus $z_{i,n_i+1}\geq 0$, since $z\in\mathcal{Z}$) for all $1\leq i\leq K'$. Let $I':=\{1,\dots,K'\}$. For handling the singular cases, we split I' into two disjoint subsets I'_1 and I'_2, where

$$\begin{aligned}
I'_1 &:= \{i\in I' \mid z_{i,n_i+1}>0 \text{ and there exist } j,k\in I',\\
&\qquad\qquad \text{such that } 2z_{i,n_i+1}=z_{j,n_j+1}+z_{k,n_k+1}\}\\
I'_2 &:= I'\setminus I'_1.
\end{aligned}$$

Hence $z_{i,n_i+1}=0$ for $i\in I'$ implies $i\in I'_2$. We have already shown in the proof of Claim 4 that in this case $a_{i,\mu;i,\mu}=0$, for all $0\leq\mu\leq n_i$. The same follows for $i\in I'_2$ with $z_{i,n_i+1}>0$, as it was demonstrated for the case $K=1$.

Now let $i\in I'_1$ and let $l,m\in I'$, such that $l\leq m$ and $2z_{i,n_i+1}=z_{l,n_l+1}+z_{m,n_m+1}$. Thus the polynomial in front of $e^{-2z_{i,n_i+1}x}$ is $q_{i,i}+q_{l,m}+\dots$, and among the contributing terms are also those in (6.25). If l or m is in I'_2, those are all zero. Write $I'_1=\{i_1,\dots,i_{r''}\}$ with $z_{i_1,n_{i_1}+1}<\dots<z_{i_{r''},n_{i_{r''}}+1}$. Then necessarily $l\in I'_2$ in the above representation for $z_{i_1,n_{i_1}+1}$. Thus the polynomial in front of $e^{-2z_{i_1,n_{i_1}+1}x}$ is q_{i_1,i_1}. It follows $a_{i_1,\mu;i_1,\mu}=0$, for all $0\leq\mu\leq n_{i_1}$, as it was demonstrated for the case $K=1$. Proceeding inductively for $i_2,\dots,i_{r''}$, we derive eventually that $a_{i,\mu;i,\mu}=0$, for all $0\leq\mu\leq n_i$ and $i\in I'$. This establishes Claim 5.

We are left with the task of determining the drift of the coefficients in $p_i(z)$. By (6.13), we have $b_{i,n_i+1} = 0$ for all $i \in I'$. Straightforward calculations show that (6.19) reduces to

$$\sum_{i=1}^{K'} q_i(x) e^{-z_{i,n_i+1}x} = 0,$$

with

$$q_i(x) = (b_{i,n_i} + z_{i,n_i+1} z_{i,n_i}) x^{n_i} + \sum_{\mu=0}^{n_i-1} (b_{i,\mu} - z_{i,\mu+1} + z_{i,n_i+1} z_{i,\mu}) x^{\mu}.$$

We conclude that for all $1 \le i \le K$ (in particular if $p_i(z) = 0$)

$$\begin{aligned} b_{i,\mu} &= z_{i,\mu+1} - z_{i,n_i+1} z_{i,\mu}, \quad 0 \le \mu \le n_i - 1 \\ b_{i,n_i} &= -z_{i,n_i+1} z_{i,n_i}. \end{aligned} \tag{6.27}$$

By continuity of Z, Claim 5 and (6.27) hold pathwise on the semi open interval $[\tau(\omega), \tau'(\omega)[$ for almost every ω. Therefore $Z_{\tau+\,.}$ is of the claimed form on $[0, \tau' - \tau[$.

Now replace $\mathcal{D}'$ by $\mathcal{D}$ and proceed as above. By (6.11) we have $\tau < \tau'$ on $\{\tau < \infty\}$, and since $\mathcal{D} \subset \mathcal{D}'$, all the above results remain valid. In addition $p_i(z) = p_{[i]}(z) \neq 0$ and thus $a_{i,n_i+1;i,n_i+1} = b_{i,n_i+1} = 0$, for all $1 \le i \le K$, by (6.12) and (6.13). Hence $Z^{i,n_i+1}_{\tau+\,.} = Z^{i,n_i+1}_{\tau}$ on $[0, \tau' - \tau[$, for all $1 \le i \le K$, up to evanescence. But this again implies $\tau' = \infty$ by the continuity of Z. $\square$

6.7. E-Consistent Itô Processes

An Itô process Z is by definition consistent with a family $\{G(\,.\,, z)\}_{z\in\mathcal{Z}}$ if and only if $\mathbb{P}$ is a martingale measure for the discounted bond price processes. We could generalize this definition and call a process Z *e-consistent* with $\{G(\,.\,, z)\}_{z\in\mathcal{Z}}$ if there exists an equivalent martingale measure $\mathbb{Q}$. Then obviously consistency implies e-consistency, and e-consistency implies the absence of arbitrage opportunities, as it is well known.

In case where the filtration is generated by the Brownian motion W, i.e. $(\mathcal{F}_t) = (\mathcal{F}^W_t)$, we can give the following stronger result:

Proposition 6.13. *Let $K \in \mathbb{N}$ and $n = (n_1, \ldots, n_K) \in \mathbb{N}_0^K$. If $(\mathcal{F}_t) = (\mathcal{F}^W_t)$, then any Itô process Z which is e-consistent with $BEP(K, n)$, is of the form as stated in Theorem 6.4.*

Proof. Let Z be an e-consistent Itô process under $\mathbb{P}$, and let $\mathbb{Q}$ be an equivalent martingale measure. Since $(\mathcal{F}_t) = (\mathcal{F}^W_t)$, we know that all $\mathbb{P}$-martingales have the representation property relative to W. By Girsanov's theorem it follows therefore that Z remains an Itô process under $\mathbb{Q}$, which is consistent with $BEP(K, n)$. The drift coefficients $b^{i,\mu}$ change under $\mathbb{Q}$ into $\tilde{b}^{i,\mu}$. Whereas $b^{i,\mu} = \tilde{b}^{i,\mu}$ on $\{a^{i,\mu;i,\mu} = 0\}$, $dt \otimes d\mathbb{P}$-a.s. The diffusion matrix a remains the same. Therefore and since the measures $dt \otimes d\mathbb{Q}$ and $dt \otimes d\mathbb{P}$ are equivalent on $\mathbb{R}_+ \times \Omega$, the Itô process Z is of the form as stated in Theorem 6.4. $\square$

Notice that in this case the expression quasi deterministic, i.e. $\mathcal{F}_0$-measurable, in Corollaries 6.7 and 6.12 can be replaced by purely deterministic.

6.8. The Diffusion Case

The main result from Section 6.3 reads much clearer for diffusion processes. In all applications the generic Itô process Z on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\le t<\infty}, \mathbb{P})$ given by (6.9) is rather the solution of a stochastic differential equation

$$Z_t^{i,\mu} = Z_0^{i,\mu} + \int_0^t b^{i,\mu}(s, Z_s)\, ds + \sum_{\lambda=1}^d \int_0^t \sigma^{i,\mu;\lambda}(s, Z_s)\, dW_s^\lambda, \tag{6.28}$$

for $0 \le \mu \le n_i + 1$ and $1 \le i \le K$, where b and σ are some Borel measurable mappings from $\mathbb{R}_+ \times \mathbb{R}^N$ to $\mathbb{R}^N$, resp. $\mathbb{R}^{N\times d}$.

The coefficients b and σ could be derived by statistical inference methods from the daily observations of the diffusion Z made by some central bank. These observations are of course made under the objective probability measure. Hence $\mathbb{P}$ is not a martingale measure in applications of this kind.

On the other hand we want a model for pricing interest rate sensitive securities. Thus the diffusion has to be e-consistent. If we assume that $(\mathcal{F}_t) = (\mathcal{F}_t^W)$, the last section applies. To stress the fact that $\mathcal{F}_0^W$-measurable functions are deterministic, we denote the initial values of the diffusion in (6.28) with small letters $z_0^{i,\mu}$.

Since all reasonable theory for stochastic differential equations requires continuity properties of the coefficients, we shall assume in the sequel that $b(t,z)$ and $\sigma(t,z)$ are continuous in z. The main result for e-consistent diffusion processes is divided into the two following theorems. The first one only requires consistency with $EP(K,n)$.

THEOREM 6.14. *Let $K \in \mathbb{N}$, $n = (n_1, \dots, n_K) \in \mathbb{N}_0^K$, $(\mathcal{F}_t) = (\mathcal{F}_t^W)$, the diffusion Z, b and σ as above. If Z is e-consistent with $BEP(K,n)$ or with $EP(K,n)$, then necessarily the exponents are constant*

$$Z^{i,n_i+1} \equiv z_0^{i,n_i+1},$$

for all $1 \le i \le K$.

PROOF. The significant difference to the proof of Theorem 6.4 is that now the diffusion matrix a and the drift b depend continuously on z.

First observe that the following sets of singular values

$$\mathcal{M} := \bigcup_{i=1}^K \left\{ z \in \mathbb{R}^N \mid p_i(z) = 0 \text{ or } p_{[i]}(z) = 0 \right\}$$
$$\mathcal{N} := \{ z \in \mathbb{R}^N \mid z_{i,n_i+1} = z_{j,n_j+1} + z_{k,n_k+1} \text{ for some } 1 \le i,j,k \le K \}$$

are contained in a finite union of hyperplanes of $\mathbb{R}^N$, see (6.10). Hence $(\mathcal{M} \cup \mathcal{N}) \subset \mathbb{R}^N$ has Lebesgue measure zero. Thus the topological closure of $\mathcal{G} := \mathbb{R}^N \setminus (\mathcal{M} \cup \mathcal{N})$ is $\mathbb{R}^N$.

