Technische Universität Wien
Vienna University of Technology

Diese Dissertation haben begutachtet:

DISSERTATION

Affine Processes – Theory and Applications in Finance

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften unter der Leitung von

a.o. Univ. Prof. Josef Teichmann,
Institut für Wirtschaftsmathematik (E105)

eingereicht an der Technischen Universität Wien,
Fakultät für Mathematik

von

Dipl.-Ing. Martin Keller-Ressel,
Matrikelnr. 9825054,
Apollogasse 20/46, 1070 Wien.

Wien, am 15. 12. 2008

Deutschsprachige Kurzfassung

Das Thema dieser Dissertation ist die Klasse der *affinen Prozesse*. Von Duffie, Filipovic und Schachermayer [2003] eingeführt, besteht diese Klasse aus allen Markov-Prozessen in stetiger Zeit und mit Wertebereich $\mathbb{R}^m_{\geqslant 0} \times \mathbb{R}^n$, deren logarithmierte charakteristische Funktion auf affine Weise vom anfänglichen Zustandsvektor des Prozesses abhängt. Im ersten Teil der Dissertation, welcher der Theorie der affinen Prozesse gewidmet ist, zeigen wir erstmals, daß jeder (stochastisch stetige) affine Prozess auch ein Feller-Prozess ist. Wir stellen einen alternativen Beweis für das Hauptresultat von Duffie et al. [2003] vor: Unter einer zusätzlichen Regularitätsvoraussetzung kann die Klasse der affinen Prozesse komplett über den infinitesimalen Generator charakterisiert werden, und die charakteristische Funktion des Prozesses erfüllt eine gewöhnliche Differentialgleichung vom verallgemeinerten Riccati-Typ. Anschließend schlagen wir zwei hinreichende Bedingungen für die Regularität eines affinen Prozesses vor. Die zweite dieser Bedingungen definiert die Unterklasse der 'analytischen affinen Prozesse'. Diese Prozesse haben interessante zusätzliche Eigenschaften: Die verallgemeinerten Riccati-Gleichungen können durch analytische Fortsetzung erweitert werden, und beschreiben dann die zeitliche Entwicklung der Momente und Kumulanten des Prozesses. Schließlich zeigen wir, daß Integration in der Zeit, Exponentielle Maßwechsel, sowie Subordination eines unabhängigen Lévy-Prozesses die affine Eigenschaft erhalten.

Im zweiten Teil der Dissertation wenden wir uns den Anwendungen affiner Prozesse zur Modellierung von stochastischer Volatilität zu. Wir definieren die Klasse der 'affinen stochastischen Volatilitätsmodelle' (ASVM), welche eine Vielzahl von stochastischen Volatilitätsmodellen einschließt, die in der Fachliteratur vorgeschlagen wurden: Darunter das Modell von Heston, die Modelle von Bates [1996, 2000] und das Barndorff-Nielsen-Shephard-Modell. Wir leiten Resultate über das Langzeitverhalten des Preis- und des stochastischen Varianzprozesses in ASVMs her, und untersuchen Momentenexplosionen (Momente, welche innerhalb endlicher Zeit unendliche Werte annehmen) des Preisprozesses. Mögliche Anwendungen dieser Resultate, beispielsweise auf die Asymptotik der impliziten Volatilitätsfläche, werden diskutiert. Wir schließen mit einigen expliziten Berechnungen für stochastische Volatilitätsmodelle, welche die Anwendung unserer Ergebnisse erlauben.

Abstract

This thesis is devoted to the study of affine processes. The class of affine processes has been introduced by Duffie, Filipovic, and Schachermayer [2003], and consists of all continuous-time Markov processes taking values in $\mathbb{R}^m_{\geqslant 0} \times \mathbb{R}^n$, whose log-characteristic function depends in an *affine* way on the initial state vector of the process. In the first part of the thesis, which is concerned with the theory of affine processes, we show for the first time that any (stochastically continuous) affine process is also a Feller process. We give an alternative proof for the main result of Duffie et al. [2003] – under an additional regularity condition, the class of affine processes can be completely characterized in terms of the infinitesimal generator, and the characteristic function of the process satisfies an ODE of the generalized Riccati type. Subsequently we introduce two sufficient conditions for regularity. The second of these conditions defines a subclass of affine processes, that we call 'analytic affine'. Not only are these processes automatically regular, but they have other interesting properties: The generalized Riccati equations can be analytically extended to a subset of the real numbers, where they describe the time-evolution of *moments* and *cumulants* of the process. Finally we collect several results on 'elementary transformations' of affine processes: We show that the operations of time-integration, exponential change of measure, subordination of an independent Lévy process, and under suitable condition also projection, preserve the affine property.

In the second part of the thesis we turn towards applications of affine processes to the modelling of stochastic volatility. We define the class of 'affine stochastic volatility models' (ASVMs), which includes a variety of stochastic volatility models that have been proposed in the literature, including the Heston model, the models of Bates [1996, 2000] and the Barndorff-Nielsen-Shephard model. We derive results on the long-term behavior of the price and the stochastic variance process in an ASVM, and study moment explosions (moments becoming infinite in finite time) of the price process. Possible applications of these results are discussed, for example to large-strike asymptotics of the implied volatility surface. We conclude with explicit calculations for several models to which the results apply.

Acknowledgements

First and foremost, I want to thank my advisor, Josef Teichmann. Not only did he draw my interest to mathematical finance and stochastic analysis in the first place, he also encouraged me to pursue a PhD in this area. His research group, funded through the prestigious START price, was a perfect environment for this, with lots of freedom, inspiration, and communication. As an advisor, he was always full of ideas, making connections between different subjects with startling creativity, providing me with high motivation and with deep insights. I also want to thank Walter Schachermayer, who was head of the Institute for Mathematical Methods in Economics, and also of the Research Group for Financial and Actuarial Mathematics (FAM) at TU Wien, during the largest part of my employment, and who played a crucial role in making FAM such an inspiring and internationally respected place for research. Despite of his eminent position and scientific merit, I got to know him as a very approachable person with an unwaning enthusiasm for research. I want to thank my second advisor, Peter Friz, for inviting me to the University of Cambridge and for fuelling my research with suggestions and challenges. My visit has resulted in fruitful scientific collaboration, that I hope will continue in the future. I want to thank my roommates of the first days at FAM, Thomas Steiner and Richard Warnung, with whom I shared some fine hours. Together with Thomas Steiner, I had the fortune to write and publish my first article, and during our time at FAM, Thomas has become a true friend and companion. My thanks also goes to my colleagues of the later days at FAM: Antonis Papapantoleon and Christa Cuchiero, who shared my passion for affine and Lévy processes; to Sara Karlsson, Takahiro Tsuchiya, and to Georg Grafendorfer, who now and then provided me with snacks from his secret drawer. I also want to thank Michael Kupper and Eberhard Mayerhofer from the Vienna Institute of Finance for interesting and inspiring discussions. I want to thank my friends and my parents, who have always supported and encouraged me, and finally I want to thank Lai Cheun, who in the last years has become a part of my life, that I could not bear to miss.

Contents

Introduction

This thesis is devoted to the study of 'affine processes', which are continuous-time Markov processes characterized by the fact that their log-characteristic function depends in an *affine* way on the initial state vector of the process. One of the first articles to discuss such a process was Kawazu and Watanabe [1971], which studies processes arising as continuous-time limit of Galton-Watson branching processes with immigration. The resulting class of processes is called CBI (continuously branching with immigration), and exhibits the mentioned 'affine property'. More recently, affine processes have attracted renewed interest, due to applications in mathematical finance. Based on the realization, that the classical bond pricing models of Vasicek and Cox-Ingersoll-Ross, as well as the stochastic volatility model of Heston, all exhibit the affine property, Duffie, Pan, and Singleton [2000] introduced the class of 'affine jump-diffusions'. This class consists of all jump-diffusion processes, whose drift vector, instantaneous covariance matrix and arrival rate of jumps all depend in an affine way on the state vector. Duffie, Filipovic, and Schachermayer [2003] subsequently extended this class of affine jump-diffusions and combined the strands of research started by Kawazu and Watanabe [1971] and Duffie et al. [2000], defining an *affine process* as a continuous-time Markov process with state space $\mathbb{R}^m_{\geqslant 0} \times \mathbb{R}^n$ and with the mentioned affine property of the log-characteristic function. It turns out that this class coincides for a large part with the class of affine jump-diffusions, but also allows for infinite activity of jumps and for killing or explosions of the process. Duffie, Filipovic, and Schachermayer aimed to give a rigorous mathematical foundation to the theory of affine processes, covering many aspects, such as the characterization of an affine process in terms of the 'admissible parameters' (comparable to the *characteristic triplet* of a Lévy process) and properties of the ordinary differential equations ('generalized Riccati equations') that are implied by the process. To do so, however, they impose a *regularity condition* on the process, which essentially corresponds to the time-differentiability of the characteristic function of the given process.

The attractiveness of affine processes for finance stems from several reasons: First, a variety of models that have been proposed in the literature, and that are used by practitioners, fall into the class of affine models. As mentioned, in the realm

of interest rate models, the classical models of Vasiček [1977] and Cox, Ingersoll, and Ross [1985], as well as multivariate extensions proposed by Dai and Singleton [2000] are all affine; these models have also found applications in intensity-based credit risk modelling (cf. Duffie [2005]). In the area of asset price modelling, the Black-Scholes model, all exponential-Lévy models (cf. Cont and Tankov [2004]), the model of Heston [1993], extensions of the Heston model, such as Bates [1996] and Bates [2000], the model of Barndorff-Nielsen and Shephard [2001] and many time-change models such as Carr and Wu [2004] are based on affine processes.
Second, affine processes exhibit a high degree of analytic tractability. The Kolmogorov PDE can be reduced to a system of ODEs (the generalized Riccati equations) by using a 'basis' of exponential functions. In many cases these generalized Riccati equations allow for explicit solutions. European-style contingent claims can be priced, and sensitivities to risk factors ('Greeks') can be calculated, in a computationally highly efficient way by using Fourier methods (see Carr and Madan [1999]).
Third, the general theory for affine processes that is at hand, allows for an integrated treatment of many different models within one theoretical framework. In particular, from the viewpoint of affine processes, there is no big difference between pure diffusion models, and models with jumps. Let us mention here that affine models allow for quite sophisticated behavior of jumps, going beyond the 'jump-diffusion'[1] paradigm usually encountered in finance: As in general Lévy models, the jumps may arrive with finite, but also with infinite intensity, and even infinite total variation. An affine process can have simultaneous jumps in multiple components (think of volatility and price jumping at the same time in a stochastic volatility model), and the arrival rate of jumps need not be constant: It may depend in an affine way on the state of any (non-negative) component, allowing for cross-excitement and self-excitement effects between factors.

The first part of this thesis, which deals with the theoretical side of affine processes, has been motivated to a great extent by the article of Duffie et al. [2003], that was mentioned above. It was our intention to take up loose ends from this article, and to tackle some of the little questions it has left open: One such 'loose end' was to consider affine processes without imposing the condition of regularity: In the main results of Section 1, we were successful in showing that every affine process is a Feller process, even without assuming regularity. Section 2 then deals with regular affine processes, and contains an alternative proof of the main result of Duffie et al. [2003] – the characterization of an affine process in terms of its

[1] By 'jump-diffusion' we understand a pure diffusion model, to which an independent jump process of finite activity has been added.

infinitesimal generator. Different from Duffie et al. [2003], our proof makes use of existing and well-known results on infinitely divisible distribution, allowing us to simplify some of the arguments. The idea behind Section 3, was to find convenient sufficient conditions for regularity of an affine process. The sufficiency of certain moment conditions has already been observed in Dawson and Li [2006]; we give a more general condition (*'Condition A'*) that allows to mix moment conditions with positivity and space-homogeneity conditions. The rest of Section 3 is concerned with a subclass of affine processes, that we call *'analytic affine'*. These processes are characterized by the fact that their moment generating function exists on a non-vanishing open subset of $\mathbb{R}^d$, and satisfies a uniform boundedness condition. Not only are these processes automatically regular, but they have other interesting properties: The generalized Riccati equations can be analytically extended to a subset of the real numbers, where they describe the time-evolution of *moments* and *cumulants* of the process. This interpretation is of great interest for applications, and is further explored in the second part of the thesis. In Section 4 we collect several results on what we call 'elementary transformations' of affine processes: We show that the operations of projection, time-integration, exponential change of measure, and subordination of an independent Lévy process all preserve the affine property, and can be represented in terms of simple transformations of the characteristics of the process. Similar results (e.g. on exponential measure change for affine processes) have recently been obtained by Kallsen and Muhle-Karbe [2008] using semi-martingale calculus. While the results are similar, our approach to prove them is entirely different and makes use only of 'elementary' methods, such as Markov theory and ODE methods, but no semi-martingale calculus.

In the second part of the thesis we turn towards applications of affine processes to the modelling of stochastic volatility. With minor modifications, this part of the thesis has been successfully submitted to the Journal of Mathematical Finance under the title 'Moment Explosions and Long-Term Behavior of Affine Stochastic Volatility Models'. We define an 'affine stochastic volatility model' as given by a log-price process $(X_t)_{t\geq 0}$ and a stochastic variance process $(V_t)_{t\geq 0}$, such that the joint process $(X_t, V_t)_{t\geq 0}$ is an *affine*. As mentioned, the class of affine processes includes a variety of stochastic volatility models, that have been proposed in the literature: The models of Heston [1993], Bates [1996, 2000] and Barndorff-Nielsen and Shephard [2001] all fall into the scope of our definition[2]. In Section 5 we derive necessary and sufficient conditions for the process to be conservative and for the martingale property of the discounted price process $S_t = \exp(X_t)$. In Section 6 we

[2]Several other stochastic volatility models are also based on affine processes, but may require a state space of more than 2 dimensions to be defined.

derive our central results on long-term properties of an affine stochastic volatility model. These results are formulated as asymptotic results for the cumulant generating function of the stock price, as time goes to infinity. Asymptotics of this type have been used by Lewis [2000] to obtain large-time-to-maturity results for the implied volatility smile of stochastic volatility models via a saddlepoint expansion. We also provide conditions for the existence of an invariant distribution of the stochastic variance process, and characterize this distribution in terms of its cumulant generating function. Both results are obtained by applying qualitative ODE theory to the generalized Riccati equations of the underlying affine process. In Section 7 we study moment explosions (moments becoming infinite in finite time) of the price process, an issue that has recently received much attention, due to the articles of Andersen and Piterbarg [2007] and Lions and Musiela [2007]. Moment explosions are intimately connected to large-strike asymptotics of the implied volatility smile via results of Lee [2004], that have later been expanded by Benaim and Friz [2006]. In Section 8 we briefly discuss these results, and other possible applications to forward-starting options. We conclude in Section 9 with explicit calculations for several models to which the results apply, such as the Heston model, a Heston model with added jumps, a model of Bates, and the Barndorff-Nielsen-Shephard model. Let us remark, that our treatment of affine stochastic volatility models makes no claim to be a complete discussion of such models. Many important aspects such as hedging, market completeness, and the relationship between physical measure and pricing measure are not discussed, but are usually highly non-trivial in jump-based models.

Part 1

Contributions to the theory of affine processes

1. Affine processes

1.1. Definition of an affine process. In simple words, an affine process can be described as a Markov process, whose log-characteristic function is an affine function of its initial state vector. With a view towards applications in finance we will consider only processes defined on the state space $D = \mathbb{R}^m_{\geqslant 0} \times \mathbb{R}^n$. This covers the typical case where economic factors with natural positivity constraints, such as volatility, interest rates or default intensities, are modelled together with factors that are unconstrained. We denote the total dimension of D by $d = m + n$. For convenient notation we define

$$I = \{1, \ldots, m\}, \quad \text{the index set of the } \mathbb{R}_{\geqslant 0}\text{-valued components,}$$
$$J = \{m+1, \ldots, m+n\}, \quad \text{the index set of the } \mathbb{R}\text{-valued components,}$$

and $M = I \cup J = \{1, \ldots, d\}$. If x is a d-dimensional vector, then $x_I = (x_i)_{i \in I}$ denotes its projection on the components with index in I, and similarly for other index sets. Also, if S is a subset of $\mathbb{R}^d$ or $\mathbb{C}^d$, then S_I denotes its projection onto the components given by I. Inequalities involving vectors are interpreted componentwise, i.e. $x \leq 0$ means that $x_i \leq 0$ for all $i \in M$, and $x < 0$ means that $x_i < 0$ for all $i \in M$. As usual $(e_i)_{i \in M}$ denote the unit vectors in $\mathbb{R}^d$. For vectors x, y in $\mathbb{R}^d$ or $\mathbb{C}^d$ we define $\langle x, y \rangle := \sum_{i=1}^d x_i y_i$; note that there is *no conjugation* in the complex case. We will often write

$$f_u(x) := \exp(\langle u, x \rangle)$$

for the exponential function with $u \in \mathbb{C}^d$ and $x \in D$. A special role will be played by the set

$$\mathcal{U} := \left\{ u \in \mathbb{C}^d : \operatorname{Re} u_I \leq 0, \quad \operatorname{Re} u_J = 0 \right\}, \tag{1.1}$$

note that $\mathcal{U}$ is precisely the set of all $u \in \mathbb{C}^d$, for which $x \mapsto f_u(x)$ is a bounded function on D. We also define

$$\mathcal{U}^\circ := \left\{ u \in \mathbb{C}^d : \operatorname{Re} u_I < 0, \quad \operatorname{Re} u_J = 0 \right\}. \tag{1.2}$$

Finally $i\mathbb{R}^d$ denotes the purely imaginary numbers in $\mathbb{C}^d$, i.e. $\left\{ u \in \mathbb{C}^d : \operatorname{Re} u = 0 \right\}$. Note that $i\mathbb{R}^d \subseteq \mathcal{U}$. We are now prepared to give a definition of an affine process:

DEFINITION 1.1 (Affine process). An affine process is a stochastically continuous[3], time-homogeneous Markov process $(X_t, \mathbb{P}^x)_{t \geq 0, x \in D}$ with state space $D = \mathbb{R}^m_{\geqslant 0} \times \mathbb{R}^n$, whose characteristic function is an exponentially-affine function of the state vector. This means that on $i\mathbb{R}^d$ there exist functions $\phi : \mathbb{R}_{\geqslant 0} \times i\mathbb{R}^d \to \mathbb{C}$ and

[3] Note that stochastic continuity is part of our definition, while in Duffie et al. [2003] it is introduced at a later stage, as a property of a *regular affine* process.

$\psi : \mathbb{R}_{\geqslant 0} \times i\mathbb{R}^d \to \mathbb{C}^d$ such that

$$\mathbb{E}^x \left[e^{\langle X_t, u \rangle} \right] = \exp\left(\phi(t,u) + \langle x, \psi(t,u) \rangle\right) \ , \tag{1.3}$$

for all $x \in D$, and for all $(t,u) \in \mathbb{R}_{\geqslant 0} \times i\mathbb{R}^d$.

It is worth to remember here that a process is called stochastically continuous, if for any sequence $t_n \to t$ in $\mathbb{R}_{\geqslant 0}$, the random variables X_{t_n} converge to X_t in probability (with respect to all $(\mathbb{P}^x)_{x \in D}$). Note also, that the existence of a filtered space $(\Omega, \mathcal{F})$, where the process $(X_t)_{t \geq 0}$ is defined, is already implicit in the notion of a *Markov process* (we largely follow Rogers and Williams [1994, Chapter III] in our notation and precise definition of a Markov process). If we do not mention specific assumptions on the filtration, it will be sufficient to take $\mathcal{F}$ as the natural filtration generated by $(X_t)_{t \geq 0}$. Recall also that $\mathbb{P}^x$ represents the law of the Markov process $(X_t)_{t \geq 0}$, *started at* x, i.e. we have that $X_0 = x$, $\mathbb{P}^x$-almost surely. As is well-known we can associate to each (time-homogeneous) Markov process $(X_t)_{t \geq 0}$ a semigroup $(P_t)_{t \geq 0}$ of operators acting on the bounded Borel functions $b\mathcal{B}(D)$, by setting

$$P_t\, f(x) = \mathbb{E}^x \left[f(X_t) \right], \quad \text{for all} \quad x \in D, t \geq 0, f \in b\mathcal{B}(D) \ .$$

It is clear that the left side of (1.3) is defined for all $u \in \mathcal{U}$. We could have made life easier by requiring in Definition 1.1 that (1.3) holds for all $u \in \mathcal{U}$ and not just in $i\mathbb{R}^d$. Though, as we will see, this is not necessary, and the exponentially-affine form of (1.3) extends automatically to a 'large enough' subset of $\mathcal{U}$. Some care has to be taken regarding points in $\mathcal{U}$ where $P_t f_u(x)$ is 0, and thus its logarithm not defined:

LEMMA 1.2. *Let $(X_t)_{t \geq 0}$ be an affine process. Then*

$$\mathcal{O} = \{(t,u) \in \mathbb{R}_{\geqslant 0} \times \mathcal{U} : P_s f_u(0) \neq 0 \ \forall\ s \in [0,t]\} \ , \tag{1.4}$$

is open in $\mathbb{R}_{\geqslant 0} \times \mathcal{U}$ and there exists a unique continuous extension of $\phi(t,u)$ and $\psi(t,u)$ to $\mathcal{O}$, such that (1.3) *holds for all* $(t,u) \in \mathcal{O}$.

For a proof of the Lemma we refer to Duffie et al. [2003, Lemma 3.1]. Note that in Duffie et al. [2003] the Lemma is shown for a *regular* affine process, but the only assumption used in the proof is the stochastic continuity of $(X_t)_{t \geq 0}$. In Section 3 we show the related extension result Lemma 3.12 by a similar proof.

PROPOSITION 1.3. *The functions ϕ and ψ have the following properties:*

(i) ϕ *maps* $\mathcal{O}$ *to* $\mathbb{C}_-$, *where* $\mathbb{C}_- := \{u \in \mathbb{C} : \operatorname{Re} u \leq 0\}$.
(ii) ψ *maps* $\mathcal{O}$ *to* $\mathcal{U}$.
(iii) $\phi(0,u) = 0$ *and* $\psi(0,u) = u$ *for all* $u \in \mathcal{U}$.

*(iv) ϕ and ψ enjoy the '**semi-flow property**':*

$$\begin{aligned}\phi(t+s,u) &= \phi(t,u) + \phi(s,\psi(t,u)),\\ \psi(t+s,u) &= \psi(s,\psi(t,u)),\end{aligned} \tag{1.5}$$

for all $t, s \geq 0$ with $(t+s,u) \in \mathcal{O}$.

(v) ϕ and ψ are jointly continuous on $\mathcal{O}$.

(vi) With the remaining arguments fixed, $u_I \mapsto \phi(t,u)$ and $u_I \mapsto \psi(t,u)$ are analytic functions in $\{u_I : \operatorname{Re} u_I < 0; (t,u) \in \mathcal{O}\}$.

(vii) Let $(t,u),(t,w) \in \mathcal{O}$ with $\operatorname{Re} u \leq \operatorname{Re} w$. Then

$$\begin{aligned}\operatorname{Re}\phi(t,u) &\leq \phi(t,\operatorname{Re} w),\\ \operatorname{Re}\psi(t,u) &\leq \psi(t,\operatorname{Re} w).\end{aligned} \tag{1.6}$$

It is obvious that $\phi(t,.)$ and $\psi(t,.)$ map real numbers to real numbers – hence the inequalities (1.6) make sense. Let us also remark that the condition $(t+s,u) \in \mathcal{O}$ in item (iv) above, guarantees by definition of $\mathcal{O}$ that also $(t,u) \in \mathcal{O}$, and via the Markov property, that $(s,\psi(s,u)) \in \mathcal{O}$.

Proof. Let $(t,u) \in \mathcal{O}$. Since P_t is a contractive semigroup we have $\|P_t f_u\|_\infty \leq \|f_u\|_\infty = 1$. On the other hand

$$P_t f_u(x) = e^{\phi(t,u)} f_{\psi(t,u)}(x)$$

by the affine property (1.3) and Lemma 1.2. Since $\|f_u\|_\infty \leq 1$ if and only if $u \in \mathcal{U}$, we conclude that $\phi(t,u) \in \mathbb{C}_-$ and $\psi(t,u) \in \mathcal{U}$ for all $(t,u) \in \mathcal{O}$ and have shown (i) and (ii). Assertion (iii) follows immediately from $P_0 f_u(x) = f_u(x)$. For (iv) we apply the semi-group property:

$$\begin{aligned}P_{t+s} f_u(x) &= P_t P_s f_u(x) = e^{\phi(s,u)} P_t f_{\psi(s,u)}(x) = \\ &= e^{\phi(s,u)+\phi(t,\psi(s,u))} f_{\psi(t,\psi(s,u))}(x), \quad \text{for all } (t+s,u) \in \mathcal{O}, x \in D\ .\end{aligned}$$

Since also

$$P_{t+s} f_u(x) = e^{\phi(t+s,u)} f_{\psi(t+s,u)}(x)$$

the assertion follows. We show (v): Let $(t_n,u_n) \to (t,u)$ in $\mathcal{O}$. Since $(X_t)_{t\geq 0}$ is stochastically continuous $X_{t_n} \to X_t$ in probability and thus also in distribution. It follows that also $\exp(\langle X_{t_n}, u_n\rangle)$ converges to $\exp(\langle X_t, u\rangle)$ in distribution as $n \to \infty$. By dominated convergence we conclude that

$$P_{t_n} f_{u_n}(x) = \mathbb{E}^x\left[\exp\left(\langle X_{t_n}, u_n\rangle\right)\right] \to \mathbb{E}^x\left[\exp\left(\langle X_t, u\rangle\right)\right] = P_t f_u(x)$$

for all $(t,u) \in \mathcal{O}$ and $x \in D$. It follows that $\phi(t,u)$ and $\psi(t,u)$ are continuous on $\mathcal{O}$. Assertion (vi) follows from analyticity properties of the extended moment generating function $\mathbb{E}^x\left[f_u(X_t)\right]$ (see Proposition B.4). Finally for (vii) note that

$$\left|\mathbb{E}^x\left[e^{\langle u,X_t\rangle}\right]\right| \le \mathbb{E}^x\left[\left|e^{\langle u,X_t\rangle}\right|\right] = \mathbb{E}^x\left[e^{\langle \operatorname{Re} u,X_t\rangle}\right] \le \mathbb{E}^x\left[e^{\langle \operatorname{Re} w,X_t\rangle}\right] ,$$

for all $x \in D$. If (t,u) and (t,w) are in $\mathcal{O}$, we deduce from the affine property (1.3) that

$$\operatorname{Re}\phi(t,u) + \langle x, \operatorname{Re}\psi(t,u)\rangle \le \phi(t,\operatorname{Re} w) + \langle x, \psi(t,\operatorname{Re} w)\rangle \ .$$

Inserting first $x = 0$ and then Ce_i with $C > 0$ arbitrarily large yields assertion (vii). □

1.2. More about the semi-flow property. One of the most interesting properties of ϕ and ψ is – at least in the author's opinion – the semi-flow property (1.5). 'Flows', that is functional equations of the type

$$f(t+s,u) = f(t, f(s,u)) \tag{1.7}$$

have been studied in several contexts, some of them quite abstract. We mention here the areas of differential equations (see e.g. Hartman [1982]), dynamical systems (see e.g. Katok and Hasselblatt [1999]), and the study of topological transformation groups by Montgomery and Zippin [1955].
By (1.5) the function $\psi(t,u)$ of any affine process has the semi-flow property

$$\psi(t+s,u) = \psi(t,\psi(s,u))$$

for all $(t+s,u) \in \mathcal{O}$, and we simply call ψ the **semi-flow of** $(X_t)_{t\ge 0}$. The function $\phi(t,u)$ satisfies the more involved functional equation

$$\phi(t+s,u) = \phi(t,u) + \phi(s,\psi(t,u)) \ .$$

A function of this type is usually called a (additive) **cocycle** of the semi-flow ψ (cf. Katok and Hasselblatt [1999]). It is often convenient to combine the semi-flow ψ and its cocycle ϕ into a 'big semi-flow' $\Upsilon(t,u)$, using the following technique: We extend $\mathcal{O}$ by one dimension and define $\widehat{\mathcal{O}} = \mathcal{O} \times \mathbb{C}$, and similarly $\widehat{\mathcal{U}} = \mathcal{U} \times \mathbb{C}$. The big semi-flow Υ now maps $\widehat{\mathcal{O}}$ to $\widehat{\mathcal{U}}$ and is given by

$$\Upsilon(t,u_1,\dots,u_d,u_{d+1}) = \begin{pmatrix} \psi(t,(u_1,\dots,u_d)) \\ \phi(t,(u_1,\dots,u_d)) + u_{d+1} \end{pmatrix} . \tag{1.8}$$

It is easy to see that Υ satisfies

(i) $\Upsilon(0,u) = u$ for all $u \in \widehat{\mathcal{U}}$,
(ii) $\Upsilon(t+s,u) = \Upsilon(t,\Upsilon(s,u))$ for all $(t+s,u) \in \widehat{\mathcal{O}}$,

and thus again constitutes a semi-flow on $\widehat{\mathcal{O}}$.

Regularity properties of the semi-flow Υ are of great importance in the study of affine processes. In fact the main results of Duffie et al. [2003] that we will present in Section 2 are based on the assumption that $\Upsilon(t,u)$ is differentiable in the time parameter t. As discussed in Montgomery and Zippin [1955], and later generalized to semi-flows by Filipović and Teichmann [2003], flows on topological spaces have the property of transferring regularity from their state variable ('u') to their (semi-)group parameter ('t'). In case of an affine process, where ϕ and ψ have an interpretation as (parts of the) characteristic function of a stochastic process, differentiability in u is equivalent to the existence of moments for $(X_t)_{t\geq 0}$. Consequently the semi-flow of an affine process $(X_t)_{t\geq 0}$ that possesses bounded moments should exhibit some regularity also in the time parameter t. This is precisely the idea that we will pursue in Section 3.

We give now some first examples of affine processes and illustrate the interplay between the semi-flow $\Upsilon(t,u)$ and the process $(X_t)_{t\geq 0}$:

Example 1.4 (Lévy Process). Suppose that $(X_t)_{t\geq 0}$ is a conservative affine process with stationary semi-flow ψ, i.e. $\psi(t,u) = u$ for all $(t,u) \in \mathcal{O}$, and thus in particular for all $(t,u) \in \mathbb{R}_{\geqslant 0} \times i\mathbb{R}^d$. Then the functional equation for the cocycle ϕ becomes

$$\phi(t+s,u) = \phi(t,u) + \phi(s,u), \quad t,s \in \mathbb{R}_{\geqslant 0}, u \in i\mathbb{R}^d .$$

This is Cauchy's first functional equation. Since ϕ is continuous and satisfies $\phi(0,u) = 0$, it is a linear function of t, i.e. of the form $\phi(t,u) = tm(u)$. On the other hand we have that

$$\mathbb{E}^0\left[e^{\langle X_t,u\rangle}\right] = e^{tm(u)}$$

such that $e^{tm(u)}$ is a characteristic function for every $t > 0$. We conclude that it is an infinitely divisible characteristic function (cf. Section C), and thus that $m(u)$ has to be of Lévy-Khintchine form. It follows that $(X_t)_{t\geq 0}$ is a Lévy process.

Example 1.5 (Ornstein-Uhlenbeck-type process). Let $(X_t)_{t\geq 0}$ be a conservative affine process on $D = \mathbb{R}$. Then, as we show in Proposition 1.9, $\psi(t,u)$ is necessarily of the form $e^{t\beta}u$ for some $\beta \in \mathbb{R}$. Consider now the Ornstein-Uhlenbeck-type process, which is defined by Sato [1999] as the unique solution of the SDE

$$dX_t = \beta X_t\, dt + dL_t, \qquad X_0 = x \in \mathbb{R} .$$

where L_t is a Lévy process with characteristic exponent $\kappa(u)$ and $\beta \in \mathbb{R}$. It can be shown that the characteristic function of $(X_t)_{t\geq 0}$ is given by

$$\mathbb{E}^x\left[e^{\langle X_t,u\rangle}\right] = \exp\left(\int_0^t \kappa\left(e^{s\beta}u\right)\, ds + \left\langle x, e^{t\beta}u\right\rangle\right) .$$

It is clear that $(X_t)_{t\geq 0}$ is an affine process with

$$\phi(t,u) = \int_0^t \kappa\left(e^{s\beta}u\right)\, ds, \qquad \psi(t,u) = e^{t\beta}u\ , \tag{1.9}$$

and thus of the type described above. Can we conclude that every affine process on $D = \mathbb{R}$ is an OU-type process? No, because it remains to show that $\phi(t,u)$ is necessarily of the form (1.9). Even though it looks similar, this problem seems to be harder than the characterization of a Lévy process in Example 1.4, and is to our knowledge still open.

The next example describes an affine process with an even more interesting and beautiful semi-flow structure:

Example 1.6 (Squared Bessel process). Consider the SDE

$$dZ_t = 2\sqrt{Z_t}\, dW_t + \delta \quad Z_0 = z \geq 0\ .$$

By Revuz and Yor [1999], there exists a unique solution, which is non-negative and has the (extended) moment generating function

$$\mathbb{E}^z\left[e^{uZ_t}\right] = \exp\left(\frac{\delta}{2}\log(1-2ut) + z\frac{u}{1-2ut}\right)\ ,$$

defined for all $u \in \mathbb{C}$ with $\operatorname{Re} u < \frac{1}{2t}$. The process $(Z_t)_{t\geq 0}$ is called squared Bessel process of dimension δ; from its moment generating function we see that it is an affine process on $D = \mathbb{R}_{\geqslant 0}$ with

$$\phi(t,u) = \frac{\delta}{2}\log(1-2ut), \qquad \psi(t,u) = \frac{u}{1-2ut}\ .$$

For every $t \geq 0$, $\psi(t,u)$ is a Möbius transformation, i.e. a bijective conformal map of the (extended) complex plane to itself. It is easily derived that $u \mapsto \psi(t,u)$ has the single fixed point 0, and that the 'left half plane' $\mathcal{U}$ is mapped to the interior of a circle, passing through 0 and $-\frac{1}{2t}$, and which is symmetric with respect to the real axis; see Figure 1 for an illustration.

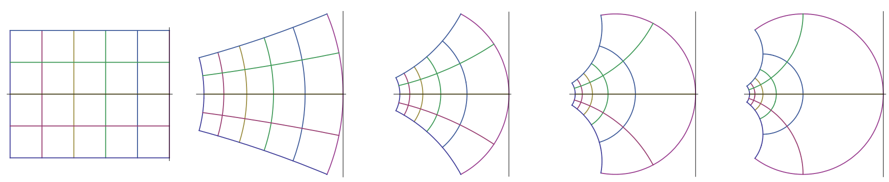

Figure 1. Illustration of the semi-flow $\psi(t,u)$ of a Squared Bessel process. Plots correspond to $t = 0, 0.5, 1, 1.5, 2$ from left to right.

1.3. The Feller property. We proceed to show that every affine process is a Feller process. Along the way, we obtain some additional results on the properties of the semi-flow ψ. In Duffie et al. [2003] the Feller property is shown under the condition that $(X_t)_{t\geq 0}$ is a *regular*[4] affine process; we give a proof that does not use any regularity assumption. We start with a simple Lemma on positive definite functions:

LEMMA 1.7. *Let Φ be a positive definite function on $\mathbb{R}^d$ with $\Phi(0) = 1$ Then*

$$|\Phi(y+z) - \Phi(y)\Phi(z)|^2 \leq \left(1 - |\Phi(y)|^2\right)\left(1 - |\Phi(z)|^2\right) \leq 1$$

for all $y, z \in \mathbb{R}^d$.

PROOF. The result follows from considering the matrix

$$M_\Phi(y,z) := \begin{pmatrix} \Phi(0) & \overline{\Phi(y)} & \Phi(z) \\ \Phi(y) & \Phi(0) & \Phi(y+z) \\ \overline{\Phi(z)} & \overline{\Phi(y+z)} & \Phi(0) \end{pmatrix}, \qquad y, z \in \mathbb{R}^d, y \neq z \;,$$

which is positive semi-definite by definition of Φ. The inequality is then derived from the fact that $\det M_\Phi(y,z) \geq 0$. See Jacob [2001, Lemma 3.5.10] for details □

The next Lemma might seem unwieldy, but it is an elaboration of the following idea: If, for arbitrarily small t, $\psi(t, .)$ maps $i\mathbb{R}^d$ to $i\mathbb{R}^d$, then $\psi(t,u)$ must be a *linear* function of u.

LEMMA 1.8. *Let $K \subseteq M$, $k \in M$, and let $(t_n)_{n\in\mathbb{N}}$ be a sequence such that $t_n \downarrow 0$. Define $\Omega_K := \left\{y \in \mathbb{R}^d : y_{M\setminus K} = 0\right\}$, and suppose that*

$$\operatorname{Re}\psi_k(t_n, iy) = 0 \qquad \textit{for all } y \in \Omega_K \textit{ and } n \in \mathbb{N}\;.$$

Then there exists $\zeta(t_n) \in \mathbb{R}^{|K|}$ and an increasing sequence of positive numbers R_n such that $R_n \uparrow \infty$ and

$$\psi_k(t_n, iy) = \langle \zeta(t_n), iy_K \rangle \;,$$

for all $y \in \Omega_K$ with $|y| < R_n$.

PROOF. As the characteristic function of the (possibly defective) random variable X_{t_n} under $\mathbb{P}^x$, the function $y \mapsto P_{t_n} f_{iy}(x)$ is positive definite for any $x \in D, n \in \mathbb{N}$. We define now for every $y \in \Omega_K$, $c > 0$, and $n \in \mathbb{N}$, the function

$$\Phi(y; n, c) := e^{-\phi(t_n,0)} P_{t_n} f_{iy}(c \cdot e_k) = \exp\Big(\phi(t_n, y) - \phi(t_n, 0) + c \cdot \psi_k(t_n, y)\Big)\;.$$

Clearly, as a function of $y \in \Omega_K$, also $\Phi(y; n, c)$ is positive definite. In addition it satisfies $\Phi(0; n, c) = \exp(c \cdot \psi_k(t_n, 0)) = 1$, since $\psi_k(t_n, 0) = 0$ by assumption.

