Finite Dimensional Realizations for Term Structure Models driven by Semimartingales

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium
(Dr. rer. nat.)
im Fach Mathematik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
der Humboldt-Universität zu Berlin

von
Herrn Dipl.-Math. Stefan Tappe
geboren am 7. Februar 1977 in Detmold

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Hans Jürgen Prömel

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:
Prof. Dr. Uwe Küchler

Gutachter:

1. Prof. Dr. Uwe Küchler
2. Prof. Tomas Björk, Ph.D.
3. Prof. Dr. Jan Kallsen

eingereicht am: 5. Juli 2005
Tag der mündlichen Prüfung: 7. November 2005

Abstract

Let $f(t,T)$ be a term structure model of Heath-Jarrow-Morton type

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dX_t,$$

driven by a multidimensional semimartingale X. Our objective is to study the existence of finite dimensional realizations for equations of this kind. Choosing the class of Grigelionis processes (including in particular Lévy processes) as driving processes, we approach this problem from two different directions.

Extending the ideas from differential geometry in Björk and Svensson (2001), we show that the criterion for the existence of finite dimensional realizations, proven in the aforementioned paper, still serves as a necessary condition in our setup. This result is applied to concrete volatility structures.

In the context of benchmark realizations, which are a natural generalization of short rate realizations, we derive integro-differential equations, suitable for the analysis of the realization problem. Generalizing Jeffrey (1995), we also prove a result stating that forward rate models, which generically possess a benchmark realization, must have a singular Hessian matrix.

Both approaches reveal that, with regard to the results known for driving Wiener processes, new phenomena emerge, as soon as the driving process X has jumps. In particular, the occurrence of jumps severely limits the range of models that admit finite dimensional realizations. For this reason we prove, for the often considered case of deterministic direction volatility structures, the existence of finite dimensional systems converging to the forward rate model.

Keywords:
HJM term structure models driven by semimartingales, Levy processes, finite dimensional realizations, Lie algebras

Zusammenfassung

Es sei ein Heath-Jarrow-Morton Zinsstrukturmodell

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dX_t$$

gegeben, angetrieben von einem mehrdimensionalen Semimartingal X. Das Ziel dieser Arbeit besteht darin, die Existenz endlich dimensionaler Realisierungen für solche Modelle zu untersuchen, wobei wir als treibende Prozesse die Klasse der Grigelionis Prozesse wählen, die insbesondere Lévy Prozesse enthält. Zur Bearbeitung der Fragestellung werden zwei veschiedene Ansätze verfolgt.

Wir dehnen die Ideen aus der Differenzialgeometrie von Björk und Svensson (2001) auf die vorliegende Situation aus und zeigen, dass das in der zitierten Arbeit bewiesene Kriterium für die Existenz endlich dimensionaler Realisierungen in unserem Fall als notwendiges Kriterium dienlich ist. Dieses Resultat wird auf konkrete Volatilitätsstrukturen angewandt.

Im Kontext von sogenannten Benchmark Realisierungen, die eine natürliche Verallgemeinerung von Short Rate Realisierungen darstellen, leiten wir Integro-Differenzialgleichungen her, die für die Untersuchung der Existenz endlich dimensionaler Realisierungen hilfreich sind. Als Verallgemeinerung eines Resultats von Jeffrey (1995) beweisen wir außerdem, dass Zinsstrukturmodelle, die eine generische Benchmark Realisierung besitzen, notwendigerweise eine singuläre Hessesche Matrix haben.

Beide Ansätze zeigen, dass neue Phänomene auftreten, sobald der treibende Prozess X Sprünge macht. Es gibt dann auf einmal nur noch sehr wenige Zinsstrukturmodelle, die endlich dimensionale Realisierungen zulassen, was ein beträchtlicher Unterschied zu solchen Modellen ist, die von einer Brownschen Bewegung angetrieben werden. Aus diesem Grund zeigen wir, dass für die in der Literatur oft behandelten Modelle mit deterministischer Richtungsvolatilität eine Folge von endlich dimensionalen Systemen existiert, die gegen das Zinsmodell konvergieren.

Schlagwörter:
von Semimartingalen angetriebene HJM Zinsstrukturmodelle, Levy Prozesse, endlich dimensionale Realisierungen, Lie Algebren

Contents

Chapter 1

Introduction

A zero coupon bond with maturity date T, also called a T-bond, is a contract which guarantees the holder one dollar (or one unit of another currency) to be paid at the date T. Denoting by $p(t,T)$ the price at time t of a bond with maturity date T, Figure 1.0.1 shows one sample path of the price process $p(\bullet, 10)$ of a zero coupon bond maturing in 10 years.

The problem of finding finite dimensional realizations

The bond price processes $p(\bullet, T)$, $T \in \mathbb{R}_+$ form a continuum of stochastic processes. Our objective of the present text is to investigate when and how a given, a priori infinite dimensional zero coupon bond model, admits a finite dimensional realization. More precisely, we wish to find a finite dimensional stochastic process Z satisfying a stochastic differential equation

$$\begin{cases} dZ_t &= \mu(t, Z_t)dt + \gamma(t, Z_{t-})dX_t \\ Z_0 &= z_0 \end{cases}, \tag{1.0.1}$$

and a mapping P such that the bond prices $p(t,T)$ can be represented as $p(t,T) = P(t,T,Z_t)$. Then, the continuum of bond price processes can be realized by means of a finite dimensional state process Z. Under appropriate conditions on the coefficients μ, γ and the driving process X in the stochastic differential equation (1.0.1), the state process is a Markov process. This is of relevance for numerical methods.

As in Heath, Jarrow, and Morton [36] we assume that the forward rates of the term structure models, for which we treat the realization problem, are specified as

$$\begin{cases} df(t,T) &= \alpha(t,T)dt + \sigma(t,T)dX_t, \\ f(0,T) &= f^*(0,T) \end{cases}. \tag{1.0.2}$$

The bond prices are then given by

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

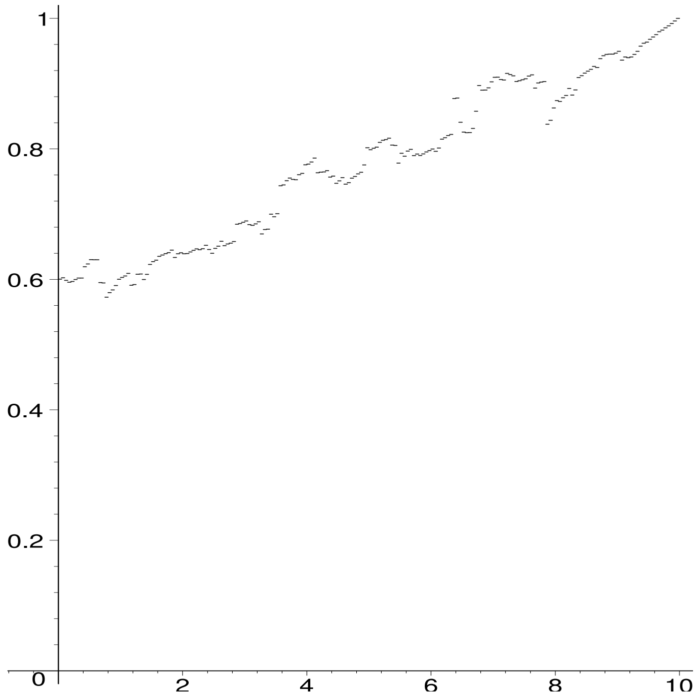

Figure 1.0.1: Price process $p(\bullet, 10)$ of a 10-year zero coupon bond.

In Heath, Jarrow, and Morton [36] the driving process X appearing in the forward rate equation (1.0.2) is a Brownian motion. As Figure 1.0.1 indicates, we are interested in models which take into account the appearance of jumps. Interest rate models driven by jump processes are strongly supported by empirical facts, see Raible [56, Chap. 5], where it is argued that empirically observed log returns of zero coupon bonds are not normally distributed, a fact, which has long before been known for the distributions of stock returns. Raible [56] therefore recommends to replace the Browninan motion by a more general Lévy process. However, despite these empirical evidences, only a few authors deal with term structure models driven by processes which allow jumps, see Section 2.2 for references, and among these authors, only a few study the existence of finite dimensional realizations.

There is a substantial literature studying the existence of finite dimensional realizations when the term structure model is driven by Wiener processes. This is done, for various special cases, in Ritchken and Sankarasubramanian [58], Cheyette [21], Bhar and Chiarella [6], Inui and Kijima [41], Björk and Gombani [11] and Chiarella and Kwon [22]. The typical assumption is that the volatility structures are functions of the short rate. A model, where the volatility is allowed to depend on a finite number of benchmark forward rates, is considered in Chiarella and Kwon [22]. The case of deterministic volatility structures was completely solved in Björk and Gombani [11].

Short rate realizations, that is realizations where the only state variable of the realization is the short rate, are considered in Carverhill [20] (for deterministic volatilities), and in Jeffrey [44], which is the deepest study of short rate realizations driven by Brownian motions. Mari [52] provides an analytical treatment of HJM interest rate models satisfying Jeffrey's constraint.

Using ideas from differential geometry, Björk and Svensson [13] have performed a systematical treatment of the realization problem, and have provided a criterion, which is necessary and sufficient for the existence of finite dimensional realizations. All above mentioned papers on that topic are indeed special cases of this framework, and their results have been rediscovered and extended in Björk and Svensson [13]. They also use their results in order to give an alternative proof of Jeffrey [44]. The paper Björk and Landén [12] provides a general method for the construction of finite dimensional realizations, and illustrates the method by constructing realizations for a number of concrete examples.

One technical problem, arising in Björk and Svensson [13], is that one chooses in this framework to a Banach space for the forward rate curves, which is very small. In Filipović and Teichmann [33], the theory was extended to a larger space. The extended theory was applied in Filipović and Teichmann [32] and Filipović and Teichmann [34].

In Björk, Landén, and Svensson [16] the theory developed in Björk and Svensson [13] was extended to stochastic volatility models.

We exhibit that all the just mentioned literature only deals with term structure models driven by Wiener processes. Björk [8] provides a survey about this topic. However, there are only a few references that deal with realizations for term structure models, in which jumps may occur.

In Eberlein and Raible [29], short rate realizations for models with deterministic volatility are studied, where the driving process is allowed to be a Lévy process with a restrictive property. The results of this paper were extended in Küchler and Naumann [46] to the most general class of Lévy processes being possible in the considered framework, and in Gapeev and Küchler [35] to term structure models driven by jump-diffusions. The latter article also deals with non-deterministic volatility structures.

Björk and Gombani [11] have solved the realization problem for term structure models with deterministic volatility, driven by finitely many Wiener processes and a marked point process.

Duffie and Kan [25] study affine realizations for interest rate models which are based on Brownian motions, but also suggest, in the jump-diffusion case, to consider partial differential equations arising from the infinitesimal generator of the jump-diffusion. Such an idea was carried out in Hyll [40], who has studied short rate realizations for term structure models driven by finitely many Wiener processes and finitely many counting processes, where the volatilities are allowed to depend on the current state of the short rate.

An approach, where techniques of geometric measure theory are applied, is turned out in Lütkebohmert [48]. In this paper, the term structure model is driven by a multidimensional Wiener process and a compensated marked point process.

In this text, we shall go two ways in order to approach the realization problem.

First, we study the Lie algebraic methodology, developed in Björk and Svensson [13] for driving Wiener processes, and extend it to our setup with driving processes allowing jumps. The second approach is to go into the framework of benchmark realizations, which are a generalization of short rate realizations, and to derive integro-differential equations which help to investigate the existence of finite dimensional realizations. Indeed, to a certain extent, this approach is inspired by the mentioned idea from Duffie and Kan [25]. We point out that, during the whole text, we focus on the existence of global realizations. Other authors, like Björk and Svensson [13], analyze the existence of local realizations, that is realizations which are valid up to a positive stopping time.

Structuring of the text

In Chapter 2 we specify the term structure models, for which we investigate the existence of finite dimensional realizations. To meet this issue, we declare the range of driving processes to be considered, namely so-called Grigelionis processes, in Section 2.1. We list their relevant properties, and explain, why they form, a priori, a class of semimartingales that is as large as possible and serves our purposes. We should, however, remark that when dealing with finite dimensional realizations, we usually have to confine ourselves to subclasses of Grigelionis processes, but Lévy processes, or more general, the non-homogeneous Lévy processes in the terminology of Eberlein, Jacod, and Raible [30], will always be contained.

Term structure models driven by Grigelionis processes are formally introduced in Section 2.2. Some facts about these models, needed later, are provided, namely the HJM drift condition (Proposition 2.2.11), which arises from the assumption that the market is free of arbitrage in the sense that $\mathbb{P}$ is a martingale measure, and the Musiela parametrization (Proposition 2.2.12) of forward rates.

In Chapter 3 we start with the treatment of the realization problem by pursuing the Lie algebraic approach from Björk and Svensson [13]. As a matter of fact, this framework, which uses ideas from differential geometry, is only designed for models driven by Wiener processes. The appearance of jumps causes new geometric aspects. In Section 3.1 we give a summary of the geometry for models driven by Wiener processes (see Björk [8] for a survey), and extend these geometric ideas for our setup. After this informal digression, which serves the main purpose of making the following theory better understandable from an intuitive point of view, we prove two fundamental results, Theorem 3.1.17 and Theorem 3.1.21, which provide formulas for the drift and the volatilities of term structure models admitting realizations. The latter result is a version of the first for Fisk-Stratonovich dynamics. The formulas are applied later in Chapter 3 as well as in Chapter 4.

Things become concrete in Section 3.2, where we derive a necessary criterion for the existence of finite dimensional realizations of Banach space valued equations, which may also be driven by jump processes (Theorem 3.2.4). This result can be regarded as one implication of the equivalence in the main result in Björk and Svensson [13] for the case of jump processes. We argue that the converse of this equivalence is not true for driving processes that have jumps.

Term structure models can be embedded in the framework of Banach space valued equations, which is done in Section 3.3. This yields a necessary criterion for the existence of finite dimensional realizations for term structure models driven by processes that may have jumps.

We apply the criterion for two types of volatility structures, namely deterministic volatility (see Section 3.4) and deterministic direction volatility (studied in Section 3.5).

For deterministic volatility we obtain that a certain differential equation, which is already known in the literature for term structure models with driving Wiener processes, must be satisfied. This a priori necessary condition for the existence of a finite dimensional realization is also sufficient with a driving process admitting jumps, which we show by providing a concrete realization (Proposition 3.4.6).

Deterministic direction volatility structures turn out to be more delicate. Lie algebraic computations support that one can, in general, only hope for realizations in an approximative sense, that is, one finds finite dimensional systems of increasing dimensions, converging to the term structure model in an appropriate sense. Such a result is given in Theorem 3.5.18.

Our second approach to the existence of finite dimensional realization is the contents of Chapter 4. There, we work within the framework of benchmark realizations, which are a natural generalization of short rate realizations. This means, we consider finite dimensional realizations where the state process consists of a set of benchmark forward rates. Although this condition seems quite restricting at first glance, a finite dimensional realization with an arbitrary state process can usually be transformed to a benchmark realization. This is shown, besides other preliminaries, in Section 4.1. It is therefore no strong restriction to deal with benchmark realizations.

The advantage of this approach is that we obtain deterministic equations, in particular suitable integro-differential equations, depending on the variables t, T, which represent time, and on a vector $r \in \mathbb{R}^d$, representing the state process. Integro-differential equations are derived in Section 4.2.

Section 4.3 is devoted to the study of deterministic volatility structures. Not surprisingly, we rediscover results of Section 3.4, but we also obtain connections to works which deal with criteria for the short rate to be qualified as a Markov process, like Eberlein and Raible [29], Küchler and Naumann [46], and others. New aspects are also obtained for short rate realizations, if the derivative of the driving process is allowed to depend on the current state of the short rate. We show in Theorem 4.3.6 that in this case the compensator of the jump measure must necessarily have an affine structure.

In Section 4.4 we consider deterministic direction volatility structures. Theorem 4.4.1 states, roughly speaking, that benchmark realizations for such term structure models cannot exist, as soon as the driving process has jumps. This result has consequences for short rate realizations. Namely, in Theorem 4.4.4 we prove that for short rate realizations with an, a priori, arbitrary volatility structure and a driving process having jumps, the volatility must be deterministic and factorizes.

Jeffrey [44] has shown, for term structure models driven by a single Brownian motion, that every generic short rate realization is affine. In Section 4.5 we generalize this result as follows. As in the whole chapter, we consider benchmark realizations rather than short rate realizations, and the driving process is multidimensional and allowed to make jumps. Our result (Theorem 4.5.3) states that forward rate models, which generically possess a benchmark realization, must have a singular Hessian matrix, which means they have to be affine in the one-dimensional case.

There are some results in this text, whose proofs are established by using standard techniques or making somewhat tedious computations. In fact, a few among these results are already known in the literature in a slightly different context. Therefore, we have attached an appendix containing the proofs, which are omitted for this reason.

Chapter 2

The term structure models

In order to prepare the grounds for the subsequent analysis of the existence of finite dimensional realizations, this chapter presents the term structure models, for which we will study the problem. In the Heath, Jarrow, and Morton [36] setting, one assumes that the forward rates are given by

$$df(t,T) = \alpha(t,T)dt + \sum_{i=1}^{n} \sigma_i(t,T)dW_i(t),$$

where $W_1, \dots, W_n$ are independent Brownian motions.

As we have exhibited in the introduction, we are interested in the analysis of models that take into account the occurrence of jumps. Some models of this kind have already been investigated in the literature. In this text, we replace the Wiener process in the forward rate equation by a so-called Grigelionis process. We summarize the relevant properties of these processes in the first section. As we will explain below, Grigelionis processes provide a class of semimartingales, as large as possible (in particular, Lévy processes are included), which is appropriate for our purposes.

In the second section, we formally introduce term structure models driven by Grigelionis processes, and gather some of their properties, like the HJM drift condition, that arises from the absence of arbitrage, and the Musiela parametrization of forward rates. The relations to other interest rate models in the literature, which allow the presence of jumps, are discussed as well.

2.1 Grigelionis processes

Our terminology is chosen as in Jacod and Shiryaev [42], which is, besides Protter [55], our main reference for facts concerning stochastic analysis. Throughout this text,

$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P})$ denotes a stochastic basis (filtered probability space) in the sense of Jacod and Shiryaev [42, Def. I.1.2], i.e. the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is right-continuous but not necessarily complete. Concerning the filtration, we moreover assume that $\mathcal{F}_0 = \{\emptyset, \Omega\}$ is the trivial σ-algebra, which guarantees that all $\mathcal{F}_0$-measurable random variables are $\mathbb{P}$-almost surely constant.

By X we usually denote n-dimensional special semimartingales Jacod and Shiryaev [42, Def. I.4.22] for some $n \in \mathbb{N}$, i.e. each component of X can uniquely be decomposed into a local martingale and a predictable process of finite variation on compact intervals. Moreover, equations, inequalities etc. are usually only meant up to indistinguishability.

Most of the upcoming facts concerning Grigelionis processes are taken from Kallsen [45, Sec. 2.3].

2.1.1 Definition. *Let X be a special semimartingale. We call $(B, C, \nu)^I$ the integral characteristics of X, where*

1. *$B \in \mathcal{V}^n$ is the predictable part of finite variation in the canonical decomposition of X;*
2. *$C \in \mathcal{V}^{n\times n}$ is the continuous process defined by $C_{ij} := \langle X_i^c, X_j^c\rangle$ for any $i, j \in \{1, \dots, n\}$;*
3. *ν is the compensator of the random measure of jumps μ^X of X.*

2.1.2 Lemma. *Let X be a special semimartingale with integral characteristics given by $(B, C, \nu)^I$. Then, there exists a predictable, real-valued process $A \in \mathcal{A}^+_{\text{loc}}$, a predictable $\mathbb{R}^n$-valued process $(\beta_t)_{t\in\mathbb{R}_+}$, a predictable $\mathbb{R}^{n\times n}$-valued process $(c_t)_{t\in\mathbb{R}_+}$ whose values are symmetric, non-negative definite matrices and a transition kernel K from $(\Omega \times \mathbb{R}_+, \mathcal{P})$ into $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ satisfying for each $t \in \mathbb{R}_+$*

$$K_t(\{0\}) = 0,$$
$$\int_{\mathbb{R}^n} (|x|^2 \wedge 1) K_t(dx) < \infty,$$

such that for any $t \in \mathbb{R}_+$ it holds

$$B_t = \int_0^t \beta_s dA_s,$$
$$C_t = \int_0^t c_s dA_s,$$
$$\nu([0,t] \times G) = \int_0^t K_s(G) dA_s \quad \text{for any } G \in \mathcal{B}(\mathbb{R}^n).$$

Proof. See Jacod and Shiryaev [42, Prop. 2.9]. □

2.1.3 Remark. *We usually drop the argument ω in the notation of transition kernels from $(\Omega \times \mathbb{R}_+, \mathcal{P})$ into $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, as is done for stochastic processes.*

2.1.4 Definition. *We call a special semimartingale X Grigelionis process or locally infinitely divisible process if A in Lemma 2.1.2 can be chosen such that its paths $A(\omega) : \mathbb{R}_+ \to \mathbb{R}$ are absolutely continuous in time.*

The following result shows that a special semimartingale is a Grigelionis process if and only if its integral characteristics are absolutely continuous.

2.1.5 Lemma. *Let X be a special semimartingale with integral characteristics given by $(B, C, \nu)^I$. Then, there is equivalence between*

1. *X is a Grigelionis process.*
2. *There exists a predictable $\mathbb{R}^n$-valued process $(\beta_t)_{t\in\mathbb{R}_+}$, a predictable $\mathbb{R}^{n\times n}$-valued process $(c_t)_{t\in\mathbb{R}_+}$ whose values are symmetric, non-negative definite matrices and a transition kernel K from $(\Omega \times \mathbb{R}_+, \mathcal{P})$ into $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ satisfying for each $t \in \mathbb{R}_+$*

$$K_t(\{0\}) = 0,$$
$$\int_{\mathbb{R}^n} (|x|^2 \wedge 1) K_t(dx) < \infty,$$

such that for any $t \in \mathbb{R}_+$ it holds

$$B_t = \int_0^t \beta_s ds,$$
$$C_t = \int_0^t c_s ds,$$
$$\nu([0,t] \times G) = \int_0^t K_s(G) ds \quad \text{for any } G \in \mathcal{B}(\mathbb{R}^n).$$

Proof. See Kallsen [45, Lemma 2.12]. □

2.1.6 Definition. *Let X be a Grigelionis process. We call any triplet $(\beta, c, K)^D$ with β, c, K as in Lemma 2.1.5 differential characteristics or a derivative of X.*

The derivative $(\beta, c, K)^D$ of a Grigelionis process is uniquely determined up to a $(\mathbb{P} \otimes \text{Leb})$-null set $N \in \mathcal{P}$ Kallsen [45, Lemma 2.14].

If the derivative $(\beta, c, K)^D$ of a Grigelionis process X is deterministic, then X is a so-called PIIAC, or non-homogeneous Lévy process, see Eberlein, Jacod, and Raible [30], Eberlein and Kluge [27] and Eberlein and Kluge [28].

Grigelionis processes with constant derivatives are Lévy processes, which are of particular interest. Our standard reference in this text concerning Lévy processes is Sato [59]. Other textbooks are Bertoin [5] and Applebaum [1].

So far, we have presented Grigelionis processes and their relevant properties for this text. At this point, we shall motivate this choice of processes.

In order to establish subsequent results (cf. Proposition 2.2.4, Theorem 3.1.17 and others), the only essential (a priori) property of the driving process, that is required, is that it has absolutely continuous characteristics. Roughly speaking, one needs, in the proof of the mentioned results, that equations can be rewritten such that disturbing terms may be added to the drift term.

Since the absolute continuity of the characteristics is the only essential property, needed in the sequel, Grigelionis processes offer, regarded from this point of view, the biggest possible class of processes.

Nevertheless, we should note that, for our purposes, we will have to impose extra conditions on the processes. In order to establish the HJM drift condition for term structure models (Proposition 2.2.11), we need that the Grigelionis process has finite exponential moments (see Assumptions 2.2.6), and for the treatment of finite dimensional realizations, we will introduce non-degenerate and linearly non-degenerate Grigelionis processes (Definition 3.1.16). However, Lévy processes will always be included as particular cases.

2.2 Term structure models driven by Grigelionis processes

After introducing Grigelionis processes, we go on to specify the term structure models. For general background on interest rate theory, we refer to Björk [9], Björk [7] or Musiela and Rutkowski [54].

Assume that the forward rates $f : \Omega \times \Delta^2 \to \mathbb{R}$, where $\Delta^2 := \{(t,T) \in \mathbb{R}_+ \times \mathbb{R}_+ : t \le T\}$, are given by the equation

$$f(t,T) = f^*(0,T) + \int_0^t \alpha(s,T)ds + \sum_{i=1}^n \int_0^t \sigma_i(s,T)dX_s^i, \quad 0 \le t \le T \qquad (2.2.1)$$

where $f^*(0,\bullet) \in C^1(\mathbb{R}_+)$ is the initial forward rate curve, and X denotes a n-dimensional Grigelionis process with derivative $(\beta, c, K)^D$ for some $n \in \mathbb{N}$. By $\int_0^t Y_s dX_s$ for a locally bounded, predictable process Y and a semimartingale X, we always mean the integral over the half open interval $(0,t]$ excluding zero, which is written $\int_{0+}^t Y_s dX_s$ in the terminology of Protter [55]. Sometimes, we use the short-hand notation $Y \bullet X$. For stochastic integration with respect to semimartingales, we refer to Jacod and Shiryaev [42, Sec. I.4d].

Concerning the drift $\alpha : \Omega \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ and the volatilities $\sigma_1, \dots, \sigma_n : \Omega \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ we make the following assumptions.

2.2.1 Assumptions.

1. *$\alpha(t,T) = 0$ and $\sigma_i(t,T) = 0$, $i = 1,\dots,n$ for all $0 \le T < t$.*

2. *The mappings $\alpha, \sigma_1, \dots, \sigma_n$ are $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}_+)$-measurable.*

3. *For all $0 \le t \le T$ it holds*

$$\int_0^T \int_t^T |\alpha(s,v)| dvds < \infty,$$
$$\int_0^T \int_t^T |\sigma_i(s,v)|^2 dvds < \infty, \quad i = 1,\dots,n.$$

4. *For each $t \in \mathbb{R}_+$, it holds $\alpha(t,\bullet), \sigma_1(t,\bullet), \dots, \sigma_n(t,\bullet) \in C^1(\mathbb{R}_+)$, the derivatives are again integrable, and for all $0 \le t \le T$ the following identities are valid.*

$$\int_0^t \frac{\partial}{\partial T}\alpha(s,T)ds = \frac{\partial}{\partial T}\int_0^t \alpha(s,T)ds,$$
$$\int_0^t \frac{\partial}{\partial T}\sigma_i(s,T)dX_s^i = \frac{\partial}{\partial T}\int_0^t \sigma_i(s,T)dX_s^i, \quad i = 1,\dots,n.$$

We have to comment on these assumptions. We impose the first condition, because for $T < t$ the drift and the volatilities are not meaningful.

The second and the third condition guarantee that for each $T \in \mathbb{R}_+$ the processes $\alpha(\bullet,T)$ and $\sigma_1(\bullet,T),\dots,\sigma_n(\bullet,T)$ are locally bounded and predictable, whence the forward rate equation (2.2.1) is well-defined. They also ensure that we may apply stochastic Fubini theorems in the proofs of Proposition 2.2.4 (the bond price equation) and Proposition 2.2.12 (the Musiela parametrization).

The fourth condition, which is rather ad hoc, is needed in order to interchange differentiation and stochastic integration. That is also needed for the proof of the Musiela parametrization, as well as our earlier assumption that $f^*(0,\bullet) \in C^1(\mathbb{R}_+)$.

In the sequel, we abbreviate forward rate equations of the form (2.2.1) as

$$\begin{cases} df(t,T) &= \alpha(t,T)dt + \sigma(t,T)dX_t, \\ f(0,T) &= f^*(0,T) \end{cases}, \tag{2.2.2}$$

and agree that, unless stated otherwise, the initial forward rate curve $f^*(0,\bullet)$ is of class $C^1(\mathbb{R}_+)$ and that Assumptions 2.2.1 are satisfied.

Heath, Jarrow, and Morton [36] were the first, who have proposed forward rate models of the type (2.2.2), with a multidimensional driving Wiener process.

In the literature, there are only a few works that deal with term structure models which take into account the possible occurrence of jumps. Some authors, Shirakawa [60], Jarrow and Madan [43] or Hyll [40] for instance, consider jump-diffusion models, where the sources of randomness are a multidimensional Wiener process, and a multidimensional point process (i.e. vectors of processes with values in $\mathbb{N}_0$, like a Poisson process).

Allowing a continuous jump spectrum, rather than the simple Poisson jump models, Björk, Kabanov, and Runggaldier [15] consider forward rate dynamics of the form

$$df(t,T) = \alpha^*(t,T)dt + \sigma^*(t,T)dW_t + \int_E \delta^*(t,x,T)\mu(dt,dx). \tag{2.2.3}$$

Term structure models of this kind are also studied in Gapeev and Küchler [35]. In the subsequent paper Björk, Di Masi, Kabanov, and Runggaldier [14] the forward rates are specified as

$$df(t,T) = \alpha^*(t,T)dt + \sigma^*(t,T)dW_t + \int_E \delta^*(t,x,T)(\mu-\nu)(dt,dx), \tag{2.2.4}$$

In both equations, (2.2.3) and (2.2.4), W denotes a multidimensional standard Wiener process, and $\mu(dt,dx)$ denotes a $\mathcal{P}\otimes\mathcal{E}$-$\sigma$-finite random measure Jacod and Shiryaev [42, Def. II.1.6] on a Lusin space $(E,\mathcal{E})$ with compensator $\nu(dt,dx)$.

Our term structure models (2.2.2) are contained in these frameworks as follows. A Grigelionis process X with derivative $(\beta,c,K)^D$ has the canonical decomposition $X_t = X_0 + X_t^c + x*(\mu^X-\nu)_t + \int_0^t \beta_s ds$. It is no further restriction to assume that the continuous martingale part X^c is given by $X_t^c = \int_0^t \sqrt{c_s}dW_s$, where W is a n-dimensional standard Wiener process and $\sqrt{c_s}$ is the square-root of the symmetric non-negative definite matrix c_s. We choose the Lusin space $(E,\mathcal{E}) = (\mathbb{R}^n,\mathcal{B}(\mathbb{R}^n))$, and the random measure $\mu=\mu^X$ associated to jumps of X. Then, the forward rate dynamics (2.2.2) are of the form (2.2.4) with coefficients

$$\begin{aligned}\alpha^*(t,T) &= \alpha(t,T) + \langle \beta_t, \sigma(t,T)\rangle,\\ \sigma^*(t,T) &= \sigma(t,T)\sqrt{c_t},\\ \delta^*(t,x,T) &= \langle x, \sigma(t,T)\rangle.\end{aligned}$$

Analogously, if the Grigelionis process X admits the representation $X = X_0 + X^c + x*\mu^X$ (for instance a Lévy process with Lévy measure K satisfying $\int_{|x|\le 1}|x|K(dx)<\infty$), then the forward rate dynamics are of the form (2.2.3).

Eberlein and Raible [29] and Küchler and Naumann [46] treat term structure models with deterministic volatility, driven by a single Lévy process. These articles, and also Gapeev and Küchler [35], start with the specification of the bond prices, and later the forward rates are derived.

This framework was considerably extended by allowing a multidimensional PIIAC, also called non-homogeneous Lévy process, as driving process in the forward rate equation (2.2.2), see Eberlein, Jacod, and Raible [30], Eberlein and Kluge [27] and Eberlein and Kluge [28].

We remark that, while in most of the mentioned articles, the term structure models are defined on an interval $[0,T^*]$ with a finite time horizon $T^*>0$, we work on $\mathbb{R}_+$.

We shall now define the bond prices $p(t,T)$ and other useful quantities. This is followed by the formal definition of a finite dimensional realization for the term structure model.

2.2.2 Definition.

1. *The* <u>*short rate*</u> *at time* $t \in \mathbb{R}_+$ *is defined as* $r_t := f(t,t)$.

2. *The* <u>*bond prices*</u> *at time* t *with maturity* T *are given by*

$$p(t,T) := \exp\left(-\int_t^T f(t,s)ds\right), \quad 0 \le t \le T,$$

3. *We define the* <u>*discounted bond prices*</u> *as*

$$z(t,T) := \exp\left(-\int_0^t r_s ds\right) p(t,T), \quad 0 \le t \le T.$$

We remark that the forward rates are thus obtained from the bond prices as

$$f(t,T) = -\frac{\partial}{\partial T} \ln p(t,T).$$

Each of the processes $p(\bullet,T), z(\bullet,T), f(\bullet,T)$ is a priori just defined on the interval $[0,T]$. For notational convenience, we agree that they are extended to $\mathbb{R}_+$ by setting $p(t,T) = p(T,T)$ for $t \ge T$, and analogously for $z(t,T)$ and $f(t,T)$.

2.2.3 Definition.

1. *Let* $d \in \mathbb{N}$. *The term structure model (2.2.2) has a* $d+1$*-dimensional* <u>*realization*</u> *if there exists a pair* (F,Z), *where* $F : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ *is a mapping, and* Z *is a* d*-dimensional semimartingale satisfying*

$$\begin{cases} dZ_t &= \mu(t, Z_{t-})dt + \gamma(t, Z_{t-})dX_t \\ Z_0 &= z_0 \end{cases} \tag{2.2.5}$$

for Borel functions $\mu : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$, $\gamma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times n}$ *and* $z_0 \in \mathbb{R}^d$, *such that* $f(t,T) = F(t,T,Z_t)$ *for all* $0 \le t \le T$. *We call the semimartingale* Z *from the pair* (F,Z) *the* <u>*state process*</u> *of the realization.*

2. *The term structure model (2.2.2) has a* <u>*finite dimensional realization*</u> *if it admits a* $d+1$*-dimensional realization for some* $d \in \mathbb{N}$.

3. *A* $d+1$*-dimensional realization is said to be* <u>*affine*</u> *if there are* $a, b_1, \ldots, b_d$: $\mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ *such that the mapping* F *of the associated pair* (F,Z) *is of the form*

$$F(t,T,z) = a(t,T) + \langle b(t,T), z\rangle \quad \text{for all } (t,T,z) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d.$$

The reason why we speak about $d+1$-dimensional realizations rather than d-dimensional realizations is that, for each $T \in \mathbb{R}_+$, the forward rates $f(\bullet, T)$ are given by applying the mapping $F(\bullet, T, \bullet)$ on the $d+1$-dimensional process (t, Z_t). In other words, we get an extra dimension by adding the time t to the state process Z.

If X is a vector of independent Lévy processes and μ, γ in (2.2.5) are Lipschitz in the sense of Protter [55, p. 236], then Z is a Markov process according to Protter [55, Thm. V.32]. When dealing with the Lie algebraic theory, one also considers Fisk-Stratonovich dynamics. Although the Itô dynamics are of major interest, we mention that the just cited result, that Z is a Markov process under appropriate conditions, remains true with Fisk-Stratonovich differentials Protter [55, Thm. V.34].

In order to guarantee that the bond market is free of arbitrage, one usually assumes the existence of a martingale measure, i.e. a probability measure $\mathbb{Q} \sim \mathbb{P}$ such that, for each $T \in \mathbb{R}_+$, the discounted bond prices $z(\bullet, T)$ are local $\mathbb{Q}$-martingales. For more details on that topic, see, e.g., Björk, Kabanov, and Runggaldier [15].

For the treatment of the realization problem, we assume that the market is arbitrage-free in the sense that the original probability measure $\mathbb{P}$ is a martingale measure.

This condition implies that the drift cannot be chosen freely, but is rather determined by the volatilities. In order to derive the condition on the drift, we need the dynamics of the discounted bond prices $z(t,T)$. We start with an equation for the bond prices $p(t,T)$, from which the dynamics of $z(t,T)$ can easily be derived. For want of a shorter notation, we define $A(t,T) := -\int_t^T \alpha(t,s)ds$ and $\Sigma(t,T) := -\int_t^T \sigma(t,s)ds$. Recall that $(\beta, c, K)^D$ denotes the derivative of the Grigelionis process X.

2.2.4 Proposition. *For each $T \in \mathbb{R}_+$, the bond prices $p(\bullet, T)$ satisfy*

$$
\begin{aligned}
dp(t,T) = p(t-,T)\left(r_t + A(t,T) + \langle \beta_t, \Sigma(t,T)\rangle + \frac{1}{2}\langle \Sigma(t,T), c_t \Sigma(t,T)\rangle\right) dt \\
+ p(t-,T)\Sigma(t,T)dX_t^c + p(t-,T)\int_{\mathbb{R}^n} \langle x, \Sigma(t,T)\rangle (\mu^X - \nu)(dt,dx) \\
+ p(t-,T)\int_{\mathbb{R}^n} \left(e^{\langle x, \Sigma(t,T)\rangle} - 1 - \langle x, \Sigma(t,T)\rangle\right)\mu^X(dt,dx). \qquad (2.2.6)
\end{aligned}
$$

Proof. See the appendix. □

2.2.5 Corollary. *For each $T \in \mathbb{R}_+$, the discounted bond prices $z(\bullet, T)$ satisfy*

$$
\begin{aligned}
dz(t,T) = z(t-,T)\left(A(t,T) + \langle \beta_t, \Sigma(t,T)\rangle + \frac{1}{2}\langle \Sigma(t,T), c_t \Sigma(t,T)\rangle\right) dt \\
+ z(t-,T)\Sigma(t,T)dX_t^c + z(t-,T)\int_{\mathbb{R}^n} \langle x, \Sigma(t,T)\rangle (\mu^X - \nu)(dt,dx) \\
+ z(t-,T)\int_{\mathbb{R}^n} \left(e^{\langle x, \Sigma(t,T)\rangle} - 1 - \langle x, \Sigma(t,T)\rangle\right)\mu^X(dt,dx). \qquad (2.2.7)
\end{aligned}
$$

Proof. See the appendix. □

Our idea, in order to derive the drift condition, is as follows. We intend to write the $\mu^X(dt, dx)$-integral in (2.2.7) as $(\mu^X - \nu)(dt, dx)$-integral plus $\nu(dt, dx)$-integral. Since the compensator ν is absolutely continuous, we can then take the latter integral to the drift term, and obtain that $\mathbb{P}$ is a martingale measure if and only if the thus derived drift term is equal to zero.

The problem, however, is that the integrand needs not to be integrable with respect to the compensated jump measure $\mu^X - \nu$. Therefore, we make, for the rest of this section, the following assumptions.

2.2.6 Assumptions.

1. *There are* $z_1^-, \ldots, z_n^- \in (-\infty, 0)$ *and* $z_1^+, \ldots, z_n^+ \in (0, \infty)$ *such that, up to a* $\mathbb{P}$*-null set, for any* $t \in \mathbb{R}_+$
$$\int_0^t \left(\int_{|x|\leq 1} |x|^2 K_s(dx) + \int_{|x|>1} e^{\langle z,x\rangle} K_s(dx) \right) ds < \infty, \quad z \in Q, \tag{2.2.8}$$
where Q *denotes the set* $Q := [z_1^-, z_1^+] \times \ldots \times [z_n^-, z_n^+]$*, and it holds for all* $(\omega, t) \in \Omega \times \mathbb{R}_+$ *(up to an evanescent set)*
$$\int_{|x|>1} e^{\langle z,x\rangle} K_{(\omega,t)}(dx) < \infty, \quad z \in Q. \tag{2.2.9}$$

2. *There are* $w_1^- \in (z_1^-, 0), \ldots, w_n^- \in (z_n^-, 0)$ *and* $w_1^+ \in (0, z_1^+), \ldots, w_n^+ \in (0, z_n^+)$ *such that, up to a* $\mathbb{P}$*-null set,*
$$\Sigma(t, T) \in Q_0 \quad \textit{for all } 0 \leq t \leq T,$$
where $Q_0 \subset Q$ *is defined as* $Q_0 := [w_1^-, w_1^+] \times \ldots \times [w_n^-, w_n^+]$.

The main restriction in the first assumption is of course (2.2.8), since then, condition (2.2.9) is fulfilled for all t with possible exception of a set of Lebesgue measure zero.

It follows from Sato [59, Thm. 25.17] that for all $(\omega, t) \in \Omega \times \mathbb{R}_+$ (up to evanescence) the function
$$\Psi_{(\omega,t)}(z) := \langle \beta_t(\omega), z\rangle + \frac{1}{2}\langle z, c_t(\omega) z\rangle + \int_{\mathbb{R}^n} \left(e^{\langle z,x\rangle} - 1 - \langle z, x\rangle\right) K_{(\omega,t)}(dx) \tag{2.2.10}$$
is definable for all $z \in \mathbb{C}$ with $\operatorname{Re} z \in Q$, and that it holds
$$e^{\Psi_{(\omega,t)}(z)} = \int_{\mathbb{R}^n} e^{\langle z,x\rangle} \mu_{(\omega,t)}(dx),$$
where $\mu_{(\omega,t)}$ denotes the infinitely divisible distribution with generating triplet $(\beta_t(\omega), c_t(\omega), K_{(\omega,t)})_1$, that is, the characteristic function of $\mu_{(\omega,t)}$ is given by
$$\hat{\mu}_{(\omega,t)}(z) = \exp\left[-\frac{1}{2}\langle z, c_t(\omega) z\rangle + i\langle \beta_t(\omega), z\rangle + \int_{\mathbb{R}^n} \left(e^{i\langle z,x\rangle} - 1 - i\langle z, x\rangle\right) K_{(\omega,t)}(dx)\right]$$

for $z \in \mathbb{R}^n$. In other words, $\Psi_{(\omega,t)}$ denotes the cumulant generating function of $\mu_{(\omega,t)}$.

This leads to a brief study of cumulant generating functions. Let μ be an infinitely divisible distribution on $\mathbb{R}^n$ with generating triplet $(\beta, c, K)_1$, satisfying

$$\int_{|x|>1} e^{\langle z,x\rangle} K(dx) < \infty, \quad z \in Q$$

where $Q \subset \mathbb{R}^n$ denotes a set as in Assumptions 2.2.6. Then, the cumulant generating function Ψ is of class $C^\infty(\operatorname{int} Q)$, and the derivatives are obtained by differentiating under the integral sign. This follows from the dominated convergence principle and its consequences, see, e.g., Bauer [4, Chap. II.16]. We obtain the following result concerning the continuity of sample paths.

2.2.7 Lemma. *Let $I \subset \mathbb{R}_+$ be a bounded or unbounded interval. There is equivalence between*

1. *The process X has ($\mathbb{P} - a.s.$) only continuous paths on the interval I.*
2. *It holds $K = 0$ $(\mathbb{P} \otimes \mathrm{Leb}|_{\mathcal{B}(I)})$-almost surely.*
3. *The functions Ψ are $(\mathbb{P} \otimes \mathrm{Leb}|_{\mathcal{B}(I)})$-almost surely polynomials on $\mathbb{R}^n$.*

If these equivalent conditions are satisfied, then the degree of each polynomial $\Psi_{(\omega,t)}$ is at most two.

Proof. The equivalence of the first two statements holds, because $dt\, K_t(dx)$ is the compensator of the random measure of jumps μ^X of X, which is unique up to a $\mathbb{P}$-null set Jacod and Shiryaev [42, Thm. II.1.8].

From the representation (2.2.10) of the cumulant generating functions $\Psi_{(\omega,t)}$ for $(\omega, t) \in \Omega \times I$, and the fact that the partial derivatives of arbitrary order are obtained by differentiating under the integral sign, it follows by evaluating the derivatives at zero that each $\Psi_{(\omega,t)}$ is a polynomial if and only if $K_{(\omega,t)} = 0$, and that its degree is in this case at most two. □

The one-dimensional case has some special features. Let μ be an infinitely divisible distribution on $\mathbb{R}$ with characteristic triplet $(\beta, c, K)_1$, satisfying

$$\int_{|x|>1} e^{\langle z,x\rangle} K(dx) < \infty, \quad z \in [z^-, z^+], \tag{2.2.11}$$

where $z^- \in (-\infty, 0)$ and $z^+ \in (0, \infty)$. Then, the cumulant generating function Ψ is holomorphic on $G := \{z \in \mathbb{C} \mid \operatorname{Re} z \in (z^-, z^+)\}$, and it holds

$$\Psi(0) = 0, \tag{2.2.12}$$

$$\Psi'(0) = \beta, \tag{2.2.13}$$

$$\Psi''(0) = c + \int_{\mathbb{R}} x^2 K(dx), \tag{2.2.14}$$

$$\Psi^{(n)}(0) = \int_{\mathbb{R}} x^n K(dx), \quad n \geq 3. \tag{2.2.15}$$

A rather short proof of these statements can be obtained by using Elstrodt [31, Thm. IV.5.8], which is also a consequence of the dominated convergence principle. Another reference on this topic is Lukacs [47].

There is an immediate consequence for infinitely divisible distributions, which will be useful later, for the proof of Theorem 4.3.6.

2.2.8 Corollary. *Let μ_1, μ_2 be two infinitely divisible distributions on $\mathbb{R}$ with generating triplets $(\beta, c, K_i)_1$, $i = 1, 2$, with Lévy measures K_i which fulfill (2.2.11). Then, the identity*

$$\int_{\mathbb{R}} x^n K_1(dx) = \int_{\mathbb{R}} x^n K_2(dx) \quad \text{for all } n \geq 2$$

implies that $\mu_1 = \mu_2$.

Proof. According to equations (2.2.12)-(2.2.15), one has $\Psi_1^{(n)}(0) = \Psi_2^{(n)}(0)$ for all $n \in \mathbb{N}_0$, where Ψ_1, Ψ_2 denote the corresponding cumulant generating functions. A standard result in complex analysis, see, e.g., Remmert [57, Thm. 8.1.1], yields that $\Psi_1 = \Psi_2$. Consequently, it holds $\mu_1 = \mu_2$. □

After this digression about cumulant generating functions, we turn back to the original problem. We impose the second condition in Assumptions 2.2.6, because we intend to insert $\Sigma(t, T)$ in Ψ and its partial derivatives. The functions Ψ are of class C^∞ on the interior of Q, but we also need that Σ lies in a compact set. For this reason, we demand that Σ takes it values in the smaller set Q_0.

Note that Assumptions 2.2.6 correspond, in a certain sense, to Assumptions ($\mathbb{INT}$) and ($\mathbb{DET}$) in Eberlein and Kluge [27, Sec. 3].

Imposing Assumptions 2.2.6, we obtain:

2.2.9 Proposition. *For all $T \in \mathbb{R}_+$, the bond prices $p(\bullet, T)$ satisfy*

$$\begin{aligned} dp(t,T) &= p(t-,T)\left(r_t + A(t,T) + \Psi_t(\Sigma(t,T))\right)dt \\ &\quad + p(t-,T)\Sigma(t,T)dX_t^c + p(t-,T)\int_{\mathbb{R}^n}\left(e^{\langle x,\Sigma(t,T)\rangle} - 1\right)(\mu^X - \nu)(dt,dx), \end{aligned} \tag{2.2.16}$$

and for the discounted bond prices $z(\bullet, T)$ it holds

$$\begin{aligned} dz(t,T) &= z(t-,T)\left(A(t,T) + \Psi_t(\Sigma(t,T))\right)dt \\ &\quad + z(t-,T)\Sigma(t,T)dX_t^c + z(t-,T)\int_{\mathbb{R}^n}\left(e^{\langle x,\Sigma(t,T)\rangle} - 1\right)(\mu^X - \nu)(dt,dx). \end{aligned} \tag{2.2.17}$$

Proof. Note that $p(\bullet, T)$ and $z(\bullet, T)$ are bounded for every $T \in \mathbb{R}_+$, because they are càdlàg and constant after time $t = T$. Thus, the Assumptions 2.2.6 imply that

$$\int_0^t p(s-,T)\int_{\mathbb{R}^n}\left|e^{\langle x,\Sigma(s,T)\rangle} - 1 - \langle x,\Sigma(s,T)\rangle\right|K_s(dx)ds < \infty \quad \text{for all } 0 \leq t \leq T,$$

$$\int_0^t z(s-,T)\int_{\mathbb{R}^n}\left|e^{\langle x,\Sigma(s,T)\rangle} - 1 - \langle x,\Sigma(s,T)\rangle\right|K_s(dx)ds < \infty \quad \text{for all } 0 \leq t \leq T.$$

Hence, for every $T \in \mathbb{R}_+$, the processes

$$t \mapsto \int_0^t p(s-,T) \int_{\mathbb{R}^n} \left| e^{\langle x, \Sigma(s,T) \rangle} - 1 - \langle x, \Sigma(s,T) \rangle \right| \nu(ds,dx),$$
$$t \mapsto \int_0^t z(s-,T) \int_{\mathbb{R}^n} \left| e^{\langle x, \Sigma(s,T) \rangle} - 1 - \langle x, \Sigma(s,T) \rangle \right| \nu(ds,dx)$$

belong to $\mathcal{A}_{\text{loc}}^+$. Applying Jacod and Shiryaev [42, Prop. II.1.28], we can write the $\mu^X(dt,dx)$-integrals in (2.2.6) and (2.2.7) as $(\mu^X - \nu)(dt,dx)$-integral plus $\nu(dt,dx)$-integral, and obtain the stated dynamics. □

By $\mathfrak{L}$ we denote the space of left-continuous functions admitting right-hand limits.

2.2.10 Lemma. *Let $f, g \in \mathfrak{L}$ be such that $g(t) > 0$, $t \in (0,\infty)$ and $\int_0^t f(s)g(s)ds = 0$ for each $t \in \mathbb{R}_+$. Then, it holds $f(t) = 0$ for all $t \in (0,\infty)$.*

Proof. See the appendix. □

One may wonder about the following left-continuity assumptions. Without this regularity, one can only deduce the following identity (2.2.18) for all t up to a set of Lebesgue measure zero, and that is exactly the reason for proving the previous lemma. In the sequel, where typically the volatility is of the form $\sigma(t,T,r_{t-})$ with a continuous mapping σ, these assumptions will be satisfied.

Note that, in the following, we drop the argument ω for a more convenient notation of the cumulant generating functions, and write Ψ_t instead of $\Psi_{(\omega,t)}$. The space $\mathbb{L}$ denotes the space of càglàd adapted processes.

2.2.11 Proposition. *Assume that $\alpha(\bullet,T), \sigma_1(\bullet,T), \ldots, \sigma_n(\bullet,T) \in \mathbb{L}$ for all $T \in \mathbb{R}_+$, and that (up to a $\mathbb{P}$-null set) the mapping $(t,z) \mapsto \Psi_t(z)$ is continuous in z and left-continuous in t. Then, $\mathbb{P}$ is a martingale measure if and only if, for every $T \in (0,\infty)$, it holds (up to a $\mathbb{P}$-null set)*

$$\int_t^T \alpha(t,s)ds = \Psi_t \left(- \int_t^T \sigma(t,s)ds \right), \quad t \in (0,T]. \tag{2.2.18}$$

Proof. It follows from equation (2.2.17) of Proposition 2.2.9 that $z(\bullet,T)$ is a local martingale if and only if the process

$$\int_0^t z(s-,T) \left(A(s,T) + \Psi_s(\Sigma(s,T)) \right) ds$$

vanishes, because a special semimartingale is a local martingale if and only if the finite variation part vanishes. Applying Lemma 2.2.10 (note in particular that $\Psi_t(\Sigma(t,T))$ has left-continuous paths) yields the assertion. □

Note that the exceptional $\mathbb{P}$-null set, where (2.2.18) does not hold, may depend on $T \in (0, \infty)$. Thus, we cannot simply differentiate with respect to T. Nevertheless, presuming we are allowed to (for instance if α, σ and Ψ are deterministic, or later, in the proof of Lemma 4.1.10), differentiating - as we have mentioned above, cumulant generating functions are of class C^∞ - yields the condition

$$\alpha(t,T) = -\Big\langle \sigma(t,T), \nabla\Psi_t \left(-\int_t^T \sigma(t,s)ds \right) \Big\rangle \tag{2.2.19}$$

on the drift. Let us look at a few examples. For a n-dimensional driving standard Wiener process X, the cumulant generating function and its gradient are given by

$$\Psi(z) = \frac{1}{2}|z|^2, \quad \nabla\Psi(z) = z,$$

and equation (2.2.19) becomes

$$\alpha(t,T) = \sum_{i=1}^n \sigma_i(t,T) \int_t^T \sigma_i(t,s)ds,$$

which is the classical drift condition known from Heath, Jarrow, and Morton [36]. Henceforth, we will also refer to equation (2.2.19) as HJM drift condition.

As a slight generalization of the preceding case, assume the Grigelionis process X has ($\mathbb{P} - a.s.$) only continuous paths. According to Lemma 2.2.7, it holds $K = 0$ ($\mathbb{P} \otimes \text{Leb}$)-almost surely, where K stems from the derivative $(\beta, c, K)^D$. Noting that the matrices c_t are symmetric, it thus holds

$$\begin{aligned} \Psi_t(z) &= \langle \beta_t, z\rangle + \frac{1}{2}\langle z, c_t z\rangle, \\ \nabla\Psi_t(z) &= \beta_t + \big(\langle c_{1\bullet}(t), z\rangle, \ldots, \langle c_{n\bullet}(t), z\rangle\big), \end{aligned}$$

and we obtain the HJM drift condition

$$\alpha(t,T) = -\langle \sigma(t,T), \beta_t\rangle + \sum_{i,j=1}^n c_{ij}(t)\sigma_i(t,T) \int_t^T \sigma_j(t,s)ds.$$

If the driving process X is a PIIAC in the terminology of Eberlein and Kluge [27], and the volatility structure is deterministic, equation (2.2.18) is just the equation from Prop. 9 of the just mentioned paper. It is also in accordance with equation (5.16) in Björk, Di Masi, Kabanov, and Runggaldier [14].

