



Dipl.-Ing. Stefan Waldenberger

# **Time-inhomogeneous affine processes and affine market models**

## **DISSERTATION**

zur Erlangung des akademischen Grades

Doktor der technischen Wissenschaften

eingereicht an der

**Technischen Universität Graz**

Betreuer/in:  
Prof. Dr. Wolfgang Müller

Institut für Statistik

Graz, September 2015

## EIDESSTATTLICHE ERKLÄRUNG

### *AFFIDAVIT*

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZ-online hochgeladene Textdokument ist mit der vorliegenden Dissertation identisch.

*I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral dissertation.*

---

Datum/Date

---

Unterschrift/Signature

## Deutschsprachige Kurzfassung

Diese Dissertation beschäftigt sich mit affinen Prozessen und deren Anwendung im Bereich der Finanzmathematik. Im ersten Teil betrachten wir die Theorie zeitinhomogener affiner Prozesse auf allgemeinen Zustandsräumen. Zuerst werden zeitinhomogene Markov Prozesse genau definiert. Für stochastisch stetige affine Prozesse zeigen wir, dass immer eine càdlàg Modifikation existiert. Anschließend betrachten wir die Regularität und die Semimartingal Eigenschaft von affinen Prozessen. Im Gegensatz zum zeithomogenen Fall sind zeitinhomogene affine Prozesse im Allgemeinen weder regulär noch Semimartingale. Der zeitinhomogene Fall führt daher zu neuen und interessanten Fragestellungen. Nimmt man an, dass ein affiner Prozess auch ein Semimartingal ist, so kann man auch ohne Regularität zeigen, dass die Parameterfunktionen verallgemeinerte Riccati Integralgleichungen lösen. Diese Aussage verallgemeinert ein wichtiges Resultat für zeithomogene affine Prozesse. Wir zeigen weiters, dass stochastisch stetige affine Semimartingale im Wesentlichen durch deterministische Zeittransformationen aus absolut stetig affinen Semimartingalen entstehen. Diese Prozesse sind eine Verallgemeinerung der zeithomogenen regulär affinen Prozesse. Im zweiten Teil betrachten wir die Klasse der affinen LIBOR Marktmodelle. Wir modifizieren das ursprüngliche Modell, sodass neben nichtnegativen affinen Prozessen auch reellwertige affine Prozesse verwendet werden können. Numerische Beispiele zeigen, dass dadurch flexiblere Volatilitätsoberflächen erzeugt werden können. Weiters führen wir die Klasse der affinen Inflationsmarktmodelle ein. Diese Modelle erweitern affine LIBOR Marktmodelle auf Inflationsmärkte. Ein Kalibrierungsbeispiel zeigt, dass beobachtete Marktpreise von Inflationsderivaten sehr gut wiedergegeben werden können.

## Abstract

This thesis is devoted to the study of affine processes and their applications in financial mathematics. In the first part we consider the theory of time-inhomogeneous affine processes on general state spaces. We present a concise setup for time-inhomogeneous Markov processes. For stochastically continuous affine processes we show that there always exists a càdlàg modification. Afterwards we consider the regularity and the semimartingale property of affine processes. Contrary to the time-homogeneous case, time-inhomogeneous affine processes are in general neither regular nor semimartingales and the time-inhomogeneous case raises many new and interesting questions. Assuming that an affine process is a semimartingale, we show that even without regularity the parameter functions satisfy generalized Riccati integral equations. This generalizes an important result for time-homogeneous affine processes. We also show that stochastically continuous affine semimartingales are essentially generated by deterministic time-changes of what we call absolutely continuously affine semimartingales. These processes generalize time-homogeneous regular affine processes.

In the second part we consider the class of affine LIBOR market models. We contribute to this class of models in two ways. First, we modify the original setup of the affine LIBOR market models in such a way that next to nonnegative affine processes real-valued affine processes can also be used. Numerical examples show that this allows for more flexible implied volatility surfaces. Second, we introduce the class of affine inflation market models, an extension of the affine LIBOR market models. A calibration example shows that these models perform very well in fitting market-observed prices of inflation derivatives.

## Acknowledgements

First and foremost I want to thank my advisor, Wolfgang Müller, who encouraged me to pursue a PhD in this area. His door was always open and I am very grateful for all the time he spent discussing and answering my many questions. I would also like to thank the whole staff of the Institute of statistics, who filled the daily coffee breaks with many interesting discussions. It was a very familiar environment and I was very happy to spend my time there. Special thanks go to Ernst Stadlober, who was very generous with my requests to attend conferences and scientific meetings.

I also want to thank Christa Cuchiero, Philipp Harms, Josef Teichmann and Thorsten Schmidt, with whom I had many interesting and helpful discussions on affine processes. I am especially grateful to Josef Teichmann for inviting me to ETH Zürich and pointing me in some interesting directions.

Special thanks go to my family for supporting all my decision and giving me the opportunity to get to this point. Finally, and perhaps most importantly, I want to thank Steffi Klinger, who encouraged me during the whole time of my PhD, unquestioningly put up with all the hours spent on writing this thesis and who always listened to my doubts and problems.

# Contents

<b>Introduction</b>	<b>1</b>
<b>1 Time-inhomogeneous affine processes</b>	<b>3</b>
1.1 Markov processes . . . . .	4
1.2 Affine processes . . . . .	11
1.3 Càdlàg modifications . . . . .	20
1.4 Regularity and semimartingale characteristics . . . . .	31
1.4.1 Affine Markov semimartingales . . . . .	35
<b>2 Affine LIBOR models driven by real-valued affine processes</b>	<b>49</b>
2.1 Affine processes . . . . .	49
2.2 Interest rate market models . . . . .	51
2.3 The modified affine LIBOR model . . . . .	54
2.3.1 Option pricing . . . . .	56
2.3.2 Examples . . . . .	60
2.4 Proofs . . . . .	65
<b>3 Affine inflation market models</b>	<b>68</b>
3.1 Inflation markets . . . . .	68
3.1.1 The inflation market model . . . . .	71
3.2 The Affine inflation market model . . . . .	72
3.2.1 Forward CPI and CPI options . . . . .	73
3.2.2 Forward Inflation and inflation caplets and floorlets . . . . .	74
3.2.3 Correlation . . . . .	76
3.3 Implementation example . . . . .	77
3.4 Appendix . . . . .	84
<b>Bibliography</b>	<b>88</b>

# Introduction

This dissertation is centered around the class of affine processes and their applications to financial markets. Time-inhomogeneous affine processes are informally described as  $E$ -valued Markov processes, where the logarithm of the conditional moment generating function is affine in the current state of the process, i.e.

$$\mathbb{E} [e^{\langle u, X_t \rangle} | \mathcal{F}_s] = \exp\left(\phi_{s,t}(u) + \langle \psi_{s,t}(u), X_s \rangle\right), \quad u \in \mathbf{i}\mathbb{R}^d, x \in E,$$

with deterministic functions  $\phi$  and  $\psi$ . This means that the transition operator of the Markov process maps exponentially affine functions into exponentially affine functions.

In the first part of this thesis we introduce time-inhomogeneous affine processes on general state spaces. While time-homogeneous affine processes ( $\phi$  and  $\psi$  depend only on  $t - s$ ) are extensively studied, time-inhomogeneous affine processes did not get the same amount of attention. We focus on continuously affine transition functions (referring to the continuity of  $\phi$  and  $\psi$ ), which basically correspond to stochastically continuous affine processes. For affine processes with a continuously affine transition function we are able to show that there always exists a càdlàg modification of the process. This extends a similar result for time-homogeneous affine processes.

We then consider the question of regularity and the semimartingale property of time-inhomogeneous affine processes. Regularity is related to differentiability of  $\phi$  and  $\psi$  with respect to the time parameters. For time-homogeneous continuously affine processes regularity always holds. It is used to infer that the parameter functions are solutions of generalized Riccati differential equations and that there exist differentiated semimartingale characteristics which are affine in the current state of the process. We give examples that in the time-inhomogeneous case neither regularity nor the semimartingale property holds automatically. However, for time-inhomogeneous continuously affine processes which are simultaneously semimartingales we can extend many results of the time-homogeneous case. We are able to show that in this case the parameter functions are solutions to generalized Riccati integral equations, where the integral is a Lebesgue-Stieltjes integral with respect to a deterministic function  $G$ . Furthermore, the semimartingale characteristics are absolutely continuous with respect to  $G$  and depend in an affine way on the current state of the process. If the functions  $\phi$  and  $\psi$  are even absolutely continuous, we show that one can choose  $G(t) = t$ . We say that such affine semimartingales have an absolutely continuously affine transition function and propose that absolutely continuously affine processes are a suitable generalization of time-homogeneous regular affine processes. The de-

terministic function  $G$  can be interpreted as a time-change. It follows that continuously affine semimartingales are essentially equal to deterministically time-changed absolutely continuously affine semimartingales.

In the second part of this dissertation we consider practical applications of affine processes based on the class of affine LIBOR market models as introduced in Keller-Ressel et al. [36]. The class of affine LIBOR market models uses the analytical tractability of affine processes to describe an analytically tractable class of LIBOR market models. In order to guarantee nonnegative interest rates only nonnegative affine processes can be used. In chapter 2 (which corresponds to Müller and Waldenberger [41]) we extend the affine LIBOR market models in such a way that real-valued affine processes can also be used without destroying the nonnegativity of interest rates. This has the advantage that the resulting models are more flexible in producing market-observed implied volatility smiles. We illustrate this by numerical examples.

Chapter 3 extends affine LIBOR market models to inflation markets. In inflation markets there exist two different types of liquidly traded inflation swaps. Market models considered so far are not able to give closed formulas for both types of swaps. Here the analytical tractability of affine inflation market models can be used to get closed formulas for both swaps. Furthermore, calls and puts on these swap rates can be priced by Fourier methods. This covers pricing for the most liquidly traded derivatives in inflation markets. Using these formulas, we are able to calibrate an affine inflation market model to market data. This chapter corresponds to Waldenberger [45].

# 1 Time-inhomogeneous affine processes

Time-homogeneous affine processes are time-homogeneous Markov processes with a state space  $E$ , such that the transition probabilities  $P_t$  satisfy

$$\int_E e^{\langle u, \xi \rangle} P_t(x, d\xi) = \Phi_t(u) e^{\langle \psi_t(u), x \rangle}$$

for  $x \in E$ ,  $t \geq 0$  and all  $u$  such that  $x \mapsto e^{\langle u, x \rangle}$  is bounded in  $E$ .

The class of affine processes includes many well-known and widely used stochastic processes. In particular, Lévy processes, Ornstein-Uhlenbeck processes (Sato [44], section 17), CIR processes (Cox et al. [9]) and Wishardt-processes (Bru [5]) belong to this class. Affine processes go back to Kawazu and Watanabe [32], where continuous-time limits of Galton-Watson branching processes were studied. Later affine processes with values in the canonical state space  $\mathbb{R}_{\geq 0}^m \times \mathbb{R}^n$  were introduced and applied in a series of papers, compare Duffie [14], Dai and Singleton [12], Duffie et al. [15]. An important break-through was presented in Duffie et al. [16], where a complete characterization of regular affine processes taking values in the canonical state space was given. Here regularity refers to the existence of the time-derivatives of  $\Phi$  and  $\psi$ . It was shown that  $\Phi$  and  $\psi$  are solutions of generalized Ricatti differential equations. Furthermore, it was proved that regular affine processes are semimartingales whose differentiated semimartingale characteristics are affine in the current state of the process. It was not until recently, that in another series of papers it was shown that the assumption of regularity is automatically satisfied for time-homogeneous stochastically continuous affine processes. For the canonical state space this was done in Keller-Ressel et al. [35], and for general state spaces in Keller-Ressel et al. [37] and Cuchiero and Teichmann [10].

Time-inhomogeneous affine processes are analogously defined with the functions  $\Phi$  and  $\psi$  depending on two time parameters. They were first introduced in Filipovic [21], where, assuming stronger forms of regularity, the results in Duffie et al. [16] were extended to the time-inhomogeneous case. In this chapter we explore how the theory of time-homogeneous affine processes with a general state space extends to the time-inhomogeneous case without such strong regularity assumptions. We show that the time-inhomogeneous case raises new and interesting questions, which do not appear in the time-homogeneous case.

We start by defining time-inhomogeneous Markov processes in a way which fits especially well with semimartingale theory. In section 1.2 we introduce time-inhomo-

geneous affine processes on general state spaces and show that stochastic continuity of the Markov process is essentially equivalent to the continuity of  $\Phi$  and  $\psi$ . This leads to the definition of a continuously affine transition function. In section 1.3 we extend the results of Cuchiero and Teichmann [10] to the time-inhomogeneous case and prove that continuously affine processes have a càdlàg modification which possesses the strong Markov property.

The last part of chapter 1 focuses on regularity questions and the semimartingale property. While in the time-homogeneous case regularity is automatically satisfied, this is not true for time-inhomogeneous affine processes. Furthermore, the semimartingale property no longer holds in general, as examples show. Under mild technical conditions we are able to show that for Markov semimartingales with a continuously affine transition function the semimartingale characteristics are absolutely continuous with respect to a deterministic increasing function  $G$ . Additionally, they depend in an affine way on the current state of the process. From this we get that the functions  $\Phi$  and  $\psi$  are solutions of generalized Riccati integral equations, where the integrals are Lebesgue-Stieltjes integrals with respect to the function  $G$ . Hence even without regularity many of the results for time-homogeneous affine processes generalize.

The function  $G$  can be used to time-change the process. Fixing the probability measure, the time-changed process turns out to be an affine process with a new affine transition function, where the new  $\Phi$  and  $\psi$  satisfy generalized Riccati integral equations with respect to the Lebesgue measure, i.e. they are even absolutely continuous. This leads to the definition of an absolutely continuously affine transition function. The time-changed process then has differentiated semimartingale characteristics, which depend on time and affinely on the current state of the process. We consider such processes as a suitable generalization of time-homogeneous regular affine processes. The class of Markov semimartingales with continuously affine transition functions is therefore essentially equal to the class of deterministically time-changed Markov semimartingales with an absolutely continuously ('regular') affine transition function.

## 1.1 Markov processes

**Definition 1.1:** A transition kernel  $N$  on a measurable space  $(E, \mathcal{E})$  is a map

$$N : E \times \mathcal{E} \rightarrow [0, \infty],$$

such that

- for each  $x \in E$  the map  $A \mapsto N(x, A)$  is a measure,
- for each  $A \in \mathcal{E}$  the map  $x \mapsto N(x, A)$  is measurable.

If furthermore  $N(x, E) = 1$  for all  $x \in E$ , then  $N$  is a transition probability. If  $N(x, E) < 1$  we can add a cemetery point  $\Delta$  to  $E$  and set  $N(x, \{\Delta\}) = 1 - N(x, E)$  and

$N(\Delta, \{\Delta\}) = 1$ , so that  $N$  is a transition probability on  $(E_\Delta, \mathcal{E}_\Delta)$ , where  $E_\Delta = E \cup \{\Delta\}$  and  $\mathcal{E}_\Delta = \sigma(\mathcal{E}, \{\Delta\})$ .

Functions  $f$  mapping from  $E$  to  $\mathbb{R}$  or  $\mathbb{C}$  are extended to  $E_\Delta$  by setting  $f(\Delta) = 0$ . To shorten notation for the rest of this section we use  $E$  for the state space, even if there exists a cemetery point. A transition probability introduces a transition operator on the set of bounded (or alternatively nonnegative) measurable functions  $f : E \rightarrow \mathbb{R}$  by

$$Nf(x) := \int_E f(y)N(x, dy).$$

Note that we use the same letter to denote the transition probability and the corresponding operator. For two kernels  $M, N$  it is possible to define  $MN$  by

$$MN(x, A) := \int_E N(y, A)M(x, dy),$$

which is again a transition kernel. Note that  $x \mapsto Nf(x)$  is measurable and that  $(MN)f(x) = M(Nf)(x)$ .

**Definition 1.2:** A transition function on  $(E, \mathcal{E})$  is a collection  $\{P_{s,t}\}_{0 \leq s \leq t}$  of transition probabilities on  $(E, \mathcal{E})$  such that

$$P_{s,u}P_{u,t} = P_{s,t}$$

for all  $0 \leq s < u < t$  and  $P_{s,s}(x, \cdot) = \delta_x(\cdot)$ , where  $\delta_x(\cdot)$  denotes the Dirac measure. The transition function is called time-homogeneous if  $P_{s,t} = P_{t-s}$  for a collection  $\{P_t\}_{t \geq 0}$  of transition probabilities.

**Definition 1.3:** A process  $X$  with state space  $(E, \mathcal{E})$  on the filtered probability space  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$  with  $\mathcal{F} := (\mathcal{F}_t)_{t \geq 0}$ , such that  $X_s = \Delta$  implies  $X_t = \Delta$  for all  $t > s$ , is a time-inhomogeneous Markov process with transition function (operator)  $\{P_{s,t}\}$  on  $(E, \mathcal{E})$ , if it is  $\mathcal{F}$ -adapted and

$$\mathbb{E}[f(X_t) | \mathcal{F}_s] = P_{s,t}f(X_s) \tag{1.1}$$

for all times  $s < t$  and bounded measurable  $f : E \rightarrow \mathbb{R}$ . It is a time-homogeneous Markov process if the transition function is time-homogeneous.

As with the analogy between transition functions and transition operators one can equivalently formulate the Markov property by requiring that for  $B \in \mathcal{E}$ ,

$$\mathbb{P}(X_t \in B | \mathcal{F}_s) = P_{s,t}(X_s, B).$$

One can always find a Markov process with a given transition function. This is the statement of the following lemma (see e.g. I.3 in Revuz and Yor [42]).

**Lemma 1.4:** Consider the coordinate process  $X$  on the filtered space  $(E^{\mathbb{R}_{\geq 0}}, \mathcal{F}_{\infty}^X, \mathcal{F}^X)$  with  $\mathcal{F}_{\infty}^X := \sigma((X_t)_{t \geq 0})$  and  $\mathcal{F}_t^X = \sigma((X_s)_{0 \leq s \leq t})$ . For a given transition function  $\{P_{s,t}\}$  and start distribution  $\mu$  on  $E$  there is a unique probability measure  $\mathbb{P}$  on  $(E^{\mathbb{R}_{\geq 0}}, \mathcal{F}_{\infty}^X)$ , such that  $X$  is a Markov process on  $(E^{\mathbb{R}_{\geq 0}}, \mathcal{F}_{\infty}^X, \mathcal{F}^X, \mathbb{P})$  and  $X_0$  has distribution  $\mu$ .

This motivates an alternative definition of a Markov process, not with respect to a fixed probability measure  $\mathbb{P}$ , but with a whole family of probability measures. We demonstrate this in the time-homogeneous case first. Consider

- i) an  $\mathcal{F}$ -adapted  $(E, \mathcal{E})$ -valued stochastic process  $X$  on  $(\Omega, \mathcal{A}, \mathcal{F})$ . If  $E$  contains a cemetery point  $\Delta$  it has to hold that  $X_s = \Delta$  implies  $X_t = \Delta$  for all  $t > s$ .
- ii) Assume that for every  $x \in E$  there exists a measure  $\mathbb{P}^x$  on  $(\Omega, \mathcal{A})$ , such that  $\mathbb{P}^x(X_0 = x) = 1$ . Additionally, assume that for all  $t \geq 0$  and  $B \in \mathcal{E}$  the map  $x \mapsto \mathbb{P}^x(X_t \in B)$  is measurable. In other words  $P_t(x, B) := \mathbb{P}^x(X_t \in B)$  is a transition probability on  $(E, \mathcal{E})$ .

Denote by  $\mathbb{E}^x$  the expectation value with respect to the measure  $\mathbb{P}^x$ .

**Definition 1.5:** The combination  $(X, \mathcal{F}, \{\mathbb{P}^x\})$  is a time-homogeneous Markov process on  $(\Omega, \mathcal{A})$ , if i) and ii) are satisfied and for each bounded measurable function  $f : E \rightarrow \mathbb{R}$  and  $x \in E$ ,  $s, t \geq 0$

$$\mathbb{E}^x [f(X_{s+t}) | \mathcal{F}_s] = P_t f(X_s). \quad (1.2)$$

Note that the tower property for conditional expectations gives that  $\{P_t\}$  is a time-homogeneous transition function. With this definition  $X$  is a homogeneous Markov process in the sense of Definition 1.3 on  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P}^x)$  for each  $x \in E$ . Contrary, for the transition function  $\{P_t\}$  of a time-homogeneous Markov process in the sense of Definition 1.3 we can construct probability measures  $\mathbb{P}^x$  using Lemma 1.4 with starting distribution  $\mu = \delta_x$ . Then  $(X, \mathcal{F}^X, \{\mathbb{P}^x\})$  is a Markov process on  $(\Omega, \mathcal{F}_{\infty}^X)$  in the sense of Definition 1.5. This is then called the canonical realization of  $\{P_t\}$ .

Sometimes the Markov property is written as

$$\mathbb{E}^x [f(X_{t+s}) | \mathcal{F}_s] = \mathbb{E}^{X_s} [f(X_t)].$$

Here the right hand side is defined as  $\mathbb{E}^{X_s} [f(X_t)] = g(X_s)$ , where

$$g(x) = \mathbb{E}^x [f(X_t)] = P_t f(x).$$

Note that  $g$  is measurable and that  $\mathbb{E}^{X_s} [f(X_t)]$  is a  $\sigma(X_s)$ -measurable random variable.

Next we generalize Definition 1.5 to time-inhomogeneous Markov processes. For this we replace ii) by

- ii') Assume that for each  $r \geq 0$  and  $x \in E$  there exists a probability measure  $\mathbb{P}^{(r,x)}$  on  $(\Omega, \mathcal{A})$  such that  $\mathbb{P}^{(r,x)}(X_0 = x) = 1$ . Additionally, assume that for all  $s, t \geq 0$ ,  $B \in \mathcal{E}$  the map  $x \mapsto \mathbb{P}^{(s,x)}(X_t \in B)$  is measurable. This implies that  $P_{s,s+t}(x, B) := \mathbb{P}^{(s,x)}(X_t \in B)$  is a transition probability.

**Definition 1.6:** *The combination  $(X, \mathcal{F}, \{\mathbb{P}^{(r,x)}\})$  is a time-inhomogeneous Markov process on  $(\Omega, \mathcal{A})$ , if i) and ii') are satisfied and for each bounded measurable function  $f : E \rightarrow \mathbb{R}$  and  $x \in E$ ,  $r, s, t \geq 0$*

$$\mathbb{E}^{(r,x)} [f(X_{t+s}) | \mathcal{F}_s] = P_{r+s,r+s+t} f(X_s) = \mathbb{E}^{(r+s, X_s)} [f(X_t)]. \quad (1.3)$$

In this case the tower property gives that  $\{P_{s,t}\}$  is a (time-inhomogeneous) transition function.

*Remark:* Definition 1.6 is not standard. For instance in Gihman and Skorohod [24] a Markov process is defined as a system  $(X, \mathcal{F}_s^r, \{\mathbb{P}^{(r,x)}\})$ , where

- $(\mathcal{F}_s^r)_{0 \leq r \leq s \leq \infty}$  is a family of  $\sigma$ -algebras, such that  $\mathcal{F}_s^r \subset \mathcal{F}_t^v$  for  $0 \leq v \leq r \leq s \leq t$ ,
- $\mathbb{P}^{(r,x)}$  are probability measures on  $(\Omega, \mathcal{F}_\infty^r)$  for  $x \in E$ , such that  $x \mapsto \mathbb{P}^{(r,x)}(X_t \in B)$  is measurable and
- $X$  is a process such that  $X_t$  is  $\mathcal{F}_t^t$ -measurable and for each bounded measurable  $f$  and  $0 \leq r \leq s \leq t < \infty$

$$\mathbb{E}^{(r,x)} [f(X_t) | \mathcal{F}_s^r] = \mathbb{E}^{(s, X_s)} [f(X_t)].$$

The canonical choice for the family of  $\sigma$ -algebras is  $\mathcal{F}_s^r = \sigma(X_t, r \leq t \leq s)$ . We believe that Definition 1.6 is more convenient, especially if one considers Markov processes which are semimartingales, as done in section 1.4.

With Definition 1.6  $X$  is a time-inhomogeneous Markov process in the sense of Definition 1.3 on  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P}^{(r,x)})$  with transition function  $\{P_{r+s,r+t}\}_{0 \leq s \leq t}$  for each  $(r, x)$ . On the other hand given a time-inhomogeneous transition function  $\{P_{s,t}\}$  we can construct probability measures  $\mathbb{P}^{(r,x)}$  by applying Lemma 1.4 using the transition function  $\{P_{r+s,r+t}\}_{0 \leq s \leq t}$  and starting measure  $\mu = \delta_x$ , such that the coordinate process  $X$  is a Markov process on  $(E^{\mathbb{R}_{\geq 0}}, \mathcal{F}_\infty^X, \mathcal{F}^X, \mathbb{P}^{(r,x)})$ . Then  $(X, \mathcal{F}^X, \{\mathbb{P}^{(r,x)}\})$  is a time-inhomogeneous Markov process in the sense of Definition 1.6. We call  $(X, \mathcal{F}^X, \{\mathbb{P}^{(r,x)}\})$  the canonical realization of  $\{P_{s,t}\}$ .

For the canonical realization of a transition function one can define the shift operators  $\theta_t : E^{\mathbb{R}_{\geq 0}} \rightarrow E^{\mathbb{R}_{\geq 0}}$ ,  $\theta_t(\omega)(s) := \omega(s+t)$ , so that

$$X_s(\theta_t(\omega)) = X_{s+t}(\omega). \quad (1.4)$$

Alternatively one could assume the existence of operators  $\theta_t : \Omega \rightarrow \Omega$ , such that (1.4) holds. In this case the Markov property can be extended. Note that this is sometimes also used to define a Markov process (see Proposition III.1.7 in Revuz and Yor [42] for the homogeneous case). Remember that  $\mathcal{F}_\infty^X = \sigma(X_t, t \geq 0)$  is the  $\sigma$ -algebra generated by the process  $X$ .

**Lemma 1.7:** Let  $(X, \mathcal{F}, \{\mathbb{P}^{(r,x)}\})$  be a Markov process and assume that shift operators  $\theta_t$  exist. Then for every bounded (or nonnegative)  $\mathcal{F}_\infty^X$ -measurable random variable  $Z$  and  $r, t \geq 0, x \in E$

$$\mathbb{E}^{(r,x)} [Z \circ \theta_t | \mathcal{F}_t] = \mathbb{E}^{(r+t, X_t)} [Z], \quad \text{a.s. on the set } \{X_t \neq \Delta\}. \quad (1.5)$$

*Proof.* For the time-homogeneous case this is proved in Revuz and Yor [42], Proposition III.1.6 and III.1.7. The time-inhomogeneous case is similar. For the right hand side to be meaningful we need measurability of  $x \mapsto \mathbb{E}^{(s,x)} [Z]$  for each  $s \geq 0$ . For  $Z = \mathbb{I}_\Gamma$ , where  $\Gamma = \{X_0 \in A_0, X_{t_1} \in A_1, \dots, X_{t_n} \in A_n\}$  this follows from the Markov property and measurability of the transition function. The family of such sets  $\Gamma$  is closed under finite intersections, an application of the monotone class theorem (Revuz and Yor [42], section 0.II) gives measurability for  $\Gamma \in \mathcal{F}_\infty^X$ . An approximation by simple functions extends measurability to  $\mathcal{F}_\infty^X$ -measurable  $Z$ .

Now let  $\Gamma = \{X_s \in A\}$ . Then equation (1.5) holds, i.e.

$$\mathbb{E}^{(r,x)} [\mathbb{I}_\Gamma \circ \theta_t | \mathcal{F}_t] = \mathbb{E}^{(r,x)} [\mathbb{I} \{X_{s+t} \in A\} | \mathcal{F}_t] = P_{r+t, r+t+s}(X_t, A) = \mathbb{E}^{(r+t, X_t)} [\Gamma].$$

With the Markov property this also follows for  $\Gamma = \{X_0 \in A_0, X_{t_1} \in A_1, \dots, X_{t_n} \in A_n\}$  and a monotone class argument as above extends this to sets  $\Gamma \in \mathcal{F}_\infty^X$  and then to  $\mathcal{F}_\infty^X$ -measurable  $Z$ .  $\square$

*Remark:* The measurability of  $x \mapsto \mathbb{E}^{(r,x)} [Z]$  for bounded  $\mathcal{F}_\infty^X$ -measurable  $Z$  allows us to define for each  $r \geq 0$  and probability measure  $\mu$  on  $(E, \mathcal{E})$  a probability measure  $\mathbb{P}^{(r,\mu)}$  on  $(\Omega, \mathcal{F}_\infty^X)$  (but not  $(\Omega, \mathcal{A})$ ) by

$$\mathbb{P}^{(r,\mu)}(\Lambda) = \int_E \mathbb{P}^{(r,x)}(\Lambda) \mu(dx).$$

Denote by  $\mathbb{E}^{(r,\mu)}$  the expectation under  $\mathbb{P}^{(r,\mu)}$ . For each bounded  $\mathcal{F}_\infty^X$ -measurable  $Z$

$$\mathbb{E}^{(r,\mu)} [Z] = \int_E \mathbb{E}^{(r,x)} [Z] \mu(dx).$$

Furthermore equation (1.5) holds with  $\mathbb{E}^{(r,x)}$  replaced by  $\mathbb{E}^{(r,\mu)}$  and  $\mathcal{F}_t$  replaced by  $\mathcal{F}_t^X \subset \mathcal{F}_t$ .

A time-inhomogeneous Markov process is time-homogeneous if the measures  $\mathbb{P}^{(r,x)}$  do not depend on  $r$ . The time-inhomogeneous case can be reduced to the time-homogeneous case by extending the state space and the probability space (see also Böttcher [6] or Wentzell [46], section 8.5.5). Consider the following:

- The extended state space  $\tilde{E} := \mathbb{R}_{\geq 0} \times E$  with the  $\sigma$ -algebra

$$\tilde{\mathcal{E}} := \{B \subset \tilde{E} : B^s := \{x : (s, x) \in B\} \in \mathcal{E} \text{ for all } s \in \mathbb{R}_{\geq 0}\}.$$

- The extended probability space  $\tilde{\Omega} := \mathbb{R}_{\geq 0} \times \Omega$  with the  $\sigma$ -algebra

$$\tilde{\mathcal{A}} := \{A \subset \tilde{\Omega} : A^s := \{\omega : (s, \omega) \in A\} \in \mathcal{A} \text{ for all } s \in \mathbb{R}_{\geq 0}\}.$$

- The space-time process  $\tilde{X} = (\Theta, X)$  on  $(\tilde{\Omega}, \tilde{\mathcal{A}})$  defined by

$$\tilde{X}_t(s, \omega) := (s + t, X_t(\omega)).$$

- The filtration  $\tilde{\mathcal{F}}$  defined by  $\tilde{\mathcal{F}}_t := \{A \subset \tilde{\mathcal{A}} : A^s \in \mathcal{F}_t \text{ for all } s \in \mathbb{R}_{\geq 0}\}$ .
- The probability measures  $\tilde{\mathbb{P}}^{(r,x)}$  on  $(\tilde{\Omega}, \tilde{\mathcal{A}})$ , which for  $A \in \tilde{\mathcal{A}}$  are defined by

$$\tilde{\mathbb{P}}^{(r,x)}(A) := \mathbb{P}^{(r,x)}(A^r).$$

Note that the probability measure  $\tilde{\mathbb{P}}^{(r,x)}$  is concentrated on the set  $\{r\} \times \Omega$ , i.e.  $\tilde{\mathbb{P}}^{(r,x)}(\{r\} \times \Omega) = 1$ .

- The transition function  $\tilde{P}_t$  on  $(\tilde{E}, \tilde{\mathcal{E}})$  defined by setting for  $B \in \tilde{\mathcal{E}}$

$$\tilde{P}_t((s, x), B) := P_{s, s+t}(x, B^{s+t}) = \mathbb{P}^{(s,x)}(X_t \in B^{s+t}).$$

- If there exist shift operators  $\theta_t$  they are extend by  $\tilde{\theta}_t(s, \omega) := (s + t, \theta_t(\omega))$ .

**Lemma 1.8:** *Let  $(X, \mathcal{F}, \{\mathbb{P}^{(r,x)}\})$  be a time-inhomogeneous Markov process on  $(\Omega, \mathcal{A})$  with state space  $(E, \mathcal{E})$ .*

- $(\tilde{X}, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(r,x)}\})$  is a time-homogeneous Markov process on  $(\tilde{\Omega}, \tilde{\mathcal{A}})$  with state space  $(\tilde{E}, \tilde{\mathcal{E}})$  and transition function  $\{\tilde{P}_t\}$ .
- $(X, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(r,x)}\})$  is a time-inhomogeneous Markov process on  $(\tilde{\Omega}, \tilde{\mathcal{A}})$  with state space  $(E, \mathcal{E})$  and transition function  $\{P_{s,t}\}$ .

*Proof.* That  $\{\tilde{P}_t\}$  is a transition function can be easily checked. The crucial point is measurability of  $\tilde{P}_t$ . For fixed  $B \in \tilde{\mathcal{E}}$  and any real Borel set  $C$

$$\Lambda := \tilde{P}_t^{-1}(\cdot, B)(C) = \bigcup_{s \in \mathbb{R}_{\geq 0}} \{s\} \times P_{s, s+t}^{-1}(\cdot, B^{s+t})(C)$$

satisfies  $\Lambda^s = P_{s, s+t}^{-1}(\cdot, B^{s+t})(C) \in \mathcal{E}$  and therefore  $\Lambda \in \tilde{\mathcal{E}}$ . Similar considerations show that the process  $\tilde{X}$  is  $\tilde{\mathcal{F}}$ -adapted. To show the Markov property of  $\tilde{X}$  one has to check that for  $B \in \tilde{\mathcal{E}}$

$$\tilde{\mathbb{E}}^{(r,x)}[\mathbb{I}\{\tilde{X}_{s+t} \in B\} | \tilde{\mathcal{F}}_s] = \tilde{\mathbb{P}}^{(r,x)}(\tilde{X}_{s+t} \in B | \tilde{\mathcal{F}}_s) = \tilde{P}_t(\tilde{X}_s, B).$$

$A \in \tilde{\mathcal{F}}_s$  implies  $A^r \in \mathcal{F}_s$  and  $\tilde{\mathbb{P}}^{(r,x)}(A) = \tilde{\mathbb{P}}^{(r,x)}(\{r\} \times A^r) = \mathbb{P}^{(r,x)}(A^r)$ . Together with the Markov property of  $X$  one obtains

$$\begin{aligned} \int_A \tilde{P}_t(\tilde{X}_s(\tilde{\omega}), B) \tilde{\mathbb{P}}^{(r,x)}(d\tilde{\omega}) &= \int_{A^r} P_{r+s, r+s+t}(X_s(\omega), B^{r+s+t}) \mathbb{P}^{(r,x)}(d\omega) \\ &= \int_{A^r} \mathbb{P}^{(r,x)}(X_{s+t} \in B^{r+s+t} | \mathcal{F}_s) \mathbb{P}^{(r,x)}(d\omega) \\ &= \int_{A^r} \mathbb{I}\{X_{s+t}(\omega) \in B^{r+s+t}\} \mathbb{P}^{(r,x)}(d\omega) \\ &= \int_A \mathbb{I}\{\tilde{X}_{s+t}(\tilde{\omega}) \in B\} \tilde{\mathbb{P}}^{(r,x)}(d\tilde{\omega}). \end{aligned}$$

This shows that  $(\tilde{X}, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(r,x)}\})$  is a time-homogeneous Markov-process. It is easy to check that  $(X, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(r,x)}\})$  is a time-inhomogeneous Markov process with transition function  $\{P_{s,t}\}$ .  $\square$

To get measurability in  $(\tilde{E}, \tilde{\mathcal{E}})$  of the map  $(s, x) \mapsto \tilde{P}_t((s, x), B)$  for  $B \in \tilde{\mathcal{E}}$  we have to consider the very large  $\sigma$ -algebra  $\tilde{\mathcal{E}}$  on  $\tilde{E}$  (and then also  $\tilde{\mathcal{A}}$  on  $\tilde{\Omega}$ ). If  $(s, x) \mapsto P_{s,s+t}(x, B)$  is  $\mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E}$ -measurable for all  $B \in \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E}$  the space-time process  $\tilde{X}$  can also be realized with respect to the smaller product  $\sigma$ -algebras  $\tilde{\mathcal{E}}' := \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E}$  on  $\tilde{E}$  and  $\tilde{\mathcal{A}}' := \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{A}$  on  $\tilde{\Omega}$  (note that  $\tilde{\mathcal{E}}' \subset \tilde{\mathcal{E}}$  and  $\tilde{\mathcal{A}}' \subset \tilde{\mathcal{A}}$ ). In this case  $\tilde{\mathbb{P}}^{(r,x)} = \delta_r \otimes \mathbb{P}^{(r,x)}$  and for  $B = [a, b] \times B'$ ,  $B' \in \mathcal{E}$ , the definition of the transition function reduces to

$$\tilde{P}_t((s, x), B) = \tilde{P}_t((s, x), [a, b] \times B') = P_{s,s+t}(x, B') \mathbb{I}_{[a,b]}(s+t). \quad (1.6)$$

If we consider a Markov process in the sense of Definition 1.3 the transformation can be simplified. In this case the  $\tilde{E}$ -valued process  $\tilde{X}_t(\omega) = (t, X_t(\omega))$  is a time-homogeneous Markov process on the original filtered probability space. The remarks regarding the measurability of the transition function are still true.

**Definition 1.9:** A transition function  $\{P_{s,t}\}$  is called *stochastically continuous* if for all  $x \in E$

$$P_{s,S}(x, \cdot) \xrightarrow{d} P_{t,T}(x, \cdot), \text{ whenever } (s, S) \rightarrow (t, T) \text{ with } 0 \leq s \leq S \text{ and } 0 \leq t \leq T.$$

For a Markov process in the sense of Definition 1.6 there is a whole family of probability measures. Hence stochastic continuity for Markov processes is formulated in terms of weak convergence of the transition function  $\{P_{s,t}\}$ . Note that this definition implies convergence in probability in the following sense. Consider the continuous functions  $f^\epsilon(x) = \frac{|x|}{\epsilon} \wedge 1$ . Under the fixed probability measure  $\mathbb{P}$  of Definition 1.3 or  $\mathbb{P}$  being equal to any of the measures  $\mathbb{P}^{(r,x)}$  in case of Definition 1.6,

$$\begin{aligned} \lim_{s \rightarrow t} \mathbb{P}(|X_t - X_s| > \epsilon) &\leq \lim_{s \rightarrow t} \mathbb{E}[\mathbb{E}[f^\epsilon(X_t - X_s) | \mathcal{F}_s]] \\ &= \mathbb{E}\left[\lim_{s \rightarrow t} \int f^\epsilon(y - X_s) P_{s,t}(X_s, dy)\right] = \mathbb{E}[f^\epsilon(0)] = 0. \end{aligned}$$

It follows that  $X$  satisfies the usual definition of stochastic continuity, i.e. that  $X_s$  converges to  $X_t$  in probability (under  $\mathbb{P}$ ) for  $s \rightarrow t$ .

In later sections we will use completed filtered probability spaces (see Revuz and Yor [42], section 0.3) and modifications of stochastic processes (see Definitions I.1.6 and I.1.7 in Revuz and Yor [42]). We want to clarify how this is to be understood.

**Definition 1.10:** Let  $(A, \mathcal{A})$  be a measurable space with a probability measure  $\mathbb{P}$ . The completion  $\mathcal{A}^{\mathbb{P}}$  is defined as

$$\mathcal{A}^{\mathbb{P}} := \{B \subset A : \exists B_1, B_2 \in \mathcal{A} : B_1 \subset B \subset B_2, \mathbb{P}(B_2 \setminus B_1) = 0\}.$$

The probability measures are extended for  $B \in \mathcal{A}^{\mathbb{P}}$  by setting  $\mathbb{P}(B) := \mathbb{P}(B_2) = \mathbb{P}(B_1)$ . A sub- $\sigma$ -algebra  $\mathcal{F}$  on  $(A, \mathcal{A})$  is completed by

$$\mathcal{F}^{\mathbb{P}} := \{A : \exists B \in \mathcal{F} : \mathbb{P}(A \setminus B \cup B \setminus A) = 0\}.$$

The completed  $\sigma$ -algebra  $\mathcal{A}^{\mathbb{P}}$  contains all subsets of sets of  $\mathbb{P}$ -measure zero and the same is true for  $\mathcal{F}^{\mathbb{P}}$ . Completing  $\sigma$ -algebras (and adding null sets in general) does not destroy the Markov property (see Lemma 1.11 below).

Let  $X$  and  $\tilde{X}$  be two  $E$ -valued stochastic process on  $(\Omega, \mathcal{A}, \mathbb{P})$ .  $X$  is a modification of  $\tilde{X}$  if  $\mathbb{P}(X_t = \tilde{X}_t) = 1$  for all  $t \geq 0$ . If  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  and  $(\tilde{X}, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  are Markov processes in the sense of Definition 1.6 they are modifications of each other, if  $\mathbb{P}^{(s,x)}(X_t = \tilde{X}_t) = 1$  for all  $(s, x)$ .

We are now prepared to formulate the following easy lemma.

**Lemma 1.11:** If  $X$  is Markov process on  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$ , it is also a Markov process on the completed space  $(\Omega, \mathcal{A}^{\mathbb{P}}, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$  with  $\mathcal{F}^{\mathbb{P}} = (\mathcal{F}_t^{\mathbb{P}})_{t \geq 0}$ . Also any modification  $\tilde{X}$  of  $X$  satisfying that  $X_s = \Delta$  implies  $\tilde{X}_t = \Delta$  for all  $t > s$  is a Markov process on  $(\Omega, \mathcal{A}^{\mathbb{P}}, \mathcal{F}^{\mathbb{P}}, \mathbb{P})$ . Furthermore the transition functions are the same.