[4]The notion of a regular affine process is discussed in detail in Section 2.

Thus we may apply Lemma 1.7 to Φ, and it holds for any $y,z \in \Omega_K$, $c > 0$ and $n \in \mathbb{N}$ that

$$|\Phi(y+z;n,c) - \Phi(y;n,c)\cdot\Phi(z;n,c)|^2 \le 1 \ . \tag{1.10}$$

For compact notation we define the abbreviations

$$\begin{aligned} C(y,z,t) :=& \mathrm{Re}\big(\phi(t,i(y+z)) + \phi(t,iy) + \phi(t,iz)\big) - 3\phi(t,0) \\ \Gamma(y,z,t,c) :=& \mathrm{Im}\big(\phi(t,i(y+z)) - \phi(t,iy) - \phi(t,iz)\big) + \\ &+ c\cdot \mathrm{Im}\left(\psi_k(t,i(y+z)) - \psi_k(t,iy) - \psi_k(t,iz)\right) \ . \end{aligned}$$

Note that (at least along the sequence t_n) $C(y,z,t)$ does not depend on c – this is where the assumption $\mathrm{Re}\,\psi_k(t_n,iy) = 0$ enters. For arbitrary real numbers a,b,α,β it holds that

$$2e^{a+b}\left(1-\cos(\alpha-\beta)\right) \le e^{2a} + e^{2b} - 2e^{a+b}\cos(\alpha-\beta) = \left|e^{a+i\alpha} - e^{b+i\beta}\right|^2 \ ,$$

which lets us rewrite inequality (1.10) as

$$e^{C(y,z,t_n)}\left(1-\cos\Gamma(y,z,t_n,c)\right) \le \frac{1}{2} \ . \tag{1.11}$$

Define now for each $n \in \mathbb{N}$

$$R_n := \sup\left\{r \ge 0 : e^{C(y,z,t_n)} > \frac{1}{2} \text{ for all } y,z \in \Omega_K \text{ with } |y| \le r, |z| \le r\right\} \ .$$

First note that $R_n > 0$: This follows from the fact that $e^{C(0,0,t)} = 1$ and $C(y,z,t)$ is continuous. Second, it holds that $R_n \uparrow \infty$: Use that $e^{C(y,z,0)} = 1$ for all $y,z \in \Omega_K$, and the continuity of $C(y,z,t)$.
Suppose that

$$\psi_k(t_n,i(y+z)) - \psi_k(t_n,iy) - \psi_k(t_n,iz) \neq 0$$

for any $n \in \mathbb{N}$ and $y,z \in \Omega_K$ with $|y| < R_n$, $|z| < R_n$. Then by definition of $\Gamma(y,z,t_n,c)$ there exists an $c > 0$ such that $\cos\Gamma(y,z,t_n,c) = -1$. Inserting into (1.11) we obtain

$$\frac{1}{2}\cdot 2 < e^{C(y,z,t_n)}\left(1-\cos\Gamma(y,z,t_n,c)\right) \le \frac{1}{2} \ ,$$

a contradiction. We conclude that

$$\psi_k(t_n,i(y+z)) - \psi_k(t_n,iy) - \psi_k(t_n,iz) = 0 \ ,$$

for all $y,z \in \Omega_K$ with $|y| < R_n$, $|z| < R_n$. Since $\psi(t,.)$ is continuous, the first Cauchy functional equation implies that ψ_k is a linear function of y_K, i.e. there exists some vector $\zeta(t)$, such that

$$\psi_k(t_n,iy) = \langle\zeta(t_n), iy_K\rangle \ . \tag{1.12}$$

for all $y \in \Omega_K$ with $|y| < R_n$. Since $\operatorname{Re}\psi(t_n, iy) = 0$ it is clear that $\zeta(t)$ is real-valued. □

Using the above Lemma, we show two propositions, that will be instrumental in proving the Feller property of an affine process.

PROPOSITION 1.9. *Let $(X_t)_{t\geq 0}$ be an affine process on $D = \mathbb{R}^m_{\geqslant 0} \times \mathbb{R}^n$ and denote by J its real-valued components. Then there exists a real $n \times n$-matrix β such that $\psi_J(t,u) = e^{t\beta} u_J$ for all $(t,u) \in \mathcal{O}$.*

PROOF. Consider the definition of $\mathcal{U}$ in (1.1). Since $\psi(t,u)$ takes by Proposition 1.3 values in $\mathcal{U}$ it is clear that $\operatorname{Re}\psi_J(t,iy) = 0$ for *any* $(t,y) \in \mathbb{R}_{\geqslant 0} \times \mathbb{R}^d$. Fix some $t_* > 0$ and define $t_n := t_*/n$ for all $n \in \mathbb{N}$. We can apply Lemma 1.8 with $K = M = \{1,\dots,d\}$ and any choice of $k \in J$, to obtain a sequence $R_n \uparrow \infty$, such that

$$\psi_J(t_n, iy) = \zeta(t_n) \cdot iy \;, \tag{1.13}$$

for all $y \in \mathbb{R}^d$ with $|y| < R_n$. Note that $\zeta(.)$ now denotes a real $n \times d$-matrix.
Let $i \in I$, $n \in \mathbb{N}$, and consider the function

$$h_n : \Omega_n := \{\omega \in \mathbb{C} : -R_n < \operatorname{Re}\omega \leq 0\} \to \mathbb{C}^n : \quad \omega \mapsto \psi_J(t_n, \omega e_i) - \zeta(t) \cdot \omega e_i \;.$$

By Prop B.4 this is an analytic function on Ω_n° and continuous on Ω_n. According to the Schwarz reflection principle h_n can be extended to an analytic function on an open superset of Ω_n. But (1.13) implies that the function h_n takes the value 0 on a subset with an accumulation point in $\mathbb{C}$. We conclude that h_n is zero everywhere. In particular we have that

$$0 = \operatorname{Re}\psi_J(t_n, \omega e_i) - \zeta(t_n) \cdot \operatorname{Re}\omega e_i = \zeta(t_n) \cdot \operatorname{Re}\omega e_i \;,$$

for all $\omega \in \Omega$. This can only hold true, if the i-th column of $\zeta(t_n)$ is zero. Since $i \in I$ arbitrary we have reduced (1.13) to

$$\psi_J(t_n, u) = \widetilde{\zeta}(t_n) \cdot u_J \;, \tag{1.14}$$

for all $(t_n, u) \in \mathcal{O}$, such that $|u_J| < R_n$. Here $\widetilde{\zeta}(t_n)$ denotes the $n \times n$-submatrix of $\zeta(t_n)$ that results from dropping the zero-columns.
Fix an arbitrary $u_* \in \mathcal{U}$ with $(t_*, u_*) \in \mathcal{O}$ and let $R := \sup\{|\psi_K(t, u_*)| : t \in [0, t_*]\}$. Since $\psi(t,u)$ is continuous, R is finite. Choose N such that $R_n > R$ for all $n \geq N$. Using the semi-flow equation we can write $\psi_J(t_*, u_*)$ as

$$\begin{aligned}\psi_J(t_*, u_*) = \psi_J\left(t_n, \psi(t_* \tfrac{n-1}{n}, u_*)\right) = \\ = \widetilde{\zeta}(t_n) \cdot \psi_J(t_* \tfrac{n-1}{n}, u_*) = \cdots = \widetilde{\zeta}(t_n)^n \cdot u_* \;;\end{aligned} \tag{1.15}$$

for any $n \geq N$. Thus, the functional equation $\psi(t,u) = \widetilde{\zeta}(t) \cdot u_J$ actually holds for all $(t,u) \in \mathcal{O}$. Another application of the semi-flow property yields then, that

$$\widetilde{\zeta}(t+s) = \widetilde{\zeta}(t)\widetilde{\zeta}(s), \qquad \text{for all } t, s \geq 0 \ .$$

Since $\widetilde{\zeta}(0) = 1$, $\widetilde{\zeta}$ is continuous and satisfies the second Cauchy functional equation, it follows that $\widetilde{\zeta}(t) = e^{\beta t}$ for some real $n \times n$-matrix β. □

The following Proposition will be crucial in proving the Feller property of $(X_t)_{t \geq 0}$. It shows that the semi-flow $\psi(t,u)$ maps the interior of $\mathcal{U}$ to the interior:

Proposition 1.10. *Suppose that* $(t,u) \in \mathcal{O}$. *If* $u \in \mathcal{U}^\circ$, *then* $\psi(t,u) \in \mathcal{U}^\circ$.

Proof. For a contradiction, assume there exists $(t,u) \in \mathcal{O}$ such that $u \in \mathcal{U}^\circ$, but $\psi(t,u) \notin \mathcal{U}^\circ$. This implies that there exists $k \in I$, such that $\operatorname{Re} \psi_k(t,u) = 0$. Let $\mathcal{O}_{t,k} = \{\omega \in \mathbb{C} : \operatorname{Re} \omega \leq 0, (t, \omega e_k) \in \mathcal{O}\}$. From the inequalities Proposition 1.3.(vii) we deduce that

$$0 = \operatorname{Re} \psi_k(t,u) \leq \psi_k(t, \operatorname{Re} \omega \cdot e_k) \leq 0 \ , \tag{1.16}$$

and thus that $\psi_k(t, \operatorname{Re} \omega \cdot e_k) = 0$ for all $\omega \in \mathcal{O}_{t,k}$, such that $\operatorname{Re} u_k \leq \operatorname{Re} \omega$. By Proposition 1.3.(vi), $\psi_k(t, \omega e_k)$ is an analytic function of ω. Since it takes the value zero on a set with accumulation point, it is zero everywhere, i.e. $\psi_k(t, \omega e_k) = 0$ for all $\omega \in \mathcal{O}_{t,k}$. We now show that the same statement holds with t replaced by $t/2$: Set $\lambda := \operatorname{Re} \psi_k(t/2, u)$. If $\lambda = 0$, we can proceed exactly as above, only with $t/2$ instead of t. If $\lambda < 0$, then we have, by another application of Proposition 1.3.(vii), that

$$0 = \operatorname{Re} \psi_k(t,u) = \operatorname{Re} \psi_k(t/2, \psi(t/2,u)) \leq \psi_k(t/2, \lambda e_k) \leq \psi_k(t/2, \operatorname{Re} \omega e_k) \leq 0 \ , \tag{1.17}$$

for all $\omega \in \mathcal{O}_{t/2,k}$ such that $\lambda \leq \operatorname{Re} \omega$. Again we can use that an analytic function that takes the value zero on a set with accumulation point, is zero everywhere, and obtain that $\psi_k(t/2, \omega e_k) = 0$ for all $\omega \in \mathcal{O}_{t/2,k}$. Repeating this argument, we finally obtain a sequence $t_n \downarrow 0$, such that

$$\operatorname{Re} \psi_k(t_n, \omega e_k) = 0 \qquad \text{for all } \omega \in \mathcal{O}_{t_n,k}.$$

We can now apply Lemma 1.8 with $K = \{k\}$, which implies that ψ_k is of the linear form

$$\psi_k(t_n, \omega e_k) = \zeta_k(t_n) \cdot \omega, \quad \text{for all} \quad \omega \in \mathcal{O}_{t_n,k} \text{ with } |\omega| \leq R_n,$$

where $\zeta_k(t_n)$ are real numbers, and $R_n \uparrow \infty$. Note that since $\zeta_k(t_n) \to 1$ as $t_n \to 0$, we have that $\zeta_k(t_n) > 0$ for n large enough. Choosing now some ω_* with $\operatorname{Re} \omega_* < 0$

it follows that $\operatorname{Re}\psi_k(t_n, \omega_* e_k) < 0$ – with strict inequality. This is a contradiction to (1.17), and the assertion is shown. □

THEOREM 1.11. *An affine process is a Feller process.*

PROOF. By stochastic continuity of $(X_t)_{t\geq 0}$ and dominated convergence, it follows immediately that $P_t f(x) = \mathbb{E}^x [f(X_t)] \to f(x)$ as $t \to 0$ for all $f \in C_0(D)$ and $x \in D$. To prove the Feller property of $(X_t)_{t\geq 0}$ it remains to show that $P_t C_0(D) \subseteq C_0(D)$:
Consider the following set of functions:

$$\Theta := \left\{ h_{(u_I,g)}(x) = e^{\langle u_I, x_I\rangle} \int_{\mathbb{R}^n} f_{iz}(x_J) g(z)\, dz : u_I \in \mathcal{U}_I^\circ, g \in C_c^\infty(\mathbb{R}^n) \right\} , \tag{1.18}$$

and denote by $\mathcal{L}(\Theta)$ the set of (complex) linear combinations of functions in Θ. From the Riemann-Lebesgue-Lemma it follows that $\int_{\mathbb{R}^n} f_{iz}(x_J) g(z)\, dz$ vanishes at infinity, and thus that $\mathcal{L}(\Theta) \subset C_0(D)$. It is easy to see that $\mathcal{L}(\Theta)$ is a subalgebra of $C_0(D)$, that is in addition closed under complex conjugation. It is easy to check that it is also point separating and vanishes nowhere (i.e. there is no $x_0 \in D$ such that $h(x_0) = 0$ for all $h \in \mathcal{L}(\Theta)$). Using a suitable version of the Stone-Weierstrass theorem (e.g. Semadeni [1971, Corollary 7.3.9]), it follows that $\mathcal{L}(\Theta)$ is dense in $C_0(D)$.
Fix some $t \in \mathbb{R}_{\geqslant 0}$ and let $h(x) \in \Theta$. By Proposition 1.9 we know that $\psi_J(t,u) = e^{\beta t} u_J$. By Lemma 1.2 it holds that $\mathbb{E}^x [f_u(X_t)] = \exp(\phi(t,u) + \langle x, \psi(t,u)\rangle)$ whenever $(t,u) \in \mathcal{O}$, and $\mathbb{E}^x[f_{(u_I,iz)}(X_t)] = 0$ whenever $(t,u) \notin \mathcal{O}$. Thus we have

(1.19)

$$\begin{aligned}
P_t h(x) &= \mathbb{E}^x \left[\int_{\mathbb{R}^n} f_{(u_I,iz)}(X_t) g(z)\, dz \right] = \int_{\mathbb{R}^n} \mathbb{E}^x [f_{(u_I,iz)}(X_t)] g(z)\, dz = \\
&= \int_{\{u \in \mathcal{U} : (t,u) \in \mathcal{O}\}} P_t f_{(u_I,iz)}(x) \cdot g(z)\, dz = \\
&= \int_{\{u \in \mathcal{U} : (t,u) \in \mathcal{O}\}} \exp\left(\phi(t,u_I,iz) + \langle x_I, \psi_I(t,u_I,iz)\rangle + \left\langle x_J, e^{t\beta} iz\right\rangle\right) g(z)\, dz \,.
\end{aligned}$$

Since $(u_I, iz) \in \mathcal{U}^\circ$ it follows by Proposition 1.10 that also $\operatorname{Re}\psi_I(t,u_I,iz) < 0$ for any $z \in \mathbb{R}^n$. This shows that $P_t h(x) \to 0$ as $|x_I| \to \infty$. In addition (1.19), as a function of x_J (1.19) can be interpreted as the Fourier transformation of a compactly supported density. The Riemann-Lebesgue-Lemma then implies that $P_t h(x) \to 0$ as $|x_J| \to \infty$, and we conclude that $P_t h \in C_0(D)$. The assertion extends by linearity to every $h \in \mathcal{L}(\Theta)$, and finally by the density of $\mathcal{L}(\Theta)$ to every $h \in C_0(D)$. This proves that the semi-group $(P_t)_{t\geq 0}$ maps $C_0(D)$ into $C_0(D)$, and hence that $(X_t)_{t\geq 0}$ is a Feller process. □

The following Corollary follows from standard results on Feller processes (see Kallenberg [1997]).

Corollary 1.12. *Any affine process* $(X_t)_{t\geq 0}$ *has a cadlag version*[5] *on* $D\cup\{\Delta\}$, *where* Δ *is an absorbing state for* X_t. *If it is conservative it has a cadlag version on* D. $(X_t)_{t\geq 0}$ *has the strong Markov property, and is characterized by its generator* $\mathcal{A}$, *a closed operator defined on a dense subset of* $C_0(D)$.

[5]In this context, Y_t is a version of X_t, if $\mathbb{P}^x[X_t = Y_t] = 1$ for all $t \geq 0$ *and for all* $x \in D$.

2. Regular Affine processes

Following Duffie et al. [2003] we impose a regularity assumption on the affine process. Under this assumption we are able to fully characterize the process in terms of its infinitesimal generator, and to obtain other crucial results.

Definition 2.1 (Regularity). An affine process is called *regular*, if the derivatives

$$F(u) := \frac{\partial \phi}{\partial t}(t,u)\Big|_{t=0+}, \qquad R(u) := \frac{\partial \psi}{\partial t}(t,u)\Big|_{t=0+} \tag{2.1}$$

exist for all $u \in \mathcal{U}$, and are continuous at $u = 0$.

Remark 2.2. Since $F(u)$ and $R(u)$, as we will see, completely characterize the process $(X_t)_{t\geq 0}$, we shall also call them **functional characteristics** of $(X_t)_{t\geq 0}$.

Note that like ϕ, the function F is scalar-valued (mapping $\mathcal{U}$ to $\mathbb{C}$), and like ψ, R is vector valued (mapping $\mathcal{U}$ to $\mathbb{C}^d$). If $(X_t)_{t\geq 0}$ is a regular affine process, we can differentiate the semi-flow equations (1.5) with respect to s and evaluate at $s = 0$, to obtain the following differential equations for ϕ and ψ, valid for $(t,u) \in \mathcal{O}$:

$$\frac{\partial}{\partial t}\phi(t,u) = F(\psi(t,u)), \quad \phi(0,u) = 0 \tag{2.2a}$$

$$\frac{\partial}{\partial t}\psi(t,u) = R(\psi(t,u)), \quad \psi(0,u) = u\ . \tag{2.2b}$$

For reasons that will become apparent later, these ODEs are called *generalized Riccati equations*. They are autonomous equations, and the variable u enters as an initial condition.

Let us remark the following detail on the derivation of the generalized Riccati equations: F and R are defined (only) as the *right-sided* derivative of ϕ and ψ, such that we should also have right-sided derivatives in (2.2). However, as we will see later, F and R are continuous functions, not just at $u = 0$, but for all $u \in \mathcal{U}$. It follows by Proposition 1.3 that the right hand sides of (2.2) are continuous functions of t. But a function with a one-sided derivative that is continuous, is continuously differentiable in the ordinary (both-sided) sense (see Yosida [1995, Section IX.3]).

The main goal that we pursue now, is to show that F and R are of a specific form, namely that they are – in case of R component-by-component – log-characteristic functions of sub-stochastic infinitely divisible measures[6], satisfying some additional *admissibility conditions*. As is well-known from the theory of Lévy-processes, the characteristic function of an infinitely divisible probability measure can be described by three parameters, the so-called Lévy-triplet (a, b, m), where a is

[6]See Appendix for an explanation of the terminology.

a positive definite matrix (also called diffusion matrix), b a vector (also called drift vector) and a (σ-finite Borel) measure $m(d\xi)$, satisfying the integrability condition $\int(1\wedge|\xi|^2)\,m(d\xi)<\infty$, which is called a Lévy measure. In the case of *sub-stochastic* infinitely divisible measures, a fourth parameter $c\in\mathbb{R}_{\geqslant 0}$ is added: It corresponds to the 'defect' of the measure μ, and is given by $c=-\log\mu(D)$ (i.e. for a probability measure $c=0$).

Since F is scalar-valued, and R d-dimensional vector-valued we end up with $(d+1)\times 4$ parameters that characterize F and R. We group them into the Lévy-quadruplet (a,b,c,m) describing F and the quadruplets $(\alpha_i,\beta_i,\gamma_i,\mu_i)_{i\in\{1,\dots,d\}}$ describing the components $R_i(u)$ respectively. As mentioned above, the parameters necessarily satisfy certain admissibility conditions, which are summarized in the following definition:

Definition 2.3 (Admissibility Conditions). A **parameter set** for an affine process is given by positive semi-definite real $d\times d$-matrices $a,\alpha^1,\dots,\alpha^d$; by $\mathbb{R}^d$-valued vectors $b,\beta^1,\dots,\beta^d$; by non-negative numbers $c,\gamma^1,\dots,\gamma^d$ and by Lévy measures $m,\mu^1,\dots\mu^d$ on $\mathbb{R}^d$.

Such a parameter set is called **admissible** for an affine process with state space D if

$$a_{kl}=0 \quad \text{if } k\in I \text{ or } l\in I\ , \tag{2.3a}$$

$$\alpha^j=0 \quad \text{for all } j\in J\ , \tag{2.3b}$$

$$\alpha^i_{kl}=0 \quad \text{if } k\in I\setminus\{i\} \text{ or } l\in I\setminus\{i\}\ , \tag{2.3c}$$

$$b\in D \tag{2.3d}$$

$$\beta^i_k\geq 0 \quad \text{for all } i\in I \text{ and } k\in I\setminus\{i\}\ , \tag{2.3e}$$

$$\beta^j_k=0 \quad \text{for all } j\in J \text{ and } k\in I\ , \tag{2.3f}$$

$$\gamma^j=0 \quad \text{for all } j\in J\ , \tag{2.3g}$$

$$\operatorname{supp} m\subseteq D \quad \text{and} \quad \int_{D\setminus\{0\}}\left\{\left(|x_I|+|x_J|^2\right)\wedge 1\right\}m(dx)<\infty \tag{2.3h}$$

$$\mu^j=0 \quad \text{for all } j\in J\ , \tag{2.3i}$$

$$\operatorname{supp}\mu^i\subseteq D \quad \text{for all } i\in I, \quad \text{and} \tag{2.3j}$$

$$\int_{D\setminus\{0\}}\left\{\left(|x_{I\setminus\{i\}}|+|x_{J\cup\{i\}}|^2\right)\wedge 1\right\}\mu_i(dx)<\infty \quad \text{for all } i\in I. \tag{2.3k}$$

The admissibility conditions certainly look unwelcoming at first sight. Note however that if $J=\emptyset$ or $I=\emptyset$ – i.e. if we have state space $\mathbb{R}^d_{\geqslant 0}$ or $\mathbb{R}^d$ – the conditions simplify considerably. Even in the case that $|I|=1$ some conditions are trivially valid, because $I\setminus\{i\}=\emptyset$. The admissibility conditions for the matrices a,α_i and the vectors b,β_i are visualized in Table 1. Note that in particular the possible structure of $R_J(u)$ is strongly constrained by the admissibility conditions: It can only take the form $R_J(u) = \beta_J u_J$, where β_J is the $n \times n$ matrix consisting of the elements $(\beta^j_k)_{j,k\in J}$.

In addition to the admissible parameters we will need the following definition:

DEFINITION 2.4 (Truncation functions). Define functions $h, \chi^1, \ldots, \chi^m$ from $\mathbb{R}^d \to [-1,1]^d$ coordinate-wise by

$$h_k(\xi) := \begin{cases} 0 & k \in I \\ \frac{\xi_k}{1+\xi_k^2} & k \in J \end{cases} \qquad \text{for all } \xi \in \mathbb{R}^d\ , \tag{2.4}$$

and

$$\chi^i_k(\xi) := \begin{cases} 0 & k \in I \setminus \{i\} \\ \frac{\xi_k}{1+\xi_k^2} & k \in J \cup \{i\} \end{cases} \qquad \text{for all } \xi \in \mathbb{R}^d, i \in I\ . \tag{2.5}$$

REMARK 2.5. The term $\frac{\xi_k}{1+\xi_k^2}$ can be replaced by $\omega(\xi_k)$, where $\omega(\xi_k)$ is any bounded continuous function from $\mathbb{R}$ to $\mathbb{R}$, that behaves like ξ_k in a neighborhood of 0. It will also be seen that the definition of h and χ^i is directly related to the integrability properties (2.3h) and (2.3k) of the Lévy measures m and μ^i.

We are now prepared to state the main result of this section, the characterization of an affine process in terms of admissible parameters. The result can also be found in Duffie et al. [2003] as Theorem 2.7.

THEOREM 2.6 (Generator of an affine process). *Let $(X_t)_{t\geq 0}$ be a regular affine process with state space D. Then there exist a set of admissible parameters $(a, \alpha^i, b, \beta^i, c, \gamma^i, m, \mu^i)_{i\in\{1,\ldots d\}}$ such that*

(a) the functions F and R defined in (2.1) are of the Lévy-Khintchine form

(2.6)

$$F(u) = \frac{1}{2}\langle u, au\rangle + \langle b, u\rangle - c + \int_{\mathbb{R}^d\setminus\{0\}} \left(e^{\langle \xi, u\rangle} - 1 - \langle h(\xi), u\rangle\right) m(d\xi)\ ,$$

$$R_i(u) = \frac{1}{2}\langle u, \alpha^i u\rangle + \langle \beta^i, u\rangle - \gamma^i + \int_{\mathbb{R}^d\setminus\{0\}} \left(e^{\langle \xi, u\rangle} - 1 - \langle \chi^i(\xi), u\rangle\right) \mu^i(d\xi)\ ;$$

and

$$a = \left(\begin{array}{c:c} 0 & 0 \\ \hdashline 0 & \succeq \end{array}\right) \qquad \underset{(i \in I)}{\alpha^i} = \left(\begin{array}{ccccccc:ccc} & & & 0 & & & & & & \\ & & & \vdots & & & & & & \\ & & & 0 & & & & & & \\ 0 & \cdots & 0 & \alpha^i_{ii} & 0 & \cdots & 0 & \star & \cdots & \star \\ & & & 0 & & & & & & \\ & & & \vdots & & & & & & \\ & & & 0 & & & & & & \\ \hdashline & & & \star & & & & & & \\ & & & \vdots & & & & & \succeq & \\ & & & \star & & & & & & \end{array}\right) \text{ where } \alpha^i_{ii} \geq 0 \qquad \underset{(j \in J)}{\alpha^j} = 0$$

$$b = \left(\begin{array}{c} \geq \\ \vdots \\ \geq \\ \hdashline \star \\ \vdots \\ \star \end{array}\right) \qquad \underset{(i \in I)}{\beta^i} = \left(\begin{array}{c} \geq \\ \vdots \\ \geq \\ \beta^i_i \\ \geq \\ \vdots \\ \geq \\ \hdashline \star \\ \vdots \\ \star \end{array}\right) \text{ where } \beta^i_i \in \mathbb{R} \qquad \underset{(j \in J)}{\beta^j} = \left(\begin{array}{c} 0 \\ \vdots \\ 0 \\ \hdashline \star \\ \vdots \\ \star \end{array}\right)$$

Table 1. Structure of a, α^i, b and β^i. Stars denote arbitrary real numbers; the small $\geq$-signs denote non-negative real numbers and the big $\geq$-signs positive semi-definite matrices. A big 0 stands for a zero-matrix, and also empty regions in a matrix denote all-zero elements. The dotted lines indicate the boundary between the first m and the last n coordinates.

(b) the generator $\mathcal{A}$ of $(X_t)_{t\geq 0}$ is given by

$$\mathcal{A}f(x) = \frac{1}{2}\sum_{k,l=1}^{d}\left(a_{kl} + \sum_{i=1}^{m}\alpha^i_{kl}x_i\right)\frac{\partial^2 f(x)}{\partial x_k \partial x_l} + \tag{2.7}$$
$$+ \langle b + \sum_{i=1}^{d}\beta^i x_i, \nabla f(x)\rangle - \left(c + \sum_{i=1}^{m}\gamma^i x_i\right) f(x) +$$
$$+ \int_{D\setminus\{0\}} \left(f(x+\xi) - f(x) - \langle h(\xi), \nabla f(x)\rangle\right) m(d\xi) +$$
$$+ \sum_{i=1}^{m}\int_{D\setminus\{0\}} \left(f(x+\xi) - f(x) - \left\langle \chi^i(\xi), \nabla f(x)\right\rangle\right) x_i \mu^i(d\xi)$$

for all $f \in C_0^2(D)$ and $x \in D$.

We start by proving part (a) of the Theorem. Our proof is new, and an alternative to the proof given in Duffie et al. [2003]. The proof will make use of elementary (and well-known) results on infinitely divisible measures, that can be found in Appendix C.

Proof of Theorem 2.6, part (a). Let $(X_t)_{t\geq 0}$ be a regular affine process. By regularity, the functions F and R, given by

$$F(u) = \left.\frac{\partial \phi}{\partial t}(t,u)\right|_{t=0+}, \qquad R(u) = \left.\frac{\partial \psi}{\partial t}(t,u)\right|_{t=0+} \tag{2.8}$$

exist, and are continuous at $u = 0$. Thus, for $u \in \mathcal{U}$ and $x \in D$,

$$\lim_{t\downarrow 0}\frac{P_t f_u(x) - f_u(x)}{t} = \lim_{t\downarrow 0}\frac{\exp(\phi(t,u) + \langle x, \psi(t,u)\rangle) - \exp(\langle x, u\rangle)}{t} = \tag{2.9}$$
$$= \big(F(u) + \langle x, R(u)\rangle\big) f_u(x) \, .$$

Since the above limit exists, the functions $\{f_u : u \in \mathcal{U}\}$ are in the domain of $\mathcal{A}$, the generator of the Markov process $(X_t)_{t\geq 0}$, and we obtain

$$\frac{\mathcal{A}f_u(x)}{f_u(x)} = F(u) + \langle x, R(u)\rangle \, . \tag{2.10}$$

Denote now by $p_t(x, d\xi)$ the transition kernel of $(X_t)_{t\geq 0}$, i.e. $p_t(x, A) := P_t \mathbf{1}_{\{A\}}(x)$ for a Borel set $A \subseteq \mathcal{B}(D)$. We can also write (2.10) as

$$\frac{\mathcal{A}f_u(x)}{f_u(x)} = \lim_{t\downarrow 0}\frac{1}{t}\left\{\int_D e^{\langle \xi - x, u\rangle} p_t(x, d\xi) - 1\right\} =$$
$$= \lim_{t\downarrow 0}\left\{\frac{1}{t}\int_D \left(e^{\langle \xi - x, u\rangle} - 1\right) p_t(x, d\xi) + \frac{p_t(x,D) - 1}{t}\right\} =$$
$$= \lim_{t\downarrow 0}\left\{\frac{1}{t}\int_{D-x} \left(e^{\langle \xi, u\rangle} - 1\right) \widetilde{p}_t(x, d\xi)\right\} + \lim_{t\downarrow 0}\frac{p_t(x,D) - 1}{t} \, , \tag{2.11}$$

where we denote by $\widetilde{p}_t(x, d\xi) := p_t(x, d\xi + x)$ the 'shifted' transition kernel. Inserting $u = 0$ into (2.11) and (2.10), shows that the last term $\lim_{t\downarrow 0}(p_t(x, D) - 1)/t$ converges to $F(0) + \langle x, R(0)\rangle$. Define $c := F(0)$ and $\gamma_i := R_i(0)$ for all $i \in K$. From the contraction property of P_t it follows that $c + \langle x, \gamma\rangle \leq 0$ for all $x \in D$. Inserting $x = \pm e_j$ (e_j denotes the j-th unit vector in $\mathbb{R}^d$) shows that $\gamma^j = 0$ for $j \in J$, and thus (2.3g).
We write $\widetilde{F}(u) = F(u) - c$ and $\widetilde{R}(u) = R(u) - \gamma$, and focus on the integral term in the last line of (2.11). This term can be interpreted as a limit of log-characteristic functions of compound Poisson distributions with intensity $1/t$ and compounding measure $\widetilde{p}_t(x, d\xi)$ (See Definition C.6). We have already established by (2.10) that the limit exists, and is given by a function continuous at 0. By Lévy's continuity theorem, this implies that the compound Poisson distributions converge weakly to a limit distribution. Since any compound Poisson distribution is infinitely divisible, and the class of infinitely divisible distributions is closed under weak convergence, $\widetilde{F}(u) + \left\langle x, \widetilde{R}(u)\right\rangle$ must then be the log-characteristic function of some infinitely divisible random variable $K(x)$ for each $x \in D$. Together with the Lévy-Khintchine representation (cf. Theorem C.4) this shows the decomposition (2.6) of F and R.

To derive the admissibility conditions, consider first $K(0)$. It is an infinitely divisible random variable with log-characteristic function $\widetilde{F}(u)$, which by the Lévy-Khintchine formula is of the form

$$\widetilde{F}(u) = \frac{1}{2}\langle u, au\rangle + \langle b, u\rangle + \int_{\mathbb{R}^d\setminus\{0\}} \left(e^{\langle \xi, u\rangle} - 1 - \left\langle \widetilde{h}(\xi), u\right\rangle\right) m(d\xi)\ , \tag{2.12}$$

for some parameters (a, b, m) and a truncation function $\widetilde{h}$. The support of the transition kernel $\widetilde{p}_t(0, .)$ is contained in D, because $\widetilde{p}_t(0, d\xi) = p_t(0, d\xi)$. We are interested in the subspace where the support is restricted to the positive orthant. Denote by T_I the projection onto the coordinates with indices in I. It holds that

$$T_I\left(\operatorname{supp}\widetilde{p}_t(0, .)\right) \subseteq \mathbb{R}^m_{\geqslant 0}\ .$$

The same must hold for the support of the compound Poisson distributions, and thus also of their limit $K(0)$, i.e. we have

$$T_I\left(\operatorname{supp} K(0)\right)\ \subseteq\ \mathbb{R}^m_{\geqslant 0}\ .$$

In other words, the projection of $K(0)$ onto the coordinates $(x_i)_{i\in I}$ is a *non-negative* random variable. Applying Lemma C.8 on linear transformations of infinitely divisible random variables and Theorem C.5, the Lévy-Khintchine representation for non-negative random variables, we get

$$T_I a T_I^* = 0 \tag{2.13}$$

$$T_I b + \int_D T_I \circ \widetilde{h}(\xi)\, m(d\xi) \in \mathbb{R}^m_{\geqslant 0} \tag{2.14}$$

$$\int_{\mathbb{R}^m_{\geqslant 0}} (|y| \wedge 1)\, m(T_I^{-1}(dy)) < \infty\ . \tag{2.15}$$

The last equation implies (2.3h) and thus that $\left(e^{\langle \xi, u\rangle} - 1 - \langle h(\xi), u\rangle\right)$ is integrable with respect to $m(d\xi)$. Consequently we can replace $\widetilde{h}$ by h in (2.12) and (2.14). Note that $T_I \circ h \equiv 0$ by definition of h, such that the second equation directly implies (2.3d). Finally the first equation implies that $a_{kl} = 0$ if $k, l \in I$. By the Cauchy-Schwarz inequality we have that $|a_{kl}| \leq \sqrt{a_{kk} a_{ll}}$ and it follows that $a_{kl} = 0$ if $k \in I$ *or* $l \in I$, which is precisely (2.3a).

Define now for each $i \in I$ the random variable

$$K^i := \lim_{n\to\infty} K(ne_i)^{*\frac{1}{n}}\ ,$$

where $*\frac{1}{n}$ denotes the $1/n$-th convolution power, and the limit is understood as a limit in distribution. Note first that the $1/n$-th convolution power is well-defined, because $K(ne_i)$ is infinitely divisible. Second, we have that the log-characteristic function of $K_n(ne_i)^{*\frac{1}{n}}$ is given by $\frac{1}{n}\widetilde{F}(u) + \widetilde{R}_i(u)$. For $n \to \infty$ this sequence converges to $\widetilde{R}_i(u)$, which is continuous at the origin. Together this shows that the limit K^i exists and is an infinitely divisible random variable with log-characteristic function $\widetilde{R}_i(u)$. We conclude that

$$\widetilde{R}_i(u) = \frac{1}{2}\left\langle u, \alpha^i u\right\rangle + \left\langle \beta^i, u\right\rangle + \int_{\mathbb{R}^d\setminus\{0\}} \left(e^{\langle \xi, u\rangle} - 1 - \left\langle \widetilde{\chi}^i(\xi), u\right\rangle\right) \mu^i(d\xi)\ , \tag{2.16}$$

for some appropriate truncation function $\widetilde{\chi}^i$. Again we are interested in conditions on the support of K^i: We have that

$$T_{I\setminus\{i\}}\left(\operatorname{supp} \widetilde{p}_t(ne_i, .)\right) \subseteq \mathbb{R}^{(m-1)}_{\geqslant 0}\ .$$

The crucial observation here is that in the projection $T_{I\setminus\{i\}}$ the i-th component has to be excluded from I. The reason for this is the shift in direction e_i that has been applied in (2.11) to the transition kernel $p_t(x, d\xi)$, and which has changed its support. As before, we conclude, that also

$$T_{I\setminus\{i\}}\left(\operatorname{supp} K(ne_i)\right) \subseteq \mathbb{R}^{(m-1)}_{\geqslant 0}\ ,$$

and the same must hold for the convolution power $K(ne_i)^{*\frac{1}{n}}$, and finally for the limit K^i. Again we apply Lemma C.8 and the characterization of non-negative infinitely divisible random variables (Theorem C.5) and obtain

$$T_{I\setminus\{i\}}\alpha^i T^*_{I\setminus\{i\}} = 0 \tag{2.17}$$

$$T_{I\setminus\{i\}}\beta^i + \int_D T_{I\setminus\{i\}} \circ \widetilde{\chi}^i(\xi)\,\mu^i(d\xi) \in \mathbb{R}^{(m-1)}_{\geqslant 0} \tag{2.18}$$

$$\int_{\mathbb{R}^{(m-1)}_{\geqslant 0}} (|y| \wedge 1)\,\mu^i(T^{-1}_{I\setminus\{i\}}(dy)) < \infty\ , \tag{2.19}$$

The last equation implies (2.3k) and shows that we can replace $\widetilde{\chi}^i$ by χ^i in (2.16) and (2.18). By definition of χ^i we have that $T_{I\setminus\{i\}} \circ \chi^i \equiv 0$, such that (2.3e) follows directly from (2.18). Finally (2.17) yields that $\alpha^i_{kl} = 0$ if $k, l \in I \setminus \{i\}$. The Cauchy-Schwartz inequality gives $|\alpha^i_{kl}| \leq \sqrt{\alpha^i_{kk}\alpha^i_{ll}}$ such that we obtain (2.3c).