Let W be a m-dimensional standard Wiener process, and let A be a n-dimensional finite variation process of the form $A_t = \sum_{s \le t} \Delta A_s = x * \mu^A_t$. Consider the interest rate model

$$\left\{ \begin{array}{rcl} df(t,T) &=& \alpha(t,T)dt + \sigma(t,T)dW_t + \eta(t,T)dA_t, \\ f(0,T) &=& f^*(0,T) \end{array} \right. .$$

The drift β from the derivative $(\beta, 0, K)^D$ of A is given by $\beta_t^j = \int_{\mathbb{R}^n} x_j K_t(dx)$ for $j = 1, \dots, n$, which is seen from the canonical decomposition of A. Concerning the cumulant generating functions of $X = (W, A)$, it therefore holds

$$\begin{aligned}
\Psi_t(w, z) &= \frac{1}{2}|w|^2 + \int_{\mathbb{R}^n} \big(e^{\langle z, x\rangle} - 1\big) K_t(dx), \\
\frac{\partial}{\partial w_i}\Psi_t(w, z) &= w_i, \quad i = 1, \dots, m, \\
\frac{\partial}{\partial z_j}\Psi_t(w, z) &= \int_{\mathbb{R}^n} x_i e^{\langle z, x\rangle} K_t(dx), \quad j = 1, \dots, n.
\end{aligned}$$

We obtain the HJM drift condition

$$\begin{aligned}
\alpha(t, T) = &\sum_{i=1}^m \sigma_i(t, T) \int_t^T \sigma_i(t, s) ds \\
&- \int_{\mathbb{R}^n} \langle x, \eta(t, T)\rangle \exp\left(-\left\langle x, \int_t^T \eta(t, s) ds\right\rangle\right) K_t(dx).
\end{aligned}$$

This equation coincides with the drift condition (28) in Björk, Kabanov, and Runggaldier [15, Prop. 3.15] and with equation (2.12) in Gapeev and Küchler [35].

Finally, we look at two concrete examples of driving processes, where the HJM drift condition is given by an explicit expression. Let $f(t, T)$ be a term structure model

$$\begin{cases} df(t, T) &= \alpha(t, T)dt + \sigma(t, T)dW_t + \eta(t, T)dN_t, \\ f(0, T) &= f^*(0, T) \end{cases},$$

where W denotes a m-dimensional Brownian motion, and where N consists of n independent Poisson processes Jacod and Shiryaev [42, Def. I.3.26] with intensities $\mathbb{E}[N_t^j] = \int_0^t \lambda_j(s) ds$. Then it holds for the cumulant generating functions of the driving process $X = (W, N)$

$$\begin{aligned}
\Psi_t(w, z) &= \frac{1}{2}|w|^2 + \sum_{j=1}^n \lambda_j(t)\big(e^{z_j} - 1\big), \\
\nabla\Psi_t(w, z) &= \big(w, \lambda_1(t)e^{z_1}, \dots, \lambda_n(t)e^{z_n}\big).
\end{aligned}$$

We derive, cf., e.g., equation (2) in Hyll [40], the HJM drift condition

$$\alpha(t, T) = \sum_{i=1}^m \sigma_i(t, T) \int_t^T \sigma_i(t, s) ds - \sum_{j=1}^n \lambda_j(t)\eta_j(t, T) \exp\left(-\int_t^T \eta_j(t, s) ds\right).$$

Term structure models driven by bilateral Gamma processes are studied in Küchler and Naumann [46, Sec. 5]. The cumulant generating function and its derivative of a bilateral Gamma process X with parameters $\alpha^+, \alpha^-, \lambda^+, \lambda^- > 0$ are given by

$$\Psi(z) = \alpha^+ \ln\left(\frac{\lambda^+}{\lambda^+ - z}\right) + \alpha^- \ln\left(\frac{\lambda^-}{\lambda^- + z}\right),$$
$$\Psi'(z) = \frac{\alpha^+}{\lambda^+ - z} - \frac{\alpha^-}{\lambda^- + z}.$$

This yields the HJM drift condition

$$\alpha(t,T) = -\sigma(t,T)\left(\frac{\alpha^+}{\lambda^+ + \int_t^T \sigma(t,s)ds} - \frac{\alpha^-}{\lambda^- - \int_t^T \sigma(t,s)ds}\right).$$

Note that it necessarily holds $Q \subset (-\lambda^-, \lambda^+)$, where Q denotes the compact set from Assumptions 2.2.6.

Sometimes, we use the so-called Musiela parametrization (see Brace and Musiela [17] and Musiela [53]) of forward rates $r_t(x) = f(t, t+x)$. It was originally formulated for term structure models driven by Brownian motions, but it is also valid in our setup.

2.2.12 Proposition. *For each $x \in \mathbb{R}_+$, the forward rates $r_t(x) = f(t, t+x)$ satisfy*

$$\begin{cases} dr_t(x) &= \left[\frac{\partial}{\partial x} r_t(x) + \alpha(t, t+x)\right] dt + \sigma(t, t+x) dX_t \\ r_0(x) &= f^*(0, x) \end{cases}.$$

Proof. See the appendix. □

Chapter 3

Finite dimensional realizations

Considering term structure models where the driving processes are allowed to make jumps, we treat in this chapter the existence of finite dimensional realizations by an extension of the Lie algebraic method, which was applied in Björk and Svensson [13] for driving Brownian motions.

In order to make the following definitions and results better understandable, we start by illustrating the relevant geometric ideas of the Lie algebraic theory. A survey about the geometric ideas for driving Brownian motion can be found in Björk [8]. We describe these ideas and extend the geometry to driving processes with jumps. Our goal is then the establishment of Theorem 3.1.17 and Theorem 3.1.21, two results, which will later provide formulas for the coefficients of forward rate equations admitting realizations.

Then, using Lie algebras, we give a necessary criterion for the existence of finite dimensional realizations for infinite dimensional equations with values in a Banach space. We also argue that the converse implication is not valid.

The applications to finance start in the third section, where we embed term structure models in the framework of infinite dimensional Banach space valued equations, and apply the results derived in the section before.

The rest of the chapter is devoted to the study of concrete volatility structures. First, we consider the case of deterministic volatility. The more interesting case of deterministic direction volatility is studied afterwards.

3.1 Geometric background and preparatory results

The first part of this section shall provide an illustration of the Lie algebraic method. This theory deals with the existence of finite dimensional realizations of Banach space valued equations. As we shall see in Section 3.3, term structure models can be incorporated into this framework.

Let $\mathcal{X}$ be a real Banach space of dimension $n \in \mathbb{N} \cup \{\infty\}$. In applications to term structure models, $\mathcal{X}$ is of course infinite dimensional, but for the sake of illustration, one may as well think of $\mathbb{R}^n$. In the sequel, we use the symbol $\mathbb{F}$ for Fréchet derivatives (see, e.g., Sec. III.5 in Werner [62]).

3.1.1 Definition.

1. *Let $U \subset \mathcal{X}$ be an open subset. A vector field on U is a mapping $f : U \to \mathcal{X}$. It is called smooth if it is infinitely often Fréchet differentiable.*

2. *For smooth vector fields f, g on U, the Lie bracket $[f, g]$ is the vector field on U, which is defined by*

$$[f, g](x) := \mathbb{F}f(x)[g(x)] - \mathbb{F}g(x)[f(x)], \quad x \in U.$$

3. *For smooth vector fields $f_1, \ldots, f_n$ on U, we define the Lie algebra generated by $f_1, \ldots, f_n$ as*

$$\{f_1, \ldots, f_n\}_{LA} := \bigcap_{L \in \mathcal{L}} L,$$

where $\mathcal{L}$ denotes the set of all linear spaces L (over $\mathbb{R}$) of smooth vector fields on U with the properties

- *$f_1, \ldots, f_n \in L$.*
- *If $g, h \in L$, so is the Lie bracket $[g, h] \in L$.*

The dimension of the Lie algebra $F = \{f_1, \ldots, f_n\}_{LA}$ at some point $x \in U$ is defined as the dimension of the linear space $\{f(x) \mid f \in F\} \subset \mathcal{X}$.

3.1.2 Remark. *A collection $f_1, \ldots, f_n$ of smooth vector fields on U provides a distribution F in the terminology of Björk and Svensson [13, Sec. 2], and the Lie algebra generated by $f_1, \ldots, f_n$ is the minimal involutive (i.e. closed under the Lie bracket) distribution containing F.*

The intuitive meanings of Lie brackets and Lie algebras will be explained below. We emphasize that the subsequent illustrations have purely motivational character, and are therefore written in an informal manner. In particular, the cited results, like the Frobenius theorem, are in general just valid in a local sense. For the sake of simplicity, we consider, in this informal passage, only time-homogeneous equations with one driving process.

Let $r = (r_t)_{t \in \mathbb{R}_+}$ be a $\mathcal{X}$-valued stochastic process satisfying

$$\begin{cases} dr_t &= \alpha(r_t)dt + \sigma(r_{t-})dX_t \\ r_0 &= r^* \end{cases},$$

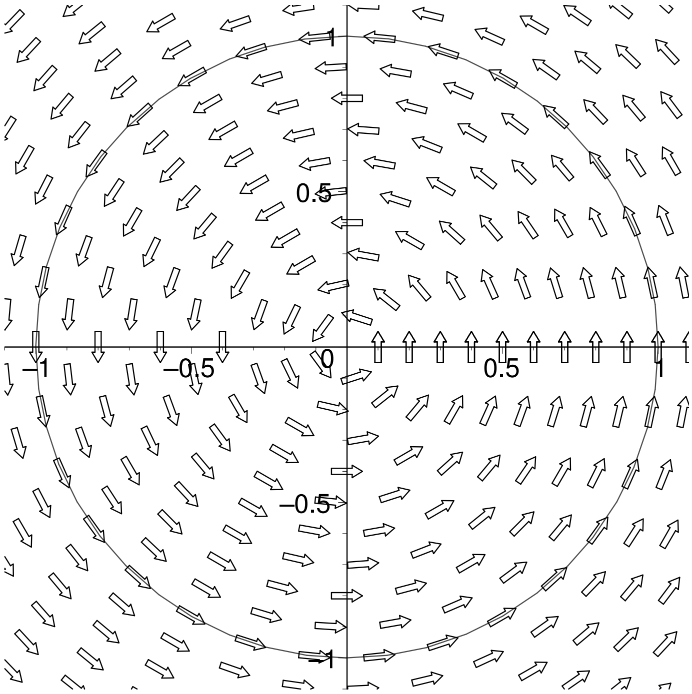

Figure 3.1.1: The vector field from equation (3.1.1) and a tangential manifold.

where $r^* \in \mathcal{X}$ and α, σ are smooth vector fields on $\mathcal{X}$. The driving process X denotes a one-dimensional semimartingale.

Fix an integer $d \in \mathbb{N}$ such that $d < n$. Recall that $n \in \mathbb{N} \cup \{\infty\}$ denotes the dimension of the Banach space $\mathcal{X}$. We are interested in finding a d-dimensional realization, i.e. a pair (G, Z), where G denotes a mapping $G : \mathbb{R}^d \to \mathcal{X}$, and Z a d-dimensional process such that $r_t = G(Z_t)$, $t \in \mathbb{R}_+$.

Consider for instance the following differential equation in $\mathcal{X} = \mathbb{R}^2$.

$$\begin{cases} dr_t &= (-r_2(t), r_1(t))dt \\ r_0 &= (1, 0) \end{cases} . \tag{3.1.1}$$

Note that the volatility vector field is equal to zero, whence (3.1.1) is an ordinary differential equation. It has the solution $r_t = (\cos t, \sin t)$. Hence, the existence of a one-dimensional realization is evident. The deterministic process r and the drift vector field are visualized in Figure 3.1.1.

Returning to the general realization problem, we rely on the following two facts. The terms "invariant" and "tangential" will be explained below.

3.1.3 Remarks.

1. *There exists a d-dimensional realization if and only if there exists a d-dimensional invariant manifold $\mathcal{G} \subset \mathcal{X}$.*

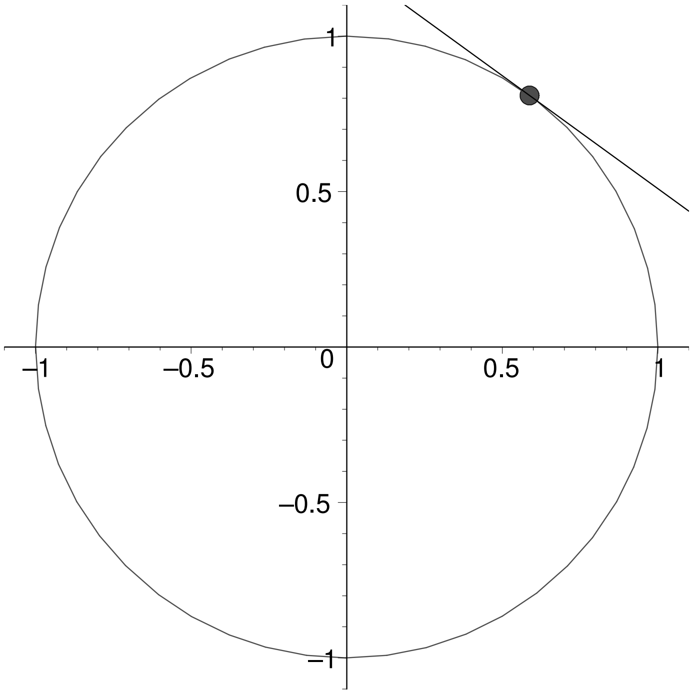

Figure 3.1.2: The tangential manifold of Figure 3.1.1 and one tangent space.

2. *There exists a d-dimensional tangential manifold* $\mathcal{G} \subset \mathcal{X}$ *if and only if the dimension of the Lie algebra generated by* α *and* σ *is at most d, i.e.* $\dim\{\alpha, \sigma\}_{LA} \leq d$.

By an invariant manifold $\mathcal{G}$, we mean a manifold with $r^* \in \mathcal{G}$ and $r_t \in \mathcal{G}$ for all $t \in \mathbb{R}_+$, i.e. the process r starts in the manifold $\mathcal{G}$, and never leaves the manifold. Hence, the term "invariant" means invariant with respect to the process r. We see from Figure 3.1.1 that $\mathcal{G} = \{r \in \mathbb{R}^2 \,|\, |r| = 1\}$ is an invariant manifold for the process r given by equation (3.1.1).

A manifold $\mathcal{G}$ is said to be tangential if $\alpha(r), \sigma(r) \in \mathcal{T}_{\mathcal{G}}(r)$ for all $r \in \mathcal{G}$, where $\mathcal{T}_{\mathcal{G}}(r)$ denotes the tangent space of the manifold $\mathcal{G}$ at point r. Thus, the term "tangential" is to be understood in the sense tangential with respect to the drift α and the volatility σ. Figure 3.1.2 shows the tangent space of the manifold $\mathcal{G} = \{r \in \mathbb{R}^2 \,|\, |r| = 1\}$ at the marked point. Regarding Figure 3.1.1, we observe that $\mathcal{G}$ is not only invariant, but also tangential.

While the first statement of Remarks 3.1.3 is geometrically obvious, the second statement relies on the Frobenius theorem, which is a classical result from differential geometry that was extended to the infinite dimensional case by Björk and Svensson [13].

We briefly explain the geometric intuition of this theorem. A more detailed survey of the geometric ideas can be found in Björk [8].

It is, at first glance, tempting to think that there exists always a two-dimensional tangential manifold, i.e. a manifold $\mathcal{G}$ such that $\alpha(r), \sigma(r) \in \mathcal{T}_{\mathcal{G}}(r)$ for all $r \in \mathcal{G}$, and that it can be constructed as follows. Assuming, for the sake of geometric intuition, that α and σ are vector fields in the Euclidian space $\mathbb{R}^3$, we start by solving the differential equation

$$\begin{cases} F'(t) &= \alpha(F(t)) \\ F(0) &= r^* \end{cases},$$

which yields a curve in $\mathbb{R}^3$. For every fixed point $F(t)$ on this curve, we solve the differential equation

$$\begin{cases} \frac{\partial}{\partial s} G(t,s) &= \sigma(G(t,s)) \\ G(t,0) &= F(t) \end{cases},$$

and obtain a surface in $\mathbb{R}^3$ parameterized by G. In the special case where α and σ are linear vector fields in $\mathbb{R}^3$, i.e. $\alpha(r) = Ar$ and $\sigma(r) = Br$ with $A, B \in \mathbb{R}^{3\times 3}$, the parametrization of the surface is given by

$$G(t,s) = \underbrace{e^{Bs} \underbrace{\left(e^{At} r^*\right)}_{\text{curve in } \mathbb{R}^3}}_{\text{surface in } \mathbb{R}^3} . \tag{3.1.2}$$

However, the thus derived manifold $\mathcal{G}$ is, in general, not a tangential manifold. One reason, which denies the tangency, is that the parametrization of the manifold should also be obtained by executing the geometric program above in reverse order. In the case of linear vector fields, we get the parametrization

$$H(t,s) = \underbrace{e^{As} \underbrace{\left(e^{Bt} r^*\right)}_{\text{curve in } \mathbb{R}^3}}_{\text{surface in } \mathbb{R}^3} , \tag{3.1.3}$$

and the corresponding manifold $\mathcal{H}$ usually differs from $\mathcal{G}$. We observe a certain lack of commutativity in the geometric program above, and this lack of commutativity is measured by the Lie bracket. This is the meaning of the Lie bracket introduced in Definition 3.1.1, which exhibits whether we can find a two-dimensional tangential manifold, namely if and only if $[\alpha, \sigma] = 0$. For linear vector fields, the Lie bracket is just $[\alpha, \sigma] = AB - BA$, and indeed, in the case where the two matrices A and B are commutative, the two manifold $\mathcal{G}$ and $\mathcal{H}$, parameterized by (3.1.2) and (3.1.3), coincide.

In general, if we are in $\mathbb{R}^n$ or a Banach space of infinite dimension, the Frobenius theorem tells us that there exists a tangential manifold with dimension at most d, if and only if $\dim\{\alpha, \sigma\}_{LA} \leq d$. It can be found by choosing a generating system $f_1, \ldots, f_m$ of $\{\alpha, \sigma\}_{LA}$, and then executing the geometric program described above with the vector fields $f_1, \dots, f_m$. This yields a tangential manifold, regardless of the order, in which we perform the construction.

With regard to Remarks 3.1.3 it arises the question when a tangential manifold is invariant, and vice versa. For deterministic equations, like (3.1.1), these two properties, tangency and invariance of a manifold, are indeed equivalent, see also Figure 3.1.1. The next step is to check what happens for a driving standard Wiener process, consider for instance

$$\begin{cases} dr_t &= (-r_2(t), r_1(t))dW_t \\ r_0 &= (1,0) \end{cases}, \tag{3.1.4}$$

which is equation (3.1.1) with dt replaced by dW_t. A natural guess is that, because of the continuity of the Wiener process, a manifold $\mathcal{G}$ is tangential if and only if it is invariant, and that a realization for the process r satisfying (3.1.4) is given by $(\cos W_t, \sin W_t)$. This conjecture, however, is not correct, since an application of Itô's formula on $(\cos W_t, \sin W_t)$ yields a drift term consisting of second order derivatives. This gap can be fixed by changing to Fisk-Stratonovich dynamics, i.e. consider instead

$$\begin{cases} dr_t &= (-r_2(t), r_1(t)) \circ dW_t \\ r_0 &= (1,0) \end{cases}, \tag{3.1.5}$$

where $\circ dW_t$ denotes the Fisk-Stratonovich integral, which will be formally introduced later (Definition 3.1.19). Indeed, the process r from (3.1.5) can be realized as $r_t = (\cos W_t, \sin W_t)$. Since we can always, by changing the drift term, convert Itô dynamics to Fisk-Stratonovich dynamics, and vice versa, considering Fisk-Stratonovich differentials is no restriction. We refer to Björk and Christensen [10] for more details on this topic.

We have now, on an intuitive level, argued that the main result of Björk and Svensson [13] is valid, namely that there exists a d-dimensional realization if and only if

$$\dim\{\alpha, \sigma\}_{LA} \leq d.$$

In this text, we are mainly interested in driving semimartingales that are allowed to make jumps. As we shall see, the appearance of jumps causes new effects, and the geometry described above changes notably.

We go on to illustrate, still using informal arguments, that for driving processes with jumps, only one implication of the just cited theorem holds, namely that the existence of a d-dimensional realization implies $\dim\{\alpha, \sigma\}_{LA} \leq d$, and that, for this implication, additional assumptions on the semimartingale X are required.

In view of Remarks 3.1.3 we have to check whether it is still true that a manifold $\mathcal{G}$ is tangential if and only if it is invariant. Consider the equation

$$\begin{cases} dr_t &= (-r_2(t), r_1(t))d(t + N_t) \\ r_0 &= (1,0) \end{cases}, \tag{3.1.6}$$

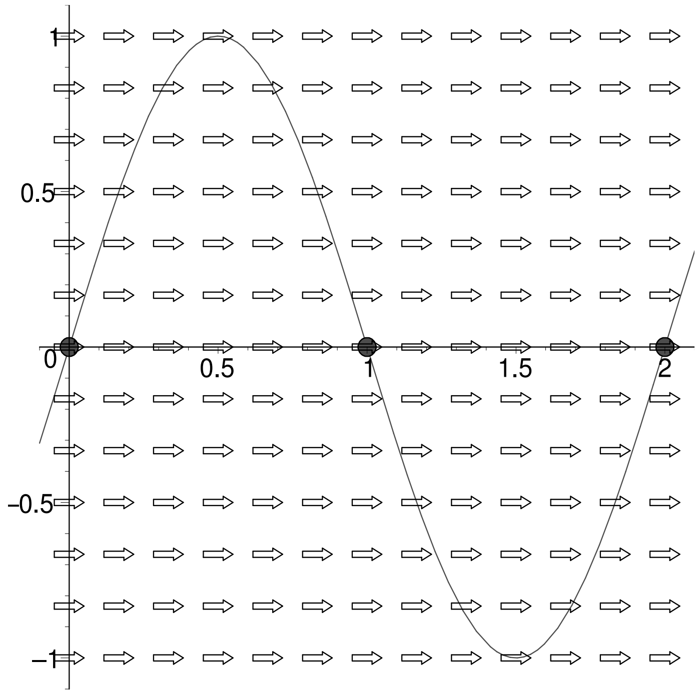

Figure 3.1.3: The vector field from equation (3.1.7) with an invariant manifold.

where N denotes a Poisson process. The manifold $\mathcal{G} = \{r \in \mathbb{R}^2 \,|\, |r| = 1\}$ provides a one-dimensional tangential manifold. Figure 3.1.2 shows the tangent space of $\mathcal{G}$ at the marked point. Assume the first jump of the Poisson process N occurs at this point. Then, the process r makes a jump in tangential direction, and consequently, it leaves the manifold $\mathcal{G}$. This example shows that a tangential manifold is, in general, not invariant. We note, however, that invariance would be implied if the tangential manifold were affine. But in general, the implication from tangency to invariance fails. It was already mentioned in Björk and Christensen [10, Sec. 6], where the situation was studied with an additional marked point process, that, besides tangency, the manifold must fulfill an extra condition, namely that it is invariant regarding all possible jumps.

Concerning the implication from invariance to tangency, consider the following equation

$$\begin{cases} dr_t &= (1,0)dN_t \\ r_0 &= (0,0) \end{cases}, \tag{3.1.7}$$

where N denotes a Poisson process. The image of the process r consists of all points $\{(i,0) \,|\, i \in \mathbb{N}_0\}$, as indicated in Figure 3.1.3.

The existence of a one-dimensional realization is apparent. There are several one-dimensional invariant manifolds, and one is plotted in Figure 3.1.3. It is evident that this invariant manifold $\mathcal{G}$ is not tangential. However, the reason, why things go wrong, is the Poisson process, which only makes jumps of size one. Replace it by a process with the property that it makes arbitrary small jumps, take for instance a Lévy process with Lévy measure K satisfying $0 \in \overline{\text{int supp}(K)}$. Then the effect shown in Figure 3.1.3, which is due to the vacancies left by the Poisson process, cannot occur, and it is plausible that invariance implies tangency.

If we insist on jump processes like the Poisson process, and impose the additional condition that the invariant manifold $\mathcal{G}$ is affine, then it is also geometrically obvious that tangency is implied. In our example (3.1.7), the condition that $\mathcal{G}$ shall be affine implies that $\mathcal{G} = \{(x, 0) \,|\, x \in \mathbb{R}\}$, which is a tangential manifold.

To sum up, we have discussed that, if the driving processes X has jumps, only one implication holds, namely if there exists a d-dimensional realization (G, Z), it holds $\dim\{\alpha, \sigma\}_{LA} \leq d$. To establish this result, we have to impose extra conditions on the semimartingale X. We need it to make arbitrary small jumps, for this reason we will introduce non-degenerate Grigelionis processes in Definition 3.1.15. Supposing it is already known that the mapping G is affine, we essentially only need that the process X makes jumps, that is why we will define linearly non-degenerate Grigelionis processes in Definition 3.1.15.

After these geometric explanations of the Lie algebraic theory, we shall turn back to the development of precise mathematical results. In this section, we proceed as follows. Next, we prove some auxiliary results, then we give the required definitions, which are motivated by the geometric aspects above, and finally arrive at Theorem 3.1.17 and Theorem 3.1.21, which provide formulas for the drift and the volatilities of forward rate equations admitting realizations.

For any measure μ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, its support $\text{supp}(\mu)$ is defined to be the set of all $x \in \mathbb{R}^n$ such that $\mu(U) > 0$ for any open set $U \subset \mathbb{R}^n$ containing x. The support $\text{supp}(\mu)$ is a closed set. For any random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, the support of $\mathbb{P}^X$ is called the support of X, denoted by $\text{supp}(X)$. It is the smallest closed set $F \subset \mathbb{R}^n$ satisfying $\mathbb{P}(X \in F) = 1$.

3.1.4 Lemma. *Let $n, m \in \mathbb{N}$. Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ be a random variable and $f, g \in C(\mathbb{R}^n, \mathbb{R}^m)$. There is equivalence between*

1. *$f \circ X = g \circ X$ $(\mathbb{P} - a.s.)$;*
2. *$f(x) = g(x)$ for all $x \in \text{supp}(X)$.*

Proof. See the appendix. □

3.1.5 Lemma. *Let $n, m \in \mathbb{N}$. Let $f \in C(\mathbb{R}^n, \mathbb{R}^m)$ and μ be a measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ such that $f = 0$ $(\mu - a.s.)$. Then it holds $f(x) = 0$ for all $x \in \text{supp}(\mu)$.*

Proof. See the appendix. □

3.1.6 Lemma. *Let Z be a d-dimensional semimartingale and $W_1, W_2 \in C(\mathbb{R}_+ \times \mathbb{R}^d)$ such that $\int_0^t W_1(s, Z_{s-})ds = \int_0^t W_2(s, Z_{s-})ds$ up to evanescence. Then, it holds*

$$W_1(t, z) = W_2(t, z), \quad t \in (0, \infty), z \in \text{supp}(Z_{t-}).$$

Proof. See the appendix. □

Let $n \in \mathbb{N}$. The set $\mathcal{M}^+(\mathbb{R}^{n\times n})$ denotes the set of all symmetric, non-negative definite $n\times n$-matrices, whereas $\mathcal{M}^{++}(\mathbb{R}^{n\times n})$ is the set of all symmetric, positive definite $n \times n$-matrices. By $\mathcal{K}^+(\mathbb{R}^n)$ we denote the set of all measures K on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ satisfying $K(\{0\}) = 0$ and $\int_{\mathbb{R}^n}(|x|^2 \wedge 1)K(dx) < \infty$. The set $\mathcal{K}^{++}(\mathbb{R}^n)$ consists of all $K \in \mathcal{K}^+(\mathbb{R}^n)$ with exception of the zero measure.

3.1.7 Definition. *Assume, given a Grigelionis process X with derivative $(\tilde{\beta}, \tilde{c}, \tilde{K})^D$, there exists a d-dimensional semimartingale Z and mappings*

$$\begin{cases} \beta & : & \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^n \\ c & : & \mathbb{R}_+ \times \mathbb{R}^d \to \mathcal{M}^+(\mathbb{R}^{n\times n}) \\ K & : & \mathbb{R}_+ \times \mathbb{R}^d \to \mathcal{K}^+(\mathbb{R}^n) \end{cases} ,$$

such that, up to evanescence,

$$\begin{cases} \tilde{\beta}_t(\omega) & = & \beta(t, Z_{t-}(\omega)) \\ \tilde{c}_t(\omega) & = & c(t, Z_{t-}(\omega)) \\ \tilde{K}_{(\omega,t)} & = & K(t, Z_{t-}(\omega)) \end{cases} .$$

Then, we call $(\beta, c, K; Z)^D$ a Z-derivative of X.

If the mappings β, c, K can be chosen such that

$$\begin{cases} \beta & : & \mathbb{R}_+ \to \mathbb{R}^n \\ c & : & \mathbb{R}_+ \to \mathcal{M}^+(\mathbb{R}^{n\times n}) \\ K & : & \mathbb{R}_+ \to \mathcal{K}^+(\mathbb{R}^n) \end{cases} ,$$

i.e. the choice of the semimartingale Z does not matter, we also say that the derivative is deterministic, and denote it by $(\beta, c, K)^D$, which is consistent with Definition 2.1.6. This situation occurs in particular, if X is a Lévy process, or, more general, a so-called PIIAC in the sense of Eberlein, Jacod, and Raible [30].

3.1.8 Definition. *Let X be a Grigelionis process with Z-derivative $(\beta, c, K; Z)^D$. We say that the derivative is of type (C) if for all $(t, z) \in (0, \infty) \times \mathbb{R}^d$*

$$\begin{cases} \beta(t, z) & = & 0 \\ c(t, z) & \in & \mathcal{M}^{++}(\mathbb{R}^{n\times n}) \\ K(t, z) & = & 0 \end{cases} ,$$

and it holds $c \in C(\mathbb{R}_+, \mathbb{R}^{n\times n})$.

3.1.9 Definition. *Let X be a Grigelionis process with Z-derivative $(\beta, c, K; Z)^D$. We say that the derivative is of type (D) if for all $(t,z) \in (0,\infty) \times \mathbb{R}^d$*

$$\begin{cases} \beta(t,z) & = & 0 \\ c(t,z) & = & 0 \\ K(t,z) & \in & \mathcal{K}^{++}(\mathbb{R}^n) \end{cases}.$$

When dealing with finite dimensional realizations, we need some assumptions about supports. They are mild, but indispensable.

3.1.10 Definition. *Let X be a Grigelionis process with Z-derivative $(\beta, c, K; Z)^D$. We say that (X, Z) has regular supports if the following conditions are satisfied.*

1. *It holds $\operatorname{supp} K(t, z_1) = \operatorname{supp} K(t, z_2)$ for all $t \in (0,\infty)$ and all $z_1, z_2 \in \mathbb{R}^d$. For every $t \in (0,\infty)$, we may therefore abbreviate $\operatorname{supp}(K_t)$.*
2. *Defining $S_X := \{(t,x) \,|\, t \in (0,\infty), x \in \operatorname{supp}(K_t)\}$, there exists, for each $(t,x) \in S_X$, an $\varepsilon > 0$ such that $\{(s,x) \,|\, s \in (t-\varepsilon, t+\varepsilon)\} \subset S_X$.*
3. *Defining $S_Z := \{(t,z) \,|\, t \in (0,\infty), z \in \operatorname{supp}(Z_{t-})\}$, there exists, for each $(t,z) \in S_Z$, an $\varepsilon > 0$ such that $\{(s,z) \,|\, s \in (t-\varepsilon, t+\varepsilon)\} \subset S_Z$.*

3.1.11 Lemma. *Let X be a Grigelionis process with derivative $(0, c, 0; Z)^D$ of type (C), and $W_1, W_2 \in C(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^n)$ be such that $\int_0^t W_1(s, Z_{s-})dX_s = \int_0^t W_2(s, Z_{s-})dX_s$ up to evanescence. Then, it holds*

$$W_1(t,z) = W_2(t,z), \quad t \in (0,\infty), z \in \operatorname{supp}(Z_{t-}).$$

Proof. Set $W := W_1 - W_2$, which is again of class $C(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^n)$. By Jacod and Shiryaev [42, Thm. I.4.40.d] the relation

$$\left\langle \int_0^\bullet W(s, Z_{s-})dX_s, \int_0^\bullet W(s, Z_{s-})dX_s \right\rangle_t = \int_0^t \langle W(s, Z_{s-}), c(s, Z_{s-})W(s, Z_{s-})\rangle ds$$

is valid. Thus, it holds $\int_0^t \langle W(s, Z_{s-}), c(s, Z_{s-})W(s, Z_{s-})\rangle ds = 0$ up to evanescence. Applying Lemma 3.1.6 we obtain

$$\langle W(t,z), c(t,z)W(t,z)\rangle = 0, \quad t \in (0,\infty), z \in \operatorname{supp}(Z_{t-}).$$

Since the matrices $c(t,z)$ are positive definite, we are done. □

The space $G_{\mathrm{loc}}(\mu^X)$ denotes the linear space of $(\mu^X - \nu)$-integrable functions, and for any $W \in G_{\mathrm{loc}}(\mu^X)$ we denote by $W * (\mu^X - \nu)$ the stochastic integral of W with respect to $\mu^X - \nu$ Jacod and Shiryaev [42, Def. II.1.27].

3.1.12 Lemma. *Let X be a n-dimensional Grigelionis process with derivative given by $(0,0,K)^D$, and $W \in G_{\text{loc}}(\mu^X)$. Then, the process $Y := W * (\mu^X - \nu)$ is a one-dimensional Grigelionis process with derivative $(0,0,K^Y)^D$, where, for fixed $(\omega,t) \in \Omega \times \mathbb{R}_+$, the measure $K^Y_{(\omega,t)}$ is given by $K^Y_{(\omega,t)}(dx) := \mathbf{1}_{\mathbb{R}\setminus\{0\}}(x) K^W_{(\omega,t)}(dx)$, with $K^W_{(\omega,t)}$ denoting the image of the measure $K_{(\omega,t)}$ under the mapping $x \mapsto W(\omega,t,x)\mathbf{1}_{\mathbb{R}^n\setminus\{0\}}(x)$.*

Proof. We define the random measure $\nu^Y(dt,dx) := dt\, K^Y_t(dx)$. To any $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})$-measurable function V, we associate the function $V_W : \Omega \times \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$, defined as

$$V_W(\omega,t,x) := V(\omega,t,W(\omega,t,x)\mathbf{1}_{\mathbb{R}^n\setminus\{0\}}(x))\mathbf{1}_{\mathbb{R}\setminus\{0\}}(W(\omega,t,x))\mathbf{1}_{\mathbb{R}^n\setminus\{0\}}(x),$$

which is $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^n)$-measurable, and if in addition, V is non-negative, then V_W is non-negative, too. Since the compensator of μ^X is of the absolutely continuous form $\nu(dt,dx) = dt\, K_t(dx)$, it holds $\nu(\omega;\{t\} \times \mathbb{R}^n) = 0$ for all $(\omega,t) \in \Omega \times \mathbb{R}_+$ (up to evanescence). Consequently, by the definition of the stochastic integral $W * (\mu^X - \nu)$ Jacod and Shiryaev [42, Def. II.1.27], the jumps of the purely discontinuous local martingale Y are given by

$$\Delta Y_t(\omega) = W(\omega,t,\Delta X_t(\omega))\mathbf{1}_{\mathbb{R}^n\setminus\{0\}}(\Delta X_t(\omega)).$$

We deduce for any $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})$-measurable function V

$$\begin{aligned} V(\omega,t,x) * \mu^Y = \sum_{0<s\le t} & V(\omega,s,W(\omega,s,\Delta X_s(\omega))\mathbf{1}_{\mathbb{R}^n\setminus\{0\}}(\Delta X_s(\omega))) \\ & \mathbf{1}_{\mathbb{R}\setminus\{0\}}(W(\omega,s,\Delta X_s(\omega)))\mathbf{1}_{\mathbb{R}^n\setminus\{0\}}(\Delta X_s(\omega)) = V_W(\omega,t,x) * \mu^X, \end{aligned} \tag{3.1.8}$$

as well as

$$\begin{aligned} V(\omega,t,x) * \nu^Y = \int_0^t \int_{\mathbb{R}^n} & V(\omega,s,W(\omega,s,x)\mathbf{1}_{\mathbb{R}^n\setminus\{0\}}(x)) \\ & \mathbf{1}_{\mathbb{R}\setminus\{0\}}(W(\omega,s,x))\mathbf{1}_{\mathbb{R}^n\setminus\{0\}}(x) K_{(\omega,t)}(dx)ds = V_W(\omega,t,x) * \nu. \end{aligned} \tag{3.1.9}$$

Since ν is predictable, equation (3.1.9) yields that the integral process $V * \nu^Y$ is predictable for any predictable function V. Thus, the random measure ν^Y is predictable. Since ν is the compensator of μ^X, we infer from (3.1.8) and (3.1.9) that for any non-negative $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})$-measurable function V it holds

$$\mathbb{E}\big[V * \mu^Y_\infty\big] = \mathbb{E}\big[V_W * \mu^X_\infty\big] = \mathbb{E}\big[V_W * \nu_\infty\big] = \mathbb{E}\big[V * \nu^Y_\infty\big].$$

Thus, ν^Y is the compensator of μ^Y. □

3.1.13 Lemma. *Let X be a Grigelionis process with derivative $(0,0,K;Z)^D$ of type (D) such that (X,Z) has regular supports, and let $W_1, W_2 \in C(\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^n)$ be such that*

$$\int_0^t \int_{\mathbb{R}^n} W_1(s,Z_{s-},x)(\mu^X - \nu)(ds,dx) = \int_0^t \int_{\mathbb{R}^n} W_2(s,Z_{s-},x)(\mu^X - \nu)(ds,dx)$$

up to evanescence. Then it holds

$$W_1(t,z,x) = W_2(t,z,x), \quad t \in (0,\infty), z \in \operatorname{supp}(Z_{t-}), x \in \operatorname{supp}(K_t).$$

Proof. Set $W := W_1 - W_2$, which is again of class $C(\mathbb{R}_+ \times \mathbb{R}^d \times \mathbb{R}^n)$. Denote by ν^Y the compensator of the random measure μ^Y of jumps of the process Y defined as

$$Y_t := \int_0^t \int_{\mathbb{R}^n} W(s, Z_{s-}, x)(\mu^X - \nu)(ds, dx).$$

According to Lemma 3.1.12 (note that $K_{t,z}(\{0\}) = 0$ for all measures $K_{t,z}$) we obtain

$$\nu^Y(\omega; [0,t] \times \mathbb{R}^n) = \int_0^t \int_{\mathbb{R}^n} \mathbf{1}_{\mathbb{R}\setminus\{0\}}(W(s, Z_{s-}(\omega), x)) K_{s, Z_{s-}(\omega)}(dx) ds.$$

Thus, there exists a $\mathbb{P}$-null set A such that the paths $t \mapsto Z_{t-}(\omega)$ are left-continuous for each $\omega \in \Omega \setminus A$ and

$$\int_0^t \int_{\mathbb{R}^n} \mathbf{1}_{\mathbb{R}\setminus\{0\}}(W(s, Z_{s-}(\omega), x)) K_{s, Z_{s-}(\omega)}(dx) ds = 0, \quad (\omega, t) \in \Omega \setminus A \times \mathbb{R}_+.$$

Let $\omega \in \Omega \setminus A$ be arbitrary. We conclude that there is a Borel set $B_\omega \subset (0,\infty)$ with $\mathrm{Leb}((0,\infty) \setminus B_\omega) = 0$ such that for each $t \in B_\omega$

$$\int_{\mathbb{R}^n} \mathbf{1}_{\mathbb{R}\setminus\{0\}}(W(t, Z_{t-}(\omega), x)) K_{t, Z_{t-}(\omega)}(dx) = 0.$$

Hence, it holds for every $t \in B_\omega$

$$W(t, Z_{t-}(\omega), x) = 0 \quad \text{for } K_{t, Z_{t-}(\omega)}\text{-almost all } x \in \mathbb{R}^n.$$

Applying Lemma 3.1.5 yields (note that the supports of $K_{t,z}$ do not depend on z, since (X, Z) has regular supports)

$$W(t, Z_{t-}(\omega), x) = 0 \quad \text{for all } t \in B_\omega \text{ and } x \in \operatorname{supp}(K_t).$$

From the left-continuity of the trajectory $t \mapsto Z_{t-}(\omega)$ and the continuity of W we deduce

$$W(t, Z_{t-}(\omega), x) = 0 \quad \text{for all } t \in (0,\infty) \text{ and } x \in \operatorname{supp}(K_t),$$

because assuming $W(t, Z_{t-}(\omega), x) \neq 0$ for some $t \in (0,\infty)$ and $x \in \operatorname{supp}(K_t)$ leads (note that $\mathrm{Leb}((0,\infty) \setminus B_\omega) = 0$ as well as the second condition of Definition 3.1.10) to the contradiction that $W(s, Z_{s-}(\omega), x) \neq 0$ for all $s \in [v, t]$ where v stems from the interval $(0, t)$. Now, we apply Lemma 3.1.4 for fixed $t \in (0,\infty)$ and $x \in \operatorname{supp}(K_t)$, which completes the proof. $\square$

Let X be a n-dimensional semimartingale for some $n \in \mathbb{N}$. We set $C(X) := \{i \in \{1,\dots,n\} \,|\, X_i^c \neq 0\}$. If X is a special semimartingale with canonical decomposition $X = X_0 + M + B$, we moreover define $D(X) := \{i \in \{1,\dots,n\} \,|\, M_i^d \neq 0\}$.

If $C(X) \neq \emptyset$, we define the vector process $\mathcal{C}(X) := (X_i^c)_{i \in C(X)}$, and if $C(X) = \emptyset$, the process $\mathcal{C}(X)$ is defined to be the (one-dimensional) zero process. If X is a special semimartingale, the process $\mathcal{D}(X)$ is defined analogously by means of the index set $D(X)$ and the purely discontinuous martingale parts M_i^d.

For matrices A and vectors b, we denote by A_c and b_c the submatrices and subvectors indexed by $C(X)$. Provided the semimartingale X is special, A_d and b_d are defined analogously with the index set $D(X)$.

Assume X is a special semimartingale. Let K be a measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. If $D(X) \neq \emptyset$, it is of the form $D(X) = \{\tau(1) < \ldots < \tau(m)\} \subset \{1,\dots,n\}$. The measure K_d on $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$ is defined by

$$K_d(B_{\tau(1)} \times \ldots \times B_{\tau(m)}) := K(\{x \in \mathbb{R}^n \,|\, x_i \in B_i \text{ for } i \in D(X), x_i = 0 \text{ for } i \notin D(X)\}),$$

where $B_{\tau(1)}, \dots, B_{\tau(m)} \in \mathcal{B}(\mathbb{R})$. If $D(X) = \emptyset$, the measure K_d is defined to be the zero measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

3.1.14 Definition. *Let X be a Grigelionis process with Z-derivative $(\beta, c, K; Z)^D$. We say that the derivative $(\beta, c, K; Z)^D$ is decomposable if*

1. *$C(X) \cup D(X) = \{1,\dots,n\}$ (but $C(X)$ and $D(X)$ need not be disjoint);*
2. *If $C(X) \neq \emptyset$, then $\mathcal{C}(X)$ has the type (C)-derivative $(0, c_c, 0; Z)^D$;*
3. *If $D(X) \neq \emptyset$, then $\mathcal{D}(X)$ has the type (D)-derivative $(0, 0, K_d; Z)^D$.*

3.1.15 Definition. *Let X be a Grigelionis process with derivative $(0, 0, K; Z)^D$ of type (D).*

1. *The process X is called non-degenerate if for all $t \in (0,\infty)$ there are linearly independent vectors $v_j(t) \in \mathbb{R}^n$, $j = 1,\dots,n$ and $\varepsilon(t) > 0$ such that $\bigcup_{j=1}^n \{\lambda v_j(t) : \lambda \in [0, \varepsilon(t)]\} \subset \operatorname{supp}(K_{t,z})$ for all $z \in \mathbb{R}^d$.*
2. *The process X is said to be linearly non-degenerate if for all $t \in (0,\infty)$ there are linearly independent vectors $v_j(t) \in \mathbb{R}^n$, $j = 1,\dots,n$ such that $v_1(t),\dots,v_n(t) \in \operatorname{supp}(K_{t,z})$ for all $z \in \mathbb{R}^d$.*

Later in practice, we assume that (X, Z) has regular supports, implying that the supports $\operatorname{supp}(K_{t,z})$ do not depend on $z \in \mathbb{R}^d$. Then, the process X is non-degenerate if the supports $\operatorname{supp}(K_t)$ contain n (which is the dimension of X) line segments of positive length, starting in zero and pointing in linearly independent directions. This corresponds to the condition of small jumps, that we have exhibited in the informal treatment of the realization problem. Formally, this condition is required in the proof of Theorem 3.1.17 in order to derive equation (3.1.26).

The process X is linearly non-degenerate, if the supports $\mathrm{supp}(K_t)$ just contain n linearly independent points. As we have pointed out in the informal geometric discussion, the conditions on the driving process X can be relaxed if one already knows that a given realization is affine, and this relaxation formally means that X is just linearly non-degenerate instead of being non-degenerate.

3.1.16 Definition. *Let X be a Grigelionis process with decomposable derivative.*

1. *The Grigelionis process X is called non-degenerate if $D(X) = \emptyset$ or $\mathcal{D}(X)$ is non-degenerate.*

2. *The Grigelionis process X is said to be linearly non-degenerate if $D(X) = \emptyset$ or $\mathcal{D}(X)$ is linearly non-degenerate.*

The class of non-degenerate Grigelionis processes is quite large. For example, let the Grigelionis process X consist of $m + n$ independent processes $X = (W_1, \dots, W_m, L_1, \dots, L_n)$, where the W_i are standard Wiener, and the L_j Lévy processes whose Lévy measures K_j satisfy $0 \in \overline{\mathrm{int}\,\mathrm{supp}(K_j)}$. Then, X is non-degenerate, because, by the independency of the processes, the Lévy measure of $(L_1, \dots, L_n)$ is concentrated on the union of the coordinate axes.

Most of the Lévy processes, that have been used in mathematical finance, have indeed the property $0 \in \overline{\mathrm{int}\,\mathrm{supp}(K)}$, where K denotes the Lévy measure. A well-known class of processes, having this property, is constituted by the generalized hyperbolic distributions, which were introduced in Barndorff-Nielsen [2]. Barndorff-Nielsen and Halgreen [3] have shown that generalized hyperbolic distributions are infinitely divisible. These processes were introduced in finance by Eberlein and Keller [26]. Hyperbolic term structure models are treated in Eberlein and Raible [29, Sec. 5]. Another class of Lévy processes, fulfilling $0 \in \overline{\mathrm{int}\,\mathrm{supp}(K)}$, is the class of bilateral Gamma processes, see Sec. 5 in Küchler and Naumann [46], where term structure models driven by bilateral Gamma processes are considered. In particular, the variance Gamma processes in the sense of Madan [49] have this property.

However, Poisson processes, or more general compound Poisson processes, are not non-degenerate. We have explained in the intuitive part of this section, why these processes, which have a lack of small jumps, must be excluded. As a matter of fact, there are works, like Shirakawa [60], Jarrow and Madan [43] or Hyll [40], which deal with term structure models driven by jump-diffusions, where one source of randomness is given by a point process. In order to integrate those models in the study of the realization problem, we have introduced the notion of linearly non-degenerate processes. If the driving process of a forward rate model is linearly non-degenerate, but not non-degenerate, like a Poisson process, we confine ourselves to the study of the realization problem to affine realizations. Since concrete realizations for term structure models, that have been constructed in the literature, see, e.g., Björk and Landén [12], have always turned out to be affine, this restriction does not seem to be severe. Anyway, in Chapter 4 we will, for other reasons, mainly study affine realizations.

We are ready to prove the first announced result, which later allows the identification of drift and volatilities of forward rate equations admitting realizations. Theorem 3.1.17, which is designed for Itô dynamics, will be applied in Chapter 4, whereas we use the later Theorem 3.1.21, which is an analogous result regarding Fisk-Stratonovich integrals, in this chapter.

We emphasize that, for convenience of notation, we do, concerning vectors, not distinguish between row and column vectors. It will always be clear from the context if a vector is meant to be a row or a column vector.

3.1.17 Theorem. *Let $d, n \in \mathbb{N}$, continuous mappings*

$$\begin{cases} \mu & : & \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \\ \gamma & : & \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times n} \\ \alpha & : & \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R} \\ \sigma & : & \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^n \end{cases},$$

and $f \in C^{2,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ be given. Let X be a Grigelionis process with decomposable derivative $(\beta, c, K; Z)^D$ such that $(\mathcal{D}(X), Z)$ has regular supports. Assume that Z_t and $f(t, Z_t)$ have the dynamics

$$dZ_t = \mu(t, Z_{t-})dt + \gamma(t, Z_{t-})dX_t, \tag{3.1.10}$$

$$df(t, Z_t) = \alpha(t, Z_{t-})dt + \sigma(t, Z_{t-})dX_t. \tag{3.1.11}$$

Then, the following statements are valid.

1. *For all $t \in (0,\infty)$, $z \in \operatorname{supp}(Z_{t-})$ and $x \in \operatorname{supp}(K^d_t)$ it holds*

$$\langle \sigma_d(t, z), x\rangle = f(t, z + \gamma_d(t, z)x) - f(t, z). \tag{3.1.12}$$

2. *If X is non-degenerate, it holds for every $t \in (0,\infty)$ and $z \in \operatorname{supp}(Z_{t-})$*

$$\begin{aligned} \alpha(t, z) &= \frac{\partial}{\partial t} f(t, z) + \langle \nabla_z f(t, z), \mu(t, z)\rangle \\ &\quad + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial z_i \partial z_j} f(t, z)\langle \gamma_{i\bullet}(t, z)^*, c(t, z)\gamma_{j\bullet}(t, z)^*\rangle, \end{aligned} \tag{3.1.13}$$

$$\sigma(t, z) = \nabla_z f(t, z)\gamma(t, z). \tag{3.1.14}$$

3. *If X is linearly non-degenerate, and there are $a, b \in C^2(\mathbb{R}_+)$ such that for all $t \in (0,\infty)$ and $z \in \operatorname{supp}(Z_{t-})$ it holds*

$$\begin{aligned} f(t, z + \gamma(t, z)x) &= a(t) + \langle b(t), z + \gamma(t, z)x\rangle, \quad x \in \operatorname{supp}(K_t) \cup \{0\}, \\ \nabla_z f(t, z) &= b(t), \end{aligned}$$

then the identities

$$\alpha(t,z) = \frac{\partial}{\partial t}a(t) + \Big\langle \frac{\partial}{\partial t}b(t), z \Big\rangle + \langle b(t), \mu(t,z)\rangle, \tag{3.1.15}$$

$$\sigma(t,z) = b(t)\gamma(t,z) \tag{3.1.16}$$

are valid for all $t \in (0,\infty)$ *and* $z \in \operatorname{supp}(Z_{t-})$.

Proof. By Itô's formula Jacod and Shiryaev [42, Thm. I.4.57] we obtain

$$\begin{aligned} f(t,Z_t) = f(0,Z_0) &+ \int_0^t \frac{\partial}{\partial s} f(s,Z_{s-})ds + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial z_i} f(s,Z_{s-})dZ_s^i \\ &+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2}{\partial z_i \partial z_j} f(s,Z_{s-}) d\langle Z_i^c, Z_j^c\rangle_s \\ &+ \sum_{0<s\le t} \Big[f(s,Z_s) - f(s,Z_{s-}) - \sum_{i=1}^d \frac{\partial}{\partial z_i} f(s,Z_{s-})\Delta Z_s^i \Big], \end{aligned} \tag{3.1.17}$$

where the last term is in $\mathcal{V}$. Taking into account the jumps $\Delta Z_t = \gamma(t,Z_{t-})\Delta X_t$, the differentials

$$\begin{aligned} dZ_t &= \mu(t,Z_{t-})dt + \gamma(t,Z_{t-})dX_t, \\ d\langle Z_i^c, Z_j^c\rangle_t &= \langle \gamma_{i\bullet}(t,Z_{t-})^*, c(t,Z_{t-})\gamma_{j\bullet}(t,Z_{t-})^*\rangle dt, \quad i,j = 1,\dots,d, \end{aligned}$$

and the associativity of the Itô integral Jacod and Shiryaev [42, I.4.37], we write equation (3.1.17) as

$$\begin{aligned} df(t,Z_t) = \frac{\partial}{\partial t} f(t,Z_{t-})dt &+ \langle \nabla_z f(t,Z_{t-}), \mu(t,Z_{t-})\rangle dt + \nabla_z f(t,Z_{t-})\gamma(t,Z_{t-})dX_t \\ &+ \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial z_i \partial z_j} f(t,Z_{t-})\langle \gamma_{i\bullet}(t,Z_{t-})^*, c(t,Z_{t-})\gamma_{j\bullet}(t,Z_{t-})^*\rangle dt \\ &+ \int_{\mathbb{R}^n} \Big(f(t, Z_{t-} + \gamma(t,Z_{t-})x) - f(t,Z_{t-}) \\ &\quad - \langle \nabla_z f(t,Z_{t-}), \gamma(t,Z_{t-})x\rangle \Big) \mu^X(dt,dx), \end{aligned} \tag{3.1.18}$$

where the $\mu^X(dt,dx)$-integral is in $\mathcal{V}$. The process $f(t,Z_t)$ is a special semimartingale, because the dynamics (3.1.11) provide a decomposition where the finite variation part is predictable. According to Prop. I.4.23 and Lemma I.3.10 in Jacod and Shiryaev [42], the $\mu^X(dt,dx)$-integral in (3.1.18) belongs to $\mathcal{A}_{\text{loc}}$. Therefore, we may integrate with respect to $(\mu^X - \nu)(dt,dx)$ plus $\nu(dt,dx)$ Jacod and Shiryaev [42, Prop. II.1.28], and write equation (3.1.18) as

$$\begin{aligned} df(t,Z_t) = \Big[&\frac{\partial}{\partial t} f(t,Z_{t-}) + \langle \nabla_z f(t,Z_{t-}), \mu(t,Z_{t-}) + \gamma(t,Z_{t-})\beta(t,Z_{t-})\rangle \\ &+ \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial z_i \partial z_j} f(t,Z_{t-}) \langle \gamma_{i\bullet}(t,Z_{t-})^*, c(t,Z_{t-})\gamma_{j\bullet}(t,Z_{t-})^*\rangle \\ &+ \int \Big(f(t,Z_{t-} + \gamma_d(t,Z_{t-})x) - f(t,Z_{t-}) \\ &\qquad - \langle \nabla_z f(t,Z_{t-}), \gamma_d(t,Z_{t-})x\rangle \Big) K^d_{t,Z_{t-}}(dx) \Big] dt \\ &+ \nabla_z f(t,Z_{t-})\gamma_c(t,Z_{t-}) d\mathcal{C}(X)_t \\ &+ \Big(f(t,Z_{t-} + \gamma_d(t,Z_{t-})x) - f(t,Z_{t-}) \Big) (\mu^{\mathcal{D}(X)} - \nu^{\mathcal{D}(X)})(dt,dx). \end{aligned} \tag{3.1.19}$$

According to Jacod and Shiryaev [42, Cor. II.2.38] the functions $W_i(\omega,t,x) = x_i$ belong to $G_{\text{loc}}(\mu^X)$. Applying Jacod and Shiryaev [42, Prop. II.1.30.b] the dynamics (3.1.11) of $f(t,Z_t)$ can be expressed as

$$\begin{aligned} df(t,Z_t) = (\alpha(t,Z_{t-}) + \langle \beta(t,Z_{t-}), \sigma(t,Z_{t-})\rangle)dt + \sigma_c(t,Z_{t-})d\mathcal{C}(X)_t \\ + \langle x, \sigma_d(t,Z_{t-})\rangle (\mu^{\mathcal{D}(X)} - \nu^{\mathcal{D}(X)})(dt,dx). \end{aligned} \tag{3.1.20}$$

Since the continuous local martingale part, the purely discontinuous local martingale part and the finite variation part of a special semimartingale are unique (which follows from Cor. I.3.16 and Thm. I.4.18 in Jacod and Shiryaev [42]), we obtain from equations (3.1.19) and (3.1.20)

$$\begin{aligned} &\int_0^t \alpha(s,Z_{s-})ds + \int_0^t \langle \beta(s,Z_{s-}), \sigma(s,Z_{s-}) - \nabla_z f(s,Z_{s-})\gamma(s,Z_{s-})\rangle ds \\ &= \int_0^t \Big[\frac{\partial}{\partial s} f(s,Z_{s-}) + \langle \nabla_z f(s,Z_{s-}), \mu(s,Z_{s-})\rangle \\ &+ \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial z_i \partial z_j} f(s,Z_{s-}) \langle \gamma_{i\bullet}(s,Z_{s-})^*, c(s,Z_{s-})\gamma_{j\bullet}(s,Z_{s-})^*\rangle \Big] ds \\ &+ \int_0^t \int \Big(f(s,Z_{s-} + \gamma_d(s,Z_{s-})x) - f(s,Z_{s-}) \\ &\qquad - \langle \nabla_z f(s,Z_{s-})\gamma_d(s,Z_{s-}), x\rangle \Big) K^d_{s,Z_{s-}}(dx) ds, \end{aligned} \tag{3.1.21}$$

as well as

$$\int_0^t \sigma_c(s, Z_{s-}) d\mathcal{C}(X)_s = \int_0^t \nabla_z f(s, Z_{s-}) \gamma_c(s, Z_{s-}) d\mathcal{C}(X)_s, \tag{3.1.22}$$

$$\begin{aligned} &\int_0^t \int \langle x, \sigma_d(s, Z_{s-}) \rangle (\mu^{\mathcal{D}(X)} - \nu^{\mathcal{D}(X)})(ds, dx) \\ &= \int_0^t \int \Big(f(s, Z_{s-} + \gamma_d(s, Z_{s-})x) - f(s, Z_{s-}) \Big) (\mu^{\mathcal{D}(X)} - \nu^{\mathcal{D}(X)})(ds, dx). \end{aligned} \tag{3.1.23}$$

Applying Lemma 3.1.11 on equation (3.1.22) we obtain

$$\sigma_c(t, z) = \nabla_z f(t, z) \gamma_c(t, z), \quad t \in (0, \infty), z \in \operatorname{supp}(Z_{t-}), \tag{3.1.24}$$

and an application of Lemma 3.1.13 on equation (3.1.23) yields the identity (3.1.12).