*Proof.*  $X$  is still  $\mathcal{F}^{\mathbb{P}}$ -adapted and conditional expectations are not influenced by adding sets of measure 0. Since the filtration contains all sets of measure 0, a modification is still adapted. Since  $\mathbb{P}(\{\tilde{X}_t \neq X_t\}) = 0$  it is still a Markov process.  $\square$

For the rest of this thesis we write  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  if we explicitly consider a Markov process in the sense of Definition 1.6 while a Markov process  $X$  can refer to either of the two definitions. If a Markov process  $X$  is to be understood in the sense of Definition 1.3, we explicitly specify the underlying probability space  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$ .

## 1.2 Affine processes

From now on consider a state space  $E_{\Delta}$ , where  $E$  is a closed subset of the  $d$ -dimensional real vector space  $V$  (think of  $\mathbb{R}^d$ , but it can also be a different space, like the space of positive definite matrices on  $\mathbb{R}^{n \times n}$ , which is isomorphic to  $\mathbb{R}^{n(n+1)/2}$ ).

**Assumption 1:**  $E$  contains an affine basis, i.e.  $d + 1$  elements  $x^0, \dots, x^d$ , such that  $x^1 - x^0, \dots, x^d - x^0$  are linearly independent.

For  $E$  consider the Borel  $\sigma$ -algebra  $\mathcal{E} = \mathcal{B}(E)$ . We use the convention that  $\|\Delta\| = \infty$ . Define

$$\mathcal{U} = \{u \in V + \mathbf{i}V \mid x \mapsto e^{\langle u, x \rangle} \text{ is a bounded function in } E\} \quad (1.7)$$

and for  $u \in \mathcal{U}$  the bounded functions  $f_u(x) := e^{\langle u, x \rangle}$  for  $x \in E$ ,  $f_u(\Delta) := 0$ .

**Definition 1.12:** A time-inhomogeneous transition function  $\{P_{s,t}\}$  is called affine if there exist functions  $\Phi_{s,t} : \mathcal{U} \rightarrow \mathbb{C}$  and  $\psi_{s,t} : \mathcal{U} \rightarrow V + \mathbf{i}V$  for  $0 \leq s \leq t$  such that for all  $u \in \mathcal{U}$ ,  $x \in E$

$$P_{s,t}f_u(x) = \Phi_{s,t}(u)e^{\langle \psi_{s,t}(u), x \rangle}. \quad (1.8)$$

A time-inhomogeneous Markov process is affine if its transition function is affine. If the transition function is time-homogeneous (i.e.  $P_{s,t} = P_{t-s}$ ), then

$$P_t f_u(x) = \Phi_t(u)e^{\langle \psi_t(u), x \rangle} \quad (1.9)$$

with  $\Phi_t(u) := \Phi_{0,t}(u)$  and  $\psi_t(u) := \psi_{0,t}(u)$ .

The affine property of a process is a property of the transition operator. Hence it does not depend on the used definition of a Markov process (Definition 1.3 versus Definition 1.6). The restrictions on the state space are not really restrictions. The affine property automatically extends to the closure of a set and one can always pass to a lower dimensional ambient vector space  $V$ , so that Assumption 1 is fulfilled (see Cuchiero and Teichmann [10]).

As long as  $\Phi_{s,t}(u) \neq 0$  we can find a function  $\phi$  (which is unique only up to adding multiples of  $2\pi\mathbf{i}$ ), such that  $\Phi_{s,t}(u) = e^{\phi_{s,t}(u)}$ . In this case equation (1.8) can be rewritten as

$$P_{s,t}f_u(x) = e^{\phi_{s,t}(u) + \langle \psi_{s,t}(u), x \rangle}. \quad (1.10)$$

This exponential affine form is one motivation for the name affine process. Note that if  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  is an affine process, then for  $r, s, t \geq 0$

$$\mathbb{E}^{(r,x)} [e^{\langle u, X_{t+s} \rangle} \mid \mathcal{F}_s] = \Phi_{r+s, r+s+t}(u)e^{\langle \psi_{r+s, r+s+t}(u), X_s \rangle},$$

which for  $s = 0$  yields

$$\mathbb{E}^{(r,x)} [e^{\langle u, X_t \rangle}] = \Phi_{r, r+t}(u)e^{\langle \psi_{r, r+t}(u), x \rangle}.$$

The exponential functions in the above equations are meaningfully defined with the convention  $f(\Delta) = 0$ , i.e.  $e^{\langle u, \Delta \rangle} = 0$ . Equations like this will be used frequently throughout this chapter.

Time-homogeneous affine process have been widely studied in the literature, see for example the pioneering work of Duffie et al. [16] and the articles Cuchiero and Teichmann [10], Keller-Ressel et al. [37], which are similar in spirit to this treatment. By Lemma 1.8 a time-inhomogeneous Markov process can be extended to the time-homogeneous space-time process. The affine property carries over to the space-time process only partially.

**Lemma 1.13:** *Let  $X$  be a Markov process with affine transition function. Then the time-homogeneous transition function  $\{\tilde{P}_t\}$  of the space-time process  $\tilde{X} := (\Theta, X)$  from Lemma 1.8 satisfies for  $u_0 \in \mathbb{C}_{\leq 0}$ ,  $u \in \mathcal{U}$  and  $f_{(u^0, u)}(s, x) := e^{u^0 s} f_u(x)$*

$$\tilde{P}_t f_{(u^0, u)}(s, x) = e^{u^0 t} \Phi_{s, s+t}(u) e^{u^0 s + \langle \psi_{s, s+t}(u), x \rangle}. \quad (1.11)$$

*Proof.* Follows directly from the definition of the time-homogeneous transition function  $\tilde{P}_t$  in Lemma 1.8, i.e.  $\tilde{P}_t f(s, x) := P_{s, s+t} f^{s+t}(x)$  with  $f^{s+t}(x) = f(s+t, x)$ .  $\square$

*Remark:*  $\tilde{X}$  is almost an affine process. The transition function  $\{\tilde{P}_t\}$  is exponentially affine in  $x$ , but not in the (basically deterministic) time component  $s$ , since  $\Phi_{s, s+t}$  and  $\psi_{s, s+t}$  depend on the state  $(s, x)$  if  $\{P_{s, t}\}$  is not time-homogeneous.

So far the functions  $\Phi$  and  $\psi$  in Definition 1.12 are not necessarily uniquely defined and can be arbitrary irregular. If we assume that the transition function is stochastically continuous, the following lemma yields continuity of  $(s, t, u) \mapsto P_{s, t} f_u(x)$ , which we later want to transfer to  $\Phi$  and  $\psi$ . Note that stochastic continuity is sometimes part of the definition of an affine process (see e.g. Cuchiero and Teichmann [10]). For an affine process consider

$$\begin{aligned} \sigma(s, u) &:= \inf\{t > s : \Phi_{s, t}(u) = 0\}, \\ o(T, u) &:= \sup\{t < T : \Phi_{t, T}(u) = 0\} \vee 0, \end{aligned} \quad (1.12)$$

where we use the conventions  $\inf\{\emptyset\} = \infty$  and  $\sup\{\emptyset\} = -\infty$ .  $\sigma$  is used for going forward in time, while  $o$  is needed for going backward in time. For  $k \in \mathbb{N}$  let

$$\begin{aligned} \mathcal{U}_k &:= \{u \in V + iV : \sup_{x \in E} |e^{\langle u, x \rangle}| \leq k\}, \\ \mathcal{Q}_k &:= \{(t, T, u) : u \in \mathcal{U}_k, o(T, u) < t \leq T\} \\ &= \{(t, T, u) : 0 \leq t \leq T, u \in \mathcal{U}_k, \Phi_{v, T}(u) \neq 0 \text{ for } t \leq v \leq T\}, \end{aligned}$$

and  $\mathcal{U} = \cup_k \mathcal{U}_k$  and  $\mathcal{Q} := \cup_k \mathcal{Q}_k$ .

**Lemma 1.14:** *Let  $\{P_{s, t}\}$  be stochastically continuous. Then*

(i)  $\Phi$  vanishes on  $\{(t, T, u), 0 \leq t \leq T, u \in \mathcal{U}\} \setminus \mathcal{Q}$ , so

$$\mathcal{Q} = \{(t, T, u) : 0 \leq t \leq T, u \in \mathcal{U}, \Phi_{t, T}(u) \neq 0\}.$$

(ii) The function  $\psi$  maps  $\mathcal{Q}$  into  $\mathcal{U}$ .

(iii) For  $x \in E$  the functions  $f^x(t, T, u) := P_{t, T} f_u(x)$  are continuous on

$$\{(t, T, u), 0 \leq t \leq T, u \in \mathcal{U}_k\}$$

for all  $k \in \mathbb{N}$ . If  $0 \in E$ , this is also true for  $\Phi$ .

(iv)  $\mathcal{Q}$  is open and  $\sigma(t, u) > t$ ,  $o(T, u) < T$  for any  $u \in \mathcal{U}$ .

*Remark:* Note that

$$\exists x \in E : f^x(t, T, u) = 0 \Leftrightarrow \Phi_{t,T}(u) = 0 \Leftrightarrow \forall x \in E_\Delta : f^x(t, T, u) = 0. \quad (1.13)$$

*Proof.* If  $\Phi_{t,T}(u) = 0$  we get that for  $s \leq t \leq T$ ,  $x \in E$

$$f^x(s, T, u) = P_{s,T}f_u(x) = P_{s,t}P_{t,T}f_u(x) = 0$$

so also  $\Phi_{s,T}(u) = 0$ . Then  $\Phi_{s,T}(u) = 0$  for all  $s < o(T, u)$ . By stochastic continuity also  $\Phi_{o(T,u),T}(u) = 0$  and (i) follows.

To get (ii) note that for  $u \in \mathcal{U}_k$  and  $x \in E$

$$|\Phi_{t,T}(u)e^{\langle \psi_{t,T}(u), x \rangle}| = |P_{t,T}f_u(x)| \leq \int_E |f_u(\xi)|P_{t,T}(x, d\xi) \leq \int_E KP_{t,T}(x, d\xi) = K.$$

If  $(t, T, u) \in \mathcal{Q}$  we can divide by  $|\Phi_{t,T}(u)|$  and  $|e^{\langle \psi_{t,T}(u), x \rangle}| \leq K|\Phi_{t,T}(u)|^{-1}$  for all  $x \in E$ . Hence  $\psi_{t,T}(u) \in \mathcal{U}$ .

Let  $x \in E$  and let  $(t_n, T_n, u_n) \rightarrow (t, T, u)$  be a converging sequence, where  $u_n, u \in \mathcal{U}_k$  and  $0 \leq t_n \leq T_n$ ,  $0 \leq t \leq T$ . There exists a function  $\rho : E \rightarrow [0, 1]$  with compact support such that

$$P_{t,T}(1 - \rho)(x) < \epsilon.$$

Choose  $N_1$  such that

$$|f_{u_n}(\xi) - f_u(\xi)| = |e^{\langle u_n, \xi \rangle} - e^{\langle u, \xi \rangle}| < \epsilon, \quad \forall n \geq N_1, \xi \in \text{supp}(\rho).$$

By stochastic continuity there is  $N_2$  such that for all  $n \geq N_2$

$$\begin{aligned} P_{t_n, T_n}(1 - \rho)(x) &< 2\epsilon, \\ |P_{t_n, T_n}f_u(x) - P_{t, T}f_u(x)| &< \epsilon. \end{aligned}$$

Then for all  $n \geq \max\{N_1, N_2\}$

$$\begin{aligned} |P_{t_n, T_n}f_{u_n}(x) - P_{t, T}f_u(x)| &\leq |P_{t_n, T_n}(\rho(f_{u_n} - f_u))(x)| \\ &\quad + |P_{t_n, T_n}((1 - \rho)(f_{u_n} - f_u))(x)| \\ &\quad + |P_{t_n, T_n}f_u(x) - P_{t, T}f_u(x)| \\ &< \epsilon + 2k\epsilon + \epsilon = 2(k + 1)\epsilon. \end{aligned}$$

Hence  $f^x(t, T, u)$  is continuous on  $\{(t, T, u) : 0 \leq t \leq T, u \in \mathcal{U}_k\}$  for each  $k \in \mathbb{N}$  and  $x \in E$ . If  $0 \in E$ ,  $\Phi_{t,T}(u) = f^0(t, T, u)$  is continuous and we get (iii). Since  $\Phi_{t,t}(u) \neq 0$  the continuity of  $f^x$  and (1.13) implies (iv).  $\square$

We want to use the continuity in Lemma 1.14 (ii) to get unique choices of  $\Phi$  and  $\psi$ . This is possible because of Assumption 1.

**Lemma 1.15:** *Let  $\{P_{s,t}\}$  be stochastically continuous. Then  $\psi$  is uniquely defined on  $\mathcal{Q}$  by requiring that  $\psi_{t,T}(u)$  is continuous on  $\mathcal{Q}_k$  for all  $k \in \mathbb{N}$  and  $\psi_{0,0}(0) = 0$ . In this case also  $\Phi$  is uniquely defined and continuous on  $\mathcal{Q}_k$  for all  $k \in \mathbb{N}$ . Furthermore, there is a unique function  $\phi_{t,T}(u)$  on  $\mathcal{Q}$ , which is continuous on  $\mathcal{Q}_k$  and satisfies  $\phi_{0,0}(0) = 0$  and  $e^{\phi_{t,T}(u)} = \Phi_{t,T}(u)$ .*

For the proof we borrow heavily from Keller-Ressel et al. [37].

*Proof.* Define the sets  $K_n := \{(t, T, u) : u \in \mathcal{U}_n, \|u\| \leq n, t \in [o(T, u) + \frac{1}{n}, T]\}$ , so that  $K_n \subset \mathcal{Q}_n$  and  $\lim_{n \rightarrow \infty} K_n = \mathcal{Q}$ . We claim that every set  $K_n$  is contractible to zero. Let  $\gamma = (t(r), T(r), u(r))_{0 \leq r \leq 1}$  be a continuous curve in  $K_n$ . Then for  $0 \leq \alpha \leq 1$  the curves  $\gamma_\alpha = (t(r) + (1 - \alpha)(T(r) - t(r)), T(r), u(r))_{0 \leq r \leq 1}$  depend continuously on  $\alpha$  and stay in  $K_n$  for each  $\alpha$ . Furthermore,  $\gamma_1 = \gamma$  and  $\gamma_0 = (T(r), T(r), u(r))_{0 \leq r \leq 1}$ . So any continuous curve in  $K_n$  is homotopically equivalent to a continuous curve in the convex set  $\mathbb{R}_{\geq 0} \times \mathcal{U}$ , where all continuous curves are contractible to zero. This proves that  $K_n$  is contractible to zero. Let  $H_n : [0, 1] \times K_n \rightarrow K_n$  be such a continuous contraction and fix  $x \in E$ . Since  $K_n$  is compact, we have  $\lim_{\alpha \rightarrow \beta} \|f^x(H_n(\alpha, \cdot)) - f^x(H_n(\beta, \cdot))\|_\infty = 0$ . Hence  $f^x \circ H_n$  is a continuous curve in  $C_b(K_n)$  from  $f^x|_{K_n}$  to the constant function 1. Here  $C_b(K_n)$  denotes the Banach space of bounded continuous functions on  $K_n$ . By Theorem 1.3 in Bucchianico [7] there exists a continuous logarithm  $g_n^x \in C_b(K_n)$ , i.e.  $f^x(t, T, u) = e^{g_n^x(t, T, u)}$  for all  $(t, T, u) \in K_n$ . We next show, that  $g_n^x$  is uniquely defined by setting  $g_n^x(0, 0, 0) = 0$ . Assume that  $\tilde{g}_n^x(t, T, u)$  is another such continuous logarithm, then  $g_n^x(t, T, u) = \tilde{g}_n^x(t, T, u) + 2\pi i \alpha(t, T, u)$  with  $\alpha(t, T, u) \in \mathbb{Z}$ . Then  $\alpha(t, T, u)$  is continuous on  $K_n$ , which is contractible to 0 and hence also connected. Hence  $\alpha(t, T, u)$  cannot depend on  $(t, T, u)$ . Inserting  $(0, 0, 0)$  gives  $\alpha(t, T, u) = 0$  and the uniqueness of  $g_n^x$ .

For arbitrary  $m \leq n$  we have

$$g_m^x(t, T, u) = g_n^x(t, T, u) + 2\pi i \alpha^x(t, T, u) \quad \text{for all } (t, T, u) \in K_m,$$

where  $\alpha^x(t, T, u)$  is a continuous function from  $K_m$  to  $\mathbb{Z}$  satisfying  $\alpha^x(0, 0, 0) = 0$ . Hence  $g_m^x(t, T, u) = g_n^x(t, T, u)$  for all  $(t, T, u) \in K_m$ . This shows that  $g_n^x$  extends  $g_m^x$ , hence there is a unique function  $g^x : \mathcal{Q} \rightarrow \mathbb{C}$  such that  $g^x(0, 0, 0) = 0$  and

$$e^{g^x(t, T, u)} = f^x(t, T, u) = \Phi_{t,T}(u) e^{\langle \psi_{t,T}(u), x \rangle}. \quad (1.14)$$

Note that like  $f^x$ ,  $g^x$  is then continuous on each  $\mathcal{Q}^k$ .

Fix a point  $x^0$  of an affine basis  $x^0, \dots, x^d$ . Then on  $\mathcal{Q}_k$

$$e^{\langle \psi_{t,T}(u), x - x^0 \rangle} = e^{g^x(t, T, u) - g^{x^0}(t, T, u)},$$

so  $\langle \psi_{t,T}(u), x - x^0 \rangle = g^x(t, T, u) - g^{x^0}(t, T, u) + 2\pi i \alpha^x(t, T, u)$ , where  $\alpha^x(t, T, u) \in \mathbb{Z}$ .  $\alpha^x(0, 0, 0) = 0$ . Continuity of  $\psi$ ,  $g^x$  and  $g^{x^0}$  on  $\mathcal{Q}_k$  gives continuity of  $\alpha^x(t, T, u)$ . Hence  $\alpha^x = 0$  and  $\langle \psi_{t,T}(u), x - x^0 \rangle$  is uniquely defined on  $\mathcal{Q}$  through the unique functions  $g^x, x \in E$ . By Assumption 1 this is also true for  $\psi$ . On  $\mathcal{Q}$  set  $\phi_{t,T}(u) = g^x(t, T, u) - \langle \psi_{t,T}(u), x \rangle$  which is equal for all  $x \in E$  and then also uniquely defined. Then  $\Phi_{t,T}(u) = e^{\phi_{t,T}(u)}$  and  $\Phi$  is also uniquely defined and continuous on  $\mathcal{Q}_k$ .  $\square$

From now when talking about continuous functions  $\Phi$  and  $\psi$ , we always mean that they are of the form described in Lemma 1.15. It suffices to require weaker properties of  $\psi$  to be sure that it is in fact a continuous version. These properties need to be sufficient to guarantee that  $\alpha^x$  in the proof of Lemma 1.15 does not depend on  $(t, T, u)$ . E.g. one could relax this by requiring that for each  $(t_0, T_0, u_0)$  and  $x \in E$  there is some open neighborhood where  $|\langle \mathfrak{S}\psi_{t,T}(u), x \rangle - \langle \mathfrak{S}\psi_{t_0, T_0}(u_0), x \rangle| < \pi$  as suggested in Cuchiero and Teichmann [10]. Note that as in the time-homogeneous case (see Cuchiero and Teichmann [10], Keller-Ressel et al. [37]) we do not get the existence of continuous versions, but only uniqueness. Under a strengthening of the assumptions on the state space, we get that the function  $\psi$  is always continuous.

**Lemma 1.16:** *Assume that  $E$  has a connected component, which contains an affine basis. If  $\{P_{s,t}\}$  is stochastically continuous, then any function  $\psi$  satisfying (1.8) is continuous on  $\mathcal{Q}_k$  with  $\psi_{0,0}(0) = 0$ .*

*Proof.* Let  $x^0, \dots, x^d$  be an affine basis in a connected component. Define the function

$$h(t, T, u, x) := e^{\langle \psi_{t,T}(u), x - x^0 \rangle}, \quad x \in V.$$

For  $x \in E$  this can be written as  $h(t, T, u, x) = \frac{f^x(t, T, u)}{f^{x^0}(t, T, u)}$ . By Lemma 1.14 iii)  $h$  is continuous in  $(t, T, u) \in \mathcal{Q}_k$  for each  $x \in E$ . Furthermore it is continuous in  $x$  for each  $(t, T, u)$  and even jointly continuous on  $\mathcal{Q}_k \times E$ . Set  $E_n = \{x \in E, \|x\| \leq n\}$ . With  $H_n$  as in the proof of Lemma 1.15,  $\alpha \rightarrow h(H_n(\alpha, (t, T, u)), x)$  is a continuous curve in  $C_b(K_n \times E_n)$  from  $h|_{K_n \times E_n}$  to the constant function 1. By Theorem 1.3 in Bucchianico [7] there is a continuous function  $g^n(t, T, u, x)$  on  $K_n \times E_n$ , such that  $h(t, T, u, x) = e^{g^n(t, T, u, x)}$ . Setting  $g(0, 0, 0, x) = 0$  for all  $x \in E_n$  uniquely defines  $g^n$ . As in the proof of Lemma 1.15 we can extend this to get a unique function  $g(t, T, u, x)$  that is continuous on  $\mathcal{Q}_k \times E$ , satisfies  $g(0, 0, 0, x) = 0$  for all  $x \in E$  and  $h(t, T, u, x) = e^{g(t, T, u, x)}$ .

Then  $\langle \psi_{t,T}(u), x - x^0 \rangle = g(t, T, u, x) - 2\pi i \alpha(t, T, u, x)$  with  $\alpha(t, T, u, x) \in \mathbb{Z}$ . Since the left hand side is continuous in  $x$  also  $\alpha$  is continuous in  $x$ . Thus  $\alpha$  is constant in  $x$ , on every connected component. Denote by  $\bar{\alpha}(t, T, u)$  the value on the connected component containing  $x^0, \dots, x^d$ . Then

$$\langle \psi_{t,T}(u), x^i - x^0 \rangle = g(t, T, u, x^i) - 2\pi i \bar{\alpha}(t, T, u), \quad 0 \leq i \leq d.$$

The case  $i = 0$  proves that  $2\pi i \bar{\alpha}(t, T, u) = g(t, T, u, x^0)$  is continuous on  $\mathcal{Q}_k$ . Then also  $\langle \psi_{t,T}(u), x^i - x^0 \rangle$  is continuous on  $\mathcal{Q}_k$ . Since  $\{x^0, \dots, x^d\}$  is an affine basis,  $\psi_{t,T}(u)$  is continuous on  $\mathcal{Q}_k$ . Finally,  $\langle \psi_{0,0}(0), x^i - x^0 \rangle = g(0, 0, 0, x^i) - 2\pi i \bar{\alpha}(0, 0, 0) = 0$  for  $1 \leq i \leq d$  implies  $\psi_{0,0}(0) = 0$ .  $\square$

This shows that for almost every state space stochastic continuity implies continuity of  $\Phi$  and  $\psi$ . Next we show that essentially also the reverse is true. We first state the following well-known fact.

**Lemma 1.17:**  $\mathcal{B}(\mathbb{R}^d) = \sigma(C_{exp}) := \sigma(\{f^{-1}(A) : A \in \mathcal{B}(\mathbb{C}), f \in C_{exp}\})$ , where

$$C_{exp} = \left\{ \sum_{k=1}^n a_k f_{u_k}(x), a_k \in \mathbb{C}, u_k \in \mathbf{i}\mathbb{R}^d, n \in \mathbb{N} \right\}.$$

*Sketch of proof.* To simplify notation assume  $d = 1$ . Take some interval  $[a, b]$  and a modified indicator function of this interval, which is 0 outside of  $[a, b]$ , 1 on  $(a, b)$  and  $1/2$  for  $\{a, b\}$ . This function can be pointwise approximated by  $C^\infty$ -functions. By Fourier methods such an approximating  $C^\infty$ -function  $g$  can be written as  $g(x) = \frac{1}{2\pi} \int_{\mathbf{i}\mathbb{R}} \hat{g}(z) e^{zx} dz$ . Cutting off the integral and approximating it with Riemann sums gives that it can be pointwise approximated by functions in  $C_{exp}$ , which then proves the result.  $\square$

**Lemma 1.18:** Let  $(t, T, u) \mapsto P_{t,T} f_u(x)$  be continuous on  $\{(t, T, u) : 0 \leq t \leq T, u \in \mathcal{U}_1\}$  for each  $x \in E$ . Then  $\{P_{t,T}\}$  is stochastically continuous.

*Proof.*  $(t, T) \mapsto P_{t,T} f_u(x)$  is continuous for all  $x \in E$  and  $u \in \mathcal{U}_1$ . Hence this also holds for  $f \in C_{exp}$ . Let  $C_b(E_\Delta)$  denote the bounded measurable functions  $f : E_\Delta \rightarrow \mathbb{C}$  and define the vector space

$$\mathcal{H} := \{f \in C_b(E_\Delta) : (t, T) \mapsto P_{t,T} f(x) \text{ is continuous for all } x \in E\},$$

which is a monotone class containing the constant function.  $C_{exp} \subset \mathcal{H}$  is closed under multiplication. An application of the functional monotone class theorem (Revuz and Yor [42], Theorem 0.2.2) together with Lemma 1.17 gives that  $\mathcal{H}$  contains all bounded measurable functions, which gives stochastic continuity of  $\{P_{t,T}\}$ .  $\square$

**Corollary 1.19:** Let  $\{P_{s,t}\}$  be an affine transition function on a state space  $E_\Delta$  which contains 0 and has a connected component containing an affine basis. Then the following are equivalent.

- $\{P_{s,t}\}$  is stochastically continuous.
- $\psi$  is continuous on  $\mathcal{Q}_k$  with  $\psi_{0,0}(0) = 0$  and  $\Phi$  and  $(t, T, u) \mapsto P_{t,T} f_u(x)$  are continuous on  $\{(t, T, u) : 0 \leq t \leq T, u \in \mathcal{U}_k\}$  for all  $x \in E$  and  $k \in \mathbb{N}$ .

*Proof.* Lemma 1.18 gives one direction. Lemma 1.16 yields the continuity of  $\psi$  on  $\mathcal{Q}_k$  and  $\psi_{0,0}(0) = 0$ . Lemma 1.14 iii) with the assumption  $0 \in E$  the rest.  $\square$

This shows that stochastic continuity and continuity of  $\Phi$  and  $\psi$  are mostly equivalent. Instead of stochastic continuity of the transition function we consider the following definition.

**Definition 1.20:** An affine transition function  $\{P_{t,T}\}$  is called a continuously affine transition function if  $\psi$  and  $\phi$  are continuous on  $\mathcal{Q}_k$  with  $\psi_{0,0}(0) = 0$ .  $\Phi$  and  $(t, T, u) \mapsto P_{t,T} f_u(x)$  are continuous on  $\{(t, T, u) : 0 \leq t \leq T, u \in \mathcal{U}_k\}$  for all  $x \in E$  and  $k \in \mathbb{N}$ . A Markov process is continuously affine, if its transition function is continuously affine.

**Lemma 1.21:** Let  $\{P_{t,T}\}$  be a continuously affine transition function.

(i)  $\Phi$ ,  $\phi$  and  $\psi$  satisfy the following semi-flow equations. For  $s \leq t \leq T$

$$\begin{aligned}\Phi_{s,T}(u) &= \Phi_{s,t}(\psi_{t,T}(u))\Phi_{t,T}(u), & \text{for all } u \in \mathcal{U}, \\ \psi_{s,T}(u) &= \psi_{s,t}(\psi_{t,T}(u)), & \text{if } \Phi_{s,T}(u) \neq 0, \\ \phi_{s,T}(u) &= \phi_{t,T}(u) + \phi_{s,t}(\psi_{t,T}(u)) & \text{if } \Phi_{s,T}(u) \neq 0.\end{aligned}\tag{1.15}$$

(ii) Let  $u \in \mathcal{U} \cap V$ . Then  $\Im\Phi_{t,T}(u) = 0$  for  $0 \leq t \leq T$  and  $\Im\psi_{t,T}(u) = 0$ ,  $\Im\phi_{t,T}(u) = 0$  for  $o(T, u) < t \leq T$ .

(iii) For all  $t \geq 0$  it holds that  $\Phi_{t,t}(u) = 1$ ,  $\phi_{t,t}(u) = 0$  and  $\psi_{t,t}(u) = u$ .

*Proof.* Regarding (i), if  $\Phi_{s,T}(u) \neq 0$ , then by Lemma 1.14 i) also  $\Phi_{t,T}(u) \neq 0$ , so by Lemma 1.14 ii)  $\psi_{t,T}(u) \in \mathcal{U}$  and

$$\begin{aligned}\Phi_{s,T}(u)e^{\langle \psi_{s,T}(u), x \rangle} &= P_{s,T}f_u(x) = P_{s,t}P_{t,T}f_u(x) = \Phi_{t,T}(u)P_{s,t}f_{\psi_{t,T}(u)}(x) \\ &= \Phi_{t,T}(u)\Phi_{s,t}(\psi_{t,T}(u))e^{\langle \psi_{s,t}(\psi_{t,T}(u)), x \rangle}.\end{aligned}\tag{1.16}$$

Hence also  $\Phi_{s,t}(\psi_{t,T}(u)) \neq 0$  and  $(s, t, \psi_{t,T}(u)) \in \mathcal{Q}$ . Thus we can write (1.16) as

$$e^{\langle \phi_{s,T}(u) + \langle \psi_{s,T}(u), x \rangle \rangle} = e^{\langle \phi_{t,T}(u) + \phi_{s,t}(\psi_{t,T}(u)) + \psi_{s,t}(\psi_{t,T}(u)), x \rangle}.$$

The previously discussed uniqueness of the logarithm together with Assumption 1 gives the semiflow equations. If  $\Phi_{s,t}(u) = 0$ , we consider two cases. If  $\Phi_{t,T}(u) = 0$  the semiflow equation for  $\Phi$  holds. Otherwise by Lemma 1.14 ii)  $\psi_{t,T}(u) \in \mathcal{U}$  and (1.16) holds. Since  $\Phi_{t,T}(u)e^{\langle \psi_{s,t}(\psi_{t,T}(u)), x \rangle} \neq 0$ , it follows that  $\Phi_{s,t}(\psi_{t,T}(u)) = 0$ . This proves the semiflow equation for  $\Phi$  in the case  $\Phi_{s,t}(u) = 0$ .

For (ii) note that for  $u \in \mathcal{U}_k \cap V$ ,  $o(T, u) < t \leq T$

$$e^{\langle \psi_{t,T}(u), x - x^0 \rangle} = \frac{P_{t,T}f_u(x)}{P_{t,T}f_u(x^0)} \in \mathbb{R}_{\geq 0}.$$

Hence  $\langle \Im\psi_{t,T}(u), x - x^0 \rangle = 2\pi i \alpha^x(t, T, u)$  with  $\alpha^x(t, T, u) \in \mathbb{Z}$ . As in the proof of Lemma 1.15 it follows that  $\langle \Im\psi_{t,T}(u), x - x^0 \rangle = 0$ . Hence also  $\Im\psi_{t,T}(u) = 0$ . This implies  $\Im\phi_{t,T}(u) = 0$  and  $\Im\Phi_{t,T}(u) = 0$ . Note that  $\Phi = 0$  outside of  $\mathcal{Q}$  by 1.14 i).

(iii) follows by the same arguments as in ii) since  $P_{t,t}f_u(x) = f_u(x)$  and therefore  $e^{\langle \psi_{t,t}(u) - u, x - x^0 \rangle} = 1$ .  $\square$

The continuity of  $\Phi$  and  $\psi$  can also be used to infer the joint measurability of the transition function  $\{P_{s,t}\}$ .

**Lemma 1.22:** For a continuously affine transition function  $(s, t, x) \mapsto P_{s,s+t}(x, B)$  is  $\mathcal{B}(\mathbb{R}_{\geq 0}^2) \otimes \mathcal{E}_\Delta$ -measurable for every  $B \in \mathcal{E}_\Delta$ .

*Proof.* Let  $C \subset \mathbb{C}$  and  $u \in \mathfrak{i}V$ . By continuity of  $\Phi$  the set  $\mathcal{Q}_u = \{(s, t) \in \mathbb{R}_{\geq 0}^2 : \Phi_{s,s+t}(u) \neq 0\}$  is open and  $\mathcal{B}(\mathbb{R}_{\geq 0}^2)$ -measurable.  $(s, t) \mapsto \psi_{s,t}(u)$  is a continuous function on  $\mathcal{Q}_u$ . Hence also  $(s, t, x) \mapsto P_{s,s+t}f_u(x) = \Phi_{s,s+t}(u)e^{\langle \psi_{s,s+t}(u), x \rangle}$  is continuous on  $\mathcal{Q}_u \times E$  and  $D = \{(s, t, x) \in \mathbb{R}_{\geq 0}^2 \times E : P_{s,s+t}f_u(x) \in C \setminus \{0\}\} \in \mathcal{B}(\mathbb{R}_{\geq 0}^2) \otimes \mathcal{E}_\Delta$ . Therefore

$$\begin{aligned} & \{(s, t, x) \in \mathbb{R}_{\geq 0}^2 \times \mathcal{E}_\Delta : P_{s,s+t}f_u(x) \in C\} \\ &= \begin{cases} 0 \notin C : & D \\ 0 \in C : & (\mathbb{R}_{\geq 0}^2 \setminus \mathcal{Q}_u \times E) \cup (\mathbb{R}_{\geq 0}^2 \times \{\Delta\}) \cup D, \end{cases} \end{aligned}$$

is in  $\mathcal{B}(\mathbb{R}_{\geq 0}^2) \otimes \mathcal{E}_\Delta$ . Hence the function  $(s, t, x) \mapsto P_{s,s+t}f_u(x)$  is  $\mathcal{B}(\mathbb{R}_{\geq 0}^2) \otimes \mathcal{E}_\Delta$ -measurable for every  $f_u$ . This extends to functions  $f \in C_{exp}$ . A monoton class argument as in the proof of Lemma 1.18 gives the result.  $\square$

This implies in particular that  $(s, x) \mapsto \tilde{P}_t((s, x), B)$  is  $\mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E}_\Delta$ -measurable for every fixed  $B \in \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E}_\Delta$ , where  $\tilde{P}_t$  is the transition function of the space-time process introduced in Lemma 1.8. To see this let  $B = [a, b] \times B'$ ,  $B' \in \mathcal{E}_\Delta$  and  $C \in \mathcal{B}([0, 1])$ . Then by Lemma 1.22

$$D = \{(s, x) : P_{s,s+t}(x, B') \in C\} \in \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E}_\Delta.$$

By (1.6)

$$\begin{aligned} \tilde{P}_t^{-1}(\cdot, [a, b] \times B')(C) &= \{(s, x) : P_{s,s+t}(x, B') \mathbb{I}_{[a,b]}(s+t) \in C\} \\ &= \begin{cases} ([a-t, b-t] \times E_\Delta) \cap D & 0 \notin C, \\ ([a-t, b-t]^C \times E_\Delta) \cup (([a-t, b-t] \times E_\Delta) \cap D) & 0 \in C, \end{cases} \end{aligned}$$

and  $\tilde{P}_t^{-1}(\cdot, [a, b] \times B')(C) \in \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E}_\Delta$ . Hence the space-time process of a Markov process with a time-inhomogeneous continuously affine transition function can be realized with respect to the smaller product  $\sigma$ -algebras  $\tilde{\mathcal{E}} = \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E}_\Delta$  and  $\tilde{\mathcal{A}} = \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{A}$  and the filtration  $\tilde{\mathcal{F}}$  given by  $\tilde{\mathcal{F}}_t := \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{F}_t$  (see the remark after Lemma 1.8). Together with Lemma 1.13 we have the following corollary.

**Corollary 1.23:** *Let  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  be a continuously affine Markov process with transition function  $\{P_{s,t}\}$ . Then  $(\tilde{X}, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(s,x)}\})$  is a time-homogeneous Markov process on  $(\tilde{\Omega}, \tilde{\mathcal{A}})$  with state space  $(\tilde{E}, \tilde{\mathcal{E}})$  and the transition function  $\{\tilde{P}_t\}$  satisfying*

$$\tilde{P}_t f_{(u^0, u)}(s, x) = e^{u^0 t} \Phi_{s,s+t}(u) e^{u^0 s + \langle \psi_{s,s+t}(u), x \rangle}.$$

*Furthermore,  $(X, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(s,x)}\})$  is a time-inhomogeneous affine process on  $(\tilde{\Omega}, \tilde{\mathcal{A}})$  with transition function  $\{P_{s,t}\}$ .*

## Conservative affine processes

A transition function is conservative if  $P_{s,t}(x, E) = 1$  for all  $t \geq s$  and  $x \in E$ . By the convention  $f(\Delta) = 0$ ,  $f_0(x) = \mathbb{I}\{x \neq \Delta\}$  this is equivalent to  $P_{s,t}f_0(x) = 1$  for all  $t \geq s$  and  $x \in E$ . For an affine process in the sense of Definition 1.6 this corresponds to  $\mathbb{P}^{(s,x)}(X_t \neq \Delta) = 1$  for all  $s, t \geq 0$  and  $x \in E$ . Note that in this case

$$\mathbb{P}^{(s,x)}(X_t \neq \Delta) = \mathbb{E}^{(s,x)} [e^{\langle 0, X_t \rangle}] = P_{s,s+t}f_0(x) = \Phi_{s,s+t}(0)e^{\langle \psi_{s,s+t}(0), x \rangle}. \quad (1.17)$$

We have the following lemma.

**Lemma 1.24:** *A continuously affine transition function is conservative if and only if  $\Phi_{s,t}(0) = 1$  and  $\psi_{s,t}(0) = 0$  for all  $t \geq s$ .*

*Proof.* If the transition function is conservative, then  $P_{s,t}f_0(x) = 1$  for all  $x \in E$  and  $t \geq s$ . Hence  $\Phi_{s,t}(0) \neq 0$  and so  $e^{\langle \psi_{s,t}(0), x - x^0 \rangle} = 1$ . As in Lemma 1.21 by the continuity of  $\psi$  and Assumption 1 it follows that  $\psi_{s,t}(0) = 0$ . Then clearly  $\Phi_{s,t}(0) = 1$ .  $\square$

## 1.3 Càdlàg modifications

In this section we prove that Markov processes with a continuously affine transition function have a càdlàg modification. For general affine transition functions this is not true as the deterministic process  $X_t = x + \mathbb{I}\{t \leq 1\}$  shows. First we consider affine processes  $X$  which are Markov in the sense of Definition 1.3, where we have a fixed probability measure  $\mathbb{P}$ . We then extend this to Markov processes  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  in the sense of Definition 1.6. For time-homogeneous processes this was proved in Cuchiero and Teichmann [10]. We follow their treatment.

We equip  $E_\Delta$  with the Alexandrov topology, where every open set with a compact complement in  $E$  is declared an open neighborhood of  $\Delta$ . Hence a sequence  $(y_k)$  in  $E_\Delta$  converges to  $\Delta$  if  $\|y_k\| \rightarrow \infty$  (here we use the convention that  $\|\Delta\| = \infty$ ). Furthermore, define the limit through rational points (see e.g. II.61 in Rogers and Williams [43]). For  $y : \mathbb{Q}_{\geq 0} \rightarrow E_\Delta$  we say that  $\lim_{q \downarrow t} y_q$  exists if  $\lim_{q_k \downarrow t} y_{q_k}$  exists in  $E_\Delta$  for every sequence  $q_k \in \mathbb{Q}_{\geq 0}$  with  $q_k \downarrow t$  and is independent of the choice of the sequence  $(q_k)$ . In an analogous way  $\lim_{q \uparrow t} y_q$  is defined.

Let  $\tilde{\Omega} \subset \Omega$  be the set where for each  $t$  the left and right limits of  $X$  through  $\mathbb{Q}$  exist in  $E_\Delta$ . We define the  $E_\Delta$ -valued càdlàg process

$$\tilde{X}_t(\omega) := \begin{cases} \lim_{q \downarrow t} X_q(\omega) & \omega \in \tilde{\Omega}, \\ \Delta & \omega \notin \tilde{\Omega}. \end{cases} \quad (1.18)$$

We show that the process  $\tilde{X}$  is a modification of the original process  $X$  by considering the martingales

$$M_t^{T,u} = \Phi_{t,T}(u)e^{\langle \psi_{t,T}(u), X_t \rangle} = \mathbb{E} [e^{\langle u, X_T \rangle} | \mathcal{F}_t], \quad 0 \leq t \leq T, u \in \mathcal{U},$$

which always have a càdlàg modification.