Consider next the random variables defined by

$$K^j_+ := \lim_{n\to\infty} K(ne_j)^{*\frac{1}{n}}, \quad \text{and} \quad K^j_- := \lim_{n\to\infty} K(-ne_j)^{*\frac{1}{n}}\ ,$$

where $j \in J$ and, as before, the limits are understood as limits in distribution. Again we obtain that K^j_+ and K^j_- are infinitely divisible; K^j_+ has the log-characteristic function $\widetilde{R}_j(u)$ which must be of the Lévy-Khintchine form (2.16), and K^j_- has log-characteristic function $-\widetilde{R}_j(u)$. Since α^j and $-\alpha^j$ cannot be both positive semi-definite, unless $\alpha^j = 0$, we conclude (2.3b). A similar argument for the Lévy measures leads to (2.3i). It is now clear that no truncation functions χ^j for $j \in J$ are necessary and $\widetilde{R}_j(u)$ are simply of the form $\widetilde{R}_j(u) = \langle \beta^j, u\rangle$, i.e. $K^j_+ = \beta^j$ a.s. and $K^j_- = -\beta^j$ a.s. As above we can deduce from $T_I(\operatorname{supp} \widetilde{p}_t(ne_j, .)) \subseteq \mathbb{R}^m_{\geqslant 0}$ that

$$\left\{T_I\beta^j\right\} = T_I\left(\operatorname{supp} K^j_+\right) \subseteq \mathbb{R}^m_{\geqslant 0} \quad \text{and} \quad \left\{-T_I\beta^j\right\} = T_I\left(\operatorname{supp} K^j_-\right) \subseteq \mathbb{R}^m_{\geqslant 0}$$

which leads to (2.3f).

It remains to show (2.3j). By Sato [1999, Theorem 8.7] convergence in law of infinitely divisible random variables X_t to a limit X, implies that for the corresponding Lévy measures, $\int f\,dm_n \to \int f\,dm$ holds for all functions $f \in C_b(\mathbb{R}^d)$ that vanish in a neighborhood of zero. Applying this to the compound Poisson approximation of $K(e_i/n)$ ($n \in \mathbb{N}$ and $i \in I$), we can choose a function g_n that is zero on $[-1/n, \infty)^m \times \mathbb{R}^n$, and have

$$\int_{\mathbb{R}^d} g_n(\xi)\,\widetilde{p}_t(e_i/n, d\xi) \to \frac{1}{n}\int_{\mathbb{R}^d} g_n(\xi)\mu^i(d\xi) + \int_{\mathbb{R}^d} g_n(\xi) m(d\xi)$$

as $t \to 0$. However, considering the support of $\widetilde{p}_t(e_i/n, d\xi)$, the left side is 0 for all $n \in \mathbb{N}$. The integral with respect to $m(d\xi)$ is zero too – condition (2.3h) has already been shown – such that also the $\mu^i(d\xi)$-integral must be zero for all $n \in \mathbb{N}$.

We conclude that the support of μ^i is a subset of D for all $i \in I$ and have shown part (a) of Theorem 2.6. □

We prepare to prove the second part of Theorem 2.6, i.e. the form of the generator. The following Lemma is not crucial, but convenient for the proof of the second part:

LEMMA 2.7. *The set $\mathcal{O}$, defined in* (1.4)*, equals $\mathbb{R}_{\geqslant 0} \times \mathcal{U}$. In particular $\phi(t,u)$ and $\psi(t,u)$ are defined on all of $\mathbb{R}_{\geqslant 0} \times \mathcal{U}$.*

We give only a sketch; for the full proof we refer to Duffie et al. [2003, Prop. 6.4]. Up to now $\phi(t,u)$ and $\psi(t,u)$ are only defined on $\mathcal{O}$, i.e. up to the point where $P_t f_u(0) = 0$. Since $\phi(t,u)$ is continuous inside $\mathcal{O}$, it is clear that for $(T,u) \in \mathcal{O}$ it must hold that $\lim_{t \to T} |\phi(t,u)| = +\infty$. According to Theorem 2.6.a, we already know that $\phi(t,u)$ and $\psi(t,u)$ must satisfy the generalized Riccati equations 2.2, with F and R given by (2.6), subject to the admissibility conditions. A comparison result for the generalized Riccati equations then shows that the solutions remain finite for all times and thus that $\lim_{t \to T} |\phi(t,u)| = +\infty$ can not happen for any finite time T and $u \in \mathcal{U}$. It follows that $\mathcal{O} = \mathbb{R}_{\geqslant 0} \times \mathcal{U}$.

By (2.10) we now know the action of the generator on the exponential functions. The idea is to extend this property to functions in $C_c^2(D)$ by applying some techniques from functional analysis. This part of the proof is essentially taken from Duffie et al. [2003], but we add explanations where we deem it appropriate.

The following function spaces will be needed:

- $C_0(D)$, the space of continuous functions on D, that vanish at infinity, i.e. for every $\epsilon > 0$ there exists a compact set $K \subset D$, such that $|f(x)| < \epsilon$ for all $x \in D \setminus K$. This space is endowed with the norm $\|.\|_\infty$, which generates the topology of uniform convergence on $C_0(D)$.
- $C_c^2(D)$, the space of functions with compact support that are twice continuously differentiable. We endow it with the norm

$$\|f\|_{D,2} := \sup_{x \in D} (1 + |x|) \left(\sum_{|\alpha| \le 2} \left| \frac{\partial^{|\alpha|}}{\partial x^\alpha} f(x) \right| \right) ,$$

  where α denotes a multi-index of length d.

PROOF OF THEOREM 2.6, PART (B). We define the operator $\mathcal{A}^\sharp$, mapping $C_c^2(D)$ into $C_0(D)$, as the integro-differential operator given by the right hand side of (2.7).

It will be convenient to split $\mathcal{A}^\sharp$ into a differential operator part and a integral operator part, i.e. $\mathcal{A}^\sharp = \mathcal{A}^\sharp_{\text{diff}} + \mathcal{A}^\sharp_{\text{int}}$, where

$$\begin{aligned}\mathcal{A}^\sharp_{\text{diff}} f(x) &:= \frac{1}{2}\sum_{k,l=1}^{d}\left(a_{kl} + \sum_{i=1}^{m}\alpha^i_{kl}x_i\right)\frac{\partial^2 f(x)}{\partial x_k \partial x_l} + \\ &\quad + \langle b + \sum_{i=1}^{d}\beta^i x_i, \nabla f(x)\rangle - \left(c + \sum_{i=1}^{m}\gamma^i x_i\right) f(x) \\ \mathcal{A}^\sharp_{\text{int}} f(x) &:= \int_{D\setminus\{0\}} \left(f(x+\xi) - f(x) - \langle h(\xi), \nabla f(x)\rangle\right) m(d\xi) + \\ &\quad + \sum_{i=1}^{m}\int_{D\setminus\{0\}} \left(f(x+\xi) - f(x) - \left\langle \chi^i(\xi), \nabla f(x)\right\rangle\right) x_i \mu^i(d\xi)\ .\end{aligned}$$

Similarly to the proof of Proposition 1.11, we define the set

$$\Theta := \left\{ h_{(u_I,g)}(x) = e^{\langle u_I, x_I\rangle}\int_{\mathbb{R}^n} f_{iz}(x_J) g(z)\, dz : \operatorname{Re} u_I < 0, g \in C_c^\infty(\mathbb{R}^n) \right\}\ , \tag{2.20}$$

and denote by $\mathcal{L}(\Theta)$ the set of complex linear combinations of elements of Θ.

The proof consists of the following steps:

(A) Show that each function in $C_c^2(D)$ can be approximated in $\|.\|_{2,D}$-norm by functions in $\mathcal{L}(\Theta)$.
(B) Show that $\mathcal{A}^\sharp$ is continuous from $C_c^2(D)$ to $C_0(D)$, i.e. $\|g_n - g\|_{2,D} \to 0$ implies that $\left\|\mathcal{A}^\sharp g_n - \mathcal{A}^\sharp g\right\|_\infty \to 0$.
(C) Show that $\mathcal{L}(\Theta) \subset D(\mathcal{A})$ and that $\mathcal{A}^\sharp g_n = \mathcal{A} g_n$ for all $g_n \in \mathcal{L}(\Theta)$.

In the proof of Proposition 1.11 we have shown that $\mathcal{L}(\Theta)$ is dense in $C_0(D)$ with respect to the $\|.\|_\infty$-norm. Since $C_c^2(D) \subset C_0(D)$, it follows that $C_c^2(D)$-functions can be approximated in $\|.\|_\infty$-norm by $\mathcal{L}(\Theta)$-functions. For point (A), however, a stronger assertion is needed: We have to approximate with respect to $\|.\|_{2,D}$. The proof is somewhat technical, and we refer to Duffie et al. [2003, Lemma 8.4].

Step (B): Since $\mathcal{A}^\sharp$ is linear, it suffices to show an estimate of the type

$$\left\|\mathcal{A}^\sharp g\right\|_\infty \le C\, \|g\|_{2,D} \quad \text{for all} \quad g \in C_c^2(D)\ . \tag{2.21}$$

The estimate is obvious for the differential-part $\mathcal{A}^\sharp_{\text{diff}}$. For the integral part, note that by a Taylor expansion it holds for any $g \in C_c^2(D)$, $x, \xi \in D$, and truncation function h, that

$$\begin{aligned}|g(x+\xi) - g(x) - \langle \nabla g(x), h(\xi)\rangle| &= |\langle \nabla g(x), (\xi - h(\xi))\rangle + M(x, g, \xi)| \le \\ &\qquad |\nabla g(x)||\xi - h(\xi)| + M(x, g, \xi)\ ,\end{aligned} \tag{2.22}$$

where

$$\|M(.,g,\xi)\|_\infty \le \|g\|_{2,D}\,|\xi|^2 \ .$$

Denoting the unit ball in $\mathbb{R}^d$ by $B_1(0)$, we can thus estimate for every $i \in I$

$$\left\| x \int_D \left(g(x+\xi) - g(x) - \langle \nabla g(x), h(\xi)\rangle\right)\, \mu^i(d\xi) \right\|_\infty \le$$
$$\le C_1 \|g\|_{2,D}\, \mu^i(D \setminus B_1(0)) + C_2 \|g\|_{2,D} \int_{B_1(0)\cap D} \left(\left|\xi - \chi^i(\xi)\xi\right| + |\xi|^2\right)\, \mu^i(d\xi) \ .$$

By the integrability condition (2.3k) for μ^i the latter integral is finite. For the Lévy measure $m(d\xi)$ a similar estimate holds, such that (2.21) with some appropriate constant C follows, and we have completed step (B).

Step (C): Let $h_{(u_I,g)} \in \Theta$. Using the notation $u(z) := (u_I, iz)$ we can write $h_{(u_I,g)}$ as

$$h_{(u_I,g)}(x) = \int_{\mathbb{R}^n} f_{u(z)}(x) g(z)\, dz \ .$$

On the exponential functions we know that $\mathcal{A}^\sharp f_u = \mathcal{A} f_u$, and we would like to justify the formal calculation

$$\mathcal{A}^\sharp h_{(u_I,g)} = \int_{\mathbb{R}^n} \mathcal{A}^\sharp f_{u(z)}(x) g(z)\, dz = \int_{\mathbb{R}^n} \mathcal{A} f_{u(z)}(x) g(z)\, dz = \mathcal{A} h_{(u_I,g)} \ .$$

For the differential part $\mathcal{A}^\sharp_{\text{diff}}$, it follows from well-known properties of the Fourier transform, that

(2.23)

$$\mathcal{A}^\sharp_{\text{diff}} \int_{\mathbb{R}^n} f_{u(z)}(x) g(z)\, dz = \int_{\mathbb{R}^n} \left(\frac{1}{2}\langle iz, a\, u(z)\rangle + \langle b, u(z)\rangle - c\right) f_{u(z)}(x) g(z)\, dz +$$
$$+ \sum_{i=1}^d x_i \int_{\mathbb{R}^n} \left(\frac{1}{2}\left\langle u(z), \alpha^i\, u(z)\right\rangle + \left\langle \beta^i, u(z)\right\rangle - \gamma^i\right) f_{u(z)}(x) g(z)\, dz \ .$$

For the integral part we calculate

$$\mathcal{A}^\sharp_{\mathrm{int}} \int_{\mathbb{R}^n} f_{u(z)}(x) g(z)\, dz = \tag{2.24}$$

$$= \int_{D\setminus\{0\}} \int_{\mathbb{R}^n} \Big(f_{u(z)}(x+\xi) - f_{u(z)}(x) - \big\langle h(\xi), u(z) f_{u(z)}(x) \big\rangle \Big) g(z)\, dz\, m(d\xi) +$$

$$+ \sum_{i=1}^d \int_{D\setminus\{0\}} \int_{\mathbb{R}^n} \Big(f_{u(z)}(x+\xi) - f_{u(z)}(x) - \big\langle \chi^i(\xi), u(z) f_{u(z)}(x) \big\rangle \Big) g(z)\, dz\, \mu^i(d\xi) =$$

$$= \int_{\mathbb{R}^n} \int_{D\setminus\{0\}} \Big(e^{\langle \xi, u(z)\rangle} - 1 - \langle h(\xi), u(z) \rangle \Big) f_{u(z)}(x)\, m(d\xi)\, g(z)\, dz +$$

$$+ \sum_{i=1}^d \int_{\mathbb{R}^n} \int_{D\setminus\{0\}} \Big(e^{\langle \xi, u(z)\rangle} - 1 - \big\langle \chi^i(\xi), u(z) \big\rangle \Big) f_{u(z)}(x)\, \mu^i(d\xi)\, g(z)\, dz\ ,$$

where the interchange of integrals is justified by Fubini's theorem in combination with the estimate (2.22). Combining (2.23) and (2.24) we obtain that for all $h_{(u_I,g)} \in \Theta$ and $x \in D$

$$\mathcal{A}^\sharp h_{(u_I,g)}(x) = \mathcal{A}^\sharp \int_{\mathbb{R}^n} f_{u(z)}(x) g(z)\, dz = \int_{\mathbb{R}^n} \left(F(u(z)) + \langle x, R(u(z)) \rangle \right) f_{u(z)}(x) g(z)\, dz$$

$$= \int_{\mathbb{R}^n} \left.\frac{\partial}{\partial t}\right|_{t=0} P_t f_{u(z)}(x) g(z)\, dz = \left.\frac{\partial}{\partial t}\right|_{t=0} P_t\, h_{(u_I,g)}(x)\ . \tag{2.25}$$

By linearity the equality clearly holds for all $h \in \mathcal{L}(\Theta)$. According to Sato [1999, Lemma 31.7], this pointwise equality for all $x \in D$ is enough to conclude that $\mathcal{L}(\Theta) \subset \mathcal{D}(\mathcal{A})$ and $\mathcal{A}^\sharp h = \mathcal{A} h$ for all $h \in \mathcal{L}(\Theta)$. Since we can approximate each $C^2_c(D)$-function in $\|.\|_{2,D}$-norm by functions in $\mathcal{L}(\Theta)$, there exists a sequence $(g_n)_{n\in\mathbb{N}}$ in $\mathcal{L}(\Theta)$ such that $\|g_n - g\|_{2,D} \to 0$ and also $\|g_n - g\|_\infty \to 0$. But $\mathcal{A}^\sharp$ is continuous from $C^2_c(D)$ to $C_0(D)$, such that we have

$$\left\| \mathcal{A} g_n - \mathcal{A}^\sharp g \right\|_\infty = \left\| \mathcal{A}^\sharp g_n - \mathcal{A}^\sharp g \right\|_\infty \to 0\ .$$

By Proposition 1.11 $(X_t)_{t\geq 0}$ is a Feller process. It is well known that this implies that its generator $\mathcal{A}$ is a closed operator on $C_0(D)$. But for a closed operator $\mathcal{A} g_n \to \mathcal{A}^\sharp g$ and $g_n \to g$ in $C_0(D)$ imply that $g \in \mathcal{D}(\mathcal{A})$ and that $\mathcal{A} g = \mathcal{A}^\sharp g$, yielding the claim of Theorem 2.6. □

3. Conditions for regularity and analyticity of affine processes

3.1. A condition for regularity.

DEFINITION 3.1. We say that an affine process satisfies **Condition A**, if the index set J of its real-valued components can be partitioned into $J = K \cup L$, and

(i) for any $x \in D, (t,u) \in \mathbb{R}_{\geqslant 0} \times \mathcal{U}$

$$\mathbb{E}^x\left[e^{\langle X_t, u\rangle}\right] = e^{\langle x_L, u_L\rangle} \cdot \mathbb{E}^{(x_{M\setminus L},0)}\left[e^{\langle X_t, u\rangle}\right] ,$$

(ii) and there exists $\epsilon > 0$ and $\tau > 0$, such that

$$\mathbb{E}^x\left[|(X_t)_K|^{1+\epsilon}\right] < \infty, \qquad \text{for all } t \in [0,\tau],\ x \in D.$$

REMARK 3.2. We will refer to condition (i) as space-homogeneity in the L-components, since it is equivalent to the statement that the law of $X_t + y_L$ under $\mathbb{P}^x$ equals the law of X_t under $\mathbb{P}^{(x+y_L)}$, for any $y_L \in \left\{y_L \in \mathbb{R}^d : y_{M\setminus L} = 0\right\}$. Note also, that together with the affine property of $(X_t)_{t\geq 0}$, it implies that $\psi_L(t,u) = u_L$ for all $(t,u) \in \mathbb{R}_{\geqslant 0} \times \mathcal{U}$. Condition (ii) can be described as existence of a moment of absolute order greater than one in the K-components.
Thus, an affine process X_t satisfies Condition A, if each of its components is either non-negative, space-homogeneous, or possesses a moment of absolute order greater one.

The motivation to introduce Condition A is the following Theorem. We postpone a discussion until after the proof.

THEOREM 3.3. *Suppose an affine process satisfies Condition A. Then it is a regular affine process.*

We will need the following Lemma:

LEMMA 3.4. *Suppose that the affine process $(X_t)_{t\geq 0}$ satisfies Condition A. Then for all $i \in I \cup K$ the derivatives*

$$\frac{\partial}{\partial u_i}\phi(t,u), \qquad \frac{\partial}{\partial u_i}\psi(t,u)$$

exist and are continuous for $(t,u) \in ([0,\tau] \times \mathcal{U}^\circ) \cap \mathcal{O}$.

REMARK 3.5. The set $([0,\tau] \times \mathcal{U}^\circ) \cap \mathcal{O}$ looks complicated, but its only property that will be used is the following: For any $u \in \mathcal{U}^\circ$, we can find $\tau(u) > 0$, such that $[0,\tau(u)) \times \{u\}$ is contained in $([0,\tau] \times \mathcal{U}^\circ) \cap \mathcal{O}$.

PROOF. Let $i,j \in I \cup K$. It holds that

$$\left|\frac{\partial}{\partial u_i}\exp\left(\langle u, X_t\rangle\right)\right| = \left|X_t^i\right| \cdot \exp\left(\langle \operatorname{Re} u, X_t\rangle\right) . \tag{3.1}$$

If $i \in I$, then the right hand side is uniformly bounded for all $t \in \mathbb{R}_{\geqslant 0}, u \in \mathcal{U}^\circ$, and thus in particular uniformly integrable. If $i \in K$, then the right hand side is by assumption 3.1.ii and by a well-known result (cf. Protter [2004, Theorem 11]) also uniformly integrable for all $t \in [0,\tau], u \in \mathcal{U}$. In any case we may conclude that $\frac{\partial}{\partial u_i}\mathbb{E}^x\left[e^{\langle X_t, u\rangle}\right]$ exists and is a continuous function of $(t,u) \in [0,\tau] \times \mathcal{U}^\circ$ for any $x \in D$. If in addition $(t,u) \in \mathcal{O}$, then Lemma 1.2 states that $\mathbb{E}^x\left[e^{\langle X_t, u\rangle}\right] = \exp\left(\phi(t,u) + \langle x, \psi(t,u)\rangle\right)$, from which the claim follows. □

Proof of Theorem 3.3. To simplify calculations we combine $\phi(t,u)$ and $\psi(t,u)$ into the 'big semi-flow' $\Upsilon(t,u)$ as described in Section 1.2. That is we set $\widehat{\mathcal{O}} := \mathcal{O} \times \mathbb{C}$ and define

$$\Upsilon : \widehat{\mathcal{O}} \to \mathbb{C}^{d+1}, \quad (t, u_1, \dots u_d, u_{d+1}) \mapsto \begin{pmatrix} \psi(t,(u_1,\dots,u_d)) \\ \phi(t,(u_1,\dots,u_d)) + u_{d+1} \end{pmatrix}. \tag{3.2}$$

Note that all vectors u have now a $(d+1)$-th component added; this component will be assigned to $\widehat{I} := I \cup \{d+1\}$. The semi-flow property is preserved by $\Upsilon(t,u)$, i.e. $\Upsilon(t+s,u) = \Upsilon(t,\Upsilon(s,u))$ for all $(t+s,u) \in \widehat{\mathcal{O}}$. The space-homogeneity condition for the components L implies that $\Upsilon_L(t,u) = u_L$ for all $(t,u) \in \mathcal{O}$. Clearly, the derivative $\frac{\partial}{\partial t}\Upsilon_L(t,u)\big|_{t=0}$ exists and is 0, for all $u \in \mathcal{U}$. In the rest of the proof we thus concentrate on the remaining (not space-homogeneous) components, whose indices we denote by $\Gamma := \widehat{I} \cup K$. Let $u \in \widehat{\mathcal{U}}^\circ := \mathcal{U}^\circ \times \mathbb{C}$ be fixed and assume that $t,s \in \mathbb{R}_{\geqslant 0}$ are small enough such that in particular $\phi(t+s,u)$, $\psi(t+s,u)$ and their u-derivatives are always well-defined (cf. Lemma 3.4). Denote by $\frac{\partial \Upsilon_\Gamma}{\partial u_\Gamma}(t,u)$ the Jacobian of Υ_Γ with respect to u_Γ. Using a Taylor expansion we have that

$$\int_0^s \Upsilon_\Gamma(r,\Upsilon(t,u))\,dr - \int_0^s \Upsilon_\Gamma(t,u)\,dr = \int_0^s \frac{\partial \Upsilon_\Gamma}{\partial u_\Gamma}(r,u)\,dr \cdot (\Upsilon_\Gamma(t,u) - u_\Gamma) + \\ + o\left(\|\Upsilon_\Gamma(t,u) - u_\Gamma\|\right) \tag{3.3}$$

On the other side, using the semi-flow property of Υ we can write the left hand side of (3.3) as

$$\begin{aligned}\int_0^s \Upsilon_\Gamma(r,\Upsilon(t,u))\,dr - \int_0^s \Upsilon_\Gamma(r,u)\,dr &= \int_0^s \Upsilon_\Gamma(r+t,u)\,dr - \int_0^s \Upsilon_\Gamma(r,u)\,dr = \\ = \int_t^{s+t} \Upsilon_\Gamma(r,u)\,dr - \int_0^s \Upsilon_\Gamma(r,u)\,dr &= \int_s^{s+t} \Upsilon_\Gamma(r,u)\,dr - \int_0^t \Upsilon_\Gamma(r,u)\,dr = \\ &= \int_0^t \Upsilon_\Gamma(r+s,u)\,dr - \int_0^t \Upsilon_\Gamma(r,u)\,dr\,.\end{aligned} \tag{3.4}$$

Denoting the last expression by $I(s,t)$ and combining (3.3) with (3.4) we obtain

$$\lim_{t\downarrow 0}\frac{\left\|\frac{1}{s}I(s,t)\right\|}{\|\Upsilon_\Gamma(t,u)-u_\Gamma\|}=\left\|\frac{1}{s}\int_0^s\frac{\partial\Upsilon_\Gamma}{\partial u_\Gamma}(r,u)\,dr\right\|\ .$$

Define $M(s,u):=\frac{1}{s}\int_0^s\frac{\partial\Upsilon_\Gamma}{\partial u_\Gamma}(r,u)\,dr$. Note that as $s\to 0$, it holds that $M(s,u)\to\frac{\partial\Upsilon_\Gamma}{\partial u_\Gamma}(0,u)=I_\Gamma$ (the identity matrix). Thus for s small enough $\|M(s,u)\|\neq 0$, and we conclude that

$$\text{(3.5)}\quad \lim_{t\downarrow 0}\frac{1}{t}\|\Upsilon_\Gamma(t,u)-u_\Gamma\|=$$
$$=\left\|\lim_{t\downarrow}\frac{I(s,t)}{st}\right\|\cdot\|M(s,u)\|^{-1}=\left\|\frac{\Upsilon_\Gamma(s,u)-u_\Gamma}{s}\right\|\cdot\|M(s,u)\|^{-1}\ .$$

The right hand side of (3.5) is well-defined and finite, implying that also the limit on the left hand side is. Thus, combining (3.3) and (3.4), dividing by st and taking the limit $t\downarrow 0$ we obtain

$$\lim_{t\downarrow 0}\frac{\Upsilon_\Gamma(t,u)-u_\Gamma}{t}=\frac{\Upsilon_\Gamma(s,u)-u_\Gamma}{s}\cdot M(s,u)^{-1}\ .$$

Again we may choose s small enough, such that $M(s,u)$ is invertible, and the right hand side of the above expression is well-defined. The existence and finiteness of the right hand side then implies the existence of the limit on the left. In addition the right hand side is a continuous function of $u\in\widehat{\mathcal{U}}^\circ$, such that also the left hand side is. Adding back the components L, for which a time derivative trivially exists ($\psi_L(t,u)=u_L$ for all $t\geq 0$), we obtain that

$$\text{(3.6)}\qquad \mathcal{R}(u):=\lim_{t\downarrow 0}\frac{\Upsilon(t,u)-u}{t}=\left.\frac{\partial}{\partial t}\Upsilon(t,u)\right|_{t=0}$$

exists and is a continuous function of $u\in\widehat{\mathcal{U}}^\circ$. Denoting the first d components of $\mathcal{R}(u)$ by $R(u)$ and the $d+1$-th component by $F(u)$ we can 'disentangle' the big semi-flow Υ, drop the $(d+1)$-th component of u, and see that

$$\text{(3.7)}\qquad F(u):=\left.\frac{\partial}{\partial t}\phi(t,u)\right|_{t=0}\quad\text{and}\quad R(u):=\left.\frac{\partial}{\partial t}\psi(t,u)\right|_{t=0}$$

are likewise well-defined and continuous on $\mathcal{U}^\circ$. To show that $(X_t)_{t\geq 0}$ is regular affine (cf. (2.1)) it remains to show that (3.7) extends continuously to $\mathcal{U}$:
To this end let $t_n\downarrow 0$, $x\in D$, $u\in\mathcal{U}^\circ$, and rewrite (3.7), as in the proof of Theorem 2.6, as

$$\text{(3.8)}\quad F(u)+\langle x,R(u)\rangle=\lim_{n\to\infty}\frac{\exp\left(\phi(t_n,u)+\langle x,\psi(t_n,u)-u\rangle\right)-1}{t_n}=$$
$$=\lim_{n\to\infty}\frac{f_{-u}(x)\,P_{t_n}f_u(x)-1}{t_n}=\lim_{n\to\infty}\frac{1}{t_n}\left\{\int_{D-x}e^{\langle x,u\rangle}\,\widetilde{p}_t(x,d\xi)-1\right\}\ ,$$

where $\widetilde{p}_t(x, d\xi)$ is the 'shifted transition kernel' of the Markov process $(X_t)_{t\geq 0}$ (see Page 22f for details). The right hand side of (3.8) can be regarded as a limit of log-characteristic functions of (infinitely divisible) sub-stochastic measures[7]. That is, there exist infinitely divisible measures $\mu_n(x, d\xi)$, such that

$$\exp\left(F(u) + \langle x, R(u)\rangle\right) = \lim_{n\to\infty} \int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, \mu_n(x, d\xi), \quad \text{for all } u \in \mathcal{U}^\circ.$$

Let now $\theta \in \mathbb{R}^d$ with $\theta_I < 0$ and $\theta_J = 0$ (note that $\theta \in \mathcal{U}^\circ$) and consider the exponentially tilted measures $e^{\langle\theta,\xi\rangle}\, \mu_n(x, d\xi)$. Their characteristic functions converge to $\exp\left(F(u+\theta) + \langle x, R(u+\theta)\rangle\right)$. Thus, by Lévy's continuity theorem, there exists $\mu_*(x, d\xi)$ such that $e^{\langle\theta,\xi\rangle}\, \mu_n(x, d\xi) \to \mu_*(x, d\xi)$ weakly. On the other hand, by Helly's selection theorem, $\mu_n(x, d\xi)$ has a vaguely convergent subsequence, which converges to some measure $\mu(x, d\xi)$. By uniqueness of the weak limit we conclude that $\mu(x, d\xi) = e^{\langle -\theta,\xi\rangle}\mu_*(x, d\xi)$. Thus we have that for all $x \in D$ and $u \in \mathcal{U}^\circ$ with $\operatorname{Re} u$ in a neighborhood $B(\theta)$ of θ,

$$\text{(3.9)} \quad \exp\left(F(u) + \langle x, R(u)\rangle\right) = \lim_{n\to\infty} \int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, \mu_n(d, d\xi) =$$

$$= \lim_{n\to\infty} \int_{\mathbb{R}^d} e^{\langle u-\theta,\xi\rangle}\, e^{\langle\theta,\xi\rangle}\mu_n(x, d\xi) = \int_{\mathbb{R}^d} e^{\langle u-\theta,\xi\rangle}\, \mu_*(x, d\xi) = \int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, \mu(x, d\xi)\,.$$

But the choice of θ was arbitrary, such that (3.9) extends to all $u \in \mathcal{U}^\circ$. Applying dominated convergence to the last term of (3.9) shows that both F and R have a continuous extension to all of $\mathcal{U}$, which we also denote by F and R respectively. It remains to show that (3.7) remains valid on $\mathcal{U}$:
Let $u \in \mathcal{U}$ and $(u_n)_{n\in\mathbb{N}} \in \mathcal{U}^\circ$ such that $u_n \to u$. Remember that by Proposition 1.10 $u_n \in \mathcal{U}^\circ$ implies that also $\psi(t, u_n) \in \mathcal{U}^\circ$ for any $t \geq 0$. Thus we have

$$\text{(3.10)} \quad \int_0^t R(\psi(s,u))\, ds = \int_0^t \lim_{u_n\to u} R(\psi(s,u_n))\, ds = \lim_{u_n\to u} \int_0^t R(\psi(s,u_n))\, ds =$$

$$= \lim_{u_n\to u} \int_0^t \lim_{r\downarrow 0} \frac{\psi(r,\psi(s,u_n)) - \psi(s,u_n)}{r}\, ds = \lim_{u_n\to u} \int_0^t \lim_{r\downarrow 0} \frac{\psi(r+s,u_n) - \psi(s,u_n)}{r}\, ds$$

$$= \lim_{u_n\to u} \int_0^t \frac{\partial}{\partial t}\psi(s,u_n)\, ds = \lim_{u_n\to u} \psi(t,u_n) - u_n = \psi(t,u) - u\,.$$

Since the left hand side of (3.10) is t-differentiable, also the right hand side is, and we obtain $R(u) = \frac{\partial}{\partial t}\psi(t,u)\big|_{t=0}$ for all $u \in \mathcal{U}$. A similar calculation as above can be made upon replacing R with F, resulting in $F(u) = \frac{\partial}{\partial t}\phi(t,u)\big|_{t=0}$ for all $u \in \mathcal{U}$, and thus showing that the affine process $(X_t)_{t\geq 0}$ is regular. $\square$

We discuss the following special cases of Theorem 3.3:

[7]More precisely as log-characteristic functions of compound Poisson distributions with intensity $1/t_n$ and kernel $\widetilde{p}_t(x, .)$; see also Definition C.6

(i) An affine process on $D = \mathbb{R}^d_{\geqslant 0}$ is regular.
(ii) A space-homogeneous affine process, or equivalently an affine process with $\psi(t, u) = u$, is regular.
(iii) An affine process that possesses an absolute moment of order greater one is regular.

Non-negative affine processes as in (i) have been studied as so-called continuously branching processes with immigration (CBI-processes) in the seventies. See e.g. Kawazu and Watanabe [1971], where the one-dimensional case ($D = \mathbb{R}_{\geqslant 0}$) is discussed, and regularity of this process is shown (using, however, a rather different approach). Is is not hard to see that a space-homogeneous affine process as in (ii) is necessarily a Lévy process (see also Example 1.4). Also in this case regularity follows by a very simple argument based on the Cauchy functional equation. The case (iii) has been studied before, for example by Dawson and Li [2006].
The contribution of our proof given above, is to provide a unified treatment of the different approaches, and – more importantly – to show that all the mentioned conditions can be mixed component-by-component. This is relevant, since a number of models in mathematical finance naturally satisfy such a 'mixed' condition of type A. We give an example:

Example 3.6. Let $(X_t, V_t^1, \dots, V_t^m)_{t \geq 0}$ be an affine process. The process $S_t = e^{X_t}$ represents the price of a risky asset, while $(V_t^1, \dots, V_t^m)$ are some latent factors describing market activity (such as stochastic variance, intensity of jumps, or 'business time') By construction, X_t will be $\mathbb{R}$-valued, while V_t is $\mathbb{R}^m_{\geqslant 0}$-valued. In addition the following homogeneity assumption is often made: The distribution of returns $S_{t+\Delta t}/S_t$ shall depend only on the current state of $(V_t^1, \dots, V_t^m)$, but not of S_t. Such a model automatically fulfills Condition A, is thus regular affine, and can be characterized by the results obtained in section 2.

3.2. 'Analytic' affine processes. We introduce now another condition implying regularity, the notion of an analytic affine process. Essentially, this is an affine process, whose moment generating function exists on a set with non-empty interior, and satisfies a uniform boundedness condition with respect to time t.

Definition 3.7 (Analytic Affine Process). Let $(X_t)_{t \geq 0}$ be an affine process. For each $t \geq 0$, define

$$\mathcal{D}_t = \left\{ y \in \mathbb{R}^d : \sup_{0 \leq s \leq t} \mathbb{E}^x \left[\exp\left(\langle X_s, y \rangle \right) \right] < \infty \quad \text{for all } x \in D \right\} ;$$

$$\mathcal{D}_{t+} := \bigcup_{s > t} \mathcal{D}_s \quad \text{and} \quad \mathcal{D} = (\mathcal{D}_{0+})^\circ \cup \{0\} \ .$$

We call $\mathcal{D}$ the **real domain** of $(X_t)_{t \geq 0}$. If $\mathcal{D}$ has non-empty interior, we call $(X_t)_{t \geq 0}$ an **analytic affine process**.

It follows immediately from Proposition B.2 and the positivity of X_t^I, that

(a) for each $t \in \mathbb{R}_{\geqslant 0}$ the sets $\mathcal{D}_t, \mathcal{D}_{t+}$ are convex and contain $[-\infty, 0]^m \times \{0\}^n$,
(b) $\mathcal{D}$ is convex and contains $[-\infty, 0]^m \times \{0\}^n$, and
(c) $t \leq s$ implies that $\mathcal{D}_t \supseteq \mathcal{D}_s$ and $\mathcal{D}_{t+} \supseteq \mathcal{D}_{s+}$.

Let us at this point introduce the following useful notation: Suppose that S is a subset of $\mathbb{R}^d$. Then the **tube domain** $S_\mathbb{C}$ associated to S is given by

$$S_\mathbb{C} := \left\{u \in \mathbb{C}^d : \operatorname{Re} u \in S\right\} \ . \tag{3.11}$$

Sometimes we shall also consider sets $S' \subset \mathbb{R}_{\geqslant 0} \times \mathbb{R}^d$, i.e. containing points (t, y), where the first coordinate represents time. In this case the 'tubification' is only applied to the last d coordinates, that is $S'_\mathbb{C} := \left\{(t, u) \in \mathbb{R}_{\geqslant 0} \times \mathbb{C}^d : (t, \operatorname{Re} u) \in S'\right\}$. For example, setting

$$\mathcal{E} := \left\{(t, y) \in \mathbb{R}_{\geqslant 0} \times \mathbb{R}^d : y \in \mathcal{D}_{t+}\right\} \ ,$$

it follows from $\left|e^{\langle X_t, u\rangle}\right| = e^{\langle X_t, \operatorname{Re} u\rangle}$ that $\sup_{0 \leq s \leq t} \left|\mathbb{E}^x\left[e^{\langle X_s, u\rangle}\right]\right| < \infty$ for all (t, u) in the tube-domain $\mathcal{E}_\mathbb{C}$. Defining also $\mathcal{E}^\circ = \left\{(t, y) \in \mathbb{R}_{\geqslant 0} \times \mathbb{R}^d : y \in \mathcal{D}^\circ_{t+}\right\}$ we will subsequently show that:

- The functions $\phi(t, u)$ and $\psi(t, u)$, characterizing the affine process $(X_t)_{t \geq 0}$ have unique extensions to analytic functions[8] on $\mathcal{E}^\circ_\mathbb{C}$. Hence the name *analytic affine*.
- Every analytic affine process is a regular affine process, and also the functions F and R defined in (2.1) have extensions to analytic functions on $\mathcal{D}^\circ_\mathbb{C}$.

We give some examples of processes that are analytic affine:

Example 3.8. Every affine process on $\mathbb{R}^m_{\geqslant 0}$ has $(-\infty, 0]^d$ a subset of its real domain, and is thus analytic affine. This follows immediately from Definition 3.7 and the positivity of X_t.

Example 3.9. Consider an affine process $(X_t, V_t^1, \ldots, V_t^m)_{t \geq 0}$ with state space $\mathbb{R} \times \mathbb{R}^m_{\geqslant 0}$ as introduced in Example 3.6. Suppose that $S_t = e^{X_t}$ models a risky asset, and that – for any $x \in D$ – $\mathbb{P}^x$ is a martingale measure for S_t. Then $\sup_{0 \leq s \leq t} \mathbb{E}^x\left[e^{X_s}\right] = e^x < \infty$ for all $t \in \mathbb{R}_{\geqslant 0}$. It is straightforward to derive that $(X_t, V_t)_{t \geq 0}$ has a real domain including the set $[0, 1] \times (-\infty, 0]^m$, and thus is analytic affine.