Now let Z be the diffusion in (6.28) which is e-consistent either with $BEP(K,n)$ or $EP(K,n)$. A closer look to the proof of Claim 4 shows that the boundedness assumption $z \in \mathcal{Z}$ was not used for the regular case $z \in \mathcal{G}$, see (6.24). Combining this with (6.12), (6.13) and Remark 6.5 we conclude that for any $1 \le i \le K$ and $1 \le \lambda \le d$

$$b^{i,n_i+1}(t, Z_t(\omega)) = \sigma^{i,n_i+1;\lambda}(t, Z_t(\omega)) = 0, \quad \text{for } (t,\omega) \in \{Z \in \mathcal{G}\} \setminus N,$$

where N is a $\mathcal{R}_+ \otimes \mathcal{F}$-measurable $dt \otimes d\mathbb{P}$-nullset. By the very definition of the product measure

$$0 = \int_N 1 \, dt \otimes d\mathbb{P} = \int_{\mathbb{R}_+} \mathbb{P}[N_t] \, dt,$$

where $N_t := \{\omega \mid (t,\omega) \in N\} \in \mathcal{F}$. Consequently $\mathbb{P}[N_t] = 0$ for a.e. $t \in \mathbb{R}_+$. Hence by continuity of $b(t, .)$ and $\sigma(t, .)$

$$b^{i,n_i+1}(t, .) = \sigma^{i,n_i+1;\lambda}(t, .) = 0 \tag{6.29}$$

on $\mathrm{supp}(Z_t) \cap \mathcal{G}$, for a.e. $t \in \mathbb{R}_+$. Here $\mathrm{supp}(Z_t)$ denotes the support of the (n.b. regular) distribution of Z_t, which is by definition the smallest closed set $A \subset \mathbb{R}^N$ with $\mathbb{P}[Z_t \in A] = 1$. Thus again by continuity of $b(t, .)$ and $\sigma(t, .)$ equality (6.29) holds for a.e. $t \in \mathbb{R}_+$ on the closure of $\mathrm{supp}(Z_t) \cap \mathcal{G}$, which is $\mathrm{supp}(Z_t)$. Hence we may replace the functions $b^{i,n_i+1}(t, .)$ and $\sigma^{i,n_i+1;\lambda}(t, .)$ by zero for almost every t without changing the diffusion Z, whence the assertion follows. ☐

The sum of two real valued diffusion processes with coefficients continuous in some argument is again a real valued diffusion with coefficients continuous in that argument. Consequently we may assume that the exponents z_0^{i,n_i+1} of the above e-consistent diffusion are mutually distinct. Since otherwise we add the corresponding polynomials to get in a canonical way an $\mathbb{R}^{\tilde{N}}$-valued diffusion $\tilde{Z}$ which is e-consistent with $BEP(\tilde{K}, \tilde{n})$, resp. $EP(\tilde{K}, \tilde{n})$, for some $\tilde{K} < K$, $\tilde{N} < N$ and some $\tilde{n} \in \mathbb{N}_0^{\tilde{K}}$. Clearly $\tilde{Z}$ provides the same interest rate model as Z and its coefficients are continuous in z.

For the second theorem we have to require e-consistency with $BEP(K, n)$. After a re-parametrization if necessary we may thus assume that

$$0 \le z_0^{1,n_1+1} < \cdots < z_0^{K,n_K+1},$$

see (6.8). The sequel of Theorem 6.14 reads now

Theorem 6.15. *If Z is e-consistent with $BEP(K, n)$, then it is non-trivial only if there exists a pair of indices $1 \le i < j \le K$, such that*

$$2z_0^{i,n_i+1} = z_0^{j,n_j+1}.$$

Proof. If there is no pair of indices $1 \le i < j \le K$ such that $2z_0^{i,n_i+1} = z_0^{j,n_j+1}$, then $\mathcal{D}' = \mathbb{R}_+ \times \Omega$. But then Z is deterministic by the second part of Theorem 6.4. ☐

The message of Theorem 6.14 is the following: There is no possibility for modelling the term structure of interest rates by exponential-polynomial families with varying exponents driven by diffusion processes. From this point of view there is no use for daily estimations of the exponents of some exponential-polynomial type functions like G_{NS} or G_S. Once the exponents are chosen, they have to be kept constant. Furthermore there is a strong restriction on this choice by Theorem 6.15. It will be shown in the next section what this means for G_{NS} and G_S in particular.

Remark 6.16. *The boundedness assumption in Theorem 6.15 – that is e-consistency with BEP(K,n) – is essential for the strong (negative) result to be valid. It can be easily checked, that $G(x, z) = z_0 + z_1 x \in EP(1, 1)$ allows for a non-trivial consistent diffusion process, see* [**26**, Section 7].

Remark 6.17. *The choice of an infinite time horizon for traded bonds is not a restriction, see (6.1). Indeed, we can limit our considerations on bonds* $P(t,T)$ *which mature within a given finite time interval* $[0,T^*]$*. Consequently, the HJM drift condition (6.3) can only be deduced for* $x \in [0, T^* - t]$*, for* $dt \otimes d\mathbb{P}$*-a.e.* $(t,\omega) \in [0,T^*] \times \Omega$*. But the functions appearing in (6.3) are analytic in* x*. Hence whenever* $t < T^*$*, relation (6.3) extends to all* $x \geq 0$*. All conclusions on e-consistent Itô processes* $(Z_t)_{0 \leq t \leq T^*}$ *can now be drawn as before.*

6.9. Applications

In this section we apply the results on e-consistent diffusion processes to the Nelson–Siegel and Svensson families.

6.9.1. The Nelson–Siegel family.

$$G_{NS}(x,z) = z_1 + (z_2 + z_3 x)e^{-z_4 x}$$

In view of Theorem 6.14 we have $z_4 > 0$. Hence it's immediate from Theorem 6.15 that there is no non-trivial e-consistent diffusion. This result has already been obtained in [**25**] for e-consistent Itô processes.

6.9.2. The Svensson family.

$$G_S(x,z) = z_1 + (z_2 + z_3 x)e^{-z_5 x} + z_4 x e^{-z_6 x}$$

By Theorems 6.14 and 6.15 there remain the two choices

i) $2z_6 = z_5 > 0$
ii) $2z_5 = z_6 > 0$

We shall identify the e-consistent diffusion process $Z = (Z^1, \dots, Z^6)$ in both cases. Let $\mathbb{Q}$ be an equivalent martingale measure. Under $\mathbb{Q}$ the diffusion Z transforms into a consistent one. Now we proceed as in the proof of Theorem 6.4. The expansion (6.19) reads as follows

$$\begin{aligned} Q_1(x) + Q_2(x)e^{-z_5 x} + Q_3(x)e^{-z_6 x} \\ + Q_4(x)e^{-2z_5 x} + Q_5(x)e^{-(z_5+z_6)x} + Q_6(x)e^{-2z_6 x} = 0, \end{aligned}$$

for some polynomials $Q_1 \dots, Q_6$. Explicitly

$$\begin{aligned} Q_1(x) &= -a_{1,1}x + \dots \\ Q_2(x) &= -a_{1,3}x^2 + \dots \\ Q_3(x) &= -a_{1,4}x^2 + \dots \\ Q_4(x) &= \frac{a_{3,3}}{z_5}x^2 + \dots \\ Q_6(x) &= \frac{a_{4,4}}{z_6}x^2 + \dots \end{aligned}$$

where $\dots$ denotes terms of lower order in x. Hence $a_{1,1} = 0$ in any case. By the usual arguments (the matrix a is nonnegative definite) the degree of Q_2 and Q_3 reduces to at most 1. Thus in both Cases i) and ii) it follows $a_{3,3} = a_{4,4} = 0$. It remains

$$\begin{aligned} Q_1(x) &= b_1 \\ Q_2(x) &= (b_3 + z_3 z_5)x + b_2 - z_3 - \frac{a_{2,2}}{z_5} + z_2 z_5 \\ Q_3(x) &= (b_4 + z_4 z_6)x - z_4 \\ Q_4(x) &= \frac{a_{2,2}}{z_5} \end{aligned} \tag{6.30}$$

while $Q_5 = Q_6 = 0$. Since in Case i) it must hold that $Q_4 = 0$, we have $a_{2,2} = 0$ and Z is deterministic. We conclude that there is no non-trivial e-consistent diffusion in Case i).

In Case ii) the condition $Q_3 + Q_4 = 0$ leads to

$$a_{2,2} = z_4 z_5. \tag{6.31}$$

Hence a possibility for a non deterministic consistent diffusion Z. We derive from (6.30) and (6.31)

$$\begin{aligned} b_1 &= 0 \\ b_2 &= z_3 + z_4 - z_5 z_2 \\ b_3 &= -z_5 z_3 \\ b_4 &= -2 z_5 z_4. \end{aligned}$$

Therefore the dynamics of $Z^1, Z^3, \dots, Z^6$ are deterministic. In particular

$$\begin{aligned} Z_t^1 &\equiv z_0^1 \\ Z_t^3 &= z_0^3 e^{-z_0^5 t} \\ Z_t^4 &= z_0^4 e^{-2 z_0^5 t}, \end{aligned} \tag{6.32}$$

while $Z_t^5 \equiv z_0^5$ and $Z_t^6 \equiv 2 z_0^5$. Denoting by $\tilde{W}$ the Girsanov transform of W, we have under the equivalent martingale measure $\mathbb{Q}$

$$Z_t^2 = z_0^2 + \int_0^t \left(\Phi(s) - z_0^5 Z_s^2 \right) ds + \sum_{\lambda=1}^{d} \int_0^t \sigma^{2,\lambda}(s)\, d\tilde{W}_s^\lambda, \tag{6.33}$$

where $\Phi(t)$ and $\sigma^{2,\lambda}(t)$ are deterministic functions in t, namely

$$\Phi(t) := z_0^3 e^{-z_0^5 t} + z_0^4 e^{-2 z_0^5 t}$$

and

$$\sum_{\lambda=1}^{d} \left(\sigma^{2,\lambda}(t) \right)^2 = z_0^4 z_0^5 e^{-2 z_0^5 t}.$$

By Lévy's characterization theorem, see [**40**, Theorem (3.6), Chapter IV], the real valued process

$$W_t^* := \sum_{\lambda=1}^{d} \int_0^t \frac{\sigma^{2,\lambda}(s)}{\sqrt{z_0^4 z_0^5 e^{-z_0^5 s}}}\, d\tilde{W}_s^\lambda, \quad 0 \le t < \infty,$$

is an $(\mathcal{F}_t)$-Brownian motion under $\mathbb{Q}$. Hence the corresponding *short rates* $r_t = G_S(0, Z_t) = z_0^1 + Z_t^2$ satisfy

$$dr_t = \left(\phi(t) - z_0^5 r_t \right) dt + \tilde{\sigma}(t)\, dW_t^*,$$

where $\phi(t) := \Phi(t) + z_0^1 z_0^5$ and $\tilde{\sigma}(t) := \sqrt{z_0^4 z_0^5} e^{-z_0^5 t}$. This is just the generalized Vasicek model. It can be easily given a closed form solution for r_t, see [**38**, p. 293].

