SCUOLA NORMALE SUPERIORE

Generalized Feller Processes and their Applications to Affine and Polynomial Processes

Candidato

dr. Tonio Möllmann

Relatori

Prof. Josef Teichmann

Prof.ssa Christa Cuchiero

Supervisore interno

Prof. Stefano Marmi

Classe di Scienze

Corso di perfezionamento in Matematica per la Finanza

XXXII Ciclo

Anno accademico 2021-2022

Contents

Acknowledgment

I thank my supervisors Christa Cuchiero and Josef Teichmann for many long, fruitful, and inspiring discussions especially during the final year of my thesis. Their assistance, intuition, and extensive knowledge of mathematical literature helped me to overcome many difficulties in the proofs of my theorems. With great passion they directed my research to interesting problems and topics. I especially enjoyed the open communication and excellent working atmosphere they were able to create.

I also thank Maria Elvira Mancino for her lectures and her and Stefano Marmi for the advice they gave me during my PhD studies.

I am very grateful to Livia Ristuccia, Danilo Vassallo, Domenico Di Gangi, Matteo Ottaviani, and especially Maddalena Ghio who started the PhD together with me for the shared experience and community at Scuola Normale that helped me to live better through the common hardships of our PhD studies.

I thank Carlo Campajola, Francesco Cordoni, Giulia Livieri for organizing many common activities in Pisa making life outside of university more pleasant. I thank Kevin Max for being an excellent fellow German expatriate that shared my curiosity for Italian culture and always gave me the possibility to speak in my familiar mother tongue when it felt like it.

I thank Akie Kobiyama, Maria Paola Bissiri, Isha Verma and Abishek Kumar for uncountable dinners at the mensa and elsewhere and for the long conversations about Japan, India, Italy and anything else that opened up new perspectives for me, made me laugh, helped me through difficult times during my research and simply made me feel at home in Pisa.

I thank Scuola Normale for funding a large part of my PhD studies and offering the facilities in Pisa that make it possible to create a good community between students of Scuola Normale.

I especially thank Akie Kobiyama and my parents who immensely supported me during very tough 18 months of pandemic, 12 of which without a workplace at Scuola Normale. Without them and their offer of shelter and workplace, it would not have been possible to finish my thesis. Thanks to them, I was able to live and work in very confined spaces for this long period while enjoying their harmonic personalities and always pleasant company.

Introduction

The aim of this thesis is to study *generalized Feller processes* and *extended Feller processes* and to connect them to the theory of affine and polynomial processes. This includes in particular a comprehensive investigation of generalized Feller semigroups.

Generalized Feller semigroups are defined in analogy to Feller semigroups where the space of functions vanishing at infinity is replaced by so-called $\mathscr{B}^\rho(E)$ -spaces. Here, $\mathscr{B}^\rho(E)$ denotes the space of functions on a completely regular topological space E that do not grow faster than a so-called admissible weight function ρ and that lie in the closure of continuous bounded functions with respect to the norm on $\mathscr{B}^\rho(E)$ which is a weighted supremum norm induced by ρ. Therefore - unlike Feller semigroups - generalized Feller semigroups act also on unbounded functions, all other properties are similar. More precisely, a generalized Feller semigroup is a family of positive linear bounded operators from $\mathscr{B}^\rho(E)$ to $\mathscr{B}^\rho(E)$ such that the semigroup properties are fulfilled, the norm of the operators remains bounded for small times and for any map f in $\mathscr{B}^\rho(E)$ the image under the semigroup converges pointwise to f as t approaches 0. In a special setting generalized Feller semigroups were introduced by Röckner and Sobol in [**36**] and generalized in 2010 by Dörsek and Teichmann in [**15**]. They proved that on $\mathscr{B}^\rho(E)$ there is a Riesz representation theorem and showed that just like Feller semigroups generalized Feller semigroups turn out to be strongly continuous. In the article [**15**], Dörsek and Teichmann used this to show convergence of splitting schemes for semigroups related to stochastic partial differential equations and acting on functions of controlled growth.

Versions of Markov processes corresponding to generalized Feller semigroups, so-called generalized Feller processes were considered by Cuchiero and Teichmann [**14**]. There they are used in order to show existence and uniqueness of solutions of certain stochastic partial differential equations corresponding to infinite dimensional affine processes whose finite dimensional projections lead to (rough) affine Volterra processes.

This thesis also treats affine processes as well as polynomial processes. Affine processes are continuous-time Markov processes that are stochastically continuous and for which the logarithm of the Fourier-Laplace transform of the marginal distributions is an affine map of the initial value. This includes for instance Lévy processes, squared Bessel processes, Ornstein-Uhlenbeck processes or Wishart processes, depending on the considered state spaces. Indeed, on $\mathbb{R}_+$ affine processes were first systematically analyzed in 1971 by Kawazu and Watanabe [**28**]. On the canonical state space $\mathbb{R}_+^n \times \mathbb{R}^m$ Duffie, Filipović, and Schachermayer provided a full characterization in [**16**]. Among many other properties they showed in particular that on the canonical state space affine processes are Feller processes. In 2013, Cuchiero and Teichmann [**13**] then considered affine processes more generally on subsets of a finite dimensional real vector space. They showed that all affine processes admit càdlàg versions. As it is (still) not known whether affine processes on general state spaces are Feller or not, this property could not be deduced therefrom as in the canonical state space, but needed to be proved by different methods. Affine processes can also be viewed as semimartingales with differential characteristics that depend in an affine way on the state of the process. This as well as their analytical tractability and flexibility make affine processes popular choices for modeling in Mathematical Finance, in particular for stochastic volatility modeling. Examples include the well-known Heston model [**22**] or Barndorff-Nielson and Shepard model [**3**]. Thanks to Fourier-inversion methods the expected value of many pay-off functions can be calculated by just solving a generalized Riccati equation, which is important in view of option pricing. For interest rate modeling so-called α-CIR models have been recently considered in [**25**] which show high flexibility in capturing persistency of low interest rates together with the presence of large jumps. An important generalization of affine processes beyond the assumption of stochastic continuity, where times of jumps can be both inaccessible and predictable, has been considered by Keller-Ressel et al. in [**29**]. There, a general theory of finite dimensional affine semimartingales (beyond dt characteristics) is developed and it is shown that the conditional characteristic function can be computed by solving measure differential equations of Riccati type.

An extension of affine processes are m-polynomial processes introduced by Cuchiero et al. in [**12**]. They are essentially continuous-time Markov processes such that for any k smaller than m the Markov semigroup maps polynomials of degree k to polynomials of the same or lower degree. They are a special class of semimartingales that includes all affine processes except some trivial cases and provided that their Lévy measure admits moments up to order m. Polynomial processes permit to calculate mixed moments in an easy way and facilitate parameter estimation via generalized methods of moments or variance reduction in the Monte Carlo simulations for pricing European claims.

In the following we explain in more detail the structure of the thesis and the main contributions. After introducing general semigroup theory in Chapter 1, as well as Markov and Feller semigroups in Chapter 2.1 and 2.2, in the first four subsections of Chapter 2.3 the foundations of generalized Feller semigroups are explained. While the first chapters, up to Chapter 2.2 provide a literature review, Chapter 2.3 contains several new results contributing to the theory of generalized Feller processes, which are highlighted in *italic* subsequently. In Proposition 2.3.46 we start by showing a *weighted space version of the Stone-Weierstraß theorem* (in the spirit of Leopoldo Nachbin) on the space $\mathscr{B}^\rho(E)$, which we apply in the existence proof of generalized Feller processes. In Proposition 2.3.54 we characterize *generalized Feller semigroups of transport type* which are generalized Feller semigroups such that at any given time the semigroup operator can be described by a composition of functions in $\mathscr{B}^\rho(E)$ with a map from E to E.

In Theorem 2.3.65, under the condition that the generalized Feller semigroup maps the constant function 1 to itself, the *existence of a generalized Feller processes* is rigorously proved. In particular, this yields stochastic processes whose conditional expectations are given by a strongly continuous semigroup even in cases when the space E is neither separable nor locally compact. This is a crucial difference to the theory of Feller processes and thus one of the main results of the thesis. Let us mention here also that generalized Feller processes are usually not classical Markov processes in the sense that the Markov property holds for *all* Borel-measurable functions. Indeed, it only holds for Baire-measurable functions, hence generalized Feller processes are strictly speaking only Markovian if the chosen σ-algebra on E is the Baire σ-algebra (for this subtle point see Remark 2.3.66).

The proof of Theorem 2.3.65 relies on a general version of the Kolmogorov Extension Theorem. In order to apply it, we construct a projective family of probability measures. For this purpose, on the sub-level sets of the admissible weight function on $E \times E$ we find a continuous linear functional which can be represented by a (sub-) probability measure via the Riesz representation for $\mathscr{B}^\rho(E \times E)$. As we let the sub-level set of the admissible weight function converge to the whole space, such a sequence of (sub-) probability measures converges to a probability measure on $E \times E$. Inductively this then yields a projective family of probability measures with the desired properties. Then the generalized Feller process is the canonical process on the product space when equipped with a product measure according to the general version of the Kolmogorov Extension Theorem. While in the case of general admissible weight functions only Dirac distributions are admitted as initial distributions, which is due to the subtle measurability issues explained above, we show in Proposition 2.3.69 that for admissible weight functions that are Baire measurable it is possible to use general Radon measures as initial distributions.

In Definition 2.3.71 we use again Baire-measurable admissible weight functions ρ to introduce a new class of stochastic processes called *extended Feller processes*. In Theorem 2.3.73 and Corollary 2.3.74 we prove existence of these processes under the condition that the generalized Feller semigroup is quasi-contractive. We *compare extended Feller processes with generalized Feller processes* and notice in Proposition 2.3.79 that if both exist, their induced laws are equivalent measures.

We also *compare generalized Feller processes and extended Feller processes with classical Feller processes.* In Proposition 2.3.84 we see that on locally compact spaces E a Feller process is a generalized Feller process if the Feller semigroup applied to the admissible weight function remains bounded for small times. Moreover, in Theorem 2.3.93 we show that for continuous admissible weight functions extended Feller processes can be reduced to Feller processes. This extends the notion of Feller processes to spaces E that are not necessarily separable but only σ-compact.

In Theorem 2.3.96 we show that generalized Feller processes admit a version that is càdlàg or càglàd if several conditions are met thereby closing a gap in a statement in [**14**]. A similar statement for extended Feller processes is proved in Theorem 2.3.99.

Chapter 3 of this thesis *relates the theory of generalized Feller processes to affine and polynomial processes.* We show in Proposition 3.2.7 that under certain conditions an m-polynomial process is a generalized Feller process. It follows then in Corollary 3.2.8 that under similar conditions also affine processes are generalized Feller processes. This adds to the existing theory of affine and polynomial processes since to date it is not known whether affine and polynomial processes on generalized state spaces are Feller or not.

In the last line of research we use the fact that the Fourier-Laplace transform of an affine process is given as the solution of an ordinary differential equation (ODE). We turn this idea around and in Theorem 3.3.2 obtain a *stochastic representation of the solution of a large class* *of ordinary differential equations via affine processes.* Since the vector fields of the involved ordinary differential equations are not necessarily locally Lipschitz continuous we obtain solutions also in cases where standard ODE theory does not apply.

Notation

$\langle \cdot, \cdot \rangle$	Dual pair (see Definition 1.4.51)
$\mathscr{B}^\rho(E;Z)$	closure of $C_b(E,Z)$ in $B^\rho(E;Z)$
$\mathcal{B}(E)$	Borelσ-algebra on the topological space E
$\mathcal{M}(T))$	space of signed Radon measures on topological Hausdorff space T
$\mathcal{M}_+(T))$	set of Radon measures on topological Hausdorff space T
$\mathcal{M}_1(\Omega,\Sigma)$	set of probability measures on measurable space (Ω,Σ)
$\mathcal{M}_\sigma(\Omega,\Sigma)$	set of σ- finite measures on measurable space (Ω,Σ)
$\mathcal{M}_c(T))$	space of signed Radon measures on topological Hausdorff space T
$B^\rho(E;Z)$	see Definition 2.3.20
$C_b(E)$	bounded continuous maps from E to $\mathbb{R}$
$C_b(E,Z)$	bounded continuous maps from E to Z
$C_c(E)$	continuous maps with compact support from E to $\mathbb{R}$
$C_c(E,Z)$	continuous maps with compact support from E to Z

CHAPTER 1

One-Parameter Semigroups

This chapter is mostly based on the book by Nagel and Engel [**18**] and on parts of the lecture notes [**20**]. For a background from functional analysis it follows [**41**] and [**39**]. It provides a literature review of semigroup theory which will be used thoughout this thesis. The exposition is kept as self-contained as possible.

1.1. Definition and Motivation

DEFINITION 1.1.1. For an index set $I = \mathbb{R}$ or $I = \mathbb{R}_+$ a one-parameter family of mappings

$$(T(t))_{t\in I}$$

that map from the state space Z into itself is said to satisfy the *functional equation* if for all $t, s \in I$

$$T(t+s) = T(t) \circ T(s). \tag{1.1.1}$$

REMARK 1.1.2. We note that $(T(t))_{t\in\mathbb{R}}$ is an (*algebraic*) *group* of maps from the state space Z onto itself equipped with the composition as group multiplication. Similarly, $(T(t))_{t\in\mathbb{R}_+}$ is a (*algebraic*) *semigroup* of maps from the state space Z onto itself equipped with the composition as semigroup multiplication. The map $t \to T(t)$ is a group homomorphism between the additive group $(\mathbb{R}, +)$ and the group mentioned above or a semigroup homomorphism between the additive semigroup $(\mathbb{R}_+, +)$ and the semigroup mentioned above.

One-parameter families of mappings that satisfy the functional equation often arise in physical systems. This is outlined in the following example taken from Nagel and Engel ([**18**], Epilogue, Section 1).