[8]'extension to analytic functions' may sound awkward, and the reader may wonder why we do not simply write 'analytic extension'. The reason is that a function is analytically extended from *within* its domain of analyticity, whereas here the function might be extended from the *boundary* – and hence from outside – of its domain of analyticity. See Proposition 3.11 for details.

Example 3.10. It might seem at first sight that Definition 3.7 is a strictly stronger requirement than Condition A introduced in Definition 3.1, and that every affine process satisfying Condition A, is also analytic affine. However, this is not the case: Consider the OU-type process $(Y_t)_{t\geq 0}$ defined by

$$dY_t = -\lambda Y_t\, dt + dL_t, \qquad Y_0 = y \in \mathbb{R}\ ,$$

where $\lambda > 0$ and L_t is a Lévy process with Lévy measure $m(d\xi) = \mathbf{1}_{\{\xi>0\}}\frac{\xi^{-3/2}}{2\sqrt{\pi}} + \mathbf{1}_{\{-1\leq\xi<0\}}$. As pointed out in Example 1.5 this process is a regular affine process with

$$\phi(t,u) = \int_0^t \kappa(e^{-\lambda s}u)\, ds, \quad \psi(t,u) = e^{-\lambda t}u\ ,$$

where

$$\kappa(u) = \int_{\mathbb{R}\setminus\{0\}} \left(e^{\xi u} - 1 - u\frac{\xi}{1+\xi^2}\, m(d\xi)\right) = \\ = -\sqrt{-u} + \frac{1-e^{-u}}{u} + u\left(\frac{\log(2)}{2} + \frac{\sqrt{2\pi}}{4}\right) - 1\ ,$$

defined for all $u \in (-\infty,0)_{\mathbb{C}}$ and for $u = 0$. The process Y_t is neither non-negative, nor space-homogeneous. It is also not of finite expectation, since $\mathbb{E}^x\left[X_t\right] = t\kappa'(0-) + xe^{-\lambda t} = +\infty$. Thus it does not satisfy Condition A. It is however analytic affine, since $\sup_{0\leq s\leq t}\mathbb{E}^x\left[e^{yX_s}\right] < \infty$ for all $t \geq 0$, and $y \leq 0$.

The following result will be needed for the extension of $\phi(t,u)$ and $\psi(t,u)$:

Proposition 3.11 (Extension Principle for moment generating functions). *Let $f : i\mathbb{R}^d \to \mathbb{C}$ be a continuous function defined on the purely imaginary numbers, and let $S \subseteq \mathbb{R}^d$ be a convex set containing 0. Suppose there exists a function $\widetilde{f} : S_{\mathbb{C}} \to \mathbb{C}$, that can be represented as*

$$\widetilde{f}(u) = \int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, \mu(d\xi), \quad u \in S_{\mathbb{C}}\ ;$$

for some Borel measure $\mu(d\xi)$, and satisfies $f = \widetilde{f}\Big|_{i\mathbb{R}^d}$. Then $\widetilde{f}$ is the unique function with these properties.

Proof. Suppose there exist functions $\widetilde{f}(u)$, and $\widehat{f}(u)$, both satisfying the properties given above. Let $y \in S$ and define $\Omega := \{\omega \in \mathbb{C} : \mathrm{Re}\,\omega \in \mathbb{R}, \mathrm{Im}\,\omega \in [0,1]\}$. Consider the function

$$h : \Omega \to \mathbb{C} : \omega \mapsto \widetilde{f}(-i\omega y) - \widehat{f}(-i\omega y)\ .$$

By Proposition B.4, h has the following properties: It is continuous on Ω, analytic on Ω°, and moreover real-valued for real ω. By the Schwarz reflection principle, it thus has an analytic extension to $\Omega_* := \Omega \cup \overline{\Omega}$. However for any $\omega \in \mathbb{R}$ the function

h takes the value zero. We conclude that (the analytic extension of) h is zero on all of Ω_*. It follows that $\widetilde{f}(u) = \widehat{f}(u)$ for all u with $\operatorname{Re} u = y$. Since y was arbitrary in S, $\widetilde{f} = \widehat{f}$ on all of $S_{\mathbb{C}}$. □

LEMMA 3.12. *(a) The function $(t, u) \mapsto P_t f_u(x)$ is continuous on $\mathcal{E}_{\mathbb{C}}^{\circ}$ for every $x \in D$.*

(b) Let $(X_t)_{t\geq 0}$ be an analytic affine process, and let

$$\mathcal{Q} := \{(t, u) \in \mathcal{E}_{\mathbb{C}} : |P_s f_u(0)| \neq 0 \, \forall \, s \in [0, t)\} \ .$$

Then $\mathcal{Q}$ is open in $\mathcal{E}_{\mathbb{C}}$, and there exist a unique extension of $\phi(t, u)$, $\psi(t, u)$ to $\mathcal{Q}$, such that

$$\mathbb{E}^x \left[e^{\langle u, X_t \rangle} \right] = \exp\left(\phi(t, u) + \langle x, \psi(t, u) \rangle \right) \tag{3.12}$$

for all $x \in D$, and (t, u) in $\mathcal{Q}$.

REMARK 3.13. By (1.3), it is clear that $\mathbb{R}_{\geqslant 0} \times i\mathbb{R}^d \subseteq \mathcal{Q}$. Also, from the continuity of $P_t f_u(0) \in \mathcal{E}_{\mathbb{C}}^{\circ}$, and the fact that $P_0 f_u(0) = f_u(0) = 1$, it follows that $\{0\} \times \mathcal{D}_{\mathbb{C}} \subseteq \mathcal{Q}$. We will eventually show in Lemma 3.19, that actually $\mathcal{Q} = \mathcal{E}_{\mathbb{C}}$, i.e. (3.12) holds for all $(t, u) \in \mathcal{E}_{\mathbb{C}}$.

PROOF. To show (a) we establish first that for all $t \geq 0$,

$$\mathcal{D}_{t+}^{\circ} = \bigcup_{s>t} \mathcal{D}_s^{\circ} \ . \tag{3.13}$$

The inclusion '$\supseteq$' should be clear from Definition 3.7. To show equality, let $y \in \mathcal{D}_{t+}^{\circ}$. There exists an open neighborhood N_y of y, such that $N_y \subseteq \mathcal{D}_{t+}$. The neighborhood N_y contains a polytope P, i.e. the convex hull of finitely many points $\{a_1, \ldots, a_M\}$, such that $y \in P^{\circ}$ (one could choose e.g. P as a L_1-ball around y). Since $\mathcal{D}_{s_1} \subseteq \mathcal{D}_{s_2}$, whenever $s_1 \geq s_2$, there exists an $\epsilon > 0$, such that $\{a_1, \ldots, a_M\} \subseteq \mathcal{D}_s$ for all $t < s \leq t + \epsilon$. Because each $\mathcal{D}_s$ is convex, it follows that $P \subseteq \mathcal{D}_s$, and thus $y \in \mathcal{D}_s^{\circ}$ for all $t < s \leq t + \epsilon$, which implies (3.13).

Fix now $x \in D$ and let $(t_n, y_n)_{n\in\mathbb{N}}$ be a sequence converging to $(t_*, y_*) \in \mathcal{E}^{\circ}$. By definition of $\mathcal{E}^{\circ}$, we have that $y_* \in \mathcal{D}_{t_*+}^{\circ}$, and by (3.13) there exists an $\epsilon > 0$ such that $y_* \in \mathcal{D}_{t_*+\epsilon}^{\circ}$. Clearly we can truncate the sequence $(t_n, y_n)_{n\in\mathbb{N}}$, such that $t_n \leq t + \epsilon$ and $y_n \in \mathcal{D}_{t_*+\epsilon}^{\circ}$ for all $n \in \mathbb{N}$. Now for arbitrary $y \in \mathcal{D}_{t_*+\epsilon}^{\circ}$, it holds that

$$\sup_{t_n} \mathbb{E}^x \left[e^{\langle y, X_{t_n} \rangle} \mathbf{1}_{\{|X_{t_n}| > r\}} \right] \leq \sup_{0\leq s\leq t+\epsilon} \mathbb{E}^x \left[e^{\langle y, X_s \rangle} \right] \mathbb{P}^x \left(|X_{t_n}| > r \right) \ .$$

The right hand side goes to 0 as $r \to \infty$, due to the stochastic continuity of X_t. Thus $e^{\langle y, X_{t_n} \rangle}$ is uniformly integrable and $\mathbb{E}^x \left[e^{\langle y, X_{t_n} \rangle} \right] \to \mathbb{E}^x \left[e^{\langle y, X_{t_n} \rangle} \right]$, pointwise for each $y \in \mathcal{D}_{t_*+\epsilon}^{\circ}$. On the other hand, $y \mapsto \log \mathbb{E}^x \left[e^{\langle y, X_{t_n} \rangle} \right]$ is, for each $n \in \mathbb{N}$ a closed convex function on $\mathcal{D}_{t_*+\epsilon}^{\circ}$ by Proposition B.2. A sequence of closed convex functions, that converges pointwise on an open set, converges uniformly on each of its compact subsets (cf. Appendix A). Thus also $\mathbb{E}^x\left[e^{\langle y_n, X_{t_n}\rangle}\right] \to \mathbb{E}^x\left[e^{\langle y_*, X_{t_*}\rangle}\right]$ and we have shown that $(t,y) \mapsto P_t f_y(x)$ is continuous in $\mathcal{E}^\circ$. The extension to complex arguments $(t,u) \in \mathcal{E}^\circ_\mathbb{C}$ follows by dominated convergence and the fact that $\mathbb{E}^x\left[\left|e^{\langle u, X_t\rangle}\right|\right] \le \mathbb{E}^x\left[e^{\langle \operatorname{Re} u, X_t\rangle}\right]$.

We turn towards assertion (b). To show that $\mathcal{Q}$ is open in $\mathcal{E}_\mathbb{C}$ write the complement $\mathcal{Q}^c := \mathcal{E}_\mathbb{C} \setminus \mathcal{Q}$ of $\mathcal{Q}$ as

$$\mathcal{Q}^c = \left\{(t,u) \in \mathcal{E}^\circ_\mathbb{C} : |P_s f_u(0)| = 0 \text{ for some } s \in [0,t)\right\} .$$

Note that $\mathcal{E}_\mathbb{C} = \mathcal{E}^\circ_\mathbb{C} \cup i\mathbb{R}^d$, but $i\mathbb{R}^d \subseteq \mathcal{Q}$ by the affine property (1.3). Thus $\mathcal{Q}^c \subseteq \mathcal{E}^\circ_\mathbb{C}$. But $P_t f_u(0)$ is a continuous function for $(t,u) \in \mathcal{E}^\circ_\mathbb{C}$. It follows that $\mathcal{Q}^c$ is a closed subset of $\mathcal{E}_\mathbb{C}$, and thus its complement $\mathcal{Q}$ open in $\mathcal{E}_\mathbb{C}$.
For all $t \in \mathbb{R}_{\geqslant 0}$, $x, \xi \in D$, define now the functions

$$f_1(u;t,x,\xi) = P_t\, f_u(x) \cdot P_t\, f_u(\xi) \quad \text{and} \quad f_2(u;t,x,\xi) = P_t\, f_u(0) \cdot P_t\, f_u(x+\xi) .$$

By the affine property (1.3) $f_1(u;t,x,\xi) = f_2(u;t,x,\xi)$ whenever $\operatorname{Re} u = 0$. Moreover, both functions can be represented as moment generating functions, defined on $(\mathcal{D}_t)_\mathbb{C}$. Thus we can apply the extension principle of Proposition 3.11 with $S = \mathcal{D}_t$, and conclude that $f_1(u;t,x,\xi) = f_2(u;t,x,\xi)$ for all $(t,u) \in \mathcal{E}_\mathbb{C}$, $x, \xi \in D$.
Suppose now that $(t,u) \in \mathcal{Q}$, such that $|P_t f_u(0)| \neq 0$. Then there exists a unique $\phi(t,u)$ such that $P_t f_u(0) = e^{\phi(t,u)}$. We can define $g(x,t,u) = e^{-\phi(t,u)} P_t f_u(x)$, which by the equality of $f_1(u)$ and $f_2(u)$, the function $g(x,t,u)$ must satisfy the Cauchy functional equation

$$g(x,t,u) g(\xi,t,u) = g(x+\xi,t,u)$$

for all $(t,u) \in \mathcal{Q}$, $x, \xi \in D$. It follows that there exists a unique $\psi(t,u)$ such that $g(x,t,u) = e^{\langle x, \psi(t,u)\rangle}$, and we have shown that the decomposition (3.12) holds for all $(t,u) \in \mathcal{Q}$. □

We make the following definition:

Definition 3.14. Let S be a convex subset of $\mathbb{R}^d$, that includes $(-\infty,0]^m \times \{0\}^n$. A function $f : S \to \mathbb{R}, y \mapsto f(y)$, is said to have **Property C**[9], if for any $y_J \in \mathbb{R}^d$ such that $(0, y_J) \in S$, the function

$$f_* : (-\infty,0]^m \times [0,1] \to \mathbb{R} : (y_I, \lambda) \mapsto f((y_I, \lambda y_J))$$

is continuous.

Note that the above definition is indirectly tied to the structure of the state space D, through the definitions of I and J (cf. Section 1.1). Property C should be

[9]'C' as in continuity

regarded as a weaker form of continuity: While the dependency of f on the first m components is truly continuous, the dependency on the last n components is only 'continuous along line segments' emanating from the origin.

LEMMA 3.15. *(a) Let $\mu(d\xi)$ be a sub-stochastic measure with support D. Define $f(y) = \int_D e^{\langle y,\xi\rangle}\mu(d\xi)$, and $S := \left\{y \in \mathbb{R}^d : f(y) < \infty\right\}$. Then $f : S \to \mathbb{R}$ has property C.*

(b) Let $\mu(d\xi)$ be a sub-stochastic infinitely divisible measure with support $\mathbb{R}^d$, and with Lévy measure ν. Suppose that ν has support D.
Define $f(y) = \int_{\mathbb{R}^d} e^{\langle y,\xi\rangle}\mu(d\xi)$, and $S := \left\{y \in \mathbb{R}^d : f(y) < \infty\right\}$. Then $f : S \to \mathbb{R}$ has property C.

PROOF. Define $S_* := (-\infty, 0]^m \times [0, 1]$. For assertion (a), we have to show that for any $y_J \in \mathbb{R}^d$ such that $(0, y_J) \in S$, the function

$$f_* : S_* \to \mathbb{R}, (y_I, \lambda) \mapsto \int \exp\left(\langle y_I, \xi_I\rangle + \lambda \langle y_J, \xi_J\rangle\right) \mu(d\xi)$$

is continuous. But for $(y_I, \lambda) \in S_*$, we can bound the integrand by $e^{\lambda\langle y_J,\xi_J\rangle}$. By Proposition B.4, the function $\lambda \mapsto \int_{\mathbb{R}^d} e^{\lambda\langle y_J,\xi_J\rangle}\,\mu(d\xi)$ is continuous on $[0, 1]$; dominated convergence implies that also f_* is.
To show assertion (b), note that by the Lévy-Khintchine formula $f(y)$ is of the form

$$f(y) := \exp\left\{\frac{1}{2}\langle y, ay\rangle + \langle b, y\rangle - \gamma + \int_D \left(e^{\langle\xi,y\rangle} - 1 - \langle h(\xi), y\rangle\right) m(d\xi)\right\}. \tag{3.14}$$

Suppose, without loss of generality, that the truncation function $h(\xi)$ is of the form $\xi\mathbf{1}_{\{|\xi|\le 1\}}$. We can decompose $f(y)$ into the product $f(y) = f_1(y)f_2(y)$, where

$$f_1(y) = \exp\left\{\frac{1}{2}\langle y, ay\rangle + \langle b, y\rangle - \gamma + \int_{D\cap\{|\xi|\le 1\}} \left(e^{\langle\xi,y\rangle} - 1 - \langle h(\xi), y\rangle\right) m(d\xi)\right\},$$

and

$$f_2(y) = \exp\left(\int_{D\cap\{|\xi|>1\}} \left(e^{\langle\xi,y\rangle} - 1\right) m(d\xi)\right).$$

By Sato [1999][Lemma 25.6], $f_1(y)$ can be extended to an entire function on $\mathbb{C}^d$, and is thus in particular continuous for $y \in S$. Thus, if $f_2(y)$ has property C, also $f(y)$ has property C. But restricted to $\{|\xi| > 1\}$, the Lévy measure $m(d\xi)$ is a *finite* Borel measure, and we can argue as in case (a) that it has property C. Thus also $f(y)$ has property C, and assertion (b) is shown. □

The following results parallels Proposition 1.3 from Section 1:

PROPOSITION 3.16. *The functions ϕ and ψ have the following properties:*

(i) $\phi(0, u) = 0$ and $\psi(0, u) = u$ for all $u \in \mathbb{C}^d$.

(ii) ϕ *and* ψ *enjoy the '**semi-flow property**':*

$$\phi(t+s,u) = \phi(s,u) + \phi(t,\psi(s,u)), \quad \psi(t+s,u) = \psi(t,\psi(s,u)), \tag{3.15}$$

for all $t,s > 0$ *with* $(t+s,u) \in \mathcal{Q}$.

(iii) ϕ *and* ψ *are jointly continuous on* $\mathcal{E}_{\mathbb{C}}^{\circ} \cap \mathcal{Q}$.

(iv) Let $t \geq 0$. *Then the function* $\psi(t,.) : \mathcal{D}_t \to \mathbb{R} : y \to \psi(t,y)$ *has property C (cf. Definition 3.14).*

Proof. Assertion (i) is trivial. For assertion (ii) apply the law of iterated expectation and the Markov property of $(X_t)_{t\geq 0}$:

$$\mathbb{E}^x\left[f_u(X_{t+s})\right] = \mathbb{E}^x\left[\mathbb{E}^x\left[f_u(X_{t+s})\middle|\,\mathcal{F}_t\right]\right] = \mathbb{E}^x\left[\mathbb{E}^{X_t}\left[f_u(X_s)\right]\right] \ , \tag{3.16}$$

where all expectations are finite, since $(t+s,u) \in \mathcal{E}_{\mathbb{C}}$. If, in addition $(t+s,u) \in \mathcal{Q}$, we can use Lemma 3.12 and rewrite (3.16) as

$$\exp\left(\phi(t+s,u) + \langle x, \psi(t+s,u)\rangle\right) = \mathbb{E}^x\left[\exp\left(\phi(s,u) + \langle X_t, \psi(s,u)\rangle\right)\right] = \\ = \exp\left(\phi(s,u) + \phi(t,\psi(s,u)) + \langle x, \psi(t,\psi(s,u))\rangle\right) \ ,$$

yielding equations (3.15). Assertion (iii) follows by combining Lemma 3.12.a and 3.12.b. For assertion (iv) consider the function $f : \mathcal{D}_t \to \mathbb{R} : y \mapsto P_t f_y(x)$. We can write $f(y)$ as $\mathbb{E}^x\left[e^{\langle y, X_t\rangle}\right] = \int_D e^{\langle y,\xi\rangle}\,\mu(x,d\xi)$, where $\mu(x,d\xi)$ is the transition kernel of the Markov process $(X_t)_{t\geq 0}$. By Lemma 3.15, $f(y)$ has property C. It follows immediately from the exponentially-affine form (3.12) of $f(y)$ that also $\psi(t,.)$ has property C. □

The next Lemma establishes the differentiability of $\phi(t,u)$ and $\psi(t,u)$ in u:

Lemma 3.17. *For all* $i \in \{1,\ldots,d\}$ *the derivatives*

$$\frac{\partial}{\partial u_i}\phi(t,u), \qquad \frac{\partial}{\partial u_i}\psi(t,u)$$

exist and are continuous for $(t,u) \in \mathcal{E}_{\mathbb{C}}^{\circ} \cap \mathcal{Q}$.

Proof. Let $[0,T] \times K_{\mathbb{C}}$ be a compact subset of $\mathcal{E}_{\mathbb{C}}^{\circ}$. By definition of $\mathcal{E}$, clearly $K \subseteq \mathcal{D}_T^{\circ}$. Choose an $\epsilon > 0$ small enough such that an ϵ-enlargement of K is still in $\mathcal{D}_T^{\circ}$. For any $x \in \mathbb{R}$ and $\epsilon > 0$ it holds that

$$\epsilon|x| \leq \frac{e^{\epsilon x} + e^{-\epsilon x}}{2} \ ,$$

such that

$$\left|\frac{\partial}{\partial u_i}\exp\left(\langle u, X_t\rangle\right)\right| = \left|X_t^i\right| \cdot \exp\left(\langle \operatorname{Re} u, X_t\rangle\right) \leq \tag{3.17}$$

$$\leq \frac{1}{2\epsilon}\Big\{\exp\big(\langle \operatorname{Re} u + \epsilon e_i, X_t\rangle\big) + \exp\big(\langle \operatorname{Re} u - \epsilon e_i, X_t\rangle\big)\Big\}\,.$$

Taking expectations, the right side equals $f(y,z) := \frac{1}{2\epsilon}\left(\mathbb{E}^x\left[e^{\langle y, X_t\rangle}\right] + \mathbb{E}^x\left[e^{\langle z, X_t\rangle}\right]\right)$, evaluated at $y = \operatorname{Re} u + \epsilon e_i$ and $z = \operatorname{Re} u + \epsilon e_i$. But $f(y,z)$ is by Proposition B.4 continuous on $\mathcal{D}_T^\circ$, and thus bounded on any compact subset. For (t,u) in $[0,T]\times K_{\mathbb{C}}$ it follows that $X_t^i \exp(\langle u, X_t\rangle$ is uniformly integrable , and thus that the derivative $\frac{\partial}{\partial u_i}\mathbb{E}^x\left[e^{\langle u, X_t\rangle}\right]$ exists and is jointly continuous. Since T and K were arbitrary the result extends to $\mathcal{E}_{\mathbb{C}}^\circ$. If in addition $(t,u) \in \mathcal{Q}$, then the exponentially-affine form of $\mathbb{E}^x\left[e^{\langle u, X_t\rangle}\right]$ yields the claim for $\phi(t,u)$ and $\psi(t,u)$. □

Following an idea from Montgomery and Zippin [1955], that has later been adapted to semi-flows by Filipović and Teichmann [2003], we can now use the semi-flow property of $\psi(t,u)$ to transfer the differentiability in the state parameter u to the semi-group parameter t:

THEOREM 3.18. *Let $(X_t)_{t\geq 0}$ be an analytic affine process. Then $(X_t)_{t\geq 0}$ is regular, the derivatives*

$$F(u) = \left.\frac{\partial}{\partial t}\phi(t,u)\right|_{t=0+} \quad \textit{and} \quad R(u) = \left.\frac{\partial}{\partial t}\psi(t,u)\right|_{t=0+}$$

exist for all $u \in \mathcal{D}_{\mathbb{C}}$, and are analytic on $\mathcal{D}_{\mathbb{C}}^\circ$. Moreover, if $(t_*, u) \in \mathcal{E}_{\mathbb{C}}$, then for all $t \in [0, t_*]$, $\phi(t,u)$ and $\psi(t,u)$ solve the generalized Riccati equations

$$\frac{\partial}{\partial t}\phi(t,u) = F(\psi(t,u)), \quad \phi(0,u) = 0 \tag{3.18a}$$

$$\frac{\partial}{\partial t}\psi(t,u) = R(\psi(t,u)), \quad \psi(0,u) = u\,. \tag{3.18b}$$

If the solution $\psi(t,u)$ stays in $\mathcal{D}_{\mathbb{C}}^\circ$, it is unique.

PROOF OF THEOREM 3.18, PART 1. To show regularity , we can repeat the proof of Theorem 3.3 almost literally: Again we combine $\phi(t,u)$ and $\psi(t,u)$ into the 'big semi-flow' $\Upsilon(t,u)$: We introduce $\widehat{\mathcal{Q}} := \mathcal{Q} \times \mathbb{C}$, $\widehat{\mathcal{D}} := \mathcal{D} \times \mathbb{R}$, and define $\Upsilon : \widehat{\mathcal{Q}} \to \mathbb{C}^{d+1}$ by

$$(t, u_1, \ldots u_d, u_{d+1}) \mapsto \begin{pmatrix} \psi(t,(u_1,\ldots,u_d)) \\ \phi(t,(u_1,\ldots,u_d)) + u_{d+1}\end{pmatrix}. \tag{3.19}$$

Υ satisfies the semi-flow property $\Upsilon(t+s,u) = \Upsilon(t,\Upsilon(s,u))$ for all $(t+s,u) \in \widehat{\mathcal{Q}}$, and $\Upsilon(0,u) = u$ for all $u \in \widehat{\mathcal{D}}_{\mathbb{C}}$. Moreover, by Lemma 3.17 the Jacobian $\frac{\partial \Upsilon}{\partial u}(t,u)$ exists and is continuous for (t,u) in $\widehat{\mathcal{E}}_{\mathbb{C}}^\circ \cap \widehat{\mathcal{Q}}$. Using exactly the same calculations as in the proof of Theorem 3.3 we arrive at the equality

$$\lim_{t\downarrow 0}\frac{\Upsilon(t,u)-u}{t}=\frac{\Upsilon(s,u)-u}{s}\cdot M(s,u)^{-1}\,, \tag{3.20}$$

valid for all $u\in\widehat{\mathcal{D}}^\circ_{\mathbb{C}}$ and small enough s, and with $M(s,u):=\frac{1}{s}\int_0^s\frac{\partial\Upsilon}{\partial u}(r,u)\,dr$. As before, the non-singularity of $M(s,u)$ is guaranteed by the fact that $\lim_{s\to 0}M(s,u)=I$ (the identity matrix). Moreover, by Lemma 3.17, $M(s,u)$ is continuous on $\widehat{\mathcal{E}}^\circ_{\mathbb{C}}\cap\widehat{\mathcal{Q}}$, such that also the left hand side of (3.20) is. But s can be chosen arbitrarily small, and it follows that the derivatives

$$F(u):=\frac{\partial}{\partial t}\phi(t,u)\Big|_{t=0}\,,\qquad R(u):=\frac{\partial}{\partial t}\psi(t,u)\Big|_{t=0} \tag{3.21}$$

exist for all $u\in\mathcal{D}^\circ_{\mathbb{C}}$ and are continuous functions there. It remains to show that (3.21) is also valid for all $u\in\mathcal{D}_{\mathbb{C}}$. Let $u\in\mathcal{D}^\circ_{\mathbb{C}}$, $t_n\downarrow 0$, $x\in D$, and rewrite (3.21) as

$$\begin{aligned}F(u)+\langle x,R(u)\rangle &=\lim_{n\to\infty}\frac{\exp\left(\phi(t_n,u)+\langle x,\psi(t_n,u)-u\rangle\right)-1}{t_n}=\\ &=\lim_{n\to\infty}\frac{f_{-u}(x)\,P_{t_n}f_y(x)-1}{t_n}=\lim_{n\to\infty}\frac{1}{t_n}\left\{\int_{D-x}e^{\langle x,u\rangle}\,\widetilde{p}_t(x,d\xi)-1\right\}\,,\end{aligned} \tag{3.22}$$

where $\widetilde{p}_t(x,d\xi)$ is the 'shifted transition kernel' of the Markov process $(X_t)_{t\geq 0}$ (see Page 22f for details). The right hand side of (3.22) can be regarded as a limit of (extended) cumulant generating functions of (infinitely divisible) sub-stochastic measures. That is, there exist infinitely divisible measures $\mu_n(x,d\xi)$, with Lévy measures $\nu_n(x,d\xi)$ supported on $D-x$, such that

$$\exp\left(F(u)+\langle x,R(u)\rangle\right)=\lim_{n\to\infty}\int_{\mathbb{R}^d}e^{\langle u,\xi\rangle}\,\mu_n(x,d\xi),\quad\text{for all } u\in\mathcal{D}^\circ_{\mathbb{C}}. \tag{3.23}$$

Let now $\theta\in\mathcal{D}^\circ$, and consider the exponentially tilted measures $\widetilde{\mu}_n(x,d\xi)=e^{\langle\theta,\xi\rangle}\,\mu_n(x,d\xi)$. They are also infinitely divisible, and their characteristic functions converge to $\exp\left(F(u+\theta)+\langle x,R(u+\theta)\rangle\right)$, a function that is continuous in a neighborhood of 0. Thus, by Lévy's continuity theorem and Sato [1999, Theorem 8.7], there exists an infinitely divisible sub-stochastic measure $\widetilde{\mu}(x,d\xi)$ with Lévy measure supported on $D-x$, such that $\widetilde{\mu}_n(x,d\xi)\to\widetilde{\mu}(x,d\xi)$ weakly. On the other hand, $\mu_n(x,d\xi)$ has a vaguely convergent subsequence, which converges to some measure $\mu(x,d\xi)$ on D. By the uniqueness of the weak limit we conclude that $\mu(x,d\xi)=e^{\langle-\theta,\xi\rangle}\widetilde{\mu}(x,d\xi)$. Again, also $\mu(x,d\xi)$ must be infinitely divisible, with Lévy measure supported on $D-x$. Summing up, we have shown that we can write for any $x\in D$, and $u\in\mathcal{D}^\circ_{\mathbb{C}}$,

$$\exp\left(F(u)+\langle x,R(u)\rangle\right)=\int_{\mathbb{R}^d}e^{\langle u,\xi\rangle}\,\mu(x,d\xi)\,, \tag{3.24}$$

where $\mu(x,d\xi)$ is an infinitely divisible measure, with Lévy measure supported on $D-x$. Clearly, the right hand side of (3.24) makes also sense for $u\in\mathcal{D}_{\mathbb{C}}$, such that we can use (3.24) to extend the functions $F(u)$ and $R(u)$ from $\mathcal{D}_\mathbb{C}^\circ$ to $\mathcal{D}_\mathbb{C}$. We have to show, however, that (3.21) remains valid, for the extended functions F and R. Note that it follows from (3.24) and Lemma 3.15 that the extended functions $F(y)$ and $R(y)$, as functions of a real parameter $y \in \mathcal{D}$, satisfy Property C (cf. Definition 3.14). Let now $y_J \in \mathbb{R}^n$ be fixed, such that $(0, y_J) \in \mathcal{D}$, let $(y_I^k)_{k\in\mathbb{N}} \in (-\infty, 0)^m$ be a sequence converging to 0, and let $\lambda_k \to 0$ be a sequence with values in $[0,1]$. Writing $y^k = (y_I^k, \lambda_k y_J)$, it holds that

$$(3.25) \quad \int_0^t R(\psi(s,0))\,ds = \int_0^t R(\lim_{k\to\infty} \psi(s,y^k))\,ds =$$
$$= \int_0^t R\Big(\lim_{k\to\infty}(\psi_I(s,y^k), \lambda_k e^{\beta s} y_J)\Big)\,ds = \lim_{k\to\infty}\int_0^t R\Big(\psi_I(s,y^k), \lambda_k e^{\beta s} y_J\Big)\,ds =$$
$$= \lim_{k\to\infty}\int_0^t \frac{\partial}{\partial s}\psi(s,y^k)\,ds = \lim_{k\to\infty}\psi(t,y^k) = \psi(t,0)\ ,$$

where we have used the continuity of $R(y)$ along the sequence $(\psi_I(t,y^k), \lambda_k e^{\beta t} y_J)$. Differentiating (3.25) on both sides with respect to t shows now that $\frac{\partial}{\partial t}\psi(t,0)\big|_{t=0}$ is well-defined and equals $R(0)$. A similar argument applied to ϕ, shows that also $\frac{\partial}{\partial t}\phi(t,0)\big|_{t=0}$ exists and equals $F(0)$. Thus (3.21) holds also at $u = 0$. Running through the chain of equations (3.22) then shows that (3.23) remains valid for $u = 0$, and thus that the measures $\mu_n(x, d\xi)$ converge *weakly* to $\mu(x, d\xi)$. Dominated convergence then allows us to extend (3.23) to all $u \in \mathcal{U}$, showing that $(X_t)_{t\geq 0}$ is regular. □

LEMMA 3.19. *It holds that $\mathcal{Q} = \mathcal{E}$. In particular the decomposition* (3.12) *is valid for all* $(t, u) \in \mathcal{E}$.

PROOF. By Duffie et al. [2003, Section 7] the law of a regular affine process is infinitely divisible for every $t \in \mathbb{R}_{\geqslant 0}$ under all $\mathbb{P}^x$ $(x \in D)$. But the extended moment generating function of an infinitely divisible distribution has no zeroes by Lemma C.3. It follows that $P_t f_u(0) \neq 0$ for all $(t, u) \in \mathcal{E}$ and thus that $\mathcal{Q} = \mathcal{E}$. □

Having obtained the differentiability of $\phi(t, u)$ and $\psi(t, u)$ at zero, we can extend the generalized Riccati equations to $\mathcal{E}$:

PROOF OF THEOREM 3.18, PART 2. Using Lemma 3.19, the Riccati equations can be derived by differentiating the semi-flow equations (3.15). The representation (3.24), together with Proposition B.4, shows that $F(u)$ and $R(u)$ are analytic in $\mathcal{D}_0^\circ$. Standard results for complex ODEs (cf. Walter [1996]) then yield the assertion on uniqueness of the Riccati solutions. □

PROPOSITION 3.20. *Suppose $(X_t)_{t\geq 0}$ is analytic affine with real domain $\mathcal{D}$.*

(a) If $0 \in \mathcal{D}^\circ$ and $F(0) = R(0) = 0$, then $(X_t)_{t\geq 0}$ is conservative.

(b) Let $i \in \{1, \dots, d\}$ *and suppose that* $(X_t)_{t \geq 0}$ *is conservative. If* $e_i \in \mathcal{D}^\circ$*, and* $F(e_i) = R(e_i) = 0$*, then* $(\exp(X_t^i))_{t \geq 0}$ *is a* $\mathbb{P}^x$*-martingale for all* $x \in D$*.*

Proof. Since $0 \in \mathcal{D}^\circ$ and $F(0) = R(0) = 0$, the generalized Riccati equations (3.18) evaluated for $u = 0$ have the unique solutions $\phi(t, 0) = \psi(t, 0) = 0$. Thus

$$\mathbb{P}^x [X_t \in D] = \mathbb{E}^x \left[e^{\langle 0, X_t \rangle} \right] = \exp \left(\phi(t, 0) + \langle x, \psi(t, 0) \rangle \right) = 1 \ ,$$

for all $t \geq 0$, showing that $(X_t)_{t \geq 0}$ is conservative.
Similarly, for $e_i \in \mathcal{D}^\circ$ and $F(e_i) = R(e_i) = 0$, the generalized Riccati equations (3.18) evaluated for $u = e_i$ have the unique solutions $\phi(t, e_i) = 0$, $\psi(t, e_i) = e_i$. Thus, for $t, h \geq 0$

$$\mathbb{E}^x \left[\exp \left(X_{t+h}^i \right) \middle| \mathcal{F}_t \right] = \exp \left(\phi(h, e_i) + \langle X_t, \psi(h, e_i) \rangle \right) = \exp(X_t^i) \ ,$$

showing that $(\exp(X_t^i))_{t \geq 0}$ is a $\mathbb{P}^x$-martingale for all $x \in D$. □

4. Elementary operations on affine processes

The goal of this section is to develop a 'calculus for affine processes', by which we understand a collection of theorems that turn operations on regular affine processes into simple transformations of its functional characteristics $F(u)$ and $R(u)$. Many of the presented results could be derived as special cases of theorems that hold in a more general setting (e.g. for semi-martingales). However, concentrating on the class of affine processes we are able to give considerably simpler, and in many cases truly 'elementary' proofs of the results. Contrary to the preceding sections, where the filtration on the underlying space Ω played only a minor role, it will appear more prominently in the subsequent results, and may in particular be different from the natural filtration. We will also relax the notational conventions from the preceding chapters in the following way:

(1) We do not require anymore that the state space D of an affine process is of the canonical form $\mathbb{R}^m_{\geqslant 0} \times \mathbb{R}^n$, but rather it can be any permutation of components such as for example $\mathbb{R}_{\geqslant 0} \times \mathbb{R}^n \times \mathbb{R}^{(m-1)}_{\geqslant 0}$.
(2) We will sometimes use superscripts instead of subscripts to denote projection (e.g. X^K_t instead of $(X_t)_K$).

4.1. Preliminary Results. The following proposition is a slight modification of Definition 1.1. The crucial part is that the Markov property of $(X_t)_{t\geq 0}$ is not assumed, but rather a consequence of the other assumptions that are made:

Proposition 4.1. *Let $(X_t)_{t\geq 0}$ be an adapted process on a filtered space $(\Omega, \mathcal{F}_t)$, taking values in $D = \mathbb{R}^m_{\geqslant 0} \times \mathbb{R}^n$. Let $(\mathbb{P}_x)_{x\in D}$ be a family of probability measures on Ω, and suppose that $(X_t)_{t\geq 0}$ is $\mathbb{P}^x$-stochastically continuous[10] for all $x \in D$. Suppose furthermore that there exist functions $\phi(t,u)$ and $\psi(t,u)$, such that for each $(t,u) \in \mathbb{R}_{\geqslant 0} \times \mathcal{U}^\circ$, $h \geq 0$ and $x \in D$*

$$\mathbb{E}^x\left[e^{\langle u, X_{t+h}\rangle}\middle|\,\mathcal{F}_t\right] = \exp(\phi(h,u) + \langle X_t, \psi(h,u)\rangle \ . \tag{4.1}$$

Then $(X_t)_{t\geq 0}$ is an affine process.

Proof. Comparing the above proposition with the definition of an affine process, we see that we have to show that $(X_t, \mathbb{P}^x)_{t\geq 0, x\in D}$ is a Markov process, and that (4.1) can be extended from $u \in \mathcal{U}^\circ$ to $\mathcal{U}$. Consider first (4.1), with $t = 0$. By dominated convergence, the left hand side is well-defined and continuous on $u \in \mathcal{U}$ for every $x \in D$, $h > 0$. Thus also the functions $\phi(h,u)$ and $\psi(h,u)$ have continuous extensions from $\mathcal{U}^\circ$ to $\mathcal{U}$, such that (4.1) holds. Let us denote these extensions also by ϕ and ψ respectively.