Provided, X is non-degenerate and $m := |D(X)| \in \{1, \dots, n\}$ (otherwise the formula (3.1.14) for $\sigma(t, z)$ follows immediately from (3.1.24), and we can continue after equation (3.1.26)), there are, for all $t \in (0, \infty)$, linearly independent vectors $v_j(t) \in \mathbb{R}^m$, $j = 1, \dots, m$ and $\varepsilon(t) > 0$ such that $\bigcup_{j=1}^m \{\lambda v_j(t) : \lambda \in [0, \varepsilon(t)]\} \subset \operatorname{supp}(K_t)$. Thus, it holds for for all $t \in (0, \infty)$, $z \in \operatorname{supp}(Z_{t-})$ and $j = 1, \dots, m$

$$\begin{aligned} \langle \gamma_d(t, z) v_j(t), \nabla_z f(t, z) \rangle &= \frac{\partial}{\partial(\gamma_d(t, z) v_j(t))} f(t, z) \\ &= \lim_{h \to 0} \frac{f(t, z + h\gamma_d(t, z) v_j(t)) - f(t, z)}{h} = \langle \sigma_d(t, z), v_j(t) \rangle. \end{aligned} \tag{3.1.25}$$

In the last step, we have used equation (3.1.12). Since the vectors $v_j(t)$, $j = 1, \dots, m$ are linearly independent, they provide a basis of $\mathbb{R}^m$, and consequently it holds for all $t \in (0, \infty)$ and $z \in \operatorname{supp}(Z_{t-})$

$$\langle \sigma_d(t, z) - \nabla_z f(t, z) \gamma_d(t, z), x \rangle = 0, \quad x \in \mathbb{R}^m.$$

We conclude that for all $t \in (0, \infty)$ and all $z \in \operatorname{supp}(Z_{t-})$

$$\sigma_d(t, z) = \nabla_z f(t, z) \gamma_d(t, z). \tag{3.1.26}$$

Together with (3.1.24), the identity (3.1.14) for $\sigma(t, z)$ is proven. Applying Lemma 3.1.4 on equations (3.1.12) and (3.1.14) we obtain that, for all $t \in (0, \infty)$ and $x \in \operatorname{supp}(K_t^d)$, it holds

$$\begin{aligned} \langle \sigma_d(t, Z_{t-}), x \rangle &= f(t, Z_{t-} + \gamma_d(t, Z_{t-})x) - f(t, Z_{t-}) \quad \mathbb{P} - a.s. \\ \sigma(t, Z_{t-}) &= \nabla_z f(t, Z_{t-}) \gamma(t, Z_{t-}) \quad \mathbb{P} - a.s. \end{aligned}$$

By the continuity assumptions on f and the coefficients, and the right-continuity of Z (notice also the second point in Definition 3.1.10), we obtain, up to a $\mathbb{P}$-null set,

$$\begin{aligned} \langle \sigma_d(t, Z_{t-}), x \rangle &= f(t, Z_{t-} + \gamma_d(t, Z_{t-})x) - f(t, Z_{t-}), \quad t \in (0, \infty), x \in \operatorname{supp}(K_t^d) \\ \sigma(t, Z_{t-}) &= \nabla_z f(t, Z_{t-}) \gamma(t, Z_{t-}), \quad t \in (0, \infty). \end{aligned}$$

Hence, equation (3.1.21) simplifies to

$$\int_0^t \alpha(s, Z_{s-})ds = \int_0^t \bigg[\frac{\partial}{\partial s} f(s, Z_{s-}) + \langle \nabla_z f(s, Z_{s-}), \mu(s, Z_{s-})\rangle$$
$$+ \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial z_i \partial z_j} f(s, Z_{s-})\langle \gamma_{i\bullet}(s, Z_{s-})^*, c(s, Z_{s-})\gamma_{j\bullet}(s, Z_{s-})^*\rangle\bigg] ds.$$

Applying Lemma 3.1.6, we obtain the stated formula (3.1.13) for $\alpha(t, z)$.

If the assumptions from the third part of the theorem are satisfied, it holds for all $t \in (0, \infty)$ and $z \in \operatorname{supp}(Z_{t-})$

$$f(t, z + \gamma_d(t, z)x) = a(t) + \langle b(t), z + \gamma_d(t, z)x\rangle, \quad x \in \operatorname{supp}(K_t^d) \cup \{0\},$$
$$\frac{\partial}{\partial t} f(t, z) = \frac{\partial}{\partial t} a(t) + \Big\langle \frac{\partial}{\partial t} b(t), z\Big\rangle.$$

The latter identity is valid, because $(\mathcal{D}(X), Z)$ has regular supports (see the third condition in Definition 3.1.10). Using equation (3.1.12), we get that for all $t \in (0, \infty)$ and $z \in \operatorname{supp}(Z_{t-})$ it holds

$$\langle \sigma_d(t, z), v_j(t)\rangle = \langle b(t), \gamma_d(t, z)v_j(t)\rangle, \quad j = 1, \dots, m$$

where $v_j(t)$, $j = 1, \dots, m$ are linearly independent vectors from $\operatorname{supp}(K_t)$, which corresponds to equation (3.1.25). Now, arguing as above, we obtain equations (3.1.15) and (3.1.16). □

3.1.18 Remark. *The assumptions of the third statement of Theorem 3.1.17 look a bit strange, and of course, they are in particular fulfilled if the whole function f is affine. The present formulation of this result will turn out to be useful in Chapter 4.*

As we have mentioned in the description of the geometric ideas, we also need the Fisk-Stratonovich integral in order to avoid second order derivative terms, which appear by application of Itô's formula.

Let M, N be two n-dimensional continuous local martingales for some $n \in \mathbb{N}$. We define $\langle M, N\rangle := \sum_{i=1}^n \langle M_i, N_i\rangle$, where for the one-dimensional components, the angle bracket denotes the predictable quadratic covariation of M_i and N_i Jacod and Shiryaev [42, Sec. I.4a], or, equivalently, the compensator of $[M_i, N_i]$ Jacod and Shiryaev [42, Prop. I.4.50.b].

3.1.19 Definition. *For two n-dimensional semimartingales X, Y the Fisk-Stratonovich integral of Y with respect to X is defined as*

$$\int_0^t Y_{s-} \circ dX_s := \int_0^t Y_{s-} dX_s + \frac{1}{2}\langle Y^c, X^c\rangle_t.$$

We also use the short-hand notation $Y_- \circ X$.

Note that in Protter [55, p. 75] the Fisk-Stratonovich integral is defined as $\int_0^t Y_{s-} \circ dX_s := \int_0^t Y_{s-} dX_s + \frac{1}{2}[Y,X]_t^c$, where $[Y,X]^c$ denotes the path by path continuous part of the quadratic co-variation $[Y,X]$. By Jacod and Shiryaev [42, Thm. I.4.52] it holds $[Y,X]_t = \langle Y^c, X^c \rangle_t + \sum_{s \le t} \Delta Y_s \Delta X_s$, implying $[Y,X]^c = \langle Y^c, X^c \rangle$, whence these two definitions of the Fisk-Stratonovich integral are consistent.

The next auxiliary result is useful in order to alternate Itô dynamics with Fisk-Stratonovich dynamics. For any semimartingale Z, we define the process $\hat{Z}$ by $\hat{Z}_t(\omega) := (t, Z_t(\omega))$.

3.1.20 Lemma. *Let X be a Grigelionis process with Z-derivative $(\beta, c, K; Z)^D$. Assume Z satisfies a stochastic differential equation*

$$dZ_t = \mu(t, Z_{t-})dt + \gamma(t, Z_{t-})dX_t, \tag{3.1.27}$$

where $\mu \in C(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^d)$ and $\gamma \in C^{2,2}(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^{d \times n})$. Then, the following identity is valid for each $i = 1, \dots, d$.

$$\langle \gamma_{i\bullet}(\hat{Z})^c, X^c \rangle_t = \sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^d \int_0^t c_{jk}(s, Z_{s-}) \left[\frac{\partial}{\partial z_l} \gamma_{ij}(s, Z_{s-}) \right] \gamma_{lk}(s, Z_{s-}) ds.$$

Proof. See the appendix. □

The following result is an analogue of Theorem 3.1.17 for Fisk-Stratonovich integrals. For this reason, the proof is contained in the appendix.

3.1.21 Theorem. *Let $d, n \in \mathbb{N}$, continuous mappings*

$$\begin{cases} \mu & : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d \\ \gamma & : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d \times n} \\ \alpha & : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R} \\ \sigma & : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^n \end{cases},$$

and $f \in C^{2,2}(\mathbb{R}_+ \times \mathbb{R}^d)$ be given. Let X be a Grigelionis process with decomposable derivative $(\beta, c, K; Z)^D$ such that $(\mathcal{D}(X), Z)$ has regular supports. Assume that Z_t and $f(t, Z_t)$ have the dynamics

$$dZ_t = \mu(t, Z_{t-})dt + \gamma(t, Z_{t-}) \circ dX_t, \tag{3.1.28}$$
$$df(t, Z_t) = \alpha(t, Z_{t-})dt + \sigma(t, Z_{t-}) \circ dX_t. \tag{3.1.29}$$

Then, the following statements are valid.

1. *For all $t \in (0, \infty)$, $z \in \operatorname{supp}(Z_{t-})$ and $x \in \operatorname{supp}(K_t^d)$ it holds*

$$\langle \sigma_d(t, z), x \rangle = f(t, z + \gamma_d(t, z)x) - f(t, z). \tag{3.1.30}$$

2. *If X is non-degenerate, it holds for every $t \in (0,\infty)$ and $z \in \operatorname{supp}(Z_{t-})$*

$$\alpha(t,z) = \frac{\partial}{\partial t} f(t,z) + \langle \nabla_z f(t,z), \mu(t,z) \rangle, \tag{3.1.31}$$

$$\sigma(t,z) = \nabla_z f(t,z) \gamma(t,z). \tag{3.1.32}$$

3. *If X is linearly non-degenerate, and there are $a, b \in C^2(\mathbb{R}_+)$ such that for all $t \in (0,\infty)$ and $z \in \operatorname{supp}(Z_{t-})$ it holds*

$$f(t, z + \gamma(t,z)x) = a(t) + \langle b(t), z + \gamma(t,z)x \rangle, \quad x \in \operatorname{supp}(K_t) \cup \{0\},$$
$$\nabla_z f(t,z) = b(t),$$

then the identities

$$\alpha(t,z) = \frac{\partial}{\partial t} a(t) + \left\langle \frac{\partial}{\partial t} b(t), z \right\rangle + \langle b(t), \mu(t,z) \rangle, \tag{3.1.33}$$

$$\sigma(t,z) = b(t) \gamma(t,z) \tag{3.1.34}$$

are valid for all $t \in (0,\infty)$ and $z \in \operatorname{supp}(Z_{t-})$.

Proof. See the appendix. □

3.2 The Lie algebraic approach

In this section, $\mathcal{X}$ denotes a real Banach space consisting of functions $r : \mathbb{R}_+ \to \mathbb{R}$. In our applications below, the space $\mathcal{X}$ will be a space of forward rate curves. We assume that the norm of this space is chosen such that for all $r \in \mathcal{X}$ and every sequence $(r_n)_{n \in \mathbb{N}}$ in $\mathcal{X}$ the convergence $r_n \to r$ in $\mathcal{X}$ implies that $r_n(x) \to r(x)$ for each $x \in \mathbb{R}_+$.

Let r be a $\mathcal{X}$-valued process whose paths are right-continuous and admit left-hand limits, abbreviated càdlàg. By this, we mean the following. For each $t \in \mathbb{R}_+$, the sequence r_s is a Cauchy sequence in $\mathcal{X}$ as $s \downarrow t$ with limit r_t, and for each $t \in (0,\infty)$ the sequence r_s is a Cauchy sequence in $\mathcal{X}$ as $s \uparrow t$, whose limit is denoted by r_{t-}. Moreover, we set $r_{0-} := r_0$.

Let X be a n-dimensional Grigelionis process whose derivative $(\beta, c, K)^D$ is deterministic (a Lévy process for instance), and $\alpha, \sigma_1, \ldots, \sigma_n : \mathbb{R}_+ \times \mathcal{X} \to \mathcal{X}$. We assume that r satisfies the $\mathcal{X}$-valued equation

$$\begin{cases} dr_t &= \alpha(t, r_{t-})dt + \sigma(t, r_{t-})dX_t \\ r_0 &= r^* \end{cases}, \tag{3.2.1}$$

where $r^* \in \mathcal{X}$. The equation (3.2.1) is to be understood as follows. The process r is a stochastic process with values in $\mathcal{X}$, and for each $x \in \mathbb{R}_+$, the process $r(x)$ is a semimartingale satisfying

$$\begin{cases} dr_t(x) &= \alpha(t, r_{t-})(x)dt + \sigma(t, r_{t-})(x)dX_t \\ r_0(x) &= r^*(x) \end{cases}.$$

In the sequel, we deal with functions $G : \mathbb{R}_+ \times \mathbb{R}^d \to \mathcal{X}$. We say that such a function is of class C^m for some $m \in \mathbb{N}$, if it is m-times Fréchet differentiable on the open subset $(0,\infty) \times \mathbb{R}^d$, and the Fréchet derivative $\mathbb{F}^m G : (0,\infty) \times \mathbb{R}^d \to L^{(m)}(\mathbb{R} \times \mathbb{R}^d, \mathcal{X})$ - the latter is the space of continuous multilinear operators from $(\mathbb{R} \times \mathbb{R}^d)^m$ to $\mathcal{X}$ - is continuous, and can be extended to a continuous function on $\mathbb{R}_+ \times \mathbb{R}^d$. Note that for any $\hat{z}_0 \in \mathbb{R}_+ \times \mathbb{R}^d$, $\hat{z} \in \mathbb{R} \times \mathbb{R}^d$ one has

$$\mathbb{F}G(\hat{z}_0)[\hat{z}] = \lim_{h \to 0} \frac{G(\hat{z}_0 + h\hat{z}) - G(\hat{z}_0)}{h}.$$

Due to the assumption that convergence in $\mathcal{X}$ implies pointwise convergence, for each $x \in \mathbb{R}_+$, the function $(t,z) \mapsto G(t,z)(x)$ is of class $C^m(\mathbb{R}_+ \times \mathbb{R}^d)$, and the partial derivatives of $G(\bullet)(x)$ at some point $\hat{z}_0$ are obtained by inserting the unit vectors in $\mathbb{F}^m G(\hat{z}_0)[\bullet](x)$.

For any $G : \mathbb{R}_+ \times \mathbb{R}^d \to \mathcal{X}$, we introduce the mapping $\hat{G} : (0,\infty) \times \mathbb{R}^d \to (0,\infty) \times \mathcal{X}$ by $\hat{G}(t,z) := (t, G(t,z))$.

The following definition of a finite dimensional realization differs a bit from Definition 2.2.3. It is adapted to the present framework.

3.2.1 Definition.

1. *Let $d \in \mathbb{N}$. The equation (3.2.1) has a $d+1$-dimensional realization if there exists a pair (G,Z), where $G : \mathbb{R}_+ \times \mathbb{R}^d \to \mathcal{X}$ is of class C^2 and such that $\hat{G}$ is a homeomorphism, and Z is a d-dimensional semimartingale with $(\mathcal{D}(X), Z)$ having regular supports, satisfying*

$$\begin{cases} dZ_t &= \mu(t, Z_{t-})dt + \gamma(t, Z_{t-})dX_t \\ Z_0 &= z_0 \end{cases},$$

   for functions $\mu \in C(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^d)$, $\gamma \in C^{2,2}(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^{d \times n})$ and $z_0 \in \mathbb{R}^d$, such that $r_t = G(t, Z_t)$ for all $t \in \mathbb{R}_+$. We call the semimartingale Z from the pair (G,Z) the state process of the realization.

2. *The equation (3.2.1) has a finite dimensional realization if it admits a $d+1$-dimensional realization for some $d \in \mathbb{N}$.*

3. *A $d+1$-dimensional realization (G,Z) is said to be affine, if there are $a, b_1, \dots, b_d : \mathbb{R}_+ \to \mathcal{X}$ of class C^2 such that*

$$G(t,z) = a(t) + \langle b(t), z \rangle \quad \text{for all } (t,z) \in \mathbb{R}_+ \times \mathbb{R}^d.$$

Provided, there exists a finite dimensional realization (G,Z), the image $\hat{\mathcal{G}} := \mathrm{Im}(\hat{G})$ is a $d+1$-dimensional manifold, parameterized by $\hat{G}$. The linear space $\mathcal{T}_{\hat{\mathcal{G}}}(\hat{r}_0) :=$

$\mathrm{Im}(\mathbb{F}\hat{G}(\hat{z}_0))$, where $\hat{z}_0 := \hat{G}^{-1}(\hat{r}_0)$, denotes the tangent space of $\hat{\mathcal{G}}$ at point $\hat{r}_0 \in \hat{\mathcal{G}}$. The tangent space has the representation

$$\mathcal{T}_{\hat{\mathcal{G}}}(\hat{r}_0) = \operatorname{span}\left\{\Big(1, \tfrac{\partial}{\partial t}G(\hat{z}_0)\Big), \Big(0, \tfrac{\partial}{\partial z_1}G(\hat{z}_0)\Big), \dots, \Big(0, \tfrac{\partial}{\partial z_d}G(\hat{z}_0)\Big)\right\}, \tag{3.2.2}$$

because for fixed $x \in \mathbb{R}_+$ it holds

$$\begin{aligned}\mathbb{F}\hat{G}(t_0, z_0)[t, z](x) &= \lim_{h\to 0} \frac{\hat{G}((t_0, z_0) + h(t, z))(x) - \hat{G}(t_0, z_0)(x)}{h} \\ &= \left(t, \lim_{h\to 0} \frac{G((t_0, z_0) + h(t, z))(x) - G(t_0, z_0)(x)}{h}\right) \\ &= t\left(1, \frac{\partial}{\partial t}G(t_0, z_0)(x)\right) + \sum_{i=1}^{d} z_i\left(0, \frac{\partial}{\partial z_i}G(t_0, z_0)(x)\right).\end{aligned}$$

Vector fields, Lie brackets and Lie algebras have already been introduced in Definition 3.1.1.

The following result will be essential in the sequel. It is a consequence of the Frobenius Theorem Björk and Svensson [13, Thm. 2.1].

3.2.2 Theorem. *Let $\mathcal{Y}$ be an arbitrary real Banach space and $U \subset \mathcal{Y}$ be an open subset. Let $f_1, \dots, f_d$ be linearly independent smooth vector fields on U, and $r_0 \in U$. The following statements are equivalent.*

1. *For all $r \in U$ from a neighborhood of r_0 there exists a d-dimensional manifold $\mathcal{G}_r \subset \mathcal{Y}$ with $r \in \mathcal{G}_r$, such that for all r^* from a neighborhood of r in $\mathcal{G}_r$*

$$f_1(r^*), \dots, f_d(r^*) \in \mathcal{T}_{\mathcal{G}_r}(r^*).$$

2. *For all $r \in U$ from a neighborhood of r_0 it holds*

$$[f_i, f_j](r) \in \operatorname{span}\{f_1(r), \dots, f_d(r)\}, \quad i, j = 1, \dots, d.$$

Proof. The statement is a particular case of Björk and Svensson [13, Thm. 2.2], see also Remark 3.1.2. □

For a differentiable function $f : (0, \infty) \times \mathcal{X} \to \mathcal{Y}$ we define the Fréchet derivatives

$$\begin{aligned}\mathbb{F}_t &: (0, \infty) \times \mathcal{X} \to L(\mathbb{R}, \mathcal{Y}), \\ \mathbb{F}_r &: (0, \infty) \times \mathcal{X} \to L(\mathcal{X}, \mathcal{Y})\end{aligned}$$

with respect to t and with respect to r as

$$\begin{aligned}\mathbb{F}_t f(t_0, r_0)[t] &:= \mathbb{F}f(t_0, r_0)[t, 0], \\ \mathbb{F}_r f(t_0, r_0)[r] &:= \mathbb{F}f(t_0, r_0)[0, r].\end{aligned}$$

Concerning the Fréchet derivative $\mathbb{F} : (0, \infty) \times \mathcal{X} \to L(\mathbb{R} \times \mathcal{X}, \mathcal{Y})$, the identity

$$\mathbb{F}f(t_0, r_0)[t, r] = \mathbb{F}_t f(t_0, r_0)[t] + \mathbb{F}_r f(t_0, r_0)[r]$$

is valid, because it holds

$$\begin{aligned} f((t_0, r_0) + h(t, r)) - f(t_0, r_0) &= \mathbb{F}f(t_0, r_0)[h(t, r)] + R(h(t, r)) \\ &= h\mathbb{F}f(t_0, r_0)[t, 0] + h\mathbb{F}f(t_0, r_0)[0, r] + R(h(t, r)), \quad \text{where } \lim_{\|u\| \to 0} \frac{R(u)}{\|u\|} = 0. \end{aligned}$$

By $\hat{r}$ we denote the $\mathbb{R}_+ \times \mathcal{X}$-valued process $\hat{r}_t(\omega) = (t, r_t(\omega))$.

3.2.3 Lemma. *Assume the $\mathcal{X}$-valued process r satisfies (3.2.1), and has a finite dimensional realization, and that one of the following conditions is satisfied.*

- *The process X is non-degenerate;*
- *The process X is linearly non-degenerate and the realization is affine.*

Then it holds for each $x \in \mathbb{R}_+$

$$\langle \sigma(\hat{r})(x)^c, X^c \rangle_t = \sum_{i,j=1}^{n} \int_0^t c_{ij}(s) \mathbb{F}_r \sigma_i(t, r_{s-})[\sigma_j(t, r_{s-})](x) ds.$$

Proof. See the appendix. □

Now we come to the announced necessary criterion for the existence of finite dimensional realizations, formulated in the terms of the drift and the volatilities of the equation (3.2.1). In order to apply the Frobenius Theorem 3.2.2, we extend $\alpha, \sigma_1, \ldots, \sigma_n$, by taking into account the time t, to mappings $\hat{\alpha}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n : (0, \infty) \times \mathcal{X} \to \mathbb{R} \times \mathcal{X}$, that is to vector fields on $(0, \infty) \times \mathcal{X}$, which is an open subset in the Banach space $\mathbb{R} \times \mathcal{X}$. With regard to Lemma 3.2.3, we define $\hat{\alpha}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n$ by

$$\hat{\alpha}(t, r) := (1, \tilde{\alpha}(t, r)), \tag{3.2.3}$$

$$\hat{\sigma}_i(t, r) := (0, \sigma_i(t, r)) \quad \text{for } i = 1, \ldots, n, \tag{3.2.4}$$

where $\tilde{\alpha} : (0, \infty) \times \mathcal{X} \to \mathcal{X}$ is defined as

$$\tilde{\alpha}(t, r) := \alpha(t, r) - \frac{1}{2} \sum_{i,j=1}^{n} c_{ij}(t) \mathbb{F}_r \sigma_i(t, r)[\sigma_j(t, r)]. \tag{3.2.5}$$

We presume that $\hat{\alpha}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n$ are smooth vector fields on $(0, \infty) \times \mathcal{X}$.

The next result can be regarded as one implication of Björk and Svensson [13, Thm. 3.2] for non-degenerate Grigelionis processes with deterministic derivative (we have argued in Section 3.1 that the converse cannot hold). However, one should note the fact that Björk and Svensson [13] consider local realizations, while we focus on global realizations, which causes minor deviations in the formulation of the theorem.

3.2.4 Theorem. *Let X be non-degenerate, and $r^0 \in \mathcal{X}$. Assume for all r from a neighborhood $U(r^0)$ of r^0 the equation (3.2.1) with initial value $r^* := r$ has a $d+1$-dimensional realization (G_r, Z_r). Let $(t_0, r_0) \in (0,\infty) \times \mathcal{X}$ be arbitrary. Assume for every (t^*, r^*) from a neighborhood of (t_0, r_0) there exist $r \in U(r^0)$ and $z^* \in \mathbb{R}^d$ such that*

$$(t^*, z^*) \in \operatorname{int}\{(t,z) \,|\, t \in (0,\infty), z \in \operatorname{supp}(Z_r(t-)\},$$

and $G_r(t^*, z^*) = r^*$. Then it holds in a neighborhood of (t_0, r_0)

$$\dim\{\hat{\alpha}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n\}_{LA} \leq d+1.$$

Proof. By assumption, for each $r \in U(r^0)$ there exists a realization (G_r, Z_r), where the semimartingale Z_r has the dynamics

$$dZ_r(t) = \mu_r(t, Z_r(t-))dt + \gamma_r(t, Z_r(t-))dX_t$$

with $\mu_r \in C(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^d)$ and $\gamma_r \in C^{2,2}(\mathbb{R}_+ \times \mathbb{R}^d, \mathbb{R}^{d\times n})$. According to Lemma 3.1.20, the processes Z_r satisfy

$$dZ_r(t) = \tilde{\mu}_r(t, Z_r(t-))dt + \gamma_r(t, Z_r(t-)) \circ dX_t, \tag{3.2.6}$$

where the components $\tilde{\mu}_r^i$, $i = 1, \ldots, d$ of $\tilde{\mu}_r$ are as follows,

$$\tilde{\mu}_r^i(t,z) = \mu_r^i(t,z) - \frac{1}{2}\sum_{j=1}^n \sum_{k=1}^n \sum_{l=1}^d c_{jk}(t)\left[\frac{\partial}{\partial z_l}\gamma_{ij}^r(t,z)\right]\gamma_{lk}^r(t,z).$$

By Lemma 3.2.3, for each $x \in \mathbb{R}_+$, the process r fulfills the Fisk-Stratonovich dynamics

$$dr_t(x) = \tilde{\alpha}(t, r_{t-})(x)dt + \sigma(t, r_{t-})(x) \circ dX_t,$$

where $\tilde{\alpha}$ was defined in (3.2.5). Inserting the realization $r_t = G_r(t, Z_t)$, we obtain for every $r \in U(r^0)$ and $x \in \mathbb{R}_+$

$$dG_r(t, Z_r(t))(x) = \tilde{\alpha}(t, G_r(t, Z_r(t-)))(x)dt + \sigma(t, G_r(t, Z_r(t-)))(x) \circ dX_t. \tag{3.2.7}$$

We have assumed at the beginning of this section that convergence in $\mathcal{X}$ implies pointwise convergence. Hence, the coefficients from (3.2.7) are continuous, and $G_r(\bullet)(x)$ is of class C^2 for each $x \in \mathbb{R}_+$. Applying Theorem 3.1.21 on the relations (3.2.6) and (3.2.7), we obtain for each $r \in U(r^0)$ and $x \in \mathbb{R}_+$

$$\tilde{\alpha}(t, G_r(t,z))(x) = \frac{\partial}{\partial t}G_r(t,z)(x) + \langle \nabla_z G_r(t,z)(x), \tilde{\mu}_r(t,z)\rangle, \tag{3.2.8}$$

$$\sigma(t, G_r(t,z))(x) = \nabla_z G_r(t,z)(x)\gamma_r(t,z), \tag{3.2.9}$$

where $t \in (0,\infty)$ and $z \in \operatorname{supp}(Z_r(t-))$.

Let us show that the first condition of the Frobenius Theorem 3.2.2 is satisfied. By hypothesis, for any (t^*, r^*) from a neighborhood of (t_0, r_0) there exist $r \in U(r^0)$ and $z^* \in \mathbb{R}^d$ such that $\hat{G}_r(t^*, z^*) = (t^*, r^*)$, i.e. (t^*, r^*) lies in the $d+1$-dimensional manifold $\hat{\mathcal{G}}_r$, and a neighborhood

$$U(t^*, z^*) \subset \{(t, z) \,|\, t \in (0, \infty), z \in \operatorname{supp}(Z_r(t-))\}$$

of (t^*, z^*). Since $\hat{G}_r$ is a homeomorphism, $\hat{G}_r(U(t^*, z^*))$ is a neighborhood of (t^*, r^*) in $\hat{\mathcal{G}}_r$, and it holds

$$\hat{G}_r(U(t^*, z^*)) \subset \bigcup_{t \in (0,\infty)} \{t\} \times G_r(t, \operatorname{supp}(Z_r(t-))).$$

Using equations (3.2.8) and (3.2.9), it follows from the representation (3.2.2) of the tangent spaces that for all $(\tilde{t}, \tilde{r}) \in \hat{G}_r(U(t^*, z^*))$ it holds

$$\hat{\alpha}(\tilde{t}, \tilde{r}), \hat{\sigma}_1(\tilde{t}, \tilde{r}), \ldots, \hat{\sigma}_n(\tilde{t}, \tilde{r}) \in \mathcal{T}_{\hat{\mathcal{G}}_r}(\tilde{t}, \tilde{r}).$$

Since the tangency of a manifold to a vector field is preserved under the Lie bracket (see, e.g., Filipović and Teichmann [34, Prop. 3.10]), an application of Theorem 3.2.2 yields that in neighborhood of (t_0, r_0)

$$\dim \{\hat{\alpha}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n\}_{LA} \leq d + 1.$$

□

Note that Theorem 3.2.4 deals with generic realizations Björk and Svensson [13, p. 213], i.e. for all r from a neighborhood $U(r^0)$ of r^0 the equation (3.2.1) with initial value $r^* := r$ admits a realization. For this reason, we demand that the derivative of the driving process X is deterministic. It only makes sense to allow that the derivative depends on the state process Z if one has just one fixed initial curve, and therefore only one state process.

With regard to the conditions of Theorem 3.2.4, one may in particular think of "large" neighborhoods $U(r^0)$, like the whole space $\mathcal{X}$, as it is done in the next result, where we focus on affine realizations. Since concrete realizations for term structure models, that have been constructed in the literature, have always turned out to be affine, this restriction seems justified.

Of course, one could prove exactly the same result as in Theorem 3.2.4 for affine realizations and driving linearly non-degenerate processes. In view of the known realizations for structure models, see in particular Björk and Landén [12], we assume a special structure of the affine realizations, which ensures that one of the technical conditions, which we impose in Theorem 3.2.4, is automatically fulfilled.

For $r \in \mathcal{X}$ and $t \in \mathbb{R}_+$ define the right-shift $\Theta_t r$ as $\Theta_t r(x) := r(t + x)$, $x \in \mathbb{R}_+$. Concerning the space $\mathcal{X}$, we furthermore assume that

- $\Theta_t r \in \mathcal{X}$ for all $r \in \mathcal{X}$ and $t \in (0, \infty)$;
- For each $r \in \mathcal{X}$ and $t \in (0, \infty)$ there exists a $\tilde{r} \in \mathcal{X}$ such that $r = \Theta_t \tilde{r}$.

3.2.5 Theorem. *Let X be linearly non-degenerate. Assume, for all $r \in \mathcal{X}$, the equation (3.2.1) with initial value $r^* := r$ has a $d+1$-dimensional affine realization (G_r, Z_r), and there is a neighborhood $U(0) \subset \mathbb{R}^d$ of zero, such that $Z_r(0) = 0$ for all $r \in \mathcal{X}$, $U(0) \subset \operatorname{supp}(Z_r(t-))$ for all $(t, r) \in (0, \infty) \times \mathcal{X}$, and*

$$G_r(t, z) = \Theta_t r + a(t) + \langle b(t), z \rangle, \quad (t, z, r) \in \mathbb{R}_+ \times \mathbb{R}^d \times \mathcal{X}$$

with $a, b_1, \dots, b_d : \mathbb{R}_+ \to \mathcal{X}$ such that $a(0) = 0$. Then, for all $(t_0, r_0) \in (0, \infty) \times \mathcal{X}$, it holds in a neighborhood of (t_0, r_0)

$$\dim \{\hat{\alpha}, \hat{\sigma}_1, \dots, \hat{\sigma}_n\}_{LA} \leq d + 1.$$

Proof. For arbitrary $(t_0, r_0) \in (0, \infty) \times \mathcal{X}$ and (t^*, r^*) from a neighborhood of (t_0, r_0) we choose $r \in \mathcal{X}$ such that $\Theta_{t^*} r = r^* - a(t^*)$. Then it holds $G_r(t^*, 0) = r^*$, and by hypothesis

$$(t^*, 0) \in \operatorname{int} \{(t, z) \mid t \in (0, \infty), z \in \operatorname{supp}(Z_r(t-)\}.$$

Now we can proceed as in the proof of Theorem 3.2.4, using the identities of Theorem 3.1.21, which are valid for linearly non-degenerate processes. □

3.3 Term structure models

We have to specify a space of forward rate curves in order to incorporate term structure models into the framework of infinite dimensional Banach space valued equations of the preceding section. We choose one of the spaces $\mathcal{H}_{\beta,\gamma}$ introduced in Björk and Svensson [13].

3.3.1 Definition. *Fix real numbers $\beta > 1$ and $\gamma > 0$. For all $f, g \in C^\infty(\mathbb{R}_+)$ define*

$$\langle f, g \rangle_{\beta,\gamma} := \sum_{n=0}^{\infty} \beta^{-n} \int_0^\infty \left(\frac{\partial^n}{\partial x^n} f(x) \right) \left(\frac{\partial^n}{\partial x^n} g(x) \right) e^{-\gamma x} dx.$$

The space $\mathcal{H}_{\beta,\gamma}$ is defined as the space of all $f \in C^\infty(\mathbb{R}_+)$ satisfying

$$\|f\|_{\beta,\gamma} := \sqrt{\langle f, f \rangle_{\beta,\gamma}} < \infty.$$

According to Björk and Svensson [13, Prop. 4.2] the space $\mathcal{H}$ (the parameters β, γ are considered to be fixed in the sequel, whence we suppress the subindices), equipped with the inner product $\langle \bullet, \bullet \rangle_{\beta,\gamma}$, is a Hilbert space, and $f_n \to f$ in $\mathcal{H}$ implies that $f_n^{(m)} \to f^{(m)}$ uniformly on compacts for every $m \in \mathbb{N}_0$, whence in particular $f_n(x) \to f(x)$ for each $x \in \mathbb{R}_+$. It also follows from Björk and Svensson [13, Prop. 4.2] that the operator $\partial/\partial x$ is a smooth vector field on this space. The smoothness of $\partial/\partial x$ is essential in order to apply the results of the preceding chapter, because it appears in the drift term of the Musiela parametrization. To sum up, so far all requirements on the space $\mathcal{H}$, imposed in Section 3.2, are fulfilled.

In order to establish Theorem 3.2.5, we have imposed further conditions on the space. For all $f \in \mathcal{H}$ and $t \in (0,\infty)$ it holds $\Theta_t f \in \mathcal{H}$, because $\|\Theta_t f\| \leq e^{\gamma t}\|f\|$, as one easily verifies by the definition of the norm in $\mathcal{H}$. By Björk and Svensson [13, Prop. 4.2] every $f \in \mathcal{H}$ can be uniquely extended to a holomorphic function on the complex plane, which ensures that for each $t \in (0,\infty)$ there exists a function $g \in C^\infty(\mathbb{R}_+)$ such that $f = \Theta_t g$, but it is not clear whether $\|g\| < \infty$.

This hints to the drawback of the space $\mathcal{H}$, namely that it is a very small space. In particular, it was pointed out by Filipović and Teichmann that it does not include the forward rate curves generated by the model introduced in Cox, Ingersoll, and Ross [24]. For this reason, Filipović and Teichmann [33] extended the theory to a larger space, see Section 6 in Björk [8] for an overview of this extension and for further references on that topic.

This extension, however, is far from trivial to carry out. The technical price, one has to pay for going into the deep parts of this so-called theory of convenient analysis, is quite high. Therefore, we follow Björk [8, Sec. 6] who formulated the main result of Filipović and Teichmann [33] in pedestrian terms for the working mathematician as follows: "When you are searching for finite dimensional realizations for equations of HJM type, you can compute the relevant Lie algebra without worrying about the space, since Filipović and Teichmann will always provide you with a convenient space to work in."

We focus on forward rate models of the form

$$\begin{cases} df(t,T) &= \alpha(t,T,r_{t-})dt + \sigma(t,T,r_{t-})dX_t \\ f(0,T) &= f^*(0,T) \end{cases} \tag{3.3.1}$$

with coefficients $\alpha, \sigma_1, \dots, \sigma_n : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{H} \to \mathbb{R}$, where r_t denotes the whole curve $r_t = f(t, t + \bullet)$ of forward rates. Unless stated otherwise, we always make the following assumptions for the term structure models of the type (3.3.1).

3.3.2 Assumptions.

1. *The derivative $(\beta, c, K)^D$ of the n-dimensional driving process X is deterministic.*
2. *There are $z_1^-, \dots, z_n^- \in (-\infty, 0)$ and $z_1^+, \dots, z_n^+ \in (0,\infty)$ such that for any $t \in \mathbb{R}_+$*

$$\int_0^t \left(\int_{|x|\leq 1} |x|^2 K_s(dx) + \int_{|x|>1} e^{\langle z,x\rangle} K_s(dx) \right) ds < \infty, \quad z \in Q,$$

where Q denotes the set $Q := [z_1^-, z_1^+] \times \ldots \times [z_n^-, z_n^+]$, and furthermore, for all $t \in \mathbb{R}_+$ it holds

$$\int_{|x|>1} e^{\langle z,x\rangle} K_t(dx) < \infty, \quad z \in Q.$$

3. *There are $w_1^- \in (z_1^-, 0), \ldots, w_n^- \in (z_n^-, 0)$ and $w_1^+ \in (0, z_1^+), \ldots, w_n^+ \in (0, z_n^+)$ such that*

$$-\int_t^T \sigma(t,s,r)ds \in Q_0, \quad (t,T,r) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{H}.$$

where $Q_0 \subset Q$ is defined as $Q_0 := [w_1^-, w_1^+] \times \ldots \times [w_n^-, w_n^+]$.

The second assumption ensures that the cumulant generating function

$$\Psi(t,z) := \langle \beta_t, z\rangle + \frac{1}{2}\langle z, c_t z\rangle + \int_{\mathbb{R}^n} \left(e^{\langle z,x\rangle} - 1 - \langle z, x\rangle\right) K_t(dx)$$

is definable for all $(t,z) \in \mathbb{R}_+ \times Q$ (see Section 2.2). The third assumption guarantees that

$$\Psi\left(t, -\int_t^T \sigma(t,s,r)ds\right)$$

exists for all $(t,T,r) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathcal{H}$.

In order to apply Theorem 3.2.4 later, we assume that X is non-degenerate (or linearly non-degenerate, if we want to apply Theorem 3.2.5, the result for affine realizations). As usual, we assume that the model is free of arbitrage (in the sense that $\mathbb{P}$ is a martingale measure). In view of Proposition 2.2.11, we presume that the drift is given by

$$\alpha(t,T,r) = -\Big\langle \sigma(t,T,r), \nabla_z \Psi\left(t, -\int_t^T \sigma(t,s,r)ds\right)\Big\rangle$$

for $t, T \in \mathbb{R}_+$ and $r \in \mathcal{H}$. In order to incorporate the forward rate model (3.3.1) in the framework of $\mathcal{H}$-valued equations, we switch to the Musiela parametrization of forward rates $r_t(x) = f(t, t+x)$, which is, according to Proposition 2.2.12, given by

$$\begin{cases} dr_t(x) &= \left[\frac{\partial}{\partial x} r_t(x) + \alpha(t, t+x, r_t)\right] dt + \sigma(t, t+x, r_{t-}) dX_t \\ r_0(x) &= f^*(0,x) \end{cases}. \tag{3.3.2}$$

We also presume $f^*(0, \bullet) \in \mathcal{H}$, that $\alpha(t, t+\bullet, r), \sigma_1(t, t+\bullet, r), \ldots, \sigma_n(t, t+\bullet, r) \in \mathcal{H}$ for all $(t,r) \in \mathbb{R}_+ \times \mathcal{H}$, and that r is a $\mathcal{H}$-valued process with càdlàg paths. Then (3.3.2) is an equation of the form (3.2.1).

In view of (3.2.3), (3.2.4) from the previous section, we define the vector fields $\hat{\alpha}, \hat{\sigma}_1, \dots \hat{\sigma}_n$ on $(0,\infty) \times \mathcal{H}$ as

$$\hat{\alpha}(t,r) := (1, \tilde{\alpha}(t,r)), \tag{3.3.3}$$

$$\hat{\sigma}_i(t,r) := (0, \tilde{\sigma}_i(t,r)) \quad \text{for } i = 1,\dots,n, \tag{3.3.4}$$

where $\tilde{\alpha}, \tilde{\sigma}_1, \dots, \tilde{\sigma}_n : (0,\infty) \times \mathcal{H} \to \mathcal{H}$ are defined by

$$\begin{aligned} \tilde{\sigma}_i(t,x,r) &:= \sigma_i(t,t+x,r) \quad \text{for } i = 1,\dots,n, \\ \tilde{\alpha}(t,x,r) &:= \frac{\partial}{\partial x} r - \left\langle \tilde{\sigma}(t,x,r), \nabla_z \Psi \left(t, -\int_0^x \tilde{\sigma}(t,y,r) dy \right) \right\rangle \\ &\quad - \frac{1}{2} \sum_{i,j=1}^n c_{ij}(t) \mathbb{F}_r \tilde{\sigma}_i(t,r)[\tilde{\sigma}_j(t,r)](x). \end{aligned}$$

Provided, all assumptions imposed in Section 3.2 are fulfilled (in particular $\hat{\alpha}, \hat{\sigma}_1, \dots \hat{\sigma}_n$ must be smooth vector fields on $(0,\infty) \times \mathcal{H}$), we can apply Theorem 3.2.4 which tells us that

$$\dim \{\hat{\alpha}, \hat{\sigma}_1, \dots, \hat{\sigma}_n\}_{LA} < \infty \tag{3.3.5}$$

is a necessary condition for the existence of a finite dimensional realization of the HJM interest rate model (3.3.1). If the driving process X is just linearly non-degenerate, (3.3.5) provides a necessary condition for the existence of an affine realization, see Theorem 3.2.5.

We have explained in Section 3.1, why, from a geometric point of view, (3.3.5) does, in general, not provide a sufficient condition for the existence of a finite dimensional realization. At this point, it is worth to mention the following. In Filipović and Teichmann [33] and the following papers Filipović and Teichmann [32], Filipović and Teichmann [34] the extended Lie algebra theory is used in order to analyze a number of concrete problems concerning forward rate equations driven by Wiener processes. In particular, Filipović and Teichmann prove the remarkable result that any forward rate model, admitting a finite dimensional realization, must necessarily have an affine term structure. It would be nice to have such a result in our setting, where the driving processes are allowed to have jumps. This would support that in the case of forward rate models the converse of Theorem 3.2.4 might be true.

In the upcoming sections, we compute the relevant Lie algebras for concrete volatility structures. As we shall see, for all considered volatilities with finite dimensional Lie algebra, there exists a finite dimensional realization, and it is indeed affine. For concrete constructions of finite dimensional realizations, we do not need the technical assumptions imposed in Section 3.2, for instance that r is a $\mathcal{H}$-valued càdlàg process or that the driving process X is non-degenerate.

3.4 Deterministic volatility

We move on to present some applications of the theory developed above. We start with the simplest case with the volatility $\sigma(t,T)$ being deterministic, that is not dependent on $r \in \mathcal{H}$. In this special situation, the fact that the driving processes are allowed to have jumps, does not change anything essential. The results known for models based on Wiener processes, see Björk and Svensson [13, Sec. 5], and Björk and Landén [12, Prop. 4.1] for a concrete realization with stationary volatility, are valid in our setting. This coincides with the results of Björk and Gombani [11, Sec. 8] where the situation of deterministic volatility is analyzed for forward rate models with an additional driving marked point processes.

Before we turn to interest rate models with deterministic volatility structures, we require an auxiliary result concerning the computation of Lie algebras. Let mappings $D, \sigma_1, \ldots, \sigma_n : \mathbb{R}_+ \to \mathcal{H}$ be given. Define the vector fields $\hat{\alpha}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n$ on $(0,\infty) \times \mathcal{H}$ by

$$\hat{\alpha}(t,r) := \left(1, \frac{\partial}{\partial x} r + D(t)\right),$$
$$\hat{\sigma}_i(t,r) := (0, \sigma_i(t)) \quad \text{for } i = 1, \ldots, n.$$

We assume that $\hat{\alpha}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n$ are smooth vector fields on $(0,\infty) \times \mathcal{H}$.

3.4.1 Lemma. *The Lie algebra generated by $\hat{\alpha}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n$ is given by*

$$\{\hat{\alpha}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n\}_{LA} = \operatorname{span}\left\{\hat{\alpha}, \left(0, \left(\tfrac{\partial}{\partial t} - \tfrac{\partial}{\partial x}\right)^j \sigma_i\right) \mid i = 1, \ldots, n,\ j \in \mathbb{N}_0\right\}.$$

Proof. See the appendix. □

Now, let a HJM term structure model of the type (3.3.1) with deterministic volatilities $\sigma_1, \ldots, \sigma_n : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ be given. We assume that the n-dimensional driving process X is non-degenerate. Recall that Assumptions 3.3.2 always have to be fulfilled. In view of Proposition 2.2.11, we presume that the drift is given by

$$\alpha(t,T) = -\left\langle \sigma(t,T), \nabla_z \Psi\left(t, -\int_t^T \sigma(t,s)ds\right)\right\rangle, \quad t, T \in \mathbb{R}_+.$$

In order to apply Theorem 3.2.4, we have to compute the Lie algebra of the vector fields $\hat{\alpha}, \hat{\sigma}_1, \ldots, \hat{\sigma}_n$ from (3.3.3) and (3.3.4), which are for deterministic volatilities given by

$$\hat{\alpha}(t,r) = (1, \tilde{\alpha}(t,r)),$$
$$\hat{\sigma}_i(t,r) = (0, \tilde{\sigma}_i(t)) \quad \text{for } i = 1, \ldots, n,$$

where $\tilde{\alpha} : (0,\infty) \times \mathcal{H} \to \mathcal{H}$ and $\tilde{\sigma}_1, \ldots, \tilde{\sigma}_n : (0,\infty) \to \mathcal{H}$ are

$$\tilde{\sigma}_i(t,x) = \sigma_i(t, t+x) \quad \text{for } i = 1, \ldots, n,$$
$$\tilde{\alpha}(t,x,r) = \frac{\partial}{\partial x} r - \left\langle \tilde{\sigma}(t,x), \nabla_z \Psi\left(t, -\int_0^x \tilde{\sigma}(t,y)dy\right)\right\rangle.$$

Note that $\mathbb{F}_r\tilde{\sigma}_i(t,r)[\tilde{\sigma}_j(t,r)] = 0$, because the volatility is deterministic. Assuming that $\hat{\alpha}, \hat{\sigma}_1, \dots, \hat{\sigma}_n$ are smooth vector fields on $(0,\infty) \times \mathcal{H}$, the relevant Lie algebra is, due to Lemma 3.4.1,

$$\{\hat{\alpha}, \hat{\sigma}_1, \dots, \hat{\sigma}_n\}_{LA} = \operatorname{span}\Big\{\hat{\alpha}, \Big(0, \big(\tfrac{\partial}{\partial t} - \tfrac{\partial}{\partial x}\big)^j \tilde{\sigma}_i, \Big) \,|\, i = 1, \dots, n,\ j \in \mathbb{N}_0\Big\}.$$

Provided, the Lie algebra is finite dimensional at some point in $(t,r) \in (0,\infty) \times \mathcal{H}$, we obtain

$$\dim \operatorname{span}\Big\{\big(\tfrac{\partial}{\partial t} - \tfrac{\partial}{\partial x}\big)^j \tilde{\sigma}_i(t,\bullet) \,|\, i = 1, \dots, n,\ j \in \mathbb{N}_0\Big\} \le d.$$

Hence, for every $i \in \{1, \dots, n\}$ there exists an integer $m_i \in \{0, \dots, d-1\}$ and a vector $(\eta_{(i,0)}(t), \eta_{(i,1)}(t), \dots, \eta_{(i,m_i+1)}(t)) \neq 0$ such that

$$\sum_{j=0}^{m_i+1} \eta_{(i,j)}(t) \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)^j \tilde{\sigma}_i(t,x) = 0, \quad x \in \mathbb{R}_+.$$

Because of the identity $(\frac{\partial}{\partial t} - \frac{\partial}{\partial x})\sigma_i(t, t+x) = \frac{\partial}{\partial t}\sigma_i(t,T)\big|_{T=t+x}$, we obtain that the mappings $\sigma_i(t,\bullet) : [t,\infty) \to \mathbb{R}$, $i = 1, \dots, n$ satisfy the differential equation

$$\sum_{j=0}^{m_i+1} \eta_{(i,j)}(t) \frac{\partial^j}{\partial t^j}\sigma_i(t,T) = 0, \quad T \in [t,\infty). \tag{3.4.1}$$

We go on to show that, under mild regularity assumptions, the differential equations (3.4.1), satisfied for each $t \in \mathbb{R}_+$, are sufficient for the existence of a finite dimensional realization for the interest rate model (3.3.1). As we have mentioned at the end of Section 3.3, we do not need the technical assumptions from Section 3.2, e.g. that r is a $\mathcal{H}$-valued càdlàg process or that the driving process X is non-degenerate.

3.4.2 Definition. *A function $f \in C^\infty(\mathbb{R}_+ \times \mathbb{R}_+)$ is called $\partial/\partial t$-regular if there are an integer $n \in \mathbb{N}$ and mappings $\eta_i \in C(\mathbb{R}_+)$, $i = 0, \dots, n$ such that, for each $T \in \mathbb{R}_+$, the mapping $f(\bullet, T) : [0,T] \to \mathbb{R}$ satisfies the differential equation*

$$\frac{\partial^{n+1}}{\partial t^{n+1}} f(t,T) + \sum_{i=0}^{n} \eta_i(t) \frac{\partial^i}{\partial t^i} f(t,T) = 0, \quad t \in [0,T].$$

3.4.3 Remarks.

- *In the proof of the upcoming Proposition 3.4.6 and Proposition 3.5.11, we define the state process Z by means of a stochastic differential equation. For this reason, we demand that the η_i in Definition 3.4.2 are continuous, since then, the existence of a unique solution Z is ensured.*

- *If a function f is $\partial/\partial t$-regular, then it also satisfies a differential equation of the kind (3.4.1) by setting $\eta_{n+1} \equiv 1$. Then, for each $t \in \mathbb{R}_+$ the vector $(\eta_0(t), \eta_1(t), \ldots, \eta_{n+1}(t))$ is non-trivial.*

An inherent class of $\partial/\partial t$-regular functions are those of the form $f(t,T) = g(T-t)$, where g is a quasi-exponential function Björk and Svensson [13, Sec. 5], that is $\operatorname{span}\{g^{(i)} \,|\, i \in \mathbb{N}_0\}$ has finite dimension. In this case, the η_i can be chosen to be constant, and are therefore continuous.