Summarizing Case ii) we have found a non-trivial e-consistent diffusion process, which is identified by (6.32) and (6.33). Actually Φ has to be replaced by a predictable process $\tilde{\Phi}$ due to the change of measure. Nevertheless this is just a one factor model. The corresponding interest rate model is the generalized Vasicek short rate model. This is very unsatisfactory since Svensson type functions $G_S(x, z)$ have six factors $z_1, \ldots, z_6$ which are observed. And it is seen that after all just one of them – that is z_2 – can be chosen to be non deterministic.

6.10. Conclusions

Bounded exponential-polynomial families like the Nelson–Siegel or Svensson family may be well suited for daily estimations of the forward rate curve. They are rather not to be used for inter-temporal interest rate modelling by diffusion processes. This is due to the facts that

- the exponents have to be kept constant
- and moreover this choice is very restricted

whenever you want to exclude arbitrage possibilities. It is shown for the Nelson–Siegel family in particular that there exists no non-trivial diffusion process providing an arbitrage free model. However there is a choice for the Svensson family, but still a very limited one, since all parameters but one have to be kept either constant or deterministic.

CHAPTER 7

Invariant Manifolds for Weak Solutions to Stochastic Equations

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.

7.1. Introduction

Consider a stochastic equation

$$\begin{cases} dX_t = (AX_t + F(t, X_t))\,dt + B(t, X_t)\,dW_t \\ X_0 = x_0 \end{cases} \tag{7.1}$$

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 (7.1) for finite dimensional regular manifolds in H. A subset $\mathcal{M}$ of H is called *(locally) invariant* for (7.1), 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 (7.1) 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*

$$\mu(t, \omega, x) := Ax + F(t, \omega, x) - \frac{1}{2}\sum_j DB^j(t, \omega, x)B^j(t, \omega, x) \in T_x\mathcal{M} \tag{7.2}$$

$$B^j(t, \omega, x) \in T_x\mathcal{M}, \quad \forall j, \tag{7.3}$$

equivalent to local invariance of $\mathcal{M}$, are derived. Here $B^j := \sqrt{\lambda_j} B e_j$, 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 (7.1) is equivalent to the set of deterministic invariance problems related to $\mu(t, \omega, \,.\,)$ and $B^j(t, \omega, \,.\,)$.

By a solution to (7.1) 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 (7.1), see [**18**]. 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 (7.1) 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 (7.1) 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 (7.1). The Hilbert space H is typically a space of functions. Suppose (7.1) 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 (7.1). That is, one observes the process $Z \in \mathcal{Z}$ which is related to X by

$$G(\cdot, Z_t) = X_t. \tag{7.4}$$

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 (7.4)? We then say that the model is consistent with $\mathcal{G}$.
ii) If so, what kind of dynamics does Z inherit from X by (7.4)? (Notice that X is a priori a weak solution to (7.1), 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 (7.2) and (7.3). 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 (locally) an Itô process. This responds to Problem ii).

Similar invariance problems have been studied in [**7**], [**10**] and [**50**]. In [**7**] and [**10**] 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 [**50**] 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 [**50**] have not yet been established for general equations (7.1). Therefore our results cannot straightly been obtained that way.

Further recent results can be found in [**3**], [**33**] for finite dimensional systems, and in [**30**], [**45**] for infinite dimensional ones. See also the references therein.

Compared to these studies we put very mild assumptions on the coefficients of (7.1), which is due to the strong structure of $\mathcal{M}$.

The paper is organized as follows. In Section 7.2 we give the exact setting for equation (7.1) and define (local) weak and strong solutions. Some classical results on stochastic equations are presented. Section 7.3 contains the main results on stochastic viability and invariance for finite dimensional C^2 submanifolds. The proofs are postponed to Sections 7.4 and 7.5, where we also recall Itô's formula for our setting. In Section 7.6 we express the consistency conditions (7.2) and (7.3) in local coordinates, making them feasible for applications. The appendix includes a detailed discussion on finite dimensional (immersed and regular) submanifolds in a Banach space. Their crucial properties are deduced.

7.2. Preliminaries on Stochastic Equations

For the stochastic background and notations we refer to [**18**]. We are given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ together with a normal filtration $(\mathcal{F}_t)_{0\le t<\infty}$. 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 (7.1) 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 [**18**, 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 (7.1) 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 (7.1) 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

$$W = \sum_j \sqrt{\lambda_j}\beta^j e_j, \tag{7.5}$$

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 [**18**, Chapt. 4].

PROPOSITION 7.1. *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

$$\|B(t,\omega,x)\|_{L_2^0}^2 = \sum_j \|B^j(t,\omega,x)\|_H^2 < \infty. \tag{7.6}$$

Then equation (7.1) may be equivalently rewritten in the form

$$\begin{cases} dX_t = (AX_t + F(t,X_t))\, dt + \sum_j B^j(t,X_t)\, d\beta_t^j \\ X_0 = x_0. \end{cases} \tag{7.1$'$}$$

Definition 7.2. *An H-valued predictable process X is a* local weak solution *to (7.1), resp. (7.1'), 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.

$$\begin{aligned} \langle \zeta, X_{t\wedge\tau}\rangle = \langle \zeta, x_0\rangle &+ \int_0^{t\wedge\tau} \Big(\langle A^\star\zeta, X_s\rangle + \langle \zeta, F(s,X_s)\rangle\Big)\, ds \\ &+ \int_0^{t\wedge\tau} \langle \zeta, B(s,X_s)\, dW_s\rangle, \end{aligned} \tag{7.7}$$

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 (7.1), resp. (7.1'), if in addition*

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

$$\mathbb{P}\Big[\int_0^{t\wedge\tau} \|AX_s\|_H \, ds < \infty\Big] = 1, \quad \forall t < \infty \tag{7.8}$$

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

$$X_{t\wedge\tau} = x_0 + \int_0^{t\wedge\tau} \Big(AX_s + F(s,X_s)\Big) ds + \int_0^{t\wedge\tau} B(s,X_s)\, dW_s, \tag{7.9}$$

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 7.3. *Let X be a weak solution to (7.1) and τ be a bounded stopping time. Then $X_{\tau+t}$ is a weak solution to*

$$\begin{cases} dY_t = (AY_t + F(\tau+t, Y_t))\,dt + B(\tau+t, Y_t)\,d\tilde{W}_t \\ Y_0 = X_\tau \end{cases} \tag{7.10}$$

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 (7.5) induces the following expansion

$$\tilde{W} = \sum_j \sqrt{\lambda_j}\tilde{\beta}^j e_j, \tag{7.11}$$

where $\tilde{\beta}^j_t := \beta^j_{\tau+t} - \beta^j_\tau$ 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 [**40**, Theorem (3.6), Chapter IV]. Let $h \in G$. Define the function $f \in C^\infty(\mathbb{R}_+ \times G; \mathbb{C})$ as follows

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

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

$$\begin{aligned} f_t(t,x) &= \frac{1}{2} f(t,x)\langle Qh, h\rangle_G \in L(\mathbb{R};\mathbb{C}) \\ f_x(t,x) &= i f(t,x)\langle h,\,.\,\rangle_G \in L(H;\mathbb{C}) \\ f_{xx}(t,x) &= -f(t,x)\langle h,\,.\,\rangle_G\langle h,\,.\,\rangle_G \in L(H\times H;\mathbb{C}) \end{aligned}$$

and Itô's formula, see [**18**, 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 $0 \le s < t < \infty$

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

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

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

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

$$\int_\tau^{\tau+t} \Phi_s\,dW_s = \int_0^t \Phi_{\tau+s}\,d\tilde{W}_s, \tag{7.12}$$

for any predictable L_2^0-valued process Φ with

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

If Φ is elementary and τ a simple stopping time, then (7.12) holds by inspection. The general case follows by a well known localization procedure, see [**18**, Lemma 4.9].

Since X is a weak solution to (7.1), for $\zeta \in D(A^\star)$ we thus have

$$\begin{aligned}\langle \zeta, X_{\tau+t}\rangle &= \langle \zeta, X_\tau\rangle + \int_\tau^{\tau+t} \Big(\langle A^\star\zeta, X_s\rangle + \langle \zeta, F(s, X_s)\rangle\Big)\,ds \\ &\quad + \int_\tau^{\tau+t} \langle \zeta, B(s, X_s)\,dW_s\rangle \\ &= \langle \zeta, X_\tau\rangle + \int_0^t \Big(\langle A^\star\zeta, X_{\tau+s}\rangle + \langle \zeta, F(\tau+s, X_{\tau+s})\rangle\Big)\,ds \\ &\quad + \int_0^t \langle \zeta, B(\tau+s, X_{\tau+s})\,d\tilde{W}_s\rangle.\end{aligned}$$

Hence it follows that $X_{\tau+t}$ is a weak solution to (7.10). □

In view of Lemma 7.3 the following definition makes sense.