**Lemma 1.25:** *There is a set  $\Omega_0$  with  $\mathbb{P}(\Omega_0) = 1$ , such that on  $\Omega_0$*

$$\tilde{M}_t^{T,u}(\omega) = \lim_{q \downarrow t} M_q^{T,u}(\omega) = \lim_{q \downarrow t} \Phi_{q,T}(u) e^{\langle \psi_{q,T}(u), X_q \rangle}(\omega) \quad (1.19)$$

*exists for  $0 \leq t < T$  and defines a  $\mathbb{C}$ -valued càdlàg function<sup>1</sup> on  $[0, T]$  for almost all (in the sense of Lebesgue)  $(T, u) \in (0, \infty) \times \mathcal{U}$ .*

*Proof.* By Doob's regularity theorem (Theorem II.65.1, Rogers and Williams [43]) (1.19) defines a càdlàg function  $\mathbb{P}$ -a.s.. However, the null set where this is not true depends on  $(T, u)$ . Define the  $\mathcal{A} \otimes \mathcal{B}((0, \infty) \times \mathcal{U})$ -measurable set

$$\Gamma = \{(\omega, T, u) \in \Omega \times (0, \infty) \times \mathcal{U} : (1.19) \text{ is not a càdlàg function}\}. \quad (1.20)$$

By the considerations above together with Fubini's theorem we obtain

$$\int_{\Omega} \int_{(0, \infty) \times \mathcal{U}} \mathbb{I}_{\Gamma}(\omega, T, u) \, d\lambda \, d\mathbb{P} = \int_{(0, \infty) \times \mathcal{U}} \int_{\Omega} \mathbb{I}_{\Gamma}(\omega, T, u) \, d\mathbb{P} \, d\lambda = 0,$$

where  $\lambda$  denotes the Lebesgue measure on  $(0, \infty) \times \mathcal{U}$ . This proves the result.  $\square$

To infer the càdlàg property of  $\tilde{X}$  from the càdlàg property of the martingales  $\tilde{M}$  we use the purely analytical Lemma 1.26. To formulate this lemma we have to introduce some notation. Let  $\Re$  and  $\Im$  be the real and imaginary part of a complex number (this is then interpreted componentwise on  $V + \mathbf{i}V$ ). Denote by  $\Pi$  the projection on  $\text{span}(\Re \mathcal{U}) \subset V$  and by  $\Pi^{\perp} = id_V - \Pi$  the projection to the orthogonal complement  $\text{span}(\Re \mathcal{U})^{\perp} \subset V$ . Also denote by  $\Pi$  and  $\Pi^{\perp}$  the extension of these projections to  $V + \mathbf{i}V$ , e.g.  $\Pi(u) = \Pi(\Re u) + \mathbf{i}\Pi(\Im u)$ . Fix  $m \in \mathbb{N}$  and  $r > 0$  and define the set  $K = \{u \in \mathcal{U}^m : \|u\| \leq r\}$ . Let  $p$  be the dimension of  $\text{span}(\Re \mathcal{U})$ . Then there exist linearly independent  $(u_1, \dots, u_p) \in K \cap \Re \mathcal{U}$  and linearly independent  $(u_{p+1}, \dots, u_d) \in \Pi^{\perp} K$ . Fix  $s > 0$ .  $\Phi$  is continuous on the compact set  $\{(t, T, u) : 0 \leq o(T, u) \leq t \leq T \leq s + 1, u \in K\}$ . Hence it is uniformly continuous. By Lemma 1.21 iii)  $\Phi_{s,s}(u) = 1$ , so for each  $\tilde{c} > 0$  there is  $\delta > 0$  such that  $|\Phi_{t,T}(u)| \geq |\Phi_{s,s}(u)| - \tilde{c} = 1 - \tilde{c}$  for all  $0 \leq s - \delta \leq t, s \leq T \leq s + \delta$ . Choosing e.g.  $\tilde{c} = \frac{1}{2}$  this gives a compact set on which also  $\psi$  is continuous and hence uniformly continuous and bounded. Hence there is  $\epsilon > 0, \eta > 0$  such that for every  $t \in \mathcal{I}_s = (s, s + \epsilon)$  and  $q$  in  $(s - \epsilon, s + \epsilon) \cap [0, t)$

$$\Pi \psi_{q,t}(\bar{u}_1), \dots, \Pi \psi_{q,t}(\bar{u}_p) \quad \text{as well as} \quad \Pi^{\perp} \psi_{q,t}(\bar{u}_{p+1}), \dots, \Pi^{\perp} \psi_{q,t}(\bar{u}_d) \quad (1.21)$$

are linearly independent for all  $\|u_i - \bar{u}_i\| < \eta, 1 \leq i \leq d$  and

$$\inf_{u \in K} |\Phi_{q,t}(u)| > c \text{ and } \sup_{u \in K} \|\psi_{q,t}(u)\|^2 < C.$$

**Lemma 1.26:** *Let  $q_k \rightarrow s$  and  $x_{q_k}$  be a sequence with values in  $E_{\Delta}$ . Then*

<sup>1</sup>For  $t = T$  we set  $M_T^{T,u} := e^{\langle u, X_T \rangle}$ .

(i) If for Lebesgue-almost all  $(t, u) \in \mathcal{I}_s \times K$

$$\lim_{k \rightarrow \infty} N_{q_k}^{t,u} := \lim_{k \rightarrow \infty} e^{\langle \psi_{q_k, t}(t, u), x_{q_k} \rangle} \in \mathbb{C} \setminus \{0\},$$

then also  $\lim_{k \rightarrow \infty} x_{q_k}$  exists in  $E$ .

(ii) If there exists some  $(t, u) \in \mathcal{I}_s \times K$  such that

$$\lim_{k \rightarrow \infty} e^{\langle \psi_{q_k, t}(t, u), x_{q_k} \rangle} = 0,$$

then  $\lim_{k \rightarrow \infty} \|x_{q_k}\| = \infty$ , i.e.  $x_{q_k} \rightarrow \Delta$ . If  $(x_{q_k})$  is even an  $E$ -valued sequence, then  $\text{relint}(\mathfrak{RU}) \neq \emptyset$  and for all  $u \in \text{relint}(\mathfrak{RU})$

$$\lim_{k \rightarrow \infty} e^{\langle u, x_{q_k} \rangle} = 0.$$

We prove this lemma at the end of this section. Using this lemma we obtain the following result.

**Lemma 1.27:**  $\tilde{X}$  is a modification<sup>2</sup> of  $X$  on  $(\Omega, \mathcal{A}^{\mathbb{P}})$ , i.e.  $\mathbb{P}(\tilde{X}_t = X_t) = 1$  for all  $t \geq 0$ .

*Proof.* By Lemma 1.25 we can fix  $\Omega_0$  with  $\mathbb{P}(\Omega_0) = 1$  such that for each  $\omega \in \Omega_0$  (1.19) is a càdlàg function for almost all  $(T, u)$ . Fix such  $\omega \in \Omega_0$ . Let  $q_k \uparrow \uparrow s$  or  $q_k \downarrow \downarrow s$ . For all  $(t, u) \in \mathcal{I}_s \times K$  and  $k$  large it holds that  $\Phi_{q_k, t}(u) \neq 0$ . For almost all  $(t, u) \in \mathcal{I}_s \times K$   $\lim_{k \rightarrow \infty} M_{q_k}^{t, u}$  exists finitely valued and since  $\Phi_{q_k, T}(u) \neq 0$  also  $\lim_{k \rightarrow \infty} e^{\langle \psi_{q_k, t}(t, u), X_{q_k}(\omega) \rangle}$  exists finitely valued. By Lemma 1.26  $\lim_{k \rightarrow \infty} X_{q_k}(\omega) \in E_{\Delta}$ . This holds for all  $s \geq 0$  and hence for each  $\omega \in \Omega_0$  the left and right limits of  $X$  through  $\mathbb{Q}$  exist in  $E_{\Delta}$  and  $\tilde{X}_s(\omega) = \lim_{q_k \downarrow \downarrow s} X_{q_k}(\omega)$  defines a càdlàg function in  $E_{\Delta}$ . Then  $\Omega_0 \subset \tilde{\Omega}$  and  $\mathbb{P}(\tilde{\Omega}) = 1$  (for  $\mathbb{P}(\tilde{\Omega})$  to be defined we consider the  $\mathbb{P}$ -completion of  $\mathcal{A}$ ).

It remains to be proved that the  $E_{\Delta}$ -valued càdlàg process  $\tilde{X}$  is a modification of  $X$ . Since  $X$  is stochastically continuous (see the remark after Definition 1.9) and convergence in probability implies almost sure convergence along a subsequence we have for each  $t$  that

$$\lim_{q \downarrow \downarrow t} X_q \xrightarrow{p} X_t \Rightarrow \lim_{k \rightarrow \infty} X_{q_k}(\omega) = X_t(\omega) \text{ for all } \omega \in \bar{\Omega}, \text{ where } \mathbb{P}(\bar{\Omega}) = 1.$$

However, on  $\tilde{\Omega}$  we have that

$$\lim_{k \rightarrow \infty} X_{q_k}(\omega) = \lim_{q \downarrow \downarrow t} X_q(\omega) = \tilde{X}_t.$$

This yields  $\tilde{X}_t = X_t$  on  $\tilde{\Omega} \cap \bar{\Omega}$ . Since  $\mathbb{P}(\tilde{\Omega} \cap \bar{\Omega}) = 1$ , this proves the lemma.  $\square$

---

<sup>2</sup>Here  $X$  is already interpreted as a process on the extended probability space  $(\Omega, \mathcal{A}^{\mathbb{P}})$ .

*Remark:* Under a single measure  $\mathbb{P}$  we could also have defined  $\tilde{X} = \Delta$  on  $\Omega_0$  to get a modification. But  $\Omega_0$  depends on  $\mathbb{P}$  while  $\tilde{\Omega}$  does not. This is important when considering a Markov process  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$ , where there is a whole family of probability measures.

Next we show that  $\tilde{X}$  a.s. stays in  $\Delta$ , as soon as it is  $\Delta$  or approaches  $\Delta$  from the left. Note that this is true for  $X$  by definition, but not necessarily for  $\tilde{X}$ . Define

$$T_\Delta := \inf\{t > 0 : \tilde{X}_{t-} = \Delta \text{ or } \tilde{X}_t = \Delta\}. \quad (1.22)$$

**Lemma 1.28:**  $\tilde{X} = \Delta$  on  $[T_\Delta, \infty)$   $\mathbb{P}$ -a.s..

*Proof.* By the definition of  $\tilde{X}$ ,  $\tilde{X}_{s-} = \Delta$  implies  $\lim_{q \uparrow s} \|X_q\| = \infty$  and  $\tilde{X}_s = \Delta$  implies  $\lim_{q \downarrow s} \|X_q\| = \infty$ . In either case if there is a subsequence  $q_k$  with  $X_{q_k}(\omega) = \Delta$  (the subsequence can depend on  $\omega$ ), then by the definition of  $X$  and  $\Delta$ ,  $X = \Delta$  on  $(s, \infty)$  and then  $\tilde{X} = \Delta$  on  $[s, \infty)$ . So without loss of generality we can restrict to  $E$ -valued sequences.

Assume there is  $\omega \in \Omega_0$  and  $q \rightarrow s$  with  $\lim_{q \rightarrow s} \|X_q(\omega)\| = \infty$ , such that  $X_q(\omega) \in E$  for all  $q$  (if not, the lemma is already true). By Lemma 1.26 i) there is a subset  $(T, u)$  of  $\mathcal{I}_s \times K$  of positive Lebesgue-measure such that  $\lim_{q \rightarrow s} e^{\langle \psi_{q,T}(u), X_q(\omega) \rangle}$  is either zero or infinite or does not exist at all. Since  $\omega \in \Omega_0$  by Lemma 1.25 the limit exists finitely valued for almost all  $(T, u) \in \mathcal{I}_s \times K$  (remember that  $\Phi$  is continuous and  $\Phi \neq 0$  on this set). Hence there are  $T(\omega), u(\omega)$  such that

$$\lim_{q \rightarrow s} e^{\langle \psi_{t,T(\omega)}(u(\omega)), X_q(\omega) \rangle} = 0.$$

Lemma 1.26 ii) then yields that  $\text{relint}(\mathfrak{R}\mathcal{U}) \neq \emptyset$  and for all  $u \in \text{relint}(\mathfrak{R}\mathcal{U})$

$$\lim_{q \rightarrow s} e^{\langle u, X_q(\omega) \rangle} = 0. \quad (1.23)$$

Hence we can fix  $u \in \text{relint}(\mathfrak{R}\mathcal{U})$ . By Lemma 1.21  $\Phi_{t,T}(u)$  and  $\psi_{t,T}(u)$  are real-valued functions. For  $T > 0$  we can choose  $s < T$  such that  $\psi_{t,T}(u) \in \text{relint}(\mathfrak{R}\mathcal{U})$  and  $\Phi_{t,T}(u) > 0$  for  $s \leq t \leq T$ . Consider the set  $\Omega_{s,T} = \{\omega \in \Omega_0 : s < T_\Delta(\omega) < T\}$  and the nonnegative càdlàg martingale from equation (1.19) which on  $\Omega_0$  satisfies

$$\tilde{M}_t^{T,u}(\omega) = \Phi_{t,T}(u) e^{\langle \psi_{t,T}(u), \tilde{X}_t(\omega) \rangle}.$$

For  $\omega \in \Omega_{s,T}$  the considerations leading to (1.23) give

$$e^{\langle \psi_{T_\Delta(\omega),T}(u), \tilde{X}_{T_\Delta(\omega)}(\omega) \rangle} = 0 \quad \text{or} \quad e^{\langle \psi_{T_\Delta(\omega),T}(u), \tilde{X}_{T_\Delta(\omega)-}(\omega) \rangle} = 0.$$

By Theorem II.78.1 in Rogers and Williams [43]  $\tilde{M}_t^{T,u}(\omega) = 0$  for all  $T_\Delta(\omega) \leq t \leq T$  a.s. on  $\Omega_{s,T}$ . Since  $\Phi_{t,T}(u) > 0$ , then  $\tilde{X}_t(\omega) = \Delta$  for  $T_\Delta(\omega) \leq t \leq T$  a.s. on  $\Omega_{s,T}$ . Since  $T > T_\Delta(\omega)$ , by the definition of  $\tilde{X}$  also  $X_t(\omega) = \Delta$  for some  $t \in (T_\Delta(\omega), T]$ . By the definition of  $\Delta$  and  $X$  then  $X_t(\omega) = \Delta$  on  $[T, \infty)$  and hence also  $\tilde{X}_t(\omega)$ . Since  $T$  was arbitrary, the lemma follows.  $\square$

Define the process

$$\bar{X} := \tilde{X}\mathbb{I}_{[0, T_\Delta)} + \Delta\mathbb{I}_{[T_\Delta, \infty)}.$$

**Corollary 1.29:** *The process  $\bar{X}$  is càdlàg in  $E_\Delta$ , a modification of  $X$  and an affine process on the completed filtered probability space  $(\Omega, \mathcal{A}^\mathbb{P}, \mathcal{F}^\mathbb{P}, \mathbb{P})$ .*

*Proof.* By Lemma 1.27 and Lemma 1.28  $\bar{X}$  is a modification of  $X$ . By Lemma 1.11 it is an affine process with the same transition function. Note that  $\bar{X}$  is adapted to  $\mathcal{F}^\mathbb{P}$ , but not necessarily to  $\mathcal{F}$ .  $\square$

So far we have worked with a single probability measure  $\mathbb{P}$ . This demonstrates the existence of a càdlàg modification for affine Markov processes in the sense of Definition 1.3. Now let  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  be an affine process in the sense of Definition 1.6. For each  $(s, x)$  we can apply Corollary 1.29 to the affine process  $X$  on  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P}^{(s,x)})$  with continuously affine transition function  $(P_{s+t, s+T})_{0 \leq t \leq T}$  and immediately get that  $\bar{X}$  is a version on  $(\Omega, \mathcal{A}^{(s,x)}, \mathcal{F}^{(s,x)}, \mathbb{P}^{(s,x)})$ , where  $\mathcal{A}^{(s,x)}$  is the completed  $\sigma$ -algebra and  $\mathcal{F}_t^{(s,x)}$  the completed filtration with respect to the measure  $\mathbb{P}^{(s,x)}$ . The  $\sigma$ -algebras in this case still depend on the measure  $\mathbb{P}^{(s,x)}$ . To get a single filtration define

$$\bar{\mathcal{A}} := \bigcap_{(s,x)} \mathcal{A}^{(s,x)}, \quad \bar{\mathcal{F}}_t := \bigcap_{(s,x)} \mathcal{F}_t^{(s,x)}, t \in [0, \infty]. \quad (1.24)$$

We can now formulate the final theorem of this section.

**Theorem 1.30:** *Let  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  be a Markov process with continuously affine transition function. Then  $(\bar{X}, \bar{\mathcal{F}}, \{\mathbb{P}^{(s,x)}\})$  on  $(\Omega, \bar{\mathcal{A}})$  is a modification of the affine process  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  which has càdlàg paths.*

*Proof.* The set  $\{X_t \neq \bar{X}_t\}$  has  $\mathbb{P}^{(r,x)}$ -measure zero for all  $(r, x)$ . Hence it is in  $\bar{\mathcal{F}}_t$ .  $X_t$  is  $\bar{\mathcal{F}}_t$ -measurable, hence  $\bar{X}_t$  is  $\bar{\mathcal{F}}_t$ -measurable and therefore adapted. By corollary 1.29  $\bar{X}$  is a Markov process on  $(\Omega, \mathcal{A}^{(s,x)}, \mathcal{F}^{(s,x)}, \mathbb{P}^{(s,x)})$  and since  $\bar{\mathcal{F}}_t \subset \mathcal{F}_t^{(s,x)}$ , it is also a Markov process on the smaller filtration. As this holds for all  $\mathbb{P}^{(s,x)}$ , this proves the claim.  $\square$

*Remark:* For a conservative Markov process with a continuously affine transition function (see Lemma 1.24) one can construct an  $E$ -valued càdlàg process by setting  $\hat{X} = \bar{X}\mathbb{I}\{T_\Delta = \infty\} + X_0\mathbb{I}\{T_\Delta < \infty\}$ . Then  $\{\hat{X}_t \neq X_t\} \subset \{T_\Delta < \infty\}$ . For a conservative process  $\mathbb{P}^{(s,x)}(T_\Delta < \infty) = 0$  for all  $(s, x)$  and hence also  $\hat{X}$  is a modification of  $X$ .

The càdlàg version of the affine process has the strong Markov property. Set  $\bar{X}_\infty := \Delta$  and remember  $f(\Delta) = 0$ .

**Theorem 1.31:** *The càdlàg process  $(\bar{X}, \bar{\mathcal{F}}, \{\mathbb{P}^{(s,x)}\})$  is a strong Markov process, i.e. for each stopping time  $\tau$ , bounded measurable  $f$  and  $s \geq 0$*

$$\mathbb{E}^{(r,x)} [f(\bar{X}_{\tau+s}) | \bar{\mathcal{F}}_\tau] = \mathbb{E}^{(r+\tau, \bar{X}_\tau)} [f(\bar{X}_s)] = P_{r+\tau, r+\tau+s} f(\bar{X}_\tau).$$

*Proof.* By Lemma 1.22 the function  $(s, x) \mapsto P_{t,t+s}f(x)$  is  $\mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E}$ -measurable. Additionally  $(\tau, \bar{X}_\tau)$  is  $\bar{\mathcal{F}}_\tau/(\mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E})$ -measurable. So  $\mathbb{E}^{(r+\tau, \bar{X}_\tau)}[f_u(X_s)]$  is  $\bar{\mathcal{F}}_\tau$ -measurable.

For a stopping time  $\tau$  define an approximating sequence of stopping times

$$\tau_n(\omega) := \begin{cases} k2^{-n} & \text{if } (k-1)2^{-n} \leq \tau(\omega) < k2^{-n}, k \in \mathbb{N}, \\ \infty & \text{if } \tau(\omega) = \infty. \end{cases}$$

Let  $\Lambda \in \bar{\mathcal{F}}_\tau$ . Then  $\Lambda_{n,k} := \{\omega : \tau_n(\omega) = k2^{-n}\} \cap \Lambda \in \bar{\mathcal{F}}_{k2^{-n}}$ . The simple Markov property yields

$$\begin{aligned} \mathbb{E}^{(r,x)} [f_u(\bar{X}_{\tau_n+s})\mathbb{I}_\Lambda] &= \sum_{k \in \mathbb{N}} \mathbb{E}^{(r,x)} [\mathbb{I}_{\Lambda_{n,k}} \mathbb{E}^{(r,x)} [f_u(\bar{X}_{k2^{-n}+s}) | \bar{\mathcal{F}}_{k2^{-n}}]] \\ &= \sum_{k \in \mathbb{N}} \mathbb{E}^{(r,x)} \left[ \mathbb{I}_{\Lambda_{n,k}} \Phi_{r+k2^{-n}, r+k2^{-n}+s}(u) e^{\langle \psi_{r+k2^{-n}, r+k2^{-n}+s}(u), \bar{X}_{k2^{-n}} \rangle} \right] \\ &= \mathbb{E}^{(r,x)} \left[ \mathbb{I}_\Lambda \Phi_{r+\tau_n, r+\tau_n+s}(u) e^{\langle \psi_{r+\tau_n, r+\tau_n+s}(u), \bar{X}_{\tau_n} \rangle} \right]. \end{aligned}$$

For  $u \in \mathfrak{iV}$  by dominated convergence, continuity of  $\Phi$  and  $\psi$  and right-continuity of  $\bar{X}$

$$\begin{aligned} \mathbb{E}^{(r,x)} [f_u(\bar{X}_{\tau+s})\mathbb{I}_\Lambda] &= \mathbb{E}^{(r,x)} \left[ \mathbb{I}_\Lambda \Phi_{r+\tau, r+\tau+s}(u) e^{\langle \psi_{r+\tau, r+\tau+s}(u), \bar{X}_\tau \rangle} \right] \\ &= \mathbb{E}^{(r,x)} \left[ \mathbb{I}_\Lambda P_{r+\tau, r+\tau+s} f_u(\bar{X}_\tau) \right] \\ &= \mathbb{E}^{(r,x)} \left[ \mathbb{I}_\Lambda \mathbb{E}^{(r+\tau, \bar{X}_\tau)} [f_u(\bar{X}_s)] \right]. \end{aligned}$$

An application of the functional monotone class theorem as in Lemma 1.22 then yields the result for bounded and measurable functions  $f$ .  $\square$

If we assume that the filtration  $\mathcal{F}$  is the natural filtration of the process  $X$ , we can show that the completed filtrations  $\mathcal{F}^{(s,x)}$  are already right-continuous. By the definition of  $\bar{\mathcal{F}}$  right-continuity transfers to  $\bar{\mathcal{F}}$ .

**Theorem 1.32:** *Let  $\mathcal{F} = \mathcal{F}^X$  be the natural filtration of the affine process  $X$  and  $\mathcal{A} = \mathcal{F}_\infty^X$ . Then  $\mathcal{F}^\mathbb{P}$  is right-continuous.*

*Proof.* To show this we use the continuity of  $\Phi$  and  $\psi$  together with the fact that  $\bar{X}$  is càdlàg. Repeating the arguments in the proof of Theorem 4 in Cuchiero and Teichmann [10] with the inhomogeneous versions of  $\Phi$  and  $\psi$  we get that for  $k \in \mathbb{N}$ ,  $u_1, \dots, u_k \in \mathfrak{iV}$ ,  $t_1, \dots, t_k \in \mathbb{R}_{\geq 0}$

$$\mathbb{E} \left[ e^{\langle u_1, \bar{X}_{t_1} \rangle + \dots + \langle u_k, \bar{X}_{t_k} \rangle} | \mathcal{F}_t^\mathbb{P} \right] = \mathbb{E} \left[ e^{\langle u_1, \bar{X}_{t_1} \rangle + \dots + \langle u_k, \bar{X}_{t_k} \rangle} | \mathcal{F}_{t+}^\mathbb{P} \right]. \quad (1.25)$$

Consider the vector space of functions

$$\mathcal{H} := \{ \text{bounded } \mathcal{A}^\mathbb{P}\text{-measurable } Z : \mathbb{E} [Z | \mathcal{F}_t^\mathbb{P}] = \mathbb{E} [Z | \mathcal{F}_{t+}^\mathbb{P}] \}.$$

This space contains the constant functions and is a monotone class. The set  $\mathcal{C} := \{\prod_{l=1}^n g_l(X_{t_l}), g_l \in C_{exp}, t_l \geq 0, n \in \mathbb{N}\}$  of functions from  $\Omega$  to  $\mathbb{R}$  is closed under multiplication, generates  $\mathcal{F}_\infty^X$  (see Lemma 1.17) and by (1.25) is a subset of  $\mathcal{H}$ . Hence we can apply the functional monotone class theorem (see Theorem 0.2.2 in Revuz and Yor [42]). This yields  $\mathbb{E}[Z | \mathcal{F}_t^{\mathbb{P}}] = \mathbb{E}[Z | \mathcal{F}_{t+}^{\mathbb{P}}]$  for bounded  $\mathcal{F}_\infty^X$ -measurable functions  $Z$ .

Now consider  $\tilde{A} \in \mathcal{F}_{t+}^{\mathbb{P}}$ . Then there is a set  $A \in \mathcal{F}_{t+}$  and  $B_1, B_2 \in \mathcal{F}_\infty^{\mathbb{P}}$ , such that  $\mathbb{P}(B_1) = \mathbb{P}(B_2) = 0$  and  $A \setminus B_1 \subset \tilde{A} \subset A \cup B_2$ . For  $Z = \mathbb{I}_A$  there is a  $\mathcal{F}_t^{\mathbb{P}}$ -measurable  $f$ , such that  $\mathbb{I}_A = f\mathbb{I}_{N^c} + \mathbb{I}_A\mathbb{I}_N$  with  $\mathbb{P}(N) = 0$ . Since  $\mathcal{F}_t^{\mathbb{P}}$  contains all null sets, it follows that  $f\mathbb{I}_{N^c}$  is  $\mathcal{F}_t^{\mathbb{P}}$ -measurable. Furthermore  $\mathbb{P}(A \cap N) = 0$ , so also  $\mathbb{I}_A\mathbb{I}_N$  is  $\mathcal{F}_t^{\mathbb{P}}$ -measurable. Then  $\mathbb{I}_A$  is  $\mathcal{F}_t^{\mathbb{P}}$ -measurable and  $A \in \mathcal{F}_t^{\mathbb{P}}$ . Hence also  $\tilde{A} \in \mathcal{F}_t^{\mathbb{P}}$  and this shows  $\mathcal{F}_t^{\mathbb{P}} = \mathcal{F}_{t+}^{\mathbb{P}}$ .  $\square$

If  $\mathcal{F} = \mathcal{F}^X$  and there exist shift operators  $\theta_t$  (e.g. for the canonical realization), the strong Markov property extends to  $\bar{\mathcal{F}}_\infty$ -measurable  $Z$ , i.e. for a stopping time  $\tau$ , for all  $s, t \geq 0$ ,  $x \in E$  and bounded  $\bar{\mathcal{F}}_\infty$ -measurable  $Z$

$$\mathbb{E}^{(s,x)} [Z \circ \theta_\tau | \bar{\mathcal{F}}_\tau] = \mathbb{E}^{(s+\tau, \bar{X}_\tau)} [Z] \quad \text{a.s. on the set } \{\bar{X}_\tau \neq \Delta\}.$$

This follows by the arguments presented in Theorem III.8 and Theorem III.9 of Rogers and Williams [43].

Furthermore the càdlàg property and possibly right continuity of the filtration carry over to the space-time process  $\tilde{X} = (\Theta, \bar{X})$  introduced in Corollary 1.23. Additionally we still have continuity of  $(s, x) \mapsto \tilde{P}_r f(u^0, u)(s, x)$  and the proof of the strong Markov property also transfers. Then the extended filtration  $\tilde{\mathcal{F}}$  is a right-continuous strong Markov filtration in the sense of Çınlar et al. [28] and affine processes can be embedded in the Markov semimartingale setting of Çınlar et al. [28] (see also section 1.4).

I.e. if  $(X, \mathcal{F}^X, \{\mathbb{P}^{(s,x)}\})$  is the coordinate process of the transition function we call the càdlàg time-homogeneous strong Markov process  $(\tilde{X}, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(s,x)}\})$  the canonical space-time realization of the continuously affine transition function  $\{P_{s,t}\}$ . Here the filtration  $\tilde{\mathcal{F}}_t = \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \bar{\mathcal{F}}_t$  is a right-continuous strong Markov filtration.

## Proof of Lemma 1.26

The proofs in this section follow closely the proofs in Cuchiero and Teichmann [10].

**Lemma 1.33:** *Let  $(x_k)$  be a  $E$ -valued sequence such that  $\lim_{k \rightarrow \infty} \Pi x_k$  exists finitely valued and  $\limsup_{k \rightarrow \infty} \|\Pi^\perp x_k\| = \infty$ . There exists a subsequence  $(x_k)$  denoted again by  $(x_k)$  and a finite number of orthonormal directions  $g_i \in \text{span}(\mathfrak{R}\mathcal{U})^\perp$  such that*

$$\lim_{k \rightarrow \infty} x_k - \sum_i \langle x_k, g_i \rangle g_i$$

*exists while  $\lim_{k \rightarrow \infty} \langle x_k, g_i \rangle = \infty$ , where the rates of divergence are non-increasing in  $i$  in the sense that  $\limsup_{k \rightarrow \infty} \frac{\langle x_k, g_{i+1} \rangle}{\langle x_k, g_i \rangle} < \infty$ .*

Fix  $v < T < \sigma(v, 0)$  and consider the set  $\mathcal{T} := \{(s, t) : v \leq s \leq t \leq T\}$ . Then there exist continuous functions  $R : \mathcal{T} \rightarrow \mathbb{R}_{>0}$  and  $\lambda_i : \mathcal{T} \rightarrow V$ , such that for all  $u \in B_{R(s,t)}^{\mathbf{i}V} := \{u \in \mathbf{i}V : \|u\| < R(s, t)\}$

$$\langle \psi_{s,t}(u), g_i \rangle = \langle \lambda_i(s, t), u \rangle. \quad (1.26)$$

*Remark:* The above lemma states that if  $x_k$  is only diverging in  $\text{span}(\mathfrak{R}\mathcal{U})^\perp$ , we can find directions of divergence  $g_i \in \text{span}(\mathfrak{R}\mathcal{U})^\perp$  and open sets in  $\mathbf{i}V$ , so that  $\langle \psi_{s,t}(u), g_i \rangle$  is linear in  $u$  for all  $i$ . Note that with  $g_i \in \langle \mathfrak{R}\mathcal{U} \rangle^\perp$  and  $\Re\psi_{s,t}(u) \in \mathcal{U}$

$$\langle \psi_{s,t}(u), g_i \rangle = \langle \Pi^\perp \Re\psi_{s,t}(u) + \mathbf{i}\Pi^\perp \Im\psi_{s,t}(u), g_i \rangle = \mathbf{i} \langle \Im\Pi^\perp \psi_{s,t}(u), g_i \rangle, \quad (1.27)$$

and it suffices to show linearity of  $\langle \Im\Pi^\perp \psi_{s,t}(u), g_i \rangle$ .

*Proof of 1.33.* For the first part we can choose directions of convergence by inductively choosing subsequences of  $(x_k)$  for which

$$g_l = \lim_{k \rightarrow \infty} \frac{x_k - \sum_{j=1}^{l-1} \langle x_k, g_j \rangle g_j}{\|x_k - \sum_{j=1}^{l-1} \langle x_k, g_j \rangle g_j\|}$$

converges. For instance  $g_1 = \lim_{k \rightarrow \infty} \frac{x_k}{\|x_k\|}$ . By construction the  $g_l$  are orthogonal. Choosing further subsequences of  $(x_k)$  one can arrange that the rates of convergence are non-increasing in  $l$ .

Next note that  $\Phi_{s,t}(0) \neq 0$  for  $(s, t) \in \mathcal{T}$ . By continuity of  $\Phi$  there is  $r > 0$ , such that  $\Phi_{s,t}(u) \neq 0$  for  $u \in B_r^{\mathbf{i}V}$  and  $(s, t) \in \mathcal{T}$ . For  $u \in B_r^{\mathbf{i}V}$  define

$$\Theta(u, x) := \frac{P_{s,t} f_u(x)}{\Phi_{s,t}(0) e^{\langle \psi_{s,t}(0), \Pi x \rangle}} = \frac{\Phi_{s,t}(u) e^{\langle \psi_{s,t}(u), x \rangle}}{\Phi_{s,t}(0) e^{\langle \Pi \psi_{s,t}(0), \Pi x \rangle}}.$$

By Lemma 1.21 ii)  $\Im\Phi_{s,t}(0) = \Im\psi_{s,t}(0) = 0$  and  $\Phi_{s,t}(0) e^{\langle \Pi \psi_{s,t}(0), \Pi x \rangle} > 0$ . Since a characteristic function is always positive definite we conclude that  $\Theta(\cdot, x)$  is a positive definite function in  $B_r^{\mathbf{i}V}$  for all  $x \in E$ . Since  $\psi_{s,t}(0) \in \mathfrak{R}\mathcal{U}$ , it follows that  $\Pi\psi_{s,t}(0) = \psi_{s,t}(0)$ . Hence  $\Theta(0, x) = 1$  and by Lemma 3.2. in Keller-Ressel et al. [35]

$$|\Theta(u + v, x) - \Theta(u, x)\Theta(v, x)|^2 \leq 1, \quad u, v \in B_{\frac{r}{2}}^{\mathbf{i}V}. \quad (1.28)$$

Using  $\Re\psi_{s,t}(u) \in \mathfrak{R}\mathcal{U} \subset \text{span}(\mathfrak{R}\mathcal{U})$  and  $\Pi\Im = \Im\Pi$  one obtains

$$P_{s,t} f_u(x) = |\Phi_{s,t}(u)| e^{\langle \Pi \Re\psi_{s,t}(u), \Pi x \rangle} e^{\mathbf{i}(\arg(\Phi_{s,t}(u)) + \langle \Im\Pi\psi_{s,t}(u), \Pi x \rangle + \langle \Im\Pi^\perp \psi_{s,t}(u), \Pi^\perp x \rangle)}.$$

Define the following continuous functions

$$\begin{aligned}
\beta_1(u, v, s, t) &:= \Im \Pi^\perp \psi_{s,t}(u+v) \\
\beta_2(u, v, s, t) &:= \Im \Pi^\perp \psi_{s,t}(u) + \Im \Pi^\perp \psi_{s,t}(v) \\
r_1(u, v, x, s, t) &:= \left| \frac{\Phi_{s,t}(u+v)}{\Phi_{s,t}(0)} \right| \exp\left(\langle \Re \Pi(\psi_{s,t}(u+v) - \psi_{s,t}(0)), \Pi x \rangle\right) \\
r_2(u, v, x, s, t) &:= \left| \frac{\Phi_{s,t}(u)\Phi_{s,t}(v)}{\Phi_{s,t}(0)^2} \right| \exp\left(\langle \Re \Pi(\psi_{s,t}(u) + \psi_{s,t}(v) - 2\psi_{s,t}(0)), \Pi x \rangle\right) \\
\alpha_1(u, v, x, s, t) &:= \arg\left(\frac{\Phi_{s,t}(u+v)}{\Phi_{s,t}(0)}\right) + \langle \Im \Pi \psi_{s,t}(u+v), \Pi x \rangle \\
\alpha_2(u, v, x, s, t) &:= \arg\left(\frac{\Phi_{s,t}(u)\Phi_{s,t}(v)}{\Phi_{s,t}(0)^2}\right) + \langle \Im \Pi(\psi_{s,t}(u) + \psi_{s,t}(v)), \Pi x \rangle.
\end{aligned}$$

Inserting in equation (1.28) (and suppressing the arguments of the functions)

$$\begin{aligned}
1 &\geq \left| r_1 e^{i(\alpha_1 + \langle \beta_1, \Pi^\perp x \rangle)} - r_2 e^{i(\alpha_2 + \langle \beta_2, \Pi^\perp x \rangle)} \right|^2 \\
&= r_1^2 + r_2^2 - 2r_1 r_2 \cos(\alpha_1 - \alpha_2 + \langle \beta_1 - \beta_2, \Pi^\perp x \rangle).
\end{aligned}$$

Since  $r_1^2 + r_2^2 \geq 2r_1 r_2$  this yields for  $u, v \in B_r^{\text{iV}}$  and  $x \in E$

$$r_1 r_2 \left(1 - \cos(\alpha_1 - \alpha_2 + \langle \beta_1 - \beta_2, \Pi^\perp x \rangle)\right) \leq \frac{1}{2}. \quad (1.29)$$

Furthermore, define

$$R(x, s, t) := \sup \left\{ \rho \leq \frac{r}{2} : r_1(u, v, x, s, t) r_2(u, v, x, s, t) \geq \frac{4}{5} \text{ for } u, v \in B_\rho^{\text{iV}} \right\}.$$

$R(x, s, t) > 0$  for  $x \in E$ , since  $r_1(0, 0, x, s, t) = r_2(0, 0, x, s, t) = 1$  and  $r_1$  and  $r_2$  are continuous. This also gives continuity of  $R(x, s, t)$ . Since  $\lim_k \Pi x_k$  exists finitely valued, also

$$R(s, t) := \frac{1}{2} \inf_k R(x_k, s, t) > 0.$$

Now suppose that  $\langle \Pi^\perp \Im \psi_{s,t}(u), g_1 \rangle$  is not linear for  $u \in B_{R(s,t)}^{\text{iV}}$ . Then by the continuity of  $\beta_1$  and  $\beta_2$  there exist  $u^*, v^* \in B_{R(s,t)}^{\text{iV}}$  and an open neighborhood  $O \subset B_{R(s,t)}^{\text{iV}}$ , such that

$$\langle \beta_1(u, v, s, t) - \beta_2(u, v, s, t), g_1 \rangle \neq 0$$

and the left-hand side is not constant on  $O$ .  $\alpha_1, \alpha_2$  are continuous in  $x$  and only depend on  $x$  via  $\Pi x$ . Since  $\Pi x_k$  is converging, while  $\Pi^\perp x_k$  is diverging with highest divergence rate in direction  $g_1$  there exist some  $(u, v) \in O$  and some  $k \in \mathbb{N}$ , such that

$$\cos(\alpha_1(u, v, x_k, s, t) - \alpha_2(u, v, x_k, s, t) + \langle \beta_1(u, v, s, t) - \beta_2(u, v, s, t), \Pi^\perp x_k \rangle) \leq \frac{1}{3}.$$

Together with  $r_1(u, v, x_k, s, t)r_2(u, v, x_k, s, t) \geq \frac{4}{5}$  this yields

$$r_1 r_2 (1 - \cos(\alpha_1 - \alpha_2 + \langle \beta_1 - \beta_2, \Pi^\perp x_k \rangle)) \geq \frac{8}{15} > \frac{1}{2},$$

which contradicts equation (1.29). Hence

$$\langle \beta_1(u, v, s, t) - \beta_2(u, v, s, t), g_1 \rangle = 0$$

for all  $(u, v) \in B_{R(s,t)}^{\text{iV}}$ . This gives linearity of  $\langle \Pi^\perp \mathfrak{S}\psi_{s,t}(u), g_1 \rangle$  for  $u \in B_{R(s,t)}^{\text{iV}}$ . By equation (1.27) this is equivalent to linearity of  $\langle \psi_{s,t}(u), g_1 \rangle$ . Since  $\psi_{s,t}(u)$  is continuous there is a continuous function  $\lambda_1(s, t)$  such that

$$\langle \psi_{s,t}(u), g_1 \rangle = \langle \lambda_1(s, t), u \rangle \quad (1.30)$$

for all  $u \in B_{R(s,t)}^{\text{iV}}$ . Proceeding inductively gives the assertion.  $\square$

*Proof of 1.26 i).* Choose  $(t, u) \in \mathcal{I}_s \times K$ , such that  $\lim_{k \rightarrow \infty} N_{q_k}^{t,u} \in \mathbb{C} \setminus \{0\}$ . This implies that there is  $N$  such that  $x_{q_k} \neq \Delta$  for all  $k \geq N$ . We pass to a subsequence of  $q_k$ , such that  $q_k \leq t$  and  $x_{q_k} \in E$  (this excludes only finitely many members of the sequence). Define

$$A := \limsup_{k \rightarrow \infty} \langle \Re \psi_{q_k, t}(u), x_{q_k} \rangle, \quad a := \liminf_{k \rightarrow \infty} \langle \Re \psi_{q_k, t}(u), x_{q_k} \rangle.$$

Since  $\lim_{k \rightarrow \infty} N_{q_k}^{t,u}$  exists finitely valued,  $A$  is finite and since the limit does not vanish,  $a$  is finite. Hence there exist subsequences, such that  $A = \lim_{m \rightarrow \infty} \Re A_m$  and  $a = \lim_{l \rightarrow \infty} \Re a_l$ , where  $A_m := \langle \psi_{q_{k_m}, t}(u), x_{q_{k_m}} \rangle$  and  $a_l := \langle \psi_{q_{k_l}, t}(u), x_{q_{k_l}} \rangle$ . Since  $\lim_{k \rightarrow \infty} N_{q_k}^{t,u}$  exists finitely valued, we get that

$$\begin{aligned} \lim_{m \rightarrow \infty} e^{\Re A_m} \cos(\Im A_m) &= \lim_{l \rightarrow \infty} e^{\Re a_l} \cos(\Im a_l), \\ \lim_{m \rightarrow \infty} e^{\Re A_m} \sin(\Im A_m) &= \lim_{l \rightarrow \infty} e^{\Re a_l} \sin(\Im a_l). \end{aligned}$$

Taking square and using  $\cos(x)^2 + \sin(x)^2 = 1$  gives  $A = a$  and the existence of

$$\lim_{k \rightarrow \infty} \langle \Re \psi_{q_k, t}(u), x_{q_k} \rangle = \lim_{k \rightarrow \infty} \langle \Re \psi_{q_k, t}(u), \Pi x_{q_k} \rangle. \quad (1.31)$$

Properly choosing  $\bar{u}_1, \dots, \bar{u}_p$  from (1.21) gives the existence of

$$\lim_{k \rightarrow \infty} \Pi x_{q_k}. \quad (1.32)$$

We now concentrate on  $\Pi^\perp x_{q_k}$ . Note that  $\Pi^\perp \psi_{s,t}(u) = \text{i}\Pi^\perp \mathfrak{S}\psi_{s,t}(u)$ . By (1.32) and (1.31) also

$$\lim_{k \rightarrow \infty} e^{\langle \Pi^\perp \psi_{q_k, t}(u), \Pi^\perp x_{q_k} \rangle} \in \mathbb{C} \setminus \{0\} \quad (1.33)$$

for almost all  $(t, u) \in \mathcal{I}_s \times K$ . Assume that

$$\limsup_{k \rightarrow \infty} \|\Pi^\perp x_{q_k}\| = \infty.$$

Then we are in the situation of the Lemma 1.33. There exist directions  $g_i$  and a subsequence of  $q_k$ , again denoted by  $q_k$ , such that  $x_{q_k} - \sum_i \langle g_i, x_{q_k} \rangle g_i$  converges. For each  $t \in \mathcal{I}_s$  after passing to a subsequence such that  $q_k < t$  set  $R(t) = \inf_k R(q_k, t)$ , which by the continuity and strict positivity of  $R$  is strictly greater than 0. Thus for  $u \in B_{R(t)}^{\mathbb{V}} \cap K$  and  $k$  large

$$\langle \Pi^\perp \psi_{q_k, t}(u), g_i \rangle = \langle \lambda_i(q_k, t), u \rangle.$$

Using this we can write the exponent in equation (1.33) (along a subsequence) as

$$\langle \Pi^\perp \psi_{q_k, t}(u), x_{q_k} - \sum_i \langle g_i, x_{q_k} \rangle g_i \rangle + \sum_i \langle g_i, x_{q_k} \rangle \langle \lambda_i(q_k, t), u \rangle.$$

By Lemma 1.21 iii),  $\psi_{s, s}(u) = u$ . Therefore we can find  $u$  and  $g_i$ , such that  $\langle \Pi^\perp \psi_{s, s}(u), g_i \rangle \neq 0$ . By the continuity of  $\psi$  we can find  $t \in \mathcal{I}_s$ ,  $N \in \mathbb{N}$  and an (open) set of positive measure  $O \subset B_{R(t)}^{\mathbb{V}} \cap K$  such that for all  $k \geq N$ ,  $u \in O$

$$\langle \lambda_i(q_k, t), u \rangle = \langle \Pi^\perp \psi_{q_k, t}(u), g_i \rangle \neq 0. \quad (1.34)$$

Furthermore  $O$  can be chosen in a way such that

$$\lim_{k \rightarrow \infty} \int_{i\mathbb{V}} \mathbb{I}_O(u) e^{\langle \Pi^\perp \psi_{q_k, t}(u), x_{q_k} - \sum_i \langle g_i, x_{q_k} \rangle g_i \rangle} e^{\langle \sum_i \langle g_i, x_{q_k} \rangle \lambda_i(q_k, t), u \rangle} du \neq 0. \quad (1.35)$$

$f_k(u) := \mathbb{I}_O(u) e^{\langle \Pi^\perp \psi_{q_k, t}(u), x_{q_k} - \sum_i \langle g_i, x_{q_k} \rangle g_i \rangle}$  converges uniformly to a bounded function  $f(u)$ , while  $\sum_i \langle g_i, x_{q_k} \rangle \lambda_i(q_k, t)$  diverges by (1.34) and Lemma 1.33. By the Riemann-Lebesgue Lemma the limit in (1.35) is zero, which gives a contradiction. We conclude that

$$\limsup_{k \rightarrow \infty} \|\Pi^\perp x_{q_k}\| < \infty.$$

An application of Cauchy-Schwarz gives

$$|\langle \Pi^\perp \psi_{q_k, t}(u), \Pi^\perp x_{q_k} \rangle| = |\mathfrak{I} \langle \mathfrak{S} \Pi^\perp \psi_{q_k, t}(u), \Pi^\perp x_{q_k} \rangle| \leq \|\mathfrak{S} \Pi^\perp \psi_{q_k, t}(u)\| \|\Pi^\perp x_{q_k}\|.$$

For large enough  $k$   $\|\Pi^\perp x_{q_k}\|$  is bounded by  $\limsup_{k \rightarrow \infty} \|\Pi^\perp x_{q_k}\| + 1$ . Furthermore,

$$\|\mathfrak{S} \Pi^\perp \psi_{q_k, t}(u)\| \leq \|\mathfrak{S} \Pi^\perp (\psi_{q_k, t}(u) - u)\| + \|\mathfrak{S} \Pi^\perp (u)\|,$$

where using the continuity of  $\psi$  the two terms can be made arbitrarily small by choosing  $u$  small,  $k$  large and  $t$  close to  $s$ . I.e. there is  $k, t, u$  so that  $-\pi < \mathfrak{S} \langle \Pi^\perp \psi_{q_k, t}(u), \Pi^\perp x_{q_k} \rangle < \pi$ . Since (1.33) converges finitely and non-zero valued,

$$\lim_{k \rightarrow \infty} \langle \Pi^\perp \psi_{q_k, t}(u), \Pi^\perp x_{q_k} \rangle \in i\mathbb{R}.$$

Using this for  $d - p$  properly chosen linearly independent vectors  $u_{p+1}, \dots, u_d$  from (1.21) we conclude

$$\lim_{k \rightarrow \infty} \Pi^\perp x_{q_k}$$

exists finitely valued. Together with (1.32) and since  $E$  is closed this proves i).