We consider a map

$$z : \mathbb{R} \to Z,$$

that maps time into the state space Z. For example, if we think of a physical system such as the motion of planets we might want to look at a map z that maps time to position and velocity of a planet. In this case the state space Z is $\mathbb{R}^6$ which is the product space of all possible positions $\mathbb{R}^3$ and all possible velocity vectors $\mathbb{R}^3$. We now consider the set of these maps and make some additional assumptions on it. Namely, we assume:

ASSUMPTION 1.1.3. *For each starting time $t_0 \in \mathbb{R}$ and each starting point x_0 there exists a unique map $z_{t_0,x_0} : \mathbb{R} \to Z$ such that*

$$z_{t_0,x_0}(t_0) = x_0.$$

We also assume:

ASSUMPTION 1.1.4. *For all maps z and all starting times $u \in \mathbb{R}$ and $v \in \mathbb{R}$ and all starting points x_0 it holds that $z_{u,x_0}(t+u) = z_{v,x_0}(t+v)$ for all $t \in \mathbb{R}$.*

In other words, our first assumption means that we can start the map at any time and any point and that this is done in a unique way and the second means that the way the map evolves after its start does not depend on its starting time. We show in the following that these two assumptions give rise to a one-parameter family of mappings that satisfy the functional equation.

If we fix a starting time $t_0 \in \mathbb{R}$ and fix a time $t \in \mathbb{R}$ the first assumption implies that for any starting point x_0 there is a unique map $z_{t_0,x_0} : \mathbb{R}_+ \to Z$ such that $z_{t_0,x_0}(t_0) = x_0$ and we can evaluate this map at $t_0 + t$. This way we can define a map by

$$\begin{aligned} T_{t_0,t_0+t} \quad &: Z \to Z \\ &x_0 \to z_{t_0,x_0}(t_0 + t) \end{aligned}$$

By the second assumption $z_{t_0,x_0}(t_0 + t)$ and thus also T_{t_0,t_0+t} depend only on t and not on t_0. This allows us to define a map $T(t) := T_{t_0,t_0+t}$ as

$$\begin{aligned} Z &\to Z \\ x_0 &\to z_{0,x_0}(t). \end{aligned}$$

We now show that $T(t) \circ T(s) = T(t+s)$. We fix some arbitrary $x \in Z$ and see that for $s, t \in \mathbb{R}$

$$T(t) \circ T(s)x = T(t)\left(z_{0,x}(s)\right) = z_{0,z_{0,x}(s)}(t).$$

By the second assumption we obtain

$$z_{0,z_{0,x}(s)}(t) = z_{s,z_{0,x}(s)}(t+s).$$

$z_{s,z_{0,x}(s)}$ is defined as the map such that its evaluation at its starting time s is its starting point $z_{0,x}(s)$. Written out this is $z_{s,z_{0,x}(s)}(s) = z_{0,x}(s)$. But since the map $z_{0,x}$ evaluated at s gives the same value as $z_{s,z_{0,x}(s)}$ evaluated at s and the map is unique by the first assumption we obtain $z_{s,z_{0,x}(s)}(\cdot) = z_{0,x}(\cdot)$ and

$$z_{s,z_{0,x}(s)}(t+s) = z_{0,x}(t+s) = T(t+s)x.$$

Since $x \in Z$ was arbitrary this implies $T(t) \circ T(s) = T(t+s)$. Hence

$$(T(t))_{t\in\mathbb{R}}$$

is a one-parameter family of mappings that satisfy the functional equation.

1.2. Linear bounded operators on a Banach space

As state spaces of the one-parameter family of maps

$$(T(t))_{t\in\mathbb{R}}$$

from Definition 1.1.1 we would like to consider Banach spaces (that is a complete normed vector space) over a field $\mathbb{K}$. The field $\mathbb{K}$ chosen is usually $\mathbb{C}$ or $\mathbb{R}$.

Additionally, in the one-parameter family of maps we will constrain ourselves to maps that are *linear* (see Definition1.2.2) and *bounded* (see Definition 1.2.4). Before working with such a one-parameter family of maps we will study this space of linear bounded maps on a Banach space. We will need the theory of functional analysis. For the convenience of the reader, we state its terminology and prove important assertions whenever we need them.

EXAMPLE 1.2.1. The probably simplest example of a Banach space is $\mathbb{R}^n$ for $n \in \mathbb{N}$ and the euclidean norm $|x| := \sqrt{(x_1)^2 + ... + (x_n)^2}$ where $x = (x_1, ..., x_n)$. Other important examples include L^p-spaces for $1 \le p < \infty$ (see Example A.4.1), L^∞-spaces (see Example A.4.2), or the space of bounded functions ℓ^∞ (see Proposition A.4.4).

We turn our attention to *linear bounded* operators. This part is based on Chapter II in [**41**].

DEFINITION 1.2.2. Let X, Y be vector spaces with a norm$\|\cdot\|$. A map

$$L : X \to Y$$

is called linear if for all $v, u \in X$ and $\lambda \in \mathbb{C}$ it holds

$$L(\lambda u + \lambda v) = \lambda L(u) + \lambda L(v).$$

A linear map is also called *operator*. We will use the expression map and operator interchangeably.

We want to look at the set of all such maps and we would like to find a norm on it.

DEFINITION 1.2.3. For a linear map $L : X \to Y$ between two normed vector spaces X and Y we define a map $\|\cdot\|$ between the space of such linear maps and $\mathbb{R}_+$ by

$$\|L\| := \inf \{M_L \in \mathbb{R}_+ : \ \|L(x)\| \leq M_L \|x\| \, \text{for all } x \in X\} . \tag{1.2.1}$$

By linearity of L an equivalent formulation is

$$\|L\| = \sup_{\|x\| \leq 1} \|L(x)\| .$$

DEFINITION 1.2.4. The space of bounded linear operators between normed vector spaces X and Y is defined as

$$L(X, Y) := \{L : X \to Y \text{ is linear and } \|L\| < \infty\}$$

where $\|\cdot\|$ is the map from Definition 1.2.3. We set

$$L(X) := L(X, X).$$

PROPOSITION 1.2.5. *Let X, Y be normed vector spaces with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$.*

(i) The map $\|\cdot\|$ from Definition 1.2.3 defines a norm on the space $L(X, Y)$.

(ii) $L(X, Y)$ is a vector space with norm $\|\cdot\|$

(iii) $L(X, Y)$ is a Banach space with respect to the norm $\|\cdot\|$ if Y is complete.

PROOF. (i) For $g, f \in L(X, Y)$ and $x \in X$

$$\begin{aligned} \|g(x) + f(x)\|_Y &\leq \|g(x)\|_Y + \|f(x)\|_Y \\ &\leq \|g\| \|x\|_X + \|f\| \|x\|_X \end{aligned}$$

hence

$$\|g + f\| \leq \|g\| + \|f\| .$$

Clearly also for all $\lambda \in \mathbb{C}$

$$\|\lambda f\| = |\lambda| \|f\|$$

and $\|f\| = 0$ if and only if $f = 0$. Thus, $\|\cdot\|$ is a norm.

(ii) For $g, f \in L(X, Y)$ for all $v, u \in X$ and $\lambda \in \mathbb{K}$ it holds

$$g(\lambda u + \lambda v) + f(\lambda u + \lambda v) = \lambda (g(u) + f(u)) + \lambda (g(v) + f(v)) .$$

Hence, $g+f$ is linear and for $\mu \in \mathbb{K}$ also μf is linear. Clearly $\|\mu f\| < \infty$ for $\mu \in \mathbb{K}$ and the triangular inequality implies $\|g+f\| < \infty$. Therefore $L(X,Y)$ is a vector space.

(iii) We have to show that each Cauchy sequence in $L(X,Y)$ converges in $L(X,Y)$. So let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $L(X,Y)$. We have to show that it converges to some element of $L(X,Y)$ and first need to find such a candidate. For any $x \in X$

$$\|f_n(x) - f_m(x)\| \leq \|f_n - f_m\| \, \|x\|$$

converges to zero for $m, n \to \infty$ hence $(f_n(x))_{n\in\mathbb{N}}$ is a Cauchy sequence in Y. Since Y is complete

$$f(x) := \lim_{n\to\infty} (f_n(x))_{n\in\mathbb{N}}$$

exists for each $x \in X$. We now show that $f \in L(X,Y)$ and that $(f_n)_{n\in\mathbb{N}}$ converges to f in $L(X,Y)$. Linearity of f holds because for all $v, u \in X$ and $\lambda \in \mathbb{C}$

$$\begin{aligned} f(\lambda u + \lambda v) :&= \lim_{n\to\infty} (f_n(\lambda u + \lambda v))_{n\in\mathbb{N}} \\ &= \lambda \lim_{n\to\infty} (f_n(u))_{n\in\mathbb{N}} + \lambda \lim_{n\to\infty} (f_n(v))_{n\in\mathbb{N}} \\ &= \lambda f(u) + \lambda f(v). \end{aligned}$$

The convergence of $(f_n)_{n\in\mathbb{N}}$ to f in $L(X,Y)$ we see in the following way. For some $\varepsilon > 0$ we choose $n_0(\varepsilon) > 0$ such that for all $m, n > n_0(\varepsilon)$ it holds

$$\|f_n - f_m\| < \varepsilon.$$

Furthermore, for each $x \in X$ we choose $n_1(x,\varepsilon) > n_0(\varepsilon)$ such that for all $n \geq n_1(x,\varepsilon)$ it holds

$$\|f(x) - f_n(x)\|_Y < \varepsilon \|x\|_X .$$

With this choice we obtain for $n > n_0(\varepsilon)$ and any $x \in X$

$$\begin{aligned} \|f(x) - f_n(x)\|_Y &\leq \left\|f(x) - f_{n_1(x,\varepsilon)}(x)\right\|_Y + \left\|f_{n_1(x,\varepsilon)}(x) - f_n(x)\right\|_Y \\ &\leq \varepsilon \|x\|_X + \varepsilon \|x\|_X . \end{aligned}$$

Hence $\|f - f_n\| \leq 2\varepsilon$ and f_n converges to f in $L(X,Y)$. Furthermore, by

$$\|f\| \leq \|f - f_n\| + \|f_n\|$$

$\|f\|$ is bounded for $n > n_0(\varepsilon)$, thus $f \in L(X,Y)$. □

REMARK 1.2.6. This proposition justifies to call $\|\cdot\|$ from Definition 1.2.3 the *operator norm*.

PROPOSITION 1.2.7. *Let X and Y be normed vector spaces and $L : X \to Y$ be a linear map. Then the following assertions are equivalent:*
(i) L is continuous.
(ii) L is continuous at $x = 0$.
(iii) There exists $M_L \geq 0$ such that $\|Lx\| \leq M_L \|x\|$ for all $x \in X$.
(iv) L is uniformly continuous.

PROOF.
(i)⇒(ii) Clear.
(iii)⇒(iv) Clear.
(iv) ⇒(i) Clear.
(ii)⇒(iii) By contradiction assume that (iii) does not hold. Then for any $n \in \mathbb{N}$ there exists $x_n \in X$ such that $\|Lx_n\| > n \|x_n\|$ holds. This implies that for any $n \in \mathbb{N}$

$$\left\| L \left(\frac{x_n}{n \|x_n\|} \right) \right\| > 1.$$

Since the sequence $\left(\frac{x_n}{n\|x_n\|} \right)_{n \in \mathbb{N}}$ converges to zero this contradicts (ii). □

REMARK 1.2.8. In particular, on a normed vector space X the norm $\|\cdot\| : X \to \mathbb{R}_+$ is a continuous linear map.

LEMMA 1.2.9. *Let X, Y, Z be vector spaces with norm $\|\cdot\|$.*
(i) For $L_1 \in L(X,Y)$ and $L_2 \in L(Y,Z)$ also $L_1 \circ L_2 \in L(X,Z)$ and it holds

$$\|L_1 \circ L_2\| \leq \|L_1\| \|L_2\| .$$

(ii) The map

$$\begin{aligned} L(X,Y) \times L(Y,Z) &\to L(X,Z) \\ (L_1, L_2) &\to L_1 \circ L_2 \end{aligned}$$

is continuous if $L(X,Y) \times L(Y,Z)$ is equipped with the product topology (see Definition A.1.7).

PROOF. (i) Linearity of $L_1 \circ L_2$ follows directly. In order to estimate the norm we see for any $M_{L_1} > \|L_1\|$ and any $M_{L_2} > \|L_2\|$ and any $x \in X$

$$\begin{aligned} \|(L_1 \circ L_2)(x)\| &\leq M_{L_1} \|L_2(x)\| \\ &\leq M_{L_1} \cdot M_{L_2} \|x\| . \end{aligned}$$

Taking the infimum on the right hand side we obtain

$$\begin{aligned}\|(L_1 \circ L_2)(x)\| &\leq \inf_{M_{L_1} > \|L_1\|} \inf_{M_{L_2} > \|L_2\|} M_{L_1} \cdot M_{L_2} \|x\| \\ &\leq \|L_1\| \cdot \|L_2\| \|x\|\end{aligned}$$

and conclude.

(ii) We equip the space $L(X,Y) \times L(Y,Z)$ with the norm

$$\begin{aligned}L(X,Y) \times L(Y,Z) &\to \mathbb{R}_+ \\ (L_1, L_2) &\to \|L_1\| + \|L_2\|.\end{aligned}$$

(It can easily be shown that this is indeed a norm.) We note that the open sets in $L(X,Y) \times L(Y,Z)$ with respect to this norm are the same as the open sets with respect to the product topology.