Definition 7.4. *For $(t_0, x_0) \in \mathbb{R}_+ \times H$ we denote by $X^{(t_0,x_0)}$ any (local) weak solution to*

$$\begin{cases} dX_t = (AX_t + F(t_0+t, X_t))\,dt + B(t_0+t, X_t)\,dW_t^{(t_0)} \\ X_0 = x_0 \end{cases} \tag{7.13}$$

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 (7.1) will straightly apply for local weak solutions $X^{(t_0,x_0)}$ to (7.13). Hence they are stated only for the case $t_0 = 0$.

In the sequel we assume X to be a (local) weak solution to (7.1). 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

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

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

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

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

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

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

(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 (7.1).

LEMMA 7.5. *Assume* **(A4)**. *Then for any x_0 there exists a local weak solution to (7.1) satisfying* **(A1)**. *Moreover any weak solution to (7.1) satisfying* **(A1)** *is unique.*

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

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

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 [**18**, Theorems 7.4 and 6.5] there exists a unique continuous weak solution $\tilde{X}$ to (7.1), when replacing F by $\tilde{F}$ and B by $\tilde{B}$. Define the stopping time $\tau := \inf\{t \geq 0 \mid \|\tilde{X}\|_H \geq l\}$. Then $\tau > 0$ and $X_t := \tilde{X}_{t\wedge\tau}$ is a local weak solution to (7.1) with lifetime τ and satisfies **(A1)**.

If X is a continuous weak solution to (7.1) then by the above arguments it is unique on $[0, \tau_n]$ for $n \geq 2$, where $\tau_n := \inf\{t \geq 0 \mid \|X\|_H \geq n\|x_0\|_H\}$. Now use that $\tau_n \uparrow \infty$. □

7.3. Invariant Manifolds

Let $\mathcal{M}$ denote an m-dimensional C^2 submanifold of H, $m \geq 1$. 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 [**11**]. 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 7.6 and 7.7 below.

Essential for our needs is the following observation. Since H is separable, by Lindelöf's Lemma [**1**, p. 4] 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.8. Using the fact that $D(A^\star)$ is dense in H, by Proposition A.10 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

$$\phi_k(\langle\zeta_{k,1}, x\rangle, \ldots, \langle\zeta_{k,m}, x\rangle) = x, \quad \forall x \in U_k \cap \mathcal{M}. \tag{7.14}$$

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 7.6. *Let X be a local weak solution to (7.1) with lifetime τ, assume* **(A1)** *and*

$$X_{t\wedge\tau} \in \mathcal{M}, \quad \forall t < \infty. \tag{7.15}$$

Then there exists a stopping time $0 < \tau' \le \tau$, such that X is a local strong solution to (7.1) with lifetime τ'.

If in addition $\tau = \infty$ and **(A3)** *holds, then X is a strong solution to (7.1).*

We can give sufficient conditions for a weak solution X to be viable in $\mathcal{M}$. Notice that $\mu(t,\omega,x)$ in (7.2) is well defined by assuming **(A2)**.

THEOREM 7.7. *Let X be a weak solution to (7.1) 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 (7.2) and (7.3) for $dt \otimes d\mathbb{P}$-a.e. $(t,\omega) \in \mathbb{R}_+ \times \Omega$, for all $x \in \mathcal{M}$. Then (7.15) holds for $\tau = \infty$.*

Moreover A is continuous on $\mathcal{M}$. Consequently (7.2) and (7.3) 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 (7.15) is too weak to imply (7.2) and (7.3). Neither does it imply that $\mathcal{M}$ lies in $D(A)$. As a link between Theorems 7.6 and 7.7 we shall formulate equivalent conditions for (7.2) and (7.3). Rather than just asking for the assumptions of Theorem 7.6 we have to require that (7.15) holds for any space-time starting point (t_0, x_0).

DEFINITION 7.8. *The submanifold $\mathcal{M}$ is called* locally invariant *for (7.1), 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 (7.13) 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 7.9. *Assume* **(A2)**, **(A4)** *and* **(A5)**. *Then the following conditions are equivalent:*

i) *$\mathcal{M}$ is locally invariant for (7.1)*
ii) *$\mathcal{M} \subset D(A)$ and (7.2), (7.3) 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 (7.2), (7.3) 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 7.10. *If $\mathcal{M}$ is an m-dimensional linear submanifold, then Theorems 7.7 and 7.9 remain valid* without *assuming* **(A2)** *and with $\mu(t,\omega,x)$ in (7.2) replaced by*

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

7.4. Proof of Theorem 7.6

We recall Itô's formula for our setting.

LEMMA 7.11 (Itô's formula). *Assume that we are given the $\mathbb{R}^m$-valued continuous 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

$$\sum_j \|\sigma_t^j(\omega)\|_{\mathbb{R}^m}^2 < \infty, \quad \forall (t,\omega) \in \mathbb{R}_+ \times \Omega \tag{7.16}$$

$$\mathbb{P}\Big[\int_0^t \Big(\|b_s\|_{\mathbb{R}^m} + \sum_j \|\sigma_s^j\|_{\mathbb{R}^m}^2\Big) ds < \infty\Big] = 1, \quad \forall t < \infty \tag{7.17}$$

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,

$$\mathbb{P}\Big[\int_0^t \Big(\|D\phi(Y_s)b_s\|_H + \frac{1}{2}\sum_j \|D^2\phi(Y_s)(\sigma_s^j, \sigma_s^j)\|_H + \sum_j \|D\phi(Y_s)\sigma_s^j\|_H^2\Big) ds < \infty\Big] = 1, \quad \forall t < \infty, \tag{7.18}$$

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

$$(\phi \circ Y)_t = (\phi \circ Y)_0 + \int_0^t \Big(D\phi(Y_s)b_s + J_s\Big) ds + \sum_j \int_0^t D\phi(Y_s)\sigma_s^j \, d\beta_s^j. \tag{7.19}$$

Proof. Point-wise convergence of the series defining J_t follows from (7.16). Hence J_t is predictable. Integrability (7.18) is a direct consequence of (7.17). Denote the right hand side of (7.19) with I_t. Clearly by (7.18) it is a well defined continuous adapted process in H, see Proposition 7.1. Choose an orthonormal basis $\{e_i\}$ in H. It is enough to show equality (7.19) 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 [**18**, Theorem 4.17]. It follows

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

□

Proof of Theorem 7.6. Assume first that X is a weak solution, that is $\tau = \infty$, and that **(A3)** holds. Fix $T > 0$. Following [**22**, Lemma (3.5)] there exists an increasing sequence of predictable stopping times $0 = T_0 \le T_1 \le T_2 \le \cdots \le T$ 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

$$\begin{aligned} b_t &:= \langle A^\star\zeta, X_t\rangle + \langle \zeta, F(t, X_t)\rangle \\ \sigma_t^j &:= \langle \zeta, B^j(t, X_t)\rangle, \end{aligned}$$

where logically $\langle A^\star\zeta, X_t\rangle$ stands for $(\langle A^\star\zeta_1, X_t\rangle, \dots, \langle A^\star\zeta_m, X_t\rangle)$. Then b and σ^j satisfy (7.16) and (7.17), see (7.6), 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 7.11) for $\phi\circ Y$ applies. Since moreover $Y_{(T_n+t)\wedge T_{n+1}}$ takes values in V on $\{T_{n+1} > T_n\}$

$$X_{(T_n+t)\wedge T_{n+1}} = (\phi\circ Y)_{(T_n+t)\wedge T_{n+1}}, \quad \text{on } \{T_{n+1} > T_n\}. \tag{7.20}$$

Now using (7.7), (7.19) and (7.20) we have for $\xi \in D(A^\star)$

$$\begin{aligned} 0 = \int_{T_n}^{(T_n+t)\wedge T_{n+1}} &\Big(\langle A^\star\xi, X_s\rangle + \langle \xi, F(s, X_s)\rangle \\ &\quad - \langle \xi, D\phi(Y_s) b_s + \frac{1}{2}\sum_j D^2\phi(Y_s)(\sigma_s^j, \sigma_s^j)\rangle\Big) ds \\ + \sum_j \int_{T_n}^{(T_n+t)\wedge T_{n+1}} &\Big(\langle \xi, B^j(s, X_s)\rangle - \langle \xi, D\phi(Y_s)\sigma_s^j\rangle\Big) d\beta_s^j. \end{aligned} \tag{7.21}$$

Notice that the series in the first integral converges point-wise and defines an H-valued predictable process, see Lemma 7.11. By Proposition 7.1 the last term in (7.21) is a continuous local martingale with respect to the filtration $(\mathcal{F}_{T_n+t})$. Therefore

$$\langle A^\star\xi, X_t\rangle = \langle \xi, -F(t, X_t) + D\phi(Y_t) b_t + \frac{1}{2}\sum_j D^2\phi(Y_t)(\sigma_t^j, \sigma_t^j)\rangle \tag{7.22}$$

on $[T_n, T_{n+1}]\cap\{T_{n+1} > T_n\}$, $dt\otimes d\mathbb{P}$-a.s. (see the proof of [**25**, Proposition 3.2]). By separability of H there exists a countable dense subset $\mathcal{D}\subset D(A^\star)$ such that (7.22) 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 (7.22). Hence it is bounded and therefore continuous on $D(A^\star)$. Since $A^{\star\star} = A$, see [**41**, Theorem 13.12], we conclude that $X_t(\omega)\in D(A)$ and – using full notation again –

$$\begin{aligned} AX_t + F(t, X_t) = D\phi(Y_t)&\Big(\langle A^\star\zeta, X_t\rangle + \langle \zeta, F(t, X_t)\rangle\Big) \\ &+ \frac{1}{2}\sum_j D^2\phi(Y_t)(\langle \zeta, B^j(t, X_t)\rangle, \langle \zeta, B^j(t, X_t)\rangle) \end{aligned} \tag{7.23}$$

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 (7.21) with β^j)

$$B^j(t, X_t) = D\phi(Y_t)\langle \zeta, B^j(t, X_t)\rangle, \quad \forall j, \tag{7.24}$$

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

In view of **(A3)**

$$
\begin{aligned}
&\mathbb{E}\Big[\int_0^T \Big(\|D\phi(Y_s)b_s\|_H + \frac{1}{2}\sum_j \|D^2\phi(Y_s)(\sigma_s^j,\sigma_s^j)\|_H + \sum_j \|D\phi(Y_s)\sigma_s^j\|_H^2\Big)ds\Big] \\
&\le C_1\mathbb{E}\Big[\int_0^T \Big(\|b_s\|_{\mathbb{R}^m} + \sum_j \|\sigma_s^j\|_{\mathbb{R}^m}^2\Big)ds\Big] \\
&\le C_2\mathbb{E}\Big[\int_0^T \Big(\|X_s\|_H + \|F(s,X_s)\|_H + \|B(s,X_s)\|_{L_2^0}^2\Big)ds\Big] < \infty,
\end{aligned}
\tag{7.25}
$$

where C_1 and C_2 depend only on n.