To prove ii) note that  $\psi_{q_k, t}(u)$  is bounded on  $\mathcal{I}_s \times K$  for  $k$  large, hence the assumption of ii) can only hold if  $\lim_{k \rightarrow \infty} \|x_{q_k}\| = \infty$ . If  $x_{q_k}$  is  $E$ -valued, furthermore

$$\lim_{k \rightarrow \infty} \langle \Re \psi_{q_k, t}(u), x_{q_k} \rangle = -\infty \quad (1.36)$$

and hence  $\Re \psi_{q_k, t}(u) \neq 0$  for a subsequence  $q_k$ . Since  $\Phi_{q_k, t}(u) > 0$  for  $k$  large, we have  $u \in \mathcal{Q}$  and  $\psi_{q_k, t}(u) \in \mathcal{U}$ . Hence  $\Re \mathcal{U} \neq \emptyset$  and since  $\Re \mathcal{U} \subset \text{relin}(\Re \mathcal{U})$ ,  $\text{relin}(\Re \mathcal{U})$  cannot be empty.

By (1.36) it follows that  $\lim_{k \rightarrow \infty} \|\Pi x_{q_k}\| = \infty$ . There is a subsequence and some direction  $g \in \text{span}(\Re \mathcal{U})$  such  $\lim_{k \rightarrow \infty} \langle g, x_{q_k} \rangle = \infty$ . For  $u \in \text{relin}(\Re \mathcal{U})$  there is  $\epsilon > 0$  such that  $u + \epsilon g \in \text{relin}(\Re \mathcal{U})$  which, by the definition of  $\mathcal{U}$ , implies  $\sup_{x \in E} \langle u + \epsilon g, x \rangle < \infty$ . Hence  $\lim_{k \rightarrow \infty} \langle u, x_{q_k} \rangle = -\infty$ , which proves the remaining parts of ii).  $\square$

## 1.4 Regularity and semimartingale characteristics

Remember that the state space  $E$  is a subset of the real vector space  $V$ . A  $E$ -valued<sup>3</sup> càdlàg process  $X$  on a filtered probability space  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$  is a semimartingale if it can be decomposed as a  $(\mathcal{F}, \mathbb{P})$ -local martingale  $M$  and a finite variation process  $A$ , i.e.  $X = X_0 + M + A$ . For an introduction to semimartingales (and semimartingale characteristics) we refer to Jacod and Shiryaev [29].

Let  $X$  be a semimartingale. The jump measure of  $X$  is defined as

$$\mu^X(dt, dx; \omega) = \sum_s \mathbb{I} \{ \Delta X_s(\omega) \neq 0 \} \delta_{(s, \Delta X_s(\omega))}(dt, dx).$$

Note that here  $\Delta X_t = X_t - X_{t-}$  refers to the jump of the process and not the cemetery state. Since both notations are well established in their respective field, we will use  $\Delta$  for both. There should be no risk of confusing this. Denote the compensator of this random measure by  $\nu(dt, dx; \omega)$ . Let  $h$  denote a truncation function for the semimartingale characteristics, i.e. a bounded function  $h : V \rightarrow V$  with  $h(x) = x$  in a neighborhood of 0. The process

$$\hat{X}_t = X_t - \sum_{s \leq t} (\Delta X_s - h(\Delta X_s))$$

then has bounded jumps and is a special semimartingale with unique canonical decomposition  $\hat{X} = X_0 + \hat{M} + \hat{B}$ , where  $\hat{M}$  is a local martingale and  $\hat{B}$  is a predictable finite

<sup>3</sup>Semimartingales are strictly speaking only defined for  $\mathbb{R}^d$ -valued processes. However, for  $V$  with dimension  $d$  we fix an isomorphism between  $\mathbb{R}^d$  and  $V$  and in this case the semimartingale property is to be understood for the corresponding process on  $\mathbb{R}^d$  defined via this isomorphism.

variation process. Furthermore consider  $C = [X^c] = [\hat{M}^c]$ , which is the quadratic variation of the continuous local martingale part of the semimartingale. The triplet  $(\hat{B}, C, \nu)$  is called a version of the semimartingale characteristics of  $X$ . Note that the characteristic  $\hat{B}$  depends on the choice of the function  $h$ . They are defined up to a set of probability zero. We also call  $(\tilde{B}, \tilde{C}, \tilde{\nu})$  a version of the characteristics if they are equal to  $(\hat{B}, C, \nu)$  up to a set of probability zero.

The semimartingale property and the above decompositions depend on the filtration and the probability measure. For a Markov process  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  there are multiple measures and we make the following definition.

**Definition 1.34:** *An  $E$ -valued càdlàg Markov process  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  is a Markov semimartingale if  $X$  is a  $(\mathcal{F}, \mathbb{P}^{(s,x)})$ -semimartingale for all  $(s, x)$ .*

Note that we restrict ourselves to  $E$ -valued Markov processes. This means that we only work with conservative affine processes (see Lemma 1.24). One approach to include a cemetery state  $\Delta$  is to consider  $\Delta$  as a point in  $V \setminus E$  (see Cheridito et al. [8]). However, this causes several difficulties (compare footnotes 7 and 8) and we therefore exclude this case. Definition 1.34 only requires that the semimartingale property holds for all probability measures  $\mathbb{P}^{(s,x)}$ . For time-homogeneous Markov processes it is shown in Çinlar et al. [28] that then there also exist processes which are versions of the the semimartingale characteristics for all measures  $\mathbb{P}^x$ .

For time-homogeneous affine processes there is the following result<sup>4</sup> by Cuchiero and Teichmann [10] and Keller-Ressel et al. [37]. Denote by  $S(V)$  the positive semidefinite matrices on  $V$  and by  $\mathcal{M}(V)$  the set of (signed) measures  $\mu$  on  $V$ . Let  $h$  denote a truncation function for the semimartingale characteristics, i.e. a bounded function  $h : V \rightarrow V$  with  $h(x) = x$  in a neighborhood of 0.

**Theorem 1.35:** *Let  $(X, \mathcal{F}, \{\mathbb{P}^x\})$  be the  $E$ -valued càdlàg canonical realization of a conservative time-homogeneous continuously affine transition function. Then there exist  $b \in V$ ,  $a \in S(V)$ ,  $m \in \mathcal{M}(V)$  and restrictions of linear maps  $\beta : E \rightarrow V$ ,  $\alpha : E \rightarrow S(V)$  and  $M : E \rightarrow \mathcal{M}(V)$ , such that for all  $u \in \mathcal{U}$*

$$F(u) := \left. \frac{\partial}{\partial t} \phi_t(u) \right|_{t=0^+}, \quad R(u) := \left. \frac{\partial}{\partial t} \psi_t(u) \right|_{t=0^+}$$

exist, i.e.  $X$  is regular.  $F$  and  $R$  are given by

$$\begin{aligned} F(u) &= \frac{1}{2} \langle u, au \rangle + \langle b, u \rangle + \int_V (e^{\langle \xi, u \rangle} - 1 - \langle h(\xi), u \rangle) m(d\xi), \\ \langle R(u), x \rangle &= \frac{1}{2} \langle u, \alpha(x)u \rangle + \langle \beta(x), u \rangle + \int_V (e^{\langle \xi, u \rangle} - 1 - \langle h(\xi), u \rangle) M(d\xi; x). \end{aligned} \tag{1.37}$$

---

<sup>4</sup>Here this result is formulated for conservative transition functions.

$\Phi$  and  $\psi$  satisfy generalized Riccati differential equations for  $t < \sigma(u) = \inf\{s \geq 0 : \Phi_s(u) = 0\}$ , i.e.

$$\begin{aligned} \frac{\partial}{\partial t} \Phi_t(u) &= \Phi_t(u) F(\psi_t(u)), & \Phi_0(u) &= 1, \\ \frac{\partial}{\partial t} \psi_t(u) &= R(\psi_t(u)), & \psi_0(u) &= u. \end{aligned} \tag{1.38}$$

Furthermore  $X$  is a Markov semimartingale, where  $(B, C, \nu)$  given by

$$\begin{aligned} B_t &= \int_0^t b + B(X_{v-}) dv, \\ C_t &= \int_0^t a + A(X_{v-}) dv, \\ \nu(dt, d\xi) &= (m(d\xi) + M(d\xi; X_{t-})) dt, \end{aligned}$$

are a version of the characteristics of  $X$  under each measure  $\mathbb{P}^x$ .

*Remark:* Note that on general state spaces  $E$  we cannot say more about the parameters  $b, \beta, a, \alpha, m, M$ . For a given state space these parameters have to satisfy additional conditions, which guarantee that the process stays within the state space. For the canonical state space  $\mathbb{R}_{\geq 0}^m \times \mathbb{R}^n$  such admissibility conditions have been formulated in Duffie et al. [16]. Using these admissibility conditions they give a characterization of time-homogeneous continuously affine processes on the canonical state space. Admissibility conditions have also been derived for matrix-valued affine processes in Cuchiero et al. [11].

We want to extend the results of Theorem 1.35 to time-inhomogeneous affine processes. The following examples show that this cannot be achieved in full generality. There are processes with a continuously affine transition function, which are not semimartingales and there are time-inhomogeneous Markov semimartingales, which are not regular. Some of the examples are presented with Markov processes in the sense of Definition 1.3, since the essential problems are easier to spot. However, given the transition function one can always construct a Markov process in the sense of Definition 1.6, which then yields a counterexample in this class of Markov processes.

*Example 1:* Every deterministic real-valued function  $f$  generates a time-inhomogeneous affine process via the transition function

$$P_{s,t}(x, d\xi) = \delta_{x+f(t)-f(s)}(d\xi).$$

By Lemma 1.4 there exist measures  $\mathbb{P}^{(s,x)}$  such that the canonical realization of the transition function,  $(X, \mathcal{F}^X, \{\mathbb{P}^{(s,x)}\})$ , is affine with

$$\begin{aligned} \Phi_{s,t}(u) &= e^{u(f(t)-f(s))}, \\ \psi_{s,t}(u) &= u. \end{aligned} \tag{1.39}$$

In this case the process  $X$  is  $\mathbb{P}^{(s,x)}$ -a.s. described by

$$X_t^{s,x} = x + f(s+t) - f(s), \quad t \geq 0.$$

If the function  $f$  is continuous, the affine transition function is continuous. If additionally  $f$  is not of finite variation, the Markov process is not a semimartingale (Proposition I.4.28, Jacod and Shiryaev [29]).

A non-deterministic example is the following.

*Example 2:* Let  $X$  be an affine process on  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$  with  $E = [a, \infty)$ ,  $a > 0$  (e.g. a shifted CIR process) and  $f : \mathbb{R}_{\geq 0} \rightarrow [1, \infty)$  a continuous function, which is not of finite variation. Consider the process  $Y_t = f(t)X_t$ , which then has the same state. Then for  $A \in \mathcal{B}(E)$ ,  $Y_t^{-1}(A) = X_t^{-1}(f(t)^{-1}A) \in \mathcal{F}_t$  and

$$\mathbb{E}[e^{\langle u, Y_t \rangle} | \mathcal{F}_s] = \mathbb{E}[e^{\langle f(t)u, X_t \rangle} | \mathcal{F}_s] = \Phi_{s,t}(f(t)u)e^{\langle \psi_{s,t}(f(t)u), X_s \rangle}.$$

Hence  $Y$  is an affine process with

$$\Phi_{s,t}^Y(u) = \Phi_{s,t}^X(f(t)u), \quad \psi_{s,t}^Y(u) = \frac{\psi_{s,t}(f(t)u)}{f(s)}.$$

The transition operator of  $Y$  is defined by

$$P_{s,t}^Y g(y) = \int_E g(f(t)\xi) P_{s,t}^X(f(s)^{-1}y, d\xi).$$

If  $X$  is a semimartingale, by Itô's Lemma  $X^{-1}$  is also a semimartingale. Assume that  $Y$  is a semimartingale. Then also  $f(t) = Y_t X_t^{-1}$  is a semimartingale. Since  $f$  is not of finite variation this cannot be true. Hence  $Y$  is not a semimartingale.

Moreover, time-inhomogeneous continuously affine transition functions may not be regular. Here regularity is defined as follows.

**Definition 1.36:** A transition function  $\{P_{s,t}\}$  on  $E$  is regular<sup>5</sup> if

$$\partial_s^- P_{s,t} e^{\langle u, x \rangle} \Big|_{s=t} = - \lim_{s \uparrow t} \frac{1}{t-s} (P_{s,t} f_u(x) - f_u(x))$$

exists for all  $(t, x, u) \in \mathbb{R}_{>0} \times E \times \mathcal{U}$  and is continuous at  $u = 0$  for all  $(t, x) \in \mathbb{R}_{>0} \times E$ .

*Remark:* For affine transition function an application of the chain rule together with Assumption 1 gives that this is equivalent to the existence of the left derivatives of  $\Phi$  and  $\psi$ , i.e.  $\partial_s^- \Phi_{s,t}(u) \Big|_{s=t}$ ,  $\partial_s^- \psi_{s,t}(u) \Big|_{s=t}$ .

---

<sup>5</sup>This is called weakly regular in Definition 2.3 of Filipovic [21].

*Example 3:* Let  $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  be a continuous, nondecreasing function and  $X$  a time-homogeneous continuously affine process on  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$ . Then the time-transformed process  $Y_t = X_{f(t)}$  satisfies

$$\mathbb{E} [e^{\langle u, Y_T \rangle} | \mathcal{F}_{f(t)}] = \Phi_{f(T)-f(t)}(u) e^{\langle \psi_{f(T)-f(t)}(u), Y_t \rangle}.$$

This corresponds to the transition function

$$P_{s,t}^Y(x, \cdot) = P_{f(t)-f(s)}^X(x, \cdot).$$

$\{P_{s,t}^Y\}$  is again affine with  $\Phi_{s,t}^Y(u) = \Phi_{f(t)-f(s)}^X(u)$  and  $\psi_{s,t}^Y(u) = \psi_{f(t)-f(s)}^X(u)$ . Since  $\Phi^X$  and  $\psi^X$  are continuous,  $\Phi^Y$  and  $\psi^Y$  are also continuous.

The Cantor function is a continuous nondecreasing function  $\tilde{f} : [0, 1] \rightarrow [0, 1]$  such that the derivative of  $\tilde{f}$  is zero almost everywhere and does not exist otherwise (this function is therefore not absolutely continuous). Set  $f = \tilde{f}$  on  $[0, 1]$  and  $f(t) = 1$  for  $t \geq 1$ . With this  $f$  the transition function  $\{P_{s,t}^Y\}$  is not regular. A concrete example is  $X_t = B_{f(t)}$ , where  $B$  is a Brownian motion.

The final example shows that regularity itself does not imply that  $X$  is a semimartingale. Hence the semimartingale property and regularity have to be treated separately.

*Example 4:* Consider the function

$$g(t) = \begin{cases} \frac{t \sin(1/t)}{\ln(t/2)} & 0 < t \leq 1, \\ 0 & t \leq 0. \end{cases}$$

$g$  is differentiable on  $[0, 1]$ , but has unbounded variation (see Guzman [27], p.294). Note that the derivative of this function diverges and oscillates for  $t \rightarrow 0$ . Using  $f(t) = g(1 - t)$  in Example 1 yields an affine process, which is regular, but not a semimartingale.

### 1.4.1 Affine Markov semimartingales

Let  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  be a Markov semimartingale in the sense of Definition 1.34 with a continuously affine transition function. So far we only required that the semimartingale property holds for all probability measures  $\mathbb{P}^{(s,x)}$ . To find processes which are versions of the characteristics under each measure  $\mathbb{P}^{(s,x)}$  we use the setup of Çinlar et al. [28]. Hence we work with the space-time process  $\tilde{X} = (\Theta, X)$  introduced in Lemma 1.13. Note that with a continuously affine transition function we can use the product  $\sigma$ -algebras  $\tilde{\mathcal{E}} := \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{E}$  on  $\tilde{E} := \mathbb{R}_{\geq 0} \times E$  and  $\tilde{\mathcal{A}} := \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{A}$  on  $\tilde{\Omega} := \mathbb{R}_{\geq 0} \times \Omega$ . The filtration  $\tilde{\mathcal{F}}$  is given by  $\tilde{\mathcal{F}}_t = \mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{F}_t$ . Then  $(\tilde{X}, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(s,x)}\})$  is a time-homogeneous Markov process on  $(\tilde{\Omega}, \tilde{\mathcal{A}})$  with state space  $(\tilde{E}, \tilde{\mathcal{E}})$ . Since on the extended space  $(\tilde{\Omega}, \tilde{\mathcal{A}})$   $X(s, \omega) = X(\omega)$ , we use the same letter  $X$  to denote the

process on  $(\Omega, \mathcal{A})$  and on  $(\tilde{\Omega}, \tilde{\mathcal{A}})$ . If unclear we specifically refer to the underlying probability space.

We assume that the filtration  $\tilde{\mathcal{F}}$  is a right-continuous strong Markov filtration (see Çinlar et al. [28]). Note that the definition of a strong Markov filtration requires the existence of shift operators  $\theta_t$  for  $X$  and  $\Omega$ , which are extended to  $\tilde{X}$  and  $\tilde{\Omega}$  by  $\tilde{\theta}_t(s, \omega) = (s + t, \theta_t(\omega))$ . For the canonical space-time realization of the continuously affine transition function this is always true (see section 1.3). With this setup we can use the results of Çinlar et al. [28].

**Definition 1.37:** *A stochastic process  $F$  is called additive, if  $\tilde{\mathbb{P}}^{(s,x)}$ -a.s. for all  $(s, x)$   $F_0 = 0$  and*

$$F_{v+t}(r, \omega) = F_v(r, \omega) + F_t \circ \tilde{\theta}_v(r, \omega) = F_v(r, \omega) + F_t(r + v, \theta_v(\omega)).$$

**Lemma 1.38:** *There exist*

1. *an adapted additive process  $F(r, \omega)$ , which is nondecreasing from  $F_0 = 0$  a.s. for each measure  $\tilde{\mathbb{P}}^{(s,x)}$ , such that  $F$  is  $\tilde{\mathbb{P}}^{(s,x)}$ -indistinguishable from a predictable process for all  $(s, x)$ .*
2. *Optional processes  $b_t(r, \omega)$  and  $c_t(r, \omega)$  with values in  $V$ , respectively  $S(V)$ ,*
3. *a positive kernel  $K_t(dy; (r, \omega))$  from  $(\tilde{\Omega}, \mathcal{O}(\tilde{\mathcal{F}}_t))^6$  giving measures in  $\mathcal{M}(V)$ ,*

*such that*

$$B = b \cdot F, \quad C = c \cdot F, \quad \nu(dt, d\xi; (r, \omega)) = K_t(d\xi; (r, \omega)) dF_t(r, \omega) \quad (1.40)$$

*are a version of the semimartingale characteristics of  $X$  on  $(\tilde{\Omega}, \tilde{\mathcal{A}})$  under each measure  $\tilde{\mathbb{P}}^{(s,x)}$ . We say that  $(b, c, K, F)$  give a version of the semimartingale characteristics of  $(X, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(s,x)}\})$ .*

*Proof.* If  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  is a Markov semimartingale, also  $(X, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(s,x)}\})$  is a Markov semimartingale. Furthermore  $X_t - X_0$  is an additive process. Applying Theorem 6.25 in Çinlar et al. [28] to  $X - X_0$  gives the result.<sup>7</sup>  $\square$

The function  $F$  from Lemma 1.38 is far from uniquely defined. Motivated by Example 3 we expect that the function  $F$  from Lemma 1.38 can even be chosen deterministic (by deterministic we here always mean independent of  $\omega$ ). The idea is to construct such a candidate and show that  $F$  is ‘absolutely continuous’ with respect to this candidate. We then go on and show that when  $F$  is deterministic,  $b$ ,  $c$  and  $K$  can be expressed as functions of  $(\Theta, X)$ , which are affine in  $X$ . Note that we cannot expect that this is true for a non-deterministic  $F$ .

<sup>6</sup> $\mathcal{O}(\tilde{\mathcal{F}}_t)$  denotes the  $\tilde{\mathcal{F}}_t$ -optional  $\sigma$ -algebra (see Çinlar et al. [28]).

<sup>7</sup> Note that  $X^\tau - X_0$  is no longer an additive process for general stopping times  $\tau$ , which is why we cannot directly use Theorem 6.25 for exploding processes by stopping the process before it explodes.

*Example 5:* Assume  $F$  to be independent of  $\omega$  and choose a bounded function  $f : E \rightarrow \mathbb{R}_{>0}$ . Define  $\tilde{F}_t = \int_0^t f(\tilde{X}_s) dF_s$ . Then  $\tilde{F}$  is additive, nondecreasing and  $\tilde{F}_0 = 0$ , but  $\tilde{F}$  depends on  $\omega$ . The characteristics of  $\tilde{X}$  can also be written as integrals with respect to any such  $\tilde{F}$ .

**Definition 1.39:** We call an affine transition function of finite variation (FV) if the maps  $t \mapsto \Phi_{t,T}(u)$  and  $t \mapsto \langle \psi_{t,T}(u), x \rangle$  for  $x \in E$  are of finite variation on  $[v, T]$  for every  $v > o(T, u)$ .

*Remark:* Let  $\{P_{s,t}\}$  be a continuously affine transition function of a Markov semimartingale. We conjecture that the affine transition function is automatically of FV (see also Lemma 1.44). However, so far we were not able to prove this.

For a Markov semimartingale with an affine transition function of finite variation we construct a deterministic candidate for  $F$  as follows. Fix an affine basis of  $E$ , which exists by Assumption 1, and set  $\psi_{t,T}^i(u) := \langle \psi_{t,T}(u), x^i - x^0 \rangle$ . Set  $\tilde{o}(T, u) = \frac{1}{2}(o(T, u) + T)$  and define the nondecreasing continuous functions

$$G^{T,u}(t) := \int_{t \wedge \tilde{o}(T,u)}^{t \wedge T} d|\Phi_{s,T}(u)| + \sum_{i=1}^d d|\psi_{s,T}^i(u)|,$$

where we integrate with respect to  $s$  and  $d|f(s)|$  denotes the total variation of  $f$ . Consider a countable dense subset of  $\mathbb{R}_{>0} \times iV$  and denote the indexed elements of this set by  $(T_i, u_i), i \in \mathbb{N}$ . From now on when writing indexed  $T_i$  or  $u_i$ , this always refers to points in this set. Define the weights

$$w_i = 2^{-i} \frac{1}{G^{T_i, u_i}(T_i)} \mathbb{I}\{G^{T_i, u_i}(T_i) > 0\}$$

and set

$$G(t) := t + \sum_{i \in \mathbb{N}} w_i G^{T_i, u_i}(t). \quad (1.41)$$

The weights guarantee that this sum converges. The function  $G$  is continuous, strictly increasing and  $|G(t)| \leq 1 + t$ . We use the function  $G$  to define the process  $G_t(r, \omega) := G_t(r) := G(r + t) - G(r)$ .  $G$  is independent of  $\omega$  and additive, i.e.  $G_{s+t}(r) = G_s(r) + G_t(s + r)$ . This is the mentioned candidate. With this construction  $d\Phi_{\cdot, T_i}(u_i)$  and  $d\psi_{\cdot, T_i}^j(u_i), j = 1, \dots, d$ , are absolutely continuous w.r.t.  $dG$  on  $[\tilde{o}(T_i, u_i), T_i]$  for all  $i$ . Denote their (deterministic) densities with respect to  $dG$  by  $s \mapsto f^\Phi(s; T_i, u_i)$  and  $s \mapsto f^\psi(s; T_i, u_i)$ .

Fix  $(s, x)$  and for  $T > s$  and  $u \in iV$  consider the uniformly bounded  $(\tilde{\mathcal{F}}, \tilde{\mathbb{P}}^{(s,x)})$ -martingales<sup>8</sup>

$$M_t^{T-s, u} = \mathbb{E}^{(s,x)} \left[ e^{\langle u, X_{T-s} \rangle} | \tilde{\mathcal{F}}_t \right] = \Phi_{s+t, T}(u) e^{\langle \psi_{s+t, T}(u), X_t \rangle}, \quad 0 \leq t \leq T - s. \quad (1.42)$$

<sup>8</sup> If one considers the cemetery state  $\Delta$  as a point in  $V \setminus E$ , the second identity in equation (1.42) no longer holds for the classical exponential function. One has to use a modified function  $f$  which is equal to  $e^{\langle u, x \rangle}$  on  $\mathcal{U} \times E$  and satisfies  $f(u, \Delta) = 0$ . However in this case equation (1.44) has to be modified to correct the jumps to  $\Delta$ .

By Theorem II.34 Jacod and Shiryaev [29] and Lemma 1.38  $X$  can be decomposed as

$$X_v = x + \int_0^v b_t dF_t + \int_0^v \int_V (\xi - h(\xi)) \mu^X(dt, d\xi) + N_v, \quad (1.43)$$

where  $\mu^X$  is the random measure associated with the jumps of  $X$  and  $N$  is a  $(\tilde{\mathcal{F}}, \tilde{\mathbb{P}}^{(s,x)})$ -local martingale (dependences on  $(r, \omega)$  are suppressed). Integration by parts and Itô's Lemma applied to  $M_t^{T-s,u}$  together with the properties of the semimartingale characteristics give for  $r \geq 0$  and  $o(T, u) - s < r \leq v \leq T - s$

$$\begin{aligned} M_v^{T-s,u} - M_r^{T-s,u} &= \int_r^v M_{t-}^{T-s,u} \frac{d\Phi_{s+t,T}(u)}{\Phi_{s+t,T}(u)} + \Phi_{s+t,T}(u) d\langle \psi_{s+t,T}(u), X_t \rangle \\ &= \int_r^v M_{t-}^{T-s,u} \left( \frac{d\Phi_{s+t,T}(u)}{\Phi_{s+t,T}(u)} + d\langle \psi_{s+t,T}(u), X_t \rangle + d[\langle \psi_{s+t,T}(u), X_t \rangle]^c \right) \\ &\quad + \int_r^v \int_V M_{t-}^{T-s,u} (e^{\langle \psi_{s+t,T}(u), \xi \rangle} - 1 - \langle \psi_{s+t,T}(u), \xi \rangle) \mu^X(dt, d\xi) \\ &= \int_r^v M_{t-}^{T-s,u} \left( \frac{d\Phi_{s+t,T}(u)}{\Phi_{s+t,T}(u)} + \langle d\psi_{s+t,T}(u), X_{t-} \rangle + \langle \psi_{s+t,T}(u), dB_t \rangle \right. \\ &\quad \left. + \langle \psi_{s+t,T}(u), dC_t \psi_{s+t,T}(u) \rangle + \langle \psi_{s+t,T}(u), dN_t \rangle \right) \\ &\quad + \int_r^v \int_V M_{t-}^{T-s,u} (e^{\langle \psi_{s+t,T}(u), \xi \rangle} - 1 - \langle \psi_{s+t,T}(u), h(\xi) \rangle) \mu^X(dt, d\xi). \end{aligned}$$

Compensating with  $\nu$  and using (1.40) gives

$$\begin{aligned} M_v^{T-s,u} &= M_r^{T-s,u} + (\tilde{N}_v - \tilde{N}_r) \\ &\quad + \int_r^v M_{t-}^{T-s,u} \left( \frac{d\Phi_{s+t,T}(u)}{\Phi_{s+t,T}(u)} + \langle d\psi_{s+t,T}(u), X_{t-} \rangle + \kappa(\psi_{s+t,T}(u)) dF_t \right), \end{aligned} \quad (1.44)$$

where integration is with respect to  $t$ ,  $\kappa$  is defined for  $(r, \omega) \in \tilde{\Omega}$  as

$$\kappa_t(u; (r, \omega)) := \langle u, b_t(r, \omega) \rangle + \frac{1}{2} \langle u, c_t(r, \omega) u \rangle + \int_V (e^{\langle u, \xi \rangle} - 1 - \langle u, h(\xi) \rangle) K_t(d\xi; (r, \omega))$$

and  $\tilde{N}_r$  is a local martingale given by

$$\begin{aligned} \tilde{N}_r &= \int_0^r M_{t-}^{T-s,u} \langle \psi_{s+t,T}(u), dN_t \rangle \\ &\quad + \int_0^r M_{t-}^{T-s,u} \int_V (e^{\langle \psi_{s+t,T}(u), \xi \rangle} - 1 - \langle \psi_{s+t,T}(u), h(\xi) \rangle) (\mu^X - \nu)(dt, d\xi). \end{aligned}$$

Since  $M_t^{T-s,u}$  is a martingale, the finite variation part of the right hand side of (1.44) has to vanish. For  $(T, u) = (T_i, u_i)$  this gives  $\tilde{\mathbb{P}}^{(s,x)}$  a.s. for  $t$  in  $[\tilde{o}(T_i, u_i) - s, T_i - s]$

$$\frac{f^\Phi(s+t; T_i, u_i)}{\Phi_{s+t, T_i}(u_i)} + \langle f^\psi(s+t; T_i, u_i), X_{t-} \rangle dG(s+t) = -\kappa_t(\psi_{s+t, T_i}(u_i)) dF_t. \quad (1.45)$$

For each  $i$  the set where equation (1.45) does not hold, is a set of  $\tilde{\mathbb{P}}^{(s,x)}$ -measure zero. Taking the countable union of these sets we find a set of  $\tilde{\mathbb{P}}^{(s,x)}$ -measure zero, where on the complement (1.45) holds for all  $i$  simultaneously.

Equation (1.45) almost gives absolute continuity of  $dF.(s, \omega)$  with respect to  $dG.(s)$ . If  $\kappa_t \neq 0$ , we could ‘divide’ by  $\kappa_t$  in equation (1.45) and would be mostly finished. The next lemmas show that we can essentially always find  $(T_i, u_i)$  for which we can do this.

**Lemma 1.40:** *Consider the sets*

$$\Gamma(s, \omega) := \{t > 0 : \kappa_t(\psi_{s+t, T_i}(u_i); (s, \omega)) = 0 \text{ for all } i \text{ with } s+t \in [\tilde{o}(T_i, u_i), T_i]\}.$$

On  $\Gamma(s, \omega)$  it holds that

$$b_t(s, \omega) = 0, \quad c_t(s, \omega) = 0, \quad K_t(d\xi, (s, \omega)) = 0.$$

*Proof.* Fix  $(s, \omega)$  and  $t \in \Gamma(s, \omega)$ ,  $u \in iV \subset \mathcal{U}_1$ . There is a sequence  $(T_k, u_k)$  with  $T_k \downarrow s+t$ ,  $u_k \rightarrow u$ . By passing to a subsequence, again denoted by  $(T_k, u_k)$ , one may assume that  $0 \leq o(T_k, u_k) < T_k \leq s+t+1$  and  $o(T_k, u_k) \rightarrow v \geq 0$ . The continuity of  $\Phi$  yields  $\Phi_{o(T_k, u_k), T_k}(u_k) \rightarrow \Phi_{v, s+t}(u)$ . Hence  $v = 0$  or  $\Phi_{v, s+t}(u) = 0$  by the definition of  $o(T_k, u_k)$ . In both cases  $v \leq o(s+t, u) < s+t$ . Set  $\epsilon := s+t-v > 0$ . There is an  $N$ , such that for all  $k \geq N$

$$o(T_k, u_k) \leq v + \frac{\epsilon}{2} = s+t - \frac{\epsilon}{2}.$$

Choosing  $N$  large enough that additionally  $s+t \leq T_k \leq s+t + \frac{\epsilon}{4}$ , it follows that for all  $k \geq N$

$$\tilde{o}(T_k, u_k) = \frac{T_k + o(T_k, u_k)}{2} \leq \frac{s+t + \frac{\epsilon}{4} + s+t - \frac{\epsilon}{2}}{2} = s+t - \frac{\epsilon}{8} < s+t.$$

Hence one obtains that  $\kappa_t(\psi_{s+t, T_k}(u_k); (s, \omega)) = 0$  or all  $k \geq N$ . By continuity of  $\psi$ ,  $\psi_{s+t, T_k}(u_k) \rightarrow \psi_{s+t, s+t}(u) = u$  and since  $u \mapsto \kappa_t(u; (s, \omega))$  is continuous for each  $t$ ,  $(s, \omega)$ , it follows that  $\kappa_t(u; (s, \omega)) = 0$ . This holds for all  $u \in i\mathbb{R}^d$  and by the uniqueness of the Lévy-Khintchine representation (Lemma II.2.44, Jacod and Shiryaev [29]) this transfers to  $b, c, K$ .  $\square$

For each  $(s, \omega)$  the measure  $dF.(s, \omega)$  restricted to the complement of  $\Gamma(s, \omega)$  is measure on  $\mathbb{R}_{\geq 0}$ . Denote by  $\tilde{F}_t(s, \omega)$  its distribution function with  $\tilde{F}_0(s, \omega) = 0$ . By Lemma 1.40 and (1.40)  $(b, c, K, \tilde{F})$  is again a version of the semimartingale characteristics of  $X$ . Define

$$\Gamma^* := \{(r, \omega) \in \tilde{\Omega} : d\tilde{F}.(r, \omega) \ll dG.(r)\}.$$

**Lemma 1.41:**  $\mathbb{P}^{(s,x)}(\Gamma^*) = 1$  for all  $(s, x)$ .

*Proof.* Fix  $(s, x)$ . Since  $\tilde{\mathbb{P}}^{(s,x)}(\{s\} \times \Omega) = 1$  it suffices to show  $\tilde{\mathbb{P}}^{(s,x)}(\Gamma_s^*) = 1$  with  $\Gamma_s^* = \{(s, \omega) : \omega \in \Omega, d\tilde{F}(s, \omega) \ll dG(s)\}$ . Using the associativity of integrals<sup>9</sup> we can ‘divide’ by  $\kappa_t(\psi_{s+t, T_i}(u_i); (s, \omega))$  in (1.45), when it is not equal to zero. For each  $i$  define

$$\Lambda^i(s, \omega) := \{t \geq 0 : \kappa_t(\psi_{s+t, T_i}(u_i); (s, \omega)) \neq 0 \text{ and } s+t \in [\tilde{o}(T_i, u_i), T_i]\}.$$

Fix  $(s, x)$ . Then (1.45) together with the construction of  $\tilde{F}$  yields  $\tilde{\mathbb{P}}^{(s,x)}$ -a.s. for all  $i \in \mathbb{N}$

$$\mathbb{I}_{\Lambda^i(s, \omega)}(t) d\tilde{F}_t(s, \omega) = H_t^i(s, \omega) dG_t(s) \quad (1.46)$$

with some function  $t \mapsto H_t^i(s, \omega)$ . The weighted sum  $\sum_i 2^{-i} \mathbb{I}_{\Lambda^i(s, \omega)} \leq 1$  is strictly positive on the complement of  $\Gamma(s, \omega)$ . Together with equation (1.46) this then yields that  $\tilde{\mathbb{P}}^{(s,x)}$ -a.s. on  $\Gamma_s^*$   $d\tilde{F}(s, \omega) \ll dG(s)$ .  $\square$

By Lemma 1.41 also  $(b, c, K, \bar{F})$  with  $\bar{F} = \tilde{F}\mathbb{I}_{\Gamma^*}$  give a version of the semimartingale characteristics of  $X$  under each  $\tilde{\mathbb{P}}^{(s,x)}$ .  $d\bar{F}(r, \omega) \ll dG(r)$  for all  $(r, \omega)$  and we denote the density by  $t \mapsto f(t; r, \omega)$ . Define

$$\tilde{b} = fb, \tilde{c} = fc \text{ and } \tilde{K}_t(d\xi, (r, \omega)) = K_t(d\xi, (r, \omega))f(t; r, \omega). \quad (1.47)$$

Then  $(\tilde{b}, \tilde{c}, \tilde{K}, G)$  give the same characteristics as  $(b, c, K, \bar{F})$ .

$G$  is continuous and hence by Proposition II.2.9 i) in Jacod and Shiryaev [29] it follows that  $X^\tau$  is quasi-left continuous under each measure  $\tilde{\mathbb{P}}^{(s,x)}$ . Hence we can also apply Theorem 6.27 in Çinlar et al. [28] to get that there exist

1. an adapted nondecreasing continuous additive process  $F$
2.  $\tilde{\mathcal{E}}$ -measurable functions  $b$  and  $c$  with values in  $V$  and  $S(V)$
3. a positive kernel  $K(dy; t, x)$  from  $(\tilde{E}, \tilde{\mathcal{E}})$  giving measures in  $\mathcal{M}(V)$ ,

such that  $(B, C, \nu)$  defined by

$$\begin{aligned} B_t(r, \omega) &= \int_0^t b(\Theta_v(r), X_{v-}(\omega)) dF_v(r, \omega), \\ C_t(r, \omega) &= \int_0^t c(\Theta_v(r), X_{v-}(\omega)) dF_v(r, \omega), \\ \nu(dt, d\xi; (r, \omega)) &= K(d\xi; \Theta_t(r), X_{t-}(\omega)) dF_t(r, \omega) \end{aligned} \quad (1.48)$$

are a version of the characteristics of  $X$  under each  $\tilde{\mathbb{P}}^{(s,x)}$ . Construct  $\bar{F}$  as before, so that  $(b, c, K, \bar{F})$  also give a version of the semimartingale characteristics and  $d\bar{F}(r, \omega) \ll dG(r)$ .  $G$  is continuous and strictly increasing. We define a time-change

$$\eta_t(r) = \eta_t(r, \omega) := \inf\{v \geq 0 : G(v+r) - G(r) \geq t\}. \quad (1.49)$$

---

<sup>9</sup>See Proposition 0.4.10 Revuz and Yor [42] or Falkner and Teschl [20] for a discussion on the rules for exchanging Stieltjes integrals.