3.4.4 Definition. *A function $f : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is $\mathfrak{L}$-Lipschitz if there exists a function $L \in \mathfrak{L}$ (the space of left-continuous functions admitting right-hand limits) such that*

$$|f(t,x) - f(t,y)| \le L(t)\|x-y\|_1, \quad t \in \mathbb{R}_+ \text{ and } x, y \in \mathbb{R}^n.$$

If f, g are $\mathfrak{L}$-Lipschitz with functions L, M, then $f + g$ is obviously $\mathfrak{L}$-Lipschitz with function $L + M$.

For a semimartingale X and a stopping time τ, the process $X^{\tau-}$, stopped at time $\tau-$, is defined as

$$X_t^{\tau-}(\omega) := X_t(\omega)\mathbf{1}_{[0,\tau(\omega))}(t) + X_{\tau(\omega)-}(\omega)\mathbf{1}_{[\tau(\omega),\infty)}(t).$$

The space $\mathbb{D}$ denotes the space of adapted càdlàg processes.

3.4.5 Lemma. *Let $f \in C(\mathbb{R}_+ \times \mathbb{R}^n)$ be $\mathfrak{L}$-Lipschitz. Then, the operator $F : \mathbb{D}^n \to \mathbb{D}$ defined as $F(X)_t := f(t, X_t)$ is functional Lipschitz in the sense of Protter [55, p. 195].*

Proof. By hypothesis, it follows immediately that the operator F is process Lipschitz, i.e. for any $X, Y \in \mathbb{D}^n$ and for any stopping time τ, the identity $X^{\tau-} = Y^{\tau-}$ implies $F(X)^{\tau-} = F(Y)^{\tau-}$, and it holds for each $t \in \mathbb{R}_+$

$$\|F(X)_t - F(Y)_t\| \le L(t)\|X_t - Y_t\|.$$

Therefore, see Protter [55, page 195], the operator F is functional Lipschitz. □

3.4.6 Proposition. *Assume $\sigma_1, \ldots, \sigma_n$ are $\partial/\partial t$-regular. Then, the interest rate model (3.3.1) admits an affine realization in the sense of Definition 2.2.3.*

Proof. By hypothesis, there exist $m_1, \ldots, m_n \in \mathbb{N}_0$ and $\eta_{(i,j)} \in C(\mathbb{R}_+)$ for $(i,j) \in V := \{(i,j) \,|\, i = 1, \ldots, n, j = 0, \ldots, m_i\}$ such that, for all $T \in \mathbb{R}_+$, the mappings $\sigma_i(\bullet, T) : [0,T] \to \mathbb{R}$, $i = 1, \ldots, n$ satisfy the differential equations

$$\frac{\partial^{m_i+1}}{\partial t^{m_i+1}}\sigma_i(t,T) + \sum_{j=0}^{m_i} \eta_{(i,j)}(t)\frac{\partial^j}{\partial t^j}\sigma_i(t,T) = 0, \quad t \in [0,T]. \tag{3.4.2}$$

Note that $d := |V| = n + \sum_{j=1}^{n} m_j$. Define the function $F : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ by

$$F(t, T, z_{(i,j)} \,|\, (i,j) \in V) := f^*(0,T) + \int_0^t \alpha(s,T)ds + \sum_{i=1}^{n} \sum_{j=0}^{m_i} \frac{\partial^j}{\partial t^j} \sigma_i(t,T) z_{(i,j)}.$$

Consider the following d-dimensional stochastic differential equation (3.4.3) for the state process $Z = (Z_{(i,j)} \,|\, (i,j) \in V)$.

$$\begin{cases} dZ_{(i,0)}(t) &= \eta_{(i,0)}(t) Z_{(i,m_i)}(t)dt + dX_t^i, & i = 1, \ldots, n \\ dZ_{(i,j)}(t) &= \left(\eta_{(i,j)}(t) Z_{(i,m_i)}(t) - Z_{(i,j-1)}(t)\right) dt, & i = 1, \ldots, n,\ j = 1, \ldots, m_i \\ Z(0) &= 0 \end{cases} . \tag{3.4.3}$$

Note that the coefficients appearing in the stochastic differential equation (3.4.3) are (up to summation) given by operators $G : \mathbb{D}^d \to \mathbb{D}$ of the form $G(Z)_t = g(t) Z_{(i,j)}(t)$ for some $(i,j) \in V$ and a continuous function g. Using Lemma 3.4.5, the operators are functional Lipschitz. According to Protter [55, Thm V.7], the stochastic differential equation (3.4.3) has a unique solution Z, which is a semimartingale. By Itô's formula Jacod and Shiryaev [42, Thm I.4.57] we obtain for fixed $T \in \mathbb{R}_+$

$$\begin{aligned} dF(t,T,Z_t) &= \frac{\partial}{\partial t} F(t,T,Z_t)dt + \sum_{i=1}^{n} \sum_{j=0}^{m_i} \frac{\partial}{\partial z_{(i,j)}} F(t,T,Z_{t-}) dZ_{(i,j)}(t) \\ &= \alpha(t,T)dt + \sum_{i=1}^{n} \sum_{j=0}^{m_i} \frac{\partial^{j+1}}{\partial t^{j+1}} \sigma_i(t,T) Z_{(i,j)}(t)dt + \sum_{i=1}^{n} \sum_{j=0}^{m_i} \frac{\partial^j}{\partial t^j} \sigma_i(t,T) dZ_{(i,j)}(t). \end{aligned}$$

Inserting the dynamics (3.4.3) of Z, we obtain

$$\begin{aligned} dF(t,T,Z_t) = \alpha(t,T)dt &+ \sum_{i=1}^{n} \sum_{j=0}^{m_i} \frac{\partial^{j+1}}{\partial t^{j+1}} \sigma_i(t,T) Z_{(i,j)}(t)dt \\ &+ \sum_{i=1}^{n} \Big(\eta_{(i,0)}(t) \sigma_i(t,T) Z_{(i,m_i)}(t)dt + \sigma_i(t,T) dX_t^i \Big) \\ &+ \sum_{i=1}^{n} \sum_{j=1}^{m_i} \frac{\partial^j}{\partial t^j} \sigma_i(t,T) \Big(\eta_{(i,j)}(t) Z_{(i,m_i)}(t) - Z_{(i,j-1)}(t) \Big) dt. \end{aligned}$$

Because of the differential equations (3.4.2) it holds for each $i = 1, \ldots, n$

$$\begin{aligned} &\sum_{j=0}^{m_i} \left(\frac{\partial^{j+1}}{\partial t^{j+1}} \sigma_i(t,T) Z_{(i,j)}(t) \right) + \eta_{(i,0)}(t) \sigma_i(t,T) Z_{(i,m_i)}(t) \\ &+ \sum_{j=1}^{m_i} \frac{\partial^j}{\partial t^j} \sigma_i(t,T) \Big(\eta_{(i,j)}(t) Z_{(i,m_i)}(t) - Z_{(i,j-1)}(t) \Big) = 0, \end{aligned}$$

whence we get

$$dF(t,T,Z_t) = \alpha(t,T)dt + \sigma(t,T)dX_t.$$

Taking into account $Z_0 = 0$, which implies $F(0,\bullet,Z_0) = f^*(0,\bullet)$, we have proven that the pair (F,Z) provides a $d+1$-dimensional affine realization. □

Note that the mapping F of the realization (F,Z) is affine. Referring to Section 3.1, this explains, geometrically, why the appearance of jumps does not matter in the present situation.

3.5 Deterministic direction volatility

We go on to study the more interesting case of deterministic direction volatility $\sigma(t,T,r) = \varphi(t,r)\lambda(t,T)$. In the time-homogeneous case $\sigma(t,x) = \varphi(r)\lambda(x)$, the vector field has constant direction, but varying length, which explains the term "deterministic direction volatility". Unlike the results of the preceding section, the occurrence of jumps now gives rise to another behavior concerning finite dimensional realizations. Consequently, our results differ from those in Björk and Svensson [13, Sec. 6], which are obtained for driving Wiener processes. The computation of Lie algebras, carried out below, leads us to prove the existence of local approximative realizations.

We start with the study of the one-dimensional case, i.e. let the driving process X with deterministic derivative $(\beta,c,K)^D$ for the interest rate model (3.3.1) be one-dimensional, and non-degenerate. As announced, the model is supposed to be of deterministic direction volatility type $\sigma(t,T,r) = \varphi(t,r)\lambda(t,T)$ with $\varphi : \mathbb{R}_+ \times \mathcal{H} \to \mathbb{R}$ and $\lambda : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$. With regard to Proposition 2.2.11, we presume that the drift is given by

$$\alpha(t,T,r) = -\sigma(t,T,r)\frac{\partial}{\partial z}\Psi\left(t, -\int_t^T \sigma(t,s,r)ds\right)$$

for all $t,T \in \mathbb{R}_+$ and $r \in \mathcal{H}$. We set for each $t,x \in \mathbb{R}_+$

$$\tilde{\lambda}(t,x) := \lambda(t,t+x),$$
$$\tilde{D}(t,x) := \int_0^x \tilde{\lambda}(t,y)dy.$$

In order to check the existence of a finite dimensional realization, we have to consider the vector fields $\hat{\alpha},\hat{\sigma}$ from (3.3.3) and (3.3.4), which are in this case

$$\hat{\alpha}(t,r) = (1,\tilde{\alpha}(t,r)),$$
$$\hat{\sigma}(t,r) = (0,\tilde{\sigma}(t,r)),$$

where $\tilde{\alpha}, \tilde{\sigma} : (0, \infty) \times \mathcal{H} \to \mathcal{H}$ are given by

$$\begin{aligned}\tilde{\sigma}(t, x, r) &= \varphi(t, r)\tilde{\lambda}(t, x),\\ \tilde{\alpha}(t, x, r) &= \frac{\partial}{\partial x} r - \varphi(t, r)\tilde{\lambda}(t, x)\frac{\partial}{\partial z}\Psi\big(t, -\varphi(t, r)\tilde{D}(t, x)\big)\\ &\quad - \frac{1}{2}c(t)\varphi(t, r)\mathbb{F}_r\varphi(t, r)[\tilde{\lambda}(t)]\tilde{\lambda}(t, x),\end{aligned}$$

because $\mathbb{F}_r\tilde{\sigma}(t, r)[\tilde{\sigma}(t, r)] = \varphi(t, r)\mathbb{F}_r\varphi(t, r)[\tilde{\lambda}(t)]\tilde{\lambda}(t)$. As usual, we assume that these are smooth vector fields on $(0, \infty) \times \mathcal{H}$. If the driving process X has discontinuities, not all of the functions $\Psi(t, \bullet)$ are polynomials (see Lemma 2.2.7), which makes the structure of the Lie algebra $\{\hat{\alpha}, \hat{\sigma}\}_{LA}$ much more complicated. However, at least each $\Psi(t, \bullet)$ is holomorphic (see Section 2.2), so we can use the power series representations

$$\frac{\partial}{\partial z}\Psi(t, z) = \sum_{i=0}^{\infty} a_i(t) z^i, \tag{3.5.1}$$

valid on any open interval $(-M, M) \subset Q$, with $a_i : \mathbb{R}_+ \to \mathbb{R}$, $i \in \mathbb{N}_0$.

At this point, we assume that for some $M > 0$ it holds $Q_0 \subset (-M, M) \subset Q$, where Q and Q_0 denote the compact sets from Assumptions 3.3.2. This ensures that the power series representation (3.5.1) is valid on Q_0.

Our idea is as follows. Consider the following interest rate models $f_m(t, T)$ for $m \in \mathbb{N}$

$$\begin{cases} df_m(t, T) &= \alpha_m(t, T, r_{t-})dt + \sigma(t, T, r_{t-})dX_t \\ f_m(0, T) &= f^*(0, T) \end{cases}, \tag{3.5.2}$$

where the only differences are the drift terms α_m, which we define as

$$\alpha_m(t, T, r) := -\sigma(t, T, r)\frac{\partial}{\partial z}\Psi_m\left(t, -\int_t^T \sigma(t, s, r)ds\right)$$

for $t, T \in \mathbb{R}_+$ and $r \in \mathcal{H}$, where the $\Psi_m : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ for $m \in \mathbb{N}$ are chosen such that $\Psi_m(t, 0) = \Psi(t, 0) = a_0(t)$, $t \in \mathbb{R}_+$ and

$$\frac{\partial}{\partial z}\Psi_m(t, z) = \sum_{i=0}^{m} a_i(t) z^i, \quad (t, z) \in \mathbb{R}_+ \times \mathbb{R}.$$

Note that $\Psi_m \to \Psi$ as $m \to \infty$, whence $\lim_{m\to\infty} \alpha_m = \alpha$. Introduce the new vector fields $\hat{\alpha}_m$, $m \in \mathbb{N}$ on $(0, \infty) \times \mathcal{H}$ by

$$\hat{\alpha}_m(t, r) := (1, \tilde{\alpha}_m(t, r)),$$

where the $\tilde{\alpha}_m : (0,\infty) \times \mathcal{H} \to \mathcal{H}$ are given by

$$\begin{aligned}\tilde{\alpha}_m(t,x,r) := \frac{\partial}{\partial x} r - \varphi(t,r)\tilde{\lambda}(t,x)\frac{\partial}{\partial z}\Psi_m(t, -\varphi(t,r)\tilde{D}(t,x)) \\ - \frac{1}{2}c(t)\varphi(t,r)\mathbb{F}_r\varphi(t,r)[\tilde{\lambda}(t)]\tilde{\lambda}(t,x).\end{aligned}$$

We assume that these are smooth vector fields on $(0,\infty) \times \mathcal{H}$. The computation of the Lie algebras $\{\hat{\alpha}_m, \hat{\sigma}\}_{LA}$ can be accomplished. We set for $m \in \mathbb{N}$

$$J_m := \{j \in \{0,1,\dots,m\} \mid a_j(t) \neq 0 \text{ for some } t \in (0,\infty)\}.$$

3.5.1 Lemma. *Assume that for all* $(t,r) \in (0,\infty) \times \mathcal{H}$

$$\begin{aligned}\varphi(t,r) &\neq 0, \\ \det\Big(a_j(t)\mathbb{F}_r^{i+1}(\varphi^{j+1}(t,r))[\tilde{\lambda}(t);\dots;\tilde{\lambda}(t)]\Big)_{\substack{i=1,\dots,|J_m| \\ j\in J_m}} &\neq 0, \quad m \in \mathbb{N}.\end{aligned}$$

Then it holds for each $m \in \mathbb{N}$

$$\{\hat{\alpha}_m, \hat{\sigma}\}_{LA} = \operatorname{span}\{(1, \tfrac{\partial}{\partial x} r), (0, (\tfrac{\partial}{\partial t} - \tfrac{\partial}{\partial x})^i(\tilde{\lambda}\tilde{D}^j)) \mid i \in \mathbb{N}_0, j \in J_m \cup \{0\}\}.$$

Proof. See the appendix. □

3.5.2 Remarks.

- *Note that the* $\mathbb{F}_r^{i+1}(\varphi^{j+1}(t,r))[\tilde{\lambda}(t);\dots;\tilde{\lambda}(t)]$ *denote the* $i+1$ *- th order Fréchet derivative of* φ^{j+1} *with respect to* r*, operating on the* $i+1$*-dimensional vector with each entry* $\tilde{\lambda}(t)$*.*

- *We need the condition* $\varphi(t,r) \neq 0$ *for the proof. It is also assumed in Björk and Svensson [13, Sec. 6].*

- *If* α *shall describe the drift of an interest rate model driven by a standard Wiener process, one has* $\frac{\partial}{\partial z}\Psi(t,z) = z$*, and the non-singularity condition of the matrix becomes*

  $$\mathbb{F}_r^2\Phi(t,r)[\tilde{\lambda}(t);\tilde{\lambda}(t)] \neq 0,$$

  where $\Phi(t,r) := \varphi^2(t,r)$*, and is thus consistent with Björk and Svensson [13]. In this case, the Lie algebra is given by*

  $$\{\hat{\alpha}, \hat{\sigma}\}_{LA} = \operatorname{span}\{(1, \tfrac{\partial}{\partial x} r), (0, (\tfrac{\partial}{\partial t} - \tfrac{\partial}{\partial x})^i\tilde{\lambda}), (0, (\tfrac{\partial}{\partial t} - \tfrac{\partial}{\partial x})^i(\tilde{\lambda}\tilde{D})) \mid i \in \mathbb{N}_0\}.$$

With regard to Lemma 3.5.1, it is reasonable to assume that all λD^j, $j \in \mathbb{N}_0$ are $\partial/\partial t$-regular, where $D(t,T) := \int_t^T \lambda(t,s)ds$ (recall the treatment of deterministic volatility in Section 3.4, where the relevant Lie algebra was of the same structural type). This looks like a very restrictive condition at first glance, but is for instance satisfied if λ is stationary, i.e. $\lambda(t,T) = \tilde{\lambda}(T-t)$, and $\tilde{\lambda}$ is quasi-exponential, because products and primitives of quasi-exponential functions are again quasi-exponential Björk and Svensson [13, Lemma 5.1].

In this case, the Lie algebras $\{\hat{\alpha}_m, \hat{\sigma}\}_{LA}$, $m \in \mathbb{N}$ are, according to Lemma 3.5.1, finite dimensional, and their dimension grows with increasing m. This suggests, but does not prove, that no finite dimensional realization for the term structure model exists if the driving processes X has discontinuous paths (by Lemma 2.2.7, the process X has continuous paths if and only if each $\Psi(t, \bullet)$ is a polynomial), and that the intrinsic reason is the infinite dimensional structure of the $\Psi(t, \bullet)$, which are not polynomials. A later result, namely Theorem 4.4.1, which is proven in the framework of benchmark realizations, supports this conjecture. Nevertheless, Lemma 3.5.1 also suggests that there might exist a sequence of finite dimensional systems (of increasing dimension) converging to the forward rate model, and we arrive at the following Definition 3.5.3. We will establish such a convergence in a local sense. Recall that for a semimartingale X and a stopping time τ, the process $X^{\tau-}$, stopped at time $\tau-$, is defined as

$$X_t^{\tau-}(\omega) := X_t(\omega)\mathbf{1}_{[0,\tau(\omega))}(t) + X_{\tau(\omega)-}(\omega)\mathbf{1}_{[\tau(\omega),\infty)}(t).$$

For a process $X \in \mathbb{D}$ (the space of adapted càdlàg processes) and $1 \le p \le \infty$ we define

$$\|X\|_{S^p} := \left\| \sup_{t \in \mathbb{R}_+} |X_t| \right\|_{L^p},$$

which is the S^p-norm defined in Protter [55, p. 188].

We say that a sequence $(X_m)_{m\in\mathbb{N}}$ of processes in $\mathbb{D}$ converges to $X \in \mathbb{D}$ in the S^p-sense, denoted by $X_m \xrightarrow{S^p} X$, if $\|X_m - X\|_{S^p} \to 0$ as $m \to \infty$.

3.5.3 Definition. *A term structure model $f(t,T)$ has a <u>local approximative realization</u> if there exist a stopping time τ, which is $\mathbb{P}-a.s.$ positive, and a sequence $(G_m, Z_m)_{m\in\mathbb{N}}$ of finite dimensional semimartingales Z_m and mappings $G_m : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^{d_m} \to \mathbb{R}$, where d_m denotes the dimension of Z_m, such that for all $1 \le p \le \infty$ and $x \in \mathbb{R}_+$*

$$G_m(\bullet, x, Z_m)^{\tau-} \xrightarrow{S^p} r(x)^{\tau-} \quad \textit{as } m \to \infty.$$

At this point, we turn to term structure models of the form (3.3.1) that are slightly more general. Namely, now the non-degenerate driving process X with deterministic derivative $(\beta, c, K)^D$ may be n-dimensional for some $n \in \mathbb{N}$, and the volatility is of the form

$$\sigma_i(t,T,r) = \sum_{j=1}^{m_i} \varphi_{(i,j)}(t,r)\lambda_{(i,j)}(t,T), \quad i = 1, \dots, n \tag{3.5.3}$$

with $m_1, \dots, m_n \in \mathbb{N}$, $\varphi_{(i,j)} : \mathbb{R}_+ \times \mathcal{H} \to \mathbb{R}$ and $\lambda_{(i,j)} : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ for $i = 1, \dots, n$ and $j = 1, \dots, m_i$. Volatility structures of this type with finitely many driving standard Wiener processes have been studied in Björk and Landén [12, Sec. 6.1]. With regard to Proposition 2.2.11, we presume that the drift is given by

$$\alpha(t,T,r) = -\Big\langle \sigma(t,T,r), \nabla_z \Psi \left(t, -\int_t^T \sigma(t,s,r)ds \right) \Big\rangle$$

for all $t, T \in \mathbb{R}_+$ and $r \in \mathcal{H}$. In view of Definition 3.5.3 we turn our attention to interest rate models

$$\begin{cases} df(t,T) &= \alpha^*(t,T,r_{t-})dt + \sigma(t,T,r_{t-})dX_t \\ f(0,T) &= f^*(0,T) \end{cases}, \tag{3.5.4}$$

which are, with exception of the drift term, like the models introduced above. We define the drift as

$$\alpha^*(t,T,r) := -\Big\langle \sigma(t,T,r), \nabla_z \Psi^* \left(t, -\int_t^T \sigma(t,s,r)ds \right) \Big\rangle$$

for $t, T \in \mathbb{R}_+$ and $r \in \mathcal{H}$, where $\Psi^* : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is of the form

$$\Psi^*(t,z) = \sum_{u=1}^{e} a_{k_u}(t) z_1^{k_u^1} \cdots z_n^{k_u^n} \tag{3.5.5}$$

with $e \in \mathbb{N}, k_1, \dots, k_e \in \mathbb{N}_0^n$ and $a_{k_1}, \dots, a_{k_e} : \mathbb{R}_+ \to \mathbb{R}$. Set for $i = 1, \dots, n$ and $j = 1, \dots, m_i$

$$\begin{aligned} \tilde{\lambda}_{(i,j)}(t,x) &:= \lambda_{(i,j)}(t, t+x), \\ \tilde{D}_{(i,j)}(t,x) &:= \int_0^x \tilde{\lambda}_{(i,j)}(t,y)dy. \end{aligned}$$

In order to check the existence of a finite dimensional realization for the modified term structure model (3.5.4), consider the vector fields $\hat{\alpha}^*, \hat{\sigma}_1, \dots, \hat{\sigma}_n$, defined in (3.3.3) and (3.3.4), which are in the present situation

$$\begin{aligned} \hat{\alpha}^*(t,r) &= (1, \tilde{\alpha}^*(t,r)), \\ \hat{\sigma}_i(t,r) &= (0, \tilde{\sigma}_i(t,r)), \quad i = 1, \dots, n \end{aligned}$$

where $\tilde{\alpha}^*, \tilde{\sigma}_1, \dots, \tilde{\sigma}_n : (0,\infty) \times \mathcal{H} \to \mathcal{H}$ are given by

$$\begin{aligned} \tilde{\sigma}_i(t,x,r) &= \sum_{j=1}^{m_i} \varphi_{(i,j)}(t,r) \tilde{\lambda}_{(i,j)}(t,x), \quad i = 1, \dots, n \\ \tilde{\alpha}^*(t,x,r) &= \frac{\partial}{\partial x} r - \Big\langle \tilde{\sigma}(t,x,r), \nabla_z \Psi^* \left(t, -\int_0^x \tilde{\sigma}(t,y,r)dy \right) \Big\rangle \\ &\quad - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{m_i} \sum_{l=1}^{m_j} c_{ij}(t) \varphi_{(j,l)}(t,r) \mathbb{F}_r \varphi_{(i,k)}(t,r) [\tilde{\lambda}_{(j,l)}(t)] \tilde{\lambda}_{(i,k)}(t,x), \end{aligned}$$

because

$$\mathbb{F}_r\tilde{\sigma}_i(t,r)[\tilde{\sigma}_j(t,r)] = \sum_{k=1}^{m_i}\sum_{l=1}^{m_j} \varphi_{(j,l)}(t,r)\mathbb{F}_r\varphi_{(i,k)}(t,r)[\tilde{\lambda}_{(j,l)}(t)]\tilde{\lambda}_{(i,k)}(t).$$

Let $V := \{(v,w) \,|\, v = 1,\dots,n,\ w = 1,\dots,m_v\}$, and denote by U the set of all (u,v,w,l), where $v = 1,\dots,n,\ w = 1,\dots,m_v,\ u = 1,\dots,e$ with $k_u^v \neq 0$, and $l_{(i,j)} \in \mathbb{N}_0$, $i = 1,\dots,n,\ j = 1,\dots,m_i$ such that

$$\sum_{j=1}^{m_i} l_{(i,j)} = \begin{cases} k_u^i, & i \neq v \\ k_u^i - 1, & i = v \end{cases}.$$

For each $(u,v,w,l) \in U$ we define

$$\varphi_{(u,v,w,l)}(t,r) := -\varphi_{(v,w)}(t,r)\prod_{i=1}^{n}\prod_{j=1}^{m_i}\left(-\varphi_{(i,j)}(t,r)\right)^{l_{(i,j)}},$$

$$\tilde{\lambda}_{(u,v,w,l)}(t,x) := \tilde{\lambda}_{(v,w)}(t,x)\prod_{i=1}^{n}\prod_{j=1}^{m_i}\left(\tilde{D}_{(i,j)}(t,x)\right)^{l_{(i,j)}}.$$

3.5.4 Lemma. *For all $(t,r) \in (0,\infty)\times\mathcal{H}$ and $x \in \mathbb{R}_+$ the identity*

$$-\left\langle \tilde{\sigma}(t,x,r), \nabla_z\Psi^*\left(t, -\int_0^x \tilde{\sigma}(t,y,r)dy\right)\right\rangle = \sum_{u\in U} b_u(t)\varphi_u(t,r)\tilde{\lambda}_u(t,x)$$

is valid, where the $(b_u)_{u\in U}$ are given by

$$b_{(u,v,w,l)}(t) := a_{k_u}(t)\left(\prod_{i=1}^{n}\frac{k_u^i!}{l_{(i,1)}!\cdots l_{(i,m_i)}!}\right).$$

Proof. See the appendix. □

As usual, we assume that the vector fields $\hat{\alpha}^*, \hat{\sigma}_1,\dots,\hat{\sigma}_n$ are smooth vector fields on $(0,\infty)\times\mathcal{H}$. In this setting, the structure of the Lie algebra becomes more complicated than in Lemma 3.5.1, and we restrict ourselves to find spaces that include the Lie algebra.

3.5.5 Lemma. *Assume for each $i \in \{1,\dots,n\}$ there exists an index $j \in \{1,\dots,m_i\}$ such that*

$$\varphi_{(i,j)}(t,r) \neq 0 \quad \textit{for all } (t,r) \in (0,\infty)\times\mathcal{H}.$$

Then, the following inclusion is valid.

$$\{\hat{\alpha}^*, \hat{\sigma}_1,\dots,\hat{\sigma}_n\}_{LA} \subset \operatorname{span}\{(1, \tfrac{\partial}{\partial x}r), (0, (\tfrac{\partial}{\partial t} - \tfrac{\partial}{\partial x})^j\tilde{\lambda}_v) \,|\, j \in \mathbb{N}_0, v \in V\cup U\}.$$

Proof. See the appendix. □

By Lemma 3.5.5, the necessary condition for the existence of a finite dimensional realization is fulfilled, provided the λ_v, $v \in V \cup U$, where the λ_u, $u \in U$ are defined later in (3.5.8), are $\partial/\partial t$-regular. As we shall see next, in this case there exists in fact a finite dimensional realization. For technical reasons, we confine ourselves to volatility structures which may depend on finitely many points of the forward rate curve r_t.

For the rest of this section, we choose $d \in \mathbb{N}$ and benchmark points $0 \leq x_1 < \ldots < x_d$. The process r denotes $r_t = (f(t, t + x_1), \ldots, f(t, t + x_d))$. Since we now intend to give concrete realizations, the technical assumptions imposed in Section 3.2 are no longer required, in particular the driving process X does not need to be non-degenerate.

Unless stated otherwise, the interest rate models appearing in the rest of this section have to satisfy the following.

3.5.6 Assumptions.

1. *The derivative* $(\beta, c, K)^D$ *of the* n*-dimensional driving process* X *is deterministic.*

2. *There are* $z_1^-, \ldots, z_n^- \in (-\infty, 0)$ *and* $z_1^+, \ldots, z_n^+ \in (0, \infty)$ *such that for any* $t \in \mathbb{R}_+$

$$\int_0^t \left(\int_{|x| \leq 1} |x|^2 K_s(dx) + \int_{|x|>1} e^{\langle z,x \rangle} K_s(dx) \right) ds < \infty, \quad z \in Q,$$

where Q *denotes the set* $Q := [z_1^-, z_1^+] \times \ldots \times [z_n^-, z_n^+]$*, and furthermore, for all* $t \in \mathbb{R}_+$ *it holds*

$$\int_{|x|>1} e^{\langle z,x \rangle} K_t(dx) < \infty, \quad z \in Q.$$

3. *There are* $w_1^- \in (z_1^-, 0), \ldots, w_n^- \in (z_n^-, 0)$ *and* $w_1^+ \in (0, z_1^+), \ldots, w_n^+ \in (0, z_n^+)$ *such that*

$$-\int_t^T \sigma(t, s, r) ds \in Q_0, \quad (t, T, r) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d.$$

where $Q_0 \subset Q$ *is defined as* $Q_0 := [w_1^-, w_1^+] \times \ldots \times [w_n^-, w_n^+]$.

The second assumption ensures that the cumulant generating function

$$\Psi(t, z) := \langle \beta_t, z \rangle + \frac{1}{2} \langle z, c_t z \rangle + \int_{\mathbb{R}^n} \left(e^{\langle z,x \rangle} - 1 - \langle z, x \rangle \right) K_t(dx)$$

is definable for all $(t, z) \in \mathbb{R}_+ \times Q$ (see Section 2.2). The third assumption guarantees that

$$\Psi\left(t, -\int_t^T \sigma(t, s, r) ds\right)$$

exists for all $(t, T, r) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d$. The term structure models, which we will study, are of the type

$$\begin{cases} df(t,T) &= \alpha(t,T,r_{t-})dt + \sigma(t,T,r_{t-})dX_t \\ f(0,T) &= f^*(0,T) \end{cases}, \tag{3.5.6}$$

with $f^*(0, \bullet) \in C(\mathbb{R}_+)$, and where the volatility structure is assumed to be of the form

$$\sigma_i(t, T, r) = \sum_{j=1}^{m_i} \varphi_{(i,j)}(t, r)\lambda_{(i,j)}(t, T), \quad i = 1, \ldots, n$$

with $m_1, \ldots, m_n \in \mathbb{N}$, $\varphi_{(i,j)} \in C(\mathbb{R}_+ \times \mathbb{R}^d)$ and $\lambda_{(i,j)} \in C(\mathbb{R}_+ \times \mathbb{R}_+)$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m_i$. We assume that $\mathbb{P}$ is a martingale measure, whence, in view of Proposition 2.2.11, we demand the drift is equal to

$$\alpha(t, T, r) = -\left\langle \sigma(t, T, r), \nabla_z \Psi\left(t, -\int_t^T \sigma(t, s, r) ds\right) \right\rangle$$

for $t, T \in \mathbb{R}_+$ and $r \in \mathbb{R}^d$. With regard to Lemma 3.5.5, we will, beside these arbitrage free forward rate models, also take into account models of the form

$$\begin{cases} df(t,T) &= \alpha^*(t,T,r_{t-})dt + \sigma(t,T,r_{t-})dX_t \\ f(0,T) &= f^*(0,T) \end{cases}, \tag{3.5.7}$$

where the only difference to the model (3.5.6) is the drift α^*, which is assumed to be given by

$$\alpha^*(t, T, r) := -\left\langle \sigma(t, T, r), \nabla_z \Psi^*\left(t, -\int_t^T \sigma(t, s, r) ds\right) \right\rangle$$

for $t, T \in \mathbb{R}_+$ and $r \in \mathbb{R}^d$, where $\Psi^* : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}$ is of the form

$$\Psi^*(t, z) = \sum_{u=1}^{e} a_{k_u}(t) z_1^{k_u^1} \cdots z_n^{k_u^n}$$

with $e \in \mathbb{N}, k_1, \ldots, k_e \in \mathbb{N}_0^n$ and $a_{k_1}, \ldots, a_{k_d} : \mathbb{R}_+ \to \mathbb{R}$. In Proposition 3.5.11 we need that Ψ^* is continuous, it is therefore wise to assume that $a_{k_1}, \ldots, a_{k_d} \in C(\mathbb{R}_+)$.

For both types of term structure models, we introduce the following notation. We set for $i = 1, \dots, n$ and $j = 1, \dots, m_i$

$$D_{(i,j)}(t,T) := \int_t^T \lambda_{(i,j)}(t,s)ds.$$

Let $V := \{(v,w) \,|\, v = 1, \dots, n,\ w = 1, \dots, m_v\}$, and denote by U the set of all (u,v,w,l), where $v = 1, \dots, n$, $w = 1, \dots, m_v$, $u = 1, \dots, e$ with $k_u^v \neq 0$, and $l_{(i,j)} \in \mathbb{N}_0$, $i = 1, \dots, n$, $j = 1, \dots, m_i$ such that

$$\sum_{j=1}^{m_i} l_{(i,j)} = \begin{cases} k_u^i, & i \neq v \\ k_u^i - 1, & i = v \end{cases}.$$

For $(u,v,w,l) \in U$ we define

$$\varphi_{(u,v,w,l)}(t,r) := -\varphi_{(v,w)}(t,r) \prod_{i=1}^{n} \prod_{j=1}^{m_i} \left(-\varphi_{(i,j)}(t,r)\right)^{l_{(i,j)}},$$

$$\lambda_{(u,v,w,l)}(t,T) := \lambda_{(v,w)}(t,T) \prod_{i=1}^{n} \prod_{j=1}^{m_i} \left(D_{(i,j)}(t,T)\right)^{l_{(i,j)}}. \tag{3.5.8}$$

3.5.7 Lemma. *For all $t, T \in \mathbb{R}_+$ and $r \in \mathbb{R}^d$ the identity*

$$\alpha^*(t,T,r) = \sum_{u \in U} b_u(t) \varphi_u(t,r) \lambda_u(t,T)$$

is valid, where the $(b_u)_{u \in U}$ are given by

$$b_{(u,v,w,l)}(t) := a_{k_u}(t) \left(\prod_{i=1}^{n} \frac{k_u^i!}{l_1^i! \cdots l_{m_i}^i!} \right).$$

Proof. This is proven exactly as Lemma 3.5.4. □

3.5.8 Definition. *The space $C_b^{\mathfrak{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$ denotes the space of all $f \in C^{0,1}(\mathbb{R}_+ \times \mathbb{R}^d)$ for which there is a function $L \in \mathfrak{L}$ such that for each $t \in \mathbb{R}_+$*

$$\sup_{x \in \mathbb{R}^d} |f(t,x)| \leq L(t),$$

$$\sup_{x \in \mathbb{R}^d} \left| \frac{\partial}{\partial x_i} f(t,x) \right| \leq L(t) \quad \textit{for all } i = 1, \dots, d.$$

Note that the product of two functions $f, g \in C_b^{\mathfrak{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$ belongs again to $C_b^{\mathfrak{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$, because $\left|\frac{\partial}{\partial x_i}(fg)(t,x)\right| \leq \left|f(t,x)\frac{\partial}{\partial x_i}g(t,x)\right| + \left|g(t,x)\frac{\partial}{\partial x_i}f(t,x)\right|$.

3.5.9 Lemma. *Assume $f \in C_b^{\mathfrak{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$. Then, for every $T > 0$ there is a constant $L > 0$ such that*

$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} |f(t,x)| \le L,$$
$$\sup_{(t,x)\in[0,T]\times\mathbb{R}^d} \left|\frac{\partial}{\partial x_i} f(t,x)\right| \le L \quad \text{for all } i = 1,\dots,d.$$

Proof. See the appendix □

3.5.10 Lemma. *Let $G \subset \mathbb{R}^n$ be an open, convex set, $f \in C^1(G)$, and $L > 0$ be a constant. Then, there is equivalence between*

1. $|f(x) - f(y)| \le L\|x - y\|_1$ *for all* $x, y \in G$.
2. $\sup_{x\in G} |\frac{\partial}{\partial x_i} f(x)| \le L$ *for each* $i = 1,\dots,n$.

Proof. See the appendix □

3.5.11 Proposition. *Assume $\Psi^* \in C(\mathbb{R}_+ \times \mathbb{R}^n)$, that all λ_v, $v \in V \cup U$ are $\partial/\partial t$-regular, and that $\varphi_v \in C_b^{\mathfrak{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$ for all $v \in V$. Then, the forward rate model (3.5.7) admits an affine realization in the sense of Definition 2.2.3.*

Proof. By the assumed $\partial/\partial t$-regularity, there exist, for each $v \in V \cup U$, integers $q_v \in \mathbb{N}_0$ and functions $\eta_{(v,k)} \in C(\mathbb{R}_+)$, $k = 0,\dots,q_v$ such that, for each $T \in \mathbb{R}_+$, the following differential equations are valid.

$$\frac{\partial^{q_v+1}}{\partial t^{q_v+1}}\lambda_v(t,T) + \sum_{k=0}^{q_v} \eta_{(v,k)}(t)\frac{\partial^k}{\partial t^k}\lambda_v(t,T) = 0, \quad t \in [0,T]. \tag{3.5.9}$$

Let $W := \{(v,k) \,|\, v \in V \cup U, k = 0,\dots,q_v\}$. We define the mapping $F : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^{|W|} \to \mathbb{R}$ by

$$F(t,T,z_{(v,k)} \,|\, v \in V \cup U, k = 0,\dots,q_v) := f^*(0,T) + \sum_{v\in V\cup U}\sum_{k=0}^{q_v} \frac{\partial^k}{\partial t^k}\lambda_v(t,T) z_{(v,k)}.$$

In order to define the state process $Z = (Z_{(v,k)} \,|\, v \in V \cup U, k = 0,\dots,q_v)$, consider the following $|W|$-dimensional stochastic differential equation (3.5.10), where the indices $v = (v_1, v_2)$ take all values from V, and the u take all values from U, and $\tilde{F}$ denotes $\tilde{F}(t,z) := (F(t,t+x_1,z),\dots,F(t,t+x_d,z))$ for $(t,z) \in \mathbb{R}_+ \times \mathbb{R}^{|W|}$.

$$\begin{cases} dZ_{(v,0)}(t) &= \eta_{(v,0)}(t)Z_{(v,q_v)}(t)dt + \varphi_v(t,\tilde{F}(t,Z_{t-}))dX_t^{v_1} \\ dZ_{(v,k)}(t) &= \left[\eta_{(v,k)}(t)Z_{(v,q_v)}(t) - Z_{(v,k-1)}(t)\right]dt, \quad k = 1,\dots,q_v \\ dZ_{(u,0)}(t) &= \left[\eta_{(u,0)}(t)Z_{(u,q_u)}(t) + b_u(t)\varphi_u(t,\tilde{F}(t,Z_{t-}))\right]dt \\ dZ_{(u,k)}(t) &= \left[\eta_{(u,k)}(t)Z_{(u,q_u)}(t) - Z_{(u,k-1)}(t)\right]dt, \quad k = 1,\dots,q_u \\ Z(0) &= 0 \end{cases}. \tag{3.5.10}$$

We wish to show that the operators appearing in the stochastic differential equation (3.5.10) for Z are functional Lipschitz. Those operators $G : \mathbb{D}^{|W|} \to \mathbb{D}$ in (3.5.10) on which we cannot directly apply Lemma 3.4.5, are (up to summation) of the form $G(Z)_t = g(t, Z_t)$ with a function $g \in C(\mathbb{R}_+ \times \mathbb{R}^{|W|})$ of the type

$$g(t, z) = f(t)\psi(t, \tilde{F}(t, z)),$$

where f is continuous (note that Ψ^* is continuous in t), and where ψ is, since all φ_v, $v \in V \cup U$ belong to $C_b^{\mathfrak{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$ (note that this space is closed under products), $\mathfrak{L}$-Lipschitz with a function $L \in \mathfrak{L}$ (apply Lemma 3.5.10). Thus it holds, by the definition of F, for all $t \in \mathbb{R}_+$ and $z_1, z_2 \in \mathbb{R}^{|W|}$

$$|g(t, z_1) - g(t, z_2)| \leq L(t)|f(t)| \sum_{i=1}^{d} \sum_{v \in V \cup U} \sum_{k=0}^{q_v} \left| \frac{\partial^k}{\partial t^k} \lambda_v(t, T) \Big|_{T=t+x_i} \right| \|z_1 - z_2\|_1,$$

that is, g is also $\mathfrak{L}$-Lipschitz. According to Lemma 3.4.5, the operator G is functional Lipschitz. By Protter [55, Thm. V.7] the stochastic differential equation (3.5.10) has a unique solution Z, which is a semimartingale. By Itô's formula Jacod and Shiryaev [42, Thm. I.4.57] we obtain for fixed $T \in \mathbb{R}_+$ the relation

$$\begin{aligned} dF(t, T, Z_t) = &\sum_{v \in V \cup U} \sum_{k=0}^{q_v} \frac{\partial^{k+1}}{\partial t^{k+1}} \lambda_v(t, T) Z_{(v,k)}(t) dt \\ &+ \sum_{v \in V \cup U} \sum_{k=0}^{q_v} \frac{\partial^k}{\partial t^k} \lambda_v(t, T) dZ_{(v,k)}(t). \end{aligned}$$

Incorporating the dynamics (3.5.10) we get

$$\begin{aligned} dF(t, T, Z_t) = &\sum_{v \in V \cup U} \sum_{k=0}^{q_v} \frac{\partial^{k+1}}{\partial t^{k+1}} \lambda_v(t, T) Z_{(v,k)}(t) dt \\ &+ \sum_{v \in V} \lambda_v(t, T) \left(\eta_{(v,0)}(t) Z_{(v,q_v)}(t) dt + \varphi_v(t, \tilde{F}(t, Z_{t-})) dX_t^{v_1} \right) \\ &+ \sum_{v \in V} \sum_{k=1}^{q_v} \frac{\partial^k}{\partial t^k} \lambda_v(t, T) \Big(\eta_{(v,k)}(t) Z_{(v,q_v)}(t) - Z_{(v,k-1)}(t) \Big) dt \\ &+ \sum_{u \in U} \lambda_u(t, T) \left(\eta_{(u,0)}(t) Z_{(u,q_u)}(t) + b_u(t) \varphi_u(t, \tilde{F}(t, Z_{t-})) \right) dt \\ &+ \sum_{u \in U} \sum_{k=1}^{q_u} \frac{\partial^k}{\partial t^k} \lambda_u(t, T) \Big(\eta_{(u,k)}(t) Z_{(u,q_u)}(t) - Z_{(u,k-1)}(t) \Big) dt. \end{aligned}$$

By the differential equations (3.5.9) it holds for every $v \in V \cup U$

$$\sum_{k=0}^{q_v} \frac{\partial^{k+1}}{\partial t^{k+1}} \lambda_v(t,T) Z_{(v,k)}(t) + \eta_{(v,0)}(t) \lambda_v(t,T) Z_{(v,q_v)}(t) + \sum_{k=1}^{q_v} \frac{\partial^k}{\partial t^k} \lambda_v(t,T) \Big(\eta_{(v,k)}(t) Z_{(v,q_v)}(t) - Z_{(v,k-1)}(t) \Big) = 0.$$

Taking into account the formula for α^* from Lemma 3.5.7, we arrive at

$$dF(t,T,Z_t) = \alpha^*(t,T,\tilde{F}(t,Z_{t-}))dt + \sigma(t,T,\tilde{F}(t,Z_{t-}))dX_t.$$

Since $Z_0 = 0$, which implies $F(0,\bullet,Z_0) = f^*(0,\bullet)$, we have shown that the pair (F,Z) provides an affine realization. □

We observe that the constructed realization is affine, like in the case of deterministic volatility. Proposition 3.5.11 can immediately be applied to term structure models driven by continuous local martingales.

3.5.12 Corollary. *Assume X has a deterministic derivative $(0,c,0)^D$, and the following conditions are fulfilled.*

- $c \in C(\mathbb{R}_+, \mathbb{R}^{n\times n})$;
- *Each* $\lambda_{(i,j)}$, $i = 1,\dots,n$, $j = 1,\dots,m_i$ *is* $\partial/\partial t$*-regular;*
- *Each* $\lambda_{(i,j)} D_{(k,l)}$, $i = 1,\dots,n$, $j = 1,\dots,m_i$, $k = 1,\dots,n$, $l = 1,\dots,m_k$ *is* $\partial/\partial t$*-regular;*
- *It holds* $\varphi_{(i,j)} \in C_b^{\mathcal{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$ *for all* $i = 1,\dots,n$, $j = 1,\dots,m_i$.

Then, the term structure model (3.5.6) has a finite dimensional realization.

Proof. By hypothesis, Ψ is given by $\Psi(t,z) = \frac{1}{2}\langle z, c(t)z\rangle$ for $(t,z) \in \mathbb{R}_+ \times \mathbb{R}^n$, and it is continuous. One also verifies that the required $\partial/\partial t$-regularity conditions of Proposition 3.5.11 are satisfied, which yields the existence of a finite dimensional realization. □

By $\mathcal{M}_{\mathcal{D}}(\mathbb{R}^{n\times n})$ we denote the set of all $n \times n$-matrices which are diagonal.

3.5.13 Corollary. *Assume X has a deterministic derivative $(0,c,0)^D$, and the following conditions are satisfied.*

- $c \in C(\mathbb{R}_+, \mathbb{R}^{n\times n})$ *and* $c(t) \in \mathcal{M}_{\mathcal{D}}(\mathbb{R}^{n\times n})$ *for all* $t \in \mathbb{R}_+$;
- *Each* $\lambda_{(i,j)}$, $i = 1,\dots,n$, $j = 1,\dots,m_i$ *is* $\partial/\partial t$*-regular;*
- *Each* $\lambda_{(i,j)} D_{(i,k)}$, $i = 1,\dots,n$, $j = 1,\dots,m_i$, $k = 1,\dots,m_i$ *is* $\partial/\partial t$*-regular;*

- *It holds* $\varphi_{(i,j)} \in C_b^{\mathcal{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$ *for all* $i = 1, \dots, n$, $j = 1, \dots, m_i$.

Then, the forward rate model (3.5.6) has a finite dimensional realization.

Proof. By, hypothesis, Ψ is given by $\Psi(t,z) = \frac{1}{2}\sum_{i=1}^n c_{ii}(t) z_i^2$ for $(t,z) \in \mathbb{R}_+ \times \mathbb{R}^n$, and it is continuous. One also verifies that the required $\partial/\partial t$-regularity conditions of Proposition 3.5.11 are satisfied, which yields the existence of a finite dimensional realization. □

3.5.14 Remark. *The assumptions in Corollary 3.5.13 on the driving process are in particular satisfied if X is a standard Wiener process. Note that the $\partial/\partial t$-regularity conditions correspond to condition (29) in Björk and Landén [12].*

So far, Proposition 3.5.11 could be applied to term structure models driven by continuous local martingales. For the general case, where we wish to find a local approximative realization, we need further results.

Recall that for a semimartingale X and a stopping time τ, the process $X^{\tau-}$, stopped at time $\tau-$, is defined as

$$X_t^{\tau-}(\omega) := X_t(\omega)\mathbf{1}_{[0,\tau(\omega))}(t) + X_{\tau(\omega)-}(\omega)\mathbf{1}_{[\tau(\omega),\infty)}(t).$$

Then, each trajectory $t \mapsto X_t^{\tau-}(\omega)$ is continuous at time $t = \tau(\omega)$, and constant thereafter.

3.5.15 Lemma. *Let $A \in \mathcal{V}$, r be a d-dimensional semimartingale, τ be a stopping time and $f \in C(\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d)$. Then the processes X and Y, defined as*

$$X_t := \left(\int_0^t f(s,t,r_{s-}) dA_s \right)^{\tau-},$$
$$Y_t := \int_0^t f(s, t \wedge \tau, r_{s-}^{\tau-})) dA_s^{\tau-},$$

coincide (up to indistinguishability).

Proof. Let $\omega \in \Omega$ be such that $\tau(\omega) \in (0,\infty)$, otherwise the assertion is trivial. Define the signed measures $\mu_\omega, \mu_\omega^{\tau-}$ on $((0,\infty), \mathcal{B}(0,\infty))$ by

$$\mu_\omega((a,b]) := A_b(\omega) - A_a(\omega),$$
$$\mu_\omega^{\tau-}((a,b]) := A_b^{\tau-}(\omega) - A_a^{\tau-}(\omega), \quad (0 \le a \le b).$$

We observe that

$$X_t(\omega) = Y_t(\omega) = \int_0^t f(s,t,r_{s-}(\omega)) \mu_\omega(ds) \quad \text{for } t < \tau(\omega). \tag{3.5.11}$$

Since f is continuous, $(s,t) \mapsto f(s,t,r_{s-}(\omega))$ is bounded on $[0,\tau(\omega)] \times [0,\tau(\omega)]$. Hence, applying Lebesgue's dominated convergence theorem yields

$$\begin{aligned} \lim_{t \uparrow \tau(\omega)} X_t(\omega) &= \lim_{t \uparrow \tau(\omega)} \int f(s,t,r_{s-}(\omega)) \mathbf{1}_{(0,t]}(s) \mu_\omega(ds) \\ &= \int f(s,\tau(\omega),r_{s-}(\omega)) \mathbf{1}_{(0,\tau(\omega))}(s) \mu_\omega(ds). \end{aligned} \tag{3.5.12}$$

Note that $\mu^{\tau-} \ll \mu$ with Radon-Nikodym derivative $\mathbf{1}_{(0,\tau(\omega))}$. Therefore, we obtain for $t \geq \tau(\omega)$

$$\begin{aligned} Y_t(\omega) &= \int f(s,\tau(\omega),r_{s-}^{\tau-}(\omega))) \mathbf{1}_{(0,t]}(s) \mu_\omega^{\tau-}(ds) \\ &= \int f(s,\tau(\omega),r_{s-}(\omega))) \mathbf{1}_{(0,\tau(\omega))}(s) \mu_\omega(ds). \end{aligned} \tag{3.5.13}$$

Since, $t \mapsto X_t(\omega)$ is continuous at $t = \tau(\omega)$, and constant thereafter, the identities (3.5.12) and (3.5.13) imply

$$X_t(\omega) = Y_t(\omega) \quad \text{for } t \geq \tau(\omega). \tag{3.5.14}$$

The assertion of the lemma follows from combining (3.5.11) and (3.5.14). □

We also need the notion of S^p-convergence in the multidimensional case. Recall that $\mathbb{D}$ denotes the space of adapted càdlàg processes. For a process $X \in \mathbb{D}^d$ and $1 \leq p \leq \infty$ we define

$$\|X\|_{S^p} := \left\| \sup_{t \in \mathbb{R}_+} \|X_t\|_1 \right\|_{L^p}.$$

One verifies that $\|\bullet\|_{S^p}$ is a norm on the space $\mathbb{D}^d$. In the one-dimensional case, it is the S^p-norm defined in Protter [55, p. 188].

We say that a sequence $(X_m)_{m\in\mathbb{N}}$ of processes in $\mathbb{D}^d$ converges to $X \in \mathbb{D}^d$ in the S^p-sense, denoted by $X_m \stackrel{S^p}{\to} X$, if $\|X_m - X\|_{S^p} \to 0$ as $m \to \infty$.

Remember that we have fixed d benchmark points $0 \leq x_1 < \ldots < x_d$, and that r denotes $r_t = (f(t,t+x_1),\ldots,f(t,t+x_d))$. This also concerns the term structure models (3.5.15) from the next result.