[10] Of course we also assume that $X_0 = x$, $\mathbb{P}^x$-almost surely, for all $x \in D$.

Clearly, (4.1) is equivalent to

$$\mathbb{E}^x\left[\, f_u(X_{t+h})\right|\mathcal{F}_t] = \mathbb{E}^{X_t}\left[f_u(X_h)\right] \ , \tag{4.2}$$

for all $u \in \mathcal{U}$, which in turn is equivalent to

$$\mathbb{E}^x\left[f_u(X_{t+h})\,\mathbf{1}_A\right] = \mathbb{E}^x\left[\mathbb{E}^{X_t}\left[f_u(X_h)\right]\mathbf{1}_A\right] \ , \tag{4.3}$$

for all $u \in \mathcal{U}$ and $A \in \mathcal{F}_t$. To show the Markov property of $(X_t)_{t\geq 0}$, it suffices to show that (4.3) holds with f_u replaced by any bounded, Borel-function f.

We will need that for any $r' > 0$ it holds that

$$\lim_{r\to\infty}\sup_{|x|\leq r'}\mathbb{E}^x\left[|X_h| > r\right] = 0 \ . \tag{4.4}$$

Denoting the law of X_h under $\mathbb{P}^x$ by $\mu_{h,x}(d\xi)$ this assertion is equivalent to the tightness of the family $(\mu_{h,x}(d\xi))_{|x|\leq r'}$. By Kallenberg [1997, Lemma 5.2], a family of measures is tight if and only if the family of their characteristic functions is equi-continuous at zero. But obviously $\sup_{|x|\leq r'}\exp\left(\phi(h,u) + \langle x, \psi(h,u)\rangle\right)$ is a continuous function for $u \in \mathcal{U}$, and thus the family $(\mu_{t,x}(d\xi))_{|x|\leq r'}$ is tight.
We continue with a standard approximation argument taken from Kallenberg [1997, Theorem 5.3]: Define $\Theta := \{f_u : u \in \mathcal{U}\}$, and let $\mathcal{L}(\Theta)$ be the set of all (complex) linear combinations of elements of Θ. Let g be a continuous function on D, bounded by $\|g\|_\infty = m$. Hold $h, t \geq 0$ and $x \in D$ fixed and let $\epsilon > 0$. Choose r' such that $\mathbb{P}^x\left(|X_t| > r'\right) < \epsilon$ and $\mathbb{P}^x\left(|X_{t+h}| > r'\right) < \epsilon$. By (4.4) we can also find $r > 0$ such that $\sup_{|x|<r'}\mathbb{P}^x\left(|X_h| > r\right) < \epsilon$, and assume without loss of generality that $r \geq r'$. Denote by g_r the restriction of g to $\{x \in D : |x| \leq r\}$, and extend g_r to a continuous function $\widetilde{g}$ on D with $\|g\|_\infty \leq m$ and period $2\pi r$ in each coordinate. The set $\mathcal{L}(\Theta)$ contains in particular all linear combinations of the functions $\sin(kx/r)$ and $\cos(kx/r)$ with $k \in \mathbb{Z}$. As a well-known consequence of the Weierstrass approximation theorem these functions approximate $\widetilde{g}$ *uniformly*, i.e. we can find $h \in \mathcal{L}(\Theta)$ such that $\|\widetilde{g} - h\|_\infty \leq \epsilon$. Thus we have that

$$\begin{aligned}
&\left|\mathbb{E}^x\left[g(X_{t+h})\mathbf{1}_A\right] - \mathbb{E}^x\left[\mathbb{E}^{X_t}\left[g(X_h)\right]\mathbf{1}_A\right]\right| \leq \\
&\leq \mathbb{E}^x\left[|g(X_{t+h}) - h(X_{t+h})|\,\mathbf{1}_A\right] + \mathbb{E}^x\left[\mathbb{E}^{X_t}\left[|g(X_h) - h(X_h)|\right]\mathbf{1}_A\right] \leq \\
&\leq \mathbb{P}^x\left[|X_{t+h}| > r\right]2m + \epsilon + \mathbb{P}^x\left[|X_t| > r\right]2m + \sup_{|x|\leq r}\mathbb{P}^x\left[|X_h| > r\right]2m + \epsilon \leq \\
&\leq (6m+2)\epsilon \ .
\end{aligned}$$

Since ϵ was arbitrary, this shows that (4.3) holds with f_u replaced by any bounded continuous function g. Pointwise monotone approximation extends the assertion to all bounded Borel functions. □

We present some basic results on the generalized Riccati equations that will be needed. The first three results can be found, along with their proofs, in Duffie et al. [2003, Section 6].

LEMMA 4.2. *Let $\psi(t,u)$ be a solution of the generalized Riccati equation* (2.2b), *and let $i \in I$. Then there exists a constant C_i, independent of (t,u), such that*

$$\operatorname{Re}\psi_i(t,u) \leq g_i(t,u), \quad \text{for all } t \geq 0,$$

where $g_i(t,u)$ is the solution of

$$\frac{\partial}{\partial t} g_i(t,u) = C_i\left(g_i^2 - g_i\right), \quad g_i(0,u) = \operatorname{Re} u_i \,. \tag{4.5}$$

LEMMA 4.3. *Let $\psi(t,u)$ be a solution of the generalized Riccati equation* (2.2b). *Then there exists a constant C, such that*

$$|\psi_I(t,u)|^2 \leq \left(|u_I|^2 + C\int_0^t \left(1+\left|e^{\beta s}u_J\right|^2\right)ds\right)\cdot \exp\left(C\int_0^t\left(1+\left|e^{\beta s}u_J\right|^2\right)ds\right)\,. \tag{4.6}$$

LEMMA 4.4. *For each $u \in \mathcal{U}^\circ$, there exists a unique solution $(\phi(t,u),\psi(t,u))$ of the generalized Riccati equation* (2.2).

DEFINITION 4.5. A function $f : S \subseteq \mathbb{R}^m \to \mathbb{R}^m$ is called quasimonotone increasing, if

$$y,z \in S;\; y \leq z;\; \text{and } y_i = z_i \qquad \text{implies} \qquad f_i(y) \leq f_i(z)\,.$$

LEMMA 4.6. *Let $(X_t)_{t\geq 0}$ be analytic affine with real domain $\mathcal{D}$. For y_J fixed, the function $y_I \mapsto R_I(y_I,y_j)$ is quasimonotone increasing on $\{y_I \in \mathbb{R}^m : (y_I,y_J)\in\mathcal{D}\}$.*

PROOF. Taking into account the admissibility conditions, we can write $R_i(y_I,y_J)$ – with $i \in I$ – as

$$\begin{aligned} R_i(y_I,y_J) &= \frac{1}{2}y_i^2\alpha_{ii}^i + y_i\left\langle \alpha_{iJ}^i, y_J\right\rangle + \frac{1}{2}\left\langle y_J, \alpha_{JJ}^i y_J\right\rangle + \\ &+ y_i\beta_i^i + \left\langle \beta_{I\setminus\{i\}}^i, y_{I\setminus\{i\}}\right\rangle - \gamma^i + \\ &+ \int_D\left(e^{\langle\xi,y\rangle} - 1 - \sum_{k\in J\cup\{i\}}\frac{\xi_k}{1+\xi_k^2}y_k\right)\mu^i(d\xi)\,. \end{aligned}$$

Assume now that $y \leq z$ with $y_i = z_i$. It follows that

$$\begin{aligned} R_i(z_I,y_J) - R_i(y_I,y_J) &= \left\langle \beta_{I\setminus\{i\}}^i, \left(z_{I\setminus\{i\}} - y_{I\setminus\{i\}}\right)\right\rangle + \\ &+ \int_D e^{\langle\xi,y\rangle}\left(\exp\left(\left\langle \xi_{I\setminus\{i\}}, z_{I\setminus\{i\}} - y_{I\setminus\{i\}}\right\rangle\right) - 1\right)\mu^i(d\xi)\,. \end{aligned}$$

Since $\beta_{I\setminus\{i\}}^i \geq 0$ it follows that $y_I \mapsto R_I(y_I,y_J)$ is quasimonotone increasing. ☐

4.2. Projection and independent products. A very simple operation is to form the cartesian product of two independent affine processes. It is not too surprising that the resulting process is again an affine process, and also that regularity is preserved:

PROPOSITION 4.7. *Let $(X_t, \mathbb{P}^x)_{(t\geq 0, x\in D_X)}$ and $(Y_t, \mathbb{P}^y)_{(t\geq 0, y\in D_Y)}$ be two independent affine processes with state spaces D_X and D_Y respectively, defined on a (common) filtered space $(\Omega, \mathcal{F}_t)$. Setting $\mathbb{P}^{(x,y)} := \mathbb{P}^x \otimes \mathbb{P}^y$, it holds that $(X_t, Y_t)_{t\geq 0}$ is an affine process on $(\Omega, \mathcal{F}_t)$ taking values in $D := D_X \times D_Y$, with respect to the probability measures $\left(\mathbb{P}^{(x,y)}\right)_{(x,y)\in D}$.*
If $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ are regular with functional characteristics $(F_X(u), R_X(u))$ and $(F_Y(w), R_Y(w))$ respectively, then also $(X_t, Y_t)_{t\geq 0}$ is regular, with functional characteristics

$$\widetilde{F}(u,w) = F_X(u) + F_Y(w), \qquad \widetilde{R}(u,w) = \begin{pmatrix} R_X(u) \\ R_Y(w) \end{pmatrix}. \tag{4.7}$$

If $\mathcal{D}_X$ and $\mathcal{D}_Y$ are the real domains of $(X_t)_{t\geq 0}$ and $(Y_t)_{t\geq 0}$ respectively, then $\mathcal{D}_X \times \mathcal{D}_Y$ is the real domain of the combined process.

PROOF. For all $(u,w) \in \mathcal{U}_X \times \mathcal{U}_Y$ and $t, h \in \mathbb{R}_{\geqslant 0}$, it holds that

$$\begin{aligned} \mathbb{E}^{(x,y)}\left[f_{(u,w)}(X_{t+h}, Y_{t+h}) \middle| \mathcal{F}_t \right] &= \mathbb{E}^{(x,y)}\left[f_u(X_{t+h}) \cdot f_w(Y_{t+h}) \middle| \mathcal{F}_t \right] = \\ &= \mathbb{E}^{x}\left[f_u(X_{t+h}) \middle| \mathcal{F}_t \right] \cdot \mathbb{E}^{y}\left[f_w(Y_{t+h}) \middle| \mathcal{F}_t \right] = \\ &= \exp\left(\phi_X(h,u) + \phi_Y(h,w) + \langle X_t, \psi_X(h,u)\rangle + \langle Y_t, \psi_Y(h,u)\rangle \right). \end{aligned} \tag{4.8}$$

By proposition 4.1 we conclude that $(X_t, Y_t)_{t\geq 0}$ is an affine process on $D_X \times D_Y$ with respect to $\mathbb{P}^{(x,y)}$. Clearly regularity of X_t and Y_t is preserved, and (4.7) follows by differentiation of (4.8). The real domain $\mathcal{D}_X \times \mathcal{D}_Y$ of the combined process, can be derived directly from definition 3.7, using the independence of X_t and Y_t. ☐

Another simple operation on an affine process is to restrict and project it onto a subspace of its state-space. We will only be interested in the case where the resulting process is again an affine process (and in particular Markovian in its own filtration). We are given an affine process $(X_t, \mathbb{P}^x)_{t\geq 0, x\in D}$ with state space D, and some subset of components $K \subset \{1, \ldots, d\}$. We set $\mathbb{P}^{x_K} = \mathbb{P}^{(x_K, 0)}$ and call $(X_t^K, \mathbb{P}^{x_K})_{t\geq 0, x_K\in D_K}$ the **projection of** $(X_t)_{t\geq 0}$ onto K. The following results gives conditions under which the affine property is preserved:

PROPOSITION 4.8. *Let $(X_t, \mathbb{P}^x)_{t\geq 0, x\in D}$ be an affine process and $K \subseteq M$. Suppose that either*
(a) $\psi_{M\setminus K}(t, u_K, 0) = 0$ for all $t \geq 0, u_K \in \mathcal{U}_K^\circ$,
(b) $(X_t)_{t\geq 0}$ is regular, $I \subseteq K$ and $R_{M\setminus K}(u_K, 0) = 0$ for all $u_K \in \mathcal{U}_K^\circ$, or

(c) $(X_t)_{t\geq 0}$ *is analytic affine,* $(y_K, 0) \in \mathcal{D}^\circ$*, and* $R_{M\setminus K}(y_K, 0) = 0$ *for all* $y_K \in \mathbb{R}$*. Then* $(X_t^K, \mathbb{P}^{x_K})_{t\geq 0, x_K \in D_K}$ *is an affine process in its own filtration. If* $(X_t)_{t\geq 0}$ *is regular with functional characteristics* $(F(u), R(u))$*, then also* $(X_t^K)_{t\geq 0}$ *is regular, with functional characteristics given by*

$$\widetilde{F}(u_K) = F(u_K, 0), \qquad \widetilde{R}(u) = R_K(u_K, 0) \; .$$

If $(X_t)_{t\geq 0}$ *is analytic affine, so is* $(X_t^K)_{t\geq 0}$*; its real domain is given by* $\mathcal{D}_K = \left\{ y_k \in \mathbb{R}^{|K|} : (y_K, 0) \in \mathcal{D} \right\}$*.*

Proof. Let $\mathcal{F}^K$ be the filtration generated by $(X_t^K)_{t\geq 0}$. The tower law and the affine property of $(X_t)_{t\geq 0}$ yield

$$\mathbb{E}^{x_K}\left[\exp\left(\left\langle u_K, X_{t+h}^K \right\rangle\right) \middle| \mathcal{F}_t^K \right] = \mathbb{E}^{x_K}\left[\mathbb{E}^{(x_K,0)}\left[\exp\left(\left\langle u_K, X_{t+h}^K \right\rangle\right) \middle| \mathcal{F}_t \right] \middle| \mathcal{F}_t^K \right] =$$
$$\mathbb{E}^{x}\left[\exp\left(\phi(h, u_K, 0) + \left\langle X_t^K, \psi_K(h, u_K, 0) \right\rangle + \left\langle X_t^{M\setminus K}, \psi_{M\setminus K}(h, u_K, 0) \right\rangle \right) \middle| \mathcal{F}_t^K \right] \; ,$$

for all $t, h \geq 0$ and $u \in \mathcal{U}^\circ$. Consider first case (a): By assumption $\psi_{M\setminus K}(h, u_K, 0) = 0$, such that the above equals

$$\mathbb{E}^{x}\left[\exp\left(\phi(h, u_K, 0) + \left\langle X_t^K, \psi_K(h, u_K, 0) \right\rangle \right) \middle| \mathcal{F}_t^K \right] =$$
$$= \exp\left(\phi(h, u_K, 0) + \left\langle X_t^K, \psi_K(h, u_K, 0) \right\rangle \right) \; ,$$

showing via Proposition 4.1 that X_t^K is an affine process in its own filtration.
To show (b), suppose that $I \subseteq K$, and $R_{M\setminus K}(u_K, 0) = 0$ for all $u_K \in \mathcal{U}_K^\circ$. Then the generalized Riccati equation

$$\frac{\partial}{\partial t}\psi_{M\setminus K}(t, u_K, 0) = R\Big(\psi_K(t, u_K, 0), \psi_{M\setminus K}(t, u_K, 0)\Big), \qquad \psi_{M\setminus K}(0, u_K, 0) = 0$$

is solved by the zero-solution $\widetilde{\psi}_{(M\setminus K)}(t, u_K, 0) \equiv 0$. The condition $I \subseteq K$ guarantees that $(u_K, 0) \in \mathcal{U}^\circ$, such that the solution is unique. Thus $\psi = \widetilde{\psi} = 0$ and the assertion follows as in case (a).
Case (c) can be shown as case (b); only the uniqueness result for the Riccati equation has to be substituted by Theorem 3.18. The real domain of $(X_t^K)_{t\geq 0}$ can be calculated directly from Definition 3.7, using the sets

$$\mathcal{D}_t^K = \left\{ y_K \in \mathbb{R}^{|K|} : \mathbb{E}^x\left[e^{\left\langle X_t^K, y_K \right\rangle} \right] < \infty \; \forall x \in D \right\} = \left\{ y_K \in \mathbb{R}^{|K|} : (y_K, 0) \in \mathcal{D}_t \right\} \; .$$

□

Corollary 4.9 ('Embedded CBI-process'). *Let* $(X_t)_{t\geq 0}$ *be an affine process. Then its non-negative component* $(X_t^I)_{t\geq 0}$ *is a regular affine process (equivalently, a CBI-process) in its own filtration.*

Proof. We apply Proposition 4.8.a with $K = I$ (implying that $M \setminus I = J$). From Proposition 1.9, we know that $\psi_J(t, u_I, 0) = e^{\beta t} \cdot 0 = 0$ for any $u_I \in \mathcal{U}_I$. Thus

$(X_t^I, \mathbb{P}^{x_I})_{t\geq 0}$ is a non-negative process in its own filtration. By Theorem 3.3 it is also regular affine. A non-negative regular affine process is also called a CBI-process (continuously branching with immigration, cf. Kawazu and Watanabe [1971]). □

4.3. Time-integration and the Feynman-Kac formula. It is more interesting to extend the state space by adding components that are *not independent* of the original process. In some important cases the resulting process will still be an affine process. The first such method is to take an affine process $(X_t)_{t\geq 0}$ and to form its time integral $Y_t := \int_0^t X_s\,ds$. Note that the integral is well-defined path-by-path, if we work with a cadlag version of the process $(X_t)_{t\geq 0}$ (which exists by Corollary 1.12). We now consider the joint process $(X_t, Y_t)_{t\geq 0}$ under the family of measures $\mathbb{P}^{(x,0)} = \mathbb{P}^x$, and extend to $\mathbb{P}^{(x,y)}$ in the following way: For each $y \in D$, let θ_y be the 'space-shift' operator, that maps a stochastic process Y_t to its shifted version $y + Y_t$. We define $\mathbb{P}^{(x,y)} = \mathbb{P}^x \circ \theta_y^{-1}$.

THEOREM 4.10. *Let $(X_t, \mathbb{P}^x)_{t\geq 0, x\in D}$ be a cadlag version of a regular affine process on D with functional characteristics (F, R). Define $Y_t = \int_0^t X_s\,ds$, and $\mathbb{P}^{(x,y)} = \mathbb{P}^x \circ \theta_y^{-1}$. Then $(X_t, Y_t)_{t\geq 0}$ is a regular affine process on D^2 under $\mathbb{P}^{(x,y)}$. Its functional characteristics $(\widetilde{F}, \widetilde{R})$ are given by*

$$\widetilde{F}(u_X, u_Y) = F(u_X), \qquad \widetilde{R}(u_X, u_Y) = \begin{pmatrix} R(u_X) + u_Y \\ 0 \end{pmatrix}.$$

The real domain of $(X_t, Y_t)_{t\geq 0}$ is given by $\mathcal{D} \times \mathbb{R}^d$; if $(X_t)_{t\geq 0}$ is analytic affine, then so is $(X_t, Y_t)_{t\geq 0}$.

It would likely be possible to approach this result through general semigroup theory and to use Trotter's product formula (see e.g. Ethier and Kurtz [1986, Corollary 6.7]) which characterizes the semigroup arising from adding the generators of two strongly continuous contraction semigroups. In the case of a regular affine process, however, there is an elementary and more direct proof of the above Theorem, which we state below:

PROOF. We want to calculate the conditional moment generating function

$$\mathbb{E}^{(x,y)}\left[\exp\left(uX_{t+s} + wY_{t+s}\right)\middle|\,\mathcal{F}_t\right] = \\ = \exp\left(w\left(y + \int_0^t X_r\,dr\right)\right)\mathbb{E}^x\left[\exp\left(uX_{t+s} + w\int_t^{t+s} X_r\,dr\right)\middle|\,\mathcal{F}_t\right], \tag{4.9}$$

for $(u, w) \in (\mathcal{U}^\circ)^2$. For each $N \in \mathbb{N}$ and $k \in \{0, \ldots, N\}$, define $s_k^N = \frac{k}{N}s$, such that $s_0^N, \ldots, s_N^N$, is an equispaced partition of $[0, s]$ into intervals of length $h := s/N$. In the following we simplify notation in the following way: s_k stands for s_k^N and wY denotes $\langle w, Y\rangle$. By writing the time-integral as a limit of Riemann sums, and using dominated convergence we have that

$$\mathbb{E}^x\left[\exp\left(uX_{t+s}+w\int_t^{t+s}X_r\,dr\right)\middle|\,\mathcal{F}_t\right]=\lim_{N\to\infty}\mathbb{E}^x\left[\exp\left(uX_{t+s}+wh\sum_{k=0}^{N}X_{t+s_k}\right)\middle|\,\mathcal{F}_t\right].$$

With the shorthand $\Sigma_N := \sum_{k=0}^N X_{t+s_k}$, and using the tower law as well as the affine property of $(X_t)_{t\geq 0}$, the expectation on the right side can be written as

$$\begin{aligned}(4.10)\quad &\mathbb{E}^x\left[\exp(uX_{t+s}+wh\Sigma_N)\middle|\,\mathcal{F}_t\right]=\\ &\mathbb{E}^x\Big[\exp\left(wh\Sigma_{N-1}\right)\cdot\mathbb{E}^x\left[\exp\left(uX_{t+s}+whX_{t+s}\right)\middle|\,\mathcal{F}_{t+s_{N-1}}\right]\Big|\,\mathcal{F}_t\Big]=\\ &\exp\left(\phi(h,u+wh)\right)\mathbb{E}^x\left[\exp(wh\Sigma_{N-1}+\psi(h,u+wh)X_{t+s_{N-1}})\middle|\,\mathcal{F}_t\right].\end{aligned}$$

Applying the tower law $(N-1)$-times in this manner (and conditioning on $\mathcal{F}_{t+s_{N-1}}, \mathcal{F}_{t+s_{N-1}}, \ldots, \mathcal{F}_{t+s_1}$ respectively) we arrive at the equation

$$\mathbb{E}^x\left[\exp\left(uX_{t+s}+wh\sum_{k=0}^{N}X_{t+s_k}\right)\middle|\,\mathcal{F}_t\right]=\exp\left(\widehat{p}_N(s,u,w)+\langle X_t,\widehat{q}_N(s,u,w)\rangle\right),$$

where the quantities $\widehat{p}_N(s,u,w)$ and $\widehat{q}_N(s,u,w)$ are defined through the following recursion:

$$(4.11\text{a})\quad \widehat{p}_0(s,u,w)=0,\qquad \widehat{p}_{k+1}(s,u,w)=\phi(h,\widehat{q}_k(s,u,w))+\widehat{p}_k(s,u,w),$$

$$(4.11\text{b})\quad \widehat{q}_0(s,u,w)=u,\qquad \widehat{q}_{k+1}(s,u,w)=\psi(h,\widehat{q}_k(s,u,w))+wh,$$

for all $k\in\{0,\ldots,N-1\}$. We claim that this recursion is in fact an Euler-like approximation scheme, which – as N tends to infinity – converges to the solutions $p(s,u,w)$ and $q(s,u,w)$ of the generalized Riccati equations

$$(4.12\text{a})\quad \frac{\partial}{\partial t}p(t,u,w)=F\left(q(t,u,w)\right)\qquad p(0,u,w)=0,$$

$$(4.12\text{b})\quad \frac{\partial}{\partial t}q(t,u,w)=R\left(q(t,u,w)\right)+w\qquad q(0,u,w)=u.$$

First, note that $\phi(t,u)$ and $\psi(t,u)$ are, by regularity of $(X_t)_{t\geq 0}$ differentiable in t, with the derivatives at $t=0$ given by $F(u)$ and $R(u)$ respectively. Thus ϕ and ψ admit the expansions

$$\phi(h,u)=h\cdot F(u)+o(h)\quad\text{and}\quad\psi(h,u)=u+h\cdot R(u)+o(h).$$

Inserting into the recursion (4.11), we obtain

$$(4.13\text{a})\quad \widehat{p}_{k+1}(s,u,w)=\widehat{p}_k+h\cdot F(q_k)+o(h)$$

$$(4.13\text{b})\quad \widehat{q}_{k+1}(s,u,w)=\widehat{q}_k+h\cdot\left(R(\widehat{q}_k)+w\right)+o(h).$$

At first sight, this is an Euler-type approximation to the ODE (4.12). However, standard proofs of convergence can not be directly applied, since F and R are not globally Lipschitz, and thus there is no obvious way to control the $o(h)$ terms uniformly for all $k \in \{0, \dots, N\}$. Nevertheless, we shall show convergence by a more elaborate argument. We modify notation a bit and subsequently write $\widehat{q}(k)$ for $\widehat{q}_k(s,u,w)$. This allows us to use subscripts to denote components, i.e. $\widehat{q}(k)$ decomposes into $(\widehat{q}_I(k), \widehat{q}_J(k))$ corresponding to the non-negative and the real-valued part of the state space respectively. By the admissibility conditions (cf. Definition 2.3) we know that $R_J(u)$ is of the form $R_J(u) = \beta u_J$, where β is a real $n \times n$ matrix. The function $R_J(u)$ satisfies a global Lipschitz condition on $\mathcal{U}^\circ$, such that the convergence of $\widehat{q}_J(N)$ to $q_J(s,u,w)$ as $N \to \infty$ follows by standard results (see e.g. Hairer, Nørsett, and Wanner [1987]). We thus turn our attention to the more delicate case of $\widehat{q}_I$. In showing convergence we follow approach (a) outlined in Hairer et al. [1987, Section II.3]: First estimate the 'local error', i.e. the difference between approximation and exact solution incurred in a single step of the recursion (4.11). Then estimate the 'transported error', i.e. the local error transported along the exact solution of (4.12) to the terminal time s. Finally the global error of the approximation will be bounded by the transported errors summed up over all time steps.

We start with some estimates that make sure that all 'relevant quantities', remain inside some compact subset K of $\mathcal{U}^\circ$. Apart from the approximations $\widehat{q}(k)$, these relevant quantities also include the terms $q(N-k, \widehat{q}(k), w)$, i.e. exact solutions of (4.12), started at an intermediate approximation $\widehat{q}(k)$; such terms are needed to estimate the transported errors. On any compact subset K of $\mathcal{U}^\circ$, F and R are in fact Lipschitz, say with Lipschitz constant L_K. Let $i \in I$. By Lemma 4.2 we can find a function $g_i(t, u_i)$ such that $\operatorname{Re}\psi_i(t,u) \le g_i(t, \operatorname{Re} u_i)$ for all $u \in \mathcal{U}^\circ$ and $t \ge 0$. Thus

$$\operatorname{Re}\widehat{q}_i(k+1) \le g_i(h, \operatorname{Re}\widehat{q}_i(k)) + h \operatorname{Re} w \le g_i(h, \operatorname{Re}\widehat{q}_i(k)) \, . \tag{4.14}$$

By Lemma 4.2 $g_i(t+s, u_i) = g_i(t, g_i(s, u_i))$; applying (4.14) recursively, we obtain

$$\operatorname{Re}\widehat{q}_i(k) \le g_i(kh, \operatorname{Re} u_i) < 0 \, . \tag{4.15}$$

From the differential equation (4.12) we also obtain that $\operatorname{Re} q_i(t,u,w) \le \operatorname{Re}\psi_i(t,u) \le g_i(t, \operatorname{Re} u_i)$, and thus

$$\operatorname{Re} q_i\Big((N-k)h, \widehat{q}(k), w\Big) \le g_i(s, \operatorname{Re} u_i) < 0 \, . \tag{4.16}$$

Next we derive an upper bound for $|\widehat{q}_I(k)|^2$: First note that $|\widehat{q}_J(k)|^2$ remains uniformly bounded for all $k \in \mathbb{N}$, say by a constant $K \ge 0$, since $\widehat{q}_J$ is a convergent approximation of the continuous function $\psi_J(s,u)$. Writing $M_h := C\int_0^h \left(1 + e^{\beta s} K \, ds\right)$

we obtain from Lemma 4.3 the estimate

$$|\widehat{q}_I(k)|^2 \leq \Big(|\widehat{q}_I(k-1)|^2 + M_h\Big)\, e^{M_h} + |w|^2 h^2 \,.$$

Applying the estimate recursively, we end up with

$$|\widehat{q}_I(k)|^2 \leq \Big(|u_I|^2 + kM_h\Big)\, e^{kM_h} + |w|^2 h^2 \frac{e^{kM_h}-1}{e^{M_h}-1} \,. \tag{4.17}$$

Letting $h \to 0$ we have that $NM_h \to (1+K)$. Similarly de L'Hospital's rule implies that $\frac{h}{e^{M_h}-1} \to \frac{1}{1+K}$, and thus that (4.17) can be bounded by a constant. We may now conclude from (4.16) and (4.17) that the quantities $\widehat{q}(k)$, as well as $q\Big((N-k)h, \widehat{q}(k), w\Big)$ remain for all $k \in \{1,\ldots,N\}$ inside some compact subset K of $\mathcal{U}^\circ$.

We can now estimate the local error incurred in the j-th step of (4.11), using a Taylor expansion in h:

$$|e(k)| = |\widehat{q}_I(k) - q(h, \widehat{q}(k-1), w)| \leq hD(h, \widehat{q}(k-1)) \tag{4.18}$$

where

$$D(h, \widehat{q}(k-1)) = \sup_{\xi \in (0,h)} |R_I(\psi(\xi, \widehat{q}(k-1))) - R_I(q(\xi, \widehat{q}(k-1), w))| \,. \tag{4.19}$$

Writing $\widetilde{q}(k) := q(h, \widehat{q}(k-1), w)$, and using a standard ODE estimate (cf. Hairer et al. [1987, Thm I.10.2]) the transported error $E(k)$ can be bounded by

$$|E(k)| \;=\; \Big|q\Big((N-k)h, \widetilde{q}(k), w\Big) - q\Big((N-k)h, \widehat{q}(k), w\Big)\Big| \;\leq\; e^{(N-k)hL_K}|e(k)| \;, \tag{4.20}$$

where L_K is the Lipschitz constant of R on K. Finally the global error E satisfies

$$|E| \leq \sum_{k=1}^{N} |E(k)| \leq e^{sL} h \sum_{k=1}^{N} D(h, \widehat{q}(k-1), w) \leq hNe^{sL} \sup_{u \in K} D(h,u) \,.$$

which converges to

$$se^{sL} \lim_{h \to 0} \sup_{u \in K} D(h,u) \quad \text{as} \quad N \to \infty \,.$$

To show that $E \to 0$ as $h \to 0$, we conclude with a compactness argument: $D(h,u)$ is by Proposition 1.3 a jointly continuous function on $\mathbb{R}_{\geqslant 0} \times \mathcal{U}$, that satisfies $D(0,u) = 0$. Thus for each $\epsilon > 0$ and $u \in K$ there exists $h_u > 0$ and a neighborhood N_u of u such that $D(\xi, w) < \epsilon$ for all $(\xi, w) \in [0, h_u] \times N_u$. But K is compact such that the open cover $(N_u)_{u \in K}$ has a finite subcover $N_{u_1}, \ldots, N_{u_M}$. Setting $h_* := \min\{h_{u_1}, \ldots, h_{u_M}\}$ it holds that $\sup_{u \in K} D(\xi, u) < \epsilon$ for all $\xi < h_*$. Since ϵ was arbitrary, the global error E goes to 0 as $N \to \infty$. □

By combining Theorems 3.18 and 4.10, we obtain the following Corollary:

Corollary 4.11 (Feynman-Kac formula for affine processes). *Let* $(X_t)_{t\geq 0}$ *be an analytic affine process with real domain* $\mathcal{D}$. *Let* $t,h \in \mathbb{R}_{\geqslant 0}$, *and* $u,w \in \mathbb{C}$ *with* $\operatorname{Re} u \in \mathcal{D}$. *Suppose that there exist unique solutions, up to time* h, *to the generalized Riccati equations*

$$\frac{\partial}{\partial t}\widetilde{\phi}(t,u,w) = F(\widetilde{\psi}(t,u,w)), \qquad \widetilde{\phi}(0,u,w) = 0$$
$$\frac{\partial}{\partial t}\widetilde{\psi}(t,u,w) = R(\widetilde{\psi}(t,u,w)) + w, \qquad \widetilde{\psi}(0,u,w) = u\ .$$

Then

$$\mathbb{E}^x\left[\exp\left(\langle u, X_{t+h}\rangle + \left\langle w, \int_t^{t+h} X_s\,ds\right\rangle\right)\middle|\,\mathcal{F}_t\right] = \exp\left(\widetilde{\phi}(h,u,w) + \left\langle X_t, \widetilde{\psi}(h,u,w)\right\rangle\right)\ .$$

Remark 4.12. If $\operatorname{Re} u \in \mathcal{D}^\circ$, then uniqueness does not have to be required; rather it follows from Theorem 3.18.

Remark 4.13. The above result is stronger than the general Feynman-Kac formula for Feller processes (cf. Rogers and Williams [1994, III.19]), since there is no requirement on $\left\langle w, \int_t^{t+h} X_s\,ds\right\rangle$ being bounded from below.

4.4. Exponential Tilting.

Theorem 4.14. *Let* $(X_t, \mathbb{P}^x)_{t\geq 0, x\in D}$ *be an analytic affine process with functional characteristics* $(F(u), R(u))$, *defined on the filtered space* $(\Omega, \mathcal{F})$, *where* $\Omega = \mathfrak{D}(\mathbb{R}_{\geqslant 0}, D)$ *is the space of cadlag paths equipped with the Skorohod topology, and* $\mathcal{F}$ *is a right-continuous,* $(\mathbb{P}^x)_{x\in D}$*-complete filtration*[11].
Let θ *be in* $\mathcal{D}^\circ$. *Then there exist measures* $\mathbb{Q}^x \sim \mathbb{P}^x$, *for each* $x \in D$, *such that* $(X_t, \mathbb{Q}^x)_{t\geq 0, x\in D}$ *is an analytic affine process with characteristics*

$$\widetilde{F}(u) = F(u+\theta) - F(\theta), \qquad \widetilde{R}(u) = R(u+\theta) - R(\theta)\ ,$$

and real domain $\mathcal{D} - \theta$. *Moreover for every* $t \in \mathbb{R}_{\geqslant 0}$, $x \in D$, *and* $A \in \mathcal{F}_t$, *it holds that*

$$\mathbb{E}^{Q_x}[A] = \mathbb{E}^{P_x}[AM_t^x]\,, \quad \textit{where}$$
$$M_t^x = \exp\left(\langle \theta, X_t - x\rangle - tF(\theta) - \left\langle R(\theta), \int_0^t X_s\,ds\right\rangle\right)\ ,$$

and M_t^x *is a* $\mathbb{P}^x$*-martingale.*

Following Palmowski and Rolski [2000], an exponential change of measure of the type seen above can be achieved for a *general Markov process* X_t with generator

[11] Note that by Corollary 1.12, every affine process has a version with cadlag paths, such that this assumption is not a big restriction. For completing a filtration with respect to the *family* $(\mathbb{P}^x)_{x\in D}$ of probability measures, see Revuz and Yor [1999, Section III.2].

$\mathcal{A}$, by setting

$$M_t = \frac{h(X_t)}{h(X_0)} \exp\left(-\int_0^t \frac{(\mathcal{A}h)(X_s)}{h(X_s)}\,ds\right) ,$$

for some non-vanishing function h in the domain of $\mathcal{A}$. Our theorem thus can be seen as the special case that $(X_t)_{t\geq 0}$ is regular affine, and $h = f_\theta$. Similar results for affine processes have also been obtained by Kallsen and Muhle-Karbe [2008], using semi-martingale methods. Our proof is original, in the sense that it does not use semi-martingale theory, and follows without much additional effort from the time-integration result Theorem 4.10.

Proof. We start by showing that M_t^x is a $\mathbb{P}^x$-martingale. Let $Y_t = \int_0^t X_s\,ds$. By Theorem 4.10, we know that (X_t, Y_t) is an affine process, and we have that

$$\mathbb{E}^{\mathbb{P}_x}\left[\exp\left(\langle \theta, X_{t+h}\rangle + \langle w, Y_{t+h}\rangle\right) \middle| \mathcal{F}_t\right] = \\ = \exp\left(p(h,\theta,w) + \langle X_t, q(h,\theta,w)\rangle + \langle Y_t, w\rangle\right) , \tag{4.21}$$

where p, q satisfy the generalized Riccati equations

$$\frac{\partial}{\partial t} p(t,\theta,w) = F(q(t,\theta,w)), \quad p(0,\theta,w) = 0 \tag{4.22a}$$

$$\frac{\partial}{\partial t} q(t,\theta,w) = R(q(t,\theta,w)) + w, \quad q(0,\theta,w) = \theta \tag{4.22b}$$

The key to create a martingale out of the process (X_t, Y_t) is to find a w, such that the second Riccati equation has a constant solution. The obvious choice is $w = -R(\theta)$, for which we obtain the solutions $q(t,\theta,-R(\theta)) = \theta$ and $p(t,\theta,-R(\theta)) = tF(\theta)$. Since $\theta \in \mathcal{U}^\circ$ the solution is unique. Inserting into (4.21), and multiplying both sides by $e^{-(t+h)F(\theta)-\langle\theta,x\rangle}$, we see that

$$\mathbb{E}^{\mathbb{P}_x}\left[\exp\left(\langle \theta, X_{t+h} - x\rangle - (t+h)F(\theta) - \langle R(\theta), Y_{t+h}\rangle\right) \middle| \mathcal{F}_t\right] = \\ = \exp\left(\langle \theta, X_t - x\rangle - tF(\theta) - \langle R(\theta), Y_t\rangle\right) .$$

But the left hand side is just $\mathbb{E}[M_{t+h}^x | \mathcal{F}_t]$ and the right hand side M_t^x, such that M_t^x is a martingale.