Summing up over $0 \le n \le N-1$ we get from (7.23) and (7.25)

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

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 (7.23) and (7.24)

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

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 (7.14). 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 (7.25) has to be replaced. We give a sketch. Define

$$
\begin{aligned}
b_t &:= \Big(\langle A^\star\zeta, X_{t\wedge\tau'}\rangle + \langle \zeta, F(t, X_{t\wedge\tau'})\rangle\Big)1_{[0,\tau']}(t) \\
\sigma_t^j &:= \langle \zeta, B^j(t, X_{t\wedge\tau'})\rangle 1_{[0,\tau']}(t).
\end{aligned}
$$

Then

$$
X_{t\wedge\tau'} = (\phi\circ Y)_t, \quad \forall t < \infty \tag{7.26}
$$

where

$$
Y_t := \langle \zeta, X_{t\wedge\tau'}\rangle = \langle \zeta, x_0\rangle + \int_0^t b_s\, ds + \sum_j \int_0^t \sigma_s^j\, d\beta_s^j. \tag{7.27}
$$

Instead of (7.25) we now have (7.17) by the very definition of X. Using (7.26) and Lemma 7.11 we get the relations (7.23) and (7.24) 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 (7.18) and (7.23) we derive (7.8) and (7.9). $\square$

7.5. Proof of Theorems 7.7, 7.9 and 7.10

A key step in proving Theorems 7.7, 7.9 and 7.10 consists in the following property.

LEMMA 7.12. *Assume* **(A2)** *and* **(A4)**. *Let $\phi : V \to U \cap \mathcal{M}$ be a parametrization satisfying (7.14). Suppose that $U \cap \mathcal{M} \subset D(A)$. Then (7.2) and (7.3) 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*

$$\begin{aligned} Ax + F(t,\omega,x) &= D\phi(y)\Big(\langle A^\star \zeta, x\rangle + \langle \zeta, F(t,\omega,x)\rangle\Big) \\ &\quad + \frac{1}{2}\sum_j D^2\phi(y)(\langle \zeta, B^j(t,\omega,x)\rangle, \langle \zeta, B^j(t,\omega,x)\rangle) \end{aligned} \tag{7.28}$$

$$B^j(t,\omega,x) = D\phi(y)\langle \zeta, B^j(t,\omega,x)\rangle, \quad \forall j, \tag{7.29}$$

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 (7.3) and (7.29), and of (7.2) and the relation

$$\begin{aligned} Ax + F(t,\omega,x) &- \frac{1}{2}\sum_j DB^j(t,\omega,x)B^j(t,\omega,x) \\ &= D\phi(y)\langle \zeta, Ax + F(t,\omega,x) - \frac{1}{2}\sum_j DB^j(t,\omega,x)B^j(t,\omega,x)\rangle \end{aligned} \tag{7.30}$$

follows from Lemma A.11. Assume (7.2) and (7.3) – and hence (7.29) and (7.30) – 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 (7.29) 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 (7.29) 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.12 applies and the last summand in (7.28) 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 (7.28) holds for $(t,\omega,x) \in \mathcal{N}^c \times (U \cap \mathcal{M})$ if and only if (7.30) does so.

It remains to show validity of (7.28) 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 (7.28) to $R(t,\omega,x)$. Then it reads

$$Ax = -F(t,\omega,x) + R(t,\omega,x) \tag{7.31}$$

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 (7.31) is continuous in x, see (7.6). 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. □

PROOF OF THEOREM 7.7. 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 (7.14) 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

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

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 7.3 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 (7.11). Hence the stochastic differential equation in $\mathbb{R}^m$

$$Y_t = Y_0 + \int_0^t b(s,Y_s)\,ds + \sum_j \int_0^t \sigma^j(s,Y_s)\,d\tilde{\beta}_s^j \tag{7.32}$$

has a unique continuous strong solution Y, see [**18**, 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

$$\tau_2 := \begin{cases} \inf\{t \geq 0 \mid Y_t \notin V_1\}, & \text{if } Y_0 \in \overline{V}_0 \text{ and } \tau_0 < K \\ 0, & \text{otherwise.} \end{cases} \tag{7.33}$$

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

$$Y \in V \quad \text{on } [0,\tau_2]\cap\{\tau_2 > 0\}. \tag{7.34}$$

From the boundedness of b and σ we derive

$$\mathbb{E}\Big[\int_0^t \Big(\|b(s,Y_s)\|_{\mathbb{R}^m} + \sum_j \|\sigma^j(s,Y_s)\|_{\mathbb{R}^m}^2\Big)ds\Big] < \infty, \quad \forall t < \infty. \tag{7.35}$$

Set $\tilde{X} := \phi \circ Y$. The assumptions of Lemma 7.11 are satisfied, hence

$$\begin{aligned} \tilde{X}_t = X_{\tau_1} &+ \int_0^t \Big(D\phi(Y_s)b(s,Y_s) + \frac{1}{2}\sum_j D^2\phi(Y_s)(\sigma^j(s,Y_s),\sigma^j(s,Y_s))\Big)ds \\ &+ \sum_j \int_0^t D\phi(Y_s)\sigma^j(s,Y_s)\,d\tilde{\beta}_s^j, \quad \text{on } [0,\tau_2]\cap\{\tau_2>0\}, \end{aligned} \tag{7.36}$$

where the series in the first integral converges point-wise and defines an H-valued predictable process. Moreover by (7.34) we have

$$\tilde{X} \in (U \cap \mathcal{M}) \subset D(A) \quad \text{on } [0,\tau_2]\cap\{\tau_2 > 0\}. \tag{7.37}$$

Hence due to **(A2)** and **(A4)** Lemma 7.12 applies and

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

on $[0,\tau_2]\cap\{\tau_2>0\}$, $dt\otimes d\mathbb{P}$-a.s. This together with (7.35) and (7.36) implies

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

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

on $\{\tau_2>0\}$. On the other hand we know by Lemmas 7.3 and 7.5 that X_{τ_1+t} is the unique continuous weak solution to

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

Whence $\tilde{X}_t=X_{\tau_1+t}$ on $[0,\tau_2]\cap\{\tau_2>0\}$. Because of (7.37) 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

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

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

The last statement of the theorem follows from Lemma 7.12 and **(A2)**. □

PROOF OF THEOREM 7.9. i)⇒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 (7.13) with lifetime τ. We proceed as in the proof of Theorem 7.6 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 (7.22) we derive the equality

$$\begin{aligned}\langle A^\star\xi,X_t\rangle=\Big\langle\xi,-F(t_0+t,X_t)+D\phi(Y_t)\Big(\langle A^\star\zeta,X_t\rangle+\langle\zeta,F(t_0+t,X_t)\rangle\Big)\\ +\frac{1}{2}\sum_j D^2\phi(Y_t)(\langle\zeta,B^j(t_0+t,X_t)\rangle,\langle\zeta,B^j(t_0+t,X_t)\rangle)\Big\rangle,\end{aligned}\tag{7.38}$$

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 (7.38) are right continuous in t. Hence for $\mathbb{P}$-a.e. ω the limit $t\downarrow 0$ exists. By (7.6) we can interchange the limiting $t\downarrow 0$ with the summation over j. Arguing as for (7.23) we conclude that $x_0\in D(A)$ and

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

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 7.12 yields condition ii).

ii)⇒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 (7.14). As in the proof of Theorem 7.7 (setting $\tau_0=\tau_1=0$) we get a unique continuous strong solution Y to

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

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 (7.11). The stopping time τ_2 given by (7.33) is strictly positive and $Y_{t\wedge\tau_2} \in V$. Analysis similar to that in the proof of Theorem 7.7 shows that $X_t^{(t_0,x_0)} := (\phi \circ Y)_{t\wedge\tau_2}$ is a continuous local strong solution to (7.13) with lifetime τ_2 and $X_t^{(t_0,x_0)} \in U \cap \mathcal{M}$, for all $t < \infty$.

ii)⇔iii): This is a direct consequence of Lemma 7.12 and **(A5)**. □

Proof of Theorem 7.10. In view of Proposition A.10 we can choose any parametrization ϕ_k in (7.14) such that $D^2\phi_k \equiv 0$ on V_k. By straightforward inspection of the proofs we see that Lemma 7.12 and Theorems 7.7 and 7.9 remain valid under the conditions stated in the theorem. □

7.6. Consistency Conditions in Local Coordinates

In this section we express the consistency conditions (7.2) and (7.3) in local coordinates and identify the underlying coordinate process Y. From now on the assumptions of Theorem 7.9 are in force and $\mathcal{M}$ is locally invariant for (7.1).