Then for  $\hat{F}_t(r, \omega) = \bar{F}_{\eta_t(r)}(r, \omega)$  we have  $d\hat{F} \ll dt$  (note that  $G_{\eta_s(r)} - G_0(r) = s$ ). The first part of the proof of Theorem 3.55 in Çinlar et al. [28] together with Proposition 3.56 applied to  $\hat{F}$  shows that there exists a nonnegative  $\tilde{\mathcal{E}}$ -measurable function  $h$ , such that

$$\bar{F}_t(r, \omega) = \int_0^t h(\Theta_v(r), X_{v-}(\omega)) dG_v(r)$$

$\tilde{\mathbb{P}}^{(s,x)}$ -a.s. under each  $(s, x)$ . Set  $\bar{b} = bh$ ,  $\bar{c} = ch$ ,  $\bar{K} = Kh$ . Then  $(\bar{b}, \bar{c}, \bar{K}, G)$  give a version of the characteristics of  $X$  via (1.48) with  $(b, c, K, F)$  replaced by  $(\bar{b}, \bar{c}, \bar{K}, G)$ . We can summarize this in the following lemma.

**Lemma 1.42:** *A version of the characteristics of  $(X, \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(s,x)}\})$  is given by  $(\bar{b}, \bar{c}, \bar{K}, G)$  via equation (1.48), where*

1.  $G$  is an additive process given by  $G_t(r, \omega) := G(r+t) - G(r)$  for a deterministic strictly increasing continuous function  $G(t)$  with  $G(0) = 0$ .
2.  $\bar{b}(t, x)$  and  $\bar{c}(t, x)$  are  $\tilde{\mathcal{E}}$ -measurable functions with values in  $V$  and  $S(V)$ ,
3.  $\bar{K}$  is a positive kernel  $\bar{K}(dy; t, x)$  from  $(\tilde{E}, \tilde{\mathcal{E}})$  giving measures in  $\mathcal{M}(V)$ .

To show a theorem in the spirit of Theorem 1.35 we use the affine basis from Assumption 1. To do this we consider independent copies of  $X$  starting at different values. Define

- the probability space  $\hat{\Omega} := \mathbb{R}_{\geq 0} \times \Omega^{d+1}$ ,
- the process  $\hat{X}(\hat{\omega}) := (\Theta(s), X(\omega^0), \dots, X(\omega^d))$ , where  $\hat{\omega} = (s, \omega^0, \dots, \omega^d)$ ,
- the filtration  $\hat{\mathcal{F}}$  defined by  $\hat{\mathcal{F}}_t := \mathcal{B}(\mathbb{R}_{\geq 0}) \times (\otimes_{i=0}^d \mathcal{F}_t)$ ,
- for  $s \geq 0$  and  $x = (x^0, \dots, x^d) \in E^{d+1}$  the probability measures

$$\hat{\mathbb{P}}^{(s,x)} := \delta_s \otimes (\otimes_{i=0}^d \mathbb{P}^{(s,x^i)}).$$

Then for each  $s \geq 0$  the processes  $\hat{X}^0, \dots, \hat{X}^d$  with  $\hat{X}^i(\hat{\omega}) = X(\omega^i)$  are independent,  $\hat{\mathbb{P}}^{(s,x)}(\hat{X}_0^i = x^i) = 1$  for  $i = 0, \dots, d$ , and each component  $\hat{X}^i$  behaves like the time-inhomogeneous affine process starting in  $(s, x^i)$ . I.e. each component  $\hat{X}^i$  has semimartingale characteristics as described in Lemma 1.42 (with the same  $G$ ) depending only on  $(s, \omega^i)$ , respectively  $(\Theta_t(s), X_{t-}(\omega^i))$  (and not  $\hat{\omega}$ )

For the affine basis  $(x^0, \dots, x^d)$  we can express  $\hat{X}_t^i$  in terms of the basis  $(x^1 - x^0, \dots, x^d - x^0)$ , i.e. we define processes  $X_{t,j}^i$  by  $\hat{X}_t^i = \sum_{j=1}^d X_{t,j}^i(x^j - x^0)$  and consider the matrix-valued stochastic process

$$H_t := H(\hat{X}_t^0, \dots, \hat{X}_t^d) := \begin{pmatrix} 1 & X_{t,1}^0 & \dots & X_{t,d}^0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{t,1}^d & \dots & X_{t,d}^d \end{pmatrix}, \quad (1.50)$$

$$Y_t = \inf_{0 \leq s \leq t} |\det H_s|. \quad (1.51)$$

Set  $\bar{x} = (x^0, \dots, x^d)$ . We have the following lemma (compare Keller-Ressel and Mayerhofer [34]).

**Lemma 1.43:** *For each  $s > 0$  there exists  $\delta(s) > 0$  such that*

$$\hat{\mathbb{P}}^{(s, \bar{x})}(\det H_t \neq 0 \text{ for all } 0 \leq t \leq \delta(s)) > 0.$$

*Proof.* Fix  $s > 0$  and set  $g(t) = \hat{\mathbb{P}}^{(s, \bar{x})}(Y_t > 0)$ . We first show that  $g(t)$  is left-continuous. Let  $t_k \uparrow t$ .  $g(0) = 1$  and  $g(t)$  is decreasing. Hence  $g(t_k)$  converges. Since  $X$  is stochastically continuous and Markov, it is also continuous in probability under  $\mathbb{P}^{(s, x)}$  and hence  $Y_t$  as well. Then there is a subsequence  $t_{k_m}$ , such that  $Y_{t_{k_m}} \downarrow Y_t$  a.s.. By dominated convergence also  $g(t_{k_m}) \downarrow g(t)$ . Hence  $g(t_k)$  also converges to  $g(t)$  and we get that  $g(t)$  is left-continuous. Since  $Y_t$  is a.s. right-continuous by dominated convergence this also holds for  $g(t)$ . Hence  $g(t)$  is continuous in  $t$  and we can define

$$\delta(s) := \inf\{t > 0 : \hat{\mathbb{P}}^{(s, \bar{x})}(Y_t > 0) = \frac{1}{2}\} > 0.$$

□

We conjecture that a uniform version of Lemma 1.43 is true for a continuously affine transition function, i.e. that for each  $T > 0$  there exists  $\delta > 0$  such that for all  $0 \leq s \leq T$   $\hat{\mathbb{P}}^{(s, \bar{x})}(\det H_t \neq 0 \text{ for all } 0 \leq t \leq \delta) > 0$ . In this case the affine transition function is of finite variation.

**Lemma 1.44:** *Assume that for each  $T > 0$  there exists  $\delta > 0$  such that for all  $0 \leq s \leq T$*

$$\hat{\mathbb{P}}^{(s, \bar{x})}(\det H_t \neq 0 \text{ for all } 0 \leq t \leq \delta) > 0.$$

*Then the functions  $t \mapsto \Phi_{t,T}(u)$  and  $t \mapsto \langle \psi_{t,T}(u), x \rangle$  are of finite variation on  $[v, T]$  for every  $v > o(T, u)$  and  $x \in E$ .*

*Proof.* Fix  $T > 0$ ,  $u \in \mathcal{U}$ . For  $o(T, u) \leq s \leq T$ ,  $0 \leq i \leq d$  consider the  $(\hat{\mathcal{F}}, \hat{\mathbb{P}}^{(s, \bar{x})})$ -martingales

$$\hat{M}_t^{T-s, u, i} = \Phi_{s+t, T}(u) e^{\langle \psi_{s+t, T}(u), \hat{X}_t^i \rangle}, \quad 0 \leq t \leq T - s.$$

By assumption there is  $\delta > 0$  and a set  $\Lambda(s)$  with  $\hat{\mathbb{P}}^{(s, \bar{x})}(\Lambda(s)) > 0$ , such that  $H_t$  is invertible on  $[0, \delta]$ . Let  $\hat{\tau}$  be a stopping time such that  $H_t$  is invertible on  $[0, \hat{\tau}]$  and  $\hat{\tau} \geq \delta$  on  $\Lambda(s)$ . We then have  $\hat{\mathbb{P}}^{(s, \bar{x})}$ -a.s. that

$$\begin{pmatrix} \ln(\Phi_{s+t \wedge \hat{\tau}, T}(u)) \\ \psi_{s+t \wedge \hat{\tau}, T}^1(u) \\ \vdots \\ \psi_{s+t \wedge \hat{\tau}, T}^d(u) \end{pmatrix} = \begin{pmatrix} 1 & X_{t \wedge \hat{\tau}, 1}^0 & \cdots & X_{t \wedge \hat{\tau}, d}^0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & X_{t \wedge \hat{\tau}, 1}^d & \cdots & X_{t \wedge \hat{\tau}, d}^d \end{pmatrix}^{-1} \begin{pmatrix} \ln(\hat{M}_{t \wedge \hat{\tau}}^{T-s, u, 0}) \\ \vdots \\ \ln(\hat{M}_{t \wedge \hat{\tau}}^{T-s, u, d}) \end{pmatrix}. \quad (1.52)$$

On  $\Lambda(s)$  and  $[0, \delta \wedge (T - s)]$  the left side is deterministic and coincides with a semimartingale. Hence  $\Phi$  and  $\psi$  are of finite variation on  $[s, (s + \delta) \wedge T]$ . Let  $v > o(T, u)$ . Then we can cover  $[v, T]$  with the open intervals  $(s, s + \delta)$ ,  $s \in [v, T]$  and by compactness there is a finite subcover. Since  $\Phi$  and  $\psi$  are of finite variation on  $[s, (s + \delta) \wedge T]$ , they are of finite variation on  $[v, T]$ .  $\square$

While it seems plausible that the assumption of Lemma 1.44 is fulfilled for a continuously affine transition function, the following counterexample shows that this is not true for all affine transition functions.

*Example 6:* Construct a time-inhomogeneous Markov processes, which is  $\mathbb{P}^{(s,x)}$ -a.s. equal to

$$X_t^{s,x} := \mathbb{I}\{s < 1\} ((1 - s - t)B_t^x \mathbb{I}\{s + t \leq 1\} + (B_t^x - B_{1-s}^x) \mathbb{I}\{s + t > 1\}) + \mathbb{I}\{s \geq 1\} B_t^x,$$

where  $B_t^x$  is a Brownian motion starting in  $x$ . This is an affine process with

$$\psi_{s,t}(u) = \begin{cases} s \leq t < 1: & u \frac{1-t}{1-s} \\ s < 1 \leq t: & 0 \\ 1 \leq s \leq t: & u \end{cases} \quad \phi_{s,t}(u) = \begin{cases} s \leq t < 1: & \frac{1}{2}u^2(t-s)(1-t)^2 \\ s < 1 \leq t: & \frac{1}{2}u^2(t-1) \\ 1 \leq s \leq t: & \frac{1}{2}u^2(t-s) \end{cases}$$

This process is not stochastically continuous in the sense of Definition 1.9. In particular, for the sequence  $(s, t)_n := (1 - \frac{1}{n}, 1) \rightarrow (1, 1)$  stochastic continuity (in  $s$ ) fails. Moreover we cannot find a uniform  $\delta$  on the interval  $[0, 1]$ , since  $\delta(s)$  from Lemma 1.43 in this case satisfies  $\delta(s) \leq 1 - s$  for  $0 \leq s \leq 1$ .

We now formulate and prove the central theorem of this section.

**Theorem 1.45:** *Let  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  be a Markov semimartingale whose transition function  $\{P_{s,t}\}$  is continuously affine of finite variation. Then there exist a deterministic  $\mathbb{R}_{\geq 0}$ -valued strictly increasing continuous function  $G$ , maps  $b, a, m$  from  $\mathbb{R}_{\geq 0}$  to  $V$ ,  $S(V)$  and  $\mathcal{M}(V)$  and maps  $\beta : \mathbb{R}_{\geq 0} \times E \rightarrow V$ ,  $\alpha : \mathbb{R}_{\geq 0} \times E \rightarrow S(V)$  and  $M : \mathbb{R}_{\geq 0} \times E \rightarrow \mathcal{M}(V)$ , which for  $dG$ -a.e.  $t \geq 0$  are restrictions of linear maps on  $E$ , such that for all  $u \in \mathcal{U}$ ,  $\Phi$  and  $\psi$  satisfy generalized Ricatti integral equations*

$$\begin{aligned} \Phi_{s,T}(u) &= 1 + \int_s^T \Phi_{t,T}(u) F(t, \psi_{t,T}(u)) dG(t), \\ \psi_{s,T}(u) &= u + \int_s^T R(t, \psi_{t,T}(u)) dG(t), \end{aligned} \tag{1.53}$$

where

$$\begin{aligned} F(t, u) &= \frac{1}{2} \langle u, a(t)u \rangle + \langle b(t), u \rangle + \int_V (e^{\langle \xi, u \rangle} - 1 - \langle h(\xi), u \rangle) m(d\xi; t), \\ \langle R(t, u), x \rangle &= \frac{1}{2} \langle u, \alpha(t, x)u \rangle + \langle \beta(t, x), u \rangle \\ &\quad + \int_V (e^{\langle \xi, u \rangle} - 1 - \langle h(\xi), u \rangle) M(d\xi; t, x). \end{aligned} \tag{1.54}$$

If furthermore  $\tilde{\mathcal{F}}$  is a right-continuous strong Markov filtration on the extended probability space  $(\tilde{\Omega}, \tilde{\mathcal{A}})^{10}$ , then

$$\begin{aligned} B_t &= \int_0^t b(\Theta_v) + \beta(\Theta_v, X_{v-}) dG(\Theta_v), \\ C_t &= \int_0^t a(\Theta_v) + \alpha(\Theta_v, X_{v-}) dG(\Theta_v), \\ \nu(dt, d\xi) &= (m(d\xi; \Theta_t) + M(d\xi; \Theta_t, X_{t-})) dt \end{aligned} \tag{1.55}$$

are a version of the semimartingale characteristics of  $X$  under each measure  $\hat{\mathbb{P}}^{(s,x)}$ . Here  $\Theta_t(r, \omega) = t + r$ .

*Remark:* Equation (1.53) reads in terms of  $\phi$

$$\phi_{s,T}(u) = \int_s^T F(t, \psi_{t,T}(u)) dG(t).$$

*Proof.* If  $\tilde{\mathcal{F}}$  is not a strong Markov filtration we consider instead of  $X$  the canonical space-time realization of section 1.3. Note that if  $(X, \mathcal{F}, \{\mathbb{P}^{(s,x)}\})$  is a Markov semimartingale, it is also a Markov semimartingale with respect to the smaller filtration generated by  $X$  and then also the canonical space-time realization is a Markov semimartingale (we only add null sets to the filtration) with a right-continuous strong Markov filtration  $\tilde{\mathcal{F}}$ . Hence without loss of generality we may assume that  $\tilde{\mathcal{F}}$  is a strong right-continuous Markov filtration.

Fix  $u \in \mathcal{U}$  and  $s \geq 0$  and consider the  $(\hat{\mathcal{F}}, \hat{\mathbb{P}}^{(s,\bar{x})})$ -martingales introduced in the proof of Lemma 1.44,

$$\hat{M}_t^{T-s,u,j} = \Phi_{s+t,T}(u) e^{\langle \psi_{s+t,T}(u), \hat{X}_t^j \rangle}, \quad 0 \leq t \leq T - s.$$

For  $(T, u) = (T_i, u_i)$  and  $j = 0, \dots, d$  we can repeat the steps leading to (1.45) using the characteristics from Lemma 1.42. This then gives that  $\hat{\mathbb{P}}^{(s,x)}$ -a.s. on  $\{s\} \times \hat{\Omega}$  (note that  $\hat{\mathbb{P}}^{(s,x)}(\{s\} \times \hat{\Omega}) = 1$ ) simultaneously for all  $j = 0, \dots, d$  and all  $i$

$$\begin{aligned} \frac{f^\Phi(s+t; T_i, u_i)}{\Phi_{s+t,T_i}(u_i)} + \langle f^\psi(s+t; T_i, u_i), X_{t-}(\omega^j) \rangle dG_t(s) = \\ - \bar{\kappa}(\psi_{s+t,T_i}(u_i); s+t, X_{t-}(\omega^j)) dG_t(s) \end{aligned}$$

on  $[\tilde{o}(T_i, u_i) - s, T_i - s]$ , where for  $t \geq 0$ ,  $u \in \mathcal{U}$ ,  $x \in E$

$$\bar{\kappa}(u; t, x) := \langle u, \bar{b}(t, x) \rangle + \frac{1}{2} \langle u, \bar{c}(t, x) u \rangle + \int_V (e^{\langle u, \xi \rangle} - 1 - \langle u, h(\xi) \rangle) \bar{K}(d\xi; t, x).$$

---

<sup>10</sup>Note that this holds for the canonical space-time realization of the Markov process  $X$ .

Formulated differently  $dG.(s)$ -a.e. on  $([\tilde{o}(T_i, u_i), T_i] - s)$

$$\frac{f^\Phi(s+t; T_i, u_i)}{\Phi_{s+t, T_i}(u_i)} + \langle f^\psi(s+t; T_i, u_i), X_{t-}(\omega^j) \rangle = -\bar{\kappa}(\psi_{s+t, T_i}(u_i); s+t, X_{t-}(\omega^j)). \quad (1.56)$$

By the proof of Lemma 1.40 we can find sequences  $(T_k, u_k)$  with  $T_k \downarrow s+t$ ,  $u_k \rightarrow u$ , such that (1.56) holds with  $(T_k, u_k)$  for all  $k \geq N$ . The right side then converges to  $-\bar{\kappa}_t(u; s+t, X_{t-}(\omega^j))$  for each  $j \in 0, \dots, d$ . Hence also the left side converges. By Lemma 1.43 there is  $\delta(s) > 0$  and a set  $\Lambda(s)$  with  $\hat{\mathbb{P}}^{(s, \bar{x})}(\Lambda(s)) > 0$ , such that  $H_t$  as defined in equation (1.50) is invertible on  $\Lambda(s)$  for  $t \in [0, \delta(s)]$ . Hence the limits

$$\begin{aligned} \tilde{f}^\Phi(s+t, u) &= \lim_{k \rightarrow \infty} f^\Phi(s+t, T_k, u_k), \\ \tilde{f}^\psi(s+t, u) &= \lim_{k \rightarrow \infty} f^\psi(s+t, T_k, u_k), \end{aligned}$$

exist. So  $\tilde{f}^\Phi(t, u)$  and  $\tilde{f}^\psi(t, u)$  are defined for  $dG.(s)$ -a.e.  $t \in [s, \delta(s)]$ . Since  $s$  was arbitrary, this limits exist  $dG$ -a.e. on  $\mathbb{R}_{\geq 0}$  and (1.56) yields

$$\tilde{f}^\Phi(s+t, u) + \langle \tilde{f}^\psi(s+t, u), X_{t-}(\omega^j) \rangle = -\bar{\kappa}(u; s+t, X_{t-}(\omega^j)) \quad (1.57)$$

The left-hand side is affine in  $X_{t-}(\omega^j)$ , hence this is also true for right hand side. By the uniqueness of the Levy-Khintchine formulas (Lemma II.2.44 in Jacod and Shiryaev [29]) it then follows that  $\bar{b}(t, x)$ ,  $\bar{c}(t, x)$ ,  $\bar{K}(d\xi; t, x)$  are also affine in  $x$   $dG$ -a.e. and that the maps  $a, b, \alpha, \beta, m, M$  as stated in the theorem are  $dG$ -a.e. given by

$$\begin{aligned} \bar{b}(t, x) &= b(t) + \beta(t, x) \\ \bar{c}(t, x) &= a(t) + \alpha(t, x) \\ \bar{K}(d\xi; t, x) &= m(d\xi; t) + M(d\xi; t, x). \end{aligned}$$

This proves the parts regarding the semimartingale characteristics.

Finally we prove the generalized Ricatti integral equations. Note that  $\bar{\kappa}(u, t, x) = F(t, u) + \langle R(t, u), x \rangle$ . Fix  $(T, u)$  and  $o(T, u) < s \leq T$ . By considering the martingales  $\hat{M}_t^{T-s, u, j}$  from equation (1.42) and applying again integration by parts, the Itô formula with the semimartingale characteristics of Lemma 1.42 one obtains

$$\frac{d\Phi_{s+t, T}(u)}{\Phi_{s+t, T}(u)} + \langle d\psi_{s+t, T}(u), X_{t-}(\omega^j) \rangle = -\bar{\kappa}(\psi_{s+t, T}(u); s+t, X_{t-}(\omega^j)) dG(s+t)$$

for  $t \in [\tilde{o}(T, u) - s, T - s]$  and  $j = 0, \dots, d$ . By Lemma 1.43 there is a  $\delta(s) > 0$  and sets  $\Lambda(s)$  with  $\hat{\mathbb{P}}^{(s, \bar{x})}(\Lambda(s)) > 0$  such that  $H_t$  is invertible on  $\Lambda(s)$  for  $t \in [0, \delta(s)]$ . Hence for all  $0 \leq r \leq \delta(s) \wedge (T - s)$ , the associativity of integrals and the fact that the integrands are affine in  $X_{t-}(\omega^j)$  yields  $d\Phi_{\cdot, T}(u) \ll dG(\cdot)$  with density  $t \mapsto -\Phi_{t, T}(u)F(t, \psi_{t, T}(u))$  and  $d\psi_{\cdot, T}(u) \ll dG(\cdot)$  with density  $t \mapsto -R(t, \psi_{t, T}(u))$  on  $[s, (s + \delta(s)) \wedge T]$ . Since  $s$  was arbitrary, this holds on  $[v, T]$  with  $v > o(T, u)$  and the generalized Ricatti integral equations in (1.53) follow.  $\square$

*Remark:* The semimartingale characteristics depend only on  $(\Theta, X)$  with  $\Theta$  depending on  $s$  and  $X$  depending on  $\omega$ . For fixed  $s$   $\tilde{\mathbb{P}}^{(s,x)}(\Theta_0 = s) = 1$  for all  $x \in E$ . Hence (1.55) with  $\Theta_t$  replaced by  $s+t$  gives a version of the characteristics of  $X$  under  $\tilde{\mathbb{P}}^{(s,x)}$  for  $s$  fixed and each  $x \in E$ . These characteristics only depend on  $\Omega$ , so they are also characteristics of  $X$  on  $(\Omega, \mathcal{A})$  with the original filtration  $\mathcal{F}_t$  under each probability measure  $\mathbb{P}^{(s,x)}$ ,  $x \in E$ .

*Remark:* This theorem relates to the results for time-homogeneous affine processes in Cuchiero and Teichmann [10] as follows. By using a full and complete class of functions it is shown in Cuchiero and Teichmann [10] that the canonical realization of a time-homogeneous continuously affine transition function is always a semimartingale and that one can always choose  $G(s) = s$ . From there one could directly use the above proof. The first part shows that the differentiated semimartingale characteristics only depend on  $X_{s-}$  and are affine. The second part then gives the functions  $F$  and  $R$ , which in this case also do not depend on time, so that we get generalized Riccati integral equations with  $G(s) = s$ . In this case the integrands are continuous, so we get differentiability and therefore regularity.

The following example shows that in the time-inhomogeneous case the local characteristics, although affine in  $X$  may be very irregular with respect to the time parameter. This similarly holds for the functions  $F$  and  $R$  in equation (1.54).

*Example 7:* Consider the Smith-Volterra-Cantor set (or fat cantor set) on  $[0, 1]$  denoted by  $A$ . This set is generated iteratively as follows. In the first step the interval  $(\frac{3}{8}, \frac{5}{8})$  is removed from  $[0, 1]$ . The iteration step removes from every remaining interval a centered open subinterval of one quarter of the length of this interval (see also Definition 2.1 in DiMartino and Urbina [13], where this is referred to as the SVC(4) set). The Smith-Volterra-Cantor set has Lebesgue measure  $\frac{1}{2}$ . We define an affine process as in Example 1 using the deterministic function  $f(t) = \int_0^t \mathbb{I}_A(v) dv$ . This gives a continuously affine process with  $\psi_{s,t}(u) = u$  and

$$\phi_{s,t}(u) = u \int_s^t \mathbb{I}_A(v) dv = u\lambda(A \cap [s, t]),$$

where  $\lambda$  denotes the Lebesgue measure. By the Lebesgue differentiation theorem  $\phi$  is differentiable almost everywhere with derivative  $\mathbb{I}_A(t)$ . Differentiable semimartingale characteristics  $(B, C, \nu)$  are given by  $B_t(s, \omega) = \int_0^t \mathbb{I}_A(s+v) dv$ ,  $C = 0$ ,  $\nu = 0$ . The function  $\mathbb{I}_A(t)$  does not have a left-handed limit for all left border points of a removed interval and no right-handed limit for all right border points of a removed interval. To see this consider a left border point (the right border point case follows analogously)  $t$  (e.g.  $t = \frac{3}{8}$ ). Then for every  $\epsilon > 0$  the sets  $A \cap [t - \epsilon, t)$  and  $A^c \cap [t - \epsilon, t)$  have positive measure. Hence we can always find sequences of times  $s_n \uparrow t$  and  $\tilde{s}_n \uparrow t$  with  $\mathbb{I}_A(s_n) = 0$  and  $\mathbb{I}_A(\tilde{s}_n) = 1$ .

Theorem 1.45 shows that although the left derivatives of  $\Phi$  and  $\psi$  might not exist everywhere (i.e. they are not regular), we can think of derivatives of  $\Phi$  and  $\psi$  existing

with respect to  $dG$  and  $dG$ -a.e.. In particular if  $G$  can be chosen as  $G(s) = s$  this then corresponds to actual derivatives, which exist almost everywhere. This motivates the following definition.

**Definition 1.46:** A continuously affine transition function  $P_{s,t}$  is called absolutely continuously affine if the functions  $s \mapsto \Phi_{s,T}(u)$  and  $s \mapsto \psi_{s,T}(u)$  are absolutely continuous on  $[v, T]$  for each  $o(T, u) < v \leq T$ .

**Corollary 1.47:** Assume that the assumptions of Theorem 1.45 hold. In this case  $((\Theta, X), \tilde{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(s,x)}\})$  is an Itô process (see Çinlar et al. [28]) if and only if the affine transition function is absolutely continuous.

*Proof.* If  $(\Theta, X)$  is a Itô process, we know from the start that we can find semimartingale characteristics of  $X$  generated by  $(b, c, K, ds)$ . Then the derivation goes through as before. On the other hand if the functions  $\Phi$  and  $\psi$  are absolutely continuous, it is easy to see that we could have used  $ds$  as candidate in the proof of Lemma 1.42. Note that  $\Theta_t(s) = \Theta_0(s) + t$  and that  $(B^\Theta, 0, 0)$  with  $B_t^\Theta = \int_0^t ds$  are differentiated characteristics of  $\Theta$ . Hence  $X$  is an Itô process.  $\square$

The following two lemmas show that we can always reduce to the absolutely continuous case by considering the deterministic time-change  $\eta_t$  of (1.49), i.e.

$$\eta_t(r) = \inf\{v \geq 0 : G(v+r) - G(r) \geq t\}.$$

**Corollary 1.48:** Assume that the assumptions of Theorem 1.45 hold. The time-changed Markov process  $((\hat{\Theta}, \hat{X}), \hat{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(s,x)}\})$  with

$$(\hat{\Theta}, \hat{X})_t(r, \omega) = (\Theta, X)_{\eta_t(r)}(r, \omega), \quad \hat{\mathcal{F}}_t = \tilde{\mathcal{F}}_{\eta_t}$$

is a Markov semimartingale with differentiable characteristics, i.e. it is an Itô process and the differentiable characteristics are affine in  $\hat{X}$ .

*Proof.* By construction  $ds \ll dG(\cdot)$ . Denote the density by  $s \mapsto f^\Theta(s)$ . Then the characteristics of  $\Theta$  can also be written as  $B_t^\Theta(r) = \int_0^t f^\Theta(r+s) dG_s(r)$ . By Proposition 7.13 in Çinlar et al. [28] the time-changed process  $((\hat{\Theta}, \hat{X}), \hat{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(s,x)}\})$  is a Markov semimartingale and a version of the semimartingale characteristics is given by (note that  $G_{\eta_s(r)} - G_0(r) = s$ )

$$\begin{aligned} \hat{B}_t &= \int_0^t b(\hat{\Theta}_v) + B(\hat{\Theta}_v, \hat{X}_{v-}) dv, \\ \hat{C}_t &= \int_0^t a(\hat{\Theta}_v) + A(\hat{\Theta}_v, \hat{X}_{v-}) dv, \\ \hat{\nu}(dt, d\xi) &= \left( m(d\xi; \hat{\Theta}_t) + M(d\xi; \hat{\Theta}_t, \hat{X}_{t-}) \right) dt. \end{aligned}$$

$\square$

**Corollary 1.49:** *Assume that the assumptions of Theorem 1.45 hold. Define the time-inhomogeneous affine transition operator  $\{\hat{P}_{s,T}\}$  by  $\hat{P}_{s,T} = P_{G^{-1}(s),G^{-1}(T)}$ , i.e.*

$$\begin{aligned}\hat{\Phi}_{s,T}(u) &= \Phi_{G^{-1}(s),G^{-1}(T)}(u), \\ \hat{\psi}_{s,T}(u) &= \psi_{G^{-1}(s),G^{-1}(T)}(u).\end{aligned}$$

Then  $\hat{P}_{s,T}$  is an absolutely continuously affine transition function.

*Proof.* By Theorem 1.45 the Ricatti equations (1.53) hold for  $\Phi$  and  $\psi$ . Together with the definition of  $\hat{\Phi}$  and  $\hat{\psi}$  this yields

$$\begin{aligned}\hat{\Phi}_{s,T}(u) &= 1 + \int_{G^{-1}(s)}^{G^{-1}(T)} \Phi_{v,G^{-1}(T)}(u) F(v, \psi_{v,G^{-1}(T)}(u)) dG(v), \\ &= 1 + \int_s^T \Phi_{G^{-1}(w),G^{-1}(T)}(u) F(G^{-1}(w), \psi_{G^{-1}(w),G^{-1}(T)}(u)) dw \\ &= 1 + \int_s^T \hat{\Phi}_{w,T}(u) F(w, \hat{\psi}_{w,T}(u)) dw \\ \hat{\psi}_{s,T}(u) &= u + \int_s^t R(w, \hat{\psi}_{w,T}(u)) dw,\end{aligned}\tag{1.58}$$

□

Note that the time-changed process  $(\hat{X}, \hat{\mathcal{F}}, \{\tilde{\mathbb{P}}^{(r,x)}\})$  is no longer a time-inhomogeneous Markov process in the sense of Definition 1.6. However, under each measure  $\tilde{\mathbb{P}}^{(r,x)}$   $\hat{X}$  is a time-inhomogeneous Markov process on  $(\tilde{\Omega}, \tilde{\mathcal{A}}, \hat{\mathcal{F}}, \tilde{\mathbb{P}}^{(r,x)})$  in the sense of Definition 1.3 with a regular affine transition function given by  $\{\hat{P}_{r+\eta_s(r),r+\eta_t(r)}\}_{0 \leq s \leq t}$ .

## 2 Affine LIBOR models driven by real-valued affine processes

Market models, the most famous example being the LIBOR market model, are very popular in the area of interest rate modeling. If these models generate nonnegative interest rates they usually do not give semi-analytic formulas for both basic interest rate derivatives, caps and swaptions. One exception is the class of affine LIBOR models proposed by Keller-Ressel et al. [36]. Using nonnegative affine processes as driving processes affine LIBOR models guarantee nonnegative forward interest rates and lead to semi-analytical formulas for caps and swaptions, so that calibration to interest rate market data is possible.

In this chapter we modify the setup of Keller-Ressel et al. [36] to allow for not necessarily nonnegative affine processes. This modification still leads to semi-analytical formulas for caps and swaptions and guarantees nonnegative forward interest rates, but allows for a wider class of driving affine processes and hence is more flexible in producing interest rate skews and smiles. Fonseca et al. [23] also propose a modification of affine LIBOR models. Their driving processes are affine processes with values in the space of positive semidefinite matrices. The approach in this paper has the advantage that a flexible class of implied volatility surfaces can be produced with a much smaller number of parameters.

The structure of this chapter is as follows. In section 2.1 affine processes and their properties are reviewed. Section 2.2 introduces the necessary notation and market setup and reviews affine LIBOR models. It concludes with some comments on practical implementation. Section 2.3 is the main section of this chapter. The first part presents the modified affine LIBOR model and semi-analytical pricing formulas for caps and swaptions are derived. The second part then gives some examples of usable affine processes with numerical calculations.

### 2.1 Affine processes

Let  $X = (X_t)_{0 \leq t \leq T}$  be a homogeneous Markov process with values in  $D = \mathbb{R}_{\geq 0}^m \times \mathbb{R}^n$  realized on a measurable space  $(\Omega, \mathcal{A})$  with filtration  $(\mathcal{F}_t)_{0 \leq t \leq T}$ , with regards to which  $X$  is adapted. Denote by  $\mathbb{P}^x[\cdot]$  and  $\mathbb{E}^x[\cdot]$  the corresponding probability and expectation when  $X_0 = x$ .  $X$  is said to be an affine process, if its characteristic function has the form

$$\mathbb{E}^x [e^{u \cdot X_t}] = \exp\left(\phi_t(u) + \psi_t(u) \cdot x\right), \quad u \in \mathbf{i}\mathbb{R}^d, x \in D, \quad (2.1)$$

where  $\phi : [0, T] \times \mathbf{i}\mathbb{R}^d \rightarrow \mathbb{C}$  and  $\psi : [0, T] \times \mathbf{i}\mathbb{R}^d \rightarrow \mathbb{C}^d$  with  $\mathbf{i}\mathbb{R}^d = \{u \in \mathbb{C}^d : \operatorname{Re}(u) = 0\}$  and  $\cdot$  denoting the scalar product in  $\mathbb{R}^d$ . By homogeneity and the Markov property the conditional characteristic function satisfies

$$\mathbb{E}^x [e^{u \cdot X_t} | \mathcal{F}_s] = \exp\left(\phi_{t-s}(u) + \psi_{t-s}(u) \cdot X_s\right).$$

Accordingly affine processes can also be defined for inhomogeneous Markov processes (see Filipovic [21]), in which case the above equality reads

$$\mathbb{E}^x [e^{u \cdot X_t} | \mathcal{F}_s] = \exp\left(\phi_{s,t}(u) + \psi_{s,t}(u) \cdot X_s\right), \quad u \in \mathbf{i}\mathbb{R}^d, x \in D,$$

with  $\phi_{s,t} : \mathbf{i}\mathbb{R}^d \rightarrow \mathbb{C}$  and  $\psi_{s,t} : \mathbf{i}\mathbb{R}^d \rightarrow \mathbb{C}^d$  for  $0 \leq s \leq t$ .

$X$  is called an analytic affine process (see Keller-Ressel [33]), if  $X$  is stochastically continuous and the interior of the set<sup>1</sup>

$$\mathcal{V} := \left\{ u \in \mathbb{C}^d : \sup_{0 \leq s \leq T} \mathbb{E}^x [e^{\operatorname{Re}(u) \cdot X_s}] < \infty \quad \forall x \in D \right\}, \quad (2.2)$$

contains  $0^2$ . In this case the functions  $\phi$  and  $\psi$  have continuous extensions to  $\mathcal{V}$ , which are analytic in the interior, such that (2.1) holds for all  $u \in \mathcal{V}$ .

The class of affine processes includes Brownian motion and more generally all Lévy processes. Since Lévy processes have stationary independent increments, in this case  $\psi_t(u) = u$  and  $\phi_t(u) = t\kappa(u)$ , where  $\kappa$  is the cumulant generating function of the Lévy process. Ornstein-Uhlenbeck processes are further important examples of affine processes. They are discussed in section 2.3.2.

The standard reference for affine processes is Duffie et al. [16]. There they give a characterization of affine processes, where  $\phi$  and  $\psi$  are specified as solutions of a system of differential equations<sup>3</sup>. Of all the rich theory of affine processes the methods in this chapter only use the specific form (2.1) of their moment generating function and the following property.

**Lemma 2.1:** *Let  $X$  be a one-dimensional analytic affine process and  $\operatorname{Re}(u) < \operatorname{Re}(w)$ ,  $u, w \in \mathcal{V}$ . Then  $\operatorname{Re}(\psi_t(u)) < \operatorname{Re}(\psi_t(w))$ , i.e.  $\psi_t|_{\mathcal{V} \cap \mathbb{R}}$  is strictly increasing.*

*Proof.* The case  $D = \mathbb{R}_+$  is already contained in Keller-Ressel et al. [36]. In case  $D = \mathbb{R}$  the lemma follows from the fact that by Proposition 3.3 in Keller-Ressel et al. [35]  $\psi_t(u) = e^{\beta t} u$  for some constant  $\beta$ .  $\square$

<sup>1</sup> $\mathcal{V}$  can be described as the (convex) set, where the extended moment generating function of  $X_t$  is defined for all times  $t \leq T$  and all starting values  $x \in E$ . By Lemma 4.2 in Keller-Ressel and Mayerhofer [34] the set  $\mathcal{V}$  is in fact equal to the seemingly smaller set  $\left\{ u \in \mathbb{C}^d : \exists x \in \operatorname{int}(D) : \mathbb{E}^x [e^{\operatorname{Re}(u) \cdot X_T}] < \infty \right\}$ .

<sup>2</sup>This also implies that  $X$  is conservative, i.e.  $\mathbb{P}^x(X_t \in D) = 1 \quad \forall x \in D$  and  $0 \leq t \leq T$ .

<sup>3</sup>The fact that this characterization holds for all stochastically continuous affine processes was first shown in Keller-Ressel et al. [35] and later for affine processes with more general state spaces in Keller-Ressel et al. [37] and Cuchiero and Teichmann [10].

*Remark:* If  $D = \mathbb{R}_+$ , it is known that both,  $\psi$  and  $\phi$ , are monotonically increasing (Keller-Ressel et al. [36]). With  $D = \mathbb{R}$  this stays true for  $\psi$ , but not  $\phi$ , as the deterministic affine process  $X_t = x_0 - t$  shows.

## 2.2 Interest rate market models

### Classical market models

Consider a tenor structure  $0 < T_1 < \dots < T_N < T_{N+1} =: T$  and a market consisting of zero coupon bonds with maturities  $T_1, \dots, T_{N+1}$ . Their price processes  $(P(t, T_k))_{0 \leq t \leq T_k}$  are assumed to be nonnegative semimartingales on a filtered probability space  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$  (here  $\mathcal{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ ), which satisfy  $P(T_k, T_k) = 1$  almost surely. If there exists an equivalent probability measure  $\mathbb{Q}^T$  such that the normalized bond price processes  $P(\cdot, T_k)/P(\cdot, T)$  are martingales<sup>4</sup>, the market is arbitrage-free. In this case we can define equivalent martingale measures  $\mathbb{Q}^{T_k}$  for the numeraires  $P(t, T_k)$  instead of  $P(t, T)$  by

$$\frac{d\mathbb{Q}^{T_k}}{d\mathbb{Q}^T} = \frac{1}{P(T_k, T)} \frac{P(0, T)}{P(0, T_k)}. \quad (2.3)$$

In particular under the measure  $\mathbb{Q}^{T_k}$  the forward bond price process  $\frac{P(\cdot, T_{k-1})}{P(\cdot, T_k)}$  and the forward interest rate process  $F^k(\cdot)$ ,

$$F^k(t) = \frac{1}{\Delta_k} \left( \frac{P(t, T_{k-1})}{P(t, T_k)} - 1 \right), \quad \Delta_k = T_k - T_{k-1}, \quad (2.4)$$

are martingales. This is the basic market setup used throughout the rest of the chapter.

In the classical LIBOR market models forward interest rate processes  $F^k$  are modeled as continuous exponential martingales under their respective martingale measure  $\mathbb{Q}^{T_k}$ . Hence forward interest rates are positive. Using driftless geometric Brownian motions as driving processes caplet prices are given by the Black formula (Black [2]) while swaption prices cannot be calculated analytically. Alternatively one can start with modeling the forward bond price processes  $P(\cdot, T_{k-1})/P(\cdot, T_k)$  instead of forward interest rate processes. Using again exponential martingales like a driftless Brownian motion it is then possible to analytically calculate caplet and swaption prices (see Eberlein and Özkan [17]). The drawback of this approach is that forward interest rates will be negative with positive probability.

---

<sup>4</sup>One can extend bond price processes to  $[0, T]$  by setting  $P(t, T_k) := \frac{P(t, T)}{P(T_k, T)}$  for  $t > T_k$ , so that  $P(\cdot, T_k)/P(\cdot, T)$  is a martingale on  $[0, T]$  if and only if it is a martingale on  $[0, T_k]$ . Economically this can be interpreted as immediately investing the payoff of a zero coupon bond into the longest-running zero coupon bond.

Keller-Ressel et al. [36] proposed the affine LIBOR models, where forward interest rates are nonnegative while swaption and caplet prices can still be calculated semi-analytically, i.e. up to a numerical integration. The above approaches model the individual forward interest rate processes (resp. forward bond price process) with respect to the individual measure  $\mathbb{Q}^{T_k}$  under which they are a martingale. Contrary Keller-Ressel et al. [36] model the price processes  $P(\cdot, T_k)/P(\cdot, T)$ , which are all martingales under the same probability measure  $\mathbb{Q}^T$ .

*Remark:* Note that all models mentioned in this chapter do not fully specify the whole term structure, but only part of it. In order to price derivatives not contained within the specified tenor structure it is necessary to specify some kind of interpolation scheme. Arbitrary interpolations may lead to arbitrage, however one can always choose an interpolation method, such that the model stays arbitrage-free (Werpachowski [47]).

## The affine LIBOR models

This section presents a summary of the affine LIBOR model introduced in Keller-Ressel et al. [36]. On the filtered probability space  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{Q}^T)$  consider a nonnegative analytic affine process  $X$  with a fixed starting value  $x_0 \in \mathbb{R}_{\geq 0}^d$ . For the tenor structure  $0 < T_1 < \dots < T_N < T_{N+1} =: T$  define for  $k = 1, \dots, N$  and  $0 \leq t \leq T_k$

$$\frac{P(t, T_k)}{P(t, T)} := \mathbb{E}^{\mathbb{Q}^T} [e^{u_k \cdot X_T} | \mathcal{F}_t] = e^{\phi_{T-t}(u_k) + \psi_{T-t}(u_k) \cdot X_t}, \quad u_k \geq 0, u_k \in \mathcal{V}, \quad (2.5)$$

where  $\mathbb{E}^{\mathbb{Q}}[\cdot]$  denotes the expectation with respect to a probability measure<sup>5</sup>  $\mathbb{Q}$ . These price processes are martingales and the resulting model is arbitrage-free.