The martingale property of M_t^x, together with the fact that $\mathbb{E}^{\mathbb{P}^x}[M_t^x] = 1$, and $M_t^x > 0$ ($\mathbb{P}^x$-a.s.) allows us to apply an extension of Kolmogorov's Existence Theorem[12] (see Kallenberg [1997, Lemma 18.18]), which guarantees the existence of measures $\mathbb{Q}^x \sim \mathbb{P}^x$ with the property that $\mathbb{E}^{Q_x}[A] = \mathbb{E}^{P_x}[AM_t^x]$ for all $A \in \mathcal{F}_t$. It remains to show that $(X_t, \mathbb{Q}^x)_{t\geq 0, x\in D}$ is an affine process with the characteristics $\widetilde{F}(u) = F(u+\theta) - F(\theta)$ and $\widetilde{R}(u) = R(u+\theta) - R(\theta)$:

[12]At this point the assumption that $(X_t)_{t\geq 0}$ is defined on Skorohod space equipped with a right-continuous, complete filtration is needed.

Similar as in (4.21), we have that

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}_x}\left[e^{\langle u, X_{t+h}\rangle}\middle|\,\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}_x}\left[e^{\langle u, X_{t+h}\rangle}\frac{M^x_{t+h}}{M^x_t}\middle|\,\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}_x}\left[e^{\langle u, X_{t+h}\rangle}M^x_{t+h}\middle|\,\mathcal{F}_t\right]\frac{1}{M^x_t}\\
&= \mathbb{E}^{\mathbb{P}_x}\left[\exp\left(\langle \theta + u, X_{t+h}\rangle - \langle \theta, x\rangle - (t+h)F(\theta) - \langle R(\theta), Y_{t+h}\rangle\right)\middle|\,\mathcal{F}_t\right]\frac{1}{M^x_t} =\\
&= \mathbb{E}^{\mathbb{P}_x}\left[\exp\left(\langle \theta + u, X_{t+h}\rangle - \langle R(\theta), Y_{t+h} - Y_t\rangle\right)\middle|\,\mathcal{F}_t\right]\exp\left(-hF(\theta) - \langle \theta, X_t\rangle\right) =\\
&= \exp\left(p(h, \theta+u, -R(\theta)) - hF(\theta) + \langle q(t, \theta+u, -R(\theta)) - \theta, X_t\rangle\right) =\\
&= \exp\left(\widetilde{\phi}(t,u) + \left\langle \widetilde{\psi}(t,u), X_t\right\rangle\right)\,,
\end{aligned}$$

where we have set $\widetilde{\phi}(t,u) = p(t, \theta+u, -R(\theta)) - hF(\theta)$ and $\widetilde{\psi}(t,u) = q(t, \theta + u, -R(\theta)) - \theta$. Comparing with (4.22), we see that $\widetilde{\phi}(t,u)$ and $\widetilde{\psi}(t,u)$ satisfy

$$\frac{\partial}{\partial t}\widetilde{\phi}(t,u) = F(\widetilde{\psi}(t,u) + \theta) - F(\theta), \quad \widetilde{\phi}(0,u) = 0 \tag{4.23a}$$

$$\frac{\partial}{\partial t}\widetilde{\psi}(t,u) = R(\widetilde{\psi}(t,u) + \theta) - R(\theta), \quad \widetilde{\phi}(0,u) = u\,, \tag{4.23b}$$

and the proof is completed. □

4.5. Subordination. The last operation we consider is the subordination of a Lévy process $(L_t)_{t\geq 0}$ by a component of an independent affine process $(X_t)_{t\geq 0}$. For better readability we write the Lévy process as $L(t)_{t\geq 0}$. The component X^i_t of the affine process acts as a time-change for the Lévy process, and the subordinated process is defined as $L(X^i_t)_{t\geq 0}$. For this operation to make sense, the component X^i_t of the affine process has to be non-decreasing almost surely. We give a sufficient condition:

Proposition 4.15. *Let $(X_t)_{t\geq 0}$ be a regular affine process and let $i \in I$. If $R_I(-ce_i) \leq 0$ for all $c > 0$, then the $(X^i_t)_{t\geq 0}$ is $\mathbb{P}^x$-almost surely non-decreasing for all $x \in D$.*

Proof. Let $t, h \geq 0$. It holds that

$$\mathbb{E}^x\left[e^{-c(X^i_{t+h} - X^i_t)}\middle|\,\mathcal{F}_t\right] = \exp\left(\phi(h, -ce_i) + \langle X_t, \psi(h, -ce_i) + ce_i\rangle\right)$$

For any $\mathbb{R}$-valued random variable Y we have that $\lim_{c\to\infty}\mathbb{E}\left[e^{-cY}\right] = +\infty\cdot\mathbb{P}(X < 0) + \mathbb{P}(X = 0)$. Thus, if we can show that

$$\phi(h, -ce_i) + \langle x, \psi(h, -ce_i) + ce_i\rangle \leq 0\,, \tag{4.24}$$

for all $c > 0$ and $x \in D$, it holds that $X^i_{t+h} - X^i_t \geq 0$ $\mathbb{P}^x$-a.s. for all $t, h \geq 0$, and thus that X^i_t is non-decreasing. Since $\psi_J(h, -ce_i) = 0$ by Proposition 1.9, and $\phi(h, -ce_i) \leq 0$ by Proposition 1.3, it is clear that

$$\psi_I(h, -ce_i) \leq -ce_i, \qquad \text{for all } c \geq 0$$

already implies (4.24). But $\psi_I(h,-ce_i)$ is the unique solution of the Riccati equation

$$\frac{\partial}{\partial h}\psi_I(h,-ce_i) = R_I(\psi_I(h,-ce_i),0), \quad \psi_I(0,-ce_i) = -ce_i \ .$$

By a comparison principle for quasi-monotonic differential equations (see Volkmann [1972, Satz 2]) $R_I(-ce_i) \le 0$ implies that also $\psi_I(h,-ce_i) \le -ce_i$, and the assertion is shown. □

Again, the affine property will be preserved by the operation of subordination, if we adjoin an additional component to the process X_t. As in Theorem 4.10 we denote by θ_y the space-shift operator, that maps each stochastic process Y_t to its shifted version $y + Y_t$.

THEOREM 4.16. *Let $(X_t)_{t\ge 0}$ be an affine process taking values in D, and let $(L(t))_{t\ge 0}$ be an independent Lévy process under all $(\mathbb{P}^x)_{x\in D}$. Let $i \in I$, and suppose that the i-th component of X_t is non-decreasing $\mathbb{P}^x$-almost surely for all $x \in D$. Define the subordinated process $Y_t = L(X_t^i)$, and for each $y \in \mathbb{R}$, let $\mathbb{P}^{(x,y)} := \mathbb{P}^x \circ \theta_y^{-1}$. Then $\left((X_t,Y_t),\mathbb{P}^{(x,y)}\right)_{t\ge 0,(x,y)\in D\times\mathbb{R}}$ is an affine process on $D\times\mathbb{R}$ in its own filtration. Suppose that in addition $(X_t)_{t\ge 0}$ is regular with functional characteristics $(F(u),R(u))$, and let $m(w)$ be the Lévy exponent[13] of L_t. Then also $(X_t,Y_t)_{t\ge 0}$ is regular affine, with functional characteristics given by*

$$\widetilde{F}(u,w) = F(u+e_i m(w)), \qquad \widetilde{R}_X(u,w) = R(u+e_i m(w)), \quad \widetilde{R}_Y(u,w) = 0 \ .$$

PROOF. Let $\mathcal{F}'$ denote the natural filtration of $(X_t,Y_t)_{t\ge 0}$. To simplify notation we write uX for $\langle u,X\rangle$. We have that

$$\begin{aligned}(4.25)\quad &\mathbb{E}^{(x,y)}\left[\exp\left(uX_{t+h}+wY_{t+h}\right)\middle|\,\mathcal{F}'_t\right] = \\ &= e^{w(y+Y_t)}\mathbb{E}^x\left[\exp\left(uX_{t+h}+w\left(L(X^i_{t+h})-L(X^i_t)\right)\right)\middle|\,\mathcal{F}'_t\right] = \\ &= e^{w(y+Y_t)}\mathbb{E}^x\left[\exp\left(uX_{t+h}+wL\left(X^i_{t+h}-X^i_t\right)\right)\middle|\,\mathcal{F}'_t\right] \ ,\end{aligned}$$

where we have used that X_t^i is non-decreasing and the increments of a Lévy process are stationary. Using the Markov property of $(X_t)_{t\ge 0}$, we can write the right hand side as

$$\begin{aligned}(4.26)\quad &e^{w(y+Y_t)}\mathbb{E}^{X_t}\left[\exp\left(uX_h+wL(X^i_h)\right)\right] = \\ &= e^{w(y+Y_t)}\mathbb{E}^{X_t}\left[\mathbb{E}^{X_t}\left[\exp\left(uX_h+wL(X^i_h)\right)\middle|\,\sigma(X_s)_{0\le s\le h}\right]\right] = \\ &= e^{w(y+Y_t)}\mathbb{E}^{X_t}\left[\exp\left(uX_h+X^i_h\cdot m(w)\right)\right] = \\ &= \exp\left(\phi\Big(t,u+e_i m(w)\Big)+\Big\langle X_h,\psi\Big(t,u+e_i m(w)\Big)\Big\rangle+(y+Y_t)w\right) \ .\end{aligned}$$

[13]The Lévy exponent of $(L_t)_{t\ge 0}$ is defined by $\mathbb{E}^x\left[e^{\langle u,L_t\rangle}\right] = e^{tm(u)}$.

By Proposition 4.1, this shows that $(X_t, Y_t)_{t\geq 0}$ is an affine process in its own filtration with respect to $\left(P^{(x,y)}\right)_{(x,y)\in D\times\mathbb{R}}$. The rest of the Theorem follows directly by differentiation of (4.26). $\square$

Operation	Input	Assumptions	Result
Independent Product (Prop. 4.7)	$(X_t)_{t\geq 0}, (Y_t)_{t\geq 0}$ reg. affine with func. char. $(F_X(u), R_X(u))$, and ; $(F_Y(u), R_Y(u))$ respectively.	–	$(X_t, Y_t)_{t\geq 0}$ reg. affine w.r.t $\mathbb{P}^(x, y) := \mathbb{P}^x \otimes \mathbb{P}^y$; $\widetilde{F}(u_X, u_Y) = F_X(u_X) + F_Y(u_Y)$, $\widetilde{R}(u_X, u_Y) = (R_X(u_X), R_Y(u_Y))$.
Projection (Prop. 4.8)	$(X_t)_{t\geq 0}$ reg. affine with func. char. $(F(u), R(u))$; $K \subseteq M$: components to project onto.	$\psi_{M\setminus K}(t, u_K, 0) = 0$; or $I \subseteq K$ and $R_{M\setminus K}(u_K, 0) = 0$; or $(y_K, 0) \in \mathcal{D}^\circ$ and $R_{M\setminus K}(y_K, 0) = 0$.	X_t^K reg. affine with func. char. $\widetilde{F}(u_K) = F(u_K, 0)$ and $\widetilde{R}(u_K) = R_K(u_K, 0)$.
Time-Integration (Thm 4.10)	$(X_t)_{t\geq 0}$ reg. affine with diff.char. $(F(u), R(u))$.	–	$(X_t, Y_t = \int_0^t X_s ds)$ reg. affine with $\widetilde{F}(u) = F(u)$, $R_X(u) = R(u) + w$, and $R_Y(u) = 0$.
Exponential Tilting (Thm. 4.14)	$(X_t, \mathbb{P}^x)_{t\geq 0, x\in D}$ analyt. affine; func. char. $(F(u), R(u))$ defined on Skorohod space of cadlag paths, with right-cont., complete filtration.	$\theta \in \mathcal{D}^\circ$	$\exists\ \mathbb{Q}^x \sim \mathbb{P}^x$, such that $(X_t, Q^x)_{t\geq 0, x\in D}$ analyt. affine; $\widetilde{F}(u) = F(u+\theta) - F(\theta)$, $\widetilde{R}(u) = R(u+\theta) - R(\theta)$.
Subordination (Thm. 4.16)	$(X_t)_{t\geq 0}$ reg. affine w. $(F(u), R(u))$; $L(t)$ independent Lévy process with char. exponent $m(u)$.	$i \in I$; component X_t^i non-decreasing increasing a.s.;	$(X_t, Y_t = L(X_t^i))$ reg. affine with $\widetilde{F}(u, w) = F(u + e_i m(w))$, $\widetilde{R}_X(u, w) = R(u + e_i m(w))$, $R_Y(u, w) = 0$.

TABLE 2. This table summarizes the results of Section 4

Part 2

Applications to stochastic volatility modelling

5. Affine Stochastic Volatility Models (ASVMs)

5.1. Definition and the generalized Riccati equations. We consider an asset-pricing model of the following kind: The interest rate r is non-negative and constant, and the asset price $(S_t)_{t\geq 0}$ is given by

$$S_t = \exp(rt + X_t) \quad t \geq 0 ,$$

such that $(X_t)_{t\geq 0}$ is the discounted log-price process starting at $X_0 \in \mathbb{R}$ a.s. The discounted price process is simply $\exp(X_t)$, such that we will assume in the remainder that $r = 0$, and that $(S_t)_{t\geq 0}$ is already discounted. Denote by $(V_t)_{t\geq 0}$ another process, starting at $V_0 > 0$ a.s., which can be interpreted as stochastic variance process of $(X_t)_{t\geq 0}$, but may also control the arrival rate of jumps. The following assumptions are made on the joint process $(X_t, V_t)_{t\geq 0}$:

- **A1:** $(X_t, V_t)_{t\geq 0}$ is a stochastically continuous, time-homogeneous Markov process.
- **A2:** The cumulant generating function $\Phi_t(u, w)$ of (X_t, V_t) is of a particular affine form: We assume that there exist functions $\phi(t, u, w)$ and $\psi(t, u, w)$ such that

$$\Phi_t(u, w) := \log \mathbb{E}\left[\exp(uX_t + wV_t)\middle|\, X_0, V_0\right] = \phi(t, u, w) + V_0\psi(t, u, w) + X_0 u$$

  for all $(t, u, w) \in \mathbb{R}_{\geqslant 0} \times \mathbb{C}^2$, where the expectation exists.

By convention, the logarithm above denotes the principal branch of the complex logarithm. Assumptions A1 and A2 make $(X_t, V_t)_{t\geq 0}$ an affine process in the sense of Duffie et al. [2003]. The term $X_0 u$ in the cumulant generating function $\Phi_t(u, w)$ corresponds to a reasonable homogeneity assumption on the model: If the starting value X_0 of the price process is shifted by x, also X_t is simply shifted by x for any $t \geq 0$. Note that Assumption A2 also implies that the variance process $(V_t)_{t\geq 0}$ is a Markov process in its own right. We do not yet make the assumption that $(S_t)_{t\geq 0}$ is conservative (i.e. without explosions or killing) or even a martingale. Instead it will be our first goal in Section 5.2 to obtain necessary and sufficient conditions for these properties.
Applying the law of iterated expectations to $\Phi_t(u, w)$ yields the following 'semi-flow equations' for ϕ and ψ: (see also Duffie et al. [2003, Eq. (3.8)–(3.9)])

$$\begin{aligned} \phi(t + s, u, w) &= \phi(t, u, w) + \phi(s, u, \psi(t, u, w)), \\ \psi(t + s, u, w) &= \psi(s, u, \psi(t, u, w)), \end{aligned} \tag{5.1}$$

for all $t, s \geq 0$ where the left hand side is defined. The following result will be crucial:

Theorem 5.1. *Suppose that* $|\phi(\tau,u,\eta)| < \infty$ *and* $|\psi(\tau,u,\eta)| < \infty$ *for some* $(\tau,u,\eta) \in \mathbb{R}_{\geqslant 0} \times \mathbb{C}^2$. *Then, for all* $t \in [0,\tau]$ *and* $w \in \mathbb{C}$ *with* $\operatorname{Re} w \le \operatorname{Re} \eta$

$$|\phi(t,u,w)| < \infty, \qquad |\phi(t,u,w)| < \infty,$$

and the derivatives

$$F(u,w) := \frac{\partial}{\partial t}\phi(t,u,w)\Big|_{t=0+}, \qquad R(u,w) := \frac{\partial}{\partial t}\psi(t,u,w)\Big|_{t=0+} \tag{5.2}$$

exist. Moreover, for $t \in [0,\tau)$, ϕ *and* ψ *satisfy the generalized Riccati equations*

$$\partial_t \phi(t,u,w) = F(u,\psi(t,u,w)), \quad \phi(0,u,w) = 0 \tag{5.3a}$$

$$\partial_t \psi(t,u,w) = R(u,\psi(t,u,w)), \quad \psi(0,u,w) = w\ . \tag{5.3b}$$

The above theorem is 'essentially' proven in Duffie et al. [2003], but under slightly different conditions[14]. Note that the differential equations (5.3) follow immediately from the semi-flow equations (5.1) by taking the derivative with respect to s, and evaluating at $s = 0$. They are called *generalized Riccati equations* since they degenerate into (classical) Riccati equations with quadratic functions F and R, if $(X_t,V_t)_{t\ge 0}$ is a pure diffusion process.
Note that the first Riccati equation is just an integral in disguise, and ϕ may be written explicitly as

$$\phi(t,u,w) = \int_0^t F(u,\psi(s,u,w))\, ds\ . \tag{5.4}$$

Also the solution ψ of the second Riccati equation can be represented at least implicitly in the following way: Suppose that $\psi(t,u,w)$ is a non-stationary local solution on $[0,\delta)$ of (5.3b). Then $R(u,\psi(t,u,w)) \neq 0$ for all $t \in [0,\delta)$, and $\psi(t,u,w)$ is a strictly monotone function of t; dividing both sides of (5.3b) by $R(u,\psi(t,u,w))$, integrating from 0 to $t < \delta$, and substituting $\eta = \psi(s,u,w)$ yields

$$\int_w^{\psi(t,u,w)} \frac{d\eta}{R(u,\eta)} ds = t\ . \tag{5.5}$$

[14]Duffie et al. assume differentiability of ϕ and ψ with respect to t ('regularity') a priori, while in our case we can deduce it directly from Assumption A2. A proof is given in the appendix.

Another important result that can be found in Duffie et al. [2003] states that F and R must be of Lévy-Khintchine form, i.e.

$$F(u,w) = (u,w)\cdot\frac{a}{2}\cdot\begin{pmatrix}u\\w\end{pmatrix} + b\cdot\begin{pmatrix}u\\w\end{pmatrix} - c + \int_{D\setminus\{0\}}\left(e^{xu+yw} - 1 - \omega_F(x,y)\cdot\begin{pmatrix}u\\w\end{pmatrix}\right) m(dx,dy)\,, \tag{5.6a}$$

$$R(u,w) = (u,w)\cdot\frac{\alpha}{2}\cdot\begin{pmatrix}u\\w\end{pmatrix} + \beta\cdot\begin{pmatrix}u\\w\end{pmatrix} - \gamma + \int_{D\setminus\{0\}}\left(e^{xu+yw} - 1 - \omega_R(x,y)\cdot\begin{pmatrix}u\\w\end{pmatrix}\right) \mu(dx,dy) \tag{5.6b}$$

where $D = \mathbb{R}\times\mathbb{R}_{\geqslant 0}$, and ω_F, ω_R are suitable truncation functions, which we fix by defining

$$\omega_F(x,y) = \begin{pmatrix}\frac{x}{1+x^2}\\0\end{pmatrix} \qquad \text{and} \qquad \omega_R(x,y) = \begin{pmatrix}\frac{x}{1+x^2}\\\frac{y}{1+y^2}\end{pmatrix}\,.$$

Moreover the parameters $(a,\alpha,b,\beta,c,\gamma,m,\mu)$ satisfy the following admissibility conditions:

- a,α are positive semi-definite 2×2-matrices, and $a_{12} = a_{21} = a_{22} = 0$.
- $b \in D$ and $\beta \in \mathbb{R}^2$.
- $c,\gamma \in \mathbb{R}_{\geqslant 0}$
- m and μ are Lévy measures on D, and $\int_{D\setminus\{0\}}\left((x^2+y)\wedge 1\right)\, m(dx,dy) < \infty$.

The affine form of the cumulant generating function, the generalized Riccati equations and finally the Lévy-Khintchine decomposition (5.6) lead to the following interpretation of F and R: F characterizes the state-independent dynamic of the process (X_t,V_t) while R characterizes its state-dependent dynamic. Both F and R decompose into a diffusion part, a drift part, a jump part and an instantaneous killing rate. Hence $a+\alpha V_t$ can be regarded as instantaneous covariance matrix of $(X_t,V_t)_{t\geq 0}$, $b+V_t\beta$ as the instantaneous drift, $m(dx,dy)+V_t\mu(dx,dy)$ as instantaneous arrival rate of jumps with jump heights in $(dx\times dy)$, and finally $c+\gamma V_t$ as the instantaneous killing rate.

The following Lemma establishes some important properties of F and R as functions of *real-valued* arguments. A proof is given in the appendix.

LEMMA 5.2. *(a) F and R are proper closed convex functions on $\mathbb{R}^2$.*
(b) F and R are analytic in the interior of their effective domain.

(c) Let U be a one-dimensional affine subspace of $\mathbb{R}^2$. Then $F|_U$ is either a strictly convex or an affine function. The same holds for $R|_U$.

(d) If $(u, w) \in \operatorname{dom} F$, then also $(u, \eta) \in \operatorname{dom} F$ for all $\eta \le w$. The same holds for R.

Remark 5.3. As usual in convex analysis, we regard F and R as functions defined on all of $\mathbb{R}^2$, that may attain values in $\mathbb{R} \cup \{+\infty\}$. The set $\{(u, w) : F(u, w) < \infty\}$ is called effective domain of F, and denoted by $\operatorname{dom} F$.

We define a function $\chi(u)$, that will appear in several conditions throughout this article. Corollary 6.5 gives an interpretation of χ as a rate of convergence for the asymptotic behavior of the cumulant generating function of $(X_t)_{t \ge 0}$.

Definition 5.4. For each $u \in \mathbb{R}$ where $R(u, 0) < \infty$, define $\chi(u)$ as

$$\chi(u) := \frac{\partial R}{\partial w}(u, w)\bigg|_{w=0} .$$

$\chi(u)$ is well-defined at least as a limit as $w \uparrow 0$, possibly taking the value $+\infty$; it can be written explicitly as

$$\chi(u) = \alpha_{12} u + \beta_1 + \int_{D \setminus \{0\}} y \left(e^{xu} - \frac{1}{1 + y^2} \right) \mu(dx, dy) .$$

Note that also $\chi(u)$ is a convex function.

5.2. Explosions and the martingale property. We are interested in conditions under which $S_t = \exp(X_t)$ is conservative and a martingale. If such conditions are satisfied, $(S_t)_{t \ge 0}$ may serve as the price process under the risk-neutral measure in an arbitrage-free asset pricing model. The following theorem gives sufficient and necessary conditions:

Theorem 5.5. *Suppose (X_t, V_t) satisfies Assumptions A1 and A2. Then the following holds:*

(a) $(S_t)_{t \ge 0}$ is conservative if and only if $F(0, 0) = R(0, 0) = 0$ and either

$$\chi(0) < \infty \quad \text{or} \quad \left(\chi(0) = \infty \ \wedge \ \int_{0-} \frac{d\eta}{R(0, \eta)} = -\infty \right) . \tag{5.7}$$

(b) $(S_t)_{t \ge 0}$ is a martingale if and only if it is conservative, $F(1, 0) = R(1, 0) = 0$ and either

$$\chi(1) < \infty \quad \text{or} \quad \left(\chi(1) = \infty \ \wedge \ \int_{0-} \frac{d\eta}{R(1, \eta)} = -\infty \right) . \tag{5.8}$$

Remark 5.6. The notation $\int_{0-}$ denotes an integral over an arbitrarily small left neighborhood of 0.

By (5.6) the condition $F(0,0) = R(0,0) = 0$ is equivalent to $c = \gamma = 0$, i.e. obviously the killing rate has to be zero for the process to be conservative. As will be seen in the proof, the integral conditions (5.7) and (5.8) are related to a uniqueness condition for non-Lipschitz ODEs, which has been discovered by Osgood [1898]. From Theorem 5.5 we derive the following easy-to-check sufficient conditions:

Corollary 5.7. *Suppose (X_t, V_t) satisfies Assumptions A1 and A2.*

(a) If $F(0,0) = R(0,0) = 0$ and $\chi(0) < \infty$ then $(S_t)_{t\geq 0}$ is conservative.

(b) If $(S_t)_{t\geq 0}$ is conservative, $F(1,0) = R(1,0) = 0$ and $\chi(1) < \infty$, then $(S_t)_{t\geq 0}$ is a martingale.

Proof. For a proof of 5.5a we refer to [Filipović, 2001, Th. 4.11]. Statement 5.5b can be shown in a similar way:
Since (X_t, V_t) is Markovian, we have for all $0 \leq s \leq t$, that

$$\mathbb{E}[S_t | \mathcal{F}_s] = S_s \exp\left(\phi(t-s,1,0) + V_s \psi(t-s,1,0)\right) .$$

We have assumed that $V_0 > 0$ a.s., such that $(S_t)_{t\geq 0}$ is a martingale if and only if $(X_t)_{t\geq 0}$ is conservative and $\psi(t,1,0) = \phi(t,1,0) \equiv 0$ for all $t \in \mathbb{R}_{\geqslant 0}$.
We show Corollary 5.7 and the first implication of 5.5b: Suppose that $(S_t)_{t\geq 0}$ is conservative and that $F(1,0) = R(1,0) = 0$. By Theorem 5.1 $\psi(t,1,w)$ solves the differential equation

$$\frac{\partial}{\partial t}\psi(t,1,w) = R(1,\psi(t,1,w)), \qquad \psi(0,1,w) = w \tag{5.9}$$

for all $w \leq 0$. Since $R(1,0) = 0$ it is clear that $\widetilde{\psi}(t,1,0) \equiv 0$ satisfies this ODE for the initial value $w = 0$. To deduce that $\widetilde{\psi}(t,1,0) = \psi(t,1,0)$ however, we need to know whether the solution is unique. Since $R(1,w)$ is continuously differentiable for $w < 0$, it satisfies a Lipschitz condition on $(-\infty,0)$. If $\chi(1) < \infty$, the Lipschitz condition can be extended to $(-\infty,0]$, and $\psi(t,1,0) \equiv 0$ is the unique solution. If $\chi(1) = \infty$, we substitute Lipschitz' condition by Osgood's condition[15]: Suppose that (5.8) holds, and there exists a non-zero solution $\widetilde{\psi}$ such that $\widetilde{\psi}(t_1,1,0) < 0$ for some $t_1 > 0$. Then for all $t < t_1$ such that ψ remains non-zero on $[t,t_1]$ we have (similarly to (5.5)) that

$$\int_{\widetilde{\psi}(t_1,1,0)}^{\widetilde{\psi}(t,1,0)} \frac{d\eta}{R(1,\eta)} = t - t_1 \ . \tag{5.10}$$

Let $t_0 = \sup\left\{t < t_1 : \widetilde{\psi}(t,1,w) = 0\right\}$. Taking the limit $t \downarrow t_0$, the left side of (5.10) tends to $-\infty$, whereas the right side remains bounded, leading to a contradiction. We conclude that $\psi(t,1,0) \equiv 0$ is the unique solution of (5.9). Finally equation (5.4) together with $F(1,0) = 0$ yields that also $\phi(t,1,0) \equiv 0$ for all $t \in \mathbb{R}_{\geqslant 0}$ and we

[15] See Osgood [1898]

have shown that $(S_t)_{t\geq 0}$ is a martingale.

For the other direction of 5.5b note that $(S_t)_{t\geq 0}$ being a martingale implies that $\phi = \psi \equiv 0$ solve the generalized Riccati equations and thus that $F(1,0) = R(1,0) = 0$. It remains to show (5.8). If $\chi(1) < \infty$, nothing is to show. For a contradiction we thus assume that $\chi(1) = \infty$, but $\int_{0-} \frac{d\eta}{R(1,\eta)} > -\infty$. Then, for each $t > 0$, (5.10) with $t_1 = 0$ implicitly defines a solution $\widetilde{\psi}(t,1,0)$ of the generalized Riccati equation (5.9), satisfying $\widetilde{\psi}(t,1,0) < 0$ for all $t > 0$. By uniqueness of the solution $\psi(t,1,w)$ for $w < 0$ and the semi-flow property (5.1), we have $\widetilde{\psi}(t+s,1,0) = \psi(t,1,\widetilde{\psi}(s,1,0))$ for t,s small enough. Letting $s \downarrow 0$ we obtain $\psi(t,1,0) = \widetilde{\psi}(t,1,0) < 0$, which is a contradiction to $\psi \equiv 0$. □

We add now two assumptions to A1 and A2 and complete our definition of an affine stochastic volatility model:

A3: The discounted price process $S_t = e^{X_t}$ is a martingale.
A4: There exists some $u \in \mathbb{R}$, such that $R(u,0) \neq 0$.

Assumption A4 excludes models where the distribution of $(X_t)_{t\geq 0}$ does not depend at all on the volatility state V_0. In such a case we can not speak of a true stochastic volatility model, and it will be beneficial to avoid these degenerate cases. We are now ready to give our definition of an affine stochastic volatility model:

Definition 5.8. The process $(X_t, V_t)_{t\geq 0}$ is called an affine stochastic volatility model, if it satisfies assumptions A1 – A4.

A simple consequence of this definition, that will often be used is the following:

Lemma 5.9. *Let $(X_t, V_t)_{t\geq 0}$ be an affine stochastic volatility model. Then $u \mapsto R(u,0)$ is a strictly convex function, satisfying $R(0,0) = R(1,0) = 0$.*

Proof. From assumption A3 and Theorem 5.5 it follows that $R(0,0) = R(1,0) = 0$. Lemma 5.2 implies that $R(u,0)$ is either strictly convex or an affine function. Assume it is affine. Then $R(u,0) = 0$ for all $u \in \mathbb{R}$. This contradicts A4, such that we conclude that $R(u,0)$ is a strictly convex function. □

6. Long-term asymptotics for ASVMs

In this section we study the behavior of an affine stochastic volatility model as $t \to \infty$. We focus first on the stochastic variance process $(V_t)_{t\geq 0}$. Under mild assumptions this process will converge in law to its invariant distribution:

6.1. Stationarity of the variance process.

Proposition 6.1. *Suppose that A1 and A2 hold, that $\chi(0) < 0$ and the Lévy measure m satisfies the logarithmic moment condition*

$$\int_{y>1} (\log y)\ m(dx, dy) < \infty.$$

Then $(V_t)_{t\geq 0}$ converges in law to its unique invariant distribution L, which has the cumulant generating function

$$l(w) = \int_w^0 \frac{F(0,\eta)}{R(0,\eta)}\, d\eta \qquad (w \leq 0)\ . \tag{6.1}$$

Keller-Ressel and Steiner [2008] show that under the given conditions the process $(V_t)_{t\geq 0}$ converges in law to a limit distribution L, whose cumulant generating function can be represented by (6.1). A short argument at the end of this paragraph shows that the limit distribution is also the unique invariant distribution of $(V_t)_{t\geq 0}$. First we make the following definition: Given some affine stochastic volatility model $(X_t, V_t)_{t\geq 0}$, we introduce the process $(\widetilde{X}_t, \widetilde{V}_t)_{t\geq 0}$, defined as the Markov process with the same transition probabilities as $(X_t, V_t)_{t\geq 0}$, but started with $X_0 = 0$ and V_0 distributed according to L. We will refer to $(\widetilde{X}_t, \widetilde{V}_t)_{t\geq 0}$ as the stochastic volatility model $(X_t, V_t)_{t\geq 0}$ 'in the stationary variance regime'. We also define the associated price process $\widetilde{S}_t := \exp(rt + \widetilde{X}_t)$. As we discuss in Section 8 the process $(\widetilde{X}_t, \widetilde{V}_t)_{t\geq 0}$ can be related to the pricing of forward-starting options, when the time until the start of the contract is large.
The cumulant generating function of $(\widetilde{X}_t, \widetilde{V}_t)$ is given by

$$\log \mathbb{E}[e^{u\widetilde{X}_t + w\widetilde{V}_t}] = \log \mathbb{E}\left[\exp\left(\phi(t,u,w) + \widetilde{V}_0 \psi(t,u,w)\right)\right] = \phi(t,u,w) + l(\psi(t,u,w))\ . \tag{6.2}$$

We verify now that L is indeed an invariant distribution of $(V_t)_{t\geq 0}$:

$$\begin{aligned} \mathbb{E}\left[\exp\left(w\widetilde{V}_t\right)\right] &= \exp\left(\phi(t,0,w) + l(\psi(t,0,w))\right) = \\ &= \exp\left(\int_0^t F(0,\psi(s,0,w))ds + \int_{\psi(t,0,w)}^0 \frac{F(0,\eta)}{R(0,\eta)}\, d\eta\right) = \\ &= \exp\left(\int_w^{\psi(t,0,w)} \frac{F(0,\eta)}{R(0,\eta)}\, d\eta + \int_{\psi(t,0,w)}^0 \frac{F(0,\eta)}{R(0,\eta)}\, d\eta\right) = \exp(l(w))\ , \end{aligned} \tag{6.3}$$

where we have used that under the conditions of the Proposition above, $\psi(t,0,w)$ is a strictly monotone function converging to 0 as $t \to \infty$. (cf. Keller-Ressel and Steiner [2008]). To see that L is unique, assume that there exists another invariant distribution L', and let $(V_t')_{t\geq 0}$ be the variance process started with V_0' distributed according to L'. Again we use that $\phi(t,u,w) \to l(w)$ and $\psi(t,0,w) \to 0$ as $t \to \infty$

(see Keller-Ressel and Steiner [2008]), and get that

$$\lim_{t\to\infty} \mathbb{E}\left[\exp(wV_t')\right] = \mathbb{E}\left[\lim_{t\to\infty} \exp\left(\phi(t,0,w) + V_0'\psi(t,0,w)\right)\right] = \mathbb{E}[\exp(l(w))] = e^{l(w)} \ ,$$

for all $w \le 0$, in contradiction to the invariance of L'.

6.2. Long-term behavior of the log-price process. We have seen that under certain conditions $(V_t)_{t\ge0}$ converges to a limit distribution, but we can not expect the same for the log-price process $(X_t)_{t\ge0}$. Nevertheless, it can be shown that the rescaled cumulant generating function $\frac{1}{t}\log\mathbb{E}\left[e^{X_t u}\right]$ converges under suitable conditions to a limit $h(u)$, that is again the cumulant generating function of some infinitely divisible random variable. This result can be interpreted such, that for large t the marginal distributions of $(X_t)_{t\ge0}$ are 'close' to the marginal distributions of a Lévy process with characteristic exponent $h(u)$. Furthermore, $h(u)$ can be directly obtained from the functions F and R, without knowledge of the explicit forms of ϕ and ψ. We start with a preparatory Lemma:

LEMMA 6.2. *Let $(X_t, V_t)_{t\ge0}$ be an affine stochastic volatility model and suppose that $\chi(0) < 0$ and $\chi(1) < 0$. Then there exist a maximal interval I and a unique function $w \in C(I) \cap C^1(I^\circ)$, such that*

$$R(u, w(u)) = 0 \qquad \textit{for all} \quad u \in I \tag{6.4}$$

and $w(0) = w(1) = 0$.
Moreover it holds that $[0,1] \subseteq I$, $w(u) < 0$ for all $u \in (0,1)$; $w(u) > 0$ for all $u \in I \setminus [0,1]$; and

$$\frac{\partial R}{\partial w}(u, w(u)) < 0 \tag{6.5}$$

for all $u \in I^\circ$.

We show Lemma 6.2 together with the next result, which makes the connection to the qualitative properties of the generalized Riccati equations.

LEMMA 6.3. *(a) For each $u \in I^\circ$, $w(u)$ is an asymptotically stable equilibrium point of the generalized Riccati equation* (5.3b).
(b) For $u \in I^\circ$, there exists at most one other equilibrium point $\widetilde{w}(u) \ne w(u)$, and if it exists, it is necessarily unstable and satisfies $\widetilde{w}(u) > \max(0, w(u))$.
(c) For $u \in \mathbb{R} \setminus I$, no equilibrium point exists.

PROOF. Define $L = \{(u,w) : R(u,w) \le 0\}$. As the level set of the closed convex function R, it is a closed and convex set. For all $u \in \mathbb{R}$, define $w(u) = \inf\{w : (u,w) \in L\}$, and $I = \{u \in \mathbb{R} : w(u) < \infty\}$. Clearly $w(u)$ is a continuous convex function, and I a subinterval of $\mathbb{R}$. We will now show that $w(u)$ and I satisfy all properties stated in Lemma 6.2. By assumption A3 and Theorem 5.5,

$R(0,0) = R(1,0) = 0$; together with Lemma 5.2 it follows that the set $[0,1]\times(-\infty,0]$ is contained in dom R. Since $R(u,0)$ is by Lemma 5.9 strictly convex, and also $\chi(u)$ is convex, we deduce that $R(u,0) < 0$ and $\frac{\partial R}{\partial w}(u,0) = \chi(u) < 0$ for all $u \in (0,1)$. In addition $R(u,w)$, as a function of w, is either affine or strictly convex, such that there exists a unique point $w(u)$, where $R(u,w(u)) = 0$, and necessarily $\frac{\partial R}{\partial w}(u,w(u)) < 0$. It is clear that for $u \in (0,1)$ $w(u)$ coincides with the function defined by (6.4), and that $w(u) < 0$. At $u = 0$ we have that $R(0,0) = 0$ and $\chi(0) < 0$, implying that $w(0) = 0$. A symmetrical argument at $u = 1$ shows that $w(1) = 0$, and thus that $[0,1] \subseteq I$.
We show next that $w(u) \in C^1(I^\circ)$: Define $u_+ = \sup I$, and $w_+ = \lim_{u\uparrow u_+} w(u)$; u_-, w_- are defined symmetrically at the left boundary of I. Note that $u_\pm$ and $w_\pm$ can take infinite values. Define the open set

$$K := \{(\lambda u_- + (1-\lambda)u_+, w) \,:\, \lambda \in (0,1),\, w < \lambda w_- + (1-\lambda)w_+\} \ .$$

Lemma 5.2 implies that K is contained in the interior of dom R. On the other hand, the graph of w, restricted to I°, i.e. the set $\{(u,w(u)) : u \in I^\circ\}$, is clearly contained in K. Since R is by Lemma 5.2 an analytic function in the interior of its effective domain, the implicit function theorem implies that $w(u) \in C^1(I^\circ)$. In addition it follows that $\frac{\partial R}{\partial w}(u,w(u)) \neq 0$ for all $u \in I^\circ$, such that the assertion $\frac{\partial R}{\partial w}(u,w(u)) < 0$, which we have shown above for $u \in (0,1)$, can be extended to all of I°. The claim that $w(u) > 0$ for $u \in I \setminus [0,1]$ can easily be derived from the convexity of $w(u)$, and the fact that $w(u) < 0$ inside $(0,1)$ and $w(0) = w(1) = 0$.