Let $\phi : V \to U \cap \mathcal{M}$ be a parametrization. We assume that (A.1) and (A.2) hold for all $y \in V$ (otherwise replace V by $V \cap V'$). Consequently

$$\|D\phi(y)^{-1}\|_{L(T_{\phi(y)}\mathcal{M};\mathbb{R}^m)} \leq \|D\Psi(\phi(y))\|, \quad \forall y \in V \tag{7.39}$$

and the latter is locally bounded since Ψ is C^2. Theorem 7.9 says that

$$\mu(t, \omega, \phi(y)) = D\phi(y)\tilde{b}(t, \omega, y) \tag{7.40}$$

$$B^j(t, \omega, \phi(y)) = D\phi(y)\rho^j(t, \omega, y), \quad \forall j, \tag{7.41}$$

for all $(t, \omega, y) \in \mathbb{R}_+ \times \Omega \times V$ (after removing a $\mathbb{P}$-nullset from Ω), where $\tilde{b}$ and ρ^j are uniquely specified mappings measurable from $(\mathbb{R}_+ \times \Omega \times V, \mathcal{P} \otimes \mathcal{B}(V))$ into $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$, see (A.3). By (7.39)

$$\rho(t, \omega, y) := D\phi(y)^{-1}B(t, \omega, \phi(y))(\cdot) \in L_2(G_0; \mathbb{R}^m).$$

To simplify notation, we suppress the argument ω whenever there is no ambiguity. In view of **(A2)** and (A.2), $\rho^j(t, \cdot) \in C^1(V; \mathbb{R}^m)$.

Fix $(t, y) \in V$. It is shown at the end of the appendix that

$$DB^j(t, \phi(y))B^j(t, \phi(y)) = \frac{d}{ds} B^j(t, c(s))_{|s=0}$$

where $c(s) := \phi(y + s\rho^j(t, y))$. On the other hand, by (7.41)

$$\begin{aligned} \frac{d}{ds} B^j(t, c(s))_{|s=0} &= \frac{d}{ds} \big(D\phi(y + s\rho^j(t, y))\rho^j(t, y + s\rho^j(t, y)) \big)_{|s=0} \\ &= D^2\phi(y)\big(\rho^j(t, y), \rho^j(t, y)\big) + D\phi(y)\big(D\rho^j(t, y)\rho^j(t, y)\big), \end{aligned}$$

compare with (A.7). Inserting this in (7.40) yields

$$A\phi(y) + F(t, \phi(y)) - \frac{1}{2}\sum_j D^2\phi(y)\big(\rho^j(t,y), \rho^j(t,y)\big) = D\phi(y)b(t,y), \tag{7.42}$$

where $b(t,\omega,y) := \tilde{b}(t,\omega,y) + \frac{1}{2}\sum_j D\rho^j(t,\omega,y)\rho^j(t,\omega,y)$. By (7.6), (7.39) and **(A2)**, the series converge absolutely and are continuous in y.

We now show how equation (7.13) can be written in local coordinates. We claim that $b(t,\omega,y) \in \mathbb{R}^m$ and $\rho(t,\omega,y)$ are locally Lipschitz in y on $[0,T]\times\Omega\times V$, for all $T<\infty$. Indeed, the right hand side of (7.28) is locally Lipschitz in y by **(A4)** and hence also the left hand side of (7.42), see (7.6) and (7.39). Solving (7.42) for b yields the claim.

Consequently, for any $(t_0, y_0) \in \mathbb{R}_+ \times V$ there exists by [**18**, Theorem 7.4] a unique continuous local strong solution $Y^{(t_0,y_0)} \in V$ to

$$\begin{cases} dY_t = b(t_0+t, Y_t)\,dt + \sum_j \rho^j(t_0+t, Y_t)\,d\beta_t^{(t_0),j} \\ Y_0 = y_0 \end{cases} \tag{7.43}$$

with lifetime $\tau(t_0,y_0)$, where $\beta_t^{(t_0),j} := \beta^j_{t_0+t} - \beta^j_{t_0}$, see (7.11). Using Itô's formula (Lemma 7.11) and proceeding as in the proof of Theorem 7.7, we see that

$$X_t^{(t_0,x_0)} := (\phi \circ Y^{(t_0,y_0)})_{t\wedge\tau(t_0,y_0)}$$

is the unique continuous local strong solution to (7.13) with lifetime $\tau(t_0,y_0)$, where $x_0 = \phi(y_0)$.

Summarizing, we have proved

THEOREM 7.13. *Under the above assumptions, the consistency conditions (7.2) and (7.3) hold for all $(t,\omega,x) \in \mathbb{R}_+ \times \Omega \times (U \cap \mathcal{M})$ if and only if (7.42) and (7.41) hold for all $(t,\omega,y) \in \mathbb{R}_+ \times \Omega \times V$.*

The coefficients b and ρ are locally Lipschitz continuous in $y \in V$: For any $T, R \in \mathbb{R}_+$ there exists a real constant $C = C(T,R)$ such that

$$\|b(t,\omega,y) - b(t,\omega,z)\|_{\mathbb{R}^m} + \|\rho(t,\omega,y) - \rho(t,\omega,z)\|_{L_2(G_0;\mathbb{R}^m)} \le C\|y-z\|_{\mathbb{R}^m} \tag{7.44}$$

for all $(t,\omega) \in [0,T]\times\Omega$ and $y,z \in V \cap B_R(\mathbb{R}^m)$.

Moreover equation (7.13) transforms locally into equation (7.43). That is, $X^{(t_0,x_0)}$ is a continuous local strong solution to (7.13) in $U \cap \mathcal{M}$ if and only if so is $\phi^{-1} \circ X^{(t_0,x_0)}$ to (7.43) in V, where $y_0 = \phi^{-1}(x_0)$.

If ϕ is linear ($D^2\phi \equiv 0$), then the same holds true without assuming **(A2)**.

Appendix: Finite Dimensional Submanifolds in Banach Spaces

For the convenience of the reader we discuss and prove some crucial properties of finite dimensional submanifolds in Banach spaces. The main tool from infinite dimensional analysis is the inverse mapping theorem, see [**1**, Theorem 2.5.2], which does not rely on a scalar product. Therefore we chose this somewhat oversized framework, although in our setting stochastic equations take place in Hilbert spaces.

Let E denote a Banach space, E' it's dual space. We write $\langle e', e\rangle$ for the duality pairing. 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.

Let $k, m \in \mathbb{N}$. We begin with an important corollary of the inverse mapping theorem, see [**1**, Theorem 2.5.12].

Proposition A.1. *Let $\phi \in C^k(V;E)$, for some open set $V \subset \mathbb{R}^m$. Suppose $D\phi(y_0)$ is one to one for some $y_0 \in V$. Then $D\phi(y_0)\mathbb{R}^m$ is m-dimensional and complemented in E*

$$E = D\phi(y_0)\mathbb{R}^m \oplus E_2.$$

Moreover, there exist two open neighborhoods V' of $(y_0, 0)$ in $V \times E_2$ and U of $\phi(y_0)$ in E, and a C^k diffeomorphism $\Psi : U \to V'$ such that

$$\Psi \circ \phi(y) = (y, 0), \quad \forall y \in V' \cap (\mathbb{R}^m \times \{0\}). \tag{A.1}$$

Furthermore, $D\phi(y)$ is one to one and

$$D\phi(y)^{-1} = D\Psi(\phi(y))|_{D\phi(y)\mathbb{R}^m}, \quad \forall y \in V' \cap (\mathbb{R}^m \times \{0\}). \tag{A.2}$$

Proof. Let $\{e_1, \dots, e_m\}$ be a basis of $D\phi(y_0)\mathbb{R}^m$. Every $z \in D\phi(y_0)\mathbb{R}^m$ has a unique representation $z = z_1e_1 + \dots z_me_m$. Set $\alpha_i(z) := z_i$. Each α_i extends to a member e_i' of E' by the Hahn–Banach theorem. Define E_2 as the intersection of the null spaces of all e_i'. Then $E = D\phi(y_0)\mathbb{R}^m \oplus E_2$ and $\Pi_{(E_2, D\phi(y_0)\mathbb{R}^m)} = \langle e_1', \cdot\rangle e_1 + \dots \langle e_m', \cdot\rangle e_m$.

Define $\Phi : V \times E_2 \to E$ by $\Phi(y, z) := \phi(y) + z$. Then $\Phi \in C^k(V \times E_2; E)$ and $D\Phi(y,z)(v_1, v_2) = D\phi(y)v_1 + v_2$ for $(v_1, v_2) \in \mathbb{R}^m \times E_2$, see [**1**, Prop. 2.4.12]. Now $D\Phi(y_0, 0) : \mathbb{R}^m \times E_2 \to E$ is linear, bijective and continuous, hence a linear isomorphism by the open mapping theorem. Accordingly, the inverse mapping theorem yields the existence of U and V' as claimed such that $\Phi : V' \to U$ is a C^k diffeomorphism with inverse Ψ.

Consequently, $\Psi \circ \phi(y) = \Psi \circ \Phi(y, 0) = (y, 0)$ for $y \in V' \cap (\mathbb{R}^m \times \{0\})$. The chain rule gives $D\Psi(\phi(y))D\phi(y) = Id_{\mathbb{R}^m}$. Hence $D\phi(y)$ is one to one and (A.2) holds. $\square$

Definition A.2. *The mapping ϕ from Proposition A.1 is called a C^k* immersion at y_0. *If ϕ is a C^k immersion at each $y_0 \in V$, we just say ϕ is a C^k* immersion.

*If $\phi : V \to E$ is an injective C^k immersion, $V \subset \mathbb{R}^m$ open, we call $\mathcal{M} := \phi(V)$ an m-*dimensional immersed C^k submanifold *of E.*

The next definition straightly extends the concept of a regular surface in $\mathbb{R}^3$.

Definition A.3. *A subset $\mathcal{M} \subset E$ is an m-*dimensional (regular) 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.*

Let $\phi : V \to E$ be an injective C^k immersion, $V \subset \mathbb{R}^m$ open, whence $\mathcal{M} = \phi(V)$ an immersed submanifold. In the notation of Proposition A.1 we have by (A.1) that $\phi : V' \cap (\mathbb{R}^m \times \{0\}) \to \phi(V' \cap (\mathbb{R}^m \times \{0\}))$ is a homeomorphism, and hence $\phi(V' \cap (\mathbb{R}^m \times \{0\}))$ is a regular submanifold. Yet in general $\phi : V \to \mathcal{M}$ is *not* a homeomorphism. In other words, being an immersed submanifold does not imply being a regular submanifold. This is the crucial difference between an immersion and an *embedding*. For examples and more details see [**1**, Section 3.5].