Let $\varepsilon > 0$ be arbitrary and let $L_1 \in L(X,Y)$ and $L_2 \in L(Y,Z)$ be arbitrary. Let $\delta_1, \delta_2 > 0$ and let $L_1' \in L(X,Y)$ such that $\left\|L_1' - L_1\right\| < \delta_1$ and $L_2' \in L(X,Y)$ such that $\left\|L_2' - L_2\right\| < \delta_2$. With

$$\delta < -\frac{\|L_2\| + \|L_1\|}{2} + \sqrt{\left(\frac{\|L_2\| + \|L_1\|}{2}\right)^2 + \varepsilon}$$

one obtains $\delta < \frac{\varepsilon}{\|L_2\| + \delta + \|L_1\|}$ and for $\delta := \delta_1 + \delta_2$

$$\begin{aligned}\left\|L_1 \circ L_2 - L_1' \circ L_2'\right\| &\leq \left\|L_1 \circ L_2 - L_1' \circ L_2\right\| + \left\|L_1' \circ L_2 - L_1' \circ L_2'\right\| \\ &\leq \left\|\left(L_1 - L_1'\right) \circ L_2\right\| + \left\|L_1' \circ \left(L_2 - L_2'\right)\right\| \\ &\leq \delta \|L_2\| + \delta \left\|L_1'\right\| \\ &\leq \delta \|L_2\| + \delta \cdot (\delta + \|L_1\|) \\ &\leq \varepsilon.\end{aligned}$$

Hence, the map $(L_1, L_2) \to L_1 \circ L_2$ is continuous with respect to the topology induced by the chosen norm on $L(X,Y) \times L(Y,Z)$ thus also with respect to the product topology. □

1.3. Uniformly continuous semigroups

As mentioned before, we would like to consider one-parameter families of mappings

$$(T(t))_{t \in \mathbb{R}_+}$$

that have a Banach space as state space. This will lead to *uniformly continuous semigroups* (Definition 1.3.4). The following presentation is taken from Engel Nagel[**18**], chapter I, section 3. X will always denote a Banach space over the field $\mathbb{C}$.

Definition 1.3.1. If for a Banach space X

$$(T(t))_{t\in\mathbb{R}_+}$$

is a one-parameter family of linear bounded operators on X that satisfies the Functional Equation (1.1.1) it is called (*one-parameter*) *semigroup on* X.

Remark 1.3.2. For a one-parameter semigroup the family of operators

$$(T(t))_{t\in\mathbb{R}_+}$$

is an (algebraic) semigroup on $(L(X), \circ)$. If this semigroup is equipped with the operator norm from Definition 1.2.3, then by Lemma 1.2.9 the composition is a continuous operation (where the product space $L(X) \times L(X)$ is equipped with the product topology). An algebraic semigroup with continuous semigroup operation is called *topological semigroup*.

Remark 1.3.3. One can also define a one-parameter *group* on the Banach space X to be a family

$$(T(t))_{t\in\mathbb{R}}$$

of linear bounded operators on X that satisfies the Functional Equation (1.1.1). In this case, by Theorem A.4.19 both group operations of $(L(X), \circ)$, composition and inversion, are continuous. Such a group where both group operations are continuous is called *topological group*.

Since we have an algebraic semigroup homomorphism between the topological semigroup $(\mathbb{R}_+, +)$ where addition is a continuous operation (with respect to the usual topology) and the topological semigroup $(L(X), \circ)$ where the composition is continuous with respect to the operator norm it is natural to ask whether the semigroup homomorphism is continuous. Such a semigroup homomorphism is called *topological semigroup homomorphism.*

Definition 1.3.4. A one-parameter semigroup

$$(T(t))_{t\in\mathbb{R}_+}$$

on a Banach space X is *uniformly continuous* if

$$\begin{aligned}\mathbb{R}_+ &\to L(X)\\ t &\to T(t)\end{aligned}$$

is continuous with respect to the operator norm on $L(X)$ as defined in Equation 1.2.1.

Written out in a more detailed fashion, uniform continuity means that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $|s-t| < \delta$ implies

$$\sup_{\|x\|\leq 1} \|T(t)x - T(s)x\| < \varepsilon.$$

In order to discuss differentiability of a one-parameter semigroup on a Banach space X, we introduce the notion of a derivative on a normed vector space which is called *Fréchet derivative*.

A map $f : U \to Y$ between the open subset U of a normed vector space $X \supset U$ and the normed vector space Y is called *Fréchet differentiable at* $x \in U$ with derivative $A(x)$ if there exists a bounded linear map $A(x) : X \to Y$ such

$$\lim_{\|h\|\to 0} \frac{\|f(x+h) - f(x) - A(x)h\|}{\|h\|} = 0.$$

It is called simply *Fréchet differentiable* with *Fréchet derivative*

$$\begin{aligned} A: \ U &\to L(X,Y) \\ x &\to A(x) \end{aligned}$$

on U if it is Fréchet differentiable at any $x \in U$ with derivative $A(x)$. If f is linear then also $x \to A(x)$ is linear. If $X = \mathbb{R}$ then derivatives are interpreted as derivatives with respect to time and will frequently be noted with a dot: $\frac{d}{dt}f(s) = \dot{f}(s)$ for $s \in \mathbb{R}$.
Let X be a Banach space and $[u,v] \subset \mathbb{R}$ be a closed interval. Let $f : [u,v] \to X$ be continuous. Then by compactness of $[u,v] f$ is in particular uniformly continuous and $(I_n(f))_{n\in\mathbb{N}}$ defined as

$$I_n(f) := 2^{-n} \sum_{i=0}^{\lfloor (v-u)\cdot 2^n \rfloor} f(u + i \cdot 2^{-n})$$

is a Cauchy sequence hence converges. Thus we can define

Definition 1.3.5. Let X be a Banach space and let $f : [u,v] \subset \mathbb{R} \to X$ be continuous. We let $\int_u^v f(s)ds$ denote the *Riemann integral*.

Familiar properties of the integral can also be shown in the same way (cf. **[38]**, Chapter 3, Exercise 23). Therefore for $u,v,w \in U$

$$\int_u^w f(s)ds = \int_u^v f(s)ds + \int_v^w f(s)ds$$

and for $f: I \to X$ and $g: I \to X$ and $\mu, \nu \in \mathbb{C}$

$$\int_u^v \mu f(s) + \nu g(s) ds = \mu \int_u^v f(s) ds + \nu \int_u^v g(s) ds.$$

Furthermore, the fundamental theorem of calculus holds also in this case. So for a continuous function $f: I \to X$ and u and h such that $u, u+h \in I$

$$\lim_{h \to 0} \frac{1}{h} \int_u^{u+h} f(s) ds = f(u) \tag{1.3.1}$$

and for a linear bounded operator $L \in L(X)$ it holds

$$L\left(\int_u^v f(s) ds\right) = \int_u^v L\left(f(s)\right) ds. \tag{1.3.2}$$

Since the integral is defined just as the limit of Riemann sums and the norm is continuous the triangle inequality yields

$$\left\| \int_u^w f(s) ds \right\| \le \int_u^w \|f(s)\| \, ds.$$

We need another tool from functional analysis:

Lemma 1.3.6. *Let U be a closed subspace of a normed vector space X and let $U \neq X$. Then for any $0 < \delta < 1$ there exists $x_\delta \in X$ with $\|x_\delta\| = 1$ such that*

$$\|x_\delta - u\| \ge 1 - \delta$$

for all $u \in U$.

Proof. Choose some $y \in X \setminus U$. Then for all $u \in U$,$\|y - u\| > 0$ and since U is closed

$$d := \inf_{u \in U} \|y - u\| > 0$$

because if not there would be a sequence in U converging to y in contradiction to $y \in X \setminus U$. For $\frac{d}{1-\delta}$ there exists $u_\delta \in U$ such that

$$\|y - u_\delta\| < \frac{d}{1-\delta}.$$

Since U is a subspace $\frac{y-u_\delta}{\|y-u_\delta\|} \in X \setminus U$. Also $\left\|\frac{y-u_\delta}{\|y-u_\delta\|}\right\| = 1$ and for any $u \in U$

$$\begin{aligned} \left\| \frac{y - u_\delta}{\|y - u_\delta\|} - u \right\| &= \frac{1}{\|y - u_\delta\|} \|y - (u_\delta + \|y - u_\delta\| \, u)\| \\ &\ge \frac{1}{\|y - u_\delta\|} d \\ &> 1 - \delta. \end{aligned}$$

Hence the assertion holds for $x_\delta := \frac{y-u_\delta}{\|y-u_\delta\|}$. □

We are now able to show that there exists a connection between a uniformly continuous semigroup and a differential equation.

PROPOSITION 1.3.7. *For a uniformly continuous semigroup*

$$(T(t))_{t\in\mathbb{R}_+}$$

on a Banach space X the map

$$t \to T(t)$$

is Fréchet differentiable and satisfies the differential equation

$$\begin{aligned} \frac{d}{dt}T(t) = AT(t) &= T(t)A \text{ for } t \in \mathbb{R}_+ \\ T(0) &= Id \end{aligned} \tag{1.3.3}$$

for some bounded operator $A \in L(X)$ which is given by

$$A = \frac{d}{dt}T(t)\bigg|_{t=0}.$$

PROOF. The proof has four steps.
First, we define a differentiable map

$$t \to V(t) := \int_0^t T(s)\,ds,$$

which is well defined since $s \to T(s)$ is uniformly continuous hence the Riemann integral exists. The differentiability follows also from the assumption of uniform continuity which ensures by the fundamental theorem of calculus that $\dot{V}(t) = T(t)$.
Second, we show that $V(t_0)$ has a continuous inverse for some small $t_0 > 0$. This means we first have to show that $\int_0^{t_0} T(s)\,ds$ is injective and surjective.
We start with injectivity. By uniform continuity of T for a given $0 < \varepsilon < 1$ there exists $t_0 > 0$ such that $\|T(s) - \mathrm{Id}\| < \varepsilon$ for all $0 < s < t_0$. Hence

$$\begin{aligned} \left\| \int_0^{t_0} \mathrm{Id}\ ds - \int_0^{t_0} T(s)\,ds \right\| &\le \int_0^{t_0} \|\mathrm{Id}\ - T(s)\|\,ds \\ &< t_0\varepsilon. \end{aligned}$$

and for any $x \in X$

(1.3.4)

$$\begin{aligned}\left\|\int_0^{t_0} T(s)x\,ds\right\| &\geq \left\|\int_0^{t_0} \mathrm{Id}\ x\,ds\right\| - \left\|\int_0^{t_0} \mathrm{Id}\ x\,ds - \int_0^{t_0} T(s)x\,ds\right\| \\ &> (t_0 - t_0\varepsilon)\,\|x\|\,.\end{aligned}$$

By linearity of $\int_0^{t_0} T(s)\,ds$ and the definition of the norm, this inequality implies that $\int_0^{t_0} T(s)\,ds$ is injective.

In order to show that $\int_0^{t_0} T(s)\,ds$ is surjective we first show that the image of $\int_0^{t_0} T(s)\,ds$ is closed. Let $(y_n)_{n\in\mathbb{N}}$ be a converging sequence that lies in the image of $\int_0^{t_0} T(s)\,ds$ and that converges to some y. $(y_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X . We choose x_n for any $n \in \mathbb{N}$ such that $y_n = \int_0^{t_0} T(s)x_n ds$. Then the inequality

$$\left\|\int_0^{t_0} T(s)u\,ds - \int_0^{t_0} T(s)v\,ds\right\| > \left\|\int_0^{t_0} T(s)\,ds\right\|\,\|u - v\|$$

for some $u, y \in X$ implies that also $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in X. Hence it converges to some $x \in X$ and by continuity of $\int_0^{t_0} T(s)\,ds$

$$\int_0^{t_0} T(s)x\,ds = \lim_{n\to\infty} \int_0^{t_0} T(s)x_n\,ds = y.$$

Thus the image of $\int_0^{t_0} T(s)\,ds$ is closed. We can use Riesz' Lemma 1.3.6 to show that if the image of $\int_0^{t_0} T(s)\,ds$ was not all of X then there would be some $z \in X$ with $\|z\| = 1$ such that

$$\left\|z - \int_0^{t_0} T(s)\frac{z}{t_0}\,ds\right\| \geq \varepsilon.$$

But this contradicts

$$\left\|\int_0^{t_0} \mathrm{Id}\ \ ds - \int_0^{t_0} T(s)\,ds\right\| < t_0\varepsilon.$$

Thus, $\int_0^{t_0} T(s)\,ds$ must be surjective and $V(t_0)$ is invertible.

Its inverse $(V(t_0))^{-1} : X \to X$ is linear because for $x_1, x_2, y_1, y_2 \in X$

such that $y_1 = \int_0^{t_0} T(s)x_1\,ds$ and $y_2 = \int_0^{t_0} T(s)x_2\,ds$ and some $\lambda \in \mathbb{C}$

$$\begin{aligned}
(V(t_0))^{-1}(\lambda y_1 + \lambda y_2) &== (V(t_0))^{-1}\left(\int_0^{t_0} T(s)(\lambda x_1 + \lambda x_2)\,ds\right) \\
&= (V(t_0))^{-1} V(t_0)(\lambda x_1 + \lambda x_2) \\
&= \lambda x_1 + \lambda x_2 \\
&= \lambda (V(t_0))^{-1} y_1 + \lambda (V(t_0))^{-1} y_2.
\end{aligned}$$

In order to show boundedness (or equivalently continuity) of $(V(t_0))^{-1}$ we use Inequality 1.3.4. So for any $y, x \in X$ with $y = \int_0^{t_0} T(s)x\,ds$

$$\begin{aligned}
\left\|(V(t_0))^{-1} y\right\| &= \|x\| \\
&< \frac{\|y\|}{t_0 - t_0\varepsilon}.
\end{aligned}$$

Third, we express T in terms of V. Because of the second step we can write $T(t)$ as

$$T(t) = (V(t_0))^{-1} V(t_0)T(t).$$

In order to reach our objective we can absorb the $T(t)$ of the right hand side in the V term in the following way:

$$\begin{aligned}
(V(t_0))^{-1} V(t_0)T(t) &= (V(t_0))^{-1}\int_0^{t_0} T(s)T(t)\,ds \\
&= (V(t_0))^{-1}\int_0^{t_0} T(s+t)\,ds \\
&= (V(t_0))^{-1}\int_t^{t+t_0} T(s)\,ds \\
&= (V(t_0))^{-1}\left(\int_0^{t+t_0} T(s)\,ds - \int_0^{t} T(s)\,ds\right) \\
&= (V(t_0))^{-1}\left(V(t+t_0) - V(t)\right).
\end{aligned}$$

Fourth, with the representation

$$T(t) = (V(t_0))^{-1}\left(V(t+t_0) - V(t)\right)$$

Fréchet differentiability of $T(t)$ follows from Fréchet differentiability of $V(t)$ and the continuity of $(V(t_0))^{-1}$. The Fréchet derivative can be calculated as

$$\begin{aligned}\frac{d}{dt}(T(t)) &= \lim_{h\searrow 0}\frac{T(t+h)-T(t)}{h}\\ &= \lim_{h\searrow 0}\frac{T(h)-\mathrm{Id}}{h}T(t)\\ &= \dot{T}(0)T(t)\end{aligned}$$

Setting $A = \dot{T}(0)$ we obtain the statement of the proposition. □

REMARK 1.3.8. $A = \dot{T}(0)$ is called generator of the semigroup

$$(T(t))_{t\in\mathbb{R}_+}.$$

In order to solve Equation 1.3.3 we define the object

$$e^{tA} := \sum_{k=0}^{\infty}\frac{(tA)^k}{k!}.$$

Here $A \in L(X)$ is a bounded linear operator on the Banach space X. The series converges because for $N, M \in \mathbb{N}$ and $M \geq N$ it holds

$$\left\|\sum_{k=0}^{M}\frac{(tA)^k}{k!} - \sum_{k=0}^{N}\frac{(tA)^k}{k!}\right\| \leq \sum_{k=N}^{M}\frac{t^k\|A\|^k}{k!}$$

which converges to zero for $M, N \to \infty$. Hence

$$\left(\sum_{k=0}^{N}\frac{(tA)^k}{k!}\right)_{N\in\mathbb{N}}$$

is a Cauchy sequence and its limit e^{tA} exists and lies in the Banach space $L(X)$.
As in the case of the exponential of complex numbers $x, y \in \mathbb{C}$ where $e^{x+y} = e^x e^y$ a similar property holds for the exponential of operators.