We have now proved most part of Lemma 6.2 (except for the uniqueness), and turn towards Lemma 6.3: Since $R(u,w(u)) = 0$ and $\frac{\partial R}{\partial w}(u,w(u)) < 0$ for all $u \in I^\circ$, $w(u)$ must be an asymptotically stable equilibrium point of the generalized Riccati equation 5.3b, showing 6.3a. Assume now that for some $u \in I^\circ$ there exists a point $\widetilde{w}(u) \neq w(u)$ such that $R(u,\widetilde{w}(u)) = 0$. By Lemma 5.2, $R(u,w)$ is, as a function of w, either strictly convex or affine. If it is affine, it has a unique root, and $\widetilde{w}(u)$ cannot exist. If it is strictly convex, there can exist a single point $\widetilde{w}(u)$ other than $w(u)$, such that $R(u,\widetilde{w}(u)) = 0$. Necessarily $\widetilde{w}(u) > w(u)$ and $\frac{\partial R}{\partial w}(u,\widetilde{w}(u)) > 0$. This shows that $\widetilde{w}(u)$ is an unstable equilibrium point of the generalized Riccati equation for ψ. In addition $\widetilde{w}(u) > w(u)$, and in particular the fact that $\widetilde{w}(0) > 0$ and $\widetilde{w}(1) > 0$ shows the uniqueness of $w(u)$ in the sense of Lemma 6.2. To see that $\widetilde{w}(u) > \max(0,w(u))$, note that we only have to show that $\widetilde{w}(u) > 0$, whenever $w(u) < 0$. This is the case only for $u \in (0,1)$. Assume that $\widetilde{w}(u) \leq 0$ for $u \in (0,1)$. Then the convexity of R and $\frac{\partial R}{\partial w}(u,\widetilde{w}(u)) > 0$ would imply that $R(u,0) \geq 0$ for some $u \in (0,1)$. This is impossible by Lemma 5.9, and we have shown 6.3b. Finally

6.3c follows directly from the definition of $w(u)$ as $w(u) = \inf\{w : (u,w) \in L\}$ and I as the effective domain of $w(u)$. $\square$

We are now ready to show our main result on the long-term properties of the log-price process $(X_t)_{t\geq 0}$.

THEOREM 6.4. *Let $(X_t, V_t)_{t\geq 0}$ be an affine stochastic volatility model and suppose that $\chi(0) < 0$ and $\chi(1) < 0$. Let $w(u)$ be given by Lemma 3.2 and define*

$$h(u) = F(u, w(u)), \qquad J = \{u \in I : F(u, w(u)) < \infty\} \ .$$

Then $[0,1] \subseteq J \subseteq I$; $w(u)$ and $h(u)$ are cumulant generating functions of infinitely divisible random variables and

$$\lim_{t\to\infty} \psi(t,u,0) = w(u) \quad \textit{for all} \quad u \in I \ ; \tag{6.6a}$$

$$\lim_{t\to\infty} \frac{1}{t}\phi(t,u,0) = h(u) \quad \textit{for all} \quad u \in J \ . \tag{6.6b}$$

COROLLARY 6.5. *Under the conditions of Theorem 6.4, the following holds:*

$$\sup_{u\in[0,1]} |\psi(t,u,0) - w(u)| \leq C\exp(-\mathfrak{X}\cdot T) \ ; \tag{6.7a}$$

$$\sup_{u\in[0,1]} \left|\frac{1}{t}\phi(t,u,0) - h(u)\right| \leq \Omega C\exp(-\mathfrak{X}\cdot T) \ ; \tag{6.7b}$$

for some constant C, and with

$$\mathfrak{X} = \inf_{u\in[0,1]} |\chi(u)| \quad \textit{and} \quad \Omega = \sup_{u\in[0,1]} \left.\frac{\partial}{\partial w}F(u,w)\right|_{w=0}$$

PROOF. Let $u \in [0,1]$. By Lemma 6.2 $(u, w(u)) \in [0,1] \times (-\infty, 0]$. By Theorem 5.5 $F(0,0) = F(1,0) = 0$, such that Lemma 5.2 guarantees that $[0,1] \times (-\infty, 0] \subseteq \operatorname{dom} F$. It follows that $[0,1] \subseteq J$. Define

$$z(t,u) = \psi(t,u,0) - w(u) \ .$$

Inserting into the generalized Riccati equation 5.3b,

$$\frac{\partial}{\partial t}z(t,u) = R(u,\psi(t,u,0)) = R(u,\psi(t,u,0)) - R(u,w(u)), \quad \text{and} \quad z(0,u) = w(u) \ .$$

If $\psi(t,u,0) \leq 0$ we can bound the right hand side by

$$R(u,\psi(t,u,0)) - R(u,w(u)) \leq z(t,u)\frac{\partial R}{\partial w}(u,0) = z(t,u)\chi(u),$$

using convexity of R. By Gronwall's inequality

$$z(t,u) \leq |w(u)| \exp\left(\chi(u)t\right) \ .$$

Since χ is convex, $\chi(0) < 0$ and $\chi(1) < 0$, we have shown (6.7a). The estimate

$$|\phi(t,u) - h(u)| = \\ = \left|\frac{1}{t}\int_0^t \left(F(u,\phi(s,u)) - F(u,w(u))\right)\,ds\right| \le \left|\frac{\partial F}{\partial w}(u,0)\right| \cdot |\psi(t,u) - w(u)|$$

yields (6.7b) and we have shown Corollary 6.5.

Let now $u \in I^\circ \setminus [0,1]$. Combining Lemma 5.9 and Lemma 6.3 we have that $R(u,w) > 0$ for all $w \in [0, w(u))$, and $R(u, w(u)) = 0$. It follows that the initial value $\psi(0,u,0) = 0$ is in the basin of attraction of the stable equilibrium point $w(u)$ and thus that $\psi(t,u,0)$ is strictly increasing and converging to $w(u)$. An additional argument may be needed at the boundary of I: Let $u_+ = \sup I$ and assume that $u_+ \in I$ (i.e. I is right-closed). Since $(u_+, w) \in \operatorname{dom} R$ for all $w \le w(u_+)$, we can define $\frac{\partial R}{\partial w}(u_+, w(u_+))$ at least as a limit for $w \uparrow w(u_+)$. By Lemma 6.2 either $\frac{\partial R}{\partial w}(u_+, w(u_+)) < 0$ or $\frac{\partial R}{\partial w}(u_+, w(u_+)) = 0$. In the first case we can argue as in the interior of I that $w(u_+)$ is an asymptotically stable equilibrium point. In the second case we use once more that by Lemma 5.2 $R(u_+, w)$ is, as a function of w, either strictly convex or affine. If it is affine, it must be equal to 0, and thus $R(u_+, 0) = 0$, in contradiction to Lemma 5.9. Hence it is strictly convex, and attains its minimum at $w(u_+)$. This implies that $R(u_+, w) > 0$ for all $w \in [0, w(u_+))$ and we conclude that $\psi(t, u_+, 0)$ converges to $w(u_+)$.[16] For $u_- = \inf I$, a symmetrical argument applies.

Assertion (6.6b) follows immediately from the representation (5.4), and

$$\lim_{t\to\infty} \frac{1}{t}\phi(t,u,0) = \lim_{t\to\infty} \frac{1}{t}\int_0^t F(u, \psi(s,u,0))\,ds = F(u, w(u))$$

for all $u \in J$.

We have shown that the sequence of infinitely divisible cumulant generating functions $\psi(t,u,0)$ converges on I to a function $w(u)$ that is continuous in a right neighborhood of 0. This is sufficient to imply that $w(u)$ is again the cumulant generating function of an infinitely divisible random variable (See Feller [1971, VIII.1, Example (e)] for the convergence part, and Sato [1999, Lemma 7.8] for the infinite divisibility.). The same argument can be applied to ϕ and $h(u)$, and we have shown Theorem 6.4. □

[16] Even though $\psi(t, u_+, 0)$ converges to $w(u_+)$, note that $w(u_+)$ is not a stable equilibrium point in the usual sense. This is due to the fact that solutions from a *right*-neighborhood $N \cap (w(u_+), \infty)$ will *diverge* from $w(u_+)$ to $+\infty$.

7. Moment explosions in ASVMs

In this section we continue to study the time evolution of moments $\mathbb{E}[S_t^u] = \mathbb{E}[e^{X_t u}]$ of the price process in an affine stochastic volatility model. We are interested in the phenomenon that in a stochastic volatility model, moments of the price process can explode (become infinite) in finite time. For stochastic volatility models of the CEV-type – a class including the Heston model, but no models with jumps – moment explosions have been studied by Andersen and Piterbarg [2007] and Lions and Musiela [2007]. In the context of option pricing, an interesting result of Lee [2004] connects the existence of moments of the stock price process to the steepness of the smile for deep in-the-money or out-of-the-money options. Our first result shows that in an affine stochastic volatility model a simple explicit expression for the time of moment explosion can be given:

7.1. Moment explosions. By definition, the u-th moment of S_t, i.e. $\mathbb{E}[S_t^u]$ is given by $S_0^u \exp\left(\phi(t,u,0) + V_0\psi(t,u,0)\right)$. We define the **time of moment explosion** for the moment of order u by

$$T_*(u) = \sup\left\{t : \mathbb{E}[S_t^u] < \infty\right\} \ .$$

It is obvious from the Markov property that $\mathbb{E}[S_t^u]$ is finite for all $t < T_*(u)$ and infinite for all $t > T_*(u)$. As in the previous section, the main result follows from a qualitative analysis of the generalized Riccati equations (5.3).

Theorem 7.1. *Suppose the conditions of Theorem 6.4 hold. Define*

$$J := \{u \in I : F(u, w(u)) < \infty\} \ ,$$
$$f_+(u) := \sup\{w \geq 0 : F(u,w) < \infty\} \ ,$$
$$r_+(u) := \sup\{w \geq 0 : R(u,w) < \infty\} \ ,$$

and suppose that $F(u,0) < \infty$, $R(u,0) < \infty$ and $\chi(u) < \infty$.

(a) If $u \in J$, then

$$T_*(u) = +\infty \ .$$

(b) If $u \in \mathbb{R} \setminus J$, then

$$T_*(u) = \int_0^{\min(f_+(u), r_+(u))} \frac{d\eta}{R(u,\eta)} \ .$$

If $F(u,0) = \infty$, $R(u,0) = \infty$ or $\chi(u) = \infty$ then

(c)

$$T_*(u) = 0 \ .$$

Proof. Suppose that $u \in J$. Then Theorem 6.4 implies that both $\psi(t,u,0)$ and $\phi(t,u,0)$ are finite for all $t \geq 0$. This proves (a). Let now $u \in \mathbb{R}\backslash J$, $F(u,0) < \infty$, $R(u,0) < \infty$ and $\chi(u) < \infty$. To prove (b) we start by analyzing the maximal lifetime of solutions to the generalized Riccati equation

$$\frac{\partial}{\partial t}\psi(t,u,0) = R(u,\psi(t,u,0)), \quad \psi(0,u,0) = 0 \ . \tag{7.1}$$

Define $M = [0, r_+(u))$ and note that $R(u,.) \in C(M)$. Since $u \notin [0,1]$, Lemma 5.9 implies that $R(u,0) > 0$. It is clear, that at least a local solution $\psi(t,u,0)$ to the ODE exists, which satisfies $0 \leq \psi(t,u,0) \leq r_+(u)$ and is an increasing function of t as long as it can be continued. Using a standard extension theorem (e.g. Hartman [1982, Lem. I.3.1]) the local solution $\psi(t,u,0)$ has a maximal extension to an interval $[0, T(u))$, such that one of the following holds:

(i) $T(u) = \infty$, or
(ii) $T(u) < \infty$ and $\psi(t,u,0)$ comes arbitrarily close to the boundary of M, i.e.

$$\limsup_{t \to T(u)} \psi(t,u,0) = r_+(u) \ .$$

Consider case (i). Since ψ is increasing, its limit for $t \to \infty$ exists, but can be infinite. Suppose $\lim_{t\to\infty} \psi(t) = \alpha < \infty$. Then α must be a stationary point, i.e. $R(u,\alpha) = 0$, but this is impossible by Lemma 6.3. The case that $\alpha = \infty$ is only possible if $r_+(u) = \infty$, such that in this case $\lim_{t\to T(u)} \psi(t,u,0) = r_+(u)$. Consider case (ii). Since ψ is increasing the limes superior can be replaced by an ordinary limit and we get $\lim_{t\to T(u)} \psi(t,u,0) = r_+(u)$ as before.

Let now T_n be a sequence such that $T_n \uparrow T(u)$. By (5.5) it holds that

$$\int_0^{\psi(T_n,u,0)} \frac{d\eta}{R(u,\eta)} ds = T_n \ . \tag{7.2}$$

Letting $n \to \infty$ we obtain that $T(u) = \int_0^{r_+(u)} \frac{d\eta}{R(u,\eta)} ds$.

We can write the time of moment explosion $T_*(u)$ as the maximum joint lifetime of $\phi(t,u,0)$ and $\psi(t,u,0)$, i.e. $T_*(u) = \sup\{t \geq 0 : \phi(t,u,0) < \infty \wedge \psi(t,u,0) < \infty\}$. By the integral representation (5.4) it is clear that if $f_+(u) \geq r_+(u)$, $\phi(t,u,0)$ is finite whenever $\psi(t,u,0)$ is finite and $T_*(u) = T(u)$. If $f_+(u) < r_+(u)$ then $\psi(T_*(u),u,0) = f_+(u)$. Inserting into the representation (7.2) yields (b).
For assertion (c), let $F(u,0) = \infty$, $R(u,0) = \infty$, or $\chi(u) = \infty$. In the first case, $\phi(t,u,0)$ does not exist beyond $t = 0$. In the other cases no local solution to the

generalized Riccati equation (7.1) exists, such that $\psi(t,u,0)$ explodes immediately. □

7.2. Moment explosions in the stationary variance regime. In Section 6.1 we have introduced $(\widetilde{X}_t, \widetilde{V}_t)_{t\geq 0}$ as the model in the stationary variance regime. The moment explosions of this process can be analyzed in a similar manner as above. We define the time of moment explosion in the stationary variance regime by

$$T_*^S(u) := \sup\left\{T \geq 0 : \mathbb{E}[\widetilde{S}_T^u] < \infty\right\} ;$$

the superscript 'S' stands for 'stationary'.

The analogue to Theorem 7.1 is the following result:

THEOREM 7.2. *Suppose the conditions of Theorem 6.4 hold. Define $f_+(u), r_+(u)$ as in Theorem 7.1, and in addition*

$$l_+ := \sup\{w > 0 : l(w) < \infty\} .$$

Suppose that $F(u,0) < \infty$, $R(u,0) < \infty$ and $\chi(0) < \infty$.

(a) If $u \in J$ and $w(u) \leq l_+$, then

$$T_*^S(u) = +\infty .$$

(b) If $u \in \mathbb{R} \setminus J$ or $w(u) > l_+$, then

$$T_*^S(u) = \int_0^{\min(f_+(u), r_+(u), l_+)} \frac{d\eta}{R(u,\eta)} .$$

If $F(u,0) = \infty$, $R(u,0) = \infty$ or $\chi(0) = \infty$, then

(c)

$$T_*^S(u) = 0 .$$

COROLLARY 7.3. *Under the conditions of Theorem 7.2,*

$$T_*^S(u) \leq T_*(u), \qquad \text{for all } u \in \mathbb{R}$$

PROOF. By equation (6.2), the moment $\mathbb{E}[\widetilde{S}_t^u]$ is given by

$$\mathbb{E}[\widetilde{S}_t^u] = \exp\left(\phi(t,u,0) + l(\psi(t,u,0))\right) .$$

This expression is finite, if $\phi(t,u,0)$ and $\psi(t,u,0)$ are finite, *and* if $\psi(t,u,0) < l_+$. It is infinite if $\phi(t,u,0)$ or $\psi(t,u,0)$ are infinite, or if $\psi(t,u,0) > l_+$. The rest of the proof can be carried out as for Theorem 7.1. Note, that now even for $u \in J$, the moment can explode, if l_+ is reached by $\psi(t,u,0)$ before the stationary point $w(u)$. Corollary 7.3 follows easily by comparing the range of integration and the conditions for case (a) and (b) between Theorem 7.1 and Theorem 7.2. □

8. Applications to the implied volatility smile

8.1. Smile behavior at extreme strikes. In the preceding section, we have kept u fixed, and looked at the first time $T_*(u)$ that the moment $\mathbb{E}[S_t^u]$ becomes infinite. It will now be more convenient to reverse the roles of T and u, and for a given time t to define the **upper critical moment** by

$$u_+(t) = \sup\{u \geq 1 : \mathbb{E}[S_t^u] < \infty\} = \sup\{u \geq 1 : T_*(u) < t\} ,$$

and the **lower critical moment** by

$$u_-(t) = \inf\{u \leq 0 : \mathbb{E}[S_t^u] < \infty\} = \inf\{u \leq 0 : T_*(u) < t\} .$$

It is seen that $u_-(T)$ and $u_+(T)$ can be defined as the generalized inverse of $T_*(u)$ on $(-\infty, 0]$ and $[1, \infty)$ respectively. In addition it is easily derived from Jensen's inequality, that

$$\begin{aligned} \mathbb{E}[S_t^u] &< \infty \qquad \text{for all } u \in (u_-(t), u_+(t)), \quad \text{and} \\ \mathbb{E}[S_t^u] &= \infty \qquad \text{for all } u \in \mathbb{R} \setminus [u_-(t), u_+(t)] . \end{aligned}$$

The results of Lee [2004] relate the explosion of moments to the 'wing behavior' of the implied volatility smile, i.e. the shape of the smile for strikes that are deep in-the-money or out-of-the-money. To give a precise statement, let ξ be the log-moneyness, which for a European option with time-to-maturity T and strike K is given by $\xi = \log\left(\frac{K}{e^{rT}S_0}\right)$.

Proposition 8.1 (Lee's moment formula). *Let $V(T,\xi)$ be the implied Black-Scholes-Variance of a European call with time-to-maturity T and log-moneyness ξ. Then*

$$\limsup_{\xi \to -\infty} \frac{V(T,\xi)}{|\xi|} = \frac{\varsigma(-u_-(T))}{T}$$

and

$$\limsup_{\xi \to \infty} \frac{V(T,\xi)}{|\xi|} = \frac{\varsigma(u_+(T) - 1)}{T}$$

where $\varsigma(x) = 2 - 4\left(\sqrt{x^2 + x} - x\right)$ and $u_\pm(T)$ are the critical moment functions.

The function ς is strictly decreasing on $\mathbb{R}_{\geqslant 0}$, mapping 0 to 2, and ∞ to 0. Thus for fixed time-to-maturity T, the steepness of the smile is decreasing as $|u_\pm(T)|$ increases. A finite critical moment $u_\pm(T)$ implies asymptotically linear behavior of $V(T,\xi)$ in ξ, and an infinite critical moment implies sublinear behavior of $V(T,\xi)$. It is also evident that $u_-(T)$ determines the 'left' side of the volatility smile, also known as small-strike, in-the-money-call or out-of-the-money-put side; $u_+(T)$ determines the 'right' side, or large-strike, out-of-the-money-call, in-the-money-put side. Finally we mention that Lee's result has been extended and strengthened by

Benaim and Friz [2006] from a 'lim sup' to a genuine limit under conditions related to regular variation of the underlying distribution function.

8.2. Forward-smile behavior. The forward smile is derived from the prices of forward-start options. For a forward-start call option – all options we consider are European – a start date τ, a strike date $T+\tau$ and a moneyness ratio M are agreed upon today (at time $t=0$). The option then yields at time $T+\tau$ a payoff of $\left(\frac{S_{T+\tau}}{S_\tau}-M\right)_+$, i.e. the relative return over the time period from τ to $\tau+T$, reduced by M and floored at 0. Under the pricing measure the value of such an option at $t=0$ is given by

$$(8.1) \qquad e^{-r(T+\tau)}\mathbb{E}\left[\left(\frac{S_{T+\tau}}{S_\tau}-M\right)_+\right]=e^{-\tau r}\mathbb{E}\left[\left(e^{X_{T+\tau}-X_\tau}-e^{\xi}\right)_+\right] ,$$

where we define the log-moneyness ξ of a forward-start option as $\xi=\log M+rT$. Forward-start options are not just interesting in their own right, but are used as building blocks of more complex derivatives, such as Cliquet options (see Gatheral [2006, Chapter 10]).

Analogously to plain vanilla options, we can define the **implied forward volatility** $\sigma(\tau,T,\xi)$, by comparing the forward option price to the price of an option with identical payoff in the Black-Scholes model. Note that the implied forward volatility depends also on τ, the starting time of the contract. For $\tau=0$, the implied volatility of a plain vanilla option is retrieved. More interesting is the behavior for $\tau>0$. Intuitively, we expect the implied volatility (and the option price) to increase with τ in a stochastic volatility model, since the uncertainty of the variance V_τ at the starting date of the option has to be priced in. In an affine stochastic volatility model, it will be seen that under mild conditions, the implied forward volatilities $\sigma(\tau,T,\xi)$ actually converge to a limit as $\tau\to\infty$. Not surprisingly, this behavior is related to the convergence of $(V_t)_{t\geq 0}$ to its invariant distribution. In the limit $\tau\to\infty$, the pricing of a forward-start option is equivalent to the pricing of a plain vanilla option *in the stationary variance regime* (cf. Section 6.1).

PROPOSITION 8.2. *Let $(X_t,V_t)_{t\geq 0}$ be an affine stochastic volatility model, satisfying the conditions of Proposition 6.1. Let $\sigma(\tau,T,\xi)$ be the implied forward volatility in this model. Then*

$$\lim_{\tau\to\infty}\sigma(\tau,T,\xi)=\widetilde{\sigma}(T,\xi) ,$$

where $\widetilde{\sigma}(T,\xi)$ is the implied volatility of a European call with payoff $\left(e^{\widetilde{X}_T}-e^{\xi}\right)_+$, and $\widetilde{X}_T$ is the log-price process of the model in the stationary variance regime.

PROOF. We can write the price of a forward-start call as

$$C(\tau,T,\xi)=e^{-\tau r}\mathbb{E}\left[\left(e^{X_{T+\tau}-X_\tau}-e^{\xi}\right)_+\right]=e^{-r\tau}\mathbb{E}\left[\mathbb{E}^{(0,V_\tau)}\left[\left(e^{X_T}-e^{\xi}\right)_+\right]\right] .$$

Denote by $C^{\mathrm{BS}}(T,\xi,\sigma)$ the (plain vanilla) call price in a Black-Scholes model with volatility σ and the normalization $S_0 = 1$. It is easy to see that the price of a forward-start option in the Black-Scholes model is just the discounted plain vanilla price, i.e. $C^{\mathrm{BS}}(\tau,T,\xi,\sigma) = e^{-r\tau}C^{\mathrm{BS}}(T,\xi,\sigma)$. By definition, the implied forward volatility of the call $C(\tau,T,\xi)$ satisfies

$$C^{\mathrm{BS}}(T,\xi,\sigma(\tau,T,\xi)) = e^{r\tau}C(\tau,T,\xi) = \mathbb{E}\left[\mathbb{E}^{(0,V_\tau)}\left[\left(e^{X_T}-e^{\xi}\right)_+\right]\right] .$$

Taking the limit $\tau\to\infty$ on both sides we obtain

$$C^{\mathrm{BS}}(T,\xi,\lim_{\tau\to\infty}\sigma(\tau,T,\xi)) = \mathbb{E}\left[\lim_{\tau\to\infty}\mathbb{E}^{(0,V_\tau)}\left[\left(e^{X_T}-e^{\xi}\right)_+\right]\right] = \mathbb{E}\left[\left(e^{\widetilde{X}_T}-e^{\xi}\right)_+\right] ,$$

using dominated convergence. It is well known that the above equation allows a unique solution in terms of the Black-Scholes implied volatility, and we get $\widetilde{\sigma}(T,\xi) = \lim_{\tau\to\infty}\sigma(\tau,T,\xi)$. □

Combining Lee's moment formula with our results on moment explosions under the stationary variance regime (Theorem 7.2), asymptotics of $\widetilde{\sigma}(T,\xi)$ for $\xi\to\pm\infty$ can be derived.

9. Examples

9.1. The Heston model with and without jumps. In the model of Heston [1993], the log-price $(X_t)_{t\geq 0}$ and the corresponding variance process $(V_t)_{t\geq 0}$ are given under the risk-neutral measure by the SDE

$$\begin{aligned} dX_t &= -\frac{V_t}{2}\,dt + \sqrt{V_t}\,dW_t^1 \\ dV_t &= -\lambda(V_t-\theta)\,dt + \zeta\sqrt{V_t}\,dW_t^2 \end{aligned}$$

where W_t^1, W_t^2 are Brownian motions with correlation parameter ρ, and $\zeta,\lambda,\theta > 0$. In affine form, the model is written as

$$F(u,w) = \lambda\theta w \tag{9.1a}$$

$$R(u,w) = \frac{1}{2}(u^2-u) + \frac{\zeta^2}{2}w^2 - \lambda w + uw\rho\zeta . \tag{9.1b}$$

It is easily calculated that χ is given by $\chi(u) = \rho\zeta u - \lambda$. We will first analyze the long term behavior of $(X_t)_{t\geq 0}$, with the help of Theorem 6.4. To satisfy the condition $\chi(1) < 0$ we need $\lambda > \zeta\rho$. Note that this condition is always satisfied if $\rho\leq 0$, the case that is typical for applications. Solving a quadratic equation we find that

$$w(u) = \frac{(\lambda-u\rho\zeta)-\sqrt{(\lambda-u\rho\zeta)^2-\zeta^2(u^2-u)}}{\zeta^2} , \quad \text{and} \quad h(u) = \lambda\theta w(u) .$$

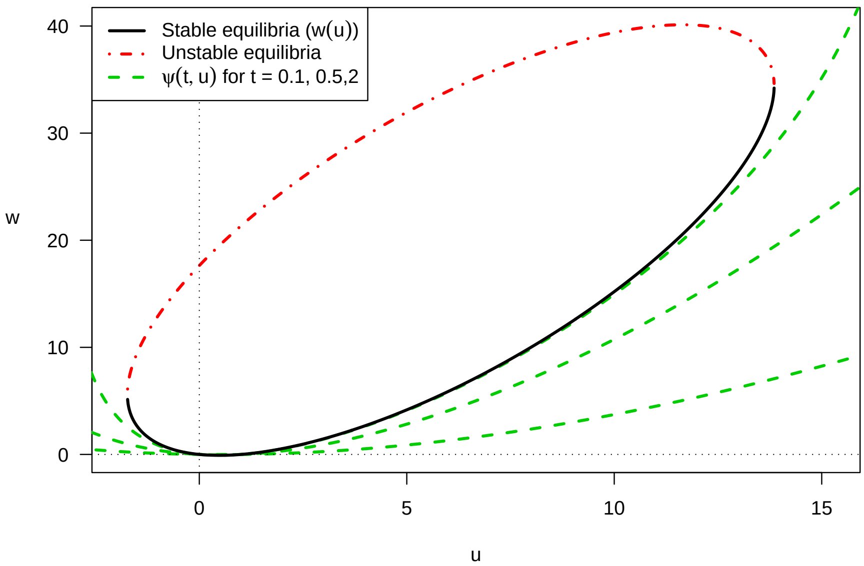

FIGURE 2. This plot shows the stable and unstable equilibria of the generalized Riccati equation of a Heston model with parameters $\rho = -0.7165$, $\zeta = 0.3877$, $\lambda = 1.3253$ and $\theta = 0.0354$ (taken from Gatheral [2006, Table 3.2]). It can be seen how the solutions $\psi(t,u)$ converge to the stable equilibrium points, which form the lower boundary of an ellipse in the (u,w)-plane.

Denoting the term under the square root by $\Delta(u)$, we see that $w(u)$ and $h(u)$ are both defined on $J = I = \{u : \Delta(u) \geq 0\}$. Since R is a second order polynomial in the Heston model, the equilibrium points of the generalized Riccati equation for ψ form an ellipse in the (u,w)-plane, and $w(u)$ is given by its lower part – see Figure 2 for an illustration. Interestingly, $w(u)$, and also $h(u)$, are cumulant generating functions of a Normal Inverse Gaussian distribution (cf. Barndorff-Nielsen [1997, Eq. (2.4)]). Thus, for large t, the price process of the Heston model is, in terms of its marginal distributions, close to a Normal-Inverse-Gaussian exponential-Lévy model.

Next we consider moment explosions in the Heston model. As mentioned above, moment explosions in the Heston model (and other models) have already been studied by Andersen and Piterbarg [2007]. Nevertheless this will provide a first test of Theorem 7.1: In the case of the Heston model it is easily determined from (9.1) that $f_+(u) = r_+(u) = \infty$. Calculating the integral in case (b) of Theorem 7.1, we obtain

$$(9.2)\qquad T_*(u) = \begin{cases} +\infty & \Delta(u) \geq 0 \\ \frac{2}{\sqrt{-\Delta(u)}} \left(\arctan \frac{\sqrt{-\Delta(u)}}{\chi(u)} + \pi \mathbf{1}_{\{\chi(u)<0\}} \right) & \Delta(u) < 0 \,. \end{cases}$$

In Figure 3 a plot of this function for typical parameter values is shown. Note that Andersen and Piterbarg [2007] distinguish an additional case where $\chi(u) < 0$, but $\Delta(u) > 0$. A little calculation shows that this can only happen if $\chi(1) \geq 0$, a case that is precluded by our assumptions in Theorem 6.4, and never occurs when $\rho \leq 0$.

We will now study the effect of adding jumps to the Heston model. The simplest case is the addition of an independent jump component with constant activity: Let $(J_t)_{t\geq 0}$ be a pure-jump Lévy process, independent of $(W_t^{1,2})_{t\geq 0}$ and define the Heston-with-jumps model by

$$dX_t = \left(\delta - \frac{V_t}{2}\right) dt + \sqrt{V_t}\, dW_t^1 + dJ_t$$
$$dV_t = -\lambda(V_t - \theta)\, dt + \zeta \sqrt{V_t}\, dW_t^2 \,.$$

The drift δ is determined by the martingale condition for $(S_t)_{t\geq 0}$. To make the example simple, we assume that $(J_t)_{t\geq 0}$ jumps only downwards. This is equivalent to saying that the Lévy measure $m(dx)$ of $(J_t)_{t\geq 0}$ is supported on $(-\infty, 0)$. The affine form of the model is

$$(9.3)\qquad F(u, w) = \lambda\theta w + \widetilde{\kappa}(u)$$

$$(9.4)\qquad R(u, w) = \frac{1}{2}(u^2 - u) + \frac{\zeta^2}{2} w^2 - \lambda w + uw\rho\zeta \,,$$

where $\widetilde{\kappa}(u)$ is the compensated cumulant generating function of the jump part, i.e.

$$\widetilde{\kappa}(u) = \int_{(-\infty,0)} \left(e^{xu} - 1 - u\,(e^x - 1)\right) m(dx) \,.$$

Let $\kappa_- < 0$ be the number such that $\widetilde{\kappa}(u)$ is finite on (κ_-, ∞) and infinite on $(-\infty, \kappa_-)$. For example, if the absolute jump heights are exponentially distributed with an expected jump size of $1/\alpha$, then $\kappa_- = -\alpha$.

To analyze the explosion times of this model, note that R, and thus $\chi(u)$, $w(u)$, I and $r_+(u)$ have not changed compared to the Heston model. As long as $u > \kappa_-$, the explosion time $T_*(u)$ is the same as in the Heston model. However, if $u \leq \kappa_-$, $F(u, 0) = \infty$ and by Theorem 7.1, $T_*(u) = 0$. Thus, the addition of jumps to the Heston model has the effect of truncating the explosion time to zero, whenever $u \leq \kappa_-$.

From the viewpoint of the critical moment functions, $u_+(t)$ does not change compared to the Heston model, but $u_-(t)$ does; in the model with jumps it is given by

$$u_-^{\text{Jump}}(t) = u_-^{\text{Heston}}(t) \vee \kappa_- \ .$$

Since u_- is increasing with t, it makes sense to define a cutoff time $T_\sharp$ by

$$T_\sharp = \sup\left\{t \geq 0 : u_-^{\text{Heston}}(t) = \kappa_-\right\} = T_*(\kappa_-) \ , \tag{9.5}$$

such that

$$\begin{aligned} u_-^{\text{Heston}}(t) &< u_-^{\text{Jump}}(t), \quad \text{if} \quad t < T_\sharp \\ u_-^{\text{Heston}}(t) &= u_-^{\text{Jump}}(t), \quad \text{if} \quad t \geq T_\sharp \ . \end{aligned} \tag{9.6}$$

In Figure 3 a comparison of the critical moment functions in the Heston model with and without jumps can be seen. By Lee's moment formula, the critical moment $u_-(t)$ moving closer to 0 will cause the left side of the implied volatility smile to become steeper. Thus the net effect of adding the jump component $(J_t)_{t\geq 0}$ to the Heston model is a steepening of the left side of the smile for maturities smaller than $T_\sharp$. For times larger than $T_\sharp$, the asymptotic behavior of the smile (in the sense of Lee's formula) is *exactly* the same as in the Heston model without jumps. This corresponds well to the frequently made observation (see e.g. Gatheral [2006, Chapter 5]) that a Heston model with jumps can be fitted well by first fitting a (jump-free) Heston model to long maturities, and then calibrating only the additional parameters to the full smile. In fact Gatheral proposes (on heuristical grounds) the concept of a 'critical time' T, after which the influence of an independent jump component on the implied volatility smile can be neglected. Equations (9.5) and (9.6) now provide a rigorous argument that this role can be attributed to the cutoff time $T_\sharp$. The analysis of the Heston model with jumps is of course easily extended to the case that $(J_t)_{t\geq 0}$ is not one-sided. In that case the effects discussed above will be seen to affect also the right side of the implied volatility smile.

9.2. A model of Bates. We consider now the model given by

$$\begin{aligned} dX_t &= \left(\delta - \frac{V_t}{2}\right) dt + \sqrt{V_t}\, dW_t^1 + \int_{\mathbb{R}} x\, \widetilde{N}(V_t, dt, dx) \\ dV_t &= -\lambda(V_t - \theta)\, dt + \zeta\sqrt{V_t}\, dW_t^2 \ . \end{aligned}$$

where as before $\lambda, \theta, \zeta > 0$ and the Brownian motions are correlated with correlation ρ. The jump component is given by $\widetilde{N}(V_t, dt, dx) = N(V_t, dt, dx) - n(V_t, dt, dx)$, where $N(V_t, dt, dx)$ is a Poisson random measure, and its intensity measure $n(V_t, dt, dx)$ is of the *state-dependent* form $V_t\mu(dx)dt$, with $\mu(dx)$ the Lévy measure given in (5.6). A model of this kind has been proposed by Bates [2000] to explain the time-variation of jump-risk implicit in observed option prices. Bates also proposes a second variance factor, which we omit in this example, in order to remain in the scope of Definition 5.8. It would however not be difficult to extend our approach to the two-factor Bates model, since the two proposed variance-factors are mutually independent, causing the corresponding generalized Riccati equations to decouple. Since it is affine, the above model can be characterized in terms of the functions F and R:

$$F(u,w) = \lambda\theta w \tag{9.7}$$

$$R(u,w) = \frac{1}{2}(u^2 - u) + \frac{\zeta^2}{2}w^2 - \lambda w + uw\rho\zeta + \widetilde{\kappa}(u)\ . \tag{9.8}$$

where $\widetilde{\kappa}(u) = \int_{(-\infty,0)} \left(e^{xu} - 1 - u\,(e^x - 1)\right) \mu(dx)$ is the compensated cumulant generating function of the Lévy measure μ. As in the Heston model we can obtain $w(u)$ and $h(u)$ explicitly, and get

$$h(u) = \frac{-\chi(u) - \sqrt{\Delta(u)}}{\zeta^2}\ , \quad \text{and} \quad h(u) = \lambda\theta w(u)\ ,$$

where $\chi(u) = \rho\zeta u - \lambda$ and $\Delta(u) = \chi(u)^2 - \zeta^2(u^2 - u + 2\widetilde{\kappa}(u))$. Both $w(u)$ and $h(u)$ are defined on $I = J = \{u : \Delta(u) \geq 0\}$. The time of moment explosion can again be calculated explicitly, and is given by

$$T_*(u) = \begin{cases} +\infty & \Delta(u) > 0 \\ \frac{2}{\sqrt{-\Delta(u)}}\left(\arctan\frac{\sqrt{-\Delta(u)}}{\chi(u)} + \pi\mathbf{1}_{\{\chi(u)<0\}}\right) & -\infty < \Delta(u) < 0 \\ 0 & \Delta(u) = -\infty\ . \end{cases} \tag{9.9}$$

9.3. The Barndorff-Nielsen-Shephard model. The Barndorff-Nielsen-Shephard (BNS) model was introduced by Barndorff-Nielsen and Shephard [2001] as a model for asset pricing. In SDE form it is given in the risk-neutral case by

$$dX_t = (\delta - \frac{1}{2}V_t)dt + \sqrt{V_t}\,dW_t + \rho\,dJ_{\lambda t}$$
$$dV_t = -\lambda V_t\,dt + dJ_{\lambda t}$$

where $\lambda > 0$, $\rho < 0$ and $(J_t)_{t\geq 0}$ is a Lévy subordinator, i.e. a pure jump Lévy process that increases a.s. The drift δ is determined by the martingale condition for $(S_t)_{t\geq 0}$. The time-scaling $J_{\lambda t}$ is introduced by Barndorff-Nielsen and Shephard to make the invariant distribution of the variance process independent of λ. The distinctive features of the BNS model are that the variance process has no diffusion component, i.e. moves purely by jumps and that the negative correlation between variance and price movements is achieved by simultaneous jumps in $(V_t)_{t\geq 0}$ and $(X_t)_{t\geq 0}$. The BNS model is an affine stochastic volatility model, and F and R are given by

$$F(u,w) = \lambda\kappa(w+\rho u) - u\lambda\kappa(\rho) \tag{9.10}$$

$$R(u,w) = \frac{1}{2}(u^2-u) - \lambda w \tag{9.11}$$

where $\kappa(u)$ is the cumulant generating function of $(J_t)_{t\geq 0}$.
We simply have $\chi(u) = -\lambda$ and $w(u)$ from Lemma 6.2 is given by

$$w(u) = \frac{1}{2\lambda}(u^2-u) \; .$$

It follows that

$$h(u) = \lambda\kappa\left(\frac{u^2}{2\lambda} + u\left(\rho - \frac{1}{2\lambda}\right)\right) - u\lambda\kappa(\rho) \; .$$

This expression can be interpreted as cumulant generating function of a Brownian motion with variance $\frac{1}{\lambda}$ and drift $\rho - \frac{1}{2\lambda}$, subordinated by the Lévy process $J_{\lambda t}$ and then mean-corrected to satisfy the martingale condition.
To analyze moment explosions in the BNS model, let $\kappa_+ := \sup\{u>0 : \kappa(u) < \infty\}$. It is easy to see that f_+ is given by $f_+ = \max(\kappa_+ - \rho u, 0)$, and that $r_+ = \infty$. Calculating the integral $\int_0^{f_+} \frac{d\eta}{R(u,\eta)}$ and applying Theorem 7.1, we see that the explosion time for the moment of order u is given by

$$T_*(u) = -\frac{1}{\lambda}\log\max\left(0, 1 - \frac{2\lambda(\max(\kappa_+ - \rho u, 0))}{u(u-1)}\right) \; .$$

The critical moment functions $u_\pm(T)$ can be obtained explicitly by solving a quadratic equation, and are given by

$$u_\pm(t) = \frac{1}{2} - \frac{\rho\lambda}{1-e^{-\lambda t}} \pm \sqrt{\frac{1}{4} + \frac{(2\kappa_+ - \rho)\lambda}{1-e^{-\lambda t}} + \frac{\rho^2\lambda^2}{(1-e^{-\lambda t})^2}} \; .$$

The large-strike asymptotics for the implied volatility smile in the sense of Lee can be explicitly calculated by inserting $u_\pm$ into Proposition 8.1.