3.5.16 Proposition. *For $m \in \mathbb{N}_0$ let $f_m(t,T)$ be a term structure model of the form*

$$\begin{cases} df_m(t,T) &= \sum_{i=1}^n \sigma_{(m,i)}(t,T,r_{t-}^m) dA_t^i \\ f_m(0,T) &= f^*(0,T) \end{cases}, \tag{3.5.15}$$

where $A \in \mathcal{V}^n$ and $f^*(0,\bullet) \in C^1(\mathbb{R}_+)$, and where $\sigma_{(m,i)} \in C(\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d)$ for $m \in \mathbb{N}_0$ and $i = 1,\dots,n$. Assume there are $T^*, L > 0$ such that it holds for all $i = 1,\dots,n$

$$\lim_{m\to\infty} \sup_{(t,T,r)\in[0,T^*+x_d]\times[0,T^*+x_d]\times\mathbb{R}^d} \left|\sigma_{(0,i)}(t,T,r) - \sigma_{(m,i)}(t,T,r)\right| = 0, \tag{3.5.16}$$

$$\left|\sigma_{(m,i)}(t,T,r_1) - \sigma_{(m,i)}(t,T,r_2)\right| \le L\|r_1 - r_2\|_1, \quad m \in \mathbb{N} \tag{3.5.17}$$

for all $t,T \in [0,T^*+x_d]$ and $r_1, r_2 \in \mathbb{R}^d$. Then, there is a stopping time τ, which is $\mathbb{P}-a.s.$ positive, such that for all $x \in \mathbb{R}_+$ and $1 \le p \le \infty$

$$r_m(x)^{\tau-} \xrightarrow{S^p} r_0(x)^{\tau-} \quad \text{as } m \to \infty.$$

Proof. Let $1 \le p \le \infty$ be arbitrary. For each $i = 1,\dots,n$ set $\tau_i := \inf\{t > 0 \,|\, \mathrm{Var}(A_i)_t > \frac{1}{2dnL}\}$, where "Var" denotes the variation processes of the A_i. The τ_i are stopping times by Protter [55, Thm. I.3]. Since the A_i have right-continuous paths, the stopping time $\tau := \tau_1 \wedge \ldots \wedge \tau_n \wedge T^*$ is $\mathbb{P}-a.s.$ positive. Note that for all $x \in \mathbb{R}_+$ and $m \in \mathbb{N}_0$ the forward rates $r_t(x)_m = f_m(t,t+x)$ are given by

$$r_t(x)_m = f^*(0,t+x) + \sum_{i=1}^n \int_0^t \sigma_{(m,i)}(s,t+x,r_m(s-))dA_s^i, \quad t \in \mathbb{R}_+.$$

Using Lemma 3.5.15, it holds for each $m \in \mathbb{N}$

$$\begin{aligned}\left\|r_0^{\tau-} - r_m^{\tau-}\right\|_{S^p} \le \sum_{i=1}^n \sum_{j=1}^d \Big\| \int_0^t &\Big(\sigma_{(0,i)}(s,(t\wedge\tau)+x_j,r_0^{\tau-}(s-)) \\ &- \sigma_{(m,i)}(s,(t\wedge\tau)+x_j,r_0^{\tau-}(s-))\Big)dA_i^{\tau-}(s)\Big\|_{S^p} \\ + \sum_{i=1}^n \sum_{j=1}^d \Big\| \int_0^t &\Big(\sigma_{(m,i)}(s,(t\wedge\tau)+x_j,r_0^{\tau-}(s-)) \\ &- \sigma_{(m,i)}(s,(t\wedge\tau)+x_j,r_m^{\tau-}(s-))\Big)dA_i^{\tau-}(s)\Big\|_{S^p}.\end{aligned} \tag{3.5.18}$$

Since it holds $\|H_-\bullet B\|_{S^p} \le \|\mathrm{Var}(B)_\infty H\|_{S^p}$ for any $H \in \mathbb{D}$ and $B \in \mathcal{V}$, for all $i = 1,\dots,n$, $j = 1,\dots,d$ the convergence

$$\begin{aligned}&\Big\| \int_0^t \Big(\sigma_{(0,i)}(s,(t\wedge\tau)+x_j,r_0^{\tau-}(s-)) - \sigma_{(m,i)}(s,(t\wedge\tau)+x_j,r_0^{\tau-}(s-))\Big)dA_i^{\tau-}(s)\Big\|_{S^p} \\ &\le \frac{1}{2dnL} \sup_{(s,t,r)\in[0,T^*]\times[0,T^*]\times\mathbb{R}^d} \left|\sigma_{(0,i)}(s,t+x_j,r) - \sigma_{(m,i)}(s,t+x_j,r)\right| \to 0 \quad \text{as } m \to \infty\end{aligned} \tag{3.5.19}$$

is valid by (3.5.16), where we take into account the definition of τ. Moreover, it holds for each $i = 1, \dots, n$, $j = 1, \dots, d$

$$\Big\| \int_0^t \Big(\sigma_{(m,i)}(s, (t \wedge \tau) + x_j, r_0^{\tau-}(s-)) - \sigma_{(m,i)}(s, (t \wedge \tau) + x_j, r_m^{\tau-}(s-)) \Big) dA_i^{\tau-}(s) \Big\|_{S^p} \leq \frac{1}{2dn} \big\| r_0^{\tau-} - r_m^{\tau-} \big\|_{S^p} \tag{3.5.20}$$

by the Lipschitz condition (3.5.17). Incorporating (3.5.19) and (3.5.20) in the inequality (3.5.18), we deduce

$$r_m^{\tau-} \xrightarrow{S^p} r_0^{\tau-} \quad \text{as } m \to \infty. \tag{3.5.21}$$

Now let $x \in \mathbb{R}_+$ be arbitrary. For each $m \in \mathbb{N}$ we obtain

$$\begin{aligned} &\big\| r_0(x)^{\tau-} - r_m(x)^{\tau-} \big\|_{S^p} \\ &\leq \sum_{i=1}^n \Big\| \int_0^t \Big(\sigma_{(0,i)}(s, t + x, r_0^{\tau-}(s-)) - \sigma_{(m,i)}(s, t + x, r_0^{\tau-}(s-)) \Big) dA_i^{\tau-}(s) \Big\|_{S^p} \\ &+ \sum_{i=1}^n \Big\| \int_0^t \Big(\sigma_{(m,i)}(s, t + x, r_0^{\tau-}(s-)) - \sigma_{(m,i)}(s, t + x, r_m^{\tau-}(s-)) \Big) dA_i^{\tau-}(s) \Big\|_{S^p} \end{aligned}$$

is valid. Arguing as before, we get that for all $i = 1, \dots, n$

$$\Big\| \int_0^t \Big(\sigma_{(0,i)}(s, t + x, r_0^{\tau-}(s-) - \sigma_{(m,i)}(s, t + x, r_0^{\tau-}(s-) \Big) dA_i^{\tau-}(s) \Big\|_{S^p}$$

tends to zero for $m \to \infty$, as well as

$$\begin{aligned} &\Big\| \int_0^t \Big(\sigma_{(m,i)}(s, t + x, r_0^{\tau-}(s-)) - \sigma_{(m,i)}(s, t + x, r_m^{\tau-}(s-)) \Big) dA_i^{\tau-}(s) \Big\|_{S^p} \\ &\leq \frac{1}{2dn} \big\| r_0^{\tau-} - r_m^{\tau-} \big\|_{S^p} \to 0 \quad \text{as } m \to \infty, \end{aligned}$$

because of the convergence (3.5.21), which completes the proof. □

3.5.17 Definition. *A set $\{f_1, \dots, f_n\}$ of functions from $\mathbb{R}_+ \times \mathbb{R}_+$ into $\mathbb{R}$ is called* <u>*$\partial/\partial t$-regular*</u> *if for each $i = 1, \dots, n$ and $m_1, \dots, m_n \in \mathbb{N}_0$ the function $f_i g_1^{m_1} \cdots g_n^{m_n}$ is $\partial/\partial t$-regular, where $g_j(t,T) := \int_t^T f_j(t,s)ds$, $j = 1, \dots, n$.*

A class of $\partial/\partial t$-regular sets are those $\{f_1, \dots, f_n\}$, where $f_i(t,T) = g_i(T-t)$ for $i = 1, \dots, n$, with quasi-exponential functions $g_1, \dots, g_n$, because products and primitives of quasi-exponential functions are again quasi-exponential Björk and Svensson [13, Lemma 5.1].

3.5.18 Theorem. *Assume the following conditions are fulfilled.*

- *$X \in \mathcal{V}^n$ has a deterministic derivative;*
- *The cumulant generating function Ψ is of class $C^{0,2}(\mathbb{R}_+ \times \operatorname{int} Q)$;*
- *The set of functions $\{\lambda_{(i,j)} \,|\, i = 1, \ldots, n, j = 1, \ldots, m_i\}$ is $\partial/\partial t$-regular;*
- *$\varphi_{(i,j)} \in C_b^{\mathcal{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$ for all $i = 1, \ldots, n,\ j = 1, \ldots, m_i$.*

Then, the interest rate model (3.5.6) has a local approximative realization.

Proof. Let $T^* > 0$ be arbitrary. Since $\Psi \in C^{0,2}(\mathbb{R}_+ \times \operatorname{int} Q)$ there exists, see, e.g., Heuser [37, Thm. 115.6], a sequence $(\Psi_m)_{m\in\mathbb{N}}$ of polynomials on $\mathbb{R}^{n+1}$ such that

$$\lim_{m\to\infty} \sup_{(t,z)\in[0,T^*]\times Q_0} \left| \frac{\partial}{\partial z_i}\Psi(t,z) - \frac{\partial}{\partial z_i}\Psi_m(t,z) \right| = 0, \quad i = 1, \ldots, n, \tag{3.5.22}$$

$$\lim_{m\to\infty} \sup_{(t,z)\in[0,T^*]\times Q_0} \left| \frac{\partial^2}{\partial z_i \partial z_j}\Psi(t,z) - \frac{\partial^2}{\partial z_i \partial z_j}\Psi_m(t,z) \right| = 0, \quad i,j = 1, \ldots, n. \tag{3.5.23}$$

Define the mappings $\alpha_m : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ for each $m \in \mathbb{N}$ by

$$\alpha_m(t,T,r) := -\Big\langle \sigma(t,T,r), \nabla_z \Psi_m\left(t, -\int_t^T \sigma(t,s,r)ds\right) \Big\rangle.$$

Using Lemma 3.5.9 on the $\varphi_{(i,j)}$, and the continuity of the $\lambda_{(i,j)}$, it follows that the σ_i, $i = 1, \ldots, n$ are bounded on $[0,T^*] \times [0,T^*] \times \mathbb{R}^d$. Therefore, it holds, by the uniform convergence (3.5.22),

$$\lim_{m\to\infty} \sup_{(t,T,r)\in[0,T^*]\times[0,T^*]\times\mathbb{R}^d} |\alpha_m(t,T,r) - \alpha(t,T,r)| = 0. \tag{3.5.24}$$

Choose $\varepsilon > 0$. By the uniform convergence (3.5.22), (3.5.23), there exists an index $m_0 \in \mathbb{N}$ such that for all $m \geq m_0$

$$\sup_{(t,z)\in[0,T^*]\times Q_0} \left| \frac{\partial}{\partial z_i}\Psi_m(t,z) - \frac{\partial}{\partial z_i}\Psi(t,z) \right| \leq \varepsilon, \quad i = 1, \ldots, n,$$
$$\sup_{(t,z)\in[0,T^*]\times Q_0} \left| \frac{\partial^2}{\partial z_i \partial z_j}\Psi_m(t,z) - \frac{\partial^2}{\partial z_i \partial z_j}\Psi(t,z) \right| \leq \varepsilon, \quad i,j = 1, \ldots, n.$$

It follows that for all $m \geq m_0$ and all $(t,z) \in [0,T^*] \times Q_0$ it holds

$$\left| \frac{\partial}{\partial z_i}\Psi_m(t,z) \right| \leq \left| \frac{\partial}{\partial z_i}\Psi(t,z) \right| + \varepsilon, \quad i = 1, \ldots, n,$$
$$\left| \frac{\partial^2}{\partial z_i \partial z_j}\Psi_m(t,z) \right| \leq \left| \frac{\partial^2}{\partial z_i \partial z_j}\Psi(t,z) \right| + \varepsilon, \quad i,j = 1, \ldots, n,$$

Since $[0,T^*] \times Q_0$ is a compact set, we conclude that there exists a constant $M > 0$ such that for each $m \geq m_0$

$$\sup_{(t,z)\in[0,T^*]\times Q_0} \left| \frac{\partial}{\partial z_i} \Psi_m(t,z) \right| \leq M, \quad i = 1,\dots,n, \tag{3.5.25}$$

$$\sup_{(t,z)\in[0,T^*]\times Q_0} \left| \frac{\partial^2}{\partial z_i \partial z_j} \Psi_m(t,z) \right| \leq M, \quad i,j = 1,\dots,n. \tag{3.5.26}$$

The partial derivatives $\frac{\partial}{\partial r_k}\sigma_i(t,T,r)$ for $i = 1,\dots,n$ and $k = 1,\dots,d$ are given by

$$\frac{\partial}{\partial r_k}\sigma_i(t,T,r) = \sum_{j=1}^{m_i} \frac{\partial}{\partial r_k}\varphi_{(i,j)}(t,r)\lambda_{(i,j)}(t,T).$$

For each $m \geq m_0$, the first order derivatives $\frac{\partial}{\partial r_k}\alpha_m(t,T,r)$, $k = 1,\dots,d$ are equal to

$$\begin{aligned}
\frac{\partial}{\partial r_k}\alpha_m(t,T,r) &= -\sum_{i=1}^{n} \left(\sum_{j=1}^{m_i} \frac{\partial}{\partial r_k}\varphi_{(i,j)}(t,r)\lambda_{(i,j)}(t,T) \right) \frac{\partial}{\partial z_i}\Psi_m\left(t, -\int_t^T \sigma(t,s,r)ds\right) \\
&- \sum_{i=1}^{n} \left[\sum_{j=1}^{m_i} \varphi_{(i,j)}(t,r)\lambda_{(i,j)}(t,T) \right] \\
&\quad \left[\sum_{j=1}^{n} \frac{\partial^2}{\partial z_i \partial z_j}\Psi_m\left(t, -\int_t^T \sigma(t,s,r)ds\right) \left(-\sum_{l=1}^{m_j} \frac{\partial}{\partial r_k}\varphi_{(j,l)}(t,r) D_{(j,l)}(t,T) \right) \right].
\end{aligned}$$

The space $C_b^{\mathcal{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$ is closed under products. Due to (3.5.25),(3.5.26), the continuity of the $\lambda_{(i,j)}$ and $D_{(j,l)}$, and Lemma 3.5.9, we obtain that there exists a constant $L > 0$ satisfying

$$\sup_{(t,T,r)\in[0,T^*]\times[0,T^*]\times\mathbb{R}^d} \left| \frac{\partial}{\partial r_k}\alpha_m(t,T,r) \right| \leq L, \quad k = 1,\dots,d,\ m \geq m_0,$$

$$\sup_{(t,T,r)\in[0,T^*]\times[0,T^*]\times\mathbb{R}^d} \left| \frac{\partial}{\partial r_k}\sigma_i(t,T,r) \right| \leq L, \quad i = 1,\dots,n,\ k = 1,\dots,d.$$

Applying Lemma 3.5.10 we obtain that for each $i = 1,\dots,n$ and $m \geq m_0$

$$\|\alpha_m(t,T,r_1) - \alpha_m(t,T,r_2)\|_1 \leq L\|r_1 - r_2\|_1, \quad t,T \in [0,T^*] \text{ and } r_1, r_2 \in \mathbb{R}^d,$$

$$\|\sigma_i(t,T,r_1) - \sigma_i(t,T,r_2)\|_1 \leq L\|r_1 - r_2\|_1, \quad t,T \in [0,T^*] \text{ and } r_1, r_2 \in \mathbb{R}^d. \tag{3.5.27}$$

Each Ψ_m is of the form $\Psi_m(t,z) = \sum_{u=1}^{e} t^{l_u} z_1^{k_u^1} \cdots z_n^{k_u^n}$, so in particular of the type (3.5.5), and it is continuous on $\mathbb{R}_+ \times \mathbb{R}^n$. Thus, Proposition 3.5.11 provides us with a finite dimensional realization (F_m, Z_m) for each forward rate model $f_m(t,T)$, $m \geq m_0$ specified as

$$\begin{cases} df_m(t,T) &= \alpha_m(t,T,r_{t-})dt + \sigma(t,T,r_{t-})dX_t \\ f_m(0,T) &= f^*(0,T) \end{cases}.$$

For the mappings $G_m(t,x,z) := F_m(t,t+x,z)$ the identity $r_m(x)_t = G_m(t,x,Z_m(t))$ is valid. Proposition 3.5.16, which may be applied by virtue of (3.5.24), (3.5.27) and the assumption $X \in \mathcal{V}^n$, gives us the existence of a $\mathbb{P}-a.s.$ positive stopping time τ such that $G_m(\bullet, x, Z_m(t))^{\tau-} \stackrel{S^p}{\to} r(x)^{\tau-}$ as $m \to \infty$ for each $x \in \mathbb{R}_+$ and $1 \leq p \leq \infty$, which proves that the interest rate model (3.5.6) has a local approximative realization. □

The range of driving processes X, for which Theorem 3.5.18 can be applied, encompasses those Lévy processes with zero Gaussian part and Lévy measure K satisfying $\int_{|x|\leq 1} |x| K(dx) < \infty$, because then X belongs to $\mathcal{V}$ Sato [59, Thm. 21.9] and Ψ does not depend on t. Examples are compound Poisson processes, bilateral Gamma processes, see Küchler and Naumann [46], and in particular variance Gamma processes, which have been used in a series of papers, see Madan and Seneta [51], Madan and Milne [50], Carr, Chang, and Madan [18], and Madan [49] for a survey. More generally, we can apply Theorem 3.5.18 if X is a generalized tempered stable process Cont and Tankov [23, Sec. 4.5], i.e. the Lévy measure of the generating triplet $(0,0,K)$ is of the form

$$K(dx) = \left(\frac{c_-}{|x|^{1+\alpha_-}} e^{-\lambda_- |x|} \mathbf{1}_{(-\infty,0)}(x) + \frac{c_+}{x^{1+\alpha_+}} e^{-\lambda_+ x} \mathbf{1}_{(0,\infty)}(x) \right) dx,$$
$$(c_- > 0,\ c_+ > 0,\ \lambda_- > 0,\ \lambda_+ > 0 \text{ and } \alpha_- < 2,\ \alpha_+ < 2),$$

and the parameters satisfy $\alpha_- < 1$, $\alpha_+ < 1$ and at least one of them is non-negative. This includes some of the CGMY processes in Carr, Geman, Madan, and Yor [19].

The reason, why we have confined ourselves to driving processes $X \in \mathcal{V}$ is that we have to establish (see Proposition 3.5.16) a stability result for a sequence of processes satisfying

$$r_t(x) = f(0,t+x) + \int_0^t \alpha(s,t+x,r_{s-})ds + \int_0^t \sigma(s,t+x,r_{s-})dX_s,$$

which are equations of Volterra type. However, Proposition 3.5.11, which provides realizations for term structure models of the kind (3.5.7), is not subject to this restriction, and of course, it is desirable to extend Theorem 3.5.18 to a more general class of driving processes X.

As an illustration of Theorem 3.5.18, we assume that the term structure model (3.5.6) is driven by a single Lévy process with zero Gaussian part and Lévy measure K satisfying $\int_{|x|\leq 1} |x| K(dx) < \infty$, where the volatility is of the form $\sigma(t,T,r) =$

$\varphi(t,r)\lambda(t,T)$. Note that we started, at the beginning of this section, with a volatility structure of this kind. Presuming that all λD^j, $j \in \mathbb{N}_0$ are $\partial/\partial t$-regular, where $D(t,T) := \int_t^T \lambda(t,s)ds$ (this is in particular fulfilled if λ is stationary, i.e. $\lambda(t,T) = \tilde{\lambda}(T-t)$, and $\tilde{\lambda}$ is quasi-exponential), and $\varphi \in C_b^{\mathcal{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$, Theorem 3.5.18 yields the existence of a local approximative realization.

To make things more concrete, assume that for some constant $M > 0$ it holds $Q_0 \subset (-M, M) \subset Q$, where Q and Q_0 denote the compact sets from Assumptions 3.5.6. Then, we can explicitly construct the desired sequence $(G_m, Z_m)_{m\in\mathbb{N}}$ by inspecting the proofs of Theorem 3.5.18 and Proposition 3.5.11. First, we note that the Taylor series representation

$$\Psi(z) = \sum_{j=0}^{\infty} \frac{1}{j!}\Psi^{(j)}(0)z^j, \quad z \in Q_0$$

is valid. Define the polynomials $\Psi_m(z) := \sum_{j=0}^{m+1} \frac{1}{j!}\Psi^{(j)}(0)z^j$ for $m \in \mathbb{N}$. Then it holds $\Psi'_m \to \Psi'$ and $\Psi''_m \to \Psi''$ uniformly on Q_0, and therefore $(\Psi_m)_{m\in\mathbb{N}}$ provides a sequence of polynomials as in the proof of Theorem 3.5.18. We verify that

$$-\sigma(t,T,r)\Psi'_m\left(-\int_t^T \sigma(t,s,r)ds\right) = \sum_{j=0}^{m} \frac{1}{j!}\Psi^{(j+1)}(0)(-\varphi(t,r))^{j+1}\lambda(t,T)D(t,T)^j.$$

For each $j \in \mathbb{N}_0$ there exists an integer $q_j \in \mathbb{N}_0$ and functions $\eta_{(j,k)} \in C(\mathbb{R}_+)$, $k = 0,\dots,q_j$ such that

$$\frac{\partial^{q_j+1}}{\partial t^{q_j+1}}(\lambda D^j)(t,T) + \sum_{k=0}^{q_j} \eta_{(j,k)}(t)\frac{\partial^k}{\partial t^k}(\lambda D^j)(t,T) = 0.$$

Looking into the proof of Proposition 3.5.11, we choose the mappings $(F_m)_{m\in\mathbb{N}}$ as

$$\begin{aligned} &F_m(t,T,z_k \,|\, k=0,\dots,q_0, z_{(j,k)} \,|\, j=0,\dots,m, k=0,\dots,q_j) \\ &:= f^*(0,T) + \sum_{k=0}^{q_0} \frac{\partial^k}{\partial t^k}\lambda(t,T)z_k + \sum_{j=0}^{m}\sum_{k=0}^{q_j} \frac{\partial^k}{\partial t^k}(\lambda D^j)(t,T)z_{(j,k)}, \end{aligned}$$

and let the state processes

$$Z_m = (Z_k^m \,|\, k=0,\dots,q_0, Z_{(j,k)}^m \,|\, j=0,\dots,m, k=0,\dots,q_j), \quad m \in \mathbb{N}$$

be the unique solutions of the following stochastic differential equation (3.5.28), in which j runs from 0 to m, and $\tilde{F}$ denotes $\tilde{F}(t,z) := (F(t,t+x_1,z),\dots,F(t,t+x_d,z))$

for $(t, z) \in \mathbb{R}_+ \times \mathbb{R}^d$.

$$\begin{cases} dZ_0^m(t) &= \eta_{(0,0)}(t)Z_{q_0}^m(t) + \varphi(t, \tilde{F}(t, Z_m(t-)))dX_t \\ dZ_k^m(t) &= \Big[\eta_{(0,k)}(t)Z_{q_0}^m(t) - Z_{k-1}^m(t)\Big]dt, \quad k = 1, \dots, q_0 \\ dZ_{(j,0)}^m(t) &= \Big[\eta_{(j,0)}(t)Z_{(j,q_j)}^m(t) + \frac{1}{j!}\Psi^{(j+1)}(0)(-\varphi(t, \tilde{F}(t, Z_m(t-))))^{j+1}\Big]\, dt \\ dZ_{(j,k)}^m(t) &= \Big[\eta_{(j,k)}(t)Z_{(j,q_j)}^m(t) - Z_{(j,k-1)}^m(t)\Big]\, dt, \quad k = 1, \dots, q_j \\ Z_m(0) &= 0 \end{cases} . \tag{3.5.28}$$

Setting $G_m(t, x, z) := F_m(t, t + x, z)$ for $m \in \mathbb{N}$, we obtain from Proposition 3.5.16 that $G_m(\bullet, x, Z_m(t))^{\tau-} \overset{S^p}{\to} r(x)^{\tau-}$ for some $\mathbb{P} - a.s.$ positive stopping time τ.

We observe that the concrete choice of the driving Lévy process X only enters via $\frac{1}{j!}\Psi^{(j+1)}(0)$ in the specification of the $Z_{(j,0)}^m$.

For a Poisson process X with intensity $\lambda > 0$ for instance, it holds $\Psi'(z) = \lambda e^z$, and thus we have to insert $\Psi^{(j+1)}(0) = \lambda$, $j \in \mathbb{N}_0$ in the stochastic differential equation (3.5.28). A more realistic term structure model is provided if the driving process X is a bilateral Gamma process Küchler and Naumann [46, Sec. 5] with parameters $\alpha^+, \alpha^-, \lambda^+, \lambda^- > 0$. Then, it necessarily holds $Q \subset (-\lambda^-, \lambda^+)$, where Q denotes the compact set from Assumptions 3.5.6, and Ψ' is given by $\Psi'(z) = \frac{\alpha^+}{\lambda^+ - z} - \frac{\alpha^-}{\lambda^- + z}$. Consequently, we have to insert

$$\frac{1}{j!}\Psi^{(j+1)}(0) = \frac{\alpha^+}{(\lambda^+)^{j+1}} + \frac{\alpha^-}{(-\lambda^-)^{j+1}}, \quad j \in \mathbb{N}_0$$

in the stochastic differential equation (3.5.28).

Chapter 4

Benchmark realizations

In the preceding chapter, we have treated the existence of finite dimensional realizations using the Lie algebraic approach. We shall now go an alternative way, in the framework of benchmark realizations. After some preliminaries, we show that this is no strong restriction, since in general, a benchmark realization can be obtained from a finite dimensional realization with an arbitrary state process by an adequate transformation. Using the fact that $\mathbb{P}$ is a martingale measure, we derive a couple of integro-differential equations, used in the sequel for the realization question.

After these preparations, we start with the study of deterministic volatility structures. Then, for deterministic direction volatility models driven by a process which has jumps, we prove, roughly speaking, that the existence of a realization implies that the volatility must be deterministic. Using this fact we show that, in principle, short rate realizations for interest rate models driven by a jump process can only exist if the volatility is deterministic. These results exhibit the restrictive nature of term structure models, which are driven by processes with jumps, concerning finite dimensional realizations.

Finally, we treat generic benchmark realizations and show, generalizing Jeffrey [44], that forward rate models with a generic benchmark realization must necessarily have a singular Hessian matrix.

4.1 Preparatory results

Throughout this chapter, we consider, for fixed integers $d, n \in \mathbb{N}$, HJM term structure models of the form

$$\begin{cases} df(t,T) &= \alpha(t,T,r_{t-})dt + \sigma(t,T,r_{t-})dX_t \\ f(0,T) &= f^*(0,T) \end{cases}, \tag{4.1.1}$$

with coefficients $\alpha, \sigma_1, \ldots, \sigma_n \in C^{0,1,0}(\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d)$, where r denotes a set of benchmark forward rates $r_t = (f(t, t + x_1), \ldots, f(t, t + x_d))$ for fixed benchmark maturities $0 \leq x_1 < \ldots < x_d$. We always make the following assumptions.

4.1.1 Assumptions.

1. *The driving process X is a n-dimensional Grigelionis process with r-derivative $(\beta, c, K; r)^D$, such that $(\mathcal{D}(X), r)$ has regular supports.*
2. *There are $z_1^-, \ldots, z_n^- \in (-\infty, 0)$ and $z_1^+, \ldots, z_n^+ \in (0, \infty)$ such that for any $t \in \mathbb{R}_+$ and $r : [0, t] \to \mathbb{R}^d$*
$$\int_0^t \left(\int_{|x| \leq 1} |x|^2 K_{s,r(s)}(dx) + \int_{|x|>1} e^{\langle z,x \rangle} K_{s,r(s)}(dx) \right) ds < \infty, \quad z \in Q,$$
where Q denotes the set $Q := [z_1^-, z_1^+] \times \ldots \times [z_n^-, z_n^+]$, and furthermore, for all $(t, r) \in \mathbb{R}_+ \times \mathbb{R}^d$ it holds
$$\int_{|x|>1} e^{\langle z,x \rangle} K_{t,r}(dx) < \infty, \quad z \in Q.$$
3. *There are $w_1^- \in (z_1^-, 0), \ldots, w_n^- \in (z_n^-, 0)$ and $w_1^+ \in (0, z_1^+), \ldots, w_n^+ \in (0, z_n^+)$ such that*
$$-\int_t^T \sigma(t, s, r) ds \in Q_0, \quad \textit{for all } (t, T, r) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d,$$
where $Q_0 \subset Q$ is defined as $Q_0 := [w_1^-, w_1^+] \times \ldots \times [w_n^-, w_n^+]$.
4. *The function $\Psi : \mathbb{R}_+ \times \mathbb{R}^d \times Q \to \mathbb{R}$ (see the definition below) is continuous.*

The second assumption ensures that the cumulant generating function
$$\Psi(t, r, z) := \langle \beta(t, r), z \rangle + \frac{1}{2} \langle z, c(t, r) z \rangle + \int_{\mathbb{R}^n} \left(e^{\langle z,x \rangle} - 1 - \langle z, x \rangle \right) K_{t,r}(dx)$$
is definable for all $(t, r, z) \in \mathbb{R}_+ \times \mathbb{R}^d \times Q$ (see Section 2.2). The function Ψ is of class C^∞ in z, but it need not be continuous in t and r. Therefore, we impose the fourth assumption. Instead of $\Psi(t, r, z)$, we will also write $\Psi_{t,r}(z)$. The third assumption guarantees that
$$\Psi \left(t, r, -\int_t^T \sigma(t, s, r) ds \right)$$
exists for all $(t, T, r) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d$, and is continuous. As usual, we assume that the model is free of arbitrage in the sense that $\mathbb{P}$ is a martingale measure.

4.1.2 Definition.

1. *A benchmark realization is given by a mapping* $F \in C^{2,1,2}(\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d)$ *such that* $f(t,T) = F(t,T,r_t)$ *for all* $0 \leq t \leq T$.

2. *In the case* $d = 1$ *and* $x_1 = 0$ *a benchmark realization is also called a short rate realization.*

A benchmark realization is (up to minor deviations, as there are the smoothness conditions on F) a finite dimensional realization in the sense of Definition 2.2.3 with state process r, i.e. the realization consists of the pair (F, r).

We remark that Chiarella and Kwon [22] also consider interest rate models of the type (4.1.1), where drift and volatilities are allowed to depend on a set of benchmark forward rates, driven by finitely many Wiener processes. They construct, under appropriate assumptions, finite dimensional realizations without imposing restrictions on the state process. We, however, demand that the state process is a set of benchmark forward rates.

In Definition 2.2.3 it is assumed that the state process satisfies a stochastic differential equation. This is automatically fulfilled for benchmark realizations, which is the content of the next lemma.

4.1.3 Lemma. *Assume the term structure model (4.1.1) admits a benchmark realization* F. *Then, the benchmark forward rate process* r *satisfies the stochastic differential equation*

$$\begin{cases} dr_t &= \mu(t, r_{t-})dt + \gamma(t, r_{t-})dX_t \\ r_0 &= r^* \end{cases}, \tag{4.1.2}$$

where $r^* \in \mathbb{R}^d$ *is given by* $r^* = (f^*(0,x_1), \ldots, f^*(0,x_d))$, *and where* $\mu : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ *has the components*

$$\mu_i(t,r) = \frac{\partial}{\partial T} F(t,T,r) \Big|_{T=t+x_i} + \alpha(t, t+x_i, r), \quad i = 1, \ldots, d, \tag{4.1.3}$$

and $\gamma : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d \times n}$ *is given by*

$$\gamma(t,r) = \begin{pmatrix} \sigma_1(t, t+x_1, r) & \cdots & \sigma_n(t, t+x_1, r) \\ \vdots & & \vdots \\ \sigma_1(t, t+x_d, r) & \cdots & \sigma_n(t, t+x_d, r) \end{pmatrix}. \tag{4.1.4}$$

Proof. The assertion is a direct consequence of the Musiela parametrization (Proposition 2.2.12). □

Affine benchmark realizations, which we will now define, are of particular interest. The following definition may seem surprising, because we do not impose that such a realization must be affine everywhere. However, the following definition is sufficient in order to establish all subsequent results, in particular the third part of Theorem 3.1.17 can be applied. In Theorem 4.4.4 we will show that an a priori arbitrary short rate realization must automatically be affine under appropriate assumptions. This can only be done if we use the following definition of an affine benchmark realization.

4.1.4 Definition.

1. *A benchmark realization F is called affine if there are $a, b_1, \dots, b_d \in C^{2,1}(\mathbb{R}_+ \times \mathbb{R}_+)$ such that for all $t \in (0, \infty)$, $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$ it holds*

$$\begin{aligned} F(t, T, r + \gamma(t, r)x) &= a(t, T) + \langle b(t, T), r + \gamma(t, r)x \rangle, \quad x \in \operatorname{supp}(K_t) \cup \{0\}, \\ \nabla_r F(t, T, r) &= b(t, T). \end{aligned}$$

2. *An affine benchmark F is said to be stationary if there are $a \in C^{2,1}(\mathbb{R}_+ \times \mathbb{R}_+)$ and $\tilde{b} \in C^2(\mathbb{R}_+, \mathbb{R}^d)$ such that for all $t \in (0, \infty)$, $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$ it holds*

$$\begin{aligned} F(t, T, r + \gamma(t, r)x) &= a(t, T) + \langle \tilde{b}(T - t), r + \gamma(t, r)x \rangle, \quad x \in \operatorname{supp}(K_t) \cup \{0\}, \\ \nabla_r F(t, T, r) &= \tilde{b}(T - t). \end{aligned}$$

We agree that, speaking about an affine benchmark realization $F(t, T, r) = a(t, T) + \langle b(t, T), r \rangle$, always means that it is affine in the weaker sense of Definition 4.1.4.

Now, we discuss the restrictions, that we have to accept, if we treat benchmark realizations. The coefficients in (4.1.1) may only depend on the state process $r_t = (f(t, t + x_1), \dots, f(t, t + x_d))$ of a benchmark realization. This is no further restriction. If drift and volatilities depend on the whole forward rate curve $f(t, t + \bullet)$, as in Chapter 3, and there exists a benchmark realization, they actually only depend on the set of forward rates $(f(t, t + x_1), \dots, f(t, t + x_d))$, because then, the forward rate curve $f(t, t + \bullet)$ can be expressed by means of $f(t, t + x_1), \dots, f(t, t + x_d)$ and the time t.

Nevertheless, the condition that the state process consists of a set of benchmark forward rates, seems rather restrictive. However, given a finite dimensional realization with an arbitrary state process, this realization can usually be transformed into a benchmark realization. See Björk and Svensson [13, Thm. 3.3] for such a result when the driving processes are Brownian motions.

We illustrate how to perform such a transformation, which also works if the driving processes have jumps, in the case of affine realizations, which are of major interest in this chapter. Our ideas follow Björk and Landén [12, Sec. 7], see also Björk and Gombani [11, Prop. 5.1] for a closely related result.

4.1.5 Proposition. *Assume, an arbitrary term structure model $f(t,T)$ has an affine $d{+}1$-dimensional realization (G,Z) in the sense of Definition 2.2.3, where $G(t,T,z) = a(t,T) + \langle b(t,T), z\rangle$. Assume there are $0 \leq x_1 < \ldots < x_d$ such that* $\det B(t) \neq 0$ *for all $t \in \mathbb{R}_+$, where*

$$B(t) := \begin{pmatrix} b_1(t,t+x_1) & \cdots & b_d(t,t+x_1) \\ \vdots & & \vdots \\ b_1(t,t+x_d) & \cdots & b_d(t,t+x_d) \end{pmatrix}.$$

Then, defining $\hat{a}(t) := (a(t,t{+}x_1),\ldots,a(t,t{+}x_d))$, $t \in \mathbb{R}_+$, an affine realization (F,r), with $r_t = (f(t,t{+}x_1),\ldots,f(t,t{+}x_d))$ as state process, is given by $F : \mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}^d \to \mathbb{R}$ defined as

$$F(t,T,r) := a(t,T) - \langle B^{-1}(t)^* b(t,T), \hat{a}(t)\rangle + \langle B^{-1}(t)^* b(t,T), r\rangle.$$

Proof. Since (G,Z) provides a realization, it holds for the forward rates $f(t,T)$ of the term structure model

$$f(t,T) = a(t,T) + \langle b(t,T), Z_t\rangle, \quad 0 \leq t \leq T, \tag{4.1.5}$$

and, in particular, for the benchmark forward rates $r_t = (f(t,t+x_1),\ldots,f(t,t+x_d))$ the identity

$$r_t = \hat{a}(t) + B(t)Z_t, \quad t \in \mathbb{R}_+$$

is valid. Since the $B(t)$, $t \in \mathbb{R}_+$ are non-singular by assumption, we deduce

$$Z_t = B^{-1}(t)(r_t - \hat{a}(t)), \quad t \in \mathbb{R}_+. \tag{4.1.6}$$

Inserting (4.1.6) into (4.1.5), we see that (F,r) gives the desired realization. □

The condition $\det B(t) \neq 0$, $t \in \mathbb{R}_+$ essentially means that (G,Z) is a realization of minimal dimension. If the realization is affine and stationary, i.e. G is of the form $G(t,T,z) = a(t,T) + \langle b(T-t), z\rangle$, which typically arises from stationary volatility structures, one only needs that

$$\det \begin{pmatrix} b_1(x_1) & \cdots & b_d(x_1) \\ \vdots & & \vdots \\ b_1(x_d) & \cdots & b_d(x_d) \end{pmatrix} \neq 0$$

for some $0 \leq x_1 < \ldots < x_d$. An immediate consequence is the following result about benchmark realizations.

4.1.6 Corollary. *Assume, an arbitrary term structure model $f(t,T)$ has an affine $d{+}1$-dimensional realization (G,Z) in the sense of Definition 2.2.3, where $G(t,T,z) = a(t,T)+\langle b(T{-}t), z\rangle$. Assume that $b_1,\ldots,b_d$ are linearly independent and real analytic. Then, for any given $T^* > 0$, there exists, apart from finitely many exceptions, for all $0 \leq x_1 < \ldots < x_d \leq T^*$ an affine realization (F,r) with $r_t = (f(t,t+x_1),\ldots,f(t,t+x_d))$ as state process.*

Proof. Let $T^* > 0$ be arbitrary. According to Björk, Kabanov, and Runggaldier [15, Prop. 5.5] it holds, apart from finitely many exceptions, for all $0 \le x_1 < \ldots < x_d \le T^*$ the relation

$$\det \begin{pmatrix} b_1(x_1) & \cdots & b_d(x_1) \\ \vdots & & \vdots \\ b_1(x_d) & \cdots & b_d(x_d) \end{pmatrix} \neq 0$$

Thus, an application of Proposition 4.1.5 finishes the proof. □

To summarize, we have now justified the upcoming investigation of finite dimensional realizations by going into the framework of benchmark realizations, since we have seen that this imposes no hard restrictions.

The advantage of this approach is that we obtain deterministic equations, in particular integro-differential equations in Section 4.2, depending on the variables t, T, which represent time, and on $r \in \mathbb{R}^d$, representing the state process. From these equations we obtain conditions concerning the existence of realizations.

The rest of this section is devoted to basic results, which are needed in the sequel.

4.1.7 Lemma.

1. *Assume the term structure model (4.1.1) admits a benchmark realization F. Then, it holds for the initial forward rate curve $f^*(0, \bullet)$ and the initial condition r^* of the benchmark forward rates*

$$f^*(0, \bullet) = F(0, \bullet, r^*). \tag{4.1.7}$$

   Moreover, the following boundary conditions are valid for $i = 1, \ldots, d$.

$$F(t, t + x_i, r) = r_i, \quad t \in (0, \infty), r \in \operatorname{supp}(r_{t-}). \tag{4.1.8}$$

2. *Assume the term structure model (4.1.1) admits an affine benchmark realization $F(t, T, r) = a(t, T) + \langle b(t, T), r\rangle$. If $t \in (0, \infty)$ is such that* $\operatorname{int} \operatorname{supp}(r_{t-}) \neq \emptyset$, *then the following boundary conditions hold for $i = 1, \ldots, d$*

$$\begin{cases} a(t, t + x_i) & = & 0 \\ b(t, t + x_i) & = & e_i \end{cases}, \tag{4.1.9}$$

   where the e_i denote the unit vectors in $\mathbb{R}^d$.

Proof.

1. It holds for all $T \in \mathbb{R}_+$ the relation $f(0, T) = F(0, T, r_0)$ ($\mathbb{P} - a.s.$), because F is a realization. Since $\mathcal{F}_0 = \{\emptyset, \Omega\}$, the identity (4.1.7) follows. By the continuity of F, we get for each $i = 1, \ldots, d$

$$r^i_{t-} = \lim_{s \uparrow t} F(s, s + x_i, r_s) = F(t, t + x_i, \lim_{s \uparrow t} r_s) = F(t, t + x_i, r_{t-}) \quad \mathbb{P} - a.s.$$

   Applying Lemma 3.1.4, the boundary condition (4.1.8) is proven.

2. We infer from equation (4.1.8) that for each $t \in (0, \infty)$ and $i = 1, \dots, d$ the identity
$$a(t, t + x_i) + \langle b(t, t + x_i), r\rangle = r_i, \quad r \in \operatorname{supp}(r_{t-})$$
is valid. Provided $\operatorname{int}\operatorname{supp}(r_{t-}) \neq \emptyset$, differentiating each of these equations with respect to $r_1, \dots, r_d$ yields the desired relation (4.1.9).

□

4.1.8 Lemma. *Assume the term structure model (4.1.1) admits a benchmark realization F. If the Grigelionis process X is non-degenerate with derivative $(\beta, c, K; r)^D$, then it holds for each $t \in (0, \infty), r \in \operatorname{supp}(r_{t-})$ and $T \geq t$*
$$\alpha(t, T, r) = \frac{\partial}{\partial t} F(t, T, r) + \langle \nabla_r F(t, T, r), \mu(t, r)\rangle + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial r_i \partial r_j} F(t, T, r) \langle \gamma_{i\bullet}(t, r)^*, c(t, r)\gamma_{j\bullet}(t, r)^*\rangle, \tag{4.1.10}$$
$$\sigma(t, T, r) = \nabla_r F(t, T, r)\gamma(t, r). \tag{4.1.11}$$
If the Grigelionis process X is linearly non-degenerate with derivative $(\beta, c, K; r)^D$, and, in addition, the realization is affine, that is $F(t, T, r) = a(t, T) + \langle b(t, T), r\rangle$, then it holds for all $t \in (0, \infty), r \in \operatorname{supp}(r_{t-})$ and $T \geq t$
$$\alpha(t, T, r) = \frac{\partial}{\partial t} a(t, T) + \left\langle \frac{\partial}{\partial t} b(t, T), r \right\rangle + \langle b(t, T), \mu(t, r)\rangle, \tag{4.1.12}$$
$$\sigma(t, T, r) = b(t, T)\gamma(t, r). \tag{4.1.13}$$

Proof. We may apply Theorem 3.1.17 for fixed $T \in (0, \infty)$, because, by hypothesis, the function $F(\bullet, T, \bullet)$ is of class C^2 in each variable. □

4.1.9 Corollary. *Assume the term structure model (4.1.1) admits a benchmark realization F. If X is non-degenerate, then it holds for all $t \in (0, \infty)$ and $r \in \operatorname{supp}(r_{t-})$*
$$\mu_i(t, r) = \alpha(t, t + x_i, r) - \frac{\partial}{\partial t} F(t, T, r)\bigg|_{T=t+x_i}, \quad i = 1, \dots, d.$$

Proof. Set $T = t + x_i$ in (4.1.10) and note that $F(t, t + x_i, r) = r_i$ by Lemma 4.1.7. □

Next, we derive the HJM drift condition.

4.1.10 Lemma. *Assume the term structure model (4.1.1) admits a benchmark realization F. Then, the following identities are valid for all $t \in (0, \infty), r \in \operatorname{supp}(r_{t-})$ and $T \geq t$.*
$$\int_t^T \alpha(t, s, r) ds = \Psi_{t,r}\left(-\int_t^T \sigma(t, s, r) ds\right), \tag{4.1.14}$$
$$\alpha(t, T, r) = -\left\langle \sigma(t, T, r), \nabla\Psi_{t,r}\left(-\int_t^T \sigma(t, s, r) ds\right)\right\rangle. \tag{4.1.15}$$

Proof. By the continuity of α and σ, and since r has càdlàg paths, $\alpha(t,T,r_{t-})$, $\sigma_1(t,T,r_{t-}),\dots,\sigma_n(t,T,r_{t-})$ belong to $\mathbb{L}$ for all $T\in\mathbb{R}_+$. Furthermore, by the continuity of Ψ (see Assumptions 4.1.1), the function $(t,z)\mapsto\Psi(t,r_{t-}(\omega),z)$ is continuous in z and left-continuous in t for all $\omega\in\Omega$, with possible exception of a $\mathbb{P}$-null set. We obtain from Proposition 2.2.11 that for fixed $T\in(0,\infty)$ it holds, up to a $\mathbb{P}$-null set,

$$\int_t^T \alpha(t,s,r_{t-})ds = \Psi\left(t, r_{t-}, -\int_t^T \sigma(t,s,r_{t-})ds\right), \quad t\in(0,T].$$

Since Ψ is continuous, we may apply Lemma 3.1.4 for fixed $t\in(0,T]$, which yields equation (4.1.14). Differentiating (4.1.14), for fixed $t\in(0,\infty)$ and $r\in\operatorname{supp}(r_{t-})$, with respect to T, we obtain the equation (4.1.15). □

4.2 Integro-differential equations

This section presents a couple of integro-differential equations that are valid for benchmark realizations.

Once again, we emphasize that, for convenience of notation, we do, concerning vectors, not distinguish between row and column vectors. It will always be clear from the context if a vector is meant to be a row or a column vector.

4.2.1 Proposition. *Assume $x_1=0$ and that the term structure model (4.1.1) admits a benchmark realization F. Denote by $(\beta,c,K;r)^D$ the derivative of the Grigelionis process X, and let $P(t,T,r)=\exp\left(-\int_t^T F(t,s,r)ds\right)$ be the bond prices. Assume furthermore that the function Φ, defined as*

$$\begin{aligned}\Phi(t,T,r) &:= \langle\nabla_r P(t,T,r),\gamma(t,r)\beta(t,r)\rangle \qquad (4.2.1)\\ &+\frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial r_i\partial r_j}P(t,T,r)\langle\gamma_{i\bullet}(t,r)^*,c(t,r)\gamma_{j\bullet}(t,r)^*\rangle\\ &+\int_{\mathbb{R}^n}\Big(P(t,T,r+\gamma(t,r)x)-P(t,T,r)-\langle\nabla_r P(t,T,r),\gamma(t,r)x\rangle\Big)K_{t,r}(dx),\end{aligned}$$

exists on $\mathbb{R}_+\times\mathbb{R}_+\times\mathbb{R}^d$, and is continuous. Then, the bond prices satisfy for each $t\in(0,\infty), r\in\operatorname{supp}(r_{t-})$ and $T\geq t$

$$-P(t,T,r)r_1+\frac{\partial}{\partial t}P(t,T,r)+\langle\nabla_r P(t,T,r),\mu(t,r)\rangle+\Phi(t,T,r)=0.$$

Proof. See the appendix. □

Similar results for short rate models can be found in the literature. Raible [56, Prop. 4.12] provides such an equation for short rate models driven by a one-dimensional Lévy process. In Björk, Kabanov, and Runggaldier [15, Prop. 6.3] the short rate equation is driven by a Wiener process and a marked point process whose intensity may depend on the short rate. Both mentioned results deal with European options. We could also extend Proposition 4.2.1 in this direction, but for our purposes, it suffices to have the equation for the bond prices. For short rate models driven by a single standard Wiener process, Proposition 4.2.1 yields the differential equation

$$\begin{cases} -P(t,T,r)r + \frac{\partial}{\partial t}P(t,T,r) + \mu(t,r)\frac{\partial}{\partial r}P(t,T,r) + \frac{1}{2}\gamma(t,r)^2\frac{\partial^2}{\partial r^2}P(t,T,r) &= 0 \\ P(T,T,r) &= 1 \end{cases},$$

which is well known from the literature that deals with interest rate theory, see, e.g., Björk [7].

4.2.2 Proposition. *Assume $x_1 = 0$ and that the term structure model (4.1.1) admits a benchmark realization F. Denote by $(\beta, c, K; r)^D$ the derivative of the Grigelionis process X, and let $P(t,T,r) = \exp\left(-\int_t^T F(t,s,r)ds\right)$ be the bond prices. Assume furthermore that the function Φ, defined in (4.2.1), exists on $\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d$, and is continuous. Then, the bond prices satisfy for each $t \in (0,\infty), r \in \operatorname{supp}(r_{t-})$ and $T \geq t$*

$$\begin{aligned} &-r_1 + \frac{\partial}{\partial t}\big(\ln P(t,T,r)\big) + \langle \nabla_r\big(\ln P(t,T,r)\big), \mu(t,r) + \gamma(t,r)\beta(t,r)\rangle \\ &+ \frac{1}{2}\sum_{i,j=1}^d \left[\frac{\partial}{\partial r_i}\big(\ln P(t,T,r)\big)\frac{\partial}{\partial r_j}\big(\ln P(t,T,r)\big) + \frac{\partial^2}{\partial r_i \partial r_j}\big(\ln P(t,T,r)\big)\right] \\ &\qquad \langle \gamma_{i\bullet}(t,r)^*, c(t,r)\gamma_{j\bullet}(t,r)^*\rangle \\ &+ \int_{\mathbb{R}^n}\Big(e^{\ln P(t,T,r+\gamma(t,r)x) - \ln P(t,T,r)} - 1 - \langle \nabla_r\big(\ln P(t,T,r)\big), \gamma(t,r)x\rangle\Big)K_{t,r}(dx) = 0. \end{aligned}$$

Proof. We insert the relation $P(t,T,r) = e^{\ln P(t,T,r)}$ in the integro-differential equation of Proposition 4.2.1, and divide, after carrying out differentiations on the exponential functions, the obtained expression by $P(t,T,r)$. □

The proofs of the next two results rely on the HJM drift condition.

4.2.3 Proposition. *Assume that the term structure model (4.1.1) admits a benchmark realization F. If X is a non-degenerate Grigelionis process with derivative $(\beta, c, K; r)^D$, then the bond prices $P(t,T,r) = \exp\left(-\int_t^T F(t,s,r)ds\right)$ satisfy for each $t \in (0,\infty), r \in \operatorname{supp}(r_{t-})$ and $T \geq t$*

$$\begin{aligned} &-r_1\mathbf{1}_{\{x_1=0\}} + \frac{\partial}{\partial t}\big(\ln P(t,T,r)\big) + \langle \nabla_r\big(\ln P(t,T,r)\big), \mu(t,r)\rangle \\ &+ \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial r_i \partial r_j}\big(\ln P(t,T,r)\big)\langle \gamma_{i\bullet}(t,r)^*, c(t,r)\gamma_{j\bullet}(t,r)^*\rangle \\ &+ \Psi_{t,r}\Big(\nabla_r\big(\ln P(t,T,r)\big)\gamma(t,r)\Big) = 0. \end{aligned}$$

Proof. Inserting the formulas (4.1.10), (4.1.11) for α and σ in equation (4.1.14) yields

$$\begin{aligned}&\int_t^T \frac{\partial}{\partial t} F(t,s,r)ds + \left\langle \nabla_r \left(\int_t^T F(t,s,r)ds \right), \mu(t,r) \right\rangle \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial r_i \partial r_j} \left(\int_t^T F(t,s,r)ds \right) \langle \gamma_{i\bullet}(t,r)^*, c(t,r)\gamma_{j\bullet}(t,r)^* \rangle \\ &= \Psi_{t,r} \left(-\nabla_r \left(\int_t^T F(t,s,r)ds \right) \gamma(t,r) \right). \qquad (4.2.2)\end{aligned}$$

Note that the interchanging of differentiation and integration is valid, because all appearing derivatives are continuous and therefore bounded on compact intervals. Note furthermore that

$$\frac{\partial}{\partial t} \int_t^T F(t,s,r)ds = \int_t^T \frac{\partial}{\partial t} F(t,s,r)ds - F(t,t,r).$$

By the boundary condition (4.1.8) of Lemma 4.1.7 we infer

$$\int_t^T \frac{\partial}{\partial t} F(t,s,r)ds = r_1 \mathbf{1}_{\{x_1=0\}} + \frac{\partial}{\partial t} \int_t^T F(t,s,r)ds. \qquad (4.2.3)$$

Inserting (4.2.3) and $\int_t^T F(t,s,r)ds = -\ln P(t,T,r)$ into (4.2.2) yields the desired equation. □

4.2.4 Proposition. *Assume that the term structure model (4.1.1) admits a benchmark realization F. If X is a non-degenerate Grigelionis process with derivative $(\beta, c, K; r)^D$, then F satisfies for all $t \in (0,\infty), r \in \operatorname{supp}(r_{t-})$ and $T \geq t$*

$$\begin{aligned}&\frac{\partial}{\partial t} F(t,T,r) + \langle \nabla_r F(t,T,r), \mu(t,r) \rangle \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2}{\partial r_i \partial r_j} F(t,T,r) \langle \gamma_{i\bullet}(t,r)^*, c(t,r)\gamma_{j\bullet}(t,r)^* \rangle \\ &\quad + \left\langle \nabla_r F(t,T,r)\gamma(t,r), \nabla \Psi_{t,r} \left(- \left(\int_t^T \nabla_r F(t,s,r)ds \right) \gamma(t,r) \right) \right\rangle = 0.\end{aligned}$$

Proof. The stated equation follows directly by inserting the formulas (4.1.10), (4.1.11) for α and σ in equation (4.1.15) □

We have derived integro-differential equations under slightly different assumptions. As we shall see next, both methods lead to the same equations when the realization is affine. Of particular interest for the subsequent analysis of the realization problem is Proposition 4.2.8.

Remember that an affine realization $F(t,T,r) = a(t,T) + \langle b(t,T), r \rangle$ is always to be understood in the weak sense of Definition 4.1.4. If we say that the realization is affine everywhere, we mean that the relation $F(t,T,r) = a(t,T) + \langle b(t,T), r \rangle$ is valid for all $(t,T,r) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d$.