LEMMA 1.3.9. *For a Banach space X and linear bounded operators $A, B \in L(X)$ that commute it holds*

$$e^{A+B} = e^A e^B$$

PROOF. By definition

$$e^{A+B}x = \sum_{k=0}^{\infty}\frac{(A+B)^k x}{k!}.$$

Since A and B commute we obtain

$$\sum_{k=0}^{\infty} \frac{(A+B)^k x}{k!} = \sum_{k=0}^{\infty}\sum_{l=0}^{k} \frac{\frac{k!}{(k-l)!l!} (A)^l (B)^{k-l} x}{k!}$$
$$= \sum_{k=0}^{\infty}\sum_{l=0}^{k} \frac{(B)^{k-l}}{(k-l)!} \frac{(A)^l x}{l!}$$

In the sum we sum up all possible combinations of powers of Ax and Bx in a fashion that is described in the following table. It indicates for which values of the indices k and l we obtain the expression $(A)^n (B)^m x$ under the sum.

	$(A)^0$	$(A)^1$	$(A)^2$	$(A)^3$
$(B)^0$	$k=0, l=0$	$k=1, l=1$	$k=2, l=2$	$k=3, l=3$
$(B)^1$	$k=1, l=0$	$k=2, l=1$	$k=3, l=2$	...
$(B)^2$	$k=2, l=0$	$k=3, l=1$	...	...
$(B)^3$	$k=3, l=0$	...	...	...

We see that for a given k we sum up a diagonal in the table. By changing the summation procedure to horizontal and vertical summation the limit remains unchanged because it exists and is unique and we are therefore able to write the sum as

$$e^{A+B}x = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{(B)^m}{m!} \frac{(A)^n x}{n!}$$
$$= e^A e^B x.$$

□

Proposition 1.3.10. *For some $A \in L(X)$ the map*

$$\mathbb{R}_+ \to L(X)$$
$$t \to e^{tA}$$

is uniformly continuous and satisfies

$$e^{(t+s)A} = e^{tA}e^{sA} \quad \text{for } t, s \in \mathbb{R}_+$$
$$e^{0\cdot A} = Id$$

thus

$$\left\{e^{tA}\right\}_{t\in\mathbb{R}_+}$$

is a uniformly continuous semigroup.

Proof. The equation

$$e^{(t+s)A} = e^{tA}e^{sA}$$

follows from Lemma 1.3.9. $e^{0\cdot A} = \mathrm{Id}$ follows from the equation above for $t = 0$ and $s = 0$ respectively.
In order to prove uniform continuity of $t \to e^{tA}$ we need to show that

$$\lim_{h\to 0} \left\| e^{(t+h)A} - e^{tA} \right\| = 0$$

for all $t \in \mathbb{R}_+$. We see that

$$\begin{aligned}
\lim_{h\to 0} \left\| e^{(t+h)A} - e^{tA} \right\| &\leq \lim_{h\to 0} \left\| e^{hA} - \mathrm{Id} \right\| \left\| e^{tA} \right\| \\
&= \lim_{h\to 0} \left\| \sum_{k=1}^{\infty} \frac{(hA)^k}{k!} \right\| \left\| e^{tA} \right\| \\
&\leq \lim_{h\to 0} \left(\sum_{k=1}^{\infty} \frac{h^k \left\| A \right\|^k}{k!} \right) \left\| e^{tA} \right\| \\
&= \lim_{h\to 0} \left(e^{h\|A\|} - 1 \right) \left\| e^{tA} \right\| \\
&= 0.
\end{aligned}$$

□

Proposition 1.3.11.
(i) For some $A \in L(X)$ the map $T(\cdot)$:

$$\begin{aligned} \mathbb{R}_+ &\to L(X) \\ t &\to e^{tA} \end{aligned} \tag{1.3.5}$$

is Fréchet differentiable and satisfies the differential equation

$$\begin{aligned} \frac{d}{dt} T(t) &= AT(t) \quad \textit{for } t \in \mathbb{R}_+ \\ T(0) &= Id. \end{aligned} \tag{1.3.6}$$

(ii) A Fréchet differentiable map $T(\cdot) : \mathbb{R}_+ \to L(X)$ that satisfies Equation 1.3.6 for some $A \in L(X)$ is of the form 1.3.5.

Proof.
First we show (i). Due to Proposition 1.3.10

$$T(\cdot) : t \to e^{tA}$$

is a uniformly continuous semigroup. Thus Proposition 1.3.7 yields that $t \to e^{tA}$ is Fréchet differentiable and satisfies

$$\begin{aligned} \frac{d}{dt} T(t) &= BT(t) \quad \text{for } t \in \mathbb{R}_+ \\ T(0) &= \mathrm{Id}. \end{aligned}$$

for some bounded operator $B \in L(X)$ which is given by $B = \frac{d}{dt}T(t)\big|_{t=0}$. We have to show that $\frac{d}{dt}T(t)\big|_0 = A$ hence $B = A$. We calculate

$$\begin{aligned}
&= \lim_{h\to 0}\frac{1}{h}\left\|T(h) - T(0) - hA\right\| \\
&= \lim_{h\to 0}\frac{1}{h}\left\|\sum_{k=0}^{\infty}\frac{(hA)^k}{k!} - \mathrm{Id} - hA\right\| \\
&= \lim_{h\to 0}\left\|\frac{1}{h}\sum_{k=2}^{\infty}\frac{(hA)^k}{k!}\right\| \\
&\le \lim_{h\to 0}\left\|hA^2\right\|\sum_{k=2}^{\infty}\frac{(h\left\|A\right\|)^{k-2}}{(k-2)!} \\
&= 0
\end{aligned}$$

which implies (i).
Concerning (ii) we assume that there is a second Fréchet differentiable map $S(\cdot)$ different from

$$T(\cdot) : t \to e^{tA}$$

that satisfies Equation 1.3.6. For an arbitrary but fixed $t \in \mathbb{R}_+$ we define

$$Q(s) = S(s)T(t-s)$$

for $0 \le s \le t$ and observe that the product rule which holds also for Fréchet derivatives yields $\frac{d}{ds}Q(s) = 0$ for all $s \in [0,t]$. Hence $Q(0) = Q(t)$. This implies $S(t) = T(t)$ equals for all $t \in \mathbb{R}_+$ since t was arbitrary. Thus $S(\cdot) = T(\cdot)$. □

To sum up, we have shown that a uniformly continuous semigroup must satisfy Equation 1.3.6 and that $t \to e^{tA}$ satisfies Equation 1.3.6 and is the only map that does so. Hence, all uniform continuous semigroups are of the form $t \to e^{tA}$ where A is called the generator of the uniformly continuous semigroup.

REMARK 1.3.12. For a Banach space X and some initial state $x_0 \in X$ and a linear bounded operator $A \in L(X)$ we can define $x(t) := e^{tA}x_0$

and see in the calculation

$$\begin{aligned}
&\lim_{h\to 0}\frac{1}{h}\left(\left\|e^{(t+h)A}x_0 - e^{tA}x_0 - Ahe^{tA}x_0\right\|\right)\\
&\leq \lim_{h\to 0}\frac{1}{h}\left(\left\|\left(e^{hA} - \mathrm{Id} - Ah\right)\right\|\left\|e^{tA}x_0\right\|\right)\\
&\leq \lim_{h\to 0}\frac{1}{h}\left(\left(e^{h\|A\|} - 1 - \|A\|\,h\right)\left\|e^{tA}x_0\right\|\right)\\
&= 0
\end{aligned}$$

that the Fréchet derivative of $t \to x(t)$ is $\dot{x}(t) = Ax(t)$. Thus $t \to e^{tA}x_0$ is a solution of the initial value problem

$$\begin{aligned}
\dot{x}(t) &= Ax(t) \quad \text{for t}\in\mathbb{R}_+\\
x(0) &= x_0.
\end{aligned}$$

We can show as before that such a solution is also unique by assuming that there is a second one $y(t)$ and defining $q(s) := x(s)y(t-s)$ for $0 \leq s \leq t$ for a fixed but arbitrary $t \in \mathbb{R}_+$. Then $\frac{d}{ds}q(s) = 0$ and $q(0) = q(t)$. Thus $y(t) = x(t)$ for all $t \in \mathbb{R}_+$.

We know that a uniformly continuous semigroup $t \to T(t)$ on a Banach space X is of the form $t \to e^{tA}$ for a bounded linear operator $A \in L(X)$. In general however, it can be difficult to calculate e^{tA} given a bounded linear operator $A \in L(X)$. The following case is an example where this calculation can be carried out.

EXAMPLE 1.3.13. For $\infty > p \geq 1$ the space

$$\ell^p := \left\{ f:\ \mathbb{N} \to \mathbb{C}:\ \sum_{k=0}^{\infty} |f(k)|^p < \infty \right\}$$

with norm $\|\cdot\|_{\ell^p}:\ f \to \left(\sum_{k=0}^{\infty} |f(k)|^p\right)^{1/p}$ is a Banach space (see for example [**41**]). On ℓ^p one can think of a linear map as an infinite dimensional matrix. Given an infinite dimensional matrix $A = (a_{ij})_{i,j\in\mathbb{N}}$ as generator with

$$a_{ij} = \begin{cases} 1 & \text{if } j - i = 1 \\ 0 & \text{otherwise} \end{cases}$$

the exponential

$$e^{tA} = \sum_{k=0}^{\infty} \frac{(A)^k\, t^k}{k!}$$

is an infinite dimensional matrix. Executing the matrix multiplication A^k shows that

$$\left(A^k\right)_{i,j} = \begin{cases} 1 & \text{if } j - i = k \\ 0 & \text{otherwise} \end{cases}.$$

Thus, the entries of e^{tA} are given by

$$\left(e^{tA}\right)_{i,j} = \begin{cases} \frac{t^{j-i}}{(j-i)!} & \text{if } j-i \geq 0 \\ 0 & \text{otherwise} \end{cases} .$$

1.4. Strongly continuous semigroups

In the following, X will always denote a $\mathbb{C}$-Banach space. The previous subsection presented some general results on uniformly continuous semigroups. However, often the semigroup

$$(T(t))_{t\in\mathbb{R}_+}$$

on a Banach space X fails to be uniformly continuous but is still *strongly continuous*. However, also in this case a mathematically rich theory can be developed which will be done in the following. This section largely follows Engel, Nagel [**18**].

1.4.1. Definition and elementary properties. One example where a semigroup on a Banach space X fails to be uniformly continuous is presented below and taken from Engel, Nagel [**18**] chapter I, section 4.

EXAMPLE 1.4.1. For a function $f : \mathbb{R} \to \mathbb{C}$ we define the *left translation* $T_l(t)\,(f)$ of f by $t \in \mathbb{R}$ as

$$T_l(t)\,(f)\,(s) := f(s+t)$$

for all $s \in \mathbb{R}$.
Similarly, one can define the *right translation* $T_r(t)\,(f)$ of f by $t \in \mathbb{R}$ as

$$T_r(t)\,(f)\,(s) := f(s-t)$$

for all $s \in \mathbb{R}$. We consider the family of maps

$$(T_l(t))_{t\in\mathbb{R}_+}$$

on the Banach space $L^p(\mathbb{R})$ for some $1 \leq p < \infty$ from Example A.4.1. We see that $f \in L^p(\mathbb{R})$ implies $T_l(t)f \in L^p(\mathbb{R})$ and $\|f\|_{L^p(\mathbb{R})} = \|T_l(t)f\|_{L^p(\mathbb{R})}$ for any $t \in \mathbb{R}_+$ and that for each $t \in \mathbb{R}_+$, $T_l(t)$ is a linear map. Hence

$$(T_l(t))_{t\in\mathbb{R}_+}$$

is a family of bounded linear operators on $L^p(\mathbb{R})$. It is also a semigroup on $L^p(\mathbb{R})$ as the Functional Equation (1.1.1) holds since for any $u, v \in \mathbb{R}$ and all $s \in \mathbb{R}$

$$\begin{aligned} T_l(u+v)(f)(s) &= f(s+u+v) \\ &= T_l(u)\,(f)\,(s+v) \\ &= T_l(v)\,(T_l(u)\,(f))\,(s). \end{aligned}$$

We show that it is not uniformly continuous. By contradiction, if it was uniformly continuous then for every $2 > \varepsilon > 0$ there would exist a $\delta > 0$ such that $|t| < \delta$ would imply

$$\sup_{\|f\|_{L^p(\mathbb{R})} \leq 1} \|T_l(t)f - f\|_{L^p(\mathbb{R})} < \varepsilon.$$

We define the function

$$f_\delta(x) := \begin{cases} (1/\delta)^{1/p} \text{ for } 0 \leq x \leq \delta \\ 0 \text{ else.} \end{cases}$$

For this function

$$\|f_\delta\|_{L^p(\mathbb{R})} = \int_0^\delta \left|(1/\delta)^{1/p}\right|^p ds = 1.$$

We see

$$(T_l(\delta)f_\delta - f_\delta)\,(x) = \begin{cases} (1/\delta)^{1/p} \text{ for } -\delta \leq x \leq 0 \\ -(1/\delta)^{1/p} \text{ for } 0 \leq x \leq \delta \\ 0 \text{ else} \end{cases}$$

and compute

$$\begin{aligned} \|T_l(\delta)f_\delta - f_\delta\|_{L^p(\mathbb{R})} &= \int_{-\delta}^{\delta} \left|(1/\delta)^{1/p}\right|^p ds \\ &= 2 \end{aligned}$$

in contradiction to uniform continuity.