Writing

$$\frac{P(t, T_{k-1})}{P(t, T_k)} = \frac{P(t, T_{k-1})}{P(t, T)} \bigg/ \frac{P(t, T_k)}{P(t, T)} \quad (2.6)$$

in (2.4) shows that forward interest rates being nonnegative is equivalent to normalized bond prices of (2.5) satisfying

$$\frac{P(t, T_1)}{P(t, T)} \geq \dots \geq \frac{P(t, T_N)}{P(t, T)} \geq 1. \quad (2.7)$$

Since for  $x \geq 0$ ,  $e^{u^T x}$  is monotonically increasing in every component of  $u$  the monotonicity for normalized bond prices in (2.7) is satisfied as long as  $u_1 \geq \dots u_N \geq 0$ .

The parameters  $u_k$  in (2.5) should be determined, so that the starting values of normalized bond prices  $P(0, T_k)/P(0, T) = \exp(\phi_T(u_k) + \psi_T(u_k) \cdot x_0)$  fit the initial term structure inferred from actual market data. For most affine processes every term

---

<sup>5</sup>Since  $x_0$  is fixed, contrary to to section 2.1 any dependence of probability measures on the starting value of the Markov process  $X$  will be suppressed from now on.

structure can be fitted and for currently nonnegative forward interest rates this can be done using an decreasing sequence  $u_1 \geq \dots \geq u_N \geq 0$  (see Keller-Ressel et al. [36]).

*Remark:* Since  $X$  is nonnegative, the  $k$ -th normalized bond price is not only greater equal to one, but is bounded from below by the time-dependent constant given by  $\exp(\phi_{T-t}(u_k))$ , which is strictly greater than one. Accordingly in the affine LIBOR models forward interest rates are bounded from below by a strictly positive time-dependent constant.

Affine LIBOR models lead to nonnegative forward interest rates. Additionally this specification is appealing because the density processes for changes of measures are again exponentially affine in  $X_t$ , i.e. inserting (2.5) into (2.3) gives

$$\frac{d\mathbb{Q}^{T_k}}{d\mathbb{Q}^T} = \frac{P(0, T)}{P(0, T_k)} e^{\phi_{T-T_k}(u_k) + \psi_{T-T_k}(u_k) \cdot X_{T_k}}.$$

Moreover normalized bond prices and because of (2.6) also forward bond prices are of exponential affine form. It follows that the moment generating function of the logarithm of normalized bond prices under  $\mathbb{Q}^{T_k}$  is also of exponential affine form and that calculation of caplet prices is possible via a one-dimensional Fourier inversion. If the dimension of the driving process is one, swaption prices can also be calculated via one-dimensional Fourier inversion (see Keller-Ressel et al. [36]). Hence this approach satisfies both, nonnegative interest rates and analytical tractability of standard interest rate market instruments. If the dimension is larger than one, the exact price of swaptions can only be calculated via higher-dimensional integration, the dimension of which is the length of the underlying swap. Alternatively Grbac et al. [25] provide approximate formulas for swaptions.

### Practical application of the affine LIBOR model

Although this framework is elegant from a theoretical point of view, a practical implementation faces several difficulties which shall be discussed here.

First, calibration of interest rates and implied volatilities cannot be separated. The initial term structure can be fitted using the  $u_k$ , but the parameters  $u_k$  also have a strong impact on implied volatilities. This can be seen by looking at the forward bond price

$$\frac{1}{P(T_{k-1}, T_k)} = \exp\left(\phi_{T-t}(u_{k-1}) - \phi_{T-t}(u_k) + (\psi_{T-t}(u_{k-1}) - \psi_{T-t}(u_k)) \cdot X_{T_{k-1}}\right), \quad (2.8)$$

which is the random variable responsible for the payoff of a caplet. The driving process  $X$  influences the distribution of this random variable through two different channels. First via the parameters of the driving process itself and second via the parameters  $u_k$  (depending on  $X$  and the initial interest rate term structure). Hence for changes

in the yield curve different parameters are required to reproduce the same implied volatility surface. If  $X$  is a Lévy process, then as mentioned in section 2.1  $\psi_t(u) = u$  and it follows that the distribution of (2.8) depends on the difference  $u_{k+1} - u_k$ , which in turn is related to the steepness of the initial yield curve<sup>6</sup>. Hence caplet implied volatilities are especially sensitive with regards the steepness of the initial yield curve.

Second, interest rates and volatilities of this model depend on the final horizon  $T$ . Changing the horizon  $T$  while using the same affine process  $X$  will lead to different results and there is no general way of rescaling the parameters of  $X$  to negate such an effect. This is rather counterintuitive, since extending the horizon of a model should not change the results for quantities already included with the shorter horizon.

Third, the types of possible volatility surfaces is rather constrained in the fully analytically tractable one-dimensional case. For example, we were only able to generate volatility skews<sup>7</sup>. This might be resolved by using higher-dimensional nonnegative processes. However, in multidimensional affine LIBOR models swaptions can no longer be calculated efficiently by Fourier methods. On the other hand allowing arbitrary affine processes destroys the nonnegativity of forward interest rates, a central property of affine LIBOR models. We propose a modification, that preserves the nonnegativity of forward interest rates without the restriction to nonnegative affine processes.

## 2.3 The modified affine LIBOR model

On the filtered probability space  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{Q}^T)$  consider an analytic one-dimensional affine process  $X$  with a fixed starting value  $x_0$ , i.e. the set  $\mathcal{V}$  defined in (3.19) contains 0 in the interior. For  $u \in \mathcal{V}$  with  $-u \in \mathcal{V}$  consider the martingales  $M^u$ ,

$$M_t^u := \mathbb{E}^{\mathbb{Q}^T} [\cosh(uX_T) | \mathcal{F}_t] = \frac{1}{2} (e^{\phi_{T-t}(u) + \psi_{T-t}(u)X_t} + e^{\phi_{T-t}(-u) + \psi_{T-t}(-u)X_t}). \quad (2.9)$$

By the symmetry of the cosinus hyperbolicus  $M^u = M^{-u}$ , hence one may restrict  $u$  to be nonnegative. For the given tenor structure  $0 < T_1 < \dots < T_N \leq T_{N+1} = T$  and the market setup of section 2.2 define the normalized bond prices for  $k = 1, \dots, N$  and  $t \leq T_k$  as

$$\frac{P(t, T_k)}{P(t, T)} := M_t^{u_k}, \quad u_k \in \{v \in \mathcal{V} : v \geq 0, -v \in \mathcal{V}\}.$$

With  $M_t^{u_k}$  being a  $\mathbb{Q}^T$ -martingale the model is arbitrage-free. For every  $x \in \mathbb{R}$  the function  $u \mapsto \cosh(ux)$  is increasing in  $u \in \mathbb{R}_{\geq 0}$  and satisfies  $\cosh(ux) \geq 1$  so that if

$$u_1 \geq u_2 \geq \dots \geq u_N \geq 0,$$

<sup>6</sup>This is similar for most affine processes, but is best visible for Lévy processes.

<sup>7</sup>The smile example of Keller-Ressel et al. [36], figure 9.2, using an Ornstein-Uhlenbeck process seems to be numerically incorrect for strikes smaller than 0.4. With the mentioned initial yield curve the underlying interest rate is always larger than the strike, which corresponds to a zero implied volatility, destroying the displayed smile.

equation (2.7) holds and forward interest rates

$$F^k(t) = \frac{1}{\Delta_k} \left( \frac{M_t^{u_{k-1}}}{M_t^{u_k}} - 1 \right), \quad 0 \leq t \leq T_{k-1}.$$

are nonnegative for all  $t$ . To fit initial market data one has to choose the sequence  $(u_k)$  so that  $M_0^{u_k} = P(0, T_k)/P(0, T)$ . The following lemma gives the condition for the affine process  $X$  under which a given initial term structure can be reproduced and shows that the  $u_k$  are uniquely determined.

**Lemma 2.2:** *If*

$$P(0, T_1)/P(0, T) < \sup_{u \in \mathcal{V}: -u \in \mathcal{V}} \mathbb{E}^{\mathbb{Q}^T} [\cosh(uX_T) | \mathcal{F}_0],$$

*then the model can fit any term structure of nonnegative forward interest rates. Additionally there exists a unique decreasing sequence  $u_1 \geq \dots \geq u_N$ , such that*

$$P(0, T_k)/P(0, T) = \mathbb{E}^{\mathbb{Q}^T} [\cosh(u_k X_T) | \mathcal{F}_0] = M_0^{u_k}.$$

*If forward interest rates are strictly positive, the sequence is strictly decreasing.*

*Proof.*  $m(u) = \mathbb{E}^{\mathbb{Q}^T} [\cosh(uX_T) | \mathcal{F}_0]$  is a continuous function which is increasing in case  $u \geq 0$ . By the assumption of the theorem there exists  $\bar{u} > 0$  with  $m(\bar{u}) > P(0, T_1)/P(0, T)$ . Furthermore  $m(0) = 1$ , which proves the lemma.  $\square$

*Remark:* Generalizing this approach to a  $d$ -dimensional driving process is possible by setting

$$M_t^u = \mathbb{E} \left[ \prod_{l=1}^d \cosh \left( u^{(l)} X_T^{(l)} \right) \middle| \mathcal{F}_t \right], \quad u = (u^{(1)}, \dots, u^{(d)}) \geq 0.$$

In this case it is guaranteed that  $M_t^u \geq M_t^w$  for  $u \geq w$ , which guarantees the nonnegativity of forward interest rates. However, the option pricing formulas in the following sections do not generalize.

As in the affine LIBOR model for a monotonically decreasing sequence  $(u_k)$  forward interest rates are not only nonnegative, but bounded below by strictly positive time-dependent constants (the bounds can be calculated numerically). This is not a big issue if these bounds are close to zero, but has to be checked during the calibration process.

In the modified affine LIBOR model the change of measure to the  $T_k$ -forward measure  $\mathbb{Q}^{T_k}$  is given by

$$\frac{d\mathbb{Q}^{T_k}}{d\mathbb{Q}^T} = \frac{P(0, T)}{P(0, T_k)} M_{T_k}^{u_k} = \frac{M_{T_k}^{u_k}}{M_0^{u_k}}. \quad (2.10)$$

Here  $M_t^{u_k}$  is a sum of exponentials of  $X_t$ , while in the affine LIBOR model the corresponding term is a single exponential. This means that contrary to the affine LIBOR model the process  $X$  is not an inhomogeneous affine process under  $\mathbb{Q}^{T_k}$  and it is not possible to calculate the moment generating function of the logarithm of forward bond prices under  $\mathbb{Q}^{T_k}$ . Nevertheless it is possible to get analytical formulas for the prices of caplets and swaptions.

### 2.3.1 Option pricing

The derivation of the pricing formulas for caplets and swaptions is based on a method first applied in Jamshidian [30]. First caplets are dealt with, swaptions follow afterwards<sup>8</sup>. Note that if  $u_k = u_{k-1}$  the corresponding forward interest rate  $F^k$  always stays zero. To exclude such pathological examples assume that the sequence  $(u_k)$  is strictly decreasing. In this section random variables are often viewed as functions of the value of the driving process  $X$ . Specifically consider the functions  $M_t^u : \mathbb{R} \rightarrow \mathbb{R}$ ,

$$x \mapsto M_t^u(x) := \frac{1}{2} \left( e^{\phi_{T-t}(u) + \psi_{T-t}(u)x} + e^{\phi_{T-t}(-u) + \psi_{T-t}(-u)x} \right). \quad (2.11)$$

The time  $t$  value of martingale  $M^u$  in (2.9) is then  $M_t^u = M_t^u(X_t)$ . In the rest of the chapter  $M_t^u$  will denote both, the function and the value of the stochastic processes, where the correct interpretation should be clear from context.

The payoff of a caplet for the  $(k+1)$ <sup>th</sup> forward rate  $F^{k+1}(T_k)$  with strike  $K$  is

$$\Delta_{k+1} (F^{k+1}(T_k) - K)_+ = \left( \frac{1}{P(T_k, T_{k+1})} - \tilde{K} \right)_+ = \left( \frac{M_{T_k}^{u_k}}{M_{T_k}^{u_{k+1}}} - \tilde{K} \right)_+,$$

where  $\tilde{K} = 1 + \Delta_{k+1}K$ . Since this payoff has to be paid at time  $T_{k+1}$  the price of the caplet and the corresponding floorlet is

$$\begin{aligned} \text{Cpl}(t, T_k, T_{k+1}, K) &= P(t, T_{k+1}) \mathbb{E}^{\mathbb{Q}^{T_{k+1}}} \left[ \left( \frac{M_{T_k}^{u_k}}{M_{T_k}^{u_{k+1}}} - \tilde{K} \right)_+ \mid \mathcal{F}_t \right], \\ \text{Flt}(t, T_k, T_{k+1}, K) &= P(t, T_{k+1}) \mathbb{E}^{\mathbb{Q}^{T_{k+1}}} \left[ \left( \tilde{K} - \frac{M_{T_k}^{u_k}}{M_{T_k}^{u_{k+1}}} \right)_+ \mid \mathcal{F}_t \right]. \end{aligned}$$

Since price processes are martingales, the put/call parity holds and prices of caplets follow from floorlets and vice versa. Because Fourier analysis is easier for floorlets, where the payoff is bounded, formulas are derived for floorlets.

Since the moment generating function of  $\ln(M_{T_k}^{u_k}/M_{T_k}^{u_{k+1}})$  is unknown, Fourier methods are not directly applicable. However, the function  $x \mapsto M_{T_k}^{u_k}(x)/M_{T_k}^{u_{k+1}}(x)$  has a unique minimum and is monotonically increasing moving away from this minimum. Using this one can get rid of the positive part and use Fourier inversion to calculate the above expectations. The above mentioned monotonicity is very fortunate and follows from a close interplay between the monotonicity of the sequence  $(u_k)$  and the function  $\psi$  with properties of the cosine hyperbolicus. Details are laid out in the proof of the following lemma, which can be found in section 2.4.

---

<sup>8</sup>Actually caplet prices coincide with prices of swaptions with only one underlying period. The difference between those two derivatives is the payoff time.

**Lemma 2.3:** For  $i = 1, \dots, n$  let  $u_0 \geq u_i \geq 0$ , where for at least one  $i$   $u_0 > u_i$ . Let  $c_i > 0$  be positive constants. Define a function  $g : \mathbb{R} \rightarrow \mathbb{R}$  by

$$g(x) := \sum_{i=1}^n c_i \frac{M_t^{u_i}(x)}{M_t^{u_0}(x)}. \quad (2.12)$$

Then  $g$  has a unique maximum at some point  $\xi \in \mathbb{R}$  and is strictly monotonically decreasing to 0 on the left and right side of  $\xi$ .

For floorlet valuation this lemma is not directly applicable as  $u_k > u_{k+1}$ , which is the wrong inequality. However, there is only one summand and the lemma can be applied to the inverse  $M_{T_k}^{u_{k+1}}(x)/M_{T_k}^{u_k}(x)$ . It follows that  $M_{T_k}^{u_k}(x)/M_{T_k}^{u_{k+1}}(x)$  has a unique minimum at some point  $\xi$  and is increasing to infinity to the left and right. Hence it is possible to write

$$\left( \tilde{K} - \frac{M_{T_k}^{u_k}(x)}{M_{T_k}^{u_{k+1}}(x)} \right)_+ = \left( \tilde{K} - \frac{M_{T_k}^{u_k}(x)}{M_{T_k}^{u_{k+1}}(x)} \right) \mathbb{I} \{ \kappa_1 < x < \kappa_2 \}, \quad (2.13)$$

where  $\kappa_1$  and  $\kappa_2$  are two uniquely determined constants satisfying  $\kappa_1 \leq \xi \leq \kappa_2$ . If  $\kappa_1 = \xi = \kappa_2$  the payoff is zero, which corresponds to  $M_{T_k}^{u_k}/M_{T_k}^{u_{k+1}} > \tilde{K}$ . This happens if the forward interest rate is bounded from below by  $\tilde{K}$ , which only happens for very low strikes  $K$ . Inserting (2.13) into the price of a floorlet it follows by a change of measure that

$$\begin{aligned} \text{Flt}(t, T_k, T_{k+1}, K) &= P(t, T_{k+1}) \mathbb{E}^{\mathbb{Q}^{T_{k+1}}} \left[ \left( \tilde{K} - \frac{M_{T_k}^{u_k}}{M_{T_k}^{u_{k+1}}} \right) \mathbb{I} \{ \kappa_1 < X_{T_k} < \kappa_2 \} \middle| \mathcal{F}_t \right] \\ &= P(t, T) \mathbb{E}^{\mathbb{Q}^T} \left[ \left( \tilde{K} M_{T_k}^{u_{k+1}} - M_{T_k}^{u_k} \right) \mathbb{I} \{ \kappa_1 < X_{T_k} < \kappa_2 \} \middle| \mathcal{F}_t \right]. \end{aligned} \quad (2.14)$$

$\tilde{K} M_{T_k}^{u_{k+1}} - M_{T_k}^{u_k}$  is the sum of exponentials of the random variable  $X_{T_k}$ . The expectation in (2.14) is calculated under the measure  $\mathbb{Q}^T$ , where the conditional moment generating function

$$\mathcal{M}_{X_t|X_s}(z) := \mathbb{E}^{\mathbb{Q}^T} [e^{zX_t} | \mathcal{F}_s] = \mathbb{E}^{\mathbb{Q}^T} [e^{zX_t} | X_s] = \exp(\phi_{t-s}(z) + \psi_{t-s}(z)X_s)$$

is known for  $z \in \mathcal{V}$ . Hence the expectation in (2.14) can be calculated via Fourier inversion. The Fourier inversion formula for terms of the above form is stated in Lemma 2.4, the proof of which is given in section 2.4.

**Lemma 2.4:** Assume that the function  $f : \mathbb{R} \rightarrow \mathbb{R}$  has the representation

$$f(x) = \sum_k C_k e^{v_k x} \mathbb{I} \{ \kappa_1 < x < \kappa_2 \}, \quad \lim_{x \downarrow \kappa_1} f(x) = \lim_{x \uparrow \kappa_2} f(x) = 0,$$

where the summation is over a finite index set and the  $C_k$  and  $v_k$  are real constants. Then for  $R \in \mathcal{V} \cap \mathbb{R}$  the Fourier inversion formula

$$\mathbb{E}[f(X_t) | \mathcal{F}_s] = \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left( \mathcal{M}_{X_t | X_s}(iu + R) \hat{f}(u - iR) \right) du$$

holds, where  $\hat{f}$  is the analytic Fourier transform given by

$$\hat{f}(z) = \frac{1}{iz} \sum_k \frac{C_k v_k}{v_k - iz} (e^{(v_k - iz)\kappa_2} - e^{(v_k - iz)\kappa_1}), \quad z \neq 0, z \neq -iv_k. \quad (2.15)$$

To calculate the price of a floorlet in (2.14) apply Lemma 2.4 to  $f_{k+1}^K(X_{T_k})$  with

$$f_{k+1}^K(x) := \left( \tilde{K} M_{T_k}^{u_{k+1}}(x) - M_{T_k}^{u_k}(x) \right) \mathbb{I}\{\kappa_1 < x < \kappa_2\}. \quad (2.16)$$

Its Fourier transform is

$$\hat{f}_{k+1}^K(z) = \frac{1}{iz} \left( (1 + \Delta_{k+1} K) h_{\kappa_1, \kappa_2}^{T_k}(-iz, u_{k+1}) - h_{\kappa_1, \kappa_2}^{T_k}(-iz, u_k) \right) \quad (2.17)$$

with

$$\begin{aligned} h_{\kappa_1, \kappa_2}^t(z, u) &:= e^{\phi_{T-t}(u)} \frac{\psi_{T-t}(u)}{2(z + \psi_{T-t}(u))} (e^{(z + \psi_{T-t}(u))\kappa_2} - e^{(z + \psi_{T-t}(u))\kappa_1}) \\ &+ e^{\phi_{T-t}(-u)} \frac{\psi_{T-t}(-u)}{2(z + \psi_{T-t}(-u))} (e^{(z + \psi_{T-t}(-u))\kappa_2} - e^{(z + \psi_{T-t}(-u))\kappa_1}). \end{aligned} \quad (2.18)$$

The case of swaptions is similar. Consider a swap which is part of the tenor structure. That is, consider  $1 \leq \alpha < \beta \leq N$  and the according interest rate swap with forward swap rate

$$S_{\alpha, \beta}(t) = \frac{P(t, T_\alpha) - P(t, T_\beta)}{\sum_{k=\alpha+1}^\beta \Delta_k P(t, T_k)}, \quad \Delta_k = T_k - T_{k-1}.$$

The payoff of a put swaption on the above swap with strike  $K$  is then

$$\begin{aligned} \sum_{k=\alpha+1}^\beta P(T_\alpha, T_k) \Delta_k (K - S_{\alpha, \beta}(T_\alpha))_+ &= \left( P(T_\alpha, T_\beta) + K \sum_{k=\alpha+1}^\beta \Delta_k P(T_\alpha, T_k) - 1 \right)_+ \\ &= \left( \frac{M_{T_\alpha}^{u_\beta}}{M_{T_\alpha}^{u_\alpha}} + \sum_{k=\alpha+1}^\beta K \Delta_k \frac{M_{T_\alpha}^{u_k}}{M_{T_\alpha}^{u_\alpha}} - 1 \right)_+. \end{aligned}$$

Since the function  $M_{T_\alpha}^{u_\beta}(x)/M_{T_\alpha}^{u_\alpha}(x) + \sum_{k=\alpha+1}^\beta K \Delta_k M_{T_\alpha}^{u_k}(x)/M_{T_\alpha}^{u_\alpha}(x)$  is of the form of Lemma 2.3, it has a unique maximum  $\xi$  and one can find constants<sup>9</sup>  $\kappa_1 \leq \xi \leq \kappa_2$  such that after a change of measure the value of a put swaption is

<sup>9</sup>As in the floorlet case if  $\kappa_1 = \kappa_2$  then the forward swap rate is always larger than the strike. Note that  $S_{\alpha, \beta}(t)$  can also be written as  $S_{\alpha, \beta}(t) = \sum_{k=\alpha+1}^\beta w_k(t) F^k(t)$  with  $w_k > 0$  (see e.g. Brigo and Mercurio [4]). It follows that if forward interest rates are bounded below by positive constants the same will be true for forward swap rates. This bound is then at most an average of the corresponding forward interest rates bounds and is therefore of the same order of magnitude, which for a meaningful model will be small enough.

$$\text{PutSwaption}(t, T_\alpha, T_\beta, K) = P(t, T) \mathbb{E}^{\mathbb{Q}^T} \left[ f_{\alpha, \beta}^K(X_{T_\alpha}) \middle| \mathcal{F}_t \right],$$

where

$$f_{\alpha, \beta}^K(x) = \left( M_{T_\alpha}^{u_\beta}(x) - M_{T_\alpha}^{u_\alpha}(x) + \sum_{k=\alpha+1}^{\beta} K \Delta_k M_{T_\alpha}^{u_k}(x) \right) \mathbb{I} \{ \kappa_1 < x < \kappa_2 \}. \quad (2.19)$$

Again this of the form in Lemma 2.4 and in this case

$$\hat{f}_{\alpha, \beta}^K(z) = \frac{1}{iz} \left( h_{\kappa_1, \kappa_2}^{T_\alpha}(-iz, u_\beta) - h_{\kappa_1, \kappa_2}^{T_\alpha}(-iz, u_\alpha) + K \sum_{k=\alpha+1}^{\beta} \Delta_k h_{\kappa_1, \kappa_2}^{T_\alpha}(-iz, u_k) \right), \quad (2.20)$$

where  $h_{\kappa_1, \kappa_2}^t(z, u)$  is defined in (2.18). The pricing formulas are summarized in the following theorem.

**Theorem 2.5:** *Let  $R \in \mathcal{V} \cap \mathbb{R}$ . In the modified affine LIBOR model prices of a forward interest rate put and a put swaption are*

$$\text{Flt}(t, T_k, T_{k+1}, K) = \frac{P(t, T)}{\pi} \int_0^\infty \text{Re} \left( \mathcal{M}_{X_{T_k} | X_t}(R + iu) \hat{f}_{k+1}^K(u - iR) \right) du, \quad (2.21)$$

$$\text{PutSwaption}(t, T_\alpha, T_\beta, K) = \frac{P(t, T)}{\pi} \int_0^\infty \text{Re} \left( \mathcal{M}_{X_{T_\alpha} | X_t}(R + iu) \hat{f}_{\alpha, \beta}^K(u - iR) \right) du, \quad (2.22)$$

The Fourier transforms  $\hat{f}_{k+1}^K$  respectively  $\hat{f}_{\alpha, \beta}^K$  are given in (2.17) respectively (2.20) for  $R \notin \{0, u_k, u_{k+1}\}$  respectively  $R \notin \{0, u_\alpha, \dots, u_\beta\}$ .

In order to calculate  $\hat{f}_i^K$  respectively  $\hat{f}_{\alpha, \beta}^K$  one has to find the roots  $\kappa_1, \kappa_2$  of the functions

$$g_k^K(x) := \tilde{K} - \frac{M_{T_k}^{u_k}(x)}{M_{T_k}^{u_{k+1}}(x)}, \quad (2.23)$$

$$g_{\alpha, \beta}^K(x) := \frac{M_{T_\alpha}^{u_\beta}(x)}{M_{T_\alpha}^{u_\alpha}(x)} + \sum_{k=\alpha+1}^{\beta} K \Delta_k \frac{M_{T_\alpha}^{u_k}(x)}{M_{T_\alpha}^{u_\alpha}(x)} - 1. \quad (2.24)$$

By Lemma 2.3 this amounts to finding the roots of a function which has a single optimum and is monotonic when moving away from this optimum. Numerical determination of the roots of such well-behaved one-dimensional functions poses no problem. Having determined those bounds valuation reduces to a one-dimensional integration of a function that is falling at least like  $1/x^2$  (depending on the moment generating function of the affine process), so also numerical integration is feasible. Note that besides caps, floors and swaptions, options like digital options or Asset-or-Nothing options can be calculated in a similar manner.

## 2.3.2 Examples

The first part of this section looks at the benchmark case of a Brownian motion, where everything can also be calculated in closed form. Afterwards Ornstein-Uhlenbeck processes are discussed. The section concludes with examples of possible volatility surfaces.

### Brownian motion

Choose  $X_t = B_t$ , a standard Brownian motion starting in 0. The conditional moment generating function is

$$\mathcal{M}_{B_T|B_t}(u) = \mathbb{E}[e^{uB_T} | \mathcal{F}_t] = \exp\left(uB_t + \frac{u^2}{2}(T-t)\right).$$

Hence this is an affine process with  $\phi_t(u) = \frac{u^2}{2}t$  and  $\psi_t(u) = u$ . Consider the time 0 price of a floorlet as given in (2.14) with  $t = 0$ . Since  $M_t^u(-x) = M_t^u(x)$  one finds that in this case  $\kappa_2 = \kappa$  and  $\kappa_1 = -\kappa$ , where  $\kappa$  is the unique positive root of (2.23) if  $g_{k+1}^K(0) < 0$  and  $\kappa = 0$  otherwise. By (2.14) the floorlet price  $\text{Flt}(0, T_k, T_{k+1}, K)$  is

$$P(0, T)\mathbb{E}^{\mathbb{Q}^T} \left[ \left( \tilde{K} e^{\frac{u_{k+1}^2}{2}(T-T_k)} \cosh(u_{k+1}B_{T_k}) - e^{\frac{u_k^2}{2}(T-T_k)} \cosh(u_k B_{T_k}) \right) \mathbb{I}\{|B_{T_k}| \leq \kappa\} \right].$$

By the symmetry of a Brownian motion starting in 0

$$\begin{aligned} \mathbb{E}[\cosh(zB_t) \mathbb{I}\{|B_t| \leq \kappa\}] &= \mathbb{E}[e^{zB_t} \mathbb{I}\{|B_t| \leq \kappa\}] = \mathbb{E}[e^{-zB_t} \mathbb{I}\{|B_t| \leq \kappa\}] \\ &= e^{\frac{1}{2}tz^2} \left( \Phi\left(\frac{\kappa}{\sqrt{t}} - z\sqrt{t}\right) - \Phi\left(-\frac{\kappa}{\sqrt{t}} - z\sqrt{t}\right) \right), \end{aligned}$$

where  $\Phi$  denotes the cumulative distribution function of a standard normal distributed random variable. Hence the floorlet price  $\text{Flt}(0, T_k, T_{k+1}, K)$  is

$$\begin{aligned} &\tilde{K}P(0, T)e^{u_{k+1}^2 \frac{T}{2}} \left( \Phi\left(\frac{\kappa}{\sqrt{T_k}} - u_{k+1}\sqrt{T_k}\right) - \Phi\left(-\frac{\kappa}{\sqrt{T_k}} - u_{k+1}\sqrt{T_k}\right) \right) \\ &- P(0, T)e^{u_k^2 \frac{T}{2}} \left( \Phi\left(\frac{\kappa}{\sqrt{T_k}} - u_k\sqrt{T_k}\right) - \Phi\left(-\frac{\kappa}{\sqrt{T_k}} - u_k\sqrt{T_k}\right) \right). \end{aligned}$$

Slightly more complicated formulas exist when  $B$  is replaced with a Brownian motion with constant drift and volatility and a starting value different from 0.

Swaptions can be treated the same way. Let  $\kappa$  be the unique positive root of (2.24)

if  $g_{\alpha,\beta}^K(0) > 0$  and  $\kappa = 0$  otherwise. Then

$$\begin{aligned} & \text{PutSwaption}(0, T_\alpha, T_\beta, K) \\ &= P(0, T) e^{u_\beta^2 \frac{T}{2}} \left( \Phi\left(\frac{\kappa}{\sqrt{T_\alpha}} - u_\beta \sqrt{T_\alpha}\right) - \Phi\left(-\frac{\kappa}{\sqrt{T_\alpha}} - u_\beta \sqrt{T_\alpha}\right) \right) \\ &- P(0, T) e^{u_\alpha^2 \frac{T}{2}} \left( \Phi\left(\frac{\kappa}{\sqrt{T_\alpha}} - u_\alpha \sqrt{T_\alpha}\right) - \Phi\left(-\frac{\kappa}{\sqrt{T_\alpha}} - u_\alpha \sqrt{T_\alpha}\right) \right) \\ &+ P(0, T) \sum_{k=\alpha+1}^{\beta} K \Delta_k e^{u_k^2 \frac{T}{2}} \left( \Phi\left(\frac{\kappa}{\sqrt{T_\alpha}} - u_k \sqrt{T_\alpha}\right) - \Phi\left(-\frac{\kappa}{\sqrt{T_\alpha}} - u_k \sqrt{T_\alpha}\right) \right). \end{aligned}$$

### Ornstein-Uhlenbeck (OU) processes

The OU process  $X$  generated by a Lévy process  $L$  is defined as the unique strong solution of (see Sato [44], section 17)

$$dX_t = -\lambda X_t dt + dL_t, \quad X_0 = x_0. \quad (2.25)$$

Then  $Y_t := e^{\lambda t} X_t = x_0 + \int_0^t e^{\lambda s} dL_s$ . Using the key formula of Lemma 3.1 in Eberlein and Raible [18] (here we assume the according integrability condition of the jumps of the Levy process) it follows that

$$\mathbb{E}[e^{uX_t}] = \mathbb{E}\left[\exp\left(e^{-\lambda t} u Y_t\right)\right] = \exp\left(e^{-\lambda t} u x_0 + \int_0^t \kappa(e^{-\lambda s} u) ds\right),$$

where  $\kappa(u) = \ln(\mathbb{E}[e^{uL_1}])$  is the cumulant generating function of the Lévy process  $L$ . Hence this process is affine with

$$\psi_t(u) = e^{-\lambda t} u \quad \text{and} \quad \phi_t(u) = \int_0^t \kappa(e^{-\lambda s} u) ds = \frac{1}{\lambda} \int_{e^{-\lambda t}}^1 \frac{\kappa(vu)}{v} dv. \quad (2.26)$$

By Corollary 2.10 in Duffie et al. [16] every affine process with state space  $\mathbb{R}$  is in fact an OU process. Hence in the context of affine processes defined on the real line OU processes are the right class to consider. For application it should be possible to calculate the integral in (2.26) analytically. Two examples where this is possible are presented below.

*Remark:* If  $L$  is a martingale, the process in (2.25) is mean-reverting to zero, however shifting the mean to  $\theta$  is easily done by using  $Z_t = \theta + X_t$ . Then  $dZ_t = \lambda(\theta - Z_t) dt + dL_t$  and

$$\mathbb{E}[e^{uZ_t}] = \exp\left((\phi_t(u) + \theta u(1 - e^{-\lambda t})) + \psi_t(u) Z_0\right).$$

Hence  $Z$  is again affine with  $\psi_t^\theta(u) = \psi_t(u)$  and  $\phi_t^\theta(u) = \phi_t(u) + \theta u(1 - e^{-\lambda t})$ . Note that this process is then generated by the Lévy process  $\tilde{L}_t = L_t + \theta \lambda t$ , i.e. the original Lévy process plus an additional drift of  $\theta \lambda$ .

The first example is the classical OU process generated by a Brownian motion  $\sigma B$ , where  $\kappa(u) = \frac{1}{2}\sigma^2 u^2$ . This process is described by

$$dX_t = -\lambda X_t dt + \sigma dB_t, \quad X_0 = x.$$

The integral in (2.26) is

$$\phi_t(u) = \frac{1}{\lambda} \int_{e^{-\lambda t}}^1 \frac{\kappa(vu)}{v} dv = \frac{\sigma^2 u^2}{4\lambda} (1 - e^{-2\lambda t}). \quad (2.27)$$

With Brownian motion describing the continuous part of Lévy processes., for the second example we consider a pure jump process, namely a Double  $\Gamma$ -OU process.  $\Gamma$ -OU processes are generated by a compound Poisson process with jump intensity  $\lambda\beta$  ( $\lambda$  being the same as in (2.25)) and exponentially distributed jumps with expectation value  $\alpha$ . The limit distribution of this process is a  $\Gamma$ -distribution, which gives the process its name. As the generating compound Poisson process is strictly increasing, the generated  $\Gamma$ -OU process is a subordinator and stays above 0. In order to find an OU process with values in  $\mathbb{R}$  consider the difference of two independent compound  $\Gamma$ -OU processes  $L^+, L^-$  with parameters  $\alpha^+, \beta^+, \alpha^-, \beta^-$  and set  $\lambda^+ = \lambda\beta^+, \lambda^- = \lambda\beta^-$ . Then  $L = L^+ - L^-$  is a compound Poisson process, where positive jumps with expected jump size  $\frac{1}{\alpha^+}$  are arriving at rate  $\lambda^+$ , while negative jumps with expected jump size  $\frac{1}{\alpha^-}$  are arriving at rate  $\lambda^-$ .

The cumulant generating function of a compound Poisson process with exponential jumps is  $\frac{\lambda\beta u}{\alpha - u}$ , which is defined for  $u < \alpha$ . Hence for  $u \in (-\alpha^-, \alpha^+)$  the moment generating function of the combined process  $L$  is

$$\mathbb{E}[e^{uL}] = \mathbb{E}[e^{uL_1^+}] \mathbb{E}[e^{-uL_1^-}] = \exp\left(\lambda \frac{(\beta^+ + \beta^-)u^2 + (\beta^+\alpha^- - \beta^-\alpha^+)u}{(\alpha^+ - u)(\alpha^- + u)}\right).$$

Inserting this into (2.26) gives

$$\begin{aligned} \phi_t(u) &= \frac{1}{\lambda} \int_{e^{-\lambda t}}^1 \frac{\kappa(vu)}{v} dv = \int_{e^{-\lambda t}}^1 \frac{(\beta^+ + \beta^-)u^2 v + (\beta^+\alpha^- - \beta^-\alpha^+)u}{(\alpha^+ - uv)(\alpha^- + uv)} dv \\ &= -\frac{\beta^+ + \beta^-}{2} \int_{e^{-\lambda t}}^1 \frac{-2u^2 v + (\alpha^+ - \alpha^-)u}{(\alpha^+ - uv)(\alpha^- + uv)} dv \\ &\quad + \int_{e^{-\lambda t}}^1 \frac{(\beta^+\alpha^- - \beta^-\alpha^+)u + (\beta^+ + \beta^-)(\alpha^+ - \alpha^-)\frac{u}{2}}{(\alpha^+ - uv)(\alpha^- + uv)} dv \end{aligned}$$

The first integral is the differentiated logarithm of the denominator, so for this integral we get

$$-\frac{\beta^+ + \beta^-}{2} \ln((\alpha^+ - uv)(\alpha^- + uv)) \Big|_{v=e^{-\lambda t}}^1 = \frac{\beta^+ + \beta^-}{2} \ln\left(\frac{(\alpha^+ - e^{-\lambda t}u)(\alpha^- + e^{-\lambda t}u)}{(\alpha^+ - u)(\alpha^- + u)}\right).$$

The second integral is

$$\frac{\beta^+ + \beta^-}{2} \int_{e^{-\lambda t}}^1 \frac{u}{\alpha^+ - vu} + \frac{u}{\alpha^- + vu} dv,$$

so that the integrands are again differentiated logarithms of their denominators. Hence we get that the integral is

$$(-\ln(\alpha^+ - vu) + \ln(\alpha^- + vu))\Big|_{v=e^{-\lambda t}}^1 = \ln\left(\frac{(\alpha^+ - e^{-\lambda t}u)(\alpha^- + u)}{(\alpha^+ - u)(\alpha^- + e^{-\lambda t}u)}\right).$$

In summary the function  $\phi$  is given by

$$\begin{aligned} \phi_t(u) = & \frac{\beta^+ + \beta^-}{2} \ln\left(\frac{(\alpha^+ - e^{-\lambda t}u)(\alpha^- + e^{-\lambda t}u)}{(\alpha^+ - u)(\alpha^- + u)}\right) \\ & + \frac{\beta^+ - \beta^-}{2} \ln\left(\frac{(\alpha^+ - e^{-\lambda t}u)(\alpha^- + u)}{(\alpha^+ - u)(\alpha^- + e^{-\lambda t}u)}\right). \end{aligned} \quad (2.28)$$

It is also possible to combine the two approaches by considering an OU process generated by a Lévy process which is the difference of two compound Poisson processes plus a Brownian motion, all of which are independent. The resulting  $\phi$  then follows by adding up the two functions (2.27) and (2.28) and for this process  $\mathcal{V} = \{u \in \mathbb{C} : -\alpha^- < \text{Re}(u) < \alpha^+\}$ . By the previous remark it is also possible to shift this process by  $\theta$ . Such OU processes are used in the following numerical examples.

## Volatility surfaces

With the OU process of the previous section it is possible to generate volatility smiles as well as volatility skews. For illustration we consider a term structure with constant interest rates of 3.5%. The tenor structure and therefore the forward interest rates are based on half year intervals. Implied volatilities are then calculated for caplets with maturities over a 5-year period and strikes ranging from 0.02 to 0.07. Figure 2.1 shows a skewed volatility surface while figure 2.2 shows a very pronounced smile, both of which are generated by an *OU* process of the just introduced type. As mentioned in the previous chapters forward interest rates in this type of model will be bounded from below. The bounds in these examples are at 1% for the forward interest rate expiring after half a year and decrease to basically 0% for the forward interest rate which expires in 5 years. Hence they are well within reasonable boundaries. For completeness an example of at-the-money implied volatilities for swaptions with maturities and underlying swap rates ranging from 2 to 7 years is displayed in figure 2.3.

## Conclusion

Classical interest rate market models are not capable of simultaneously allowing for semi-analytical pricing formulas for caplets and swaptions and guaranteeing nonnegative forward interest rates. One exception are the affine LIBOR models presented in Keller-Ressel et al. [36]. We modify their approach to also allow for driving processes

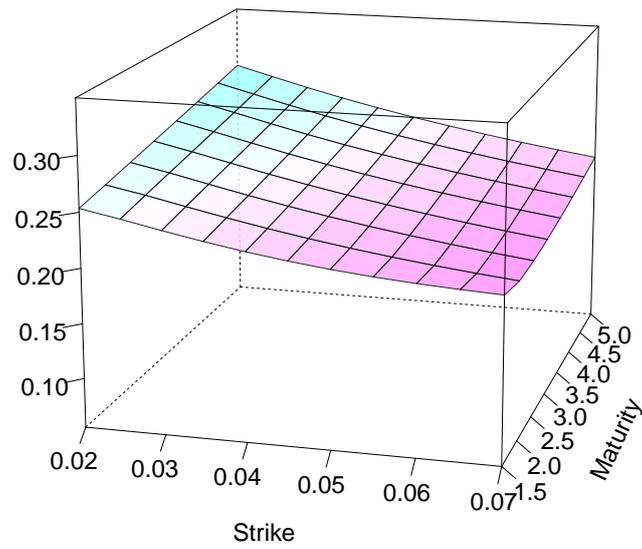


Figure 2.1: Implied volatility skew of caplets generated by an OU process with parameters  $\lambda = 0.02, \alpha^+ = 12, \alpha^- = 10, \beta^+ = 50, \beta^- = 5, \sigma = 0.3, \theta = 0.5, x = 0.7$  and  $T = 10$ .

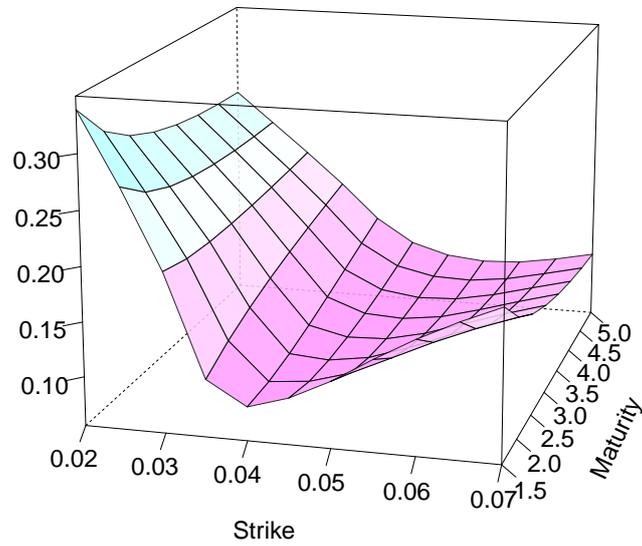


Figure 2.2: Implied volatility smile of caplets generated by an OU process with parameters  $\lambda = 0.02, \alpha^+ = 50, \alpha^- = 5, \beta^+ = 50, \beta^- = 10, \sigma = 0, \theta = 0, x = 1$  and  $T = 10$ .