9.4. The Heston model in the stationary variance regime. In the Heston model the limit distribution of the variance process $(V_t)_{t\geq 0}$ is a Gamma distribution with parameters $(\frac{2\lambda\theta}{\zeta^2}, \frac{2\lambda}{\zeta^2})$. This is well-known, but can also be obtained by applying Proposition 6.1. The cumulant generating function $l(w)$ is thus given by

$$l(w) = -\frac{2\lambda\theta}{\zeta^2}\log\left(1 - \frac{\zeta^2}{2\lambda}w\right) \; ,$$

defined on $(-\infty, \frac{2\lambda}{\zeta^2})$, such that $l_+ = \frac{2\lambda}{\zeta^2}$. As before we have that $\chi(u) = \rho\zeta u - \lambda$, and we assume that $\chi(1) < 0$. In addition we define $\chi^+(u) = \rho\zeta u + \lambda$. By Theorem 7.2,

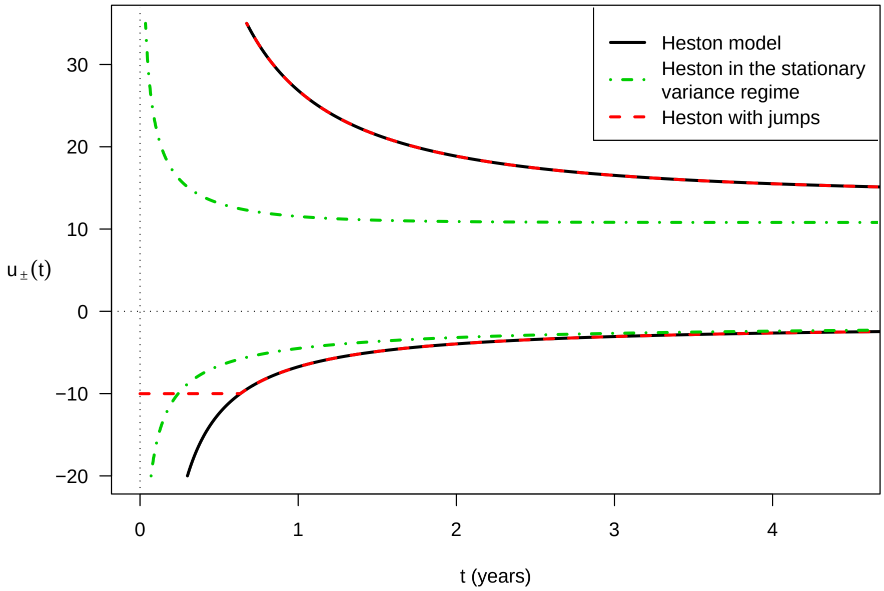

FIGURE 3. This plot shows the critical moment functions $u_\pm(t)$ for a Heston model with the same parameters as in Figure 2. Also shown are $u^S_\pm(t)$ for the model in the stationary variance regime, and $u^{\mathrm{Jump}}_\pm(t)$ for the Heston model with an independent jump component, whose negative jump heights are exponentially distributed with mean $\alpha = -0.1$. Note that $u^{\mathrm{Jump}}_\pm(t)$ coincides with $u_\pm(t)$ everywhere except in the lower left corner of the plot.

the explosion time in the stationary regime is given by

$$(9.12)\quad T^S_*(u) = \int_0^{2\lambda/\zeta^2} \frac{d\eta}{R(u,\eta)} =$$

$$= \begin{cases} \infty & \sqrt{\Delta(u)} > -\chi^+(u), \\ \frac{1}{\sqrt{\Delta}} \log \left| \frac{\chi^+\chi + 2\lambda\sqrt{\Delta} - \Delta}{\chi^+\chi - 2\lambda\sqrt{\Delta} - \Delta} \right| & 0 < \sqrt{\Delta(u)} < -\chi^+(u), \\ \frac{2}{\sqrt{-\Delta}} \arctan\left(\frac{2\lambda\sqrt{-\Delta}}{(\chi^+\chi) - \Delta} + \pi \mathbf{1}_{\{\chi^+\chi < \Delta\}} \right) & \Delta(u) < 0\,. \end{cases}$$

In Figure 3 $T^S_*(u)$ is plotted together with $T_*(u)$ for the Heston model.

9.5. The BNS model in the stationary variance regime. In the BNS model, the cumulant generating function of the limit distribution L of the variance process is given by Proposition 6.1 by

$$l(w) = \int_0^w \frac{\kappa(\eta)}{\eta} d\eta\,,$$

provided the log-moment condition $\int_{y>1}(\log y)\,\mu(dy) < \infty$ holds for the Lévy measure of $(J_t)_{t\geq 0}$. The above integral is finite as long as $w \in (-\infty, \kappa_+)$, and infinite outside. Thus $l_+ = \kappa_+$. In Section 9.3 we obtained that $f_+(u) = \kappa_+ - \rho u$, such that the time of moment explosion under stationary variance is given by

$$T_*^S(u) = -\frac{1}{\lambda}\log\max\left(0, 1 - \frac{2\lambda k(u)}{u(u-1)}\right) ,$$

where $k(u) = \kappa_+$ for $u \geq 1$ and $k(u) = \max(\kappa_+ - \rho u, 0)$ for $u \leq 0$. Again, this expression can be inverted to give the critical moment functions in the stationary variance case. By definition $\rho \leq 0$, such that we obtain

$$u_-^S(T) = \frac{1}{2} - \frac{\rho\lambda}{1-e^{-\lambda T}} - \sqrt{\frac{1}{4} + \frac{(2\kappa_+ - \rho)\lambda}{1-e^{-\lambda T}} + \frac{\rho^2\lambda^2}{(1-e^{-\lambda T})^2}}$$

$$u_+^S(T) = \frac{1}{2} + \sqrt{\frac{1}{4} + \frac{2\kappa_+\lambda}{1-e^{-\lambda T}}} .$$

10. Additional proofs for Part 2

Proof of Theorem 5.1. Let $t \leq \tau$. By the semi-flow equation we can write

$$\phi(\tau, u, \eta) = \phi(t, u, \eta) + \phi(\tau - t, u, \psi(t, u, \eta))$$
$$\psi(\tau, u, \eta) = \psi(\tau - t, u, \psi(t, u, \eta)) .$$

Since the left sides are finite by assumption, it follows that also $\phi(t, u, \eta)$ and $\psi(t, u, \eta)$ are. V_t is non-negative, such that

$$|\mathbb{E}[\exp(uX_t + wV_t)]| \leq |\mathbb{E}[\exp(uX_t + \eta V_t)]| ,$$

whenever $\operatorname{Re} w \leq \operatorname{Re}\eta$. Thus $\phi(t, u, w)$ and $\psi(t, u, w)$ exist for all $w \in \mathbb{C}$ with $\operatorname{Re} w \leq \operatorname{Re}\eta$. As a particular case we can conclude that $\phi(t, u, w)$ and $\psi(t, u, w)$ exist for all (u, w) in $\mathcal{U} := \left\{(u, w) \in \mathbb{C}^2 : \operatorname{Re} u = 0, \operatorname{Re} w \leq 0\right\}$.
We also define $\mathcal{U}^\circ := \left\{(u, w) \in \mathbb{C}^2 : \operatorname{Re} u = 0, \operatorname{Re} w < 0\right\}$, and show next that $\phi(t, u, w)$ and $\psi(t, u, w)$ are (right-)differentiable at $t = 0$ for all $(u, w) \in \mathcal{U}^\circ$. The key idea of our proof is originally due to Montgomery and Zippin [1955], and has also been presented in Filipović and Teichmann [2003] and Dawson and Li [2006]. First note that the identity

$$\mathbb{E}\left[wV_t e^{uX_t + wV_t}\right] = \left(\frac{\partial}{\partial w}\phi(t, u, w) + V_0\frac{\partial}{\partial w}\psi(t, u, w)\right) \cdot$$
$$\cdot \exp\left(\phi(t, u, w) + V_0\psi(t, u, w) + X_0 u\right)$$

shows that $\frac{\partial}{\partial w}\phi(t,u,w)$ and $\frac{\partial}{\partial w}\psi(t,u,w)$ exist, and are continuous for all $t \le \tau$ and $(u,w) \in \mathcal{U}^\circ$. By Taylor expansion it holds that

$$\int_0^s \psi(r,u,\psi(t,u,w))\,dr - \int_0^s \psi(r,u,w)\,dr = \int_0^s \frac{\partial}{\partial w}\psi(r,u,w)\,dr\,(\psi(t,u,w)-w) + o\big(|\psi(t,u,w)-w|\big)\ . \tag{10.1}$$

On the other side, using the semi-flow property, we calculate

$$\int_0^s \psi(r,u,\psi(t,u,w))\,dr - \int_0^s \psi(r,u,w)\,dr = \int_0^s \psi(r+t,u,w)\,dr - \int_0^s \psi(r,u,w)\,dr = \tag{10.2}$$

$$= \int_t^{s+t} \psi(r,u,w)\,dr - \int_0^s \psi(r,u,w)\,dr = \int_0^t \psi(r+s,u,w)\,dr - \int_0^t \psi(r,u,w)\,dr\ .$$

Denoting the last expression by $I(s,t)$, and putting (10.1) and (10.2) together, we obtain

$$\lim_{t\to 0} \frac{\left|\frac{1}{s}I(s,t)\right|}{|\psi(t,u,w)-w|} = \left|\frac{1}{s}\int_0^s \frac{\partial}{\partial w}\psi(t,u,w)\,dr\right|\ .$$

Thus, writing $M_s = \frac{1}{s}\int_0^s \frac{\partial}{\partial w}\psi(t,u,w)\,dr$, we have

$$\lim_{t\to 0}\frac{1}{t}\,|\psi(t,u,w)-w| = \left|\lim_{t\to 0}\frac{I(s,t)}{st}\right|\cdot |M_s|^{-1} = \left|\frac{\psi(s,u,w)-w}{s}\right| |M_s|^{-1}\ .$$

But M_s is a continuous function of s, and $\lim_{s\to 0} M_s = \frac{\partial}{\partial w}\psi(0,u,w) = 1$, such that for s small enough $M_s \neq 0$. We conclude that the left hand side is finite, and using (10.1) we obtain that

$$\lim_{t\to 0}\frac{\psi(t,u,w)-w}{t} = \left(\frac{\psi(s,u,w)-w}{s}\right)\cdot\left(\frac{1}{s}\int_0^s \frac{\partial}{\partial w}\psi(r,u,w)\,dr\right)^{-1}\ .$$

The finiteness of the right hand side implies the existence of the limit on the left. In addition the right hand side is continuous for $(u,w) \in \mathcal{U}^\circ$, showing that also the left hand side is. A similar calculation for $\phi(t,u,w)$ shows that

$$\lim_{t\to 0}\frac{\phi(t,u,w)}{t} = \frac{\phi(s,u,w)}{s} - \lim_{t\to 0}\left(\frac{\psi(t,u,w)-w}{t}\right)\cdot\left(\frac{1}{s}\int_0^s \frac{\partial}{\partial w}\phi(r,u,w)\,dr\right)\ ,$$

allowing the same conclusions for $\phi(t,u,w)$. We have thus shown that the time-derivatives of $\phi(t,u,w)$ and $\psi(t,u,w)$ at $t=0$ exist, and are continuous in $\mathcal{U}^\circ$. Combining Duffie et al. [2003, Proposition 7.2] and Duffie et al. [2003, Proposition 6.4] the differentiability can be extended from $\mathcal{U}^\circ$ to $\mathcal{U}$, and we have shown that $(X_t, V_t)_{t\ge 0}$ is a *regular* affine process. The rest of Theorem 5.1 follows now as in Duffie et al. [2003, Theorem 2.7] □

Proof of Lemma 5.2. We prove the assertions of Lemma 5.2 for F; they follow analogously for R. By the Lévy-Khintchine representation (5.6), $F(u,w)+c$ is the cumulant generating functions of some infinitely divisible random variables, say X. Writing $z = (u, w) \in \mathbb{R}^2$, and using Hölder's inequality it holds for any $\lambda \in [0, 1]$ that

$$F(\lambda z_1 + (1-\lambda)z_2) = \log \mathbb{E}\left[e^{\lambda\langle z_1, X\rangle} e^{(1-\lambda)\langle z_2, X\rangle}\right] - c \leq \tag{10.3}$$
$$\leq \lambda \log \mathbb{E}\left[e^{\langle z_1, X\rangle}\right] + (1-\lambda)\mathbb{E}\left[e^{\langle z_2, X\rangle}\right] - c = \lambda F(z_1) + (1-\lambda)F(z_2) ,$$

showing convexity of F. In addition equality in (10.3) holds if and only if $ke^{\langle z_1, X\rangle} = e^{\langle z_2, X\rangle}$ a.s. for some $k > 0$. This in turn is equivalent to $\langle z_1 - z_2, X\rangle$ being constant a.s. Choosing now z_1 and $z_2 \neq z_1$ from some one-dimensional affine subspace $U = \{p + \langle q, x\rangle : x \in \mathbb{R}\}$ of $\mathbb{R}^2$, we see that either $\langle q, X\rangle$ is constant a.s. in which case $F|_U$ is affine, or it is not constant, in which case strict inequality holds in (10.3) for all $z_1, z_2 \in U$, showing (c).

Let $L_\alpha = \{z : F(z) \leq \alpha\}$ be a level set of F, and $z_n \in L_\alpha$ a sequence converging to z. Then by Fatou's Lemma

$$\log \mathbb{E}[e^{\langle z, X\rangle}] - c \leq \liminf_{n\to\infty} \log \mathbb{E}[e^{\langle z_n, X\rangle}] - c \leq \alpha ,$$

showing that $z \in L_\alpha$ and thus that F is a closed convex function. Finally F is proper, because $F(0, 0) = c > -\infty$, showing (a).

Next we show analyticity: Consider the random variables $X_n := X\mathbf{1}_{\{|X|\leq n\}}$. Since they are bounded, their Laplace transforms, and hence also their cumulant generating functions are entire functions on $\mathbb{C}^2$, and thus analytic on $\mathbb{R}^2$. As a uniform limit of analytic functions $F(u, w)$ is analytic in the interior of dom F, showing (b).

Assertion (d) follows directly from Theorem 5.1. □

Appendix

A. Convex Analysis

A non-empty set $C \subset \mathbb{R}^d$, is convex, if for any two points $y, y' \in C$, it also contains the line segment connecting y and y', that is for any $\lambda \in (0,1)$

$$\lambda y + (1-\lambda) y' \in C \ .$$

A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, not identically $+\infty$, is called a **convex function**, if for all $y, y' \in \mathbb{R}^d$ and $\lambda \in (0,1)$, it holds that

$$f(\lambda x + (1-\lambda) x') \leq \lambda f(x) + (1-\lambda) f(x') \ . \tag{A.1}$$

The function f is called **strictly convex** if (A.1) holds with strict inequality whenever $y \neq y'$. For a convex function, we define its **effective domain**, often simply called 'domain', as the set where is takes finite values:

$$\operatorname{dom} f = \left\{ y \in \mathbb{R}^d : f(y) < +\infty \right\} \ .$$

It is easy to see that the domain of a convex function is also a convex set. A convex function that is lower semicontinuous on all of $\mathbb{R}^d$, is called **closed convex**. Closed convex functions have stronger regularity properties than convex functions and occur in probability theory as cumulant generating functions (see below).
We list some important results on convex functions, the proofs of which can all be found in Hiriart-Urruty and Lemaréchal [1993]:

Lemma A.1. *Any convex function on $\mathbb{R}^d$ is locally Lipschitz, and in particular continuous, on the interior of its domain*[17]*.*

Lemma A.2. *Any convex function on $\mathbb{R}$ is upper semicontinuous on its domain. It follows that a closed convex function on $\mathbb{R}$ is continuous not just in the interior, but also on the boundary of its domain. Moreover, any closed convex function on $\mathbb{R}^d$ is continuous along any line segment contained in its domain.*

Lemma A.3. *Let f_n be a sequence of convex functions on $\mathbb{R}^d$, converging pointwise to f. Then f is convex, and the convergence is uniform on every compact subset of the interior of* $\operatorname{dom} f$*.*

B. (Extended) cumulant and moment generating functions

Definition B.1 (sub-stochastic measure). A (Borel) measure μ on $\mathbb{R}^d$ is called sub-stochastic measure, if $0 < \mu(\mathbb{R}^d) \leq 1$.

[17]This and most other results can be generalized from the interior of the domain to the so-called *relative interior*; see Hiriart-Urruty and Lemaréchal [1993] for details.

For any sub-stochastic measure μ on $\mathbb{R}^d$ we introduce the following integral transforms:

$$\Phi(y) : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}, y \mapsto \int_{\mathbb{R}^d} e^{\langle y,\xi \rangle} \, d\xi$$

is called the **moment generating function** of μ, and

$$\kappa(y) : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}, y \mapsto \log \int_{\mathbb{R}^d} e^{\langle y,\xi \rangle} \, d\xi$$

is called its **cumulant generating function**.

PROPOSITION B.2. *(1) The moment generating function $\Phi(y)$ is a strictly positive, lower semi-continuous, log-convex function.*

(2) The cumulant generating function $\kappa(y)$ is a closed convex function with $\kappa(0) \leq 0$.

(3) Let $y \in \mathbb{R}^d$. Then the function $\mathbb{R} \ni s \mapsto \kappa(sy)$ is either strictly convex or linear.

PROOF. Let $y_n \to y$. Then $e^{\langle y_n,\xi \rangle}$ is a sequence of positive functions, such that by Fatou's Lemma

$$\liminf_{y_n \to y} \Phi(y) = \liminf_{y_n \to y} \int_{\mathbb{R}^d} e^{\langle y,\xi \rangle} \, d\xi \geq \int_{\mathbb{R}^d} \liminf_{y_n \to y} e^{\langle y,\xi \rangle} \, d\xi = \Phi(y)$$

showing lower semi-continuity of Φ. The positivity of the exponential function implies that also $\Phi(y)$ is positive. Consider now $\kappa(y)$. By Hölder's inequality

$$\kappa\left(\lambda y + (1-\lambda)y'\right) = \log \int_{\mathbb{R}^d} e^{\lambda \langle y,\xi \rangle} e^{(1-\lambda)\left\langle y',\xi \right\rangle} \, d\xi \leq$$

$$\leq \lambda \log \int_{\mathbb{R}^d} e^{\langle y,\xi \rangle} \, d\xi + (1-\lambda) \log \int_{\mathbb{R}^d} e^{\left\langle y',\xi \right\rangle} \, d\xi = \lambda\kappa(y) + (1-\lambda)\kappa(y')$$

for all $y, y' \in \mathbb{R}^d$ and $\lambda \in (0,1)$. since $\kappa(0) = \log \nu(\mathbb{R}^d) \in (-\infty, 0]$ we conclude that $\kappa(y)$ is convex and $\Phi(y)$ is log-convex. As a lower semicontinuous convex function κ is closed convex.

Consider now the function $\mathbb{R} \ni s \mapsto \kappa(sy)$ for some fixed $y \in \mathbb{R}^d$. Clearly this is a convex function. Applying Hölder's inequality to $\kappa((\lambda s + (1-\lambda)s')y)$ with $s \neq s'$, we see that equality holds if and only if there exists $k \in \mathbb{R}$ such that $ke^{s\langle y,\xi \rangle} = e^{s'\langle y,\xi \rangle}$ for μ-almost every $\xi \in \mathbb{R}^d$. This is equivalent to the function $\xi \mapsto \langle y, \xi \rangle$ being constant ($:= C$) almost everywhere. In this case

$$\kappa(sy) = \log \int_{\mathbb{R}^d} e^{s\langle y,\xi \rangle} = \log(C^s \mu(\mathbb{R}^d)) = s \log C + \log(\mu(\mathbb{R}^d)) \, ,$$

which is a linear function of s. In any other case strict inequality holds in Hölder's inequality, such that $s \mapsto \kappa(sy)$ is strictly convex. $\square$

Definition B.3. Let $\mathcal{D}_{\mathbb{C}} = \{u \in \mathbb{C}^d : \Phi(\operatorname{Re} u) < \infty\}$. Then we can define the **extended moment generating function**

$$\Phi(u) : \mathcal{D}_{\mathbb{C}} \to \mathbb{C} : u \to \int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, \mu(d\xi)\ .$$

If μ is infinitely divisible[18], then by Lemma C.3 $\Phi(u)$ cannot have zeroes in $\mathbb{C}^n$, and we can take the logarithm to define the **extended cumulant generating function**

$$\Phi(u) : \mathcal{D}_{\mathbb{C}} \to \mathbb{C} : u \to \log \int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, \mu(d\xi)\ .$$

Proposition B.4. *(i) The extended moment generating function and the extended cumulant generating functions are analytic on $\mathcal{D}_{\mathbb{C}}^{\circ}$.*

(ii) Let $y \in \mathbb{R}^d$ such that $\Phi(y) < \infty$, and set $\Omega = \{z \in \mathbb{C} : \operatorname{Re} z \in \mathbb{R}, \operatorname{Im} z \in [0,1]\}$. Then $h : \Omega \to \mathbb{C} : z \mapsto \Phi(izy)$ is analytic on Ω° and continuous on Ω.

Proof. Define $\mu_n(d\xi) := \mathbf{1}_{\{|\xi|\le n\}}\mu(d\xi)$. It holds that $\left|e^{\langle u,\xi\rangle}\right| = e^{\langle \operatorname{Re} u,\xi\rangle}$, such that dominated convergence allows us to conclude

$$\Phi_n(u) := \int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, \mu_n(d\xi) \to \int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, \mu(d\xi) = \Phi(u)$$

pointwise for each $u \in \mathcal{D}_{\mathbb{C}}^{\circ}$ as $n \to \infty$. We show now that the convergence is uniform on every compact subset K of $u \in \mathcal{D}_{\mathbb{C}}^{\circ}$:

$$\text{(B.1)}\qquad \left|\int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, \mu(d\xi) - \int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, \mu_n(d\xi)\right| = \left|\int_{\mathbb{R}^d} e^{\langle u,\xi\rangle}\, (\mu-\mu_n)\,(d\xi)\right| \le$$
$$\le \int_{\mathbb{R}^d} e^{\langle \operatorname{Re} u,\xi\rangle}\, (\mu-\mu_n) = \Phi(\operatorname{Re} u) - \Phi_n(\operatorname{Re} u).$$

Since the exponential function is locally Lipschitz, we can estimate the last term on the compact set K by

$$\sup_{u\in K} |\Phi(\operatorname{Re} u) - \Phi_n(\operatorname{Re} u)| \le L_K \sup_{u\in K} |\kappa(\operatorname{Re} u) - \kappa_n(\operatorname{Re} u)|\ .$$

The right hand side is the difference of two convex functions converging pointwise. But pointwise convergence of convex functions implies uniform convergence on compact subset of the interior of their domains (cf.Hiriart-Urruty and Lemaréchal [1993].) We conclude that also Φ_n converges uniformly on compact subsets of $\mathcal{D}_{\mathbb{C}}^{\circ}$ to Φ. On the other hand, Φ_n is – as the Fourier-Laplace transform of a measure with bounded support – an entire function by the Paley-Wiener-Schwartz theorem. In particular $\Phi_n(u)$ is analytic on $\mathcal{D}_{\mathbb{C}}^{\circ}$, and since analyticity is preserved by uniform convergence on compacts, so is $\Phi(u)$.

Regarding claim (ii), it is clear that h is analytic on Ω°, as the composition of

[18] See Section C.

analytic maps. To show continuity note that

$$(B.2)\qquad |\Phi(izy)| = \left|\int_{\mathbb{R}^d} e^{iz\langle y,\xi\rangle}\,\mu(d\xi)\right| \le \int_{\mathbb{R}^d} e^{\operatorname{Im} z\langle y,\xi\rangle}\,\mu(d\xi) = \Phi(\operatorname{Im} z\cdot y)\ .$$

Taking the logarithm on the right side we know that $s \mapsto \log\Phi(su)$ is a finite closed convex function on $[0,1]$. By Lemma (A.2), it is therefore continuous, taking a maximum at some point in $[0,1]$. This means that the right hand side of (B.2) can be bounded uniformly for $\operatorname{Im} z \in [0,1]$. Dominated convergence then proves the continuity of h on Ω. □

C. Infinite Divisibility and related notions

Definition C.1. An $\mathbb{R}^d$-valued random variable X is called infinitely divisible, if for any $n \in \mathbb{N}$ there exist iid random variables $X_1, \dots, X_n$, independent of X, such that

$$X = X_1 + \cdots + X_n$$

The term 'truncation function' has no consistent definition throughout the literature on infinitely divisible distributions. We will use a more refined terminology than most texts and define the following:

Definition C.2. (1) Let ν be a Lévy measure on $\mathbb{R}^d$. We say that a function $h : \mathbb{R}^d \to \mathbb{R}^d$ is a **truncation function for** ν, if it is bounded, continuous, and satisfies

$$\int_{\mathbb{R}^d} (|h(\xi) - \xi| \wedge 1)\ \nu(d\xi) < \infty\ .$$

(2) We say that $h : \mathbb{R}^d \to \mathbb{R}^d$ is a **(universal) truncation function**, if it is a truncation function for any Lévy measure on $\mathbb{R}^d$, or equivalently, if h is bounded, continuous, and satisfies

$$h(\xi) = \xi + \mathcal{O}(|\xi|^2)\ .$$

Frequently used truncation functions are

$$h(\xi) = \mathbf{1}_{\{|\xi|\le 1\}}\xi \quad \text{and} \quad h(\xi) = \frac{\xi}{1+|\xi|^2}\ .$$

Note that the latter one is in $C^\infty(\mathbb{R}^d)$.

Lemma C.3. *The extended moment generating function $\Phi(u) : \mathcal{D}^{\mathbb{C}} \to \mathbb{C}$ of an infinitely divisible sub-stochastic measure has no zeroes.*

Proof. Define $\widetilde{\Phi}(y) = \frac{\Phi(u)}{\Phi(0)}$. The function $\widetilde{\Phi}$ is the extended moment generating function of the infinitely divisible probability measure $\mu(d\xi)/\mu(\mathbb{R}^d)$. Clearly $\widetilde{\Phi}$ has the same zeroes as Φ. Let $\Xi(u) := \lim_{n\to\infty} \widetilde{\Phi}(u)^{1/n}$. By infinite divisibility, $\widetilde{\Phi}(u)^{1/n}$ is a well-defined moment generating function of some probability measures for each n. By continuity of the moment generating function there exists a neighborhood N of 0 such that $\widetilde{\Phi}(y) \neq 0$ in N. It follows that $\Xi(u) = 1$ in N, and thus that also $\Xi(u)$ is a moment generating function. But a moment generating function that takes the value 1 on a set with an accumulation point, must equal 1 everywhere. It follows that $\widetilde{\Phi}(y)$ can have no zeros, and the same must hold for $\Phi(y)$. □

Theorem C.4 (Lévy-Khintchine formula). *Let X be an infinitely divisible random variable on $\mathbb{R}^d$, then*

$$\mathbb{E}[e^{\langle u,X\rangle}] = \exp\left[\frac{1}{2}\langle u, Au\rangle + \langle b, u\rangle + \int_{\mathbb{R}^d}\left(e^{\langle x,u\rangle} - 1 - \langle h(x), u\rangle\right) m(dx)\right]$$

for all $u \in i\mathbb{R}^d$, where A is a symmetric nonnegative definite $d \times d$-matrix, $b \in \mathbb{R}^d$, and m is a Borel measure on $\mathbb{R}^d$ satisfying

$$m(\{0\}) = 0 \quad \textit{and} \quad \int_{\mathbb{R}^d}\left(|x|^2 \wedge 1\right) m(dx) < \infty\,.$$

Proof. Sato [1999, Theorem 8.1]. □

Theorem C.5. *Let X be an infinitely divisible random variable on $\mathbb{R}^d_{\geqslant 0}$, then*

$$\mathbb{E}[e^{\langle u,X\rangle}] = \exp\left[\langle \gamma, u\rangle + \int_{\mathbb{R}^d_{\geqslant 0}}\left(e^{\langle x,u\rangle} - 1\right) m(dx)\right]$$

for all $u \in \mathbb{C}^d_-$, where $\gamma \in \mathbb{R}^d_{\geqslant 0}$, and m is a Borel measure on $\mathbb{R}^d_{\geqslant 0}$ satisfying

$$m(\{0\}) = 0 \quad \textit{and} \quad \int_{\mathbb{R}^d_{\geqslant 0}}\left(|x| \wedge 1\right) m(dx) < \infty\,.$$

Proof. Sato [1999, Chapter 24]. □

Definition C.6. Let $\lambda > 0$ and p a probability measure on $\mathbb{R}^d$. A distribution is called compound Poisson distribution with intensity λ and kernel p if its characteristic function is given by

$$\phi(u) = \exp\left(\lambda \int_{\mathbb{R}^d}\left(e^{\langle \xi,u\rangle} - 1\right) p(d\xi)\right)$$

Proof. Sato [1999, Corollary 8.8]. □

Proposition C.7. *(a) A compound Poisson distribution is infinitely divisible with triplet $(0, 0, \lambda p)$.*

(b) A distribution is infinitely divisible if and only if it is the limit of compound Poisson distributions.

Lemma C.8 (Linear Transformation Lemma). *Let T be a linear map from $\mathbb{R}^d$ to $\mathbb{R}^k$ (in particular T could be a projection), let h and h' be truncation functions on $\mathbb{R}^d$ and $\mathbb{R}^k$ respectively, and let K be a infinitely divisible random variable on $\mathbb{R}^d$*

with Lévy triplet $(A, B, M)_h$. *Then* $T \cdot K$ *is a infinitely divisible random variable on* $\mathbb{R}^k$ *with Lévy triplet* $(A', B', M')_{h'}$, *where*

$$A' = TAT^* \tag{C.1a}$$

$$B' = TB + \int_{\mathbb{R}^d} (T \circ h - h' \circ T)(\xi)\, m(d\xi) \tag{C.1b}$$

$$M'(d\xi) = M(T^{-1} d\xi) \tag{C.1c}$$

PROOF. Sato [1999, Prop. 11.10]. □

Bibliography

Leif B. G. Andersen and Vladimir V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11:29–50, 2007.

Ole E. Barndorff-Nielsen. Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2:41 – 68, 1997.

Ole E. Barndorff-Nielsen and Neil Shephard. Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society B*, 63:167–241, 2001.

David S. Bates. Jump and stochastic volatility: exchange rate processes implicit in Deutsche Mark options. *The Review of Financial Studies*, 9:69–107, 1996.

David S. Bates. Post-'87 crash fears in the S&P 500 futures option market. *Journal of Econometrics*, 94:181–238, 2000.

Shalom Benaim and Peter Friz. Regular variation and smile asymptotics. arXiv:math/0603146v2, to appear in Mathematical Finance, 2006.

P. Carr and D. B. Madan. Option valuation using the fast fourier transform. *Journal of Computational Finance*, 2(4), 1999.

Peter Carr and Liuren Wu. Time-changed Lévy processes and option pricing. *Journal of Financial Economics*, 71(1):113–141, 2004.

Rama Cont and Peter Tankov. *Financial Modelling with Jump Processes*. Financial Mathematics Series. Chapman & Hall/CRC, 2004.

John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross. A theory on the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

Qiang Dai and Kenneth J. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, 55(5):1943–1977, 2000.

D. A. Dawson and Zenghu Li. Skew convolution semigroups and affine markov processes. *The Annals of Probability*, 34(3):1103 – 1142, 2006.

D. Duffie, D. Filipovic, and W. Schachermayer. Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053, 2003.

Darrell Duffie. Credit risk modeling with affine processes, cattedra galileana lectures, scuola normale, pisa. *Journal of Banking and Finance*, 29:2751–2802, 2005.

Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343 – 1376, 2000.

Steward N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley, 1986.

William Feller. *An introduction to probability theory and its applications*. John Wiley & Sons, 2nd edition, 1971.

Damir Filipović. A general characterization of one factor affine term structure models. *Finance and Stochastics*, 5:389–412, 2001.

Damir Filipović and Josef Teichmann. Regularity of finite-dimensional realizations for evolution equations. *Journal of Functional Analysis*, 197:433–446, 2003.

Jim Gatheral. *The Volatilty Surface*. Wiley Finance, 2006.

E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I*. Springer, 1987.

Philip Hartman. *Ordinary Differential Equations*. Birkhäuser, 1982.

S. Heston. A closed-form solution of options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6:327–343, 1993.

Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Analysis and Minimization Algorithms I*. Number 305 in Grundlehren der mathematischen Wissenschaften. Springer, 1993.

Niels Jacob. *Pseudo Differential Operators and Markov Processes*, volume I. Imperial College Press, 2001.

Olav Kallenberg. *Foundations of Modern Probability*. Springer, 1997.

Jan Kallsen and Johannes Muhle-Karbe. Exponentially affine martingales, affine measure changes and exponential moments of affine processes. Preprint, 2008.

Anatole Katok and Boris Hasselblatt. *Modern Theory of Dynamical Systems*. Cambridge University Press, 1999.

Kiyoshi Kawazu and Shinzo Watanabe. Branching processes with immigration and related limit theorems. *Theory of Probability and its Applications*, XVI(1):36–54, 1971.

Martin Keller-Ressel and Thomas Steiner. Yield curve shapes and the asymptotic short rate distribution in affine one-factor models. *Finance and Stochastics*, 12 (2):149 – 172, 2008.

Roger Lee. The moment formula for implied volatility at extreme strikes. *Mathematical Finance*, 14(3):469–480, 2004.

Alan L. Lewis. *Option Valuation under Stochastic Volatility*. Finance Press, 2000.

Pierre-Louis Lions and Marek Musiela. Correlations and bounds for stochastic volatility models. *Annales de l'Institut Henri Poincaré*, 24:1–16, 2007.

Deane Montgomery and Leo Zippin. *Topological Transformation Groups*. Interscience Publishers, Inc., 1955.

W. F. Osgood. Beweis der Existenz einer Lösung der Differentialgleichung $\frac{dy}{dx} = f(x, y)$ ohne Hinzunahme der Cauchy-Lipschitz'schen Bedingung. *Monatshefte* *für Mathematik und Physik*, 9:331–345, 1898.

Z. Palmowski and T. Rolski. The technique of the exponential change of measure for Markov processes. *Reports of the Mathematical Institute at the University of Wroclaw*, 119, 2000.

Philip E. Protter. *Stochastic Integration and Differential Equations*. Springer, 2004.

Daniel Revuz and Marc Yor. *Continuous Martingales and Brownian Motion*. Springer, 3rd edition, 1999.

L.C.G. Rogers and David Williams. *Diffusions, Markov Processes and Martingales, Volume 1*. Cambridge Mathematical Library, 2nd edition, 1994.

Ken-Iti Sato. *Lévy processes and infinitely divisible distributions*. Cambridge University Press, 1999.

Zbigniew Semadeni. *Banach spaces of continuous functions*. Polish Scientific Publishers, 1971.

Oldrich Vasiček. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5:177–188, 1977.

Peter Volkmann. Gewöhnliche Differentialungleichungen mit quasimonoton wachsenden Funktionen in topologischen Vektorräumen. *Mathematische Zeitschrift*, 127:157–164, 1972.

Wolfgang Walter. *Ordinary Differential Equations*. Springer, 6th edition, 1996.

Kosaku Yosida. *Functional Analysis*. Springer, 6th edition, 1995.