Thus, instead of focusing on uniform continuity we will look at the following, weaker, concept of continuity:

Definition 1.4.2. A one-parameter semigroup

$$(T(t))_{t \in \mathbb{R}_+}$$

on a Banach space X is called *strongly continuous* (or $\mathcal{C}_0$−semigroup) if for all $x \in X$ the *orbit map* ξ_x :

$$\begin{aligned} \mathbb{R}_+ &\to X \\ t &\to T(t)x \end{aligned}$$

is continuous.

The natural objective is to obtain similar statements as in the case of uniform continuity. As a tool from functional analysis we need to use the *uniform boundedness principle*. In order to show it we first prove *Baire's theorem*. The proofs for both theorems are based on the versions that can be found in Werner [**41**], chapter 4.

DEFINITION 1.4.3. A set M is called *convex* if for all $x, y \in M$ and all $0 \leq \lambda \leq 1$ also $\lambda x + (1-\lambda)\, y \in M$.

THEOREM 1.4.4. *(Baire's theorem)*
In a complete metric space (X, d) for a sequence of open, dense subsets $(O_n)_{n\in\mathbb{N}}$ also $\bigcap_{n\in\mathbb{N}} O_n$ is dense in X.

PROOF. We need to show that for an arbitrary open ball

$$U_\varepsilon(x_0) := \{x \in X : \, d(x, x_0) < \varepsilon\}$$

with radius $\varepsilon > 0$ and center $x_0 \in X$ the intersection

$$U_\varepsilon(x_0) \cap \left(\bigcap_{n\in\mathbb{N}} O_n \right)$$

is nonempty. In order to show this, we will construct a sequence of open balls $(U_{\varepsilon_k}(x_k))_{k\in\mathbb{N}}$ such that $\varepsilon_k < \frac{\varepsilon_{k-1}}{2}$ and

$$U_{\varepsilon_k}(x_k) \subset \left(\bigcap_{n=1}^{k-1} U_{\varepsilon_k}(x_k) \right)$$

and

$$U_{\varepsilon_k}(x_k) \subset U_\varepsilon(x_0) \cap \left(\bigcap_{n=1}^{k} O_n \right)$$

for all $k \in \mathbb{N}$. If we have found such a sequence of open balls $(U_{\varepsilon_k}(x_k))_{k\in\mathbb{N}}$ then its centers $(x_k)_{k\in\mathbb{N}}$ form a Cauchy sequence since $\varepsilon_k < \frac{\varepsilon_{k-1}}{2}$. Because of the completeness on X we see that $(x_k)_{k\in\mathbb{N}}$ converges to some limit called x which then lies in

$$U_\varepsilon(x_0) \cap \left(\bigcap_{n\in\mathbb{N}} O_n \right)$$

as desired.
In order to construct the sequence of open balls $(U_{\varepsilon_k}(x_k))_{k\in\mathbb{N}}$ with the mentioned properties we start out with $U_\varepsilon(x_0)$ and want to find $U_{\varepsilon_1}(x_1)$. Since O_1 is dense $U_\varepsilon(x_0)\cap O_1$ is nonempty. It is also open. Thus there is an open ball $U_{\varepsilon_1}(x_1)$ contained in $U_\varepsilon(x_0)\cap O_1$. We can easily choose $\varepsilon_1 < \frac{\varepsilon_0}{2}$. By the same reasoning as before, the intersection $U_{\varepsilon_1}(x_1)\cap O_2$ is nonempty and open so there is an open ball $U_{\varepsilon_2}(x_2)$ which lies in $U_{\varepsilon_1}(x_1)\cap O_2$ where again we can choose $\varepsilon_2 < \frac{\varepsilon_1}{2}$. Carrying out this procedure inductively we obtain $(U_{\varepsilon_k}(x_k))_{k\in\mathbb{N}}$ with the desired properties. □

We call a set $N\subset X$ *nowhere dense* if $\overset{\circ}{\overline{N}}=\emptyset$. We call a set $F\subset X$ of *first category* if there exists a sequence $(N_i)_{i\in\mathbb{N}}$ of nowhere dense sets $N_i\subset X$ such that $F=\bigcup_{i=1}^\infty N_i$. We call a set $S\subset X$ of *second category* if it is not of first category.

THEOREM 1.4.5. *(Baire's category theorem)*
Let (X,d) be a complete metric space and $F\subset X$ a set of first category. Then $\overline{X\setminus F}=X$.

PROOF. Let $(N_i)_{i\in\mathbb{N}}$ be a sequence of nowhere dense sets $N_i\subset X$ such that $F=\bigcup_{i=1}^\infty N_i$. Then

$$X\setminus F=\bigcap_{i=1}^{\infty}(X\setminus N_i)\supset\bigcap_{i=1}^{\infty}\left(X\setminus\overline{N_i}\right).$$

$X\setminus\overline{N_i}$ is open. By assumption $\overset{\circ}{\overline{N_i}}=\emptyset$ for any $i\in\mathbb{N}$ which means that for any open set $O\subset X$ the intersection

$$O\cap\left(X\setminus\overline{N_i}\right)=O\setminus\overline{N_i}$$

is nonempty for any $i\in\mathbb{N}$ or put differently that $\left(X\setminus\overline{N_i}\right)$ is dense in X for any $i\in\mathbb{N}$. Therefore we can apply Baire's theorem (Theorem 1.4.4) to $\bigcap_{i=1}^\infty\left(X\setminus\overline{N_i}\right)$ and obtain that it is dense in X. Thus also $X\setminus F$ is dense in X. □

THEOREM 1.4.6. *(Uniform boundedness principle)*
Let X be a Banach space, Y a normed vector space and I an index set. Let $(T_i)_{i\in I}$ be a family of bounded linear operators mapping from X to Y. If for all $x\in X$

$$\sup_{i\in I}\|T_ix\|<\infty$$

holds, then also

$$\sup_{i\in I}\|T_i\|<\infty.$$

PROOF. We will assume that X is nonempty, the statement being trivial otherwise. We want to show that there exists $N_0 \in \mathbb{N}$ such that we can find a small open ball $U_\delta(0)$ around 0 with radius $\delta > 0$ such that for all $x \in U_\delta(0)$ it holds

$$\sup_{i \in I} \|T_i x\| \leq N_0.$$

If we are able to find this we can conclude by computing

$$\begin{aligned}\sup_{i \in I} \|T_i\| &= \sup_{i \in I} \left(\sup_{\|x\| \leq 1} \|T_i x\| \right) \\ &\leq \sup_{i \in I} \left(\sup_{y \in U_\delta(0)} \left\| T_i \left(\frac{y}{\delta} \right) \right\| \right) \\ &\leq \frac{N_0}{\delta}.\end{aligned}$$

The proof that there exist $\delta > 0$ and $N_0 \in \mathbb{N}$ such that

$$\sup_{i \in I} \|T_i x\| < N_0$$

holds for all $x \in U_\delta(0)$ has three steps.
First, we find a candidate $N \in \mathbb{N}$ to be the N_0 mentioned above. For this purpose, we use Baire's category theorem. The assumption that

$$\sup_{i \in I} \|T_i x\| < \infty$$

holds for all $x \in X$ means that $X = \bigcup_{n \in \mathbb{N}} E_n$ where

$$E_n := \left\{ x \in X : \sup_{i \in I} \|T_i x\| \leq n \right\}.$$

By Baire's category theorem (Theorem 1.4.5) X cannot be of first category because if it was then $\overline{X \setminus X} = X$ which is impossible for a nonempty set. Hence X is of second category which means it cannot be written as the countable union of nowhere dense sets. Thus there is some $N \in \mathbb{N}$ such that E_N is not nowhere dense.
Second, we show that there is an open ball contained in E_N. We can see that for each $n \in \mathbb{N}$ E_n is closed when we write it as

$$E_n = \bigcap_{i \in I} \{x \in X : \|T_i x\| \leq n\}$$

and keep in mind that $\|T_i(\cdot)\|$ is continuous which implies that

$$\{x \in X : \|T_i x\| \leq n\} = \|T_i(\cdot)\|^{-1}([0, n])$$

is closed. Therefore the fact that E_N is not nowhere dense means that $\overline{\mathring{E_N}} = \mathring{E_N}$ is nonempty. This allows us to find $y \in E_N$ and $\varepsilon > 0$ such that the open ball $U_\varepsilon(y)$ is contained in E_N.
Third, we show that $U_\varepsilon(0)$ is contained in E_N. $U_\varepsilon(y) \subset E_N$ means that $\|z - y\| < \varepsilon$ for some $z \in X$ implies that

$$\sup_{i \in I} \|T_i z\| \leq N.$$

We see that also $U_\varepsilon(-y)$ must be contained in E_N because

$$\|z - (-y)\| < \varepsilon$$

for some $z \in X$ implies

$$\|(-z) - y\| < \varepsilon$$

which yields

$$\sup_{i \in I} \|T_i\,(-z)\| \leq N$$

hence

$$\sup_{i \in I} \|T_i z\| \leq N.$$

Therefore also $U_\varepsilon(0)$ is contained in E_N which can be seen by representing $x \in U_\varepsilon(0)$ as

$$x = \frac{1}{2}\left(\underbrace{(x - y)}_{\in E_N} + \underbrace{(x - (-y))}_{\in E_N}\right)$$

and by observing that

$$E_N = \left\{x \in X : \sup_{i \in I} \|T_i x\| \leq N\right\}$$

is convex (see Definition A.3.71).
By setting $N_0 := N$ and $\delta := \varepsilon$ we conclude. □

We want to establish statements that are equivalent to strong continuity. For this we will follow Engel, Nagel, chapter I, section 5[**18**]. We start with the following lemma.

LEMMA 1.4.7. *Let X be a Banach space and $K \subset \mathbb{R}$ be a compact set and*

$$F\ : K \to L(X)$$

be some function. Then the following three assertions are equivalent.
(i) For all $x \in X$ the map

$$\begin{aligned} K &\to X \\ t &\to F(t)x \end{aligned}$$

is continuous.
(ii) There is a dense subset D of X such that for all $x \in D$ the map

$$K \to X$$
$$t \to F(t)x$$

is continuous. Additionally,

$$\sup_{t \in K} \|F(t)\| < \infty.$$

(iii) For any compact subset $C \subset X$ the map

$$K \times C \to X$$
$$(t, x) \to F(t)x$$

is uniformly continuous.

PROOF.
(iii) $\Rightarrow$ (i) This is clear if we choose the compact set C to be $\{x\}$.
(i) $\Rightarrow$ (ii) All we need to show is

$$\sup_{t \in K} \|F(t)\| < \infty.$$

The continuity of $t \to \|F(t)x\|$ for each $x \in X$ implies that for each $x \in X$ on the compact set $K \subset \mathbb{R}$ the function $t \to \|F(t)x\|$ attains its maximum. Therefore

$$\sup_{t \in K} \|F(t)x\| < \infty$$

for each $x \in X$ and we conclude by applying the uniform boundedness principle from Theorem 1.4.6.
(ii) $\Rightarrow$ (iii) We have to show that for an arbitrary compact set $C \subset X$ and for an arbitrary $\varepsilon > 0$ there exists a $\delta > 0$ such that for all $s, t \in K$ and $x, y \in C$

$$|s - t| + \|x - y\| < \delta$$

implies

$$\|F(s)x - F(t)y\| < \varepsilon.$$

Using the inequality

$$\|F(s)x - F(t)y\| < \|F(s)x - F(t)x\| + \|F(t)x - F(t)y\|$$

we see that it is enough to find $\delta > 0$ such that for all $t \in K$ and all $x, y \in C$ such that $\|x - y\| < \delta$

$$\|F(t)x - F(t)y\| < \frac{\varepsilon}{2}$$

and for all $x \in C$ and $s, t \in K$ such that $|s-t| < \delta$

$$\|F(s)x - F(t)x\| < \frac{\varepsilon}{2}.$$

For the first inequality, we see that by assumption

$$M := \sup_{t \in K} \|F(t)\| < \infty$$

so

$$\|F(t)x - F(t)y\| < \frac{\varepsilon}{2}$$

holds if $\|x - y\| < \frac{\varepsilon}{2M}$.
For the second inequality, we need to use the continuity of $t \to F(t)y$ which however holds only for $y \in D$. Therefore we approximate any $x \in C$ by some $y \in D$ and see that the compactness of C yields that we can approximate all $x \in C$ with a given accuracy by just *finitely* many $(y_i)_{i=1,\dots,n} \in D$. Precisely, given $\gamma > 0$ the density of D in X implies that

$$C \subset \bigcup_{y \in D} U_\gamma(y)$$

where $U_\gamma(y)$ are open balls with radius $\gamma > 0$ and center $y \in D$. Because of the compactness of C already finitely many of these balls suffice to cover C. The centers of these finitely many balls we call $(y_i)_{i=1,\dots,n}$. Therefore, if we choose $\gamma = \frac{\varepsilon}{6M}$ then for any $x \in X$ we choose some $y_i \in (y_i)_{i=1,\dots,n}$ such that $\|x - y_i\| < \gamma$ which implies

$$\|F(s)x - F(s)y_i\| < \frac{\varepsilon}{6}$$

and

$$\|F(t)x - F(t)y_i\| < \frac{\varepsilon}{6}.$$

With this choice

$$\begin{aligned}\|F(s)x - F(t)x\| &< \underbrace{\|F(s)x - F(s)y_i\|}_{<\frac{\varepsilon}{6}} + \|F(s)y_i - F(t)y_i\| \\ &+ \underbrace{\|F(t)y_i - F(t)x\|}_{<\frac{\varepsilon}{6}} \\ &= \frac{2\varepsilon}{6} + \|F(s)y_i - F(t)y_i\|\end{aligned}$$

Using the continuity of $t \to F(t)y_i$ for $y_i \in D$ we choose $\tau > 0$ small enough such that $|s-t| < \tau$ implies

$$\|F(s)y_i - F(t)y_i\| < \frac{\varepsilon}{6}$$

for all (finitely many!) $(y_i)_{i=1,\dots,n}$ and conclude that

$$\|F(s)x - F(t)x\| < \frac{\varepsilon}{2}.$$

Hence for $\delta := \min\left\{\tau, \frac{\varepsilon}{2M}\right\}$ we see that for all $s, t \in K$ and $x, y \in C$

$$|s-t| + \|x-y\| < \delta$$

implies

$$\|F(s)x - F(t)y\| < \varepsilon$$

which is what we had to show. □

PROPOSITION 1.4.8. *For a semigroup*

$$(T(t))_{t\in\mathbb{R}_+}$$

on a Banach space X the following three assertions are equivalent.
(i)

$$(T(t))_{t\in\mathbb{R}_+}$$

is strongly continuous.
(ii) For all $x \in X$

$$\lim_{t\searrow 0} T(t)x = x.$$

(iii)
(a) There is a dense subset D of X such that for all $x \in D$

$$\lim_{t\searrow 0} T(t)x = x.$$

(b) Additionally, there is $\delta > 0$ and $M \geq 1$ such that $\|T(t)\| \leq M$ for all $0 \leq t \leq \delta$.