In what follows, $\mathcal{M}$ denotes an m-dimensional C^k submanifold of E. As an immediate consequence of Proposition A.1 and Definition A.3 we show that $\mathcal{M}$ shares in fact the characterizing property of a C^k manifold:

Lemma A.4. *Let $\phi_i : V_i \to U_i \cap \mathcal{M}$, $i = 1, 2$, be two parametrizations such that $W := U_1 \cap U_2 \cap \mathcal{M} \neq \emptyset$. Then the change of parameters*

$$\phi_1^{-1} \circ \phi_2 : \phi_2^{-1}(W) \to \phi_1^{-1}(W)$$

is a C^k diffeomorphism.

Proof. Of course, $\phi_1^{-1} \circ \phi_2 : \phi_2^{-1}(W) \to \phi_1^{-1}(W)$ is a homeomorphism.

Now let $y \in \phi_2^{-1}(W)$ and $r = \phi_1^{-1} \circ \phi_2(y)$. Then $E = D\phi_1(r)\mathbb{R}^m \oplus E_2$ by Proposition A.1. Moreover there exist two open neighborhoods $V_1' \subset V_1 \times E_2$ of $(r, 0)$ and $U' \subset E$ of $\phi_1(r)$, and a C^k diffeomorphism $\Psi : U' \to V_1'$ such that $\phi_1 = \Psi^{-1}$ on $V_1' \cap (\mathbb{R}^m \times \{0\})$.

Now again we use the fact that ϕ_1 is a homeomorphism. Hence there exists an open set $U'' \subset E$ with $\phi_1(V_1' \cap (\mathbb{R}^m \times \{0\})) = U'' \cap \mathcal{M}$. Set $U := U' \cap U''$. Then $\phi_1^{-1} = \Psi$ on $U \cap \mathcal{M}$. By continuity of ϕ_2 there is an open neighborhood $V_2' \subset \phi_2^{-1}(W)$ of y with $\phi_2(V_2') \subset U \cap \mathcal{M}$. But then $\phi_1^{-1} \circ \phi_{2|V_2'} = \Psi \circ \phi_{2|V_2'}$ is C^k in V_2'.

Since y was arbitrary, $\phi_1^{-1} \circ \phi_2$ is C^k in $\phi_2^{-1}(W)$. By symmetry the assertion follows. $\square$

Definition A.5. *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.

By Lemma A.4, the definition of $T_x\mathcal{M}$ is independent of the choice of the parametrization.

DEFINITION A.6. *A* vector field X *on* $\mathcal{M}$ *is a mapping assigning to each point* x *of* $\mathcal{M}$ *an element* $X(x)$ *of* $T_x\mathcal{M}$.

A vector field X can be represented locally as

$$X(x) = D\phi(y)\alpha(y), \quad y = \phi^{-1}(x), \quad \forall x \in U \cap \mathcal{M}, \tag{A.3}$$

where $\phi : V \to U \cap \mathcal{M}$ is a parametrization and α is an $\mathbb{R}^m$-valued vector field on V (uniquely determined by ϕ). This is how smoothness of vector fields on $\mathcal{M}$ can be defined in terms of smoothness of $\mathbb{R}^m$-valued vector fields.

DEFINITION A.7. *The vector field* X *is* of class C^r, $0 \le r < k$, *if for any parametrization* ϕ *the corresponding* $\mathbb{R}^m$*-valued vector field* α *in (A.3) is of class* C^r.

Again by Lemma A.4 this is a well defined concept.

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 [**11**, Theorem (5.1), Chapt. II]. Since ϕ is a homeomorphism there exists an open 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.8. *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 $x \in \mathcal{M}$. As shown in Proposition A.1, the tangent space $T_x\mathcal{M}$ is complemented in E by $E_2 = \cap_{i=1}^m \mathcal{N}(e_i')$, where $\{e_1, \dots, e_m\}$ is a basis for $T_x\mathcal{M}$ and $\langle e_i', e_j \rangle = \delta_{ij}$ (Kronecker delta). The following lemma states that, locally in x, $\Pi_{(E_2, T_x\mathcal{M})}$ is a diffeomorphism on $\mathcal{M}$ into $T_x\mathcal{M}$.

LEMMA A.9. *There exists a parametrization* $\phi : V \to U \cap \mathcal{M}$ *in* x *such that*

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

PROOF. Let $\psi : V_1 \to U_1 \cap \mathcal{M}$ be a parametrization in x and write $y = \psi^{-1}(x)$. Denote by $I_{T_x\mathcal{M}} : T_x\mathcal{M} \to \mathbb{R}^m$ the canonical isomorphism $I_{T_x\mathcal{M}} := (\langle e_1', \cdot\rangle, \dots, \langle e_m', \cdot\rangle)$. Then the map $h := I_{T_x\mathcal{M}} \circ \Pi_{(E_2, T_x\mathcal{M})} \circ \psi : V_1 \to \mathbb{R}^m$ is C^k and $Dh(y) = I_{T_x\mathcal{M}} \circ D\psi(y)$ is one to one. By the inverse mapping theorem there exist open 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 an open 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, $\phi^{-1} = I_{T_x\mathcal{M}} \circ \Pi_{(E_2, T_x\mathcal{M})}$ on $U \cap \mathcal{M}$. □

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

Proposition A.10. *Let $D \subset E'$ be a dense subset. Then for any $x \in \mathcal{M}$ there exist elements $f_1', \ldots, f_m'$ in D and a parametrization $\phi : V \to U \cap \mathcal{M}$ in x such that*

$$\phi(\langle f_1', z\rangle, \ldots, \langle f_m', z\rangle) = z, \quad \forall z \in U \cap \mathcal{M}.$$

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 $E = F_1 \oplus F_2$, $\dim F_1 = m$, such that F_1 is "not too far" from $T_x\mathcal{M}$ and such that $\Pi_{(F_2,F_1)} = \langle f_1', \cdot\rangle f_1 + \cdots + \langle f_m', \cdot\rangle f_m$, with $f_1' \ldots, f_m' \in D$. Thereby the expression "not too far" means that $\Pi_{(F_2,F_1)|T_x\mathcal{M}} : T_x\mathcal{M} \to F_1$ is an isomorphism.

Let $\{e_1, \ldots, e_m\}$ be a basis for $T_x\mathcal{M}$ and $\psi : V_1 \to U_1 \cap \mathcal{M}$ the parametrization in x given by Lemma A.9. Set $y = \psi^{-1}(x)$. We may assume that $\|e_i\|_E = 1$, for $1 \leq i \leq m$. Since D is dense in E', there exist elements $f_1', \ldots, f_m'$ in D with $\|f_i' - e_i'\|_{E'} < 2^{-m}$. Notice that by definition $\langle e_i', e_j\rangle = \delta_{ij}$. Hence

$$|\langle f_i', e_j\rangle| \leq |\langle e_i', e_j\rangle| + |\langle f_i' - e_i', e_j\rangle| < 2^{-m}, \quad \text{if } i \neq j$$

and

$$|\langle f_i', e_i\rangle| \geq |\langle e_i', e_i\rangle| - |\langle f_i' - e_i', e_i\rangle| > 1 - 2^{-m}.$$

The matrix $M := (\langle f_i', e_j\rangle)_{1 \leq i,j \leq m}$ is therefore invertible, which follows from the theorem of Gerschgorin, see [**43**]. Consequently, the family $\{f_1', \ldots, f_m'\}$ is linearly independent in E'. Since E is reflexive, by the Hahn–Banach theorem there exists a linearly independent set $\{f_1, \ldots, f_m\}$ in E such that $\langle f_i', f_j\rangle = \delta_{ij}$. Write $F_1 := \operatorname{span}\{f_1, \ldots, f_m\}$. The f_i' and f_i induce then a decomposition $E = F_1 \oplus F_2$ with projection $\Pi_{(F_2,F_1)} = \langle f_1', \cdot\rangle f_1 + \cdots + \langle f_m', \cdot\rangle f_m$. Since

$$\Pi_{(F_2,F_1)} e_j = \sum_{i=1}^m M_{ij} f_i, \quad \text{for } 1 \leq j \leq m,$$

we see that $\Pi_{(F_2,F_1)|T_x\mathcal{M}} : T_x\mathcal{M} \to F_1$ is an isomorphism. Let $I_{F_1} : F_1 \to \mathbb{R}^m$ denote the canonical isomorphism $I_{F_1} := (\langle f_1', \cdot\rangle, \ldots, \langle f_m', \cdot\rangle)$. Then the map $h := I_{F_1} \circ \Pi_{(F_2,F_1)} \circ \psi : V_1 \to \mathbb{R}^m$ is C^k and $Dh(y) = I_{F_1} \circ \Pi_{(F_2,F_1)} \circ D\psi(y)$ is one to one. Now proceed as in the proof of Lemma A.9 to get the desired parametrization ϕ in x with $\phi^{-1} = I_{F_1} \circ \Pi_{(F_2,F_1)}$ on $U \cap \mathcal{M}$, for some neighborhood U of x.

If $\mathcal{M}$ is linear, we choose $\psi(u) = x + \sum_{i=1}^m u_i e_i$. An easy computation shows that $\phi(v) = (x - (\Pi_{(F_2,F_1)|T_x\mathcal{M}})^{-1} \circ \Pi_{(F_2,F_1)} x) + \sum_{i=1}^m (\sum_{j=1}^m M_{ij}^{-1} v_j) e_i$. Setting $N := M^{-1}$ and $e_0 := x - (\Pi_{(F_2,F_1)|T_x\mathcal{M}})^{-1} \circ \Pi_{(F_2,F_1)} x$ completes the proof. ☐

Lemma A.9 asserts the existence of a particular parametrization $\phi : V \to U \cap \mathcal{M}$ in $x \in \mathcal{M}$ related to the decomposition $E = T_x\mathcal{M} \oplus E_2$. Actually, E_2 is complemented in E simultaneously by all tangent spaces $T_z\mathcal{M}$, $z \in U \cap \mathcal{M}$.