4.2.5 Proposition. *Assume $x_1 = 0$ and that the term structure model (4.1.1) admits an affine benchmark realization $F(t,T,r) = a(t,T) + \langle b(t,T), r\rangle$, which is affine everywhere, such that $-(\int_t^T b(t,s)ds)\gamma(t,r) \in Q_0$ (the set from Assumptions 4.1.1) for each $(t,T,r) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d$. Setting*

$$A(t,T) := \int_t^T a(t,s)ds \quad \text{and} \quad B(t,T) := \int_t^T b(t,s)ds,$$

it holds for all $t \in (0,\infty), r \in \operatorname{supp}(r_{t-})$ and $T \geq t$

$$\begin{aligned} r_1 + \frac{\partial}{\partial t}A(t,T) + \Big\langle \frac{\partial}{\partial t}B(t,T), r\Big\rangle + \langle B(t,T), \mu(t,r)\rangle \\ - \Psi_{t,r}(-B(t,T)\gamma(t,r)) = 0. \end{aligned}$$

Proof. Inserting $\ln P(t,T,r) = -(A(t,T) + \langle B(t,T), r\rangle)$ in the definition (4.2.1) of Φ, we obtain for each $(t,T,r) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d$

$$\Phi(t,T,r) = P(t,T,r)\Psi(t,r,-B(t,T)\gamma(t,r)).$$

This is seen by writing $P(t,T,r) = e^{\ln P(t,T,r)}$, as in the proof of Proposition 4.2.2. Therefore, Φ exists and is continuous by Assumptions 4.1.1. The stated equation follows from Proposition 4.2.2. □

4.2.6 Proposition. *Let X be linearly non-degenerate. Assume that the term structure model (4.1.1) admits an affine benchmark realization $F(t,T,r) = a(t,T) + \langle b(t,T), r\rangle$. Set*

$$A(t,T) := \int_t^T a(t,s)ds \quad \text{and} \quad B(t,T) := \int_t^T b(t,s)ds.$$

Let $t \in (0,\infty)$ be such that $\operatorname{int}\operatorname{supp}(r_{t-}) \neq \emptyset$. Then it holds for all $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$

$$\begin{aligned} r_1\mathbf{1}_{\{x_1=0\}} + \frac{\partial}{\partial t}A(t,T) + \Big\langle \frac{\partial}{\partial t}B(t,T), r\Big\rangle + \langle B(t,T), \mu(t,r)\rangle \\ - \Psi_{t,r}(-B(t,T)\gamma(t,r)) = 0. \end{aligned}$$

Proof. The assertion follows from inserting the representations (4.1.12), (4.1.13) for α and σ into (4.1.14), by taking into account

$$\begin{aligned} \int_t^T \frac{\partial}{\partial t}a(t,s)ds &= a(t,t) + \frac{\partial}{\partial t}\int_t^T a(t,s)ds, \\ \int_t^T \frac{\partial}{\partial t}b_i(t,s)ds &= b_i(t,t) + \frac{\partial}{\partial t}\int_t^T b_i(t,s)ds, \quad i = 1,\ldots,d \end{aligned}$$

and the boundary condition (4.1.9) of Lemma 4.1.7, which may be applied due to the hypothesis $\operatorname{int}\operatorname{supp}(r_{t-}) \neq \emptyset$. □

4.2.7 Proposition. *Assume $x_1 = 0$ and that the term structure model (4.1.1) admits an affine benchmark realization $F(t,T,r) = a(t,T) + \langle b(t,T), r\rangle$, which is affine everywhere, such that $-(\int_t^T b(t,s)ds)\gamma(t,r) \in Q_0$ (the set from Assumptions 4.1.1) for each $(t,T,r) \in \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d$. Then it holds for all $t \in (0,\infty), r \in \operatorname{supp}(r_{t-})$ and $T \geq t$*

$$\frac{\partial}{\partial t}a(t,T) + \Big\langle \frac{\partial}{\partial t}b(t,T), r\Big\rangle + \langle b(t,T), \mu(t,r)\rangle$$
$$+ \Big\langle b(t,T)\gamma(t,r), \nabla\Psi_{t,r}\Big(-\int_t^T b(t,s)ds\,\gamma(t,r)\Big)\Big\rangle = 0.$$

Proof. The claimed equation follows by differentiating the equation of Proposition 4.2.5 with respect to T. □

4.2.8 Proposition. *Let X be linearly non-degenerate. Assume that the term structure model (4.1.1) admits an affine benchmark realization $F(t,T,r) = a(t,T) + \langle b(t,T), r\rangle$. Let $t \in (0,\infty)$ be such that* $\operatorname{int}\operatorname{supp}(r_{t-}) \neq \emptyset$. *Then it holds for all $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$*

$$\frac{\partial}{\partial t}a(t,T) + \Big\langle \frac{\partial}{\partial t}b(t,T), r\Big\rangle + \langle b(t,T), \mu(t,r)\rangle$$
$$+ \Big\langle b(t,T)\gamma(t,r), \nabla\Psi_{t,r}\Big(-\int_t^T b(t,s)ds\,\gamma(t,r)\Big)\Big\rangle = 0.$$

Proof. The claimed equation follows by differentiating the equation of Proposition 4.2.6 with respect to T. □

4.3 Deterministic volatility

This section is devoted to the study of affine benchmark realizations of term structure models with deterministic volatility. Most of the time (see Theorem 4.3.6 for an exception) we assume that $\beta(t,\bullet)$, $c(t,\bullet)$ and $K(t,\bullet)$ from the derivative $(\beta, c, K; r)^D$ of the driving process X are constant on $\operatorname{supp}(r_{t-})$, which essentially means that X is a process with independent increments.

4.3.1 Proposition. *Assume the term structure model (4.1.1) has an affine benchmark realization $F(t,T,r) = a(t,T) + \langle b(t,T), r\rangle$. Let $t \in (0,\infty)$ be such that*

- *For each $T \geq t$, the mapping $\sigma(t,T,\bullet)$ is constant on* $\operatorname{supp}(r_{t-})$;
- *The mappings $\beta(t,\bullet), c(t,\bullet), K(t,\bullet)$ are constant on* $\operatorname{supp}(r_{t-})$.

Then, $\gamma(t,\bullet)$ is constant on $\operatorname{supp}(r_{t-})$, *and, for each $T \geq t$, the mapping $\alpha(t,T,\bullet)$ is constant on* $\operatorname{supp}(r_{t-})$. *Moreover, there are $\mu_1(t) \in \mathbb{R}^d$ and $\mu_2(t) \in \mathbb{R}^{d\times d}$ such that*

$$\mu(t,r) = \mu_1(t) + \mu_2(t)r, \quad r \in \operatorname{supp}(r_{t-}).$$

Proof. By Lemma 4.1.3, it holds for $r \in \mathbb{R}^d$

$$\gamma(t,r) = \begin{pmatrix} \sigma_1(t,t+x_1,r) & \cdots & \sigma_n(t,t+x_1,r) \\ \vdots & & \vdots \\ \sigma_1(t,t+x_d,r) & \cdots & \sigma_n(t,t+x_d,r) \end{pmatrix},$$

showing that $\gamma(t,\bullet)$ is constant on $\operatorname{supp}(r_{t-})$. According to Lemma 4.1.10, it holds for all $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$

$$\alpha(t,T,r) = -\Big\langle \sigma(t,T,r), \nabla\Psi_{t,r}\Big(-\int_t^T \sigma(t,s,r)ds \Big) \Big\rangle.$$

Thus, $\alpha(t,T,\bullet)$ is for every $T \geq t$ constant on $\operatorname{supp}(r_{t-})$, because $\beta(t,\bullet)$, $c(t,\bullet)$ and $K(t,\bullet)$ are constant on $\operatorname{supp}(r_{t-})$. Since the following identities concern only those $r \in \mathbb{R}^d$ from $\operatorname{supp}(r_{t-})$, we write $\alpha(t,T)$. By Lemma 4.1.3 it holds for $i = 1,\dots,d$

$$\mu_i(t,r) = \frac{\partial}{\partial T}a(t,T)\Big|_{T=t+x_i} + \Big\langle \frac{\partial}{\partial T}b(t,T)\Big|_{T=t+x_i}, r \Big\rangle + \alpha(t,t+x_i).$$

Defining $\mu_1(t) \in \mathbb{R}^d$ as

$$\mu_1^i(t) := \frac{\partial}{\partial T}a(t,T)\big|_{T=t+x_i} + \alpha(t,t+x_i), \quad i = 1,\dots,d,$$

and $\mu_2(t) \in \mathbb{R}^{d\times d}$ by

$$\mu_2(t) := \begin{pmatrix} \frac{\partial}{\partial T}b_1(t,T)|_{T=t+x_1} & \cdots & \frac{\partial}{\partial T}b_d(t,T)|_{T=t+x_1} \\ \vdots & & \vdots \\ \frac{\partial}{\partial T}b_1(t,T)|_{T=t+x_d} & \cdots & \frac{\partial}{\partial T}b_d(t,T)|_{T=t+x_d} \end{pmatrix}, \tag{4.3.1}$$

it follows that $\mu(t,r) = \mu_1(t) + \mu_2(t)r$ for all $r \in \operatorname{supp}(r_{t-})$. □

4.3.2 Proposition. *Assume X is linearly non-degenerate, and the term structure model (4.1.1) has an affine benchmark realization $F(t,T,r) = a(t,T) + \langle b(t,T), r\rangle$. Let $t \in (0,\infty)$ be such that*

- *For each $T \geq t$, the mapping $\sigma(t,T,\bullet)$ is constant on* $\operatorname{supp}(r_{t-})$*;*
- *The mappings $\beta(t,\bullet), c(t,\bullet), K(t,\bullet)$ are constant on* $\operatorname{supp}(r_{t-})$*;*
- $\operatorname{int}\operatorname{supp}(r_{t-}) \neq \emptyset$.

Then, the following identity is valid.

$$\frac{\partial}{\partial t}b(t,T) = -\mu_2(t)^* b(t,T), \quad T \geq t.$$

Proof. Since X is linearly non-degenerate, we may use equation (4.1.12) from Lemma 4.1.8, which yields that for all $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$

$$\alpha(t,T,r) = \frac{\partial}{\partial t} a(t,T) + \left\langle \frac{\partial}{\partial t} b(t,T), r \right\rangle + \langle b(t,T), \mu(t,r) \rangle.$$

By Proposition 4.3.1 there are $\mu_1(t) \in \mathbb{R}^d$ and $\mu_2(t) \in \mathbb{R}^{d\times d}$ such that $\mu(t,r) = \mu_1(t) + \mu_2(t) r$ for all $r \in \operatorname{supp}(r_{t-})$. Therefore, it holds for all $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$

$$\alpha(t,T,r) = \frac{\partial}{\partial t} a(t,T) + \langle b(t,T), \mu_1(t) \rangle + \left\langle \mu_2(t)^* b(t,T) + \frac{\partial}{\partial t} b(t,T), r \right\rangle$$

Since, by Proposition 4.3.1, $\alpha(t,T,\bullet)$ is constant on $\operatorname{supp}(r_{t-})$ for each $T \geq t$, and $\operatorname{int}\operatorname{supp}(r_{t-}) \neq \emptyset$ by assumption, differentiating with respect to $r_1, \dots, r_d$ yields the desired equation. □

The next result deals with the special case of short rate realizations, i.e. we consider term structure models of the type (4.1.1) with $d = 1$ and $x_1 = 0$.

4.3.3 Corollary. *Let X be linearly non-degenerate. Assume the term structure model (4.1.1) has an affine short rate realization $F(t,T,r) = a(t,T) + b(t,T)r$. Suppose for all $t \in (0,\infty)$ the following conditions are satisfied.*

- *For each $T \geq t$, the mapping $\sigma(t,T,\bullet)$ is constant on* $\operatorname{supp}(r_{t-})$*;*
- *The mappings $\beta(t,\bullet), c(t,\bullet), K(t,\bullet)$ are constant on* $\operatorname{supp}(r_{t-})$*;*
- $\operatorname{int}\operatorname{supp}(r_{t-}) \neq \emptyset$.

Then, there are $\tau_1, \dots, \tau_n : (0,\infty) \to \mathbb{R}$ and $\zeta \in C^1(0,\infty)$ with $\zeta(T) > 0$, $T \in (0,\infty)$, such that for each $i = 1, \dots, n$

$$\sigma_i(t,T,r) = \tau_i(t)\zeta(T), \quad t \in (0,\infty), r \in \operatorname{supp}(r_{t-}) \text{ and } T \geq t.$$

Proof. According to Proposition 4.3.2 and the boundary condition (4.1.9) from Lemma 4.1.7, for each fixed $T \in (0,\infty)$, the function $b(\bullet,T) : (0,T] \to \mathbb{R}$ satisfies the differential equation

$$\begin{cases} \frac{\partial}{\partial t} b(t,T) &= -\mu_2(t) b(t,T) \\ b(T,T) &= 1 \end{cases}.$$

The mapping μ_2, which is given by $\mu_2(t) = \frac{\partial}{\partial T} b(t,T)|_{T=t}$, is continuous, because $b \in C^{2,1}(\mathbb{R}_+ \times \mathbb{R}_+)$. Thus, the unique solution of the differential equation is given by

$$b(t,T) = \exp\left(\int_0^T \mu_2(s) ds - \int_0^t \mu_2(s) ds \right), \quad t \in (0,T].$$

Using the formula (4.1.13) from Lemma 4.1.8, we get for $t \in (0, \infty)$, $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$

$$\sigma(t, T, r) = \gamma(t, r) b(t, T) = \gamma(t, r) \exp\left(\int_0^T \mu_2(s) ds - \int_0^t \mu_2(s) ds \right).$$

Note that, for each $t \in (0, \infty)$, the mapping $\gamma(t, \bullet)$ is constant on $\operatorname{supp}(r_{t-})$ by Proposition 4.3.1. Choosing some $r(t) \in \operatorname{supp}(r_{t-})$ for every $t \in (0, \infty)$, the desired factorization is obtained by setting for $t, T \in (0, \infty)$ and $i = 1, \ldots, n$

$$\tau_i(t) := \gamma_i(t, r(t)) \exp\left(- \int_0^t \mu_2(s) ds \right) \quad \text{and} \quad \zeta(T) := \exp\left(\int_0^T \mu_2(s) ds \right).$$

□

4.3.4 Remark. *As apparent from the proof, mild regularity assumptions on the set* $\{(t, r) \,|\, t \in (0, \infty), r \in \operatorname{supp}(r_{t-})\}$ *imply that* $(r(t))_{t \in (0,\infty)}$ *can be chosen such that* $t \mapsto \gamma(t, r(t)) = \sigma(t, t, r(t))$ *is continuous. Then, the functions* τ_i *are also continuous. Alternatively, one can assume that* $\sigma(t, T)$ *depends nowhere on* r.

If the driving process X is a single Lévy process, such a factorization $\sigma(t, T) = \tau(t)\zeta(T)$ for deterministic volatility structures is known as necessary and sufficient condition for the Markov property of the short rate. This was first shown in Carverhill [20] for term structure models driven by Brownian motions. In Eberlein and Raible [29] this result was proven for a restricted class of Lévy processes, and extended in Küchler and Naumann [46] to the general case. This topic was further investigated in Gapeev and Küchler [35], and Corollary 4.3.3 may be regarded as an analogy of Gapeev and Küchler [35, Thm. 3.5] for the two-dimensional case, where the term structure model is driven by a Brownian motion and a purely discontinuous Lévy process with paths of finite variation on compacts. It is mentioned in Eberlein and Kluge [27, Sec. 5] that, with a multidimensional driving PIIAC, a factorization as in Corollary 4.3.3 implies that the short rate is Markovian, even though the driving process does not necessarily possess stationary increments. This may be seen as an analogy of Corollary 4.3.3 for several driving processes.

In order to avoid technicalities, we assume for the next result that $\sigma : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}^n$ does not depend on $r \in \mathbb{R}^d$, instead of only presuming that $\sigma(t, T, \bullet)$ is constant for all $r \in \operatorname{supp}(r_{t-})$, and we assume that the derivative of the driving process X is deterministic.

4.3.5 Proposition. *Let* X *be a linearly non-degenerate Grigelionis process with deterministic derivative, and* $\sigma \in C^{m,m}(\mathbb{R}_+ \times \mathbb{R}_+, \mathbb{R}^n)$ *for some* $m \in \mathbb{N}_0 \cup \{\infty\}$. *Assume the term structure model (4.1.1) has an affine* $d+1$*-dimensional benchmark realization* $F(t, T, r) = a(t, T) + \langle b(t, T), r \rangle$ *for some* $d \in \mathbb{N}$ *with* $b \in C^{m \vee 2, m \vee 1}(\mathbb{R}_+ \times \mathbb{R}_+, \mathbb{R}^d)$.

Then, for any $t \in (0,\infty)$ *with* $\operatorname{int}\operatorname{supp}(r_{t-}) \neq \emptyset$ *it holds concerning the mappings* $\sigma_1(t,\bullet),\ldots,\sigma_n(t,\bullet) : [t,\infty) \to \mathbb{R}$ *and* $b_1(t,\bullet),\ldots,b_d(t,\bullet) : [t,\infty) \to \mathbb{R}$

$$\frac{\partial}{\partial t^j}\sigma_i(t,\bullet) \in \operatorname{span}\{b_1(t,\bullet),\ldots,b_d(t,\bullet)\}, \quad i = 1,\ldots,n,\ j = 0,1,\ldots,m.$$

Proof. It is sufficient to show that for each $j = 0,1,\ldots,m$ there exists a mapping $M_j \in C^{m-j}((0,\infty),\mathbb{R}^{n\times d})$ such that

$$\frac{\partial^j}{\partial t^j}\sigma(t,T) = M_j(t)b(t,T) \quad \text{for all } T \in [t,\infty). \tag{4.3.2}$$

We prove (4.3.2) by induction on j.

1. By Lemma 4.1.3, the mapping $\gamma : \mathbb{R}_+ \to \mathbb{R}^{d\times n}$ is given by

$$\gamma(t) = \begin{pmatrix} \sigma_1(t,t+x_1) & \cdots & \sigma_n(t,t+x_1) \\ \vdots & & \vdots \\ \sigma_1(t,t+x_d) & \cdots & \sigma_n(t,t+x_d) \end{pmatrix},$$

and is thus of class C^m. By the formula (4.1.13) of Lemma 4.1.8, it holds

$$\sigma(t,T) = \gamma(t)^* b(t,T).$$

Thus, for $j = 0$, equation (4.3.2) holds with $M_0(s) := \gamma(s)^*$, $s \in (0,\infty)$.

2. Assume for some $j = 0,1,\ldots,m-1$ there is $M_j \in C^{m-j}((0,\infty),\mathbb{R}^{n\times d})$ such that (4.3.2) is satisfied. Since $\frac{\partial}{\partial t}b(t,T) = -\mu_2(t)^* b(t,T)$ by Proposition 4.3.2, we obtain

$$\frac{\partial^{j+1}}{\partial t^{j+1}}\sigma(t,T) = \left[\frac{\partial}{\partial t}M_j(t) - M_j(t)\mu_2(t)^*\right] b(t,T).$$

Hence, (4.3.2) is satisfied choosing $M_{j+1}(s) := \frac{\partial}{\partial s}M_j(s) - M_j(s)\mu_2(s)^*$, $s \in (0,\infty)$. Note that the demanded smoothness is satisfied, because $\frac{\partial}{\partial s}M_j(s)$ is of class C^{m-j-1}, and μ_2, which is given by equation (4.3.1) in the proof of Proposition 4.3.1, is of class C^{m-1} by the smoothness assumption on b.

□

For the case that σ and b are infinitely often differentiable in t and T, Proposition 4.3.5 yields that

$$\dim\operatorname{span}\left\{\tfrac{\partial^j}{\partial t^j}\sigma_i(t,\bullet)\,\middle|\, i = 1,\ldots,n,\ j \in \mathbb{N}_0\right\} \leq d,$$

implying that for every $i \in \{1,\dots,n\}$ there exists an integer $m_i \in \{0,\dots,d-1\}$ and a non-trivial vector $(\eta_{(i,0)}(t), \eta_{(i,1)}(t), \dots, \eta_{(i,m_i+1)}(t))$ such that

$$\sum_{j=0}^{m_i+1} \eta_{(i,j)}(t) \frac{\partial^j}{\partial t^j} \sigma_i(t,T) = 0 \quad \text{for all } T \in [t,\infty).$$

Note that this is exactly the differential equation (3.4.1) from Section 3.4.

We will briefly explore the situation when the volatility is stationary, that is $\sigma(t,T,r) = \tilde{\sigma}(T-t,r)$ with $\tilde{\sigma} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^n$, and one has an affine, stationary benchmark realization $F(t,T,r) = a(t,T) + \langle \tilde{b}(T-t), r \rangle$. For the sake of simplicity, we assume that the support $\operatorname{supp}(r_{t-})$ does not depend on $t \in (0,\infty)$. Denoting these supports by S, we assume that, for each $x \in \mathbb{R}_+$, the mapping $\tilde{\sigma}(x,\bullet)$ is constant on S. We also presume that $\beta(t,\bullet)$, $c(t,\bullet)$ and $K(t,\bullet)$ are constant on S.

Arguing similarly as in the proof of Proposition 4.3.1, we find out the following. The mapping γ is constant on $(0,\infty) \times S$, namely it holds

$$\gamma(t,r) = \begin{pmatrix} \tilde{\sigma}_1(x_1,r) & \cdots & \tilde{\sigma}_n(x_1,r) \\ \vdots & & \vdots \\ \tilde{\sigma}_1(x_d,r) & \cdots & \tilde{\sigma}_n(x_d,r) \end{pmatrix}.$$

The mappings $\alpha(t,T,\bullet)$ are constant on S. If, moreover, the process X has stationary increments in the sense that $\Psi_{t,r}$ does not depend on t, then there is a function $\tilde{\alpha} : \mathbb{R}_+ \to \mathbb{R}$ such that $\alpha(t,T,r) = \tilde{\alpha}(T-t)$ for all $(t,r) \in (0,\infty) \times S$ and $T \geq t$. There are $\mu_1 : (0,\infty) \to \mathbb{R}^d$ and $\mu_2 \in \mathbb{R}^{d \times d}$ such that

$$\mu(t,r) = \mu_1(t) + \mu_2 r, \quad \text{for all } (t,r) \in (0,\infty) \times S,$$

namely, letting $r \in S$ be an arbitrary point from the support, choose μ_1 as

$$\mu_1^i(t) := \frac{\partial}{\partial T} a(t,T)\big|_{T=t+x_i} + \alpha(t, t+x_i, r), \quad i = 1,\dots,d,$$

and the $d \times d$ matrix μ_2 as

$$\mu_2 := \begin{pmatrix} \tilde{b}_1'(x_1) & \cdots & \tilde{b}_d'(x_1) \\ \vdots & & \vdots \\ \tilde{b}_1'(x_d) & \cdots & \tilde{b}_d'(x_d) \end{pmatrix}.$$

The analogous result to Proposition 4.3.2 is the differential equation

$$\tilde{b}'(x) = \mu_2^* \tilde{b}(x), \quad x \in \mathbb{R}_+.$$

Using the identity $\tilde{\sigma}(x) = \tilde{b}(x)\gamma$, it is immediately seen from solving this differential equation that $\tilde{\sigma}$ is quasi-exponential Björk and Svensson [13, Sec. 5], since it is of the form $\tilde{\sigma}(x) = Ae^{Bx}C$ with $A \in \mathbb{R}^{1\times d}, B \in \mathbb{R}^{d\times d}$ and $C \in \mathbb{R}^{d\times n}$. For short rate realizations (i.e. $d = 1$), we obtain in particular $\tilde{\sigma}_i(x) = \eta_i e^{\lambda x}$ with constants $\eta_1, \dots, \eta_n \in \mathbb{R}$ and $\lambda \in \mathbb{R}$, which is a version of Corollary 4.3.3 for stationary volatilities. For a single driving process, i.e. $n = 1$, this is a coincidence with Eberlein and Raible [29, Thm. 4.4]. If $\lambda \neq 0$, one has the Hull-White extension Hull and White [39] of the Vasiček model Vasiček [61], and the Ho-Lee model Ho and Lee [38] for $\lambda = 0$.

The reformulation of Proposition 4.3.5 for stationary volatilities is that the derivatives $\tilde{\sigma}_i^{(j)}$, $j = 0, 1, \dots, m$, where $m \in \mathbb{N}_0 \cup \{\infty\}$ depends on the degrees of smoothness of $\tilde{\sigma}$ and $\tilde{b}$, belong to the linear space $\operatorname{span}\{\tilde{b}_1, \dots, \tilde{b}_d\}$. In particular, if $\tilde{\sigma}$ and $\tilde{b}$ are of class C^∞, it follows that

$$\dim \operatorname{span}\left\{\tilde{\sigma}_i^{(j)} \,|\, i = 1, \dots, n,\ j \in \mathbb{N}_0\right\} \leq d,$$

implying that for every $i \in \{1, \dots, n\}$ there exists an integer $m_i \in \{0, \dots, d-1\}$ and a non-trivial vector $(\eta_{(i,0)}, \eta_{(i,1)}, \dots, \eta_{(i,m_i+1)})$ such that

$$\sum_{j=0}^{m_i+1} \eta_{(i,j)} \tilde{\sigma}_i^{(j)}(x) = 0 \quad \text{for all } x \in \mathbb{R}_+.$$

Now, we deal with short rate realizations, that is term structure models of the type (4.1.1) with $d = 1$ and $x_1 = 0$, which are driven by a one-dimensional process, i.e. $n = 1$. We assume that the one-dimensional X is a linearly non-degenerate Grigelionis process with derivative $(\beta, c, K; r)^D$.

4.3.6 Theorem. *Assume the term structure model (4.1.1) has an affine short rate realization $F(t,T,r) = a(t,T) + b(t,T)r$. Let $t \in (0,\infty)$ such that the following conditions are satisfied.*

- $a(t,\bullet), b(t,\bullet) \in C^\infty([t,\infty))$;
- *For each $T \geq t$, the mapping $\sigma(t,T,\bullet)$ is constant on* $\operatorname{supp}(r_{t-})$;
- $\sigma(t,t,r) \neq 0$ *for all* $r \in \operatorname{supp}(r_{t-})$;
- *There are $\beta_1(t), \beta_2(t), c_1(t), c_2(t) \in \mathbb{R}$ such that $\beta(t,r) = \beta_1(t) + \beta_2(t)r$ and $c(t,r) = c_1(t) + c_2(t)r$ for all $r \in \operatorname{supp}(r_{t-})$;*
- $\operatorname{int}\operatorname{supp}(r_{t-}) \neq \emptyset$.

Then, the mapping $\gamma(t,\bullet)$ is constant on $\operatorname{supp}(r_{t-})$, and there are $\alpha_1(t,\bullet), \alpha_2(t,\bullet) : [t,\infty) \to \mathbb{R}$, $\mu_1(t), \mu_2(t) \in \mathbb{R}$ and signed measures $K_1(t), K_2(t)$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that, for all $r \in \operatorname{supp}(r_{t-})$

$$\begin{aligned}\alpha(t,T,r) &= \alpha_1(t,T) + \alpha_2(t,T)r, \quad T \geq t\\ \mu(t,r) &= \mu_1(t) + \mu_2(t)r,\\ K(t,r) &= K_1(t) + K_2(t)r.\end{aligned}$$

Proof. By Lemma 4.1.3, it holds $\gamma(t,r) = \sigma(t,t,r)$ for $r \in \mathbb{R}$, showing that $\gamma(t,\bullet)$ is constant on $\operatorname{supp}(r_{t-})$. Since all forthcoming equations in this proof are valid for $r \in \operatorname{supp}(r_{t-})$, we denote it by $\gamma(t)$. Note that by hypothesis it holds $\gamma(t) \neq 0$. By Proposition 4.2.8, the equation

$$\frac{\partial}{\partial t}a(t,T) + \frac{\partial}{\partial t}b(t,T)r + b(t,T)\mu(t,r) \\ + b(t,T)\gamma(t)\Psi'_{t,r}\left(-\gamma(t)\int_t^T b(t,s)ds\right) = 0 \tag{4.3.3}$$

is valid for all $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$. By the boundary condition (4.1.9) of Lemma 4.1.7, it holds $b(t,t) = 1$, and, by hypothesis, $b(t,\bullet)$ is continuous (in fact, even of class C^∞). Thus, there is a neighborhood $U_t \subset (0,\infty)$ of t such that

$$b(t,T) \neq 0 \quad \text{for all } T \in U_t.$$

Hence, the following sequence $\Gamma^i_k(t,T)$ is well defined for $T \in [t,\infty) \cap U_t$. For $i = 1,2$ and $k \in \mathbb{N}$ set

$$\Gamma^i_k(t,T) := \begin{cases} -\frac{1}{b(t,T)}\frac{\partial}{\partial t}a(t,T), & i = 1 \text{ and } k = 1 \\ -\frac{1}{b(t,T)}\frac{\partial}{\partial t}b(t,T), & i = 2 \text{ and } k = 1 \\ -\frac{1}{b(t,T)}\frac{\partial}{\partial T}\Gamma^i_{k-1}(t,T), & k \neq 1 \end{cases}.$$

Since $a(t,\bullet), b(t,\bullet) \in C^\infty([t,\infty))$ by assumption, it holds $\Gamma^i_k(t,\bullet) \in C^\infty([t,\infty) \cap U_t)$ for $i = 1,2$ and $k \in \mathbb{N}$. Dividing (4.3.3) by $b(t,T)$ for $T \in [t,\infty) \cap U_t$, we get

$$\mu(t,r) + \gamma(t)\Psi'_{t,r}\left(-\gamma(t)\int_t^T b(t,s)ds\right) = \Gamma^1_1(t,T) + \Gamma^2_1(t,T)r \tag{4.3.4}$$

for all $r \in \operatorname{supp}(r_{t-})$ and $T \in [t,\infty) \cap U_t$. Taking $T = t$ in (4.3.4) yields, since $\Psi'_{t,r}(0) = \beta(t,r)$ by equation (2.2.13)

$$\mu(t,r) + \gamma(t)(\beta_1(t) + \beta_2(t)r) = \Gamma^1_1(t,t) + \Gamma^2_1(t,t)r, \quad r \in \operatorname{supp}(r_{t-}).$$

Setting $\mu_i(t) := \Gamma^i_1(t,t) - \gamma(t)\beta_i(t)$, $i = 1,2$, we obtain

$$\mu(t,r) = \mu_1(t) + \mu_2(t)r, \quad r \in \operatorname{supp}(r_{t-}).$$

Since the driving process X is linearly non-degenerate, we may apply Lemma 4.1.8, and obtain from equation (4.1.12) for all $T \geq t$ and $r \in \operatorname{supp}(r_{t-})$

$$\alpha(t,T,r) = \frac{\partial}{\partial t}a(t,T) + \frac{\partial}{\partial t}b(t,T)r + b(t,T)\mu(t,r).$$

Defining $\alpha_1(t,T) := \frac{\partial}{\partial t}a(t,T) + b(t,T)\mu_1(t)$ and $\alpha_2(t,T) := \frac{\partial}{\partial t}b(t,T) + b(t,T)\mu_2(t)$ for $T \in [t,\infty)$ yields

$$\alpha(t,T,r) = \alpha_1(t,T) + \alpha_2(t,T)r, \quad r \in \operatorname{supp}(r_{t-}), T \geq t.$$

Differentiating (4.3.4) with respect to T and dividing by $b(t,T)$, $T \in [t,\infty) \cap U_t$, we obtain

$$\gamma(t)^2 \Psi''_{t,r}\left(-\gamma(t)\int_t^T b(t,s)ds\right) = \Gamma_2^1(t,T) + \Gamma_2^2(t,T)r \tag{4.3.5}$$

for all $r \in \operatorname{supp}(r_{t-})$ and $T \in [t,\infty) \cap U_t$. Taking $T = t$ in (4.3.5) yields, since $\Psi''_{t,r}(0) = c(t,r) + \int_{\mathbb{R}} x^2 K_{t,r}(dx)$ by equation (2.2.14)

$$\int_{\mathbb{R}} x^2 K_{t,r}(dx) = \frac{\Gamma_2^1(t,t)}{\gamma(t)^2} - c_1(t) + \left(\frac{\Gamma_2^2(t,t)}{\gamma(t)^2} - c_2(t)\right) r, \quad r \in \operatorname{supp}(r_{t-}), \tag{4.3.6}$$

Differentiating (4.3.5) several times with respect to T and dividing by $b(t,T)$, $T \in [t,\infty) \cap U_t$, we get

$$\gamma(t)^k \Psi^{(k)}_{t,r}\left(-\gamma(t)\int_t^T b(t,s)ds\right) = \Gamma_k^1(t,T) + \Gamma_k^2(t,T)r, \quad k \geq 3 \tag{4.3.7}$$

for all $r \in \operatorname{supp}(r_{t-})$ and $T \in [t,\infty) \cap U_t$. Taking $T = t$ in (4.3.7) yields, since $\Psi^{(k)}_{t,r}(0) = \int_{\mathbb{R}} x^k K_{t,r}(dx)$, $k \geq 3$ by equation (2.2.15)

$$\int_{\mathbb{R}} x^k K_{t,r}(dx) = \frac{\Gamma_k^1(t,t)}{\gamma(t)^k} + \frac{\Gamma_k^2(t,t)}{\gamma(t)^k} r, \quad k \geq 3 \tag{4.3.8}$$

for all $r \in \operatorname{supp}(r_{t-})$. Let

$$r^-(t) := \inf \operatorname{supp}(r_{t-}) \quad \text{and} \quad r^+(t) := \sup \operatorname{supp}(r_{t-}).$$

It holds $r^-(t) < r^+(t)$, because int $\operatorname{supp}(r_{t-}) \neq \emptyset$ by hypothesis. Define the sequences $r_n^-(t), r_n^+(t)$ for $n \in \mathbb{N}$ as

$$r_n^-(t) := \begin{cases} r^-(t), & r^-(t) > -\infty \\ \inf\left(\operatorname{supp}(r_{t-}) \cap [-n,\infty)\right), & r^-(t) = -\infty \end{cases},$$

and

$$r_n^+(t) := \begin{cases} r^+(t), & r^+(t) < \infty \\ \sup\left(\operatorname{supp}(r_{t-}) \cap (-\infty,n]\right), & r^+(t) = \infty \end{cases}.$$

Note that $r_n^-(t), r_n^+(t) \in \operatorname{supp}(r_{t-})$ for all $n \in \mathbb{N}$, because the support of r_{t-} is a closed set. For $i = 1, 2$ and $n \in \mathbb{N}$ define the signed measures

$$K_n^i(t) := \begin{cases} -\frac{r_n^-(t)}{r_n^+(t)-r_n^-(t)} K(t, r_n^+(t)) + \frac{r_n^+(t)}{r_n^+(t)-r_n^-(t)} K(t, r_n^-(t)), & i = 1 \\ \frac{1}{r_n^+(t)-r_n^-(t)} K(t, r_n^+(t)) - \frac{1}{r_n^+(t)-r_n^-(t)} K(t, r_n^-(t)), & i = 2 \end{cases}. \tag{4.3.9}$$

For arbitrary $n \in \mathbb{N}$ and $r \in \mathbb{R}$ it holds

$$K_n^1(t) + K_n^2(t) r = \frac{r - r_n^-(t)}{r_n^+(t) - r_n^-(t)} K(t, r_n^+(t)) + \frac{r_n^+(t) - r}{r_n^+(t) - r_n^-(t)} K(t, r_n^-(t)), \tag{4.3.10}$$

which shows that, for each $r \in [r_n^-(t), r_n^+(t)]$, the measures $K_n^1(t) + K_n^2(t)r$ are true (not just signed) measures. It is also clear that they satisfy the condition (2.2.11) concerning finite exponential moments, because they are defined by means of the measures $K_{t,r}$, which have this property. By relations (4.3.6), (4.3.8) and (4.3.10) we obtain for all $n \in \mathbb{N}$ and $r \in \operatorname{supp}(r_{t-})$

$$\int_{\mathbb{R}} x^k K_{t,r}(dx) = \int_{\mathbb{R}} x^k (K_n^1(t) + K_n^2(t) r)(dx), \quad k \geq 2.$$

Applying Corollary 2.2.8, it holds for every $n \in \mathbb{N}$

$$K(t, r) = K_n^1(t) + K_n^2(t) r, \quad r \in \operatorname{supp}(r_{t-}) \cap [r_n^-(t), r_n^+(t)]. \tag{4.3.11}$$

As, we have mentioned above, it holds $r_n^-(t), r_n^+(t) \in \operatorname{supp}(r_{t-})$, $n \in \mathbb{N}$. Inserting $K(t, r_n^+(t)) = K_{n+1}^1(t) + K_{n+1}^2(t) r_n^+(t)$ and $K(t, r_n^-(t)) = K_{n+1}^1(t) + K_{n+1}^2(t) r_n^-(t)$ in (4.3.10), we obtain for every $n \in \mathbb{N}$

$$K_n^1(t) + K_n^2(t) r = K_{n+1}^1(t) + K_{n+1}^2(t) r, \quad r \in \operatorname{supp}(r_{t-}) \cap [r_{n+1}^-(t), r_{n+1}^+(t)]. \tag{4.3.12}$$

Since $\operatorname{int} \operatorname{supp}(r_{t-}) \neq \emptyset$, and $r_n^-(t) \downarrow r^-(t)$, $r_n^+(t) \uparrow r^+(t)$, there is an index $n_0 \in \mathbb{N}$ such that

$$K_{n_0}^1(t) = K_n^1(t) \quad \text{and} \quad K_{n_0}^2(t) = K_n^2(t), \quad n \geq n_0. \tag{4.3.13}$$

This follows by inserting any arbitrary Borel set in (4.3.12), and differentiating with respect to r (if n is large enough, the interval $[r_n^-(t), r_n^+(t)]$ contains a point from the interior of $\operatorname{supp}(r_{t-})$, which is non-empty by assumption). We conclude from (4.3.11) and (4.3.13) that for all $n \geq n_0$

$$K(t, r) = K_{n_0}^1(t) + K_{n_0}^2(t) r, \quad r \in \operatorname{supp}(r_{t-}) \cap [r_n^-(t), r_n^+(t)]. \tag{4.3.14}$$

Since $r_n^-(t) \downarrow r^-(t)$ and $r_n^+(t) \uparrow r^+(t)$, equation (4.3.14) actually holds for all $r \in \operatorname{supp}(r_{t-})$. □

The assumptions of Theorem 4.3.6 concerning the derivative $(\beta, c, K; r)^D$ are in particular satisfied if X is the sum of a Brownian motion and a purely discontinuous local martingale, that is $X = \sqrt{c}W + x * (\mu^X - \nu)$, where $c \geq 0$ is a constant, and W is a standard Wiener process. The derivative of X is then given by $(0, c, K; r)^D$, and applying Theorem 4.3.6 yields that $K(t, r) = K_1(t) + K_2(t)r$.

As an example, let N be a point process with compensator $\int_0^t \lambda(s, r_{s-})ds$, i.e. the intensities are given by a deterministic function $\lambda : \mathbb{R}_+ \times \mathbb{R} \to (0, \infty)$. Define the driving process as $X_t := \sqrt{c}W_t + N_t - \int_0^t \lambda(s, r_{s-})ds$, where $c \geq 0$ is a constant, and W is a standard Wiener process. Then, the intensities have the affine structure

$$\lambda(t, r) = \lambda_1(t) + \lambda_2(t)r.$$

This is a coincidence with equation (29) in Hyll [40, Thm. 1], where, in a similar context, an affine structure of the intensities of driving counting processes is shown. The marked point process, considered in Björk, Kabanov, and Runggaldier [15, Prop. 6.5], also has an absolutely continuous compensator, which depends on the short rate in an affine manner.

4.4 Deterministic direction volatility

The results of Section 3.5 have suggested that for deterministic direction volatility models, with driving processes that make jumps, there exists no finite dimensional realization. Within the present framework, we provide a rigorous proof. The following result concerns any fixed time point $t \in (0, \infty)$. It can as well be applied for intervals $I \subset \mathbb{R}_+$, and this should be viewed in connection with Lemma 2.2.7.

We assume that the driving process in (4.1.1) is a one-dimensional linearly non-degenerate Grigelionis process with derivative $(\beta, c, K; r)^D$. An inherent class of processes which satisfy the assumptions of the following theorem are Lévy processes with non-trivial Lévy measure.

4.4.1 Theorem. *Assume $x_1 = 0$ and that the term structure model (4.1.1) has an affine $d + 1$-dimensional benchmark realization $F(t, T, r) = a(t, T) + \langle b(t, T), r\rangle$. Let $t \in (0, \infty)$ be such that*

- $a(t, \bullet), b_1(t, \bullet), \ldots, b_d(t, \bullet) \in C^\infty([t, \infty))$;
- $\overline{\text{int supp}(r_{t-})} = \text{supp}(r_{t-})$;
- *The mapping $K(t, \bullet)$ is non-zero and constant on* $\text{supp}(r_{t-})$;
- *There are $\lambda(t, \bullet) : [t, \infty) \to \mathbb{R}$ with $\lambda(t, t) \neq 0$, and $\varphi(t, \bullet) : \mathbb{R}^d \to \mathbb{R}$ of class $C^\infty(\text{int supp}(r_{t-}))$, such that*

$$\sigma(t, T, r) = \varphi(t, r)\lambda(t, T), \quad r \in \text{supp}(r_{t-}), T \geq t.$$

Then, $\varphi(t,\bullet)$ *is constant on* $\operatorname{supp}(r_{t-})$.

Proof. Our goal is to establish that for all $i = 1, \dots, d$

$$\frac{\partial}{\partial r_i}\varphi(t,r) = 0, \quad r \in \operatorname{int}\operatorname{supp}(r_{t-}) \tag{4.4.1}$$

which, due to the assumption $\overline{\operatorname{int}\operatorname{supp}(r_{t-})} = \operatorname{supp}(r_{t-})$, then proves that $\varphi(t,\bullet)$ is constant on $\operatorname{supp}(r_{t-})$. By Proposition 4.2.8 the following equation is satisfied for all $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$.

$$\begin{aligned}\frac{\partial}{\partial t}a(t,T) + \Big\langle \frac{\partial}{\partial t}b(t,T), r\Big\rangle + \langle b(t,T), \mu(t,r)\rangle \\ + \langle \gamma(t,r), b(t,T)\rangle \Psi'_{t,r}\left(-\Big\langle \gamma(t,r), \int_t^T b(t,s)ds\Big\rangle\right) = 0.\end{aligned} \tag{4.4.2}$$

Set $\lambda(t) := (\lambda(t, t+x_1), \dots, \lambda(t, t+x_d)) \in \mathbb{R}^d$. By Lemma 4.1.3 it holds

$$\gamma(t,r) = (\sigma(t, t+x_1, r), \dots, \sigma(t, t+x_d, r)) = \varphi(t,r)\lambda(t), \quad r \in \operatorname{supp}(r_{t-}).$$

Together with the smoothness assumptions on $a(t,\bullet)$ and $b(t,\bullet)$ we can rewrite equation (4.4.2) as

$$\sum_{i=1}^{2d+1} \psi_i(t,r)\eta_i(t,T) + \varphi(t,r)\langle \lambda(t), b(t,T)\rangle \Psi'_{t,r}\left(-\varphi(t,r)\Big\langle \lambda(t), \int_t^T b(t,s)ds\Big\rangle\right) = 0, \tag{4.4.3}$$

valid for all $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$, with functions $\psi_i(t,\bullet) : \mathbb{R}^d \to \mathbb{R}$ and $\eta(t,\bullet) \in C^\infty([t,\infty))$ for $i = 1, \dots, 2d+1$. Noting that $x_1 = 0$, it follows from the hypothesis $\lambda(t,t) \neq 0$ and the boundary condition (4.1.9) from Lemma 4.1.7 that $\langle \lambda(t), b(t,t)\rangle = \lambda(t,t) \neq 0$. By the continuity of $b(t,\bullet)$ there is a neighborhood $U_t \subset (0,\infty)$ of t, such that

$$\langle \lambda(t), b(t,T)\rangle \neq 0 \quad \text{for all } T \in U_t.$$

Dividing (4.4.3) by $\langle \lambda(t), b(t,T)\rangle$ for $T \in [t,\infty) \cap U_t$, we get

$$\sum_{i=1}^{2d+1} \psi_i(t,r)\kappa_i(t,T) + \varphi(t,r)\Psi'_{t,r}\left(-\varphi(t,r)\Big\langle \lambda(t), \int_t^T b(t,s)ds\Big\rangle\right) = 0 \tag{4.4.4}$$

for all $r \in \operatorname{supp}(r_{t-})$ and $T \in [t,\infty) \cap U_t$, where $\kappa_i(t,\bullet) \in C^\infty([t,\infty) \cap U_t)$, $i = 1, \dots, 2d+1$. Differentiating (4.4.4) with respect to T and dividing by $\langle \lambda(t), b(t,T)\rangle$, and repeating this procedure two times, we arrive at an equation of the form

$$\sum_{i=1}^{2d+1} \psi_i(t,r)\tau_i(t,T) + \varphi(t,r)^4\Psi^{(4)}_{t,r}\left(-\varphi(t,r)\Big\langle \lambda(t), \int_t^T b(t,s)ds\Big\rangle\right) = 0, \tag{4.4.5}$$

valid for all $r \in \operatorname{supp}(r_{t-})$ and $T \in [t,\infty) \cap U_t$, with $\tau_i(t,\bullet) \in C^\infty([t,\infty) \cap U_t)$, $i = 1,\ldots,2d+1$. Observe that equation (4.4.5) is of the form

$$\sum_{i\in I} \psi_i(t,r)\zeta_i(t,T) + \sum_{k=4}^{n} \varphi(t,r)^k \Gamma_k(t,T)\Psi_{t,r}^{(k)}\left(-\varphi(t,r)\Big\langle \lambda(t), \int_t^T b(t,s)ds\Big\rangle\right) = 0, \tag{4.4.6}$$

fulfilled for all $r \in \operatorname{supp}(r_{t-})$ and $T \in [t,\infty) \cap V_t$, where the following boundary conditions are satisfied.

$$\begin{cases} I \subset \mathbb{N} \text{ is a finite (possibly empty) set} \\ n \geq 4 \text{ is an even integer} \\ V_t \subset (0,\infty) \text{ is a neighborhood of } t \\ \zeta_i(t,\bullet), \Gamma_k(t,\bullet) \in C^\infty([t,\infty)\cap V_t) \text{ for each } i \in I \text{ and each } k = 4,\ldots,n \\ \Gamma_n(t,t) \neq 0 \end{cases}. \tag{4.4.7}$$

Suppose in equation (4.4.6) there is an index $j \in I$ such that $\zeta_j(t,t) \neq 0$. Then, there is a neighborhood $W_t \subset V_t$ of t such that $\zeta_j(t,T) \neq 0$ for all $T \in [t,\infty) \cap W_t$. Dividing (4.4.6) by $\zeta_j(t,T)$, $T \in [t,\infty) \cap W_t$ gives an equation of the form

$$\begin{aligned} &\psi_j(t,r) + \sum_{i\in I\setminus\{j\}} \psi_i(t,r)\tilde{\zeta}_i(t,T) \\ &\quad + \sum_{k=4}^{n} \varphi(t,r)^k \tilde{\Gamma}_k(t,T)\Psi_{t,r}^{(k)}\left(-\varphi(t,r)\Big\langle \lambda(t), \int_t^T b(t,s)ds\Big\rangle\right) = 0, \end{aligned} \tag{4.4.8}$$

valid for all $r \in \operatorname{supp}(r_{t-})$ and $T \in [t,\infty) \cap W_t$, with functions $\tilde{\zeta}_i(t,\bullet) \in C^\infty([t,\infty)\cap W_t)$, $i \in I \setminus \{j\}$ and $\tilde{\Gamma}_k(t,\bullet) \in C^\infty([t,\infty) \cap W_t)$, $k = 4,\ldots,n$. Differentiating (4.4.8) with respect to T we obtain

$$\begin{aligned} &\sum_{i\in I\setminus\{j\}} \psi_i(t,r)\frac{\partial}{\partial T}\tilde{\zeta}_i(t,T) \\ &\quad + \sum_{k=4}^{n} \Bigg[\varphi(t,r)^k \frac{\partial}{\partial T}\tilde{\Gamma}_k(t,T)\Psi_{t,r}^{(k)}\left(-\varphi(t,r)\Big\langle \lambda(t), \int_t^T b(t,s)ds\Big\rangle\right) \\ &\quad - \varphi(t,r)^{k+1}\tilde{\Gamma}_k(t,T)\langle\lambda(t), b(t,T)\rangle\Psi_{t,r}^{(k+1)}\left(-\varphi(t,r)\Big\langle \lambda(t), \int_t^T b(t,s)ds\Big\rangle\right)\Bigg] = 0 \end{aligned} \tag{4.4.9}$$

for all $r \in \operatorname{supp}(r_{t-})$ and $T \in [t,\infty) \cap W_t$. We recognize (4.4.9) as an equation of the form (4.4.6). Concerning (4.4.7), we note that the term $\tilde{\Gamma}_n(t,t)\langle\lambda(t), b(t,t)\rangle$, which is obtained for $k = n+1$, is unequal to zero, and that all other boundary conditions from (4.4.7) are fulfilled with the exception that $n+1$ is an odd number. Therefore, we differentiate (4.4.9) with respect to T. Arguing as above, this yields an equation of the form (4.4.6) satisfying all boundary conditions from (4.4.7).

Repeating the described procedure, dividing by some $\zeta_j(t,T)$ with $\zeta_j(t,t) \neq 0$ and differentiating two times with respect to T, sufficiently often, we arrive at an equation of the form (4.4.6), satisfying the boundary conditions (4.4.7), and the additional condition

$$\zeta_i(t,t) = 0 \quad \text{for all } i \in I. \tag{4.4.10}$$

Recall that $K(t,\bullet)$ is assumed to be constant on $\operatorname{supp}(r_{t-})$. Since the equations in this proof are valid for $r \in \operatorname{supp}(r_{t-})$, we denote it by K_t. Taking $T = t$ in (4.4.6), we obtain, due to (4.4.10), and since $\Psi_{t,r}^{(k)}(0) = \int_{\mathbb{R}} x^k K_t(dx)$, $k \geq 4$ for $r \in \operatorname{supp}(r_{t-})$ by equation (2.2.15),

$$\sum_{k=4}^{n} \varphi(t,r)^k \Gamma_k(t,t) \int_{\mathbb{R}} x^k K_t(dx) = 0, \quad r \in \operatorname{supp}(r_{t-}). \tag{4.4.11}$$

Now fix an arbitrary $r_0 \in \operatorname{int} \operatorname{supp}(r_{t-})$. In order to establish (4.4.1), assume, on the contrary, that $\frac{\partial}{\partial r_i}\varphi(t,r_0) \neq 0$ for some index $i \in \{1,\ldots,d\}$. Then, there is a neighborhood $U(r_0) \subset \operatorname{int} \operatorname{supp}(r_{t-})$ of r_0 such that $\frac{\partial}{\partial r_i}\varphi(t,r) \neq 0$ for all $r \in U(r_0)$. Differentiating (4.4.11) with respect to r_i gives

$$\frac{\partial}{\partial r_i}\varphi(t,r) \sum_{k=4}^{n} k\varphi(t,r)^{k-1} \Gamma_k(t,t) \int_{\mathbb{R}} x^k K_t(dx) = 0, \quad r \in U(r_0). \tag{4.4.12}$$

We can divide (4.4.12) by $\frac{\partial}{\partial r_i}\varphi(t,r)$, $r \in U(r_0)$, and then differentiate once more with respect to r_i. Repeating the procedure of dividing by $\frac{\partial}{\partial r_i}\varphi(t,r)$ and then differentiating with respect to r_i sufficiently many times, we arrive at

$$\frac{\partial}{\partial r_i}\varphi(t,r) n! \Gamma_n(t,t) \int_{\mathbb{R}} x^n K_t(dx) = 0, \quad r \in U(r_0). \tag{4.4.13}$$

Since $\int_{\mathbb{R}} x^n K_t(dx) > 0$ (the measure K_t is non-trivial by assumption, and n is an even integer) and $\Gamma_n(t,t) \neq 0$ according to (4.4.7), equation (4.4.13) yields the contradiction that $\frac{\partial}{\partial r_i}\varphi(t,r_0) = 0$. This completes the proof. □

4.4.2 Remark. *As it is seen from the proof, the intrinsic reason, which causes the non-existence of a finite dimensional realization, is the fact that the $\Psi_{t,r}$ are no polynomials. The argumentation of the proof fails if, for instance, X is a Brownian motion.*

At this juncture, we shall take a look at the results of Gapeev and Küchler [35, Sec. 4], where finite dimensional realizations for term structure models with non-deterministic volatilities, driven by jump-diffusions, are studied. Let a HJM forward rate model with two driving processes

$$\begin{cases} df(t,T) &= \alpha(t,T,r_{t-})dt + \sigma_1(t,T,r_{t-})dW_t + \sigma_2(t,T,r_{t-})dL_t \\ f(0,T) &= f^*(0,T) \end{cases} \tag{4.4.14}$$

be given, where r denotes the short rate, W denotes a standard Wiener process, and L a compound Poisson process. Gapeev and Küchler [35, Thm. 4.1] provides a finite dimensional realization, which is affine, for forward rate models of the type

$$\begin{cases} df(t,T) &= \alpha(t,T,r_{t-})dt + \sigma(t,T,r_{t-})dW_t + \delta(t,x,T,r_{t-})\mu(dt,dx) \\ f(0,T) &= f^*(0,T) \end{cases},$$

where μ denotes a homogeneous Poisson random measure, and the volatilities are of the form $\sigma(t,T,r) = \varphi(t,r)\lambda(T)$ and $\delta(t,x,T,r) = \zeta(T)$. Under further conditions, the short rate is one component of a finite dimensional Markov process. As a matter of fact, this result cannot be applied to term structure models of the type (4.4.14), because δ is not of the form $\delta(t,x,T,r) = x\eta(t,T,r)$.

If the volatilities in (4.4.14) are of the form $\sigma_i(t,T,r) = \varphi_i(t,r)\lambda_i(T)$ for $i = 1,2$, then, according to Gapeev and Küchler [35, Remark 4.3], the short rate process does not need to be a component of a finite dimensional Markov process, but it is a component of an infinite dimensional one. This infinite dimensional process consists, besides finitely many other components, of the sequence of processes (see equation (4.34) in Gapeev and Küchler [35])

$$\xi_n(t) = \lambda_2(t) \int_0^t \varphi_2(s,r_s)^n \int_{\mathbb{R}} x^n e^{x\varphi_2(s,r_s)\int_s^t \lambda_2(v)dv} K(dx)ds, \quad n \in \mathbb{N}$$

where K denotes the Lévy measure of the compound Poisson process. Presuming that λ_2 and φ_2 are non-trivial, we see that the sequence $(\xi_n)_{n\in\mathbb{N}}$ can only have finitely many non-zero components if the Lévy measure K is trivial, and in this case all ξ_n are zero. As in the proof of Theorem 4.4.1, the crucial point is whether the Lévy measure K is zero, or equivalently, if the cumulant generating function Ψ is a polynomial.

Another reference that deals with volatility structures depending on the short rate is Ritchken and Sankarasubramanian [58]. Consider term structure models

$$\begin{cases} df(t,T) &= \alpha(t,T,r_{t-})dt + \sigma(t,T,r_{t-})dX_t \\ f(0,T) &= f^*(0,T) \end{cases},$$

where r denotes the short rate, X denotes a one-dimensional Brownian motion, and the volatility is of the form $\sigma(t,T,r) = \varphi(t,r)\lambda(t,T)$. Ritchken and Sankarasubramanian [58] show that for some models of this type there exists a three-dimensional (one dimension is for the time t) affine realization

$$F(t,T,r,z) = a(t,T) + b_1(t,T)r + b_2(t,T)z,$$

namely if $\sigma(t,T,r) = \varphi(t,r)\lambda(T)$, see condition (2.11) in Ritchken and Sankarasubramanian [58]. In this case, the two-dimensional state process (r,Z) is a Markov process, and one component is the short rate. Without stressing the technical details, we shall now see that this is, in principle, impossible for non-deterministic volatilities if the driving process X admits jumps (for instance if it is a Lévy process with non-trivial Lévy measure). Assume there exists $x \in \mathbb{R}_+$ such that $b_2(t, t+x) \neq 0$ for each $t \in \mathbb{R}_+$. Then we conclude from

$$r_t(x) = a(t, t+x) + b_1(t, t+x)r_t + b_2(t, t+x)Z_t, \quad t \in \mathbb{R}_+$$

that the process Z can be expressed as

$$Z_t = \frac{1}{b_2(t, t+x)}\Big(r_t(x) - a(t, t+x) - b_1(t, t+x)r_t\Big), \quad t \in \mathbb{R}_+.$$

Inserting this identity into

$$f(t, T) = a(t, T) + b_1(t, T)r_t + b_2(t, T)Z_t, \quad t \in \mathbb{R}_+$$

we see that the term structure model has an affine benchmark realization. Applying Theorem 4.4.1 yields that $\sigma(t, T, r)$ does not depend on r.