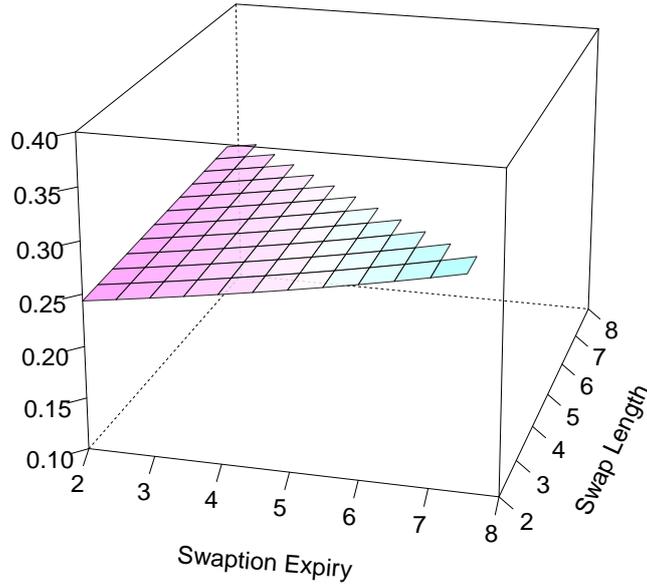


Figure 2.3: Swaption implied volatilities generated by an OU process with parameters  $\lambda = 0.02$ ,  $\alpha^+ = 12$ ,  $\alpha^- = 10$ ,  $\beta^+ = 50$ ,  $\beta^- = 5$ ,  $\sigma = 0.3$ ,  $\theta = 0.5$ ,  $x = 0.7$  and  $T = 10$ .

which are not necessarily nonnegative. Caplet and swaption valuation is possible via one-dimensional numerical integration. This allows for a fast calculation of implied volatilities for these types of interest rate derivatives. With the additional flexibility of real-valued affine processes this type of model is capable of producing skewed implied volatility surfaces as well as implied volatility surfaces with pronounced smiles.

## 2.4 Proofs

*Proof of Lemma 2.3.* For a function  $f(x)$  denote its even and odd part by

$$f^e(x) = \frac{1}{2}(f(x) + f(-x)), \quad f^o(x) = \frac{1}{2}(f(x) - f(-x)).$$

Note that if  $f$  is monotonically increasing, the same is true for  $f^o$ . Then (2.11) can be written as

$$\begin{aligned} M_t^u(x) &= \frac{1}{2} \left( e^{\phi_{T-t}(u) + \psi_{T-t}(u)x} + e^{\phi_{T-t}(-u) + \psi_{T-t}(-u)x} \right) \\ &= e^{\phi_{T-t}^e(u) + \psi_{T-t}^e(u)x} \cosh(\phi_{T-t}^o(u) + \psi_{T-t}^o(u)x) \end{aligned}$$

and

$$\frac{M_t^{u_i}(x)}{M_t^{u_0}(x)} = e^{(\phi_{T-t}^e(u_i) - \phi_{T-t}^e(u_0)) + (\psi_{T-t}^e(u_i) - \psi_{T-t}^e(u_0))x} \frac{\cosh(\phi_{T-t}^o(u_i) + \psi_{T-t}^o(u_i)x)}{\cosh(\phi_{T-t}^o(u_0) + \psi_{T-t}^o(u_0)x)}. \quad (2.29)$$

If  $u_i = u_0$ , then (2.29) is constant and has no influence regarding monotonicity or maxima. Hence from now on assume  $u_0 > u_i$  for all  $i$ . The function  $g$  of equation (2.12) can be written as

$$g(x) = \sum_{i=1}^n c_i e^{A_i} e^{a_i x} \frac{\cosh(B_i + b_i x)}{\cosh(B_0 + b_0 x)},$$

where for  $i = 0, \dots, n$

$$\begin{aligned} A_i &= (\phi_{T-t}^e(u_i) - \phi_{T-t}^e(u_0)), & B_i &= \phi_{T-t}^o(u_i), \\ a_i &= (\psi_{T-t}^e(u_i) - \psi_{T-t}^e(u_0)), & b_i &= \psi_{T-t}^o(u_i). \end{aligned}$$

Since  $\psi$  is monotonically increasing (see Lemma 2.1), also  $\psi^o$  is monotonically increasing. With  $\psi^o(0) = 0$  it follows that  $b_i \geq 0$  for all  $i$ . Furthermore note that  $a_i < b_0 - b_i$  is equivalent to  $\psi_{T-t}(u_i) < \psi_{T-t}(u_0)$  and  $-a_i < b_0 - b_i$  is equivalent to  $\psi_{T-t}(-u_0) < \psi_{T-t}(-u_i)$ . Since  $u_0 > u_i \geq 0$  the monotonicity of  $\psi$  yields

$$|a_i| < b_0 - b_i. \quad (2.30)$$

An elementary calculation gives

$$g'(x) = \frac{1}{\cosh(B_0 + b_0 x)^2} \sum_{i=1}^n c_i e^{A_i} e^{a_i x} f_i(x),$$

where

$$\begin{aligned} f_i(x) &= a_i \cosh(B_i + b_i x) \cosh(B_0 + b_0 x) + b_i \sinh(B_i + b_i x) \cosh(B_0 + b_0 x) \\ &\quad - b_0 \cosh(B_i + b_i x) \sinh(B_0 + b_0 x). \end{aligned}$$

The derivative of  $f_i$  is

$$\begin{aligned} f_i'(x) &= b_i \cosh(B_0 + b_0 x) \left( a_i \sinh(B_i + b_i x) + (b_i - b_0) \cosh(B_i + b_i x) \right) \\ &\quad + b_0 \cosh(B_i + b_i x) \left( a_i \sinh(B_0 + b_0 x) + (b_i - b_0) \cosh(B_0 + b_0 x) \right). \end{aligned} \quad (2.31)$$

Using (2.30) the terms inside the brackets of each row in (2.31) are strictly less than

$$|a_i| (\operatorname{sgn}(a_i) \sinh(B_j + b_j x) - \cosh(B_j + b_j x)) \leq 0, \quad (j = i, 0).$$

The last inequality is true since  $\cosh(x) \pm \sinh(x) \geq 0$ . The terms outside of the brackets in (2.31) are all positive. Hence  $f' \geq 0$  and the  $f_i$  are monotonically decreasing. Using (2.30) a simple calculation shows that  $\lim_{x \rightarrow -\infty} f_i(x) = \infty$  and  $\lim_{x \rightarrow \infty} f_i(x) = -\infty$ . Since  $c_i > 0$  for all  $i$  the same is true for  $\sum_{i=1}^n c_i e^{a_i x} f_i(x)$ . Hence  $g$  has a single maximum and is decreasing to the left and right of it. Furthermore  $g(x) \geq 0$  and again using (2.30)  $\lim_{x \rightarrow \infty} g(x) = \lim_{x \rightarrow -\infty} g(x) = 0$ .  $\square$

*Proof of Lemma 2.4.*  $f$  is continuous with compact support. Hence the extended Fourier transform  $\hat{f}(z) = \int_{\mathbb{R}} f(x)e^{-izx} dx$  exists for all  $z \in \mathbb{C}$  and is analytic. For  $z \neq 0, z \neq -iv_k$ , it is given by

$$\begin{aligned}\hat{f}(z) &= \int_{\kappa_1}^{\kappa_2} e^{-izx} f(x) dx = \frac{1}{iz} \int_{\kappa_1}^{\kappa_2} e^{-izx} f'(x) dx \\ &= \frac{1}{iz} \sum_k \frac{C_k v_k}{v_k - iz} (e^{(v_k - iz)\kappa_2} - e^{(v_k - iz)\kappa_1}).\end{aligned}$$

Since  $\hat{f}(u - iR) = O(u^{-2})$  for fixed  $R$ , it is absolutely integrable. By Fourier inversion

$$f(x) = \frac{1}{2\pi} \int_{\text{Im}(z)=-R} e^{izx} \hat{f}(z) dz = \frac{1}{2\pi} \int_0^\infty \text{Re} \left( e^{(iu+R)x} \hat{f}(u - iR) \right) du,$$

where the last equation follows from the fact that  $f$  is real valued and the symmetry  $\overline{\hat{f}(z)} = \hat{f}(-\bar{z})$ . Since  $\int \mathbb{E}[|e^{(iz+R)X_t}| | \mathcal{F}_s] |\hat{f}(z)| dz = \mathcal{M}_{X_t|X_s}(R) \int |\hat{f}(z)| dz$  is bounded if  $R \in \mathcal{V} \cap \mathbb{R}$ , conditional expectation and integration can be interchanged.  $\square$

## 3 Affine inflation market models

Arbitrage-free inflation models have first been rigorously introduced in Jarrow and Yildirim [31]. Since then several inflation models have been proposed. Similar to interest rate models one can distinguish between short rate models and market models. While short rate models in the spirit of Jarrow and Yildirim [31] aim at modeling the unobservable continuous nominal and real short rate, market models use discrete observable rates as the basis for modeling (see Belgrade et al. [1], Mercurio [38]). These observable rates are the basis of liquidly traded inflation swaps, zero coupon inflation-indexed swaps and year-on-year inflation-indexed swaps. While there have been several extensions of these models (e.g. Mercurio and Moreni [39, 40]) all of these models suffer the problem that there exist analytical formulas for only one type of swap, but not both. The model in this chapter leads to closed formulas for both types.

Based on the ideas in Keller-Ressel et al. [36] one can use affine processes to describe analytically highly tractable models. Affine processes are Markov process, where the characteristic function is of exponentially affine form, i.e.

$$\mathbb{E} \left[ e^{uX_t} | X_s \right] = e^{\phi_{t-s}(u) + \psi_{t-s}(u)X_s}.$$

The class of affine processes contains a large number of processes, e.g. every Lèvy process is affine. Using the longest-dated nominal zero coupon bond as a numeraire we can model the normalized bond prices as “exponential martingales” with respect to an affine process  $X$ . In this type of models we are not only able to price both types of inflation swaps, but can also derive semianalytical formulas for calls and puts on the underlying inflation rates, another liquidly traded inflation derivative.

The structure of this chapter is as follows. The first part describes inflation markets and typical traded derivatives. Afterwards we outline the setup of an inflation market model. The second part introduces the affine inflation market model and derives pricing formulas for the introduced inflation derivatives. The third part provides a concrete model specification including the calibration to actual market data. In section 3.4 we collect the properties of affine processes needed in this chapter. Furthermore we specify the affine processes used in the numerical section.

### 3.1 Inflation markets

Denote the time  $t$  price of the nominal zero coupon bond with maturity  $T$  by  $P(t, T)$  and consider an inflation index with time  $t$  value  $I(t)$ . Typically inflation indexes

are so-called consumer price indexes (CPI). To shorten notation we will use the term CPI synonymous for inflation index, nevertheless the reader can think of an arbitrary inflation index. The basic mathematical instruments in inflation-linked markets are so called inflation-linked zero coupon bonds (corresponding to zero coupon bonds for nominal interest rate markets). An inflation-linked zero coupon bond with maturity  $T$  is a bond paying  $I(T)$  at time  $T$ . Denote its price by  $P_{ILB}(t, T)$ .

In actual markets governments issue inflation-linked coupon bonds. Such a bond pays a fixed coupon on the variable basis  $I(T_k)/I(T_0)$  at some fixed number of pre-determined dates  $T_k \leq T$  (typically annually), where  $T_0$  is the time of issue. Additionally to the coupons such a bond redeems at maturity  $T$  with value  $\max\{I(T)/I(T_0), 1\}$ . Such a bond can therefore be described as a combination of inflation-linked zero coupon bonds plus an included option with payoff  $(1 - I(T)/I(T_0))_+$ . In general these bonds are issued with maturities of several years and inflation is positive. In this case the included option has little influence on the total price, which is why it is market practice to mostly ignore it. In particular if one ignores these options it is possible to strip inflation-linked zero coupon bond prices out of actually traded inflation-linked coupon bonds by the same methods used for nominal quantities.

Consider the quantity

$$P_R(t, T) := \frac{P_{ILB}(t, T)}{I(t)}, \quad (3.1)$$

which is called the price of a real zero-coupon bond. Note that this is not the price<sup>1</sup> of a traded asset, but a theoretical quantity. The usage of the term price is motivated by the fact that this quantity can be viewed as the price of a zero coupon bond in a fictitious economy, where everything is measured in terms of the inflation index  $I(t)$ <sup>2</sup>. Given real zero-coupon bond prices continuously compounded real interest rates are defined by  $R(t, T) := -\ln(P_R(t, T))/(T - t)$ . Accordingly one can define real counterparts to other nominal quantities such as forward interest rates or the short rate. This quantities are sometimes used as a starting point for inflation option pricing models (see e.g. Jarrow and Yildirim [31], Mercurio [38]).

Next to inflation-linked bond markets there exist several liquidly traded inflation-linked derivatives. First consider the forward price of the inflation index (forward CPI), i.e. the at time  $t$  fixed value  $\mathcal{I}(t, T)$ , which at time  $T$  can be exchanged against  $I(T)$  without additional costs. Since  $P_{ILB}(t, T)$  is the current price of  $I(T)$ , the time  $t$  forward CPI for maturity  $T$  is

$$\mathcal{I}(t, T) := \frac{P_{ILB}(t, T)}{P(t, T)}. \quad (3.2)$$

---

<sup>1</sup>Here we mean price in terms of money which has to be paid at a transaction. In fact it is essentially this number that is quoted on trading screens. However, in case such a bond is traded, the cash-flow is then the quoted number multiplied by the according index ratio.

<sup>2</sup>One could interpret  $I(t)$  as a numeraire, but one has to be careful not to use this as a mathematical numeraire, since  $I(t)$  is not actually traded.

In a zero coupon inflation-indexed swap (ZCIIS) two parties exchange the realized inflation  $\frac{I(T)}{I(t)}$  against a fixed amount  $(1 + K)^{T-t}$ . ZCIIS are mostly traded for full year maturities  $M$ . For  $T = t + M$  the value of such a payer swap can be expressed as

$$P(t, T) \left( \frac{\mathcal{I}(t, T)}{I(t)} - (1 + K)^M \right). \quad (3.3)$$

The rate  $K$ , for which equation (3.3) is zero is then called the ZCIIS rate  $ZCIIS(t; M)$ . These ZCIIS rates are quoted in the market for several full-year maturities.

*Remark:* Note that ZCIIS rates and inflation-linked bonds are closely related via (3.2) and (3.3). In reality this relationship is not observed. This is partly due to different creditworthiness of counterparties in bond and swap markets. A more detailed analysis of this difference can be found in Fleckenstein et al. [22]. For model calibration one has to choose one market, usually the swap market.

Next to ZCIIS there is a second important type of swap in inflation markets, the year-on-year inflation-indexed swaps (YYIIS). These swaps exchange the annualized inflation against a fixed rate  $K$ , i.e. consider an annually spaced tenor structure  $T_k = t + k, k = 0, \dots, M$ . The netted payment of a payer YYIIS at time  $T_k$  is  $(I(T_k)/I(T_{k-1}) - 1) - K$ . Hence the inflation leg consists of payoffs of the form

$$\frac{1}{T - S} \left( \frac{I(T)}{I(S)} - 1 \right).$$

Denote the forward value of such a payoff, the annualized forward inflation rate, by  $F_I(t, S, T)$ . Then the value of a payer YYIIS with maturity  $M$  and strike  $K$  can be expressed as

$$\sum_{k=1}^M P(t, T_k) (F_I(t, T_{k-1}, T_k) - K).$$

The YYIIS rate  $YYIIS(t; M)$  is the rate  $K$  such that the corresponding YYIIS has zero value. Note that given YYIIS rates for all annual maturities one can calculate annual forward inflation rates  $F_I(t, T_{k-1}, T_k)$  and vice versa.

Forward CPIs  $\mathcal{I}(t, T)$ , respectively forward inflation rates  $F_I(t, S, T)$  are the mathematical quantities underlying the market-traded ZCIIS, respectively YYIIS. Inflation market models aim at modeling these quantities. In existing inflation market models either  $\mathcal{I}(t, T)$  or  $F_I(t, S, T)$  can be expressed by analytical formulas, but not both. Consider a market, where price processes are assumed to be semimartingales on a filtered probability space  $(\Omega, \mathcal{A}, (\mathcal{F}_t), \mathbb{P})$ . Fix a  $T$ -forward measure  $\mathbb{Q}^T$ , i.e. a probability measure equivalent to  $\mathbb{P}$  such that asset prices normalized with the numeraire price  $P(t, T)$  are  $\mathbb{Q}^T$ -martingales. Then

$$\begin{aligned} \mathcal{I}(t, T) &= \mathbb{E}^{\mathbb{Q}^T} [I(T) | \mathcal{F}_t] \\ F_I(t, S, T) &= \mathbb{E}^{\mathbb{Q}^T} \left[ \frac{1}{T - S} \left( \frac{I(T)}{I(S)} - 1 \right) \middle| \mathcal{F}_t \right]. \end{aligned}$$

Calculating the expectations of the inflation index, as well as the fraction of the inflation index at two different times proves difficult. In the model of this chapter both are exponentially affine in the underlying driving stochastic process. For an affine process (see section 3.4) such expectations can be calculated and we are able to give semianalytical formulas for many standard options like caps and floors of forward inflation rates.

### 3.1.1 The inflation market model

We now introduce the general setup of an (inflation) market model. Consider a tenor structure  $0 < T_1 < \dots < T_N =: T$  and a market consisting of zero coupon bonds with maturities  $T_k$  and prices  $P(t, T_k)$ . The price processes  $(P(t, T_k))_{0 \leq t \leq T_k}$  are assumed to be positive semimartingales on a filtered probability space  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$  (here  $\mathcal{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ ), which satisfy  $P(T_k, T_k) = 1$  almost surely. If there exists an equivalent probability measure  $\mathbb{Q}^T$  such that the normalized bond price processes  $P(\cdot, T_k)/P(\cdot, T)$  are martingales<sup>3</sup>, the market is arbitrage-free. This setup describes the class of interest rate market models like the classical LIBOR market model (Brace et al. [3]) and its extensions.

To extend this setup to inflation markets consider an inflation index  $I$ , where we assume w.l.o.g. that  $I(0) = 1$ . Assume there exist inflation-linked zero-coupon bonds with the same maturities<sup>4</sup>  $T_1, \dots, T_N$  and price processes  $P_{ILB}(t, T_k)_{0 \leq t \leq T_k}$ , all of which are positive semimartingales<sup>5</sup>. If there exists an equivalent probability measure  $\mathbb{Q}^T$  such that all normalized price processes

$$\left( \frac{P(t, T_k)}{P(t, T)} \right)_{0 \leq t \leq T_k}, \quad \left( \frac{P_{ILB}(t, T_k)}{P(t, T)} \right)_{0 \leq t \leq T_k} \quad (3.4)$$

are  $\mathbb{Q}^T$ -martingales, the extended market model is arbitrage-free. For given  $\mathbb{Q}^T$  define the  $T_k$ -forward measures  $\mathbb{Q}^{T_k}$  by

$$\frac{d\mathbb{Q}^{T_k}}{d\mathbb{Q}^T} = \frac{1}{P(T_k, T)} \frac{P(0, T)}{P(0, T_k)}. \quad (3.5)$$

Under  $\mathbb{Q}^{T_k}$  the forward interest rate

$$F^k(t) := \frac{1}{\Delta_k} \left( \frac{P(t, T_{k-1})}{P(t, T_k)} - 1 \right), \quad \Delta_k := T_k - T_{k-1},$$

---

<sup>3</sup>One can extend bond price processes to  $[0, T]$  by setting  $P(t, T_k) := \frac{P(t, T)}{P(T_k, T)}$  for  $t > T_k$ , so that  $P(\cdot, T_k)/P(\cdot, T)$  is a martingale on  $[0, T]$  if and only if it is a martingale on  $[0, T_k]$ . Economically this can be interpreted as immediately investing the payoff of a zero coupon bond into the longest-running zero coupon bond.

<sup>4</sup>The assumption that for each zero coupon maturity there is a ILB with the same maturity is used only for notional convenience.

<sup>5</sup>The inflation index is only described through the bond prices  $P_{ILB}$ . I.e. the distribution of  $I(t)$  is only given at times  $T_k$ , where it coincides with the distribution of  $P_{ILB}(T_k, T_k)$ .

the earlier introduced forward CPI

$$\mathcal{I}(t, T_k) = \frac{P_{ILB}(t, T_k)}{P(t, T_k)},$$

and for  $j < k$  the forward inflation rates  $F_I(t, T_j, T_k)$  given by

$$1 + (T_k - T_j)F_I(t, T_j, T_k) = \mathbb{E}^{\mathbb{Q}^{T_k}} \left[ \frac{I(T_k)}{I(T_j)} \mid \mathcal{F}_t \right]$$

are all martingales. Modeling (some of) these martingales is the starting point of inflation market models in the literature (see e.g. Mercurio [38]). In contrast we start by modeling the normalized bond prices in (3.4) and derive the above quantities thereof.

### 3.2 The Affine inflation market model

Let  $(X_t)_{0 \leq t \leq T}$  with  $X_0 = x$  be an analytic affine process with state space  $\mathbb{R}_{\geq 0}^m \times \mathbb{R}^n$ ,  $m > 0$ ,  $n \geq 0$  on the probability space  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{Q}^T)$  and define for  $k = 1, \dots, N$

$$\begin{aligned} \frac{P(t, T_k)}{P(t, T)} &:= M_t^{u_k}, & u_k &\in (\mathbb{R}_{\geq 0}^m \times \{0\}^n) \cap \mathcal{V}, \\ \frac{P_{ILB}(t, T_k)}{P(t, T)} &:= M_t^{v_k}, & v_k &\in \mathbb{R}^{m+n} \cap \mathcal{V}, \end{aligned} \tag{3.6}$$

where

$$M_t^u := \mathbb{E}^{\mathbb{Q}^T} [e^{u \cdot X_T} \mid \mathcal{F}_t] = \exp\left(\phi_{T-t}(u) + \psi_{T-t}(u) \cdot X_t\right), \quad u \in \mathcal{V}, \tag{3.7}$$

with  $\mathcal{V}$  defined in (3.19). The processes  $M_t^u$  are  $\mathbb{Q}^T$ -martingales by the definition of an affine process. Hence this model is arbitrage-free. Note that in (3.6) the parts of  $u_k$  corresponding to real-valued components of  $X$  are chosen to be zero. For a decreasing sequence  $u_1 \geq \dots \geq u_N \geq 0$  one then has  $M_t^{u_{k-1}} \geq M_t^{u_k}$ , so that forward interest rates  $F^k(t)$  are guaranteed to be nonnegative for all  $k$ . Contrary to interest rates<sup>6</sup> inflation rates are not required to be nonnegative, which is why we do not restrict  $v_k$  in (3.6). The values of  $u_k$  and  $v_k$  should be calibrated to fit the initial term structures, i.e.  $M_0^{u_k} = P(0, T_k)/P(0, T)$  and  $M_0^{v_k} = P_{ILB}(0, T_k)/P(0, T)$ . By Lemma 3.2 it follows that parameters  $u_k$  fitting a current term structure with nonnegative forward interest rates can always be chosen to be decreasing. For multidimensional affine processes such sequences are far from unique. Concrete specifications how to choose  $u_k$  and  $v_k$  will be presented in section 3.3.

<sup>6</sup>Although interest rates are currently negative in certain countries, interest rates are still bounded below by the costs of physically keeping money. We can incorporate bounds different from 0 by setting  $\frac{P(t, T_k)}{P(t, T)} := c_k M_t^{u_k}$ .

The big advantage of this setup is that (3.5) in this case reads

$$\frac{d\mathbb{Q}^{T_k}}{d\mathbb{Q}^T} = \frac{M_{T_k}^{u_k}}{M_0^{u_k}} = \frac{1}{M_0^{u_k}} \exp\left(\phi_{T-T_k}(u_k) + \psi_{T-T_k}(u_k) \cdot X_{T_k}\right) \quad (3.8)$$

which is exponentially affine in  $X$ . In particular it is easy to check (see Keller-Ressel et al. [36]) that for  $0 \leq s \leq r$  and  $\psi_{T-r}(u_k) + w \in \mathcal{V}$

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}^{T_k}} [e^{w \cdot X_r} | \mathcal{F}_s] &= \exp\left(\phi_{r-s}(\psi_{T-r}(u_k) + w) - \phi_{r-s}(\psi_{T-r}(u_k))\right) \\ &\quad \exp\left((\psi_{r-s}(\psi_{T-r}(u_k) + w) - \psi_{r-s}(\psi_{T-r}(u_k))) \cdot X_s\right). \end{aligned} \quad (3.9)$$

Hence the moment generating function of  $X$  is also known<sup>7</sup> under different measures  $\mathbb{Q}^{T_k}$ . Together with the exponential affine form of basic quantities this is the reason why this model is analytically highly tractable. For example the forward rates  $F^k$  satisfy

$$(1 + \Delta_k F^k(t)) = \frac{M_t^{u_{k-1}}}{M_t^{u_k}} = e^{A(t, u_{k-1}, u_k) + B(t, u_{k-1}, u_k) \cdot X_t},$$

with

$$\begin{aligned} A(t, v, u) &:= \phi_{T-t}(v) - \phi_{T-t}(u), \\ B(t, v, u) &:= \psi_{T-t}(v) - \psi_{T-t}(u). \end{aligned} \quad (3.10)$$

Hence the  $\mathbb{Q}^{T_k}$ -extended moment generating function of  $\ln(1 + \Delta_k F^k(t))$  can be calculated explicitly using (3.9). The price of a caplet then follows using a Fourier-inversion formula (see Keller-Ressel et al. [36]). Swaptions can also be dealt with (Keller-Ressel et al. [36], Grbac et al. [25]) and so the most common interest rates derivatives can be calculated efficiently. We can use similar methods for inflation derivatives.

### 3.2.1 Forward CPI and CPI options

As mentioned before, the main advantage of this model is that for several important quantities the moment generating function is known under all forward measures  $\mathbb{Q}^{T_k}$ . Start by looking at the forward CPI

$$\mathcal{I}(t, T_k) = \frac{P_{ILB}(t, T_k)}{P(t, T_k)} = \frac{P_{ILB}(t, T_k)}{P(t, T)} \frac{P(t, T)}{P(t, T_k)} = \frac{M_t^{v_k}}{M_t^{u_k}} = e^{A(t, v_k, u_k) + B(t, v_k, u_k) \cdot X_t}, \quad (3.11)$$

with  $A$  and  $B$  defined in (3.10). Hence the forward CPI is of exponential affine form and therefore the  $\mathbb{Q}^{T_k}$ -moment generating function of its logarithm can be calculated

---

<sup>7</sup>This also shows that  $X$  is a time-inhomogeneous affine process under  $\mathbb{Q}^{T_k}$ .

using (3.9). In particular, setting  $A_I^k := A(T_k, v_k, u_k)$ ,  $B_I^k := B(T_k, v_k, u_k)$  and using  $I(T_k) = \mathcal{I}(T_k, T_k)$  one has

$$\begin{aligned}
\mathcal{M}_{\ln(I(T_k))|\mathcal{F}_s}^{\mathbb{Q}^{T_k}}(z) &:= \mathbb{E}^{\mathbb{Q}^{T_k}} [I(T_k)^z | \mathcal{F}_s] = \mathbb{E}^{\mathbb{Q}^{T_k}} \left[ \exp\left(zA_I^k + zB_I^k \cdot X_{T_k}\right) | \mathcal{F}_s \right] \\
&= \exp\left(zA_I^k + \phi_{T_k-s}(\psi_{T-T_k}(u_k) + zB_I^k) - \phi_{T_k-s}(\psi_{T-T_k}(u_k))\right) \\
&\quad \exp\left((\psi_{T_k-s}(\psi_{T-T_k}(u_k) + zB_I^k) - \psi_{T_k-s}(\psi_{T-T_k}(u_k))) \cdot X_s\right) \\
&= \exp\left(z\phi_{T-T_k}(v_k) + (1-z)\phi_{T-T_k}(u_k)\right) \\
&\quad \exp\left(\phi_{T_k-s}(z\psi_{T-T_k}(v_k) + (1-z)\psi_{T-T_k}(u_k))\right) \\
&\quad \exp\left(\psi_{T_k-s}(z\psi_{T-T_k}(v_k) + (1-z)\psi_{T-T_k}(u_k)) \cdot X_s\right) / M_s^{u_k}.
\end{aligned}$$

Here the last equality follows using (3.20). Note that this function is well defined and analytic in  $z$  if  $z\psi_{T-T_k}(v_k) + (1-z)\psi_{T-T_k}(u_k) \in \text{int}(\mathcal{V})$ .

Given the moment generating function of  $\ln(I(T_k))$  CPI calls and puts can be calculated using the following well-known Fourier inversion formula (see e.g. Eberlein et al. [19]). If  $R \in (1, \infty)$  such that  $\mathcal{M}_{X|\mathcal{F}}(R) < \infty$ , then

$$\mathbb{E}[(e^X - K)_+ | \mathcal{F}] = \frac{K}{\pi} \int_0^\infty \text{Re} \left( \mathcal{M}_{X|\mathcal{F}}(iu + R) \frac{K^{-(iu+R)}}{(iu + R)(iu + R - 1)} \right) du. \tag{3.12}$$

Thus the price of a forward CPI call with maturity  $T_k$  and payoff  $(I(T_k) - K)_+$  is

$$\begin{aligned}
&\text{CPICall}(t, T_k, K) \\
&= \frac{KP(t, T_k)}{\pi} \int_0^\infty \text{Re} \left( \mathcal{M}_{\ln(I(T_k))|\mathcal{F}_s}^{\mathbb{Q}^{T_k}}(iu + R) \frac{K^{-(iu+R)}}{(iu + R)(iu + R - 1)} \right) du,
\end{aligned}$$

where  $R > 1$  is chosen to satisfy  $R\psi_{T-T_k}(v_k) + (1-R)\psi_{T-T_k}(u_k) \in \text{int}(\mathcal{V})$ .

*Remark:* In section 3.1 it is mentioned that ILBs usually come with an included option guaranteeing a redemption of at least the original nominal amount. For an ILB issued at  $S$  and with maturity  $T_k$  this translates into an option  $(1 - I(T_k)/I(S))_+$  which corresponds to  $1/I(S)$  CPI puts with strike  $I(S)$ .

### 3.2.2 Forward Inflation and inflation caplets and floorlets

Typically inflation market models are not able to handle both forward CPI and forward inflation products analytically. With the approach presented here forward inflation rates are of a similar form as forward CPIs. The annualized inflation  $F_I(T_k, T_{k-j}, T_k)$  satisfies

$$1 + (T_k - T_{k-j})F_I(T_k, T_{k-j}, T_k) = \frac{I(T_k)}{I(T_{k-j})} = e^{A_I^k + B_I^k \cdot X_{T_k} - A_I^{k-j} - B_I^{k-j} \cdot X_{T_{k-j}}} =: e^{Y^k}. \tag{3.13}$$

The following Lemma gives the moment generating functions for random variables of this type.

**Lemma 3.1:** *Let  $s \leq r \leq t \leq T$  and*

$$\psi_{T-t}(u_k) + w \in \mathcal{V} \text{ and } \psi_{t-r}(\psi_{T-t}(u_k) + w) - \psi_{T-r}(u_k) + u \in \mathcal{V}.$$

Then

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}^{T_k}} [e^{u \cdot X_r + w \cdot X_t} | \mathcal{F}_s] &= \exp\left(\left(\psi_{r-s}(\psi_{t-r}(\psi_{T-t}(u_k) + w) + u) - \psi_{T-s}(u_k)\right) \cdot X_s\right) \\ &\quad \exp\left(\phi_{t-r}(\psi_{T-t}(u_k) + w) + \phi_{r-s}(\psi_{t-r}(\psi_{T-t}(u_k) + w) + u) - \phi_{t-s}(\psi_{T-t}(u_k))\right). \end{aligned} \quad (3.14)$$

*Proof.* Using the tower property and applying (3.9) twice it follows

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}^{T_k}} [e^{u \cdot X_r + w \cdot X_t} | \mathcal{F}_s] &= \mathbb{E}^{\mathbb{Q}^{T_k}} \left[ \mathbb{E}^{\mathbb{Q}^{T_k}} [e^{w \cdot X_t} | \mathcal{F}_r] e^{u \cdot X_r} | \mathcal{F}_s \right] \\ &= \mathbb{E}^{\mathbb{Q}^{T_k}} \left[ \exp\left(\phi_{t-r}(\psi_{T-t}(u_k) + w) - \phi_{t-r}(\psi_{T-t}(u_k))\right) \right. \\ &\quad \left. \exp\left(\left(\psi_{t-r}(\psi_{T-t}(u_k) + w) - \psi_{t-r}(\psi_{T-t}(u_k)) + u\right) \cdot X_r\right) | \mathcal{F}_s \right] \\ &= \exp\left(\phi_{t-r}(\psi_{T-t}(u_k) + w) - \phi_{t-r}(\psi_{T-t}(u_k))\right) \\ &\quad \exp\left(\phi_{r-s}(\psi_{t-r}(\psi_{T-t}(u_k) + w) + u) - \phi_{r-s}(\psi_{T-r}(u_k))\right) \\ &\quad \exp\left(\left(\psi_{r-s}(\psi_{t-r}(\psi_{T-t}(u_k) + w) + u) - \psi_{r-s}(\psi_{T-r}(u_k))\right) \cdot X_s\right) \\ &= \exp\left(\phi_{t-r}(\psi_{T-t}(u_k) + w) + \phi_{r-s}(\psi_{t-r}(\psi_{T-t}(u_k) + w) + u) - \phi_{t-s}(\psi_{T-t}(u_k))\right) \\ &\quad \exp\left(\left(\psi_{r-s}(\psi_{t-r}(\psi_{T-t}(u_k) + w) + u) - \psi_{r-s}(\psi_{T-r}(u_k))\right) \cdot X_s\right). \end{aligned}$$

Here we used the semiflow property (3.20) to simplify the expression.  $\square$

By Lemma 3.1 the  $\mathbb{Q}^{T_k}$ -moment generating function of  $Y^k$  defined in (3.13) is

$$\begin{aligned} \mathcal{M}_{Y^k | \mathcal{F}_s}^{\mathbb{Q}^{T_k}}(z) &= \mathbb{E}^{\mathbb{Q}^{T_k}} \left[ \exp\left(z A_I^k + z B_I^k \cdot X_{T_k} - z A_I^{k-j} - z B_I^{k-j} \cdot X_{T_{k-j}}\right) | \mathcal{F}_s \right] \\ &= \exp\left(z A_I^k - z A_I^{k-j} + \phi_{T_k - T_{k-j}}(\psi_{T-T_k}(u_k) + z B_I^k)\right) \\ &\quad \exp\left(\phi_{T_{k-j} - s}(\psi_{T_k - T_{k-j}}(\psi_{T-T_k}(u_k) + z B_I^k) - z B_I^{k-j}) - \phi_{T_k - s}(\psi_{T-T_k}(u_k))\right) \\ &\quad \exp\left(\left(\psi_{T_{k-j} - s}(\psi_{T_k - T_{k-j}}(\psi_{T-T_k}(u_k) + z B_I^k) - z B_I^{k-j}) - \psi_{T-s}(u_k)\right) \cdot X_s\right), \end{aligned}$$

which is well-defined if

$$\left\{ \psi_{T-T_k}(u_k) + z B_I^k, \psi_{T_k - T_{k-j}}(\psi_{T-T_k}(u_k) + z B_I^k) - \psi_{T-T_{k-j}}(u_k) - z B_I^{k-j} \right\} \subset \mathcal{V}. \quad (3.15)$$

The forward inflation rate is then given by  $1 + (T_k - T_{k-j})F_I(t, T_{k-j}, T_k) = \mathcal{M}_{Y^k | \mathcal{F}_t}^{\mathbb{Q}^{T_k}}(1)$ . So for  $\psi_{T-T_{k-j}}(v_k) - B_I^{k-j} \in \mathcal{V}$  it is

$$1 + (T_k - T_{k-j})F_I(t, T_{k-j}, T_k) = \exp\left(\left(\psi_{T_{k-j}-t}(\psi_{T-T_{k-j}}(v_k) - B_I^{k-j}) - \psi_{T-t}(u_k)\right) \cdot X_t\right) \exp\left(\phi_{T-T_{k-j}}(u_{k-j}) + \phi_{T_{k-j}-t}(\psi_{T-T_{k-j}}(v_k) - B_I^{k-j}) + \phi_{T-t}(u_k)\right).$$

Furthermore the payoff of an inflation caplet with strike  $K$  is

$$(T_k - T_{k-j})(F_I(T_k, T_{k-j}, T_k) - K)_+ = \left(\frac{I(T_k)}{I(T_{k-j})} - \tilde{K}\right)_+$$

where  $\tilde{K} = 1 + (T_k - T_{k-j})K$ . With the Fourier inversion formula (3.12) one can calculate the price of an inflation caplet. In particular, we have for  $R > 1$

$$\begin{aligned} \text{InflCpl}(t, T_{k-j}, T_k, K) \\ = \frac{\tilde{K}P(t, T_k)}{\pi} \int_0^\infty \text{Re} \left( \mathcal{M}_{Y^k | \mathcal{F}_t}^{\mathbb{Q}^{T_k}}(iu + R) \frac{\tilde{K}^{-(iu+R)}}{(iu + R)(iu + R - 1)} \right) du, \end{aligned}$$

provided that  $\psi_{T-T_k}(u_k) + RB_I^k \in \text{int}(\mathcal{V})$  and  $\psi_{T_k-T_{k-j}}(\psi_{T-T_k}(u_k) + RB_I^k) - \psi_{T-T_{k-j}}(u_k) - RB_I^{k-j} \in \text{int}(\mathcal{V})$ .

### 3.2.3 Correlation

So far we considered the pricing of typical market traded options. Another important aspect is the correlation structure. The relevant quantities

$$\begin{aligned} \ln(1 + \Delta_k F_n^k(t)) &= \ln\left(\frac{M_t^{u_{k-1}}}{M_t^{u_k}}\right) = A(t, u_{k-1}, u_k) + B(t, u_{k-1}, u_k)X_t, \\ \ln(\mathcal{I}^j(t)) &= \ln\left(\frac{M_t^{v_j}}{M_t^{u_j}}\right) = A(t, v_j, u_j) + B(t, v_j, u_j)X_t. \end{aligned}$$

$$\ln(1 + \Delta_k F_I(t, T_{k-j}, T_k)) = \text{const} + \left(\psi_{T_{k-j}-t}(\psi_{T-T_k}(v_k) - B_I^{k-j}) - \psi_{T-t}(u_k)\right) \cdot X_t. \quad (3.16)$$

are all affine transformation of  $X_t$  and the correlation for two such terms is

$$\text{Cor}[A_t + B_t \cdot X_t, \tilde{A}_t + \tilde{B}_t \cdot X_t] = \frac{\text{Var}[B_t \cdot X_t, \tilde{B}_t \cdot X_t]}{\sqrt{\text{Var}[B_t \cdot X_t]} \sqrt{\text{Var}[\tilde{B}_t \cdot X_t]}}.$$

For independent components of  $X_t$  this simplifies to<sup>8</sup>

$$\frac{\sum_{i=1}^d B_t^i \tilde{B}_t^i \text{Var}[X_t^i]}{\sqrt{\sum_{i=1}^d (B_t^i)^2 \text{Var}[X_t^i]} \sqrt{\sum_{i=1}^d (\tilde{B}_t^i)^2 \text{Var}[X_t^i]}}. \quad (3.17)$$

Hence correlations strongly depend on  $B(t, u_{k-1}, u_k) = \psi_{T-t}(u_{k-1}) - \psi_{T-t}(u_k)$  and  $B(t, v_j, u_j) = \psi_{T-t}(v_j) - \psi_{T-t}(u_j)$ , respectively the structure of the  $v_k$  and  $u_k$ . The exact correlation depends on the used measure (e.g.  $\mathbb{Q}^{T_k}, \mathbb{P}$ ), but choosing  $v_k$  and  $u_k$  cleverly, one can guarantee that the correlation structure, i.e. the correlation signs, stay the same. Concrete specifications for meaningful correlation structures will be given in the next section. Similar observations can also be made for instantaneous correlations of the corresponding quantities in the case of continuous affine processes (see Grbac et al. [25] for the general idea).

### 3.3 Implementation example

We design the structure of the affine inflation market model in such a way that the calibration can be separated into the calibration to nominal market data and the calibration to inflation market data afterwards. The method used to calibrate to nominal market data is based on the ideas in Grbac et al. [25]. There they fit a multiple curve affine LIBOR market model by using a common driving process  $X^0$  plus additional driving processes  $X^1, \dots, X^M$ , all of which are independent, affine and nonnegative. For calibration they use caplets with full-year maturities and underlying forwards of tenors less than a year. Using one individual driving process for each year one can then use an iterative procedure to calibrate to market data. Their approach can also be used in this setup.

In particular consider a semiannual tenor structure  $T_k = k/2, k = 1, \dots, N$ ,  $N$  even, and a driving affine process consisting of  $M + 1 = N/2 + 1$  components  $X^0, X^1, \dots, X^M$ , all of which are independent analytic affine processes with functions  $\phi^i$  and  $\psi^i$ ,  $i = 0, \dots, M$ . Then by (3.23)

$$\begin{aligned} \phi_t(u) &= \sum_{i=0}^M \phi_t^i(u^i) \\ \psi_t(u) &= (\psi_t^0(u^0), \psi_t^1(u^1), \dots, \psi_t^M(u^M)), \end{aligned}$$

where  $u^i, i = 0, \dots, M$  denotes the corresponding component of  $u \in \mathbb{R}^{M+1}$ . To describe an affine inflation market model we also have to specify the vectors  $u_k$ . The structure of the vectors should be so that the following points are satisfied.