PROOF.
(i) ⇒(ii) is clear.
(ii) ⇒ (iii) (a) is clear.
(ii) ⇒ (iii) (b) By contradiction we assume that for all $\delta > 0$ and all $M \geq 1$ there is $0 \leq t \leq \delta$ such that $\|T(t)\| > M$. Hence we can choose a sequence $(t_n)_{n\in\mathbb{N}}$ converging to 0 such that

$$\lim_{n\to\infty} \|T(t_n)\| = \infty.$$

On the other hand by (ii)

$$\lim_{n\to\infty} \|T(t_n)x\| = \|x\|$$

for all $x \in X$ which implies that

$$\sup_{n\in\mathbb{N}} \|T(t_n)x\| < \infty$$

for all $x \in X$. Hence, by the uniform boundedness principle (Theorem 1.4.6)

$$\sup_{n\in\mathbb{N}} \|T(t_n)\| < \infty$$

in contradiction to

$$\lim_{n\to\infty} \|T(t_n)\| = \infty.$$

(iii) $\Rightarrow$ (ii) If we choose $\delta > 0$ provided by (iii) (b) we obtain that $T(t)$ is uniformly bounded on the compact interval $[0,\delta]$. Together with (iii)(a) $t \to T(t)$ satisfies the assertion of Lemma 1.4.7 (ii) which is equivalent to Lemma 1.4.7 (i) that yields that $t \to T(t)x$ is continuous on the compact interval $[0,\delta]$ for all $x \in X$. Thus (ii) follows.
(ii) $\Rightarrow$ (i) Fix an arbitrary $t_0 \in \mathbb{R}_+$. We have to show right continuity

$$\lim_{h\searrow 0} T(t_0+h)x \to T(t_0)x$$

and left continuity

$$\lim_{h\nearrow 0} T(t_0+h)x \to T(t_0)x.$$

Regarding right continuity the functional equality yields

$$\lim_{h\searrow 0} T(t_0+h)x = \lim_{h\searrow 0} T(h)T(t_0)x$$

and by (ii) we obtain

$$\lim_{h\searrow 0} T(h)T(t_0)x = T(t_0)x.$$

In order to show left continuity, we apply the functional equality and obtain

$$\lim_{h\nearrow 0} \|T(t_0+h)x - T(t_0)x\| = \lim_{h\nearrow 0} \|T(t_0+h)\left(x - T(-h)x\right)\|.$$

By (ii), we know

$$\lim_{h\nearrow 0} \|(x - T(-h)x)\| = 0.$$

We need an estimate for $T(t_0+h)$ and use the fact that (ii) implies (iii)(b). Therefore, we choose $t_0 > \delta > 0$ and $M \geq 1$ such that

$$\|T(t)\| \leq M$$

for $0 \leq t \leq \delta$. Thus,

$$\begin{aligned} \|T(t_0+h)\| &\leq \|T(t_0-\delta)\| \cdot \|T(\delta+h)\| \\ &\leq \|T(t_0-\delta)\| \cdot M \end{aligned}$$

for $-\delta \leq h \leq 0$. Hence, we obtain for the limit

$$\begin{aligned}\lim_{h\nearrow 0} \|T(t_0+h)x - T(t_0)x\| &\leq \lim_{h\nearrow 0} \|T(t_0+h)\| \, \|(x - T(-h)x)\| \\ &\leq \lim_{h\nearrow 0} \|T(t_0-\delta)\| \cdot M \cdot \|(x - T(-h)x)\| \\ &= 0\end{aligned}$$

and conclude. □

PROPOSITION 1.4.9. *For a strongly continuous semigroup*

$$(T(t))_{t\in\mathbb{R}_+}$$

on a Banach space X there exist constants $\omega \in \mathbb{R}$ and $M \geq 1$ such that for all $t \geq 0$

$$\|T(t)\| \leq Me^{\omega t}.$$

PROOF. The map $t \to T(t)x$ is continuous for all $x \in X$. Hence $t \to \|T(t)x\|$ attains its maximum on the compact interval $[0,1]$ for all $x \in X$ and

$$\sup_{t\in[0,1]} \|T(t)x\| < \infty$$

for all $x \in X$. By the uniform boundedness principle (Theorem 1.4.6)

$$M := \sup_{t\in[0,1]} \|T(t)\| < \infty.$$

If we now want to bound $\|T(t_0)\|$ for a fixed but arbitrary $t_0 \in \mathbb{R}_+$ we can write $t_0 = n + s$ for $n \in \mathbb{N}$ and $s \in [0,1]$ and obtain by the functional equation for T that

$$\|T(t_0)\| = \|(T(1))^n T(s)\| \leq \|(T(1))\|^n \cdot \|T(s)\|.$$

Thus

$$\|T(t_0)\| \leq M^{n+1} = M \cdot M^n = M \cdot e^{n\cdot \ln M} \leq M \cdot e^{t_0 \cdot \ln M}.$$

Setting $\omega := \ln M$ and observing that $t_0 \in \mathbb{R}_+$ was arbitrary we obtain the statement of the proposition. □

DEFINITION 1.4.10. In the following we will call a strongly continuous semigroup for which $\omega = 0$ and $M = 1$ are possible in Proposition 1.4.9 *contraction semigroup*.

EXAMPLE 1.4.11. We take a second look at the left translation semigroup

$$(T_l(t))_{t\in\mathbb{R}_+}$$

from Example 1.4.1. We have already shown in Example 1.4.1 that the semigroup is not uniformly continuous on $L^p(\mathbb{R})$. We now show that it is instead strongly continuous on $L^p(\mathbb{R})$. For this purpose we want to use Proposition 1.4.8. Since for any $f \in L^p(\mathbb{R})$

$$\begin{aligned}\|T_l(t)f\|_{L^p(\mathbb{R})} &= \left(\int_{-\infty}^{\infty} |f(s+t)|^p \, ds\right)^{1/p} \\ &= \left(\int_{-\infty}^{\infty} |f(s)|^p \, ds\right)^{1/p} \\ &= \|f\|_{L^p(\mathbb{R})}\end{aligned}$$

$\|T_l(t)\|_{L(L^p(\mathbb{R}))} = 1$ and statement (iii)(b) of Proposition 1.4.8 holds. It remains to show (iii)(a). Therefore we have to find a dense subset $D \subset L^p(\mathbb{R})$.

Since for all $f \in L^p(\mathbb{R})$ it holds $\int_{-\infty}^{\infty} |f(s)|^p \, ds < \infty$ for some $\varepsilon > 0$ there also exist $a, b \in \mathbb{R}$ with $a < b$ such that for the function

$$f_{[a,b]}(x) := \begin{cases} f(x) & \text{for } a \le x \le b \\ 0 & \text{elsewhere} \end{cases}$$

the estimate

$$\left\|f - f_{[a,b]}\right\|_{L^p(\mathbb{R})} < \varepsilon$$

holds. On the other hand we can approximate the function $f_{[a,b]}$ in the following way. We define the so-called *mollifiers*

$$\eta_\rho(x) := \begin{cases} \frac{1}{\rho} C \exp\left(\frac{1}{|x/\rho|^2 - 1}\right) & \text{if } |x| < \rho \\ 0 & \text{if } |x| \ge \rho \end{cases}$$

which are in $C_c^\infty(\mathbb{R})$ which means they are infinitely differentiable functions with compact support. It can be shown (see **[19]**) that the convolution $\left(f_{[a,b]}\right)^\rho$ defined by

$$\left(f_{[a,b]}\right)^\rho(x) := \int_{\mathbb{R}} \eta_\rho(x-y) f_{[a,b]}(y) dy$$

for each $x \in \mathbb{R}$ is also in $C_c^\infty(\mathbb{R})$. Also it is known (see **[19]**) that since $f_{[a,b]} \in L^p(\mathbb{R})$

$$\left(f_{[a,b]}\right)^\rho \to f_{[a,b]}$$

on $L^p([c,d])$ for any real $c < d$. Hence $\left(f_{[a,b]}\right)^\rho$ approximates $f_{[a,b]}$ also in $L^p(\mathbb{R})$ thus f as well and since $f \in L^p(\mathbb{R})$ was arbitrary $C_c^\infty(\mathbb{R})$ is dense in $L^p(\mathbb{R})$.

In order to show (iii)(a) of Proposition 1.4.8 it is therefore enough to show that

$$\lim_{t \searrow 0} \|T_l(t)f - f\|_{L_p(\mathbb{R})} = 0$$

for all $f \in C_c^\infty(\mathbb{R})$. Since a continuous function f with compact support is also uniformly continuous. Therefore for any $\varepsilon > 0$ there exists $t > 0$ such that $s < t$ implies $|f(r+s) - f(r)| < \varepsilon$ for all $r \in \mathbb{R}$. Since the support of f is bounded by some constant

$$C := \sup_{x \in \text{supp}\{f\}} \|x\|$$

we obtain

$$\|T_l(s)f - f\|_{L_p(\mathbb{R})} \leq ((C+t)\,\varepsilon^p)^{1/p}$$

for all $s < t$. Hence

$$\lim_{t \searrow 0} \|T_l(t)f - f\|_{L_p(\mathbb{R})} = 0$$

and we conclude by applying Proposition 1.4.8.

1.4.2. Generators of strongly continuous semigroups. We recall that for *uniformly continuous* semigroups $(T(t))_{t\in\mathbb{R}_+}$ on a Banach space X the map

$$\begin{aligned} \mathbb{R}_+ &\to L(X) \\ t &\to T(t) \end{aligned}$$

is Fréchet differentiable as shown in Proposition 1.3.7. We hope to show some kind of differentiability also in the case of *strongly continuous* semigroups. Since in this case we only have continuity of the orbit maps

$$\xi_x : t \to T(t)x \in X$$

we can hope to obtain differentiability at most for the orbit maps. Regarding the continuity of the orbit maps we have seen in Proposition 1.4.8 that right continuity of all orbit maps at $t = 0$ already implies continuity of all orbit maps at any $t \in \mathbb{R}_+$. The following lemma shows that a similar statements holds also for differentiability. It its taken from Nagel, Engel **[18]**, chapter II, section 1.

LEMMA 1.4.12. *For a strongly continuous semigroup $(T(t))_{t\in\mathbb{R}_+}$ on a Banach space X and $x \in X$ it holds: If the orbit map*

$$\xi_x : t \to T(t)x \in X$$

is right differentiable at $t = 0$ with $\dot{\xi}_x(0)$ then it is also differentiable on $\mathbb{R}_+$ and its derivative is given by

$$\begin{aligned} \dot{\xi}_x(t) &= T(t)\dot{\xi}_x(0) \\ &= \dot{\xi}_x(0)T(t) \end{aligned}$$

PROOF. We fix some arbitrary $t_0 \in \mathbb{R}_+$ and have to show right differentiability

$$\lim_{h \searrow 0} \frac{1}{h} \left\| T(t_0+h)x - T(t_0)x - \partial_+ \left(T(t_0)x\right) \right\| = 0$$

and left differentiability

$$\lim_{h \nearrow 0} \frac{1}{h} \left\| T(t_0+h)x - T(t_0)x - \partial_- \left(T(t_0)x\right) \right\| = 0$$

and have to show that left and right derivative coincide

$$\partial_+ \left(T(t_0)x\right) = \partial_- \left(T(t_0)x\right).$$

For right differentiability we use the semigroup property:

$$\begin{aligned} &\lim_{h \searrow 0} \frac{1}{h} \left(T(t_0+h)x - T(t_0)x\right) \\ &= T(t_0) \left(\lim_{h \searrow 0} \left(\frac{1}{h} T(h)x - x \right) \right) \\ &= T(t_0)\dot{\xi}_x(0) \end{aligned}$$

By the same token

$$\begin{aligned} &\lim_{h \searrow 0} \frac{1}{h} \left(T(t_0+h)x - T(t_0)x\right) \\ &= \dot{\xi}_x(0)T(t_0). \end{aligned}$$

Hence $\xi_x : t \to T(t)x \in X$ is right differentiable at $t_0 \in \mathbb{R}_+$. For left continuity we use the semigroup property to reduce the problem to the one of right continuity:

$$\begin{aligned} &\lim_{h \nearrow 0} \frac{1}{h} \left(T(t_0+h)x - T(t_0)x\right) \\ &= \lim_{h \nearrow 0} T(t_0+h) \left(\left(-\frac{1}{h} \right) \left(T(-h)x - x\right) \right) \end{aligned}$$

By the first part of the proof we know the limit of $\left(-\frac{1}{h}\right)\left(T(-h)x - x\right)$. By Proposition 1.4.9 we see that $\|T(t_0+h)\|$ is bounded for h on a compact interval. Thus we obtain

$$\lim_{h \nearrow 0} T(t_0+h) \left(\left(-\frac{1}{h} \right) \left(T(-h)x - x\right) - \dot{\xi}_x(0) \right) = 0,$$

and conclude

$$\begin{aligned} & \lim_{h \nearrow 0} \frac{1}{h} \left(T(t_0 + h)x - T(t_0)x \right) \\ = \; & T(t_0)\dot{\xi}_x(0) \\ = \; & \dot{\xi}_x(0)T(t_0). \end{aligned}$$

□

We do not know for which $x \in X$ the orbit $t \to \xi_x(t)$ is differentiable at $t = 0$ but we can define a map that maps to this derivative whenever possible:

DEFINITION 1.4.13. The *generator* A of a strongly continuous semigroup

$$(T(t))_{t \in \mathbb{R}_+}$$

on a Banach space X is a map

$$A : \; \mathcal{D}(A) \subset X \to X$$

defined by

$$Ax := \dot{\xi}_x(0) = \lim_{h \searrow 0} \frac{T(h)x - x}{h}$$

on its domain

$$\mathcal{D}(A) := \left\{ x \in X | \quad \begin{matrix} \text{there exists } \dot{\xi}_x(0) \in \text{X such that} \\ \lim\limits_{h \searrow 0} \left\| \frac{T(h)x - x}{h} - \dot{\xi}_x(0) \right\| = 0 \end{matrix} \right\}.$$

We also say that A *generates* $(T(t))_{t \in \mathbb{R}_+}$.