Theorem 4.4.1 has further consequences for short rate realizations. We assume that the driving process X is one-dimensional, that is we consider forward rate models of the form (4.1.1) with $d = n = 1$ and $x_1 = 0$.

4.4.3 Corollary. *Let X be linearly non-degenerate. Assume the term structure model (4.1.1) has an affine short rate realization $F(t, T, r) = a(t, T) + b(t, T)r$ with $a, b \in C^{2,\infty}(\mathbb{R}_+ \times \mathbb{R}_+)$. Suppose, for all $t \in (0, \infty)$, the following conditions are satisfied.*

- *The mappings $\beta(t, \bullet), c(t, \bullet), K(t, \bullet)$ are constant on $\operatorname{supp}(r_{t-})$;*
- $K_t \neq 0$
- $\overline{\operatorname{int}\operatorname{supp}(r_{t-})} = \operatorname{supp}(r_{t-})$*;*
- $\sigma(t, t, \bullet) \in C^\infty(\operatorname{int}\operatorname{supp}(r_{t-}))$*.*

Then, $\gamma(t, \bullet)$ is constant on $\operatorname{supp}(r_{t-})$ for every $t \in (0, \infty)$, and there are $\tau : (0, \infty) \to \mathbb{R}$ and $\zeta \in C^1(0, \infty)$ with $\zeta(T) > 0$, $T \in (0, \infty)$, such that

$$\sigma(t, T, r) = \tau(t)\zeta(T), \quad t \in (0, \infty), r \in \operatorname{supp}(r_{t-}) \text{ and } T \geq t.$$

Proof. First, we observe that $\gamma(t, r) = \sigma(t, t, r)$ for all $(t, r) \in (0, \infty) \times \mathbb{R}$ by Lemma 4.1.3. According to equation (4.1.13) of Lemma 4.1.8 it holds

$$\sigma(t, T, r) = \gamma(t, r)b(t, T)$$

for all $t \in (0, \infty)$, $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$. Note that all required conditions of Theorem 4.4.1 are fulfilled, in particular, it holds $b(t, t) = 1$ by the boundary condition (4.1.9) of Lemma 4.1.7, and $\gamma(t, \bullet) \in C^\infty(\operatorname{int}\operatorname{supp}(r_{t-}))$ by hypothesis. Applying

Theorem 4.4.1 yields that, for each $t \in (0, \infty)$, the mapping $\gamma(t, \bullet)$ is constant on $\operatorname{supp}(r_{t-})$. Consequently, for all $t \in (0, \infty)$ and $T \geq t$, the mapping $\sigma(t, T, \bullet)$ is constant on $\operatorname{supp}(r_{t-})$. We conclude from Corollary 4.3.3 the factorization

$$\sigma(t, T, r) = \tau(t)\zeta(T), \quad t \in (0, \infty), r \in \operatorname{supp}(r_{t-}) \text{ and } T \geq t,$$

with functions $\tau : (0, \infty) \to \mathbb{R}$ and $\zeta \in C^1(0, \infty)$ such that $\zeta(T) > 0$, $T \in (0, \infty)$. □

We remark another coincidence with Hyll [40, Thm. 1]. Equation (28) of the cited theorem shows that, for short rate realizations driven by finitely many Wiener processes and finitely many point processes, the short rate volatilities of the point processes do necessarily not depend on the short rate.

Corollary 4.4.3 can be improved if we demand that the driving process makes arbitrary small jumps. The idea is to show that the given short rate realization must, because of the small jumps, necessarily be affine. Thinking of the geometry, described at the beginning of Chapter 3, this is plausible, because otherwise, the process jumps out of a non-affine manifold. While Corollary 4.4.3 can be applied to Poisson or compound Poisson processes for instance, the next result can, for example, be used if the driving process is a bilateral Gamma process or a generalized hyperbolic process.

4.4.4 Theorem. *Let X be non-degenerate. Assume the term structure model (4.1.1) has a short rate realization $F \in C^{2,\infty,2}(\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R})$. Suppose, for all $t \in (0, \infty)$, the following conditions are satisfied.*

- *The mappings $\beta(t, \bullet), c(t, \bullet), K(t, \bullet)$ are constant on* $\operatorname{supp}(r_{t-})$*;*
- $0 \in \overline{\operatorname{int}\operatorname{supp}(K_t)}$*;*
- *The support* $\operatorname{supp}(r_{t-})$ *is a bounded or unbounded interval, and* $\inf \operatorname{supp}(r_{t-}) < \sup \operatorname{supp}(r_{t-})$*;*
- $\sigma(t, t, \bullet) \in C^\infty(\operatorname{int}\operatorname{supp}(r_{t-}))$ *and* $\sigma(t, t, r) \neq 0$ *for all* $r \in \operatorname{supp}(r_{t-})$*.*

Then, $\gamma(t, \bullet)$ is constant on $\operatorname{supp}(r_{t-})$ *for every $t \in (0, \infty)$, and there are $\tau : (0, \infty) \to \mathbb{R}$ and $\zeta \in C^1(0, \infty)$ with $\tau(t) \neq 0$, $t \in (0, \infty)$ and $\zeta(T) > 0$, $T \in (0, \infty)$, such that*

$$\sigma(t, T, r) = \tau(t)\zeta(T), \quad t \in (0, \infty), r \in \operatorname{supp}(r_{t-}) \text{ and } T \geq t.$$

Proof. First, we observe that $\gamma(t, r) = \sigma(t, t, r)$ for all $(t, r) \in (0, \infty) \times \mathbb{R}$ by Lemma 4.1.3. Fix $t \in (0, \infty)$ and $T \geq t$. Equations (3.1.12) and (3.1.14) of Theorem 3.1.17 yield that for all $r \in \operatorname{supp}(r_{t-})$ and all $x \in \operatorname{supp}(K_t)$

$$\gamma(t, r)x \frac{\partial}{\partial r} F(t, T, r) = F(t, T, r + \gamma(t, r)x) - F(t, T, r). \tag{4.4.15}$$

Let $I \subset \operatorname{supp}(r_{t-})$ be an arbitrary, non-empty, compact interval. By the assumptions on $\sigma(t, t, \bullet)$, it holds $\inf_{r \in I} |\gamma(t, r)| \in (0, \infty)$, and since $0 \in \overline{\operatorname{int}\operatorname{supp}(K_t)}$, there is an $\varepsilon > 0$ such that one of the intervals $[0, \varepsilon]$, $[-\varepsilon, 0]$ is contained in $\operatorname{supp}(K_t)$. Let $M := \varepsilon \inf_{r \in I} |\gamma(t,r)|$. Then, by equation (4.4.15), there exists, for each $r_0 \in I$, a compact interval $J(r_0) \subset \mathbb{R}$ of length M, with $r_0 \in \{\inf J(r_0), \sup J(r_0)\}$, such that

$$(r - r_0)\frac{\partial}{\partial r}F(t,T,r_0) = F(t,T,r) - F(t,T,r_0) \quad \text{for all } r \in J(r_0).$$

Consequently, $r \mapsto \frac{\partial}{\partial r}F(t,T,r)$ is constant on $J(r_0)$. Since each interval $J(r_0)$ is of constant, positive length, the continuity of the mapping $r \mapsto \frac{\partial}{\partial r}F(t,T,r)$ implies that it is constant on I. The compact interval $I \subset \operatorname{supp}(r_{t-})$ was allowed to be arbitrary. Hence, it follows that $r \mapsto \frac{\partial}{\partial r}F(t,T,r)$ is constant on $\operatorname{supp}(r_{t-})$, say

$$\frac{\partial}{\partial r}F(t,T,r) = b(t,T) \quad \text{for all } r \in \operatorname{supp}(r_{t-}). \tag{4.4.16}$$

Consequently, there are mappings $a, b : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ such that for all $t \in (0,\infty)$, $r \in \operatorname{supp}(r_{t-})$ and $T \geq t$

$$F(t,T,r) = a(t,T) + b(t,T)r. \tag{4.4.17}$$

Differentiating (4.4.17) with respect to r shows, by the assumption $F \in C^{2,\infty,2}(\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R})$, that $a, b \in C^{2,\infty}(\mathbb{R}_+ \times \mathbb{R}_+)$. Inserting (4.4.16) and (4.4.17) in (4.4.15), we obtain for all $t \in (0,\infty)$, $r \in \operatorname{supp}(r_{t-})$, $x \in \operatorname{supp}(K_t)$ and $T \geq t$

$$F(t,T,r + \gamma(t,r)x) = a(t,T) + b(t,T)(r + \gamma(t,r)x). \tag{4.4.18}$$

Taking together (4.4.16), (4.4.17) and (4.4.18), we conclude that F is an affine short rate realization in the sense of Definition 4.1.4. Therefore, Corollary 4.4.3 yields the factorization

$$\sigma(t,T,r) = \tau(t)\zeta(T), \quad t \in (0,\infty), r \in \operatorname{supp}(r_{t-}) \text{ and } T \geq t,$$

with functions $\tau : (0,\infty) \to \mathbb{R}$ and $\zeta \in C^1(0,\infty)$ such that $\zeta(T) > 0$, $T \in (0,\infty)$. Since, by hypothesis, $\sigma(t,t,r) = \tau(t)\zeta(t) \neq 0$ for every $t \in (0,\infty)$ and $r \in \operatorname{supp}(r_{t-})$, we infer that $\tau(t) \neq 0$ for all $t \in (0,\infty)$. □

4.4.5 Remark. *Assuming that the support S of r_{t-} does not depend on $t \in (0,\infty)$, an analogous argumentation reveals that for short rate models with an, a priori arbitrary, stationary volatility $\sigma(t,T,r) = \tilde{\sigma}(T-t,r)$, it follows that γ is constant on $(0,\infty) \times S$, and that there are constants $c, \lambda \in \mathbb{R}$ such that $\tilde{\sigma}(x,r) = ce^{\lambda x}$, $(x,r) \in \mathbb{R}_+ \times S$.*

The result that the short rate volatility $\gamma(t, \bullet)$ is constant on $\operatorname{supp}(r_{t-})$ differs notably from what is known for short rate models driven by a Brownian motion. We briefly sketch this setup. Presuming, there exists an affine short rate realization $F(t,T,r) = a(t,T) + b(t,T)r$, Proposition 4.2.8 yields an equation of the form

$$b(t,T)\mu(t,r) - \gamma(t,r)^2 b(t,T)\int_t^T b(t,s)ds = \eta_1(t,T) + \kappa_1(t,T)r.$$

Dividing this equation by $b(t,T)$ and differentiating with respect to T, we get an equation of the type

$$\gamma(t,r)^2 b(t,T) = \eta_2(t,T) + \kappa_2(t,T)r.$$

Taking $T = t$ we obtain that the short rate volatility must be of the form

$$\gamma(t,r)^2 = \gamma_1(t) + \gamma_2(t)r.$$

This condition is well known in the literature that deals with mathematical finance, see, e.g., Björk [7]. An example is the model in Cox, Ingersoll, and Ross [24], where one has a short rate volatility of the form $\gamma(t,r) = c\sqrt{r}$. We refer to Jeffrey [44] for further details.

We will now discuss a converse of Theorem 4.4.4, which gives us the opportunity to show connections to some results in the literature. Assume that the derivative of the driving process X is deterministic, and that the volatility $\sigma : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ is deterministic and factorizes, more precisely $\sigma(t,T) = \tau(t)\zeta(T)$ with $\tau, \zeta : \mathbb{R}_+ \to \mathbb{R}$ such that $\tau(t) \neq 0$, $t \in \mathbb{R}_+$ and $\zeta(T) \neq 0$, $T \in \mathbb{R}_+$. We moreover assume that $\tau \in C^1(\mathbb{R}_+)$ and $\zeta \in (\mathbb{R}_+)$. Then, the volatility satisfies the differential equation

$$\frac{\partial}{\partial t}\sigma(t,T) - \frac{\tau'(t)}{\tau(t)}\sigma(t,T) = 0.$$

According to Proposition 3.4.6, there exists a two-dimensional realization (G,Z), which is (see the proof of Proposition 3.4.6) given by the mapping $G : \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, defined as

$$G(t,T,z) := f^*(0,T) + \int_0^t \alpha(s,T)ds + \sigma(t,T)z,$$

and the one-dimensional state process Z, which is the unique solution of the stochastic differential equation

$$\begin{cases} dZ_t &= -\frac{\tau'(t)}{\tau(t)} Z_t dt + dX_t \\ Z_0 &= 0. \end{cases}.$$

Since $\sigma(t,t) \neq 0$ for all $t \in \mathbb{R}_+$, we may apply Proposition 4.1.5, which yields the existence of a short rate realization (F,r) for the term structure model. For further computations, we assume that the driving process X is a Lévy process with cumulant generating function Ψ. Applying Proposition 4.1.5 again, the short rate realization (F,r) is given by the mapping

$$\begin{aligned} F(t,T,r) = f^*(0,T) - \int_0^t \left[\Psi'\left(-\int_s^T \sigma(s,v)dv \right) - \Psi'\left(-\int_s^t \sigma(s,v)dv \right) \right] \sigma(s,T)ds \\ + \frac{\zeta(T)}{\zeta(t)}\Big(r - f^*(0,t)\Big), \end{aligned}$$

which corresponds to equation (4.12) in Eberlein and Raible [29]. Using Lemma 4.1.3, it turns out that the dynamics of the short rate are

$$\begin{aligned} dr_t = \Bigg[& \frac{\partial}{\partial t} f^*(0,t) - \Psi'(0)\sigma(t,t) + \int_0^t \frac{\partial^2}{\partial t^2} \Psi \left(- \int_s^t \sigma(s,v)dv \right) ds \\ & - \frac{\zeta'(t)}{\zeta(t)} \left(\int_0^t \frac{\partial}{\partial t} \Psi \left(- \int_s^t \sigma(s,v)dv \right) ds + f^*(0,t) - r_t \right) \Bigg] dt + \sigma(t,t)dX_t. \end{aligned} \tag{4.4.19}$$

This expression is in accordance with equation (4.7) in Eberlein and Raible [29]. If the function $\frac{\zeta'(t)}{\zeta(t)}$, $t \in \mathbb{R}_+$ is bounded (which is for instance satisfied for stationary forward rate volatilities, because then $\zeta(t) = e^{\lambda t}$ for some $\lambda \in \mathbb{R}$), the coefficients of (4.4.19) are Lipschitz in the sense of Protter [55, p. 236]. Then, by Protter [55, Thm. V.32], the short rate r is a Markov process, i.e. for $u \geq t$ and any bounded, Borel measurable function f

$$\mathbb{E}[f(r_u) \,|\, \mathcal{G}_t] = \mathbb{E}[f(r_u) \,|\, \sigma(r_t)],$$

where $(\mathcal{G}_t)$ denotes the natural filtration $\mathcal{G}_t = \sigma(r_s, s \leq t)$, generated by the short rate process. This is a certain coincidence with the results of Eberlein and Raible [29] and Küchler and Naumann [46], where the Markov property of the short rate was investigated with respect to the filtration generated by the Lévy process X.

4.5 Generic benchmark realizations

Assume there is a benchmark realization F for the interest rate model (4.1.1). It arises the question whether a slightly different model $\tilde{f}(t,T)$, namely one with the same drift α and volatility σ, but with another initial forward rate curve, say

$$\begin{cases} d\tilde{f}(t,T) &= \alpha(t,T,r_{t-})dt + \sigma(t,T,r_{t-})dX_t \\ \tilde{f}(0,T) &= \tilde{f}^*(0,T) \end{cases},$$

still admits a benchmark realization. In other words, we are looking for generic benchmark realizations.

For technical reasons, we assume $d \leq n$, i.e. the number of benchmark forward rates is less than or equal to the number of driving processes, which essentially means that we deal with minimal realizations, as it is the case for short rate realizations. We moreover presume $x_1 = 0$, i.e. the first benchmark forward rate is the short rate, and we assume that there exists a convex neighborhood $U(r^*) \subset \mathbb{R}^d$ of r^* (the initial condition of the benchmark forward rates r), such that

$$U(r^*) \subset \operatorname{supp}(r_{t-}) \quad \text{for all } t \in (0,\infty). \tag{4.5.1}$$

Furthermore, we assume the existence of a subset $I = \{i_1 < \ldots < i_d\} \subset \{1, \ldots, n\}$ such that the mapping $\hat{\sigma} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d \times d}$, defined as

$$\hat{\sigma}(t,r) := \begin{pmatrix} \sigma_{i_1}(t, t + x_1, r) & \cdots & \sigma_{i_d}(t, t + x_1, r) \\ \vdots & & \vdots \\ \sigma_{i_1}(t, t + x_d, r) & \cdots & \sigma_{i_d}(t, t + x_d, r) \end{pmatrix},$$

satisfies

$$\det \hat{\sigma}(t,r) \neq 0 \quad \text{for all } (t,r) \in (0,\infty) \times U(r^*). \tag{4.5.2}$$

We introduce the d-dimensional vector $\hat{\sigma} := (\sigma_i)_{i \in I}$. As we have argued in Section 3.2, if one considers more than one initial forward rate curve, it does not make sense to allow the derivative of the driving process to depend on the state process. Therefore, we assume that the driving process X is a n-dimensional non-degenerate Grigelionis process with deterministic derivative $(\beta, c, K)^D$. We define the set S of allowable realizations as follows.

4.5.1 Definition. *Denote by S the set of all mappings $\tilde{F} \in C^{2,1,2}(\mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}^d)$ such that*

1. *$\tilde{F}$ is a benchmark realization for an interest rate model of the form*

$$\begin{cases} d\tilde{f}(t,T) &= \alpha(t,T,r_{t-})dt + \sigma(t,T,r_{t-})dX_t \\ \tilde{f}(0,T) &= \tilde{f}^*(0,T) \end{cases},$$

where $\tilde{f}^*(0,\bullet) \in C^1(\mathbb{R}_+)$ is chosen such that $r^* = (\tilde{f}^*(0,x_1), \ldots, \tilde{f}^*(0,x_d))$.

2. *It holds $U(r^*) \subset \operatorname{supp}(\tilde{r}_{t-})$ for all $t \in (0,\infty)$ and $(\mathcal{D}(X), \tilde{r})$ has regular supports, where $\tilde{r}$ denotes the benchmark forward rates $\tilde{r}_t = (\tilde{f}(t,t+x_1), \ldots, \tilde{f}(t,t+x_d))$.*

4.5.2 Lemma. *The following statements are valid.*

1. *It holds $F \in S$.*

2. *For each $\tilde{F} \in S$, the identity*

$$\nabla_r \tilde{F}(t,T,r) = \hat{\sigma}(t,T,r)\hat{\sigma}^{-1}(t,r)$$

is valid for all $(t,r) \in (0,\infty) \times U(r^*)$ and $T \geq t$.

3. *For all $\tilde{F} \in S$, the associated benchmark forward rates $\tilde{r}$ satisfy the stochastic differential equation*

$$\begin{cases} d\tilde{r}_t &= \tilde{\mu}(t,\tilde{r}_{t-})dt + \gamma(t,\tilde{r}_{t-})dX_t \\ \tilde{r}_0 &= r^* \end{cases},$$

where $\tilde{\mu} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ *has the components*

$$\tilde{\mu}_i(t,r) = \frac{\partial}{\partial T}\tilde{F}(t,T,r)\Big|_{T=t+x_i} + \alpha(t,t+x_i,r), \quad i = 1,\dots,d.$$

Proof. The first statement is obvious, the second follows from equation (4.1.11) of Lemma 4.1.8 and Lemma 4.1.3, and the third is a consequence of Lemma 4.1.3. □

In the sequel, $\mathcal{J}f$ and $\mathcal{H}f$ denote the Jacobian and the Hessian matrix of a function f. The proof of the upcoming result is based on methods used in Jeffrey [44].

4.5.3 Theorem. *Assume for every* $t \in (0,\infty)$ *there are* $r \in U(r^*)$ *and* $T \geq t$ *such that* $\det \mathcal{H}_r F(t,T,r) \neq 0$*, or, equivalently,* $\det \mathcal{J}_r \hat{\sigma}(t,T,r)\hat{\sigma}^{-1}(t,r) \neq 0$*. Then, it holds* $f^*(0,\bullet) = \tilde{f}^*(0,\bullet)$ *for every* $\tilde{F} \in S$*.*

Proof. First of all, $\det \mathcal{H}_r F(t,T,r) \neq 0$ is equivalent to $\det \mathcal{J}_r \hat{\sigma}(t,T,r)\hat{\sigma}^{-1}(t,r) \neq 0$ by the second statement of Lemma 4.5.2.

Let $t \in (0,\infty)$ be arbitrary, and choose an arbitrary $\tilde{F} \in S$. According to Lemma 4.5.2 and equation (4.1.10) of Lemma 4.1.8, it holds for all $r \in U(r^*)$ and $T \geq t$

$$\begin{aligned}\alpha(t,T,r) = {} & \frac{\partial}{\partial t}\tilde{F}(t,T,r) + \langle \nabla_r F(t,T,r), \tilde{\mu}(t,r)\rangle \\ & + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial r_i \partial r_j} F(t,T,r)\langle \gamma_{i\bullet}(t,r)^*, c(t)\gamma_{j\bullet}(t,r)^*\rangle.\end{aligned}$$

Integrating $\int_t^T \alpha(s,T,r)ds$, we obtain, since $\tilde{F}(T,T,r) = r_1$, $r \in U(r^*)$ by the boundary condition (4.1.8) of Lemma 4.1.7,

$$\begin{aligned}\tilde{F}(t,T,r) = {} & \int_t^T \langle \nabla_r F(s,T,r), \tilde{\mu}(s,r)\rangle ds - \int_t^T \alpha(s,T,r)ds + r_1 \\ & + \frac{1}{2}\sum_{i,j=1}^{d}\int_t^T \frac{\partial^2}{\partial r_i \partial r_j} F(s,T,r)\langle \gamma_{i\bullet}(s,r)^*, c(s)\gamma_{j\bullet}(s,r)^*\rangle ds \end{aligned} \tag{4.5.3}$$

for $r \in U(r^*)$ and $T \geq t$. For fixed $T \geq t$ and $r \in U(r^*)$ define $\tilde{h}_{T,r} : [0,1] \to \mathbb{R}$ by

$$\tilde{h}_{T,r}(x) := \tilde{F}(t,T,r^* + x(r-r^*)), \quad x \in [0,1].$$

Note that $r^* + x(r - r^*) \in U(r^*)$ for each $x \in [0,1]$, because $U(r^*)$ is a convex neighborhood of r^* by assumption. Therefore, it holds for $r \in U(r^*)$ and $T \geq t$

$$\begin{aligned}\tilde{F}(t,T,r) - \tilde{F}(t,T,r^*) &= \tilde{h}_{T,r}(1) - \tilde{h}_{T,r}(0) = \int_0^1 \tilde{h}'_{T,r}(x)dx \\ &= \int_0^1 \sum_{i=1}^{d}(r_i - r_i^*)\frac{\partial}{\partial r_i}\tilde{F}(t,T,r^* + x(r-r^*))dx \\ &= \Big\langle r - r^*, \int_0^1 \nabla_r F(t,T,r^* + x(r-r^*))dx\Big\rangle. \end{aligned} \tag{4.5.4}$$

For $i = 1, \dots, d$ we define the vectors

$$\tilde{v}_i(t) := -\frac{\partial}{\partial t}\tilde{F}(t,T,r^*)\Big|_{T=t+x_i},$$

$$w_i(t,r) := \alpha(t,t+x_i,r) - \Big\langle r - r^*, \frac{\partial}{\partial t}\int_0^1 \nabla_r F(t,T,r^* + x(r-r^*))dx\Big\rangle\Big|_{T=t+x_i}.$$

Using Corollary 4.1.9, we infer from equation (4.5.4)

$$\tilde{\mu}(t,r) = \tilde{v}(t) + w(t,r), \quad r \in U(r^*).$$

Inserting this identity into equation (4.5.3) yields for $r \in U(r^*)$ and $T \geq t$

$$\begin{aligned}\tilde{F}(t,T,r) &= \int_t^T \langle \nabla_r F(s,T,r), \tilde{v}(s) + w(s,r)\rangle ds - \int_t^T \alpha(s,T,r)ds + r_1 \\ &+ \frac{1}{2}\sum_{i,j=1}^d \int_t^T \frac{\partial^2}{\partial r_i \partial r_j}F(s,T,r)\langle \gamma_{i\bullet}(s,r)^*, c(s)\gamma_{j\bullet}(s,r)^*\rangle ds. \end{aligned} \tag{4.5.5}$$

Incorporating formula (4.5.4), and differentiating with respect to t, we obtain

$$\begin{aligned}&\frac{\partial}{\partial t}\tilde{F}(t,T,r^*) + \Big\langle r - r^*, \frac{\partial}{\partial t}\int_0^1 \nabla_r F(t,T,r^* + x(r-r^*))dx\Big\rangle \\ &\quad + \langle \nabla_r F(t,T,r), \tilde{v}(t) + w(t,r)\rangle - \alpha(t,T,r) \\ &\quad + \frac{1}{2}\sum_{i,j=1}^d \frac{\partial^2}{\partial r_i \partial r_j}F(t,T,r)\langle \gamma_{i\bullet}(t,r)^*, c(t)\gamma_{j\bullet}(t,r)^*\rangle = 0 \end{aligned} \tag{4.5.6}$$

for $r \in U(r^*)$ and $T \geq t$. Equation (4.5.6) is valid for every $\tilde{F} \in S$, in particular for F itself. We deduce that

$$\frac{\partial}{\partial t}F(t,T,r^*) - \frac{\partial}{\partial t}\tilde{F}(t,T,r^*) = \langle \nabla_r F(t,T,r), \tilde{v}(t) - v(t)\rangle$$

for $r \in U(r^*)$ and $T \geq t$. Differentiating with respect to $r_1, \dots, r_d$ we obtain

$$\mathcal{H}_r F(t,T,r)(\tilde{v}(t) - v(t)) = 0 \quad \text{for all } r \in U(r^*) \text{ and } T \geq t.$$

Since $t \in (0,\infty)$ was allowed to be arbitrary, and, by hypothesis, there are $r \in U(r^*)$ and $T \geq t$ such that $\det \mathcal{H}_r F(t,T,r) \neq 0$, we conclude that

$$\tilde{v}(t) - v(t) = 0, \quad t \in (0,\infty).$$

Equation (4.5.5) is valid for every $\tilde{F} \in S$, in particular for F itself. So, we deduce that for all $(t,r) \in (0,\infty) \times U(r^*)$ and $T \geq t$

$$F(t,T,r) - \tilde{F}(t,T,r) = 0.$$

By the continuity of F and $\tilde{F}$ it follows that for all $T \in (0, \infty)$

$$F(0, T, r^*) - \tilde{F}(0, T, r^*) = \lim_{t \downarrow 0} \Big(F(t, T, r^*) - \tilde{F}(t, T, r^*) \Big) = 0.$$

By the boundary condition (4.1.7) of Lemma 4.1.7 it follows that $f^*(0, \bullet) = \tilde{f}^*(0, \bullet)$. $\square$

Theorem 4.5.3 generalizes Jeffrey [44] concerning the driving processes, and in the direction that we deal with benchmark realizations, instead of short rate realizations. In the case of short rate realizations, i.e. $d = 1$ and $x_1 = 0$, hypothesis (4.5.2) means that there has to exist an index $i \in \{1, \dots, n\}$ such that $\sigma_i(t, t, r) \neq 0$ for all $(t, r) \in (0, \infty) \times U(r^*)$. Then, we obtain:

4.5.4 Corollary. *Assume for every $t \in (0, \infty)$ there are $r \in U(r^*)$ and $T \geq t$ such that $\frac{\partial^2}{\partial r^2} F(t, T, r) \neq 0$, or, equivalently, $\frac{\partial}{\partial r} \left[\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)} \right] \neq 0$. Then, it holds $f^*(0, \bullet) = \tilde{f}^*(0, \bullet)$ for every $\tilde{F} \in S$.*

Proof. The assertion follows directly from Theorem 4.5.3. $\square$

4.5.5 Remark. *The relation $\frac{\partial^2}{\partial r^2} F(t, T, r) \neq 0$ means that the realization is not affine, and $\frac{\partial}{\partial r} \left[\frac{\sigma_i(t,T,r)}{\sigma_i(t,t,r)} \right] \neq 0$ means that the volatility is not of the form $\sigma_i(t, T, r) = \varphi(t, T) \sigma_i(t, t, r)$ for some function $\varphi(t, T)$ which is independent of r.*

The statement of Corollary 4.5.4 is found in Jeffrey [44] for a single standard Wiener process, and in Hyll [40] for finitely many standard Wiener and finitely many counting processes, all mutually independent.

4.5.6 Remarks.

- *Theorem 4.5.3 yields a necessary criterion, in terms of the volatilities, for the existence of a generic benchmark realization, namely $\det \mathcal{J}_r \hat{\sigma}(t, T, r) \hat{\sigma}^{-1}(t, r) = 0$. This criterion is by no means sufficient. As we have seen in the previous sections, it is in general not even enough for a non-generic realization. For instance, all deterministic direction volatilities $\sigma(r, T, r) = \varphi(t, r) \lambda(t, T)$ satisfy $\frac{\partial}{\partial r} \left[\frac{\sigma(t,T,r)}{\sigma(t,t,r)} \right] = 0$. For details on this topic in the case of one driving Wiener process, we refer to Jeffrey [44].*
- *As mentioned at the end of Section 3.3, we conjecture that every term structure model admitting a finite dimensional realization must have an affine term structure. With regard to this conjecture it would be desirable to show that Theorem 4.5.3 also holds with $\mathcal{H}_r F(t, T, r) \neq 0$ instead of $\det \mathcal{H}_r F(t, T, r) \neq 0$ (this is only equivalent in the one-dimensional case). It would also be nice to relax the assumption $d \leq n$, i.e. that the number of benchmark forward rates is less than or equal to the number of driving processes.*

To understand the geometric picture one can think of the following hypothetical program:

1. Assume we have a model for the benchmark forward rates

$$dr_t = \mu(t, r_{t-})dt + \gamma(t, r_{t-})dX_t,$$

   and we observe the initial forward rate curve $\tilde{f}^*(0, \bullet)$ at the market. Choose $r^* := (\tilde{f}^*(0, x_1), \ldots, \tilde{f}^*(0, x_d)) \in \mathbb{R}^d$ as initial condition, that is the benchmark forward rates are specified by

$$\begin{cases} dr_t &= \mu(t, r_{t-})dt + \gamma(t, r_{t-})dX_t \\ r_0 &= r^* \end{cases}.$$

2. Solve the integro-differential equation of Proposition 4.2.4 in order to compute forward rates $F(t, T, r)$. This will give us a new initial curve $f^*(0, T) := F(0, T, r^*)$.

3. Derive volatilities $\sigma(t, T, r)$ from the forward rates.

4. Now take the HJM forward rate model

$$\begin{cases} df(t, T) &= \alpha(t, T, r_{t-})dt + \sigma(t, T, r_{t-})dX_t \\ f(0, T) &= \tilde{f}^*(0, T) \end{cases},$$

   where α is determined by the HJM drift condition (4.1.15) of Lemma 4.1.10.

The question is now whether the thus constructed forward rate model will have a benchmark realization. Obviously, if the initial forward rate curve is $\tilde{f}^*(0, \bullet) = f^*(0, \bullet)$, then a benchmark realization is given by the mapping F. If, however, the initial forward rate curve $\tilde{f}^*(0, \bullet)$ differs from $f^*(0, \bullet)$, then it is no longer clear that there exists a realization. What Theorem 4.5.3 says is that we can only fit one initial forward rate curve, if the Hessian matrix $\mathcal{H}_r F(t, T, r)$ is non-singular.

Chapter 5

Conclusion

We have studied the existence of finite dimensional realizations for term structure models driven by processes, which are allowed to make jumps, by approaching the problem from two directions. We have used the Lie algebraic methodology from Björk and Svensson [13], and adapted this framework to our setting, taking into account the possible occurrence of jumps. The other way was to consider the problem within the framework of benchmark realization, which enabled us to derive useful integro-differential equations.

The main insight of these investigations is the following. As for term structure models driven by Wiener processes, everything works fine provided the volatility is deterministic. If, however, the volatility is non-deterministic, new phenomena emerge, as soon as the driving processes have jumps. In particular, the occurrence of jumps severely limits the range of forward rate models having finite dimensional realizations. The intrinsic reason, and this comes out from both approaches, is that the cumulant generating functions of the driving processes are no polynomials, as soon as discontinuities appear, thus having an infinite dimensional structure. For deterministic direction volatility models, which are frequently considered in the literature, we have, for this reason, shown the existence of finite dimensional systems converging locally to the forward rate model.

In Chapter 3, we have started with the investigation of Banach space valued equations, and later applied the results to term structure models. We have shown that for Banach space valued equations the existence of finite dimensional realizations implies that the Lie algebra generated by the vector fields appearing in the equation is finite dimensional (Theorem 3.2.4). This may be regarded as one implication of the main theorem in Björk and Svensson [13] for our setting. It is clear, from a geometric point of view, that the converse is not valid as soon as the driving processes admit jumps.

A principal observation is that all known realizations for term structure models, this concerns those constructed in the literature (see, e.g., Björk and Landén [12]) as well as those appearing in this text, are affine. Indeed, it is known from works of Filipović and Teichmann that Wiener driven models that admit a finite dimensional realization, must necessarily have an affine term structure. This gives rise to the conjecture that for term structure models the converse of Theorem 3.2.4 is true, that is there exists a realization if and only if the relevant Lie algebra is finite dimensional. This conjecture, whose prove would require to go into the theory of convenient analysis (Filipović and Teichmann [33]), is left open in the text.

However, our necessary criterion has served reasonably well in order to exclude the existence of finite dimensional realizations, when the driving processes admit jumps. A technical detail that has not been settled is the exact determination of the relevant Lie algebra (or at least its dimension) for deterministic direction volatility models. Nevertheless, Lemma 3.5.1, which deals with closely related Lie algebras, suggested that the Lie algebra is, in general, infinite dimensional. This was supported by the later result Theorem 4.4.1. Moreover, Lemma 3.5.1 suggested that there exists, at least, a sequence of finite dimensional systems converging to the forward rate model. This was proven in Theorem 3.5.18, where the convergence was established in a local sense. We have assumed for this result that the driving processes have finite variation paths on compact intervals. Although this is a reasonable class of processes, it is of course desirable to extend Theorem 3.5.18 to a more general class of driving processes.

The benchmark realization approach in Chapter 4 has proven to be quite efficient, see Theorem 4.4.1 and Theorem 4.4.4. We have shown for certain a priori non-deterministic volatility structures that the existence of a finite dimensional realization implies that the volatility must be deterministic. In other words, the existence of a finite dimensional realization is excluded for those non-deterministic models as soon as the driving process admits jumps. Both results exhibit the restrictive nature of term structure models, which are driven by processes with jumps, concerning finite dimensional realizations.

We have also derived a structural result concerning the driving process, namely in Theorem 4.3.6 we have proven that for short rate realizations with deterministic volatility, the compensator of the jump measure must necessarily have an affine structure. For term structure models with deterministic volatility we have discovered connections to Küchler and Naumann [46], Gapeev and Küchler [35], and other papers that deal with the question when the short rate is a Markov process.

Finally, we have generalized the result of Jeffrey [44] that every generic short rate realization is affine. Admitting that the driving processes make jumps, we have proven that any generic benchmark realization must have a singular Hessian matrix. Nevertheless, it arises the question whether this result can be improved by stating that the Hessian matrix is zero rather than only being non-singular. This would give evidence that the conjecture formulated at the end of Section 3.3, namely that every term structure model, admitting a finite dimensional realization, must have an affine term structure, is true.

Bibliography

[1] D. Applebaum. *Lévy processes and stochastic calculus.* Cambridge University Press, Cambridge, 2004.

[2] O. E. Barndorff-Nielsen. Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society London*, Series A, Vol. 353: 401–419, 1977.

[3] O. E. Barndorff-Nielsen and O. Halgreen. Infinite divisibility of the hyperbolic and generalized inverse Gaussian distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 38:309–312, 1977.

[4] H. Bauer. *Wahrscheinlichkeitstheorie.* Walter de Gruyter, Berlin, 1990.

[5] J. Bertoin. *Lévy processes.* Cambridge University Press, Cambridge, 1996.

[6] R. Bhar and C. Chiarella. Transformation of Heath-Jarrow-Morton models to Markovian systems. *The European Journal of Finance*, 3:1–26, 1997.

[7] T. Björk. *Arbitrage theory in continuous time.* Oxford University Press, New York, 2004.

[8] T. Björk. On the geometry of interest rate models. In W. Runggaldier, editor, *Financial Mathematics.* Springer Lecture Notes in Mathematics, Springer Verlag, 2003.

[9] T. Björk. Interest rate theory. In W. Runggaldier, editor, *Financial Mathematics.* Springer Lecture Notes in Mathematics, Vol. 1656, Springer Verlag, 1997.

[10] T. Björk and B. Christensen. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9(4):323–348, 1999.

[11] T. Björk and A. Gombani. Minimal realizations of interest rate models. *Finance and Stochastics*, 3(4):413–432, 1999.

[12] T. Björk and C. Landén. On the construction of finite dimensional realizations for nonlinear forward rate models. *Finance and Stochastics*, 6(3):303–331, 2002.

[13] T. Björk and L. Svensson. On the existence of finite dimensional realizations for nonlinear forward rate models. *Mathematical Finance*, 11(2):205–243, 2001.

[14] T. Björk, G. Di Masi, Y. Kabanov, and W. Runggaldier. Towards a general theory of bond markets. *Finance and Stochastics*, 1(2):141–174, 1997.

[15] T. Björk, Y. Kabanov, and W. Runggaldier. Bond market structure in the presence of marked point processes. *Mathematical Finance*, 7(2):211–239, 1997.

[16] T. Björk, C. Landén, and L. Svensson. Finite dimensional Markovian realizations for stochastic volatility forward rate models. *Proceedings of the Royal Society London*, Series A, Vol. 460:53–83, 2004.

[17] A. Brace and M. Musiela. A multifactor Gauss Markov implementation of Heath, Jarrow, and Morton. *Mathematical Finance*, 4:259–283, 1994.

[18] P. Carr, E. Chang, and D. Madan. The variance gamma process and option pricing. *European Finance Review*, 2(1):79–105, 1998.

[19] P. Carr, H. Geman, D. Madan, and M. Yor. The fine structure of asset returns: an empirical investigation. *Journal of Business*, 75(2):305–332, 2002.

[20] A. Carverhill. When is the spot rate Markovian? *Mathematical Finance*, 4(4): 305–312, 1994.

[21] O. Cheyette. Markov representation of the Heath-Jarrow-Morton model. Preprint, BARRA Inc., 1996.

[22] C. Chiarella and O. K. Kwon. Forward rate dependent Markovian transformations of the Heath-Jarrow-Morton term structure model. *Finance and Stochastics*, 5 (2):237–257, 2001.

[23] R. Cont and P. Tankov. *Financial modelling with jump processes.* Chapman and Hall / CRC Press, London, 2003.

[24] J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–408, 1985.

[25] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6(4):379–406, 1996.

[26] E. Eberlein and U. Keller. Hyperbolic distributions in finance. *Bernoulli*, 1: 281–299, 1995.

[27] E. Eberlein and W. Kluge. Exact pricing formulae for caps and swaptions in a Lévy term structure model. Fdm preprint 86, University of Freiburg, 2004.

[28] E. Eberlein and W. Kluge. Valuation of floating range notes in Lévy term structure models. *Mathematical Finance*, 2005.

[29] E. Eberlein and S. Raible. Term structure models driven by general Lévy processes. *Mathematical Finance*, 9(1):31–53, 1999.

[30] E. Eberlein, J. Jacod, and S. Raible. Lévy term structure models: no-arbitrage and completeness. *Finance and Stochastics*, 9:67–88, 2005.

[31] J. Elstrodt. *Maß- und Integrationstheorie.* Springer, Berlin, 2002.

[32] D. Filipović and J. Teichmann. On finite dimensional term structure models. Working paper, 2002.

[33] D. Filipović and J. Teichmann. Existence of finite dimensional realizations for stochastic equations. *Journal of Functional Analysis*, 2001.

[34] D. Filipović and J. Teichmann. On the geometry of the term structure of interest rates. *Proceedings of the Royal Society London*, Series A, Vol. 460:129–167, 2004.

[35] P. Gapeev and U. Küchler. On Markovian short rates in term structure models driven by jump-diffusion processes. *Statistics and Decisions*, 99:901–919, 2004.

[36] 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(1):77–105, 1992.

[37] H. Heuser. *Lehrbuch der Analysis 2.* Teubner, Stuttgart, 1991.

[38] T. Ho and S. Lee. Term structure movements and pricing interest rate contingent claims. *Journal of Finance*, 41:1011–1029, 1986.

[39] J. Hull and A. White. Pricing interest-rate derivative securities. *Review of Financial Studies*, 3:573–592, 1990.

[40] M. Hyll. Affine term structures and short-rate realizations of forward rate models driven by jump-diffusion processes. In *Essays on the term structure of interest rates.* PhD thesis, Stockholm School of Economics, 2000.

[41] K. Inui and M. Kijima. A Markovian framework in multi-factor Heath-Jarrow-Morton models. *Journal of Financial and Quantitative Analysis*, 33(3):423–440, 1998.

[42] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes.* Springer, Berlin, 1987.

[43] R. Jarrow and D. Madan. Option pricing using the term structure of interest rates to hedge systematic discontinuities in asset returns. *Mathematical Finance*, 5(4):311–336, 1995.

[44] A. Jeffrey. Single factor Heath-Jarrow-Morton term structure models based on Markov spot interest rate dynamics. *Journal of Financial and Quantitative Analysis*, 30(4):619–642, 1995.

[45] J. Kallsen. *Semimartingale modelling in finance.* PhD thesis, University of Freiburg, 1998.

[46] U. Küchler and E. Naumann. Markovian short rates in a forward rate model with a general class of Lévy processes. Discussion paper 6, Sonderforschungsbereich 373, Humboldt University Berlin, 2003.

[47] E. Lukacs. *Characteristic functions.* Griffin, London, 1970.

[48] E. Lütkebohmert. Finite dimensional realizations for Heath, Jarrow and Morton type forward interest rate models with jumps. Preprint, University of Bonn, 2004.

[49] D. Madan. Purely discontinuous asset pricing processes. In E. Jouini, J. Cvitanič, and M. Musiela, editors, *Option Pricing, Interest Rates and Risk Management.* Cambridge University Press 105-153, Cambridge, 2001.

[50] D. Madan and F. Milne. Option pricing with V.G. martingale components. *Mathematical Finance*, 1(4):39–55, 1991.

[51] D. Madan and E. Seneta. The variance gamma model for share market returns. *Journal of Business*, 63(4):511–524, 1990.

[52] C. Mari. Single factor models with Markovian spot interest rate: an analytical treatment. *Decisions in Economics and Finance*, 26(1):39–52, 2003.

[53] M. Musiela. Stochastic PDE's and term structure models. Preprint, 1993.

[54] M. Musiela and M. Rutkowski. *Martingale methods in financial modelling.* Springer, Berlin, 1997.

[55] P. Protter. *Stochastic integration and differential equations.* Springer, Berlin, 1992.

[56] S. Raible. *Lévy processes in finance: theory, numerics, and empirical facts.* PhD thesis, University of Freiburg, 2000.

[57] R. Remmert. *Funktionentheorie 1.* Springer, Berlin, 1992.

[58] P. Ritchken and L. Sankarasubramanian. Volatility structures of forward rates and the dynamics of the term structure. *Mathematical Finance*, 5(1):55–72, 1995.

[59] K. Sato. *Lévy processes and infinitely divisible distributions.* Cambridge University Press, Cambridge, 1999.

[60] H. Shirakawa. Interest rate option pricing with Poisson-Gaussian forward rate curve processes. *Mathematical Finance*, 1(4):77–94, 1991.

[61] O. Vasiček. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

[62] D. Werner. *Funktionalanalysis.* Springer, Berlin, 2002.

Appendix A

Attached proofs

There are some results in this text, whose proofs are established by using standard techniques or making somewhat tedious computations. In fact, a few among these results are known in the literature in a slightly different context. The proofs, which we have omitted for this reason, are gathered in this appendix.

A.1 Proofs of Chapter 2

Some of the results of Section 2.2 still need to be proven. The proof of the bond price equation in Proposition 2.2.4 is similar to that of the third part of Björk, Kabanov, and Runggaldier [15, Prop. 2.4]. The main ingredients are the stochastic Fubini theorems, Itô's formula and the fact that a Grigelionis process has absolutely continuous characteristics.

Proof. (of Proposition 2.2.4) Define

$$F(t,T) := -\int_t^T f(t,s)ds.$$

From the forward rate equation (2.2.1) we obtain

$$F(t,T) = -\int_t^T f^*(0,s)ds - \int_t^T \int_0^t \alpha(v,s)dvds - \int_t^T \int_0^t \sigma(v,s)dX_v ds. \tag{A.1.1}$$

Note that

$$r_s = f^*(0,s) + \int_0^s \alpha(v,s)dv + \int_0^s \sigma(v,s)dX_v. \tag{A.1.2}$$

The following changes of the order of integration are justified by virtue of Assumptions 2.2.1 and the stochastic Fubini theorems in the appendix of Björk, Di Masi, Kabanov, and Runggaldier [14]. We proceed as follows. We change the order of integration $\int_t^T \int_0^t dv ds = \int_0^t \int_t^T ds dv$ in (A.1.1), afterwards we split the integrals $\int_t^T = \int_v^T - \int_v^t$, and then change the order of integration $\int_0^t \int_v^t ds dv = \int_0^t \int_0^s dv ds$. Incorporating the short rate equation (A.1.2) in the then derived identity yields

$$F(t,T) = F(0,T) + \int_0^t r_s ds - \int_0^t \int_v^T \alpha(v,s) ds dv - \int_0^t \int_v^T \sigma(v,s) ds dX_v.$$

Using the canonical decomposition $X_t = X_0 + X_t^c + x * (\mu^X - \nu)_t + \int_0^t \beta_s ds$, we obtain the dynamics

$$\begin{aligned} dF(t,T) &= (r_t + A(t,T) + \langle \beta_t, \Sigma(t,T) \rangle)\, dt \\ &\quad + \Sigma(t,T) dX_t^c + \int_{\mathbb{R}^n} \langle x, \Sigma(t,T) \rangle (\mu^X - \nu)(dt,dx). \end{aligned} \tag{A.1.3}$$

Now we apply Itô's formula Jacod and Shiryaev [42, Thm. I.4.57] on $p(t,T) = e^{F(t,T)}$ for fixed $T \in \mathbb{R}_+$, and obtain

$$\begin{aligned} p(t,T) &= p(0,T) + \int_0^t p(s-,T) dF(s,T) + \frac{1}{2} \int_0^t p(s-,T) d\langle F(\bullet,T)^c, F(\bullet,T)^c \rangle_t \\ &\quad + \sum_{0<s\le t} \Big[e^{F(s,T)} - e^{F(s-,T)} - e^{F(s-,T)} \Delta F(s,T) \Big]. \end{aligned} \tag{A.1.4}$$

Inserting the dynamics (A.1.3) for $F(t,T)$ as well as

$$\begin{aligned} d\langle F(\bullet,T)^c, F(\bullet,T)^c \rangle_t &= \langle \Sigma(t,T), c_t \Sigma(t,T) \rangle dt, \\ \Delta F(t,T) &= \langle \Delta X_t, \Sigma(t,T) \rangle \end{aligned}$$

into (A.1.4), we obtain the claimed dynamics (2.2.6) for the bond prices. □

The equation for the discounted bond prices is an immediate consequence.

Proof. (of Corollary 2.2.5) The discounted bond prices are given by $z(t,T) = B_t p(t,T)$, where $B_t := \exp(-\int_0^t r_s ds)$. Since B has continuous paths, the quadratic co-variation of B_t and $p(t,T)$ vanishes Jacod and Shiryaev [42, Prop. I.4.49.d], and we get

$$dz(t,T) = B_{t-} dp(t,T) + p(t-,T) dB_t$$

Incorporating $dB_t = -r_t B_t dt$ and the bond price dynamics from Proposition 2.2.4 yields the asserted equation (2.2.7) for the discounted bond prices. □

Proof. (of Lemma 2.2.10) The left-continuity of $t \mapsto f(t)g(t)$ yields that $f(t)g(t) = 0$ for all $t \in (0,\infty)$, because, assuming $f(t)g(t) \neq 0$ for some $t \in (0,\infty)$ leads to the contradiction $\int_u^t f(s)g(s) ds \neq 0$ for some $u \in (0,t)$. The assertion follows by virtue of the assumption $g > 0$. □

Now, we prove the Musiela parametrization. Our proof is essentially a copy of the proof of the second part of Björk, Kabanov, and Runggaldier [15, Prop. 2.4], where an analogous equation for the short rate is derived, see also Appendix 1 in Jeffrey [44].

Proof. (of Proposition 2.2.12) All changes of order of integration, and the interchangings of differentiation and integration, performed in this proof, are valid due to Assumptions 2.2.1 and the stochastic Fubini theorems in the appendix of Björk, Di Masi, Kabanov, and Runggaldier [14]. Fix $x \in \mathbb{R}_+$. From the forward rate dynamics (2.2.1) we obtain

$$r_t(x) - r_0(x) = r_0(t+x) - r_0(x) + \int_0^t \alpha(s, t+x)ds + \sum_{i=1}^n \int_0^t \sigma_i(s, t+x)dX_s^i. \tag{A.1.5}$$

Let us consider the terms in (A.1.5) in detail. First, we observe that

$$r_0(t+x) - r_0(x) = \int_0^t \frac{\partial}{\partial x} r_0(s+x)dx.$$

For the stochastic integral term, we get

$$\begin{aligned}
\sum_{i=1}^n \int_0^t \sigma_i(s, t+x)dX_s^i &= \sum_{i=1}^n \int_0^t \left(\int_s^t \frac{\partial}{\partial v}\sigma_i(s, v+x)dv + \sigma_i(s, s+x) \right) dX_s^i \\
&= \int_0^t \left(\sum_{i=1}^n \int_0^t \frac{\partial}{\partial x}\sigma_i(s, v+x)dX_s^i \right) dv + \sum_{i=1}^n \int_0^t \sigma_i(s, s+x)dX_s^i.
\end{aligned}$$

Analogously, the identity

$$\int_0^t \alpha(s, t+x)ds = \int_0^t \left(\int_0^t \frac{\partial}{\partial x}\alpha(s, v+x)ds \right) dv + \int_0^t \alpha(s, s+x)ds$$

is valid. Inserting the three previous equations into (A.1.5), we obtain

$$\begin{aligned}
dr_t(x) &= \left[\frac{\partial}{\partial x} \left(r_0(t+x) + \int_0^t \alpha(s, t+x)ds + \int_0^t \sigma(s, t+x)dX_s \right) \right] dt \\
&\quad + \alpha(t, t+x)dt + \sigma(t, t+x)dX_t \\
&= \left[\frac{\partial}{\partial x} r_t(x) + \alpha(t, t+x) \right] dt + \sigma(t, t+x)dX_t,
\end{aligned}$$

as we have stated. The initial condition $r_0(x) = f^*(0, x)$ follows from the definition of the $r_t(x)$. $\square$

A.2 Proofs of Chapter 3

In Section 3.1, we have formulated a few auxiliary results about supports.

Proof. (of Lemma 3.1.4)

1. Assume, it holds $\mathbb{P}(A) = 1$, where $A \in \mathfrak{F}$ is the event $A = \{f \circ X = g \circ X\}$. By hypothesis $f(x) = g(x)$ for all $x \in X(A)$, and therefore, due to the continuity of f and g, the identity $f(x) = g(x)$ is valid for all $x \in \overline{X(A)}$. The support $\mathrm{supp}(X)$ is the smallest closed set $F \subset \mathbb{R}^n$ satisfying $\mathbb{P}(X \in F) = 1$. Thus, the relation $\mathbb{P}(X \in \overline{X(A)}) \geq \mathbb{P}(A) = 1$ implies that $\mathrm{supp}(X) \subset \overline{X(A)}$, whence $f(x) = g(x)$ for all $x \in \mathrm{supp}(X)$.

2. Assume $f(x) = g(x)$ for all $x \in \mathrm{supp}(X)$. Then we obtain $\mathbb{P}(f \circ X = g \circ X) \geq \mathbb{P}(X \in \mathrm{supp}(X)) = 1$, i.e. $f \circ X = g \circ X$ ($\mathbb{P} - a.s.$)

□

Proof. (of Lemma 3.1.5) Assuming $f(x) \neq 0$ for some $x \in \mathrm{supp}(\mu)$ implies, by the continuity of f, that $f(y) \neq 0$ for all y from a neighborhood $U(x) \subset \mathbb{R}^n$ of x. The support of μ is the set of all $x \in \mathbb{R}^n$ such that $\mu(U) > 0$ for all open sets $U \subset \mathbb{R}^n$ with $x \in U$. Consequently, it holds $\mu(U(x)) > 0$, which contradicts $f = 0$ ($\mu - a.s.$). □

Proof. (of Lemma 3.1.6) Set $W := W_1 - W_2$, which is again of class $C(\mathbb{R}_+ \times \mathbb{R}^d)$. By hypothesis, there exists a $\mathbb{P}$-null set A such that the paths $t \mapsto Z_{t-}(\omega)$ are left-continuous for each $\omega \in \Omega \setminus A$ and

$$\int_0^t W(s, Z_{s-}(\omega))ds = 0, \quad (\omega, t) \in \Omega \setminus A \times \mathbb{R}_+.$$

For each $\omega \in \Omega \setminus A$, the left-continuity of the trajectory $t \mapsto W(t, Z_{t-}(\omega))$ yields, by virtue of Lemma 2.2.10, that $W(t, Z_{t-}(\omega)) = 0$ for all $t \in (0, \infty)$. Applying, for each $t \in (0, \infty)$, Lemma 3.1.4 on the mapping $\omega \mapsto W(t, Z_{t-}(\omega))$, the proof is done. □

Next, we give the omitted proofs of the results that deal with Fisk-Stratonovich integration.