---

<sup>8</sup>Up to some technical conditions the variance of a one-dimensional affine process is

$$\text{Var}[X_t^i] = \frac{\partial^2}{\partial^2 u} \Big|_{u=0} (\phi_t^i(u) + \psi_t^i(u) X_0^i).$$

	$X^0$	$X^1$	$X^2$	$\dots$	$X^M$	$X^{M+1}$	$X^{M+2}$	$\dots$	$X^{2M}$
$u_1$	$\tilde{u}_1$	$\bar{u}_1$	$\bar{u}_3$	$\dots$	$\bar{u}_{N-1}$	0	0	$\dots$	0
$u_2$	$\tilde{u}_2$	$\bar{u}_2$	$\bar{u}_3$	$\dots$	$\bar{u}_{N-1}$	0	0	$\dots$	0
$u_3$	$\tilde{u}_3$	0	$\bar{u}_3$	$\dots$	$\bar{u}_{N-1}$	0	0	$\dots$	0
$u_4$	$\tilde{u}_4$	0	$\bar{u}_4$	$\dots$	$\bar{u}_{N-1}$	0	0	$\dots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$u_{N-2}$	$\tilde{u}_{N-2}$	0	0	$\dots$	$\bar{u}_{N-1}$	0	0	$\dots$	0
$u_{N-1}$	$\tilde{u}_{N-1}$	0	0	$\dots$	$\bar{u}_{N-1}$	0	0	$\dots$	0
$u_N$	$\tilde{u}_N$	0	0	$\dots$	$\bar{u}_N$	0	0	$\dots$	0

Table 3.1: Description of the parameter structure  $u_k$ . Each row corresponds to one vector with the column names denoting the process the position in the vector corresponds to.

- Forward interest rates are nonnegative. This is the case if  $0 \leq u_k \leq u_{k-1}$ .
- The model matches the initial interest rate term structure. This basically fixes one component of each vector  $u_k$ .
- Calibration to market data is possible by using an iterative procedure.
- The model has a meaningful correlation structure.

This can be achieved by choosing the vectors  $u_k$  in the following way. They depend on  $2N$  real parameters

$$\tilde{u}_1 \geq \dots \geq \tilde{u}_N \geq 0, \quad \bar{u}_1 \geq \dots \geq \bar{u}_N \geq 0.$$

For  $1 \leq j \leq M$  set (compare table 3.1, ignoring the zero columns in the rightmost side for now)

$$u_k = \tilde{u}_k e^0 + \bar{u}_k e^{\lceil \frac{k}{2} \rceil} + \sum_{l=\lceil \frac{k}{2} \rceil + 1}^M \bar{u}_{2l-1} e^l$$

where  $e^0, e^1, \dots, e^M$  denote the base vectors  $(1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$  of  $\mathbb{R}^{M+1}$ . Note that with this choice the vectors  $(u_k)$  are decreasing. Given  $\tilde{u}_1, \dots, \tilde{u}_n$  and the processes  $X^0, X^1, \dots, X^M$  by Lemma 3.2 the parameters  $\bar{u}_1, \dots, \bar{u}_N$  are determined by fitting the current term structure. I.e. we require that

$$\frac{P(0, T_k)}{P(0, T)} = \mathbb{E}^{\mathbb{Q}^T} [e^{u_k \cdot X_T}] = \mathbb{E}^{\mathbb{Q}^T} [e^{\tilde{u}_k X_T^0}] \mathbb{E}^{\mathbb{Q}^T} \left[ e^{\bar{u}_k X_T^{\lceil \frac{k}{2} \rceil}} \right] \prod_{i=\lceil \frac{k}{2} \rceil + 1}^M \mathbb{E}^{\mathbb{Q}^T} [e^{\bar{u}_{2i-1} X_T^i}],$$

so that the parameters  $\bar{u}_1, \dots, \bar{u}_N$  can be calculated using backwards iteration. Furthermore the semiannual forward interest rates with full-year maturities  $F^{2k}$  only

depend on  $u_{2k-1}, u_{2k}$  and the processes  $X^0, X^k, \dots, X^M$ . Hence if  $X^0$  and  $\tilde{u}_1, \dots, \tilde{u}_N$  are already specified, one can fit  $X^M$  to caplets on the forward interest rate  $F^N$  and then go backwards to iteratively fit the processes  $X^k$  to caplets on forward interest rates  $F^{2k}$ . Hence if  $X^0$  and  $\tilde{u}_1, \dots, \tilde{u}_N$  are fixed, all the remaining parameters can be calibrated to the yield curve and caplet prices. As stated in Grbac et al. [25] and confirmed by our numerical tests the concrete choice of  $X^0$  and  $\tilde{u}_k$  (within a meaningful range) has no qualitative impact on the resulting calibration quality. Henceforth  $X^0$  is fixed as a CIR process (specifications of this process can be found equation (3.24)). So far we have said nothing about resulting correlations. This is where the choice of the parameters  $\tilde{u}_1, \dots, \tilde{u}_n$  comes in. The relevant functions for correlations of forward interest rates  $F^{2k}$  are

$$B(t, u_{2k-1}, u_{2k}) = (\psi_{T-t}^0(\tilde{u}_{2k-1}) - \psi_{T-t}^0(\tilde{u}_{2k}), 0, \dots, 0, \psi_{T-t}^k(\bar{u}_{2k-1}) - \psi_{T-t}^k(\bar{u}_{2k}), 0, \dots, 0).$$

These functions are “orthogonal” except for the first component<sup>9</sup>. Hence by equation (3.17) the correlation structure mainly depends on the sequence  $(\tilde{u}_k)$ . Since  $X^0$  is nonnegative the function  $\psi_t^0(u)$  is increasing in  $u$  (see Keller-Ressel et al. [36]). With  $(\tilde{u}_k)$  being a decreasing sequence this results in nonnegative correlations. Setting  $\tilde{u}_k = \tilde{u}$  for all  $k$  results in zero correlation<sup>10</sup>. The magnitude of correlations depends on how much of the variance of forward bond prices is explained by  $\tilde{u}_k X^0$ . Here we choose the factors  $\tilde{u}_k$  so that  $\mathbb{E}^{\mathbb{Q}^T} [e^{2\tilde{u}_k X_T^c}] = P(0, T_k)/P(0, T)$ . The idea behind this choice is that approximately half of the variance should be explained by the common factor. Alternative if one has additional information on correlations (e.g. through market data), this could be incorporated in  $\tilde{u}_k$ .

For calibration we used market data from September, 29th 2011. The yield curve is bootstrapped from LIBOR and swap rates and is displayed in figure 3.1. The used caplet implied volatilities for 6-month forward interest rates with strikes 1% to 6% and maturities from 1 to 10 years are bootstrapped from cap data. The resulting implied volatilities can be found in figure 3.2. We fixed the time horizon  $T = 10$  and chose the parameters of the common CIR process as  $\lambda = 0.026, \theta = 0.65, \eta = 0.5, x = 3.45$ . The  $M$  individual driving processes are chosen to be CIR processes with added jumps (see equation (3.25)). Their parameters and the parameters  $\bar{u}_k$  are calibrated with the mentioned recursive method. In each step the parameters are chosen so that the mean squared errors of implied volatilities are minimized. In contradiction with Grbac et al. [25] we were not able to produce a similar calibration quality as described in their paper<sup>11</sup>. The resulting calibration can be found in figure 3.2. Especially for long dated caplet volatilities the pronounced skew could not be reproduced. Nevertheless the model provides a reasonable fit of the caplet volatility surface, especially since the focus is on inflation derivatives.

<sup>9</sup>Although only stated for even forward rates for notional simplicity this is true for all forward rates.

<sup>10</sup>Negative correlations in this setup are only possible if the sequence  $(\tilde{u}_k)$  is not decreasing which means that forward interest rates can become negative.

<sup>11</sup>Several requests for clarification with the authors have resulted in the answer that their results are currently not in a state to be shared.

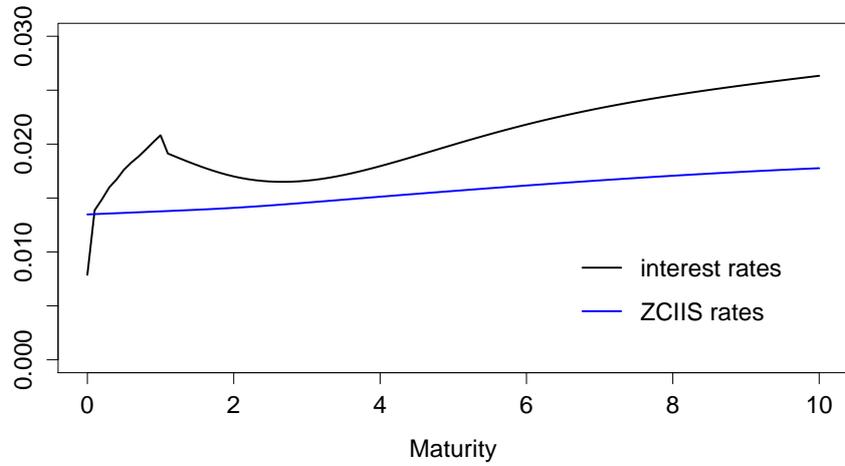


Figure 3.1: Yield curve and ZCIS curve from September, 29th 2011

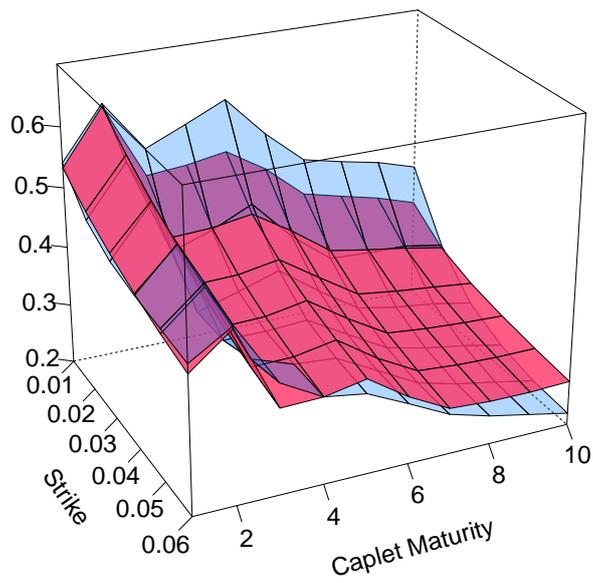


Figure 3.2: Market and model caplet implied volatilities of 6-month forward interest rates with strikes 1% to 6% and maturities from 1 to 10 years. Market volatilities are in a transparent blue, while model volatilities are displayed in red.

	$X^0$	$X^1$	$X^2$	$\dots$	$X^M$	$X^{M+1}$	$X^{M+2}$	$\dots$	$X^{2M}$
$v_1$	$\tilde{v}_1$	$\bar{u}_1$	$\bar{u}_3$	$\dots$	$\bar{u}_{N-1}$	$\bar{v}_1$	0	$\dots$	0
$v_2$	$\tilde{v}_2$	$\bar{u}_2$	$\bar{u}_3$	$\dots$	$\bar{u}_{N-1}$	$\bar{v}_2$	0	$\dots$	0
$v_3$	$\tilde{v}_3$	0	$\bar{u}_3$	$\dots$	$\bar{u}_{N-1}$	0	$\bar{v}_3$	$\dots$	0
$v_4$	$\tilde{v}_4$	0	$\bar{u}_4$	$\dots$	$\bar{u}_{N-1}$	0	$\bar{v}_4$	$\dots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$v_{N-2}$	$\tilde{v}_{N-2}$	0	0	$\dots$	$\bar{u}_{N-1}$	0	0	$\dots$	0
$v_{N-1}$	$\tilde{v}_{N-1}$	0	0	$\dots$	$\bar{u}_{N-1}$	0	0	$\dots$	$\bar{v}_{N-1}$
$v_N$	$\tilde{v}_N$	0	0	$\dots$	$\bar{u}_N$	0	0	$\dots$	$\bar{v}_N$

Table 3.2: Description of the inflation parameter structure  $v_k$ . Each row corresponds to one vector with the column names denoting the process the position in the vector corresponds to. Note that for annual inflation rate option pricing the vectors  $v_j$  with  $j$  odd do not matter.

The next step is to extend the calibration to inflation markets. Additionally to the  $M+1$  driving processes used for modeling interest rates we use another  $M$  independent analytic affine processes (one for each year) driving inflation-related quantities. Since the individual processes are independent by (3.23) the setup without inflation can be embedded by setting the additional components of  $u_k$  equal to zero (see table 3.1). So from now on we assume that the processes  $X^0, X^1, \dots, X^M$  and the vectors  $u_1, \dots, u_N$  are already fixed.

The choice of the inflation parameters  $v_k$  focuses on the same aspects as the choice of  $u_k$  without the restriction that inflation rates have to be nonnegative. We again assume that the vectors  $v_k$  are determined by  $2N$  parameters<sup>12</sup>  $\tilde{v}_1, \dots, \tilde{v}_N, \bar{v}_1, \dots, \bar{v}_N$ . In particular, we choose (see also table 3.2)

$$v_k = \tilde{v}_k e^0 + \bar{u}_k e^{\lceil \frac{k}{2} \rceil} + \sum_{l=\lceil \frac{k}{2} \rceil + 1}^M \bar{u}_{2l-1} e^l + \bar{v}_k e^{M + \lceil \frac{k}{2} \rceil}.$$

Choosing the nominal components  $v_k^i = u_k^i$  for  $i = 1, \dots, M$  has the advantage that forward CPIs and therefore also forward inflation rates do not depend on the nominal processes  $X^1, \dots, X^M$ . Together with the choice of  $v_k$  with respect to the inflation processes  $X^{M+1}, \dots, X^{2M}$  this implies that the forward CPI  $\mathcal{I}(t, T_{2k})$  depends only on  $X^0$  and  $X^{M+k}$ . In particular the function  $B(t, v_{2k}, u_{2k})$  corresponding to the forward CPI  $\mathcal{I}(t, T_{2k})$  defined in (3.11) is

$$B(t, v_{2k}, u_{2k}) = (\psi_{T-t}^0(\tilde{v}_{2k}) - \psi_{T-t}^0(\tilde{u}_{2k}), 0, \dots, 0, \psi_{T-t}^{M+k}(\bar{v}_{2k}), 0, \dots, 0).$$

<sup>12</sup>We actually only require  $N$  parameters, since the odd rows in table 3.2 do not contribute for the considered annual inflation rates. For notional simplicity we nevertheless consider  $2N$  parameters.

From this it follows that for different forward CPIs these functions are “orthogonal” except for the first component. Furthermore except for the first component they are also “orthogonal” to the functions  $B(t, u_{j-1}, u_j)$  relevant for forward interest rates. Hence the correlation structure mainly depends on  $\tilde{u}_k$  and  $\tilde{v}_k$ . Since  $\psi_t^0$  is monotonically increasing, one should choose  $\tilde{v}_k > \tilde{u}_k$  if the corresponding forward CPI should be positively correlated with nominal interest rates or  $\tilde{v}_k < \tilde{u}_k$  if it should be negatively correlated<sup>13</sup>. Also note that two annual forward CPIs are positively correlated if  $\text{sgn}(\tilde{v}_k - \tilde{u}_k) = \text{sgn}(\tilde{v}_j - \tilde{u}_j)$ . This therefore gives us some criteria how to determine the parameters  $\tilde{u}_k$  from correlation assumptions and from now on we assume that the parameters  $\tilde{v}_1, \dots, \tilde{v}_N$  are given. Assuming also a fixed process  $X^{M+k}$  we would like to determine  $\bar{v}_k$  from the current term structure  $P_{ILB}(0, T_k)/P(0, T)$ . We differentiate between two cases. First consider a nonnegative affine process  $X^{M+k}$ . In this case by Lemma 3.2 there exists a unique  $\bar{v}_k$ , so that  $M_0^{v_k} = P_{ILB}(0, T_k)/P(0, T)$ . For  $\tilde{v}_k \approx \tilde{u}_k$  this typically results in  $\bar{v}_k > 0$ . In this case zero coupon inflation for  $[t, T_k]$  is always positive. To avoid this alternatively consider an affine process, which takes negative and positive values. In this case  $M_0^v$  is not necessarily increasing in  $\bar{v}_k$ . However, by Lemma 3.2 it is still a convex function in  $\bar{v}_k$ . This means that there are at most two possible choices for  $\bar{v}_k$ , in which case we need to pick one. For the results in this chapter we only used results where one choice was smaller than zero and the other one larger than zero, which we then picked. To determine the processes  $X^{M+1}, \dots, X^{2M}$  notice that the annual forward inflation rate  $F_I(t, T_{2k-2}, T_{2k})$  depends only the processes  $X^0, X^{M+k-1}, X^k$  and  $F_I(t, T_0, T_2)$  depends only on  $X^0, X^{M+1}$ . By starting with inflation options on  $F_I(T_2, T_0, T_2)$  one can calibrate the parameters of  $X^{M+1}$ . Then one can iteratively calibrate the parameters of  $X^{M+k}$  using inflation options on  $F_I(T_{2k}, T_{2(k-1)}, T_{2k})$ .

For the calibration example we used ZCIIS rates from 1 to 10 years (see figure 3.1) and inflation options for strikes ranging from  $-2\%$  to  $6\%$ , the prices of which are displayed in figure 3.4. We chose  $\tilde{v}_k = \tilde{u}_k(1 + ck)$  with  $c \approx 0.08$ , so that  $ck$  is between 1 and 1.15. This means that correlation are positive as usually observed. For the processes  $X^{M+1}, \dots, X^{2M}$  we used Ornstein-Uhlenbeck processes with added jumps (see equation (3.26)). The parameters of these processes and the sequence  $(\bar{v}_k)$  are then calibrated as described. In each step the mean squared errors of option prices are minimized. The resulting fit is displayed in figure 3.4 and this shows that the calibration is very accurate.

We would also like to display the fit in terms of (shifted lognormal) implied volatilities. Typically one only has quotes on ZCIIS rates, which do not directly translate to forward inflation rates. However, annual forward inflation rates can be approximated by  $F_I(t, T_{k-j}, T_k) \approx \mathcal{I}(t, T_k)/\mathcal{I}(t, T_{k-j}) - 1$  (see figure 3.3). It is market practice to use this as forward value in the shifted Black formula to calculate approximate market

---

<sup>13</sup>We decided to introduce correlation for the inflation part only via the common factor process. We could also have changed the parameters  $\bar{u}_k$  in order to introduce correlation. In this case even more complicated correlation patterns could be created.

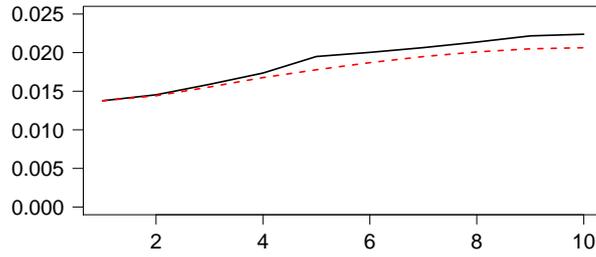


Figure 3.3: Linear interpolated annual forward inflation rates  $F_I(0, T_{2(k-1)}, T_{2k})$  (black) and its approximation  $\mathcal{I}(t, T_{2k})/\mathcal{I}(t, T_{2(k-1)}) - 1$  (dashed red) for maturities from 1 to 10 years.

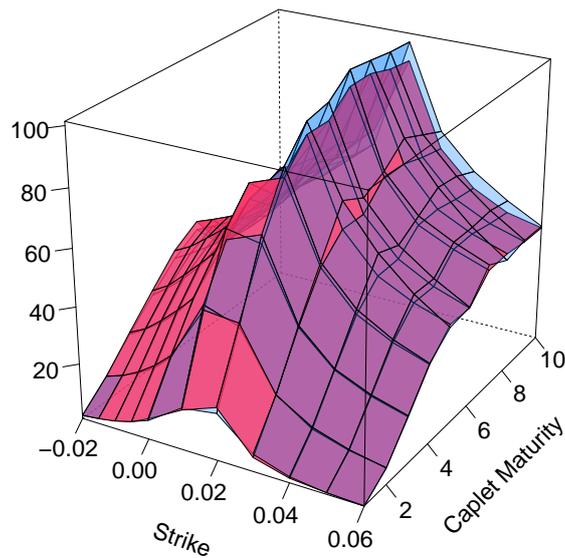


Figure 3.4: Market and model caplet/floorlet prices in basis points for annual forward inflation with strikes  $-2\%$  to  $6\%$  and maturities from 1 to 10 years. Market prices are in a transparent blue, while model prices are displayed in red. For strikes between  $-2\%$  and  $1\%$  prices are quoted for floorlets, for strikes between  $2\%$  and  $6\%$  prices are quoted for caplets. Market prices are bootstrapped from corresponding cap/floor data.

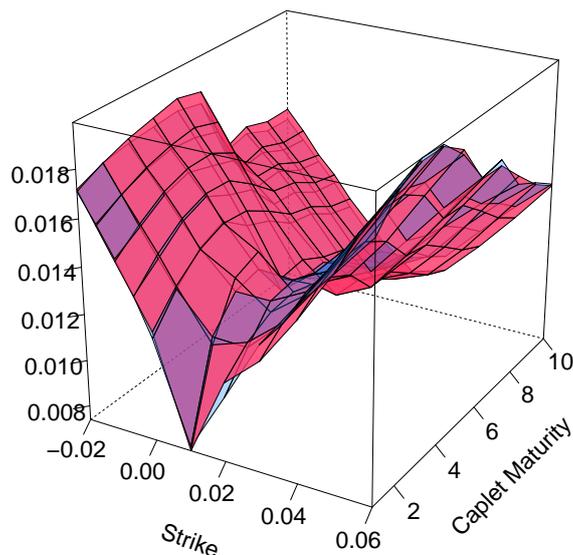


Figure 3.5: Market and model implied volatilities for annual forward inflation with strikes  $-2\%$  to  $6\%$  and maturities from 1 to 10 years. Market volatilities are in a transparent blue, while model volatilities are displayed in red.

volatilities. The resulting fit in terms of shifted implied volatilities is then displayed in figure 3.5. This shows that implied volatilities are closely reproduced by the model across all maturities and smiles and that the model is very capable in fitting the observed market data.

## Conclusion

We have introduced a highly tractable inflation market model, where we are able to derive analytical formulas for both types of inflation-indexed swaps. Furthermore inflation caps and floors as well as CPI caps and floors can be calculated with a one-dimensional Fourier inversion formula. Hence prices for liquidly traded inflation derivatives can be calculated quickly and accurately. Additionally the proposed model is able to price classical interest rate derivatives like caps and floors. Using these formulas we are able to calibrate the model to market data. The calibration example shows that the model can be calibrated to inflation market data very accurately.

## 3.4 Appendix on affine processes

Let  $X = (X_t)_{0 \leq t \leq T}$  be a homogeneous Markov process with values in  $D = \mathbb{R}_{\geq 0}^m \times \mathbb{R}^n$  on a measurable space  $(\Omega, \mathcal{A})$  with filtration  $(\mathcal{F}_t)_{0 \leq t \leq T}$ , with regards to which  $X$  is adapted. Denote by  $\mathbb{P}^x$  and  $\mathbb{E}^x[\cdot]$  the corresponding probability and expectation when

$X_0 = x$ .  $X$  is said to be an affine process, if its characteristic function has the form

$$\mathbb{E}^x [e^{u \cdot X_t}] = \exp\left(\phi_t(u) + \psi_t(u) \cdot x\right), \quad u \in \mathfrak{i}\mathbb{R}^d, x \in D,$$

where  $\phi : [0, T] \times \mathfrak{i}\mathbb{R}^d \rightarrow \mathbb{C}$  and  $\psi : [0, T] \times \mathfrak{i}\mathbb{R}^d \rightarrow \mathbb{C}^d$  and  $\mathfrak{i}\mathbb{R}^d = \{u \in \mathbb{C}^d : \operatorname{Re}(u) = 0\}$ . By homogeneity and the Markov property the conditional characteristic function satisfies

$$\mathbb{E}^x [e^{u \cdot X_t} | \mathcal{F}_s] = \exp\left(\phi_{t-s}(u) + \psi_{t-s}(u) \cdot X_s\right). \quad (3.18)$$

Accordingly affine processes can also be defined for inhomogeneous Markov processes (see Filipovic [21]). In this case the affine property reads

$$\mathbb{E}^x [e^{u \cdot X_t} | \mathcal{F}_s] = \exp\left(\phi_{s,t}(u) + \psi_{s,t}(u) \cdot X_s\right), \quad u \in \mathfrak{i}\mathbb{R}^d, x \in D,$$

with  $\phi_{s,t} : \mathfrak{i}\mathbb{R}^d \rightarrow \mathbb{C}$  and  $\psi_{s,t} : \mathfrak{i}\mathbb{R}^d \rightarrow \mathbb{C}^d$  for  $0 \leq s \leq t$ .

$X$  is called an analytic affine process, if  $X$  is stochastically continuous and the interior of the set<sup>14</sup>

$$\mathcal{V} := \left\{ u \in \mathbb{C}^d : \sup_{0 \leq s \leq T} \mathbb{E}^x [e^{\operatorname{Re}(u) \cdot X_s}] < \infty \quad \forall x \in D \right\}, \quad (3.19)$$

contains 0. In this case the functions  $\phi$  and  $\psi$  have continuous extensions to  $\mathcal{V}$ , which are analytic in the interior, such that (3.18) holds for all  $u \in \mathcal{V}$  (see Keller-Ressel [33]).

The class of affine processes includes Brownian motion and more generally all Lévy processes. Since Lévy processes have stationary independent increments, it follows that  $\psi_t(u) = u$ , while  $\phi_t(u) = t\kappa(u)$ , where  $\kappa$  is the cumulant generating function of the Lévy process. Ornstein-Uhlenbeck processes are further important examples of affine processes. The affine processes used in this work are described at the end of this section.

The standard reference for affine processes is Duffie et al. [16]. There they give a characterization of affine processes, where  $\phi$  and  $\psi$  are specified as solutions of a system of differential equations. To motivate this consider an affine process  $X$ . By the tower property for conditional expectations it holds for all  $x \in D$

$$\mathbb{E}^x [e^{u \cdot X_{t+s}}] = \mathbb{E}^x [\mathbb{E}^x [e^{u \cdot X_{t+s}} | \mathcal{F}_s]] = \mathbb{E}^x [e^{\phi_t(u) + \psi_t(u) \cdot X_s}].$$

Using equation (3.18) it follows that  $\phi$  and  $\psi$  satisfy the so-called semi-flow equations

$$\begin{aligned} \phi_{t+s}(u) &= \phi_t(u) + \phi_s(\psi_t(u)), & \phi_0(u) &= 0, \\ \psi_{t+s}(u) &= \psi_s(\psi_t(u)), & \psi_0(u) &= u. \end{aligned} \quad (3.20)$$

---

<sup>14</sup> $\mathcal{V}$  can be described as the (convex) set, where the extended moment generating function of  $X_t$  is defined for all times  $t$  and all starting values  $x$ . By Lemma 4.2 in Keller-Ressel and Mayerhofer [34] the set  $\mathcal{V}$  is in fact equal to the seemingly smaller set  $\{u \in \mathbb{C}^d : \exists x \in \operatorname{int}(D) : \mathbb{E}^x [e^{\operatorname{Re}(u) \cdot X_T}] < \infty\}$ .

For a stochastically continuous affine process  $X$  it was shown in Keller-Ressel et al. [35] that the functions

$$F(u) := \frac{\partial}{\partial t} \phi_t(u) \Big|_{t=0^+}, \quad R(u) := \frac{\partial}{\partial t} \psi_t(u) \Big|_{t=0^+}$$

exist<sup>15</sup>. Rewriting (3.20) in terms of difference quotients and letting  $s \rightarrow 0$  we get that  $\phi$  and  $\psi$  satisfy generalized Ricatti equations

$$\begin{aligned} \frac{\partial}{\partial t} \phi_t(u) &= F(\psi_t(u)), & \phi_0(u) &= 0, \\ \frac{\partial}{\partial t} \psi_t(u) &= R(\psi_t(u)), & \psi_0(u) &= u. \end{aligned} \quad (3.21)$$

The functions  $F$  and  $R$  have a specific form of Levy-Khintchine type as first described in Duffie et al. [16]. There it is also shown that for every  $F$  and  $R$  of this form (3.21) has a unique solution. Specifying the functions  $F$  and  $R$  is an alternative way to specify an affine process. Keller-Ressel and Mayerhofer [34] give conditions on  $F$  and  $R$  under which a solution of (3.21) defines an analytic affine process. Note that in order to evaluate  $\phi$  and  $\psi$  one would like to have closed form solutions to the system (3.21), which in general is not the case.

Coupling independent affine processes is very tractable. For two independent affine processes  $X$  and  $Y$  and all starting values  $x, y$  one obtains

$$\mathbb{E}^{(x,y)} [e^{(u_X, u_Y) \cdot (X_t, Y_t)}] = \mathbb{E}^{(x,y)} [e^{u_X \cdot X_t} e^{u_Y \cdot Y_t}] = \mathbb{E}^{(x,y)} [e^{u_X \cdot X_t}] \mathbb{E}^{(x,y)} [e^{u_Y \cdot Y_t}]. \quad (3.22)$$

Hence  $(X, Y)$  is an affine process with

$$\begin{aligned} \phi_t^{(X,Y)}(u_X, u_Y) &= \phi_t^X(u_X) + \phi_t^Y(u_Y), \\ \psi_t^{(X,Y)}(u_X, u_Y) &= (\psi_t^X(u_X), \psi_t^Y(u_Y)). \end{aligned} \quad (3.23)$$

This fact together with the following Lemma is used in section 3.3.

**Lemma 3.2:** *Let  $X$  be an analytic affine processes comprised of  $m + n$  independent affine processes, where the first  $m$  are nonnegative. For  $(u_1, \dots, u_k, v, u_{k+1}, \dots, u_n) \in \text{int}(\mathcal{V}) \cap \mathbb{R}^{m+n}$  define the function*

$$f^k(v) := \mathbb{E}^x [e^{(u_1, \dots, u_{k-1}, v, u_{k+1}, \dots, u_{m+n}) \cdot X_t}].$$

*Then  $f^k$  is monotonically increasing if  $k \leq m$  and  $f^k$  is convex for all  $k$ .*

*Proof.* For each  $x \in D$  the term inside the expectation is convex in  $v$  and monotonically increasing in  $v$  if  $k \leq m$ . This then also holds after taking the expectation.  $\square$

---

<sup>15</sup>This was also shown for affine processes with general state spaces in Keller-Ressel et al. [37] and Cuchiero and Teichmann [10].

The last part of this section describes the affine processes used in this chapter. One classical example is the CIR process, which is the unique solution of

$$dX_t = -\lambda(X_t - \theta) dt + 2\eta\sqrt{X_t} dW_t, \quad X_0 = x. \quad (3.24)$$

For this process the functions  $\phi$  and  $\psi$  are defined for  $\operatorname{Re}(u) < \frac{\lambda}{2\eta^2}(1 - e^{-\lambda t})^{-1}$ ,

$$\begin{aligned} \phi_t(u) &= -\frac{\lambda\theta}{2\eta^2} \ln \left( 1 - \frac{2\eta^2}{\lambda}(1 - e^{-\lambda t})u \right), \\ \psi_t(u) &= \frac{e^{-\lambda t}u}{1 - \frac{2\eta^2}{\lambda}(1 - e^{-\lambda t})u}. \end{aligned}$$

The CIR process almost surely stays nonnegative. It is strictly positive if  $\frac{\lambda\theta}{2} > \eta^2$ . One can add jumps to this process by adding the differential of a compound Poisson process  $L_t$  to the dynamics of  $X$ .

$$dX_t = -\lambda(X_t - \theta) dt + 2\eta\sqrt{X_t} dW_t + dL_t, \quad X_0 = x. \quad (3.25)$$

If  $L_t$  has exponentially distributed jumps with expectation values  $\frac{1}{\alpha}$  arriving at rate  $\lambda\beta$  the functions  $\phi$  and  $\psi$  are (see Grbac and Papapantoleon [26])

$$\begin{aligned} \phi_t(u) &= -\frac{\lambda\theta}{2\eta^2} \ln \left( 1 - \frac{2\eta^2}{\lambda}(1 - e^{-\lambda t})u \right) \\ &\quad - \frac{\lambda\beta}{\lambda - 2\eta^2\alpha} \ln \left( \frac{\alpha - u}{\alpha - u(e^{-\lambda t} + (1 - e^{-\lambda t})\frac{1}{\lambda}2\eta^2\alpha)} \right), \\ \psi_t(u) &= \frac{e^{-\lambda t}u}{1 - \frac{2\eta^2}{\lambda}(1 - e^{-\lambda t})u}, \end{aligned}$$

where

$$\operatorname{Re}(u) < \min \left\{ \frac{\lambda}{2\eta^2}(1 - e^{-\lambda t})^{-1}, \alpha \left( e^{-\lambda t} + (1 - e^{-\lambda t})\frac{1}{\lambda}2\eta^2\alpha \right)^{-1}, \alpha \right\}.$$

Since  $L_t$  has only positive jumps this process also stays nonnegative. As a third example consider the real-valued affine process defined by

$$dX_t = -\lambda(X_t - \theta) dt + \sigma dW_t + d\tilde{L}_t, \quad X_0 = x, \quad (3.26)$$

where  $\tilde{L}_t$  is a compound Poisson process with positive jumps with mean  $\frac{1}{\alpha^+}$  arriving at rate  $\lambda\beta^+$  and negative jumps with mean  $\frac{1}{\alpha^-}$  arriving at rate  $\lambda\beta^-$ . The functions  $\phi$  and  $\psi$  in this case read (see Müller and Waldenberger [41])

$$\begin{aligned} \phi_t(u) &= \frac{\sigma^2 u^2}{4\lambda}(1 - e^{-2\lambda t}) + \theta u(1 - e^{-\lambda t}) + \frac{\beta^+ + \beta^-}{2} \ln \left( \frac{(\alpha^+ - e^{-\lambda t}u)(\alpha^- + e^{-\lambda t}u)}{(\alpha^+ - u)(\alpha^- + u)} \right) \\ &\quad + \frac{\beta^+ - \beta^-}{2} \ln \left( \frac{(\alpha^+ - e^{-\lambda t}u)(\alpha^- + u)}{(\alpha^+ - u)(\alpha^- + e^{-\lambda t}u)} \right), \\ \psi_t(u) &= e^{-\lambda t}u, \end{aligned}$$

for  $-\alpha^- < \operatorname{Re}(u) < \alpha^+$ .

# Bibliography

- [1] Nabyl Belgrade, Eric Benhamou, and Etienne Koehler. A Market Model for Inflation. *SSRN eLibrary*, 2004.
- [2] Fischer Black. The pricing of commodity contracts. *Journal of Financial Economics*, 3(1-2):167–179, 1976.
- [3] Alan Brace, Dariusz Gatarek, and Marek Musiela. The market model of interest rate dynamics. *Mathematical Finance*, 7(2). ISSN 1467-9965.
- [4] Damiano Brigo and Fabio Mercurio. *Interest Rate Models - Theory and Practice: With Smile, Inflation and Credit (Springer Finance)*. Springer, 2nd edition, 2006.
- [5] Marie-France Bru. Wishart processes. *Journal of Theoretical Probability*, 4(4): 725–751, 1991.
- [6] Björn Böttcher. Feller evolution systems: Generators and approximation. *Stochastics and Dynamics*, 14(03):1350025, 2014.
- [7] A. Di Bucchianico. Banach algebras, logarithms, and polynomials of convolution type. *Journal of Mathematical Analysis and Applications*, 156(1):253 – 273, 1991.
- [8] Patrick Cheridito, Damir Filipović, and Marc Yor. Equivalent and absolutely continuous measure changes for jump-diffusion processes. *Ann. Appl. Probab.*, (3):1713–1732, 08 . doi: 10.1214/105051605000000197.
- [9] John C Cox, Jonathan E Ingersoll, and Stephen A Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.
- [10] Christa Cuchiero and Josef Teichmann. Path properties and regularity of affine processes on general state spaces. In *Séminaire de Probabilités XLV*, Lecture Notes in Mathematics, pages 201–244. Springer International Publishing, 2013.
- [11] Christa Cuchiero, Damir Filipović, Eberhard Mayerhofer, and Josef Teichmann. Affine processes on positive semidefinite matrices. *Ann. Appl. Probab.*, 21(2): 397–463, 04 2011.
- [12] Qiang Dai and Kenneth J. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, 55(5):1943–1978, 2000.

- [13] R. DiMartino and W. Urbina. On Cantor-like sets and Cantor-Lebesgue singular functions. *ArXiv e-prints*, March 2014.
- [14] Darrell Duffie. A yield-factor model of interest rates. *Math. Finance*, pages 379–406, 1996.
- [15] Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.
- [16] Darrell Duffie, Damir Filipović, and Walter Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13:984–1053, 2003.
- [17] Ernst Eberlein and Fehmi Özkan. The Lévy LIBOR model. *Finance and Stochastics*, 9(3):327–348, 2005.
- [18] Ernst Eberlein and Sebastian Raible. Term structure models driven by general Lévy processes. *Mathematical Finance*, 9(1), 1999.
- [19] Ernst Eberlein, Kathrin Glau, and Antonis Papapantoleon. Analysis of fourier transform valuation formulas and applications. *Applied Mathematical Finance*, 17(3):211–240, 2010.
- [20] Neil Falkner and Gerald Teschl. On the substitution rule for lebesgue–stieltjes integrals. *Expositiones Mathematicae*, 30(4):412 – 418, 2012.
- [21] Damir Filipovic. Time-inhomogeneous affine processes. *Stochastic Processes and their Applications*, 115(4):639–659, 2005.
- [22] Matthias Fleckenstein, Francis A. Longstaff, and Hanno N. Lustig. Why Does the Treasury Issue TIPS? The TIPS-Treasury Bond Puzzle. *SSRN eLibrary*, 2010.
- [23] José Da Fonseca, Alessandro Gnoatto, and Martino Grasselli. A flexible matrix LIBOR model with smiles. *Journal of Economic Dynamics and Control*, 37(4): 774–793, 2013.
- [24] II Gihman and AV Skorohod. *The Theory of Stochastic Processes II, Vol. 218 of Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen*. Springer, Berlin, 1975.
- [25] Z. Grbac, A. Papapantoleon, J. Schoenmakers, and D. Skovmand. Affine LIBOR models with multiple curves: theory, examples and calibration. *ArXiv e-prints*, May 2014.
- [26] Zorana Grbac and Antonis Papapantoleon. A tractable libor model with default risk. *Mathematics and Financial Economics*, 7(2):203–227, 2013. ISSN 1862-9679.

- [27] Alberto Guzman. *Derivatives and integrals of multivariable functions*. Springer Science & Business Media, 2012.
- [28] E. Çinlar, J. Jacod, P. Protter, and M.J. Sharpe. Semimartingales and markov processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 54 (2):161–219, 1980.
- [29] Jean Jacod and Albert N Shiryaev. *Limit theorems for stochastic processes, Volume 288 of Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, second edition, 2013.
- [30] Farshid Jamshidian. An exact bond option formula. *The Journal of Finance*, 44 (1):205–209, 1989.
- [31] R. Jarrow and Y. Yildirim. Pricing treasury inflation protected securities and related derivatives using an HJM model. *Journal of Financial and Quantitative Analysis*, 38(2):337–359, 2003.
- [32] Kiyoshi Kawazu and Shinzo Watanabe. Branching processes with immigration and related limit theorems. *Theory of Probability & Its Applications*, 16(1):36–54, 1971.
- [33] Martin Keller-Ressel. *Affine processes - Theory and application in finance*. PhD-thesis, Wien University of Technology, 2008.
- [34] Martin Keller-Ressel and Eberhard Mayerhofer. Exponential moments of affine processes. *Ann. Appl. Probab.*, 25(2):714–752, 04 2015.
- [35] Martin Keller-Ressel, Josef Teichmann, and Walter Schachermayr. Affine processes are regular. *Journal of Probability Theory and Related Fields*, 151(3-4): 591–611, 2011.
- [36] Martin Keller-Ressel, Antonis Papapantoleon, and Josef Teichmann. The affine LIBOR models. *Mathematical Finance*, 23(4):627–658, 2013.
- [37] Martin Keller-Ressel, Walter Schachermayer, and Josef Teichmann. Regularity of affine processes on general state spaces. *Electron. J. Probab.*, 18:no. 43, 1–17, 2013.
- [38] Fabio Mercurio. Pricing inflation-indexed derivatives. *Journal of Quantitative Finance*, 5(3):289–302, 2005.
- [39] Fabio Mercurio and Nicola Moreni. Inflation with a smile. *Risk magazine*, 2006.
- [40] Fabio Mercurio and Nicola Moreni. A multi-factor SABR model for forward inflation rates. *SSRN eLibrary*, 2009.

- [41] Wolfgang Müller and Stefan Waldenberger. Affine LIBOR models driven by real-valued affine processes. submitted, arXiv:1503.00864, 2015.
- [42] Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion (Grundlehren der mathematischen Wissenschaften)*. Springer, 12 2010.
- [43] L. C. G. Rogers and David Williams. *Diffusions, Markov Processes, and Martingales: Volume 1, Foundations (Cambridge Mathematical Library)*. Cambridge University Press, 2 edition, 5 2000.
- [44] Ken-iti Sato. *Lévy Processes and Infinitely Divisible Distributions (Cambridge Studies in Advanced Mathematics)*. Cambridge University Press, 1st edition, 1999.
- [45] Stefan Waldenberger. The affine inflation market models. submitted, arXiv:1503.04979, 2015.
- [46] Alexander D Wentzell. *Theorie zufälliger Prozesse*, volume 65. Springer-Verlag, 2014.
- [47] Roman Werpachowski. Arbitrage-Free Rate Interpolation Scheme for Libor Market Model with Smooth Volatility Term Structure. *SSRN eLibrary*, 2010.