Lemma A.11. *Let $\phi : V \to U \cap \mathcal{M}$ be a parametrization. If there exist elements $e_1', \ldots, e_m'$ in E' with the property that*

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

then $e_1', \ldots, e_m'$ are linearly independent in E' and

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

where $E_2 := \bigcap_{i=1}^m \mathcal{N}(e_i')$. Moreover, the induced projections are given by

$$\Pi_{(E_2,T_x\mathcal{M})} = D\phi(y)(\langle e_1', \cdot\rangle, \ldots, \langle e_m', \cdot\rangle), \quad y = \phi^{-1}(x), \quad \forall x \in U \cap \mathcal{M}. \tag{A.4}$$

PROOF. Define $\Lambda := (\langle e_1', \cdot\rangle, \dots, \langle e_m', \cdot\rangle) \in L(E;\mathbb{R}^m)$. 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, $e_1', \dots, e_m'$ 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

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

i.e. $B_{|\mathcal{M}}$ is a C^1 vector field on $\mathcal{M}$.

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

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

is well defined and satisfies

$$c : (-\delta, \delta) \to U \cap \mathcal{M}, \quad c \in C^1((-\delta,\delta); E), \quad c(0) = x, \quad c'(0) = B(x). \tag{A.5}$$

Hence

$$\frac{d}{dt} B(c(t))_{|t=0} = DB(x)B(x). \tag{A.6}$$

In terms of differential geometry $B(x)$ can be identified with the equivalence class $[c]$ of all curves satisfying (A.5). By the very definition of $DB(x)$ as a mapping from $T_x\mathcal{M}$ into $T_{B(x)}E \equiv E$, the vector $DB(x)B(x)$ is in the same manner identified with $[B \circ c]$, see [**1**, Section 3.3]. This is just (A.6) and means that the differentiable structure on $\mathcal{M}$ is the one induced by E.

Suppose now that ϕ satisfies the assumptions of Lemma A.11. To shorten notation, we write $\langle e', \cdot\rangle$ instead of $(\langle e_1', \cdot\rangle, \dots, \langle e_m', \cdot\rangle)$. By (A.4)

$$B(c(t)) = D\phi(\langle e', c(t)\rangle)\langle e', B(c(t))\rangle, \quad \forall t \in (-\delta, \delta).$$

Differentiation with respect to t gives

$$\frac{d}{dt} B(c(t))_{|t=0} = D^2\phi(y)(\langle e', B(x)\rangle, \langle e', B(x)\rangle) + D\phi(y)\langle e', DB(x)B(x)\rangle. \tag{A.7}$$

Combining (A.4), (A.6) and (A.7) we get the following proposition.

PROPOSITION A.12. *Let $\mathcal{M}$, B and $\phi : V \to U \cap \mathcal{M}$ be as above. Then*

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

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

Bibliography

1. R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, tensor analysis, and applications*, Springer-Verlag, Berlin-Heidelberg-New York, 1988.
2. H. Amann, *Ordinary differential equations*, Walter de Gruyter & Co., Berlin, 1990, An introduction to nonlinear analysis, Translated from the German by Gerhard Metzen.
3. 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.
4. H. Bauer, *Maß- und Integrationstheorie*, second ed., Walter de Gruyter & Co., Berlin, 1992.
5. B. M. Bibby and M. Sørensen, *On estimation for discretely observed diffusions: A review*, Research Report No. 334, Department of Theoretical Statistics, Institute of Mathematics, University of Aarhus, 1995.
6. BIS, *Zero-coupon yield curves: Technical documentation*, Bank for International Settlements, Basle, March 1999.
7. T. Björk and B. J. Christensen, *Interest rate dynamics and consistent forward rate curves*, Math. Finance **9** (1999), 323–348.
8. T. Björk and A. Gombani, *Minimal realization of forward rates*, Finance and Stochastics **3** (1999), no. 4, 413–432.
9. T. Björk, Y. Kabanov, and W. Runggaldier, *Bond market structure in the presence of marked point processes*, Math. Finance **7** (1997), no. 2, 211–239.
10. 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.
11. W. M. Boothby, *An introduction to differentiable manifolds and Riemannian geometry*, second ed., Academic Press, 1986.
12. A. Brace, D. Gątarek, and M. Musiela, *The market model of interest rate dynamics*, Math. Finance **7** (1997), no. 2, 127–155.
13. A. Brace and M. Musiela, *A multifactor Gauss Markov implementation of Heath, Jarrow, and Morton*, Math. Finance **4** (1994), 259–283.
14. H. Brézis, *Analyse fonctionnelle*, fourth ed., Masson, Paris, 1993, Théorie et applications. [Theory and applications].
15. A. Chojnowska-Michalik, *Stochastic differential equations in Hilbert spaces*, Probability theory (Papers, VIIth Semester, Stefan Banach Internat. Math. Center, Warsaw, 1976), PWN, Warsaw, 1979, pp. 53–74.
16. R. Cont, *Modeling term structure dynamics, an infinite dimensional approach*, Working paper, Ecole Polytechnique Paris, 1998.
17. J. Cox, J. Ingersoll, and S. Ross, *A theory of the term structure of interest rates*, Econometrica **53** (1985), 385–408.
18. G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, Cambridge University Press, 1992.
19. F. Delbaen and W. Schachermayer, *A general version of the fundamental theorem of asset pricing*, Math. Ann. **300** (1994), 463–520.
20. ———, *The no-arbitrage property under a change of numéraire*, Stochastics Stochastics Rep. **53** (1995), no. 3-4, 213–226.
21. D. Duffie and R. Kan, *A yield-factor model of interest rates*, Math. Finance **6** (1996), 379–406.
22. M. Emery, *Stochastic calculus in manifolds*, Springer-Verlag, Berlin-Heidelberg-New York, 1989.
23. D. Filipović, *Exponential-polynomial families and the term structure of interest rates*, to appear in BERNOULLI, Chapter 6 in this thesis.

24. ———, *Invariant manifolds for weak solutions to stochastic equations*, to appear in Probab. Theor. Relat. Fields, Chapter 7 in this thesis.
25. ———, *A note on the Nelson–Siegel family*, Math. Finance **9** (1999), 349–359, Chapter 5 in this thesis.
26. ———, *A general characterization of affine term structure models*, Working paper, ETH Zürich, 2000.
27. B. Goldys and M. Musiela, *Infinite dimensional diffusions, Kolmogorov equations and interest rate models*, Tech. report, University of New South Wales, 1998.
28. 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.
29. N. Ikeda and S. Watanabe, *Stochastic differential equations and diffusion processes*, Noth-Holland, Amsterdam, 1981.
30. W. Jachimiak, *Stochastic invariance in infinite dimension*, Working paper, Polish Academy of Sciences, 1998.
31. J. Jacod and A. N. Shiryaev, *Limit theorems for stochastic processes*, Grundlehren der mathematischen Wissenschaften, vol. 288, Springer-Verlag, Berlin-Heidelberg-New York, 1987.
32. I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*, second ed., Springer-Verlag, New York, 1991.
33. A. Milian, *Invariance for stochastic equations with regular coefficients*, Stochastic Anal. Appl. **15** (1997), no. 1, 91–101.
34. ———, *Comparison theorems for stochastic evolution equations*, Working paper, University of Technology Krakow, 1999.
35. K. R. Miltersen, *An arbitrage theory of the term structure of interest rates*, Ann. Appl. Probab. **4** (1994), no. 4, 953–967.
36. A. J. Morton, *Arbitrage and martingales*, Ph.D. thesis, Cornell University, 1989.
37. M. Musiela, *Stochastic PDEs and term structure models*, Journées Internationales de Finance, IGR-AFFI, La Baule, 1993.
38. M. Musiela and M. Rutkowski, *Martingale methods in financial modelling*, Applications of Mathematics, vol. 36, Springer-Verlag, Berlin-Heidelberg, 1987.
39. C. Nelson and A. Siegel, *Parsimonious modeling of yield curves*, J. of Business **60** (1987), 473–489.
40. D. Revuz and M. Yor, *Continuous martingales and Brownian motion*, Grundlehren der mathematischen Wissenschaften, vol. 293, Springer-Verlag, Berlin-Heidelberg-New York, 1994.
41. W. Rudin, *Functional analysis*, second ed., McGraw-Hill, 1991.
42. H. R. Schwarz, *Numerische Mathematik*, second ed., Teubner, Stuttgart, 1986.
43. J. Stoer and R. Bulirsch, *Einführung in die Numerische Mathematik*, second ed., vol. II, Springer-Verlag, Berlin-Heidelberg-New York, 1978.
44. L. E. O. Svensson, *Estimating and interpreting forward interest rates: Sweden 1992-1994*, IMF Working Paper No. 114, September 1994.
45. G. Tessitore and J. Zabczyk, *Comments on transition semigroups and stochastic invariance*, Preprint, 1998.
46. T. Vargiolu, *Calibration of the Gaussian Musiela model using the Karhunen–Loeve expansion*, Working paper, Scuola Normale Superiore Pisa, 1998.
47. ———, *Invariant measures for the Musiela equation with deterministic diffusion term*, Finance Stochast. **3** (1999), no. 4, 483–492.
48. O. Vasicek, *An equilibrium characterization of the term structure*, J. Finan. Econom. **5** (1977), 177–188.
49. M. Yor, *Existence et unicité de diffusions à valeurs dans un espace de Hilbert*, Ann. Inst. H. Poincaré Sect. B (N.S.) **10** (1974), 55–88.
50. J. Zabczyk, *Stochastic invariance and consistency of financial models*, Preprint, Polish Academy of Sciences, Warsaw, 1999.