REMARK 1.4.14. The convergence

$$\lim_{h \to 0} \frac{T(h)x - x}{h}$$

is to be taken with respect to the norm of the Banach space which means that

$$\dot{\xi}_x(0) = \lim_{h \to 0} \frac{T(h)x - x}{h}$$

if and only if

$$\lim_{h \to 0} \left\| \dot{\xi}_x(0) - \frac{T(h)x - x}{h} \right\| = 0.$$

The domain $\mathcal{D}(A)$ is an important part of the operator A and the generator should also be written as pair $(A, \mathcal{D}(A))$ even though this is often omitted. It is also worth noting that the operator $A : \mathcal{D}(A) \to X$ does not have to be continuous. For $x, y \in \mathcal{D}(A)$ and $\mu \in \mathbb{C}$ we obtain

$$\begin{aligned}
&\lim_{h\to 0} \left\| \frac{T(h)\,(\mu x + \mu y) - (\mu x + \mu y)}{h} - \mu \dot{\xi}_x(0) - \mu \dot{\xi}_y(0) \right\| \\
&\leq \mu \lim_{h\to 0} \left\| \frac{T(h)\,(x) - (x)}{h} - \dot{\xi}_x(0) \right\| + \mu \lim_{h\to 0} \left\| \frac{T(h)(y) - (y)}{h} - \dot{\xi}_y(0) \right\| \\
&= 0,
\end{aligned}$$

thus $\mathcal{D}(A)$ is a vector space.

In the following result we see among other things that just like uniformly continuous semigroups also strongly continuous semigroups are linked to a differential equation. It is again taken from Nagel, Engel [**18**], chapter II, section 1.

Proposition 1.4.15. *Let*

$$(T(t))_{t\in\mathbb{R}_+}$$

be a strongly continuous semigroup on a Banach space X. The generator $(A, \mathcal{D}(A))$ has the properties:
(i) The operator $A : \ \mathcal{D}(A) \to X$ is linear.
(ii) For all $x \in \mathcal{D}(A)$ and all $t \in \mathbb{R}_+$ also $T(t)x \in \mathcal{D}(A)$ and

$$\begin{aligned}
\frac{d}{dt} T(t)x &= AT(t)x \\
&= T(t)Ax.
\end{aligned}$$

(iii) For all $x \in X$ and all $t \in \mathbb{R}_+$

$$\int_0^t T(s)x\,ds \in \mathcal{D}(A)$$

and

$$T(t)x - x = A \int_0^t T(s)x\,ds.$$

For all $x \in \mathcal{D}(A)$ and all $t \in \mathbb{R}_+$

$$T(t)x - x = \int_0^t T(s)Ax\,ds.$$

PROOF.
(i) This follows from the linearity of $T(t) \in L(X)$ for all $t \in \mathbb{R}_+$ and the definition of the operator

$$Ax := \lim_{h \to 0} \frac{T(h)x - x}{h}.$$

(ii) By $x \in \mathcal{D}(A)$ the map

$$\xi_x \ : t \to T(t)x \in X$$

is right differentiable at $t = 0$ and by Lemma 1.4.12 differentiable on $\mathbb{R}_+$. Thus, by the semigroup property $T(t)x \in A$ and by Lemma 1.4.12 the derivative is given by

$$\begin{aligned} \dot{\xi}_x(t) &= T(t)\dot{\xi}_x(0) \\ &= T(t)A \\ &= AT(t). \end{aligned}$$

(iii) For the first part of (iii) we have to show $\int_0^t T(s)xds \in \mathcal{D}(A)$ and to compute $A\int_0^t T(s)xds$. Therefore we need to find the limit

$$\lim_{h \searrow 0} \frac{1}{h}\left(T(h)\int_0^t T(s)x\,ds - \int_0^t T(s)x\,ds\right).$$

Since we can pull continuous linear operators in the integral (see Equation 1.3.2) we can use the Functional Equation 1.1.1 and obtain

$$\lim_{h \searrow 0} \frac{1}{h}\left(\int_t^{t+h} T(s)x\,ds - \int_0^h T(s)x\,ds\right).$$

By continuity of $s \to T(s)x$ for all $x \in X$ the fundamental theorem of calculus (Equation 1.3.1) for Banach space valued integrals yields

$$\lim_{h \searrow 0} \frac{1}{h}\left(\int_t^{t+h} T(s)x\,ds - \int_0^h T(s)x\,ds\right) = T(t)x - x.$$

For the second part of (iii) when $x \in \mathcal{D}(A)$ we already know

$$T(t)x - x = \lim_{h \searrow 0} \left(\int_0^t \frac{T(h+s)x - T(s)x}{h}\,ds\right)$$

and need to show that

$$\lim_{h \searrow 0} \left(\int_0^t T(s)\frac{T(h)x - x}{h}\,ds\right) = \int_0^t T(s)Ax\,ds.$$

This follows directly from (ii) if we can pull the limit in the integral. We can do so because

$$\lim_{h \searrow 0} \left\| \frac{T(h)x - x}{h} - Ax \right\| = 0$$

by assumption so for some $\varepsilon > 0$ there exists $h_0 > 0$ such that for all $h_0 > h > 0$

$$\left\| \frac{T(h)x - x}{h} - Ax \right\| < \varepsilon.$$

This implies that for all $0 \leq s \leq t$

$$\begin{aligned} \left\| T(s) \frac{T(h)x - x}{h} - T(s)Ax \right\| &\leq \|T(s)\| \cdot \left\| \frac{T(h)x - x}{h} - Ax \right\| \\ &\leq Me^{\omega t} \cdot \varepsilon \end{aligned}$$

by Proposition 1.4.9 for some $\omega \in \mathbb{R}$ and $M \geq 1$ and we are allowed to pull the limit in the integral by the dominated convergence which holds true also for integrals of Banach space valued functions. □

In the next lemma we see how we can rescale a strongly continuous semigroup in a way such that it remains a strongly continuous semigroup and what the generator of the rescaled semigroup looks like. (See also Nagel, Engel, chapter II, section 2.2 l[**18**])

LEMMA 1.4.16. *(Rescaled Semigroup) For a strongly continuous semigroup*

$$(T(t))_{t \in \mathbb{R}_+}$$

on a Banach space X with generator $(A, \mathcal{D}(A))$, $\lambda \in \mathbb{C}$ fixed, $t \geq 0$ and

$$S(t) := e^{-\lambda t} T(t)$$

i)

$$(S(t))_{t \in \mathbb{R}_+}$$

is a strongly continuous semigroup on a Banach space X (called rescaled semigroup) and
ii) its generator $(B, \mathcal{D}(B))$ is given by $(A - \lambda, \mathcal{D}(A))$.

PROOF.
i) If $T(t) \in L(X)$ then also

$$S(t) = e^{-\lambda t} T(t) \in L(X).$$

$(S(t))_{t\in\mathbb{R}_+}$satisfies the Functional Equation 1.1.1 because

$$\begin{aligned} S(s+t) &= e^{-\lambda(s+t)}T(s+t) \\ &= e^{-\lambda t}T(t)e^{-\lambda s}T(s) \\ &= S(t)S(s). \end{aligned}$$

$(S(t))_{t\in\mathbb{R}_+}$ is also strongly continuous because for all $x \in X$ the orbit map

$$t \to e^{-\lambda t}T(t)x$$

is the product of the two continuous maps $t \to e^{-\lambda t}$ and $t \to T(t)x$, hence continuous.
ii) In order to determine the generator B of $(S(t))_{t\in\mathbb{R}_+}$ for some $x \in \mathcal{D}(A)$ we have to find the limit

$$\lim_{h\searrow 0} \frac{1}{h}\left(S(h)x - x\right) = \lim_{h\searrow 0} \frac{1}{h}\left(e^{-\lambda h}T(h)x - x\right)$$

It follows from the fact that for small $h > 0$ $\left\|e^{-\lambda h}\right\|$ is bounded by some C that

$$\begin{aligned} &\lim_{h\searrow 0} \left\|\frac{1}{h}\left(e^{-\lambda h}T(h)x - x\right) - (-\lambda x + Ax)\right\| \\ \le &\lim_{h\searrow 0} \left\|\frac{1}{h}\left(e^{-\lambda h}T(h)x - e^{-\lambda h}x\right) - e^{-\lambda h}e^{\lambda h}Ax\right\| \\ &+ \lim_{h\searrow 0} \left\|\frac{1}{h}\left(e^{-\lambda h}x - x\right) - (-\lambda x)\right\| \\ \le &\lim_{h\searrow 0} \underbrace{\left\|e^{-\lambda h}\right\|}_{\le C} \left\|\frac{1}{h}\left(T(h)x - x\right) - e^{\lambda h}Ax\right\| \\ &+ 0 \\ = &\ 0 \end{aligned}$$

Therefore $\mathcal{D}(A) \subset \mathcal{D}(B)$ and $B = A - \lambda$. The representation

$$T(t) = e^{-(-\lambda)t}S(t)$$

and the same reasoning as above yield that also $\mathcal{D}(B) \subset \mathcal{D}(A)$. □

The generator $(A, \mathcal{D}(A))$ has also several other important properties. One of them requires the following definition:

DEFINITION 1.4.17. For a linear operator $A:\ X \supset \mathcal{D}(A) \to X$ on a normed space X its *graph* $\mathrm{gr}(A)$ is defined as

$$\mathrm{gr}(A) := \left\{(x, Ax) \subset \mathcal{D}(A) \times X\right\}.$$

On $X \times X$ we define the norm

$$\begin{aligned} \|\cdot\|_{X\times X} : X \times X &\to \mathbb{R}_+ \\ (x,y) &\to \|x\| + \|y\| , \end{aligned}$$

and call a linear operator $A: X \supset \mathcal{D}(A) \to X$ *closed* if its graph is closed in the norm $\|\cdot\|_{X\times X}$.

DEFINITION 1.4.18. For a linear operator $(A, \mathcal{D}(A))$ on a normed space X the graph norm $\|\cdot\|_A$ is defined as

$$\begin{aligned} \|\cdot\|_A : \mathcal{D}(A) \subset X &\to \mathbb{R}_+ \\ x &\to \|x\| + \|Ax\| . \end{aligned}$$

LEMMA 1.4.19. *Let $A : \mathcal{D}(A) \subset X \to X$ be a closed linear map on a Banach space X. Then $\mathcal{D}(A)$ is a Banach space with respect to $\|\cdot\|_A$ and A is continuous with respect to $\|\cdot\|_A$.*

PROOF. Clearly, $(\mathcal{D}(A), \|\cdot\|_A)$ is a normed vector space with respect to $\|\cdot\|_A$. If $(x_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $(\mathcal{D}(A), \|\cdot\|_A)$, then $\left((x_n, Ax_n)_{n\in\mathbb{N}}\right)$ is a Cauchy sequence in $\left(X \times X, \|\cdot\|_{X\times X}\right)$ which converges in$\|\cdot\|_{X\times X}$ to some $(x,y) \in X \times X$ due to completeness of $X \times X$. Since A is closed $(x,y) \in \mathrm{gr}(A)$, hence $(\mathcal{D}(A), \|\cdot\|_A)$ is complete. Continuity follows immediately from $\|Ax\| \le \|x\|_A$. □

DEFINITION 1.4.20. For a linear operator $A: X \supset \mathcal{D}(A) \to X$ on a Banach space X the subspace $D \subset \mathcal{D}(A)$ is called a *core for* A if D is closed in $\mathcal{D}(A)$ in the graph norm $\|\cdot\|_A$.

LEMMA 1.4.21. *Let $\mathcal{D}(A)$ be closed. A linear operator $A: X \supset \mathcal{D}(A) \to X$ on a Banach space X is closed if the following implication holds:*
If a sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{D}(A)$ converges to zero as n tends to infinity and for some $y \in X$ $\lim\limits_{n\to\infty} \|Ax_n - y\| = 0$, then $y = 0$.

PROOF. Let $(z_n)_{n\in\mathbb{N}} \subset \mathcal{D}(A)$ be some sequence. In order to show closedness of A we have to prove that from $\lim\limits_{n\to\infty} \|z_n - \hat{z}\| = 0$ for some $\hat{z} \in X$ and $\lim\limits_{n\to\infty} \|Az_n - \hat{y}\| = 0$ for some $\hat{y} \in X$ it follows $\hat{z} \in \mathcal{D}(A)$ and $\hat{y} = A\hat{z}$.