Proof. (of Lemma 3.1.20) For each $i, j = 1, \ldots, n$ the continuous martingale part of $\gamma_{ij}(t, Z_t)$ is, by Itô's formula Jacod and Shiryaev [42, Thm. I.4.57], given by

$$\gamma_{ij}(t, Z_t)^c = \sum_{l=1}^d \int_0^t \frac{\partial}{\partial z_l} \gamma_{ij}(s, Z_{s-}) dZ_s^{l,c}.$$

Since $Z_t^c = \int_0^t \gamma(s, Z_{s-})dX_s^c$, which follows from the dynamics (3.1.27), we deduce from the associativity of the Itô integral Jacod and Shiryaev [42, I.4.37]

$$\gamma_{ij}(t, Z_t)^c = \sum_{l=1}^{d} \int_0^t \left[\frac{\partial}{\partial z_l}\gamma_{ij}(s, Z_{s-})\right] \gamma_{l\bullet}(s, Z_{s-})dX_s^c.$$

Thus, we can compute $\langle \gamma_{i\bullet}(\hat{Z})^c, X^c\rangle_t$ by using Jacod and Shiryaev [42, Thm. I.4.40.d], and obtain the asserted identity. □

For the proof of Theorem 3.1.21, we need the associativity of Fisk-Stratonovich integrals. Since this result was not readily available in the literature Protter [55], we first derive this fact before proving the theorem.

A.2.1 Lemma. *For one-dimensional semimartingales X, Y, Z it holds*

$$X_- \circ (Y_- \circ Z) = (XY)_- \circ Z.$$

Proof. By the definition of the Fisk-Stratonovich integral it holds

$$\begin{aligned} X_- \circ (Y_- \circ Z) &= X_- \circ (Y_- \bullet Z + \tfrac{1}{2}\langle Y^c, Z^c\rangle) \\ &= X_- \bullet (Y_- \bullet Z + \tfrac{1}{2}\langle Y^c, Z^c\rangle) + \tfrac{1}{2}\langle X^c, (Y_- \bullet Z + \tfrac{1}{2}\langle Y^c, Z^c\rangle)^c\rangle \\ &= X_- \bullet (Y_- \bullet Z) + \tfrac{1}{2} X_- \bullet \langle Y^c, Z^c\rangle + \tfrac{1}{2}\langle X^c, Y_- \bullet Z^c\rangle. \end{aligned} \tag{A.2.1}$$

By the associativity of the Itô integral (see I.4.37 in Jacod and Shiryaev [42]) we obtain

$$X_- \bullet (Y_- \bullet Z) = (XY)_- \bullet Z. \tag{A.2.2}$$

Moreover, according to Jacod and Shiryaev [42, Thm. I.4.40.d] it holds

$$X_- \bullet \langle Y^c, Z^c\rangle + \langle X^c, Y_- \bullet Z^c\rangle = \langle (X_- \bullet Y + Y_- \bullet X)^c, Z^c\rangle. \tag{A.2.3}$$

Since $[X, Y] \in \mathcal{V}$ Jacod and Shiryaev [42, Thm. I.4.47.b], where $[X, Y]$ denotes the quadratic co-variation

$$[X, Y] := XY - X_0Y_0 - X_- \bullet Y - Y_- \bullet X,$$

we can write equation (A.2.3) as

$$X_- \bullet \langle Y^c, Z^c\rangle + \langle X^c, Y_- \bullet Z^c\rangle = \langle (XY)^c, Z^c\rangle. \tag{A.2.4}$$

The asserted identity follows by inserting (A.2.2) and (A.2.4) into (A.2.1). □

Proof. (of Theorem 3.1.21) By Itô's formula for Fisk-Stratonovich integrals Protter [55, Thm. V.21] we obtain

$$f(t,Z_t) = f(0,Z_0) + \int_0^t \frac{\partial}{\partial s} f(s,Z_{s-})ds + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial z_i} f(s,Z_{s-}) \circ dZ_s^i \\ + \sum_{0<s\le t} \Big[f(s,Z_s) - f(s,Z_{s-}) - \sum_{i=1}^d \frac{\partial}{\partial z_i} f(s,Z_{s-})\Delta Z_s^i \Big], \tag{A.2.5}$$

where the last term is in $\mathcal{V}$. Taking into account the jumps $\Delta Z_t = \gamma(t,Z_{t-})\Delta X_t$, the dynamics (3.1.28) of Z and the associativity of the Fisk-Stratonovich integral (Lemma A.2.1), we write equation (A.2.5) as

$$df(t,Z_t) = \frac{\partial}{\partial t} f(t,Z_{t-})dt + \langle \nabla_z f(t,Z_{t-}), \mu(t,Z_{t-})\rangle dt + \nabla_z f(t,Z_{t-})\gamma(t,Z_{t-}) \circ dX_t \\ + \int_{\mathbb{R}^n} \Big(f(t, Z_{t-} + \gamma(t,Z_{t-})x) - f(t,Z_{t-}) \\ - \langle \nabla_z f(t,Z_{t-}), \gamma(t,Z_{t-})x\rangle \Big) \mu^X(dt,dx), \tag{A.2.6}$$

where the $\mu^X(dt,dx)$-integral is in $\mathcal{V}$. Recall that $\hat{Z}$ is defined by $\hat{Z}_t(\omega) := (t, Z_t(\omega))$. The process $\langle \sigma(\hat{Z})^c, X^c\rangle$ belongs to $\mathcal{V}$ and is predictable (in fact, even continuous). Thus, (3.1.29) provides a decomposition of the semimartingale $f(t,Z_t)$ where the finite variation part is predictable. Consequently, $f(t,Z_t)$ is a special semimartingale. According to Prop. I.4.23 and Lemma I.3.10 in Jacod and Shiryaev [42], the $\mu^X(dt,dx)$-integral in (A.2.6) belongs to $\mathcal{A}_{\text{loc}}$. Therefore, we may integrate with respect to $(\mu^X - \nu)(dt,dx)$ plus $\nu(dt,dx)$ Jacod and Shiryaev [42, Prop. II.1.28], and obtain

$$df(t,Z_t) = \Big[\frac{\partial}{\partial t} f(t,Z_{t-}) + \langle \nabla_z f(t,Z_{t-}), \mu(t,Z_{t-}) + \gamma(t,Z_{t-})\beta(t,Z_{t-})\rangle \\ + \int \Big(f(t,Z_{t-} + \gamma_d(t,Z_{t-})x) - f(t,Z_{t-}) \\ - \langle \nabla_z f(t,Z_{t-}), \gamma_d(t,Z_{t-})x\rangle \Big) K^d_{t,Z_{t-}}(dx) \Big] dt \\ + \nabla_z f(t,Z_{t-})\gamma_c(t,Z_{t-}) \circ d\mathcal{C}(X)_t \\ + \Big(f(t,Z_{t-} + \gamma_d(t,Z_{t-})x) - f(t,Z_{t-}) \Big)(\mu^{\mathcal{D}(X)} - \nu^{\mathcal{D}(X)})(dt,dx). \tag{A.2.7}$$

According to Jacod and Shiryaev [42, Cor. II.2.38] the functions $W_i(\omega,t,x) = x_i$ belong to $G_{\text{loc}}(\mu^X)$. Applying Jacod and Shiryaev [42, Prop. II.1.30.b] the dynamics (3.1.29) of $f(t,Z_t)$ can be expressed as

$$df(t,Z_t) = (\alpha(t,Z_{t-}) + \langle \beta(t,Z_{t-}), \sigma(t,Z_{t-})\rangle)dt + \sigma_c(t,Z_{t-}) \circ d\mathcal{C}(X)_t \\ + \langle x, \sigma_d(t,Z_{t-})\rangle(\mu^{\mathcal{D}(X)} - \nu^{\mathcal{D}(X)})(dt,dx). \tag{A.2.8}$$

Since the continuous local martingale part, the purely discontinuous local martingale part and the finite variation part of a special semimartingale are unique (which follows from Cor. I.3.16 and Thm. I.4.18 in Jacod and Shiryaev [42]), we obtain from equations (A.2.7) and (A.2.8)

$$\begin{aligned}&\int_0^t \alpha(s, Z_{s-})ds + \int_0^t \langle \beta(s, Z_{s-}), \sigma(s, Z_{s-}) - \nabla_z f(s, Z_{s-})\gamma(s, Z_{s-})\rangle ds \\ &+ \frac{1}{2}\langle \sigma(\hat{Z})^c, X^c\rangle_t = \int_0^t \left(\frac{\partial}{\partial s} f(s, Z_{s-}) + \langle \nabla_z f(s, Z_{s-}), \mu(s, Z_{s-})\rangle\right) ds \\ &\quad + \int_0^t \int \Big(f(s, Z_{s-} + \gamma_d(s, Z_{s-})x) - f(s, Z_{s-}) \\ &\qquad - \langle \nabla_z f(s, Z_{s-}), \gamma_d(s, Z_{s-})x\rangle \Big) K^d_{s,Z_{s-}}(dx)ds + \frac{1}{2}\langle (\nabla_z f(\hat{Z})\gamma(\hat{Z}))^c, X^c\rangle_t,\end{aligned} \tag{A.2.9}$$

as well as

$$\int_0^t \sigma_c(s, Z_{s-})d\mathcal{C}(X)_s = \int_0^t \nabla_z f(s, Z_{s-})\gamma_c(s, Z_{s-})d\mathcal{C}(X)_s, \tag{A.2.10}$$

$$\begin{aligned}&\int_0^t \int \langle x, \sigma_d(s, Z_{s-})\rangle (\mu^{\mathcal{D}(X)} - \nu^{\mathcal{D}(X)})(ds, dx) \\ &= \int_0^t \int \Big(f(s, Z_{s-} + \gamma_d(s, Z_{s-})x) - f(s, Z_{s-})\Big)(\mu^{\mathcal{D}(X)} - \nu^{\mathcal{D}(X)})(ds, dx).\end{aligned} \tag{A.2.11}$$

Applying Lemma 3.1.11 on equation (A.2.10) we obtain

$$\sigma_c(t, z) = \nabla_z f(t, z)\gamma_c(t, z), \quad t \in (0, \infty), z \in \operatorname{supp}(Z_{t-}), \tag{A.2.12}$$

and an application of Lemma 3.1.13 on equation (A.2.11) yields the identity (3.1.30).

Provided, X is non-degenerate, we obtain, arguing exactly as in the proof of Theorem 3.1.17,

$$\sigma_d(t, z) = \nabla_z f(t, z)\gamma_d(t, z), \quad t \in (0, \infty), z \in \operatorname{supp}(Z_{t-})$$

from which, together with equation (A.2.12), we conclude the identity (3.1.32) for $\sigma(t, z)$. Applying Lemma 3.1.4 on equations (3.1.30) and (3.1.32) we obtain that, for all $t \in (0, \infty)$ and $x \in \operatorname{supp}(K^d_t)$, it holds

$$\begin{aligned}\langle \sigma_d(t, Z_{t-}), x\rangle &= f(t, Z_{t-} + \gamma_d(t, Z_{t-})x) - f(t, Z_{t-}) \quad \mathbb{P} - a.s. \\ \sigma(t, Z_{t-}) &= \nabla_z f(t, Z_{t-})\gamma(t, Z_{t-}) \quad \mathbb{P} - a.s.\end{aligned}$$

By the continuity assumptions on f and the coefficients, and the right-continuity of Z (notice also the second point in Definition 3.1.10), we obtain, up to a $\mathbb{P}$-null set,

$$\begin{aligned}\langle \sigma_d(t, Z_{t-}), x\rangle &= f(t, Z_{t-} + \gamma_d(t, Z_{t-})x) - f(t, Z_{t-}), \quad t \in (0, \infty), x \in \operatorname{supp}(K^d_t) \\ \sigma(t, Z_{t-}) &= \nabla_z f(t, Z_{t-})\gamma(t, Z_{t-}), \quad t \in (0, \infty) \\ \langle \sigma(\hat{Z})^c, X^c\rangle_t &= \langle (\nabla_z f(\hat{Z})\gamma(\hat{Z}))^c, X^c\rangle_t, \quad t \in (0, \infty).\end{aligned}$$

Hence, equation (A.2.9) simplifies to

$$\int_0^t \alpha(s, Z_{s-})ds = \int_0^t \left(\frac{\partial}{\partial s} f(s, Z_{s-}) + \langle \nabla_z f(s, Z_{s-}), \mu(s, Z_{s-}) \rangle \right) ds.$$

Applying Lemma 3.1.6, we obtain the stated formula (3.1.31) for $\alpha(t, z)$.

If the assumptions from the third part of the theorem hold, we obtain equations (3.1.33) and (3.1.34) by an argumentation analogous to that of the proof of Theorem 3.1.17. □

Proof. (of Lemma 3.2.3) By assumption, there is a $d+1$-dimensional realization (G, Z). For each $x \in \mathbb{R}_+$ and $j = 1, \ldots, n$, the continuous martingale part of the process $\sigma_j(t, r_t)(x)$ is, according to Itô's formula Jacod and Shiryaev [42, Thm. I.4.57], given by

$$\begin{aligned} \sigma_j(t, r_t)(x)^c &= \sigma_j(t, G(t, Z_t))(x)^c \\ &= \sum_{i=1}^d \int_0^t \frac{\partial}{\partial z_i} \Big(\sigma_j(s, G(s, Z_{s-}))(x) \Big) dZ_s^{i,c}. \end{aligned}$$

Because of the dynamics of Z

$$dZ_t = \mu(t, Z_{t-})dt + \gamma(t, Z_{t-})dX_t,$$

and the associativity of the Itô integral Jacod and Shiryaev [42, I.4.37], we obtain

$$\sigma_j(t, r_t)(x)^c = \sum_{i=1}^d \sum_{l=1}^n \int_0^t \frac{\partial}{\partial z_i} \Big(\sigma_j(s, G(s, Z_{s-}))(x) \Big) \gamma_{il}(s, Z_{s-}) dX_s^{l,c} \tag{A.2.13}$$

Define the mappings $\hat{G} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}_+ \times \mathcal{X}$ by $\hat{G}(t, z) := (t, G(t, z))$, and $\sigma_j^x : \mathbb{R}_+ \times \mathcal{X} \to \mathbb{R}$ by $\sigma_j^x(t, r) := \sigma_j(t, r)(x)$. Observe that for all $(t, z) \in \mathbb{R}_+ \times \mathbb{R}^d$ and $l = 1, \ldots, n$

$$\begin{aligned} \sum_{i=1}^d \frac{\partial}{\partial z_i} \Big(\sigma_j(t, G(t, z))(x) \Big) \gamma_{il}(t, z) &= \mathbb{F}(\sigma_j^x \circ \hat{G})(t, z)[0, \gamma_{\bullet l}(t, z)] \\ &= \mathbb{F}\sigma_j^x(\hat{G}(t, z)) \circ \mathbb{F}\hat{G}(t, z)[0, \gamma_{\bullet l}(t, z)]. \end{aligned} \tag{A.2.14}$$

Let us determine the Fréchet derivative $\mathbb{F}\hat{G}(t, z)[0, \gamma_{\bullet l}(t, z)]$. It holds for each $x \in \mathbb{R}_+$

$$\begin{aligned} \mathbb{F}\hat{G}(t, z)[0, \gamma_{\bullet l}(t, z)](x) &= \left(0, \lim_{h \to 0} \frac{G(t, z + h\gamma_{\bullet l}(t, z))(x) - G(t, z)(x)}{h} \right) \\ &= (0, \langle \gamma_{\bullet l}(t, z), \nabla_z G(t, z)(x) \rangle). \end{aligned}$$

Applying Theorem 3.1.17 on the dynamics (note that the coefficients are continuous and $G(\bullet)(x)$ of class C^2, because convergence in the Banach space $\mathcal{X}$ implies pointwise convergence)

$$\begin{aligned} dZ_t &= \mu(t, Z_{t-})dt + \gamma(t, Z_{t-})dX_t, \\ dG(t, Z_t)(x) &= \alpha(t, G(t, Z_{t-}))(x)dt + \sigma(t, G(t, Z_{t-}))(x)dX_t, \end{aligned}$$

we obtain for all $t \in (0, \infty)$ and $z \in \operatorname{supp}(Z_{t-})$

$$\sigma_l(\hat{G}(t, z))(x) = \langle \gamma_{\bullet l}(t, z), \nabla_z G(t, z)(x) \rangle.$$

Applying Lemma 3.1.4 on this equation we obtain that, for all $t \in (0, \infty)$, it holds $\sigma_l(\hat{G}(t, Z_t))(x) = \langle \gamma_{\bullet l}(t, Z_t), \nabla_z G(t, Z_t)(x) \rangle$ ($\mathbb{P} - a.s.$) Since the trajectories of Z are right-continuous, we obtain, up to an evanescent set,

$$\mathbb{F}\hat{G}(t, Z_t)[0, \gamma_{\bullet l}(t, Z_t)](x) = (0, \sigma_l(\hat{G}(t, Z_t))(x)). \tag{A.2.15}$$

Taking together (A.2.13), (A.2.14) and (A.2.15), we get that the continuous martingale part of $\sigma_j(t, r_t)(x)$ equals

$$\sigma_j(t, r_t)(x)^c = \sum_{l=1}^{n} \int_0^t \mathbb{F}_r \sigma_j(s, r_{s-})[\sigma_l(s, r_{s-})](x) dX_s^{l,c}.$$

This identity and Jacod and Shiryaev [42, Thm. I.4.40] imply the stated equation. □

In the text, we have omitted the proofs of those results, where Lie algebras are computed.

Proof. (of Lemma 3.4.1) The vector fields $\hat{\alpha}$ and $\hat{\sigma}_1, \ldots, \hat{\sigma}_n$ have the Fréchet derivatives

$$\begin{aligned} \mathbb{F}\hat{\alpha}(t_0, r_0)[t, r] &= \mathbb{F}_r\hat{\alpha}(t_0, r_0)[r] + \mathbb{F}_t\hat{\alpha}(t_0, r_0)[t] \\ &= \left(0, \frac{\partial}{\partial x} r + t \frac{\partial}{\partial t} D(t_0)\right), \qquad \text{(A.2.16)} \\ \mathbb{F}\hat{\sigma}_i(t_0, r_0)[t, r] &= \mathbb{F}_r\hat{\sigma}_i(t_0, r_0)[r] + \mathbb{F}_t\hat{\sigma}_i(t_0, r_0)[t] \\ &= \left(0, t \frac{\partial}{\partial t} \sigma_i(t_0)\right) \quad \text{for } i = 1, \ldots, n. \qquad \text{(A.2.17)} \end{aligned}$$

According to (A.2.17) all Lie brackets $[\hat{\sigma}_i, \hat{\sigma}_j]$ are zero, and by (A.2.16) the brackets $[\hat{\sigma}_i, \hat{\alpha}]$ are given by

$$[\hat{\sigma}_i, \hat{\alpha}](t, r) = \left(0, \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) \sigma_i(t)\right), \quad i = 1, \ldots, n.$$

Their Fréchet derivatives are

$$\mathbb{F}[\hat{\sigma}_i, \hat{\alpha}](t_0, r_0)[t, r] = \left(0, t \frac{\partial}{\partial t} \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right) \sigma_i(t_0)\right), \quad i = 1, \ldots, n.$$

Thus, we obtain the Lie brackets

$$[[\hat{\sigma}_i, \hat{\alpha}], \hat{\alpha}](t,r) = \left(0, \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)^2 \sigma_i(t)\right), \quad i = 1, \ldots, n.$$

Proceeding in this fashion, we see that all relevant brackets are either zero or of the form $(0, (\frac{\partial}{\partial t} - \frac{\partial}{\partial x})^j \sigma_i)$ with $i \in \{1, \ldots, n\}$ and $j \in \mathbb{N}_0$. □

In order to establish the next results that deal with the computation of Lie algebras, we cite an auxiliary result from Björk and Svensson [13].

A.2.2 Lemma. *Take the vector fields $f_1, \ldots, f_k$ as given. The Lie algebra $\{f_1, \ldots, f_k\}_{LA}$ remains unchanged under the following operations.*

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

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

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

Proof. This is Björk and Svensson [13, Lemma 3.1]. □

Now, we can continue our Lie algebraic computations.

Proof. (of Lemma 3.5.1) Defining the vector field $\hat{\lambda}$ on $(0,\infty) \times \mathcal{H}$ as $\hat{\lambda}(t,r) := (0, \tilde{\lambda}(t))$, we can express the fields $\hat{\alpha}_m$ and $\hat{\sigma}$ as

$$\begin{aligned}\hat{\alpha}_m(t,r) &= \left(1, \frac{\partial}{\partial x} r - \varphi(t,r)\tilde{\lambda}(t)\frac{\partial}{\partial z}\Psi_m(t, -\varphi(t,r)\tilde{D}(t))\right) \\ &\quad - \frac{1}{2}c(t)\varphi(t,r)\mathbb{F}_r\varphi(t,r)[\tilde{\lambda}(t)]\hat{\lambda}(t), \quad m \in \mathbb{N}, \\ \hat{\sigma}(t,r) &= \varphi(t,r)\hat{\lambda}(t).\end{aligned}$$

Fix $m \in \mathbb{N}$ and set $n := |J_m|$. Since it is assumed $\varphi(t,r) \neq 0$ for all $(t,r) \in (0,\infty) \times \mathcal{H}$, we may apply Lemma A.2.2 on the vector fields $\hat{\alpha}_m, \hat{\sigma}$, and obtain

$$\{\hat{\alpha}_m, \hat{\sigma}\}_{LA} = \{\hat{f}, \hat{\lambda}\}_{LA}, \tag{A.2.18}$$

where $\hat{f}$ is the vector field on $(0,\infty) \times \mathcal{H}$ defined as

$$\begin{aligned}\hat{f}(t,r) &:= \left(1, \frac{\partial}{\partial x} r - \varphi(t,r)\tilde{\lambda}(t)\frac{\partial}{\partial z}\Psi_m(t, -\varphi(t,r)\tilde{D}(t))\right) \\ &= \left(1, \frac{\partial}{\partial x} r\right) + \left(0, \sum_{j \in J_m} (-1)^{j+1} a_j(t) \varphi^{j+1}(t,r)(\tilde{\lambda}\tilde{D}^j)(t)\right). \end{aligned} \tag{A.2.19}$$

Define inductively the brackets $\hat{g}_i$, $i \in \mathbb{N}$ by

$$\hat{g}_i := \begin{cases} [\hat{\lambda}, [\hat{\lambda}, \hat{f}]], & i = 1 \\ [\hat{\lambda}, \hat{g}_{i-1}], & i \geq 2 \end{cases}.$$

We see from (A.2.18) that the relevant Lie algebra equals

$$\{\hat{\alpha}_m, \hat{\sigma}\}_{LA} = \{\hat{f}, \hat{\lambda}, [\hat{\lambda}, \hat{f}], \hat{g}_1, \ldots, \hat{g}_n\}_{LA}. \tag{A.2.20}$$

We go on to compute the brackets appearing in (A.2.20). We only need the Fréchet derivative of $\hat{f}$ with respect to r for the computation of the Lie bracket $[\hat{\lambda}, \hat{f}]$, because $\hat{\lambda}$ has t-component zero. It is given by

$$\mathbb{F}_r \hat{f}(t_0, r_0)[r] = \left(0, \frac{\partial}{\partial x} r + \sum_{j \in J_m} (-1)^{j+1} a_j(t_0) \mathbb{F}_r(\varphi^{j+1}(t_0, r_0))[r](\tilde{\lambda}\tilde{D}^j)(t_0)\right).$$

The derivative of $\hat{\lambda}$ is $\mathbb{F}\hat{\lambda}(t_0, r_0)[t, r] = (0, t\frac{\partial}{\partial t}\tilde{\lambda}(t_0))$. We obtain the Lie bracket

$$[\hat{\lambda}, \hat{f}](t, r) = \left(0, \left(\frac{\partial}{\partial t} - \frac{\partial}{\partial x}\right)\tilde{\lambda}(t) - \sum_{j \in J_m} (-1)^{j+1} a_j(t) \mathbb{F}_r(\varphi^{j+1}(t, r))[\tilde{\lambda}(t)](\tilde{\lambda}\tilde{D}^j)(t)\right). \tag{A.2.21}$$

Similarly, we find out that for each $i \in \mathbb{N}$

$$\hat{g}_i(t, r) = \left(0, (-1)^{i+1} \sum_{j \in J_m} (-1)^{j+1} a_j(t) \mathbb{F}_r^{i+1}(\varphi^{j+1}(t, r))[\tilde{\lambda}(t); \ldots; \tilde{\lambda}(t)](\tilde{\lambda}\tilde{D}^j)(t)\right).$$

Recall that $n = |J_m|$. For each $(t, r) \in (0, \infty) \times \mathcal{H}$ define the $n \times n$ matrix $A(t, r)$ by

$$A_{ij}(t, r) := (-1)^{i+j} a_j(t) \mathbb{F}_r^{i+1}(\varphi^{j+1}(t, r))[\tilde{\lambda}(t); \ldots; \tilde{\lambda}(t)], \quad i = 1, \ldots, n, \; j \in J_m.$$

Denoting by $\hat{h}$ the column vector with entries $(0, \tilde{\lambda}\tilde{D}^j)$, $j \in J_m$, and by $\hat{g}$ the column vector consisting of $\hat{g}_1, \ldots, \hat{g}_n$, the relation

$$A\hat{h} = \hat{g} \tag{A.2.22}$$

is valid. We proceed with the Gaussian algorithm in equation (A.2.22), which replaces A by the $n \times n$-unit matrix, because A is non-singular by hypothesis. The mutual exchange of two rows in A changes the order of $\hat{g}$, multiplication of a row in A with a non-zero real number leads to a multiplication of one of the vector fields $\hat{g}_i$ with a smooth non-zero scalar field, and adding the multiple of a row to another row in A implies that one vector field $\hat{g}_i$ is replaced by $\hat{g}_i + \rho \hat{g}_j$, where ρ is a smooth scalar field. Consequently, each step of the Gaussian algorithm leads to a replacement of the vector fields of $\hat{g}_1, \ldots, \hat{g}_n$, which is valid by Lemma A.2.2. So, we get from (A.2.20)

$$\{\hat{\alpha}_m, \hat{\sigma}\}_{LA} = \{\hat{f}, \hat{\lambda}, [\hat{\lambda}, \hat{f}], \hat{\lambda}\tilde{D}^j \,|\, j \in J_m\}_{LA}.$$

A further application of Lemma A.2.2, where we notice the representations (A.2.19) of $\hat{f}$ and (A.2.21) of $[\hat{\lambda}, \hat{f}]$, simplifies the Lie algebra to

$$\{\hat{\alpha}_m, \hat{\sigma}\}_{LA} = \{(1, \tfrac{\partial}{\partial x} r), (0, \tilde{\lambda}), (0, (\tfrac{\partial}{\partial t} - \tfrac{\partial}{\partial x})\tilde{\lambda}), (0, \tilde{\lambda}\tilde{D}^j) \,|\, j \in J_m\}_{LA}.$$

Now, applying Lemma 3.4.1 yields the desired identity for the Lie algebra generated by $\hat{\alpha}_m$ and $\hat{\sigma}$. □

Proof. (of Lemma 3.5.4) Inserting the definitions of $\tilde{\sigma}$ and Ψ^* we obtain

$$\begin{aligned}
&-\left\langle \tilde{\sigma}(t,x,r), \nabla_z \Psi^* \left(t, -\int_0^x \tilde{\sigma}(t,y,r)dy\right)\right\rangle \\
&= -\sum_{v=1}^{n} \left[\sum_{w=1}^{m_v} \varphi_{(v,w)}(t,r)\tilde{\lambda}_{(v,w)}(t,x)\right] \\
&\quad \left[\sum_{\substack{u=1\\ k_u^v \neq 0}}^{e} k_u^v a_{k_u}(t) \prod_{\substack{i=1\\ i\neq v}}^{n} \left(-\sum_{j=1}^{m_i} \varphi_{(i,j)}(t,r)\tilde{D}_{(i,j)}(t,x)\right)^{k_u^i} \right.\\
&\quad \left.\left(-\sum_{j=1}^{m_v} \varphi_{(v,j)}(t,r)\tilde{D}_{(v,j)}(t,x)\right)^{k_u^v-1}\right].
\end{aligned}$$

Since, in general, one has for integers $k, m \in \mathbb{N}$ and real numbers $a_1, \ldots, a_m \in \mathbb{R}$

$$(a_1 + \ldots + a_m)^k = \sum_{\substack{l_1,\ldots,l_m \in \mathbb{N}_0\\ l_1+\ldots+l_m = k}} \frac{k!}{l_1! \cdots l_m!} a_1^{l_1} \cdots a_m^{l_m},$$

we get the equation

$$\begin{aligned}
&-\left\langle \tilde{\sigma}(t,x,r), \nabla_z \Psi^* \left(t, -\int_0^x \tilde{\sigma}(t,y,r)dy\right)\right\rangle \\
&= \sum_{v=1}^{n} \left[\sum_{w=1}^{m_v} -\varphi_{(v,w)}(t,r)\tilde{\lambda}_{(v,w)}(t,x)\right] \\
&\quad \left[\sum_{\substack{u=1\\ k_u^v \neq 0}}^{e} k_u^v a_{k_u}(t) \prod_{\substack{i=1\\ i\neq v}}^{n} \left(\sum_{\substack{l_{(i,1)},\ldots,l_{(i,m_i)} \in \mathbb{N}_0\\ l_{(i,1)}+\ldots+l_{(i,m_i)} = k_u^i}} \frac{k_u^i!}{l_{(i,1)}! \cdots l_{(i,m_i)}!} \right.\right.\\
&\quad \left.\left(\prod_{j=1}^{m_i} (-\varphi_{(i,j)}(t,r))^{l_{(i,j)}}\right)\left(\prod_{j=1}^{m_i} \tilde{D}_{(i,j)}(t,x)^{l_{(i,j)}}\right)\right) \\
&\quad \left(\sum_{\substack{l_{(v,1)},\ldots,l_{(v,m_v)} \in \mathbb{N}_0\\ l_{(v,1)}+\ldots+l_{(v,m_v)} = k_u^v-1}} \frac{(k_u^v-1)!}{l_{(v,1)}! \cdots l_{(v,m_v)}!} \right.
\end{aligned}$$

$$\left(\prod_{j=1}^{m_v}(-\varphi_{(v,j)}(t,x))^{l_{(v,j)}}\right)\left(\prod_{j=1}^{m_v}\tilde{D}_{(v,j)}(t,x)^{l_{(v,j)}}\right)\Bigg)\Bigg].$$

It holds, in general, for real numbers $a_{(i,j)} \in \mathbb{R}$, $i = 1, \dots, n$, $j = 1, \dots, m_i$

$$\prod_{i=1}^{n}\left(\sum_{j=1}^{m_i} a_{(i,j)}\right) = \sum_{j_1=1}^{m_1}\cdots\sum_{j_n=1}^{m_n}\left(\prod_{i=1}^{n} a_{(i,j_i)}\right).$$

Denoting by $L_{(u,v)}$ the set of all $l_{(i,j)} \in \mathbb{N}_0$, $i = 1, \dots, n$, $j = 1, \dots, m_i$ such that

$$\sum_{j=1}^{m_i} l_{(i,j)} = \begin{cases} k_u^i, & i \neq v \\ k_u^i - 1, & i = v \end{cases},$$

we obtain the identity

$$\begin{aligned}
&-\left\langle \tilde{\sigma}(t,x,r), \nabla_z \Psi^*\left(t, -\int_0^x \tilde{\sigma}(t,y,r)dy\right)\right\rangle \\
&= \sum_{v=1}^{n}\left[\sum_{w=1}^{m_v} -\varphi_{(v,w)}(t,r)\tilde{\lambda}_{(v,w)}(t,x)\right] \\
&\quad \Bigg[\sum_{\substack{u=1 \\ k_u^v \neq 0}}^{e} k_u^v a_{k_u}(t) \sum_{l \in L_{(u,v)}}\left(\prod_{\substack{i=1 \\ i\neq v}}^{n}\frac{k_u^i!}{l_{(i,1)}!\cdots l_{(i,m_i)}!}\right)\frac{(k_u^v-1)!}{l_{(v,1)}!\cdots l_{(v,m_v)}!} \\
&\quad \left(\prod_{i=1}^{n}\prod_{j=1}^{m_i}(-\varphi_{(i,j)}(t,r))^{l_{(i,j)}}\right)\left(\prod_{i=1}^{n}\prod_{j=1}^{m_i}\tilde{D}_{(i,j)}(t,x)^{l_{(i,j)}}\right)\Bigg].
\end{aligned}$$

Expanding this identity, the stated equation follows. □

Proof. (of Lemma 3.5.5) The drift and the volatility vector fields are given by

$$\begin{aligned}
\hat{\alpha}^*(t,r) &= \left(1, \frac{\partial}{\partial x}r\right) + \sum_{u\in U} b_u(t)\varphi_u(t,r)\left(0, \tilde{\lambda}_u(t)\right) \\
&\quad - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{m_i}\sum_{l=1}^{m_j} c_{ij}(t)\varphi_{(j,l)}(t,r)\mathbb{F}_r\varphi_{(i,k)}(t,r)[\tilde{\lambda}_{(j,l)}(t)]\left(0, \tilde{\lambda}_{(i,k)}(t)\right), \\
\hat{\sigma}_i(t,r) &= \sum_{j=1}^{m_i}\varphi_{(i,j)}(t,r)\left(0, \tilde{\lambda}_{(i,j)}(t)\right), \quad i = 1, \dots, n.
\end{aligned}$$

Taking into account the assumptions on the $\varphi_{(i,j)}(t,r)$, the inclusion

$$\begin{aligned}
\{\hat{\alpha}^*, \hat{\sigma}_1, \dots, \hat{\sigma}_n\}_{LA} &\subset \{\hat{\alpha}^*, \hat{\sigma}_1, \dots, \hat{\sigma}_n, \left(1, \tfrac{\partial}{\partial x}r\right), (0, \tilde{\lambda}_v) \,|\, v \in V \cup U\}_{LA} \\
&= \{\left(1, \tfrac{\partial}{\partial x}r\right), (0, \tilde{\lambda}_v) \,|\, v \in V \cup U\}_{LA}
\end{aligned}$$

is valid by an application of Lemma A.2.2. Using Lemma 3.4.1 yields the assertion. □

Proof. (of Lemma 3.5.9) Since $\sup_{t\in[0,T]} |L(t)| < \infty$ for every function $L \in \mathcal{L}$, the assertion of the lemma follows immediately. □

The next proof is an exercise in multidimensional analysis.

Proof. (of Lemma 3.5.10) Assume it holds $|f(x) - f(y)| \leq L\|x - y\|_1$ for all $x, y \in G$. Let $x \in G$ and $i \in \{1, \dots, n\}$ be arbitrary. Since G is open, there exists $\varepsilon > 0$ such that $\{x + he_i \,|\, h \in [0, \varepsilon]\} \subset G$, where the e_i denote the unit vectors in $\mathbb{R}^n$. We obtain

$$\left|\frac{\partial}{\partial x_i} f(x)\right| = \lim_{h\to 0} \frac{|f(x + he_i) - f(x)|}{|h|} \leq L.$$

Conversely, assume the relation $\sup_{x\in G} |\frac{\partial}{\partial x_i} f(x)| \leq L$ is valid for each $i = 1, \dots, n$. For fixed $x, y \in G$ define the function $g : [0, 1] \to \mathbb{R}$ as $g(t) := f(y + t(x - y))$. It holds $\{y + t(x - y) \,|\, t \in [0, 1]\} \subset G$, because G is convex. Thus, we get

$$\begin{aligned}
|f(x) - f(y)| = |g(1) - g(0)| &= \left|\int_0^1 g'(t)dt\right| \\
&= \left|\sum_{i=1}^n \left(\int_0^1 \frac{\partial}{\partial z_i} f(y + t(x - y))dt\right)(x_i - y_i)\right| \leq L\|x - y\|_1.
\end{aligned}$$

□

A.3 Proofs of Chapter 4

The only proof, left open in Chapter 4, is that of Proposition 4.2.1. It is similar to the proof of Raible [56, Prop. 4.12].

Proof. (of Proposition 4.2.1) Let $T \in (0, \infty)$ be arbitrary. Itô's formula Jacod and Shiryaev [42, Thm. I.4.57] yields

$$\begin{aligned}
P(t, T, r_t) = P(0, T, r_0) &+ \int_0^t \frac{\partial}{\partial s} P(s, T, r_{s-})ds + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial r_i} P(s, T, r_{s-})dr_s^i \\
&+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2}{\partial r_i \partial r_j} P(s, T, r_{s-})d\langle r_i^c, r_j^c\rangle_s \\
&+ \sum_{0<s\leq t} \Big[P(s, T, r_s) - P(s, T, r_{s-}) - \sum_{i=1}^d \frac{\partial}{\partial r_i} P(s, T, r_{s-})\Delta r_s^i\Big], \qquad \text{(A.3.1)}
\end{aligned}$$

where the last term is in $\mathcal{V}$. Taking into account $\Delta r_t = \gamma(t, r_{t-})\Delta X_t$ and the differentials

$$\begin{aligned}
dr_t &= \mu(t, r_{t-})dt + \gamma(t, r_{t-})dX_t, \\
d\langle r_i^c, r_j^c\rangle_t &= \langle \gamma_{i\bullet}(t, r_{t-})^*, c(t, r_{t-})\gamma_{j\bullet}(t, r_{t-})^*\rangle dt, \quad i, j = 1, \dots, d,
\end{aligned}$$

and the associativity of the Itô integral Jacod and Shiryaev [42, I.4.37], we write equation (A.3.1) as

$$\begin{aligned} dP(t,T,r_t) = {} & \frac{\partial}{\partial t} P(t,T,r_{t-})dt + \langle \nabla_r P(t,T,r_{t-}), \mu(t,r_{t-})\rangle dt \\ & + \nabla_r P(t,T,r_{t-})\gamma(t,r_{t-})dX_t \\ & + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial r_i \partial r_j} P(t,T,r_{t-})\langle \gamma_{i\bullet}(t,r_{t-})^*, c(t,r_{t-})\gamma_{i\bullet}(t,r_{t-})^*\rangle dt \\ & + \int_{\mathbb{R}^n} \Big(P(t,T,r_{t-} + \gamma(t,r_{t-})x) - P(t,T,r_{t-}) \\ & \qquad - \langle \nabla_r P(t,T,r_{t-}), \gamma(t,r_{t-})x\rangle \Big) \mu^X(dt,dx), \end{aligned} \tag{A.3.2}$$

where the $\mu^X(dt,dx)$-integral is in $\mathcal{V}$. The process $P(t,T,r_t)$ is a special semimartingale, because the dynamics (2.2.16) from Proposition 2.2.9 provide a decomposition where the finite variation part is predictable. According to Prop. I.4.23 and Lemma I.3.10 in Jacod and Shiryaev [42], the $\mu^X(dt,dx)$-integral in (A.3.2) belongs to $\mathcal{A}_{\text{loc}}$. Therefore, we may integrate with respect to $(\mu^X - \nu)(dt,dx)$ plus $\nu(dt,dx)$ Jacod and Shiryaev [42, Prop. II.1.28], and write equation (A.3.2) as

$$P(t,T,r_t) = P(0,T,r_0) + M_t + \int_0^t Y_{s-} ds, \tag{A.3.3}$$

where M is the local martingale

$$\begin{aligned} M_t = {} & \int_0^t \nabla_r P(s,T,r_{s-})\gamma(s,r_{s-})dX_s^c \\ & + \int_0^t \int_{\mathbb{R}^n} \Big(P(s,T,r_{s-} + \gamma(s,r_{s-})x) - P(s,T,r_{s-}) \Big) (\mu^X - \nu)(ds,dx), \end{aligned}$$

and Y is the process given by

$$\begin{aligned} Y_t = {} & \frac{\partial}{\partial t} P(t,T,r_t) + \langle \nabla_r P(t,T,r_t), \mu(t,r_t) + \gamma(t,r_t)\beta(t,r_t)\rangle \\ & + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2}{\partial r_i \partial r_j} P(t,T,r_t)\langle \gamma_{i\bullet}(t,r_t)^*, c(t,r_t)\gamma_{j\bullet}(t,r_t)^*\rangle \\ & + \int_{\mathbb{R}^n} \Big(P(t,T,r_t + \gamma(t,r_t)x) - P(t,T,r_t) \\ & \qquad - \langle \nabla_r P(t,T,r_t), \gamma(t,T,r_t)x\rangle \Big) K_{t,r_t}(dx). \end{aligned}$$

Let $B_t := \exp(-\int_0^t r_s^1 ds)$. Since B has continuous paths, the quadratic co-variation of B_t and $p(t,T)$ vanishes Jacod and Shiryaev [42, Prop. I.4.49.d], and we get

$$B_t P(t,T,r_t) = P(0,T,r_0) + \int_0^t B_{s-} dP(s,T,r_s) + \int_0^t P(s,T,r_{s-}) dB_s.$$

Inserting (A.3.3) and $dB_t = -r_t^1 B_t dt$, we obtain

$$B_t P(t,T,r_t) = P(0,T,r_0) + \int_0^t B_{s-} dM_s + \int_0^t B_{s-}(Y_{s-} - P(s,T,r_{s-})r_{s-}^1)ds.$$

This is the canonical decomposition of the special semimartingale $B_t P(t,T,r_t)$. However, since $x_1 = 0$ by assumption, the process r_t^1 is just the short rate. Due to the fact that $\mathbb{P}$ is a martingale measure, the discounted bond price process $B_t P(t,T,r_t)$ is a local martingale. We deduce that the finite variation part $\int_0^t B_{s-}(Y_{s-} - P(s,T,r_{s-})r_{s-}^1)ds$ vanishes. By Lemma 2.2.10 we conclude that for each $t \in (0,\infty)$ it holds

$$Y_{t-} - P(t,T,r_{t-})r_{t-}^1 = 0 \quad (\mathbb{P} - a.s.)$$

Applying Lemma 3.1.4, for any fixed $t \in (0,\infty)$, yields the desired integro-differential equation. □

Appendix B

Notation

The notation follows the usual conventions, nevertheless the general mathematical symbols that will be used are gathered in the first table. The second table contains standard terminology from probability theory. This is followed by a list of notation concerning stochastic processes. We mainly use the notation of Jacod and Shiryaev [42], and sometimes that of Protter [55]. Afterwards, we collect the relevant notation from interest rate theory. Finally, other quantities, used in the text, are summarized in the last table.

General symbols

$A := B$	A is defined by B
$[a, b]$, (a, b)	closed, open interval from a to b
$\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{Z}$	$\{1, 2, \ldots\}$, $\{0, 1, \ldots\}$, $\{0, +1, -1, +2, -2, \ldots\}$
$\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_-$, $\mathbb{C}$	$(-\infty, \infty)$, $[0, \infty)$, $(-\infty, 0]$, complex numbers
$\operatorname{Re} z, \operatorname{Im} z, \bar{z}$	real part, imaginary part, complex conjugate of $z \in \mathbb{C}$
$a \vee b$, $a \wedge b$	maximum, minimum of a and b
$\|x\|_p$ $(1 \le p < \infty)$	$\left(\sum_{i=1}^n \|x_i\|^p\right)^{1/p}$
$\|x\|_\infty$	$\sup_{i=1,\ldots,n} \|x_i\|$
$\langle x, y \rangle$	inner product $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$ for $x, y \in \mathbb{R}^n$
$\|x\|$	modulus of $x \in \mathbb{R}$ or Euclidean norm $\|x\|_2 = \sqrt{\langle x, x \rangle}$ of $x \in \mathbb{R}^n$
$A_{i\bullet}, A_{\bullet j}$	i-th row or j-th column of a matrix A
A^*	transpose of A
$\det A$	determinant of A
$\|S\|$, S a set	cardinality of S
$A \subset B$	A is contained in B or $A = B$
$\inf A$, $\sup B$	infimum and supremum of sets $A, B \subset \mathbb{R}$
$\operatorname{int} X$, $X \subset \mathbb{R}^n$	interior of X in $\mathbb{R}^n$
$\overline{X}$, $X \subset \mathbb{R}^n$	closure of X in $\mathbb{R}^n$
$\operatorname{span}\{f_1, \ldots, f_n\}$	the subspace spanned by $f_1, \ldots, f_n$
$\dim V$	linear dimension of V

$f(\bullet)$, $g(\bullet, x_2)$	the functions $x \mapsto f(x)$, $x_1 \mapsto g(x_1, x_2)$
f', f'', $f^{(m)}$	first, second, m-fold derivative of f
∇f, $\mathcal{J} f$, $\mathcal{H} f$	gradient of f, Jacobian matrix of f, Hessian matrix of f
ln	natural logarithm
$C(I) = C^0(I)$	$\{f : I \to \mathbb{R} \mid f \text{ is continuous}\}$
$C^m(I)$	$\{f : I \to \mathbb{R} \mid f^{(m)} \text{ exists and is continuous}\}$
$C^\infty(I)$	$\{f : I \to \mathbb{R} \mid f \in C^m(I) \text{ for each } m \in \mathbb{N}\}$
$C^{m,n}(I \times J)$	$\{f : I \times J \to \mathbb{R} \mid \frac{\partial^{m+n}}{\partial x^m \partial y^n} f \text{ exists and is continuous}\}$
$C^{m,n}(I \times J, \mathbb{R}^d)$	$\{f : I \times J \to \mathbb{R}^d \mid f_1, \ldots, f_d \in C^{m,n}(I \times J)\}$

Terminology from probability theory

$(\Omega, \mathcal{F}, \mathbb{P})$	probability space
$\mathbb{E}[X]$	expected value of a random variable X
$\mathbb{E}[X \mid \mathcal{G}]$	conditional expectation of X given the sub σ-algebra $\mathcal{G} \subset \mathcal{F}$
$\mathbb{P} \ll \mathbb{Q}$, $\mathbb{P} \sim \mathbb{Q}$	absolute continuity and equivalence between measures
$\hat{\mu}$	characteristic function $\hat{\mu}(z) = \int_{\mathbb{R}^n} e^{i\langle z,x\rangle} \mu(dx)$
Ψ	cumulant generating function $\Psi(z) = \ln\left(\int_{\mathbb{R}^n} e^{\langle z,x\rangle} \mu(dx)\right)$
$\mu\|_{\mathcal{G}}$	the measure μ restricted to the sub σ-algebra $\mathcal{G}$
$\mu_1 \otimes \mu_2$	product measure from μ_1 and μ_2
$\mathrm{supp}(X)$, $\mathrm{supp}(\mu)$	support of the random variable X or of the measure μ
Leb	Lebesgue measure
$\mathbf{1}_A$	indicator function of the set A
$\mathcal{E} \otimes \mathcal{F}$	the σ-algebra $\sigma(\mathcal{E} \times \mathcal{F})$
$\sigma(Z_i, i \in I)$	σ-algebra generated by $(Z_i)_{i \in I}$
$\mathcal{B}(I)$	Borel σ-algebra of $I \subset \mathbb{R}^n$

Notation concerning stochastic processes

$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in \mathbb{R}_+}, \mathbb{P})$	stochastic basis
$\mathbb{D}$, $\mathbb{L}$	càdlàg and càglàd adapted processes
$\mathcal{A}_{\mathrm{loc}}$, $\mathcal{A}^+_{\mathrm{loc}}$	locally integrable, and locally integrable increasing processes
$\mathcal{V}$	processes of finite variation on compact intervals
$\mathcal{O}$, $\mathcal{P}$	optional σ-algebra, predictable σ-algebra
X_-, ΔX	càglàd modification of X, jumps of X
$\mathrm{Var}(A)$	variation process of $A \in \mathcal{V}$
X^c, M^d	continuous and purely discontinuous martingale parts
$\mathcal{C}(X)$, $\mathcal{D}(X)$	continuous and purely discontinuous parts, excluding zeros
X^τ, $X^{\tau-}$	process X stopped at times τ and $\tau-$
$\\|X\\|_{S^p}$	S^p-norm of a process $X \in \mathbb{D}$
$X_m \xrightarrow{S^p} X$	S^p-convergence $\lim_{m\to\infty} \\|X_m - X\\|_{S^p} = 0$
$\langle M, N\rangle$, $[X, Y]$	predictable quadratic covariation, quadratic co-variation
$X \bullet Y$, $X_- \circ Y$	stochastic integral, Fisk-Stratonovich integral

μ^X, ν	jump measure of X and its compensator
$W * \mu^X$	integral process $W * \mu_t^X = \sum_{s \le t} W(s, \Delta X_s) \mathbf{1}_{\{\Delta X_s \neq 0\}}$
$G_{\mathrm{loc}}(\mu^X)$	$(\mu^X - \nu)$-integrable functions
$W * (\mu^X - \nu)$	stochastic integral of $W \in G_{\mathrm{loc}}(\mu^X)$ with respect to $\mu^X - \nu$
$(B, C, \nu)^I$	integral characteristics of a special semimartingale
$(\beta, c, K)^D$	derivative of a Grigelionis process
$(\beta, c, K; Z)^D$	Z-derivative of a Grigelionis process

Quantities from interest rate theory

$f(t,T)$, $f^*(0,\bullet)$	forward rates, initial forward rate curve
r_t, $r_t(x)$	short rate $r_t = f(t,t)$, forward rates $r_t(x) = f(t, t+x)$
$p(t,T)$, $z(t,T)$	bond prices, discounted bond prices
$\alpha(t,T)$, $\sigma(t,T)$	drift and volatilities of an interest rate model
$A(t,T)$	integrated drift term $A(t,T) = -\int_t^T \alpha(t,s)ds$
$\Sigma(t,T)$	integrated volatilities $\Sigma(t,T) = -\int_t^T \sigma(t,s)ds$
(F, Z)	realization with state process Z

Other symbols

$\mathcal{M}^+(\mathbb{R}^{n\times n})$	symmetric, non-negative definite $n \times n$-matrices
$\mathcal{M}^{++}(\mathbb{R}^{n\times n})$	symmetric, positive definite $n \times n$-matrices
$\mathcal{M}_{\mathcal{D}}(\mathbb{R}^{n\times n})$	$n \times n$-diagonal matrices
$\mathcal{K}^+(\mathbb{R}^n)$, $\mathcal{K}^{++}(\mathbb{R}^n)$	Lévy measures on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, non-zero Lévy measures
$\mathcal{D}$, $\mathcal{L}$	$\{f : \mathbb{R}_+ \to \mathbb{R} \mid f$ is càdlàg$\}$, $\{f : \mathbb{R}_+ \to \mathbb{R} \mid f$ is càglàd$\}$
$C_b^{\mathcal{L}}(\mathbb{R}_+ \times \mathbb{R}^d)$	$\{f \in C^{0,1}(\mathbb{R}_+ \times \mathbb{R}^d) \mid f, \frac{\partial}{\partial x_1} f, \ldots, \frac{\partial}{\partial x_d} f$ are bounded by a $L \in \mathcal{L}\}$
Δ^n, $n \in \mathbb{N}$	$\{x \in \mathbb{R}^n : x_1 \le \ldots \le x_n\}$
$L(\mathcal{X}, \mathcal{Y})$	space of continuous linear operators from $\mathcal{X}$ to $\mathcal{Y}$
$L^{(n)}(\mathcal{X}, \mathcal{Y})$	space of continuous multilinear operators from $\mathcal{X}^n$ to $\mathcal{Y}$
$\mathbb{F}f$	Fréchet derivative of f
$[f, g]$	Lie bracket $[f,g](x) = \mathbb{F}f(x)[g(x)] - \mathbb{F}g(x)[f(x)]$
$\{f_1, \ldots, f_n\}_{LA}$	Lie algebra generated by smooth vector fields
$\mathcal{T}_{\mathcal{G}}(x_0)$	tangent space of a manifold $\mathcal{G}$ at point $x_0 \in \mathcal{G}$
$\langle f, g \rangle_{\beta,\gamma}$	$\sum_{n=0}^{\infty} \beta^{-n} \int_0^{\infty} \left(\frac{\partial^n}{\partial x^n} f(x)\right) \left(\frac{\partial^n}{\partial x^n} g(x)\right) e^{-\gamma x} dx$ for $\beta > 1, \gamma > 0$
$\mathcal{H}_{\beta,\gamma}$	$\{f \in C^\infty(\mathbb{R}_+) : \langle f, f \rangle_{\beta,\gamma} < \infty\}$
$\Theta_t r$	right-shift $\Theta_t r(x) = r(t+x)$

Appendix C

Acknowledgement

Writing this thesis proved to be challenging, and many people have contributed to the completion of this text.

I gratefully thank Prof. U. Küchler, who supervised and encouraged me throughout the years. Without his support and encouragement, in particular during the first year when my efforts seemed to fail, I surely would have given up.

Prof. T. Björk and Prof. U. Küchler provided me with the opportunity for a stay at the Stockholm School of Economics (November 2004), which was financed by the Deutsche Forschungsgemeinschaft (DFG), Graduiertenkolleg 251 "Stochastische Prozesse und probabilistische Analysis". I enjoyed the hospitality of the Department of Finance, and am grateful to Prof. T. Björk, who sacrificed quite some time discussing my manuscripts. His generous sharing of his broad mathematical knowledge, his invaluable comments and constructive criticism have been very rewarding. Thanks to his illustrations, I got a much better geometric understanding of the Lie algebraic theory.

This text would not have been written at all, if it wasn't for Prof. H. Dietz who sparked my interest for mathematical research in the field of probability theory, and helped in many ways to get me started with this project.

I would like to thank my colleagues at the Humboldt University of Berlin for the pleasant company during the last years: My office roommate Robert Lorenz, Markus Riedle and Markus Reiß from the neighboring office (the latter is a former member of the research group), Hagen Gilsing, Evelyn Buckwar and Onno van Gaans.

Sometimes, emotional support is more important than any scientific advice. Writing this thesis would have been impossible without the support, motivation and consultation of my family and closest friends. I am deeply indebted to you, and dedicate this thesis to all of you. My gratitude goes to my mother Monika, my father Jörn, and my brother Christian, for their constant support in all situations. I take the opportunity to thank Sascha Nickel for his friendship over the years. When Sascha and Heyka visited me in Berlin at the end of May, it seemed impossible to submit the thesis just within one month, but, after an enjoyable weekend, the strangest things can happen...

Very special thanks go to Stephanie Hof.

I gratefully acknowledge that the research has been financed by the Deutsche Forschungsgemeinschaft (DFG), Graduiertenkolleg 251 "Stochastische Prozesse und probabilistische Analysis".

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Arbeit selbständig ohne fremde Hilfe verfaßt und nur die angegebene Literatur und die angegebenen Hilfsmittel verwendet zu haben.

Stefan Tappe
5. Juli 2005