$\hat{z} \in \mathcal{D}(A)$ holds due to closedness of $\mathcal{D}(A)$. For $x_n := z_n - \hat{z}$ for all $n \in \mathbb{N}$ it holds $(x_n)_{n\in\mathbb{N}} \subset \mathcal{D}(A)$ and we obtain $\lim\limits_{n\to\infty} \|Ax_n - (\hat{y} - A\hat{z})\| = 0$ hence $\hat{y} = A\hat{z}$ by the assumed implication and A is closed. □

DEFINITION 1.4.22. If for a linear operator $A: X \supset \mathcal{D}(A) \to X$ on a Banach space X the closure of the graph $\mathrm{gr}(A)$ in the norm $\|\cdot\|_{X\times X}$ is the graph of another linear operator $\bar{A}: X \supset \mathcal{D}(\bar{A}) \to X$ then $\bar{A}$ is called the *closure* of A and A is called *closable*.

LEMMA 1.4.23. *A linear operator $A: X \supset \mathcal{D}(A) \to X$ on a Banach space X is closable if the following implication holds:*
If a sequence $(x_n)_{n\in\mathbb{N}} \subset \mathcal{D}(A)$ converges to zero as n tends to infinity and for some $y \in X$ $\lim\limits_{n\to\infty} \|Ax_n - y\| = 0$, then $y = 0$.

PROOF. Let $(z_n)_{n\in\mathbb{N}} \subset \mathcal{D}(A)$ be some sequence. If $\lim\limits_{n\to\infty} \|z_n - \hat{z}\| = 0$ for some $\hat{z} \in X$ and $\lim\limits_{n\to\infty} \|Az_n - \hat{y}\| = 0$ for some $\hat{y} \in X$ we can define

$$\bar{A}\hat{z} := \lim_{n\to\infty} Az_n = \hat{y}.$$

Such an operator is well-defined because if for a different sequence $(z'_n)_{n\in\mathbb{N}} \subset \mathcal{D}(A)$ such that $\lim\limits_{n\to\infty} \|z'_n - \hat{z}\| = 0$ the sequence $(Az_n)_{n\in\mathbb{N}}$ converges, then by setting $z''_n := z'_n - z_n$ we obtain a sequence $(z''_n)_{n\in\mathbb{N}} \subset \mathcal{D}(A)$ such that $\lim\limits_{n\to\infty} \|z''_n - 0\| = 0$, hence by the assumed implication $\lim\limits_{n\to\infty} \|Az'_n - Az_n\| = 0$. $\bar{A}$ inherits linearity from A and is closed by definition. □

Following the a proof presented in Nagel, Engel [**18**], chapter II, section 1, we can now show some important properties of the generator.

PROPOSITION 1.4.24. *The generator $(A, \mathcal{D}(A))$ of a strongly continuous semigroup*

$$(T(t))_{t\in\mathbb{R}_+}$$

is
(i) a closed operator
(ii) densely defined
(iii) determines the strongly continuous semigroup uniquely.

PROOF.
(i) We have to show that for a sequence $(x_n)_{n\in\mathbb{N}}$ with $x_n \in \mathcal{D}(A)$ that converges to some $x \in X$ and for a sequence $(Ax_n)_{n\in\mathbb{N}}$ that converges to some $y \in X$ it holds that

$$(x, y) \in \{(x, Ax) \subset \mathcal{D}(A) \times X\}.$$

This is the case if we can show $x \in \mathcal{D}(A)$ and $y = Ax$.
Given such two sequences $(x_n)_{n\in\mathbb{N}}$ and $(Ax_n)_{n\in\mathbb{N}}$, in order to establish a relationship between their respective limits, we first establish one between x_n and Ax_n. This is possible thanks to Proposition 1.4.15(iii) which states

$$T(t)x_n - x_n = \int_0^t T(s)Ax_n\,ds.$$

Passing to the limit we obtain

$$T(t)x - x = \lim_{n\to\infty} \int_0^t T(s)Ax_n\,ds.$$

As already seen in the proof of Proposition 1.4.15(iii), by dominated convergence we can pull the limit inside the integral thanks to the inequality

$$\|T(s)Ax_n - T(s)y\| \le Me^{\omega t}\,\|Ax_n - y\|$$

for all $0 \le s \le t$ and for some $\omega \in \mathbb{R}$ and $M \ge 1$ which holds because by Proposition 1.4.9. Doing so we obtain

$$T(t)x - x = \int_0^t T(s)y\,ds.$$

Dividing by t and letting $t \to 0$ yields thanks to the strong continuity of $(T(s))_{s\in\mathbb{R}_+}$ and the fundamental theorem of calculus (Equation 1.3.1)

$$\lim_{t\searrow 0} \frac{T(t)x - x}{t} = y.$$

Hence $x \in \mathcal{D}(A)$ and $Ax = y$.
(ii) In order to show that $\mathcal{D}(A)$ is dense in X we choose $x \in X$ arbitrarily and construct a sequence in $\mathcal{D}(A)$ that converges to x. By Proposition 1.4.15(iii) we know that

$$\frac{1}{1/n} \int_0^{1/n} T(s)x\,ds \in \mathcal{D}(A).$$

Strong continuity implies

$$\frac{1}{1/n} \int_0^{1/n} T(s)x\,ds \to x$$

hence we have found the sequence we were looking for and $\mathcal{D}(A)$ is densely defined.
(iii) In order to show that $(A, \mathcal{D}(A))$ uniquely determines the strongly continuous semigroup $(T(t))_{t\in\mathbb{R}_+}$, we assume that it is also the generator of a different strongly continuous semigroup $\{S(t)\}_{t\in\mathbb{R}_+}$and show

that both semigroups are equal. For this end, we fix some arbitrary $t_0 \in \mathbb{R}_+$ and $x \in \mathcal{D}(A)$ and for $0 \leq s \leq t_0$ we define the orbit

$$s \to \eta_x(s) := T(t_0 - s)S(s)x.$$

If we are able to show that

$$\frac{d}{ds}\eta_x(s) = 0$$

for all $0 \leq s \leq t_0$ it follows immediately that $\eta_x(0) = \eta_x(t_0)$. This implies $T(t_0)x = S(t_0)x$ and we are done since $t_0 \in \mathbb{R}_+$ and $x \in \mathcal{D}(A)$ were arbitrary.
In order to show

$$\frac{d}{ds}\eta_x(s) = 0$$

we try to write the difference quotient of $\eta_x(s)$ in a way that makes use of the difference quotients of the orbit maps $s \to S(s)x$ and $s \to T(s)x$ of which we know already the limit by Lemma 1.4.12. We obtain

$$\begin{aligned} &\frac{1}{h}\left(\eta_x(s+h) - \eta_x(s)\right) \\ = &\frac{1}{h}\left(T(t_0 - s - h)\left(S(s+h)x - S(s)x\right)\right) \\ &+ \frac{1}{h}\left(\left(T(t_0 - s - h) - T(t_0 - s)\right)S(s)x\right). \end{aligned}$$

For the second term we see that

$$\begin{aligned} &\lim_{h\to 0}\frac{1}{h}\left(\left(T(t_0 - s - h) - T(t_0 - s)\right)S(s)x\right) \\ = &-\lim_{h\to 0}\frac{1}{h}\left(T(t_0 - s) - \left(T(t_0 - s - h)\right)S(s)x\right) \end{aligned}$$

can be computed thanks to Proposition 1.4.15 (ii) which yields that $S(s)x \in \mathcal{D}(A)$. Thus, the derivative of the orbit map $t \to -T(t)S(s)x$ at $t_0 - s$ is

$$-T(t_0 - s)AS(s)x$$

by Lemma 1.4.12. For the first term we define the continuous function f :

$$\begin{aligned} &[0,1] \to X \\ &h \to \begin{cases} \frac{S(s+h)x - S(s)x}{h} & \text{if } h \in (0,1] \\ AS(s)x & \text{if } = 0. \end{cases} \end{aligned}$$

Since continuous functions map compact sets to compact sets the image $f([0,1])$ is compact. Therefore, if we set $x_h := f(h)$ we see that by

Proposition 1.4.8 the map

$$\begin{aligned}[0,t] \times f([0,1]) &\to X \\ (h, x_h) &\to T(t-s-h)x_h\end{aligned}$$

is uniformly continuous. Hence, for $x_h \to AS(s)x$ and $h \to 0$ we obtain

$$T(t-s-h)x_h \to T(t-s)AS(s)x.$$

Thus the difference quotient is

$$\begin{aligned}\frac{d}{ds}\eta_x(s) &= T(t-s)AS(s)x - T(t-s)AS(s)x \\ &= 0.\end{aligned}$$

and we conclude. □

The concept of a *spectrum* of a closed linear operator on a Banach space defined below generalizes the concept of eigenvalues of matrices.

Definition 1.4.25. The *resolvent set* $\rho(A)$ of a densely defined linear operator $(A, \mathcal{D}(A))$ on a Banach space is defined as

$$\rho(A) := \{\lambda \in \mathbb{C} : \ \lambda - A : X \supset \mathcal{D}(A) \to X \text{ is bijective with continuous inverse}\}.$$

For$\lambda \in \rho(A)$ the *resolvent* $R(\lambda, A)$ is defined as:

$$\begin{aligned}X &\to \mathcal{D}(A) \\ x &\to (\lambda - A)^{-1}x.\end{aligned}$$

The set

$$\sigma(A) := \mathbb{C} \setminus \rho(A)$$

is called *spectrum*.

Remark 1.4.26. A closed bijective operator $A : \mathcal{D}(A) \subset X \to X$ is by Lemma 1.4.19 continuous with respect to $(\mathcal{D}(A), \|\cdot\|_A)$ (see Definition 1.4.18) and by the open mapping theorem (Theorem A.4.9) the preimage of an open set in X under A is an open set in $(\mathcal{D}(A), \|\cdot\|_A)$. By $\|\cdot\| \leq \|\cdot\|_A$ the identity

$$(\mathcal{D}(A), \|\cdot\|_A) \to (X, \|\cdot\|)$$

is a bounded linear operator between Banach spaces (compare Lemma 1.4.19) hence again by the open mapping theorem is an open map. Thus, any open set in $(\mathcal{D}(A), \|\cdot\|_A)$ is open in X and the inverse of a closed bijective operator is continuous.

On the other hand, if the densely defined linear operator $(A, \mathcal{D}(A))$ is not closed, then $\rho(A) = \emptyset$. In order to see this, observe that by non-closedness there is $(x_n)_{n\in\mathbb{N}} \subset \mathcal{D}(A)$ such that $x_n \to x \in X$ and

$Ax_n \to y \in X$ but either $x \notin \mathcal{D}(A)$ or $x \in \mathcal{D}(A)$ but $Ax \neq y$. If there was some$\lambda \in \rho(A)$ then due to continuity of $R(\lambda, A)$

$$R(\lambda, A)\,(\lambda x - y) = \lim_{n\to\infty} R(\lambda, A)\,(\lambda x_n - Ax_n) = x,$$

which would yield $x \in \mathcal{D}(A)$ and by

$$x = R(\lambda, A)\,(\lambda x - Ax)\,,$$

and injectivity of $R(\lambda, A)$ we would obtain $Ax = y$ in contradiction to non-closedness of A. Hence $\rho(A) = \emptyset$.

However, thanks to closedness of the generator A we can hope for $\rho(A) \neq \emptyset$ and the following proposition yields that the resolvent set is open (the proof follows [**18**], Proposition IV.1.3):

PROPOSITION 1.4.27. *For a closed operator $A : X \supset \mathcal{D}(A) \to X$ for $\mu \in \rho(A)$ and $\lambda \in \mathbb{C}$ such that $|\mu - \lambda| < \frac{1}{\|R(\mu,A)\|}$ it holds $\lambda \in \rho(A)$ and*

$$R(\lambda, A) = \sum_{n=0}^{\infty} (\mu - \lambda)^n\, R(\mu, A)^{n+1}.$$

PROOF. We write $(\lambda - A)$ in terms of $(\mu - A)$ as

$$\begin{aligned} \lambda - A &= \mu - A - \mu + \lambda \\ &= [\mathrm{Id} + (-\mu + \lambda)\, R\,(\mu, A)]\,(\mu - A)\,. \end{aligned} \tag{1.4.1}$$

By assumption $\mu - A$ is bijective. Since

$$\|(-\mu + \lambda)\, R\,(\mu, A)\| < 1$$

the equation

$$(\mathrm{Id} + (-\mu + \lambda)\,(R(\mu, A))\, x = 0$$

for some $x \in X$ implies $x = 0$ thus

$$\mathrm{Id} + (-\mu + \lambda)\, R\,(\mu, A)$$

is injective which yields injectivity of $\lambda - A$ by Equation 1.4.1. Regarding surjectivity of $\lambda - A$; it is enough to show surjectivity of

$$\mathrm{Id} + (-\mu + \lambda)\, R\,(\mu, A)\,.$$

Let $y \in X$. Then for

$$z := \sum_{i=0}^{\infty} \left[(\mu - \lambda)\, R\,(\mu, A)\right]^i y \tag{1.4.2}$$

the series converges thanks to

$$\|(\mu - \lambda)\, R\,(\mu, A)\| < 1,$$

and we obtain

$$\begin{aligned}
&(\mathrm{Id}+(-\mu+\lambda)\,(R(\mu,A))\,z\\
=&(\mathrm{Id}+(-\mu+\lambda)\,(R(\mu,A))\left(y+\sum_{i=1}^{\infty}\left[(\mu-\lambda)\,R\,(\mu,A)\right]^{i}y\right)\\
=&y+(-\mu+\lambda)\,(R(\mu,A)\,y\\
&+\sum_{i=1}^{\infty}\left[(\mu-\lambda)\,R\,(\mu,A)\right]^{i}y\\
&+(-\mu+\lambda)\,(R(\mu,A)\sum_{i=1}^{\infty}\left[(\mu-\lambda)\,R\,(\mu,A)\right]^{i}y\\
=&y.
\end{aligned}$$

Hence, $\mathrm{Id}+(-\mu+\lambda)\,R\,(\mu,A)$ is a linear bounded bijective operator and its continuous inverse is given by

$$(\mathrm{Id}+(-\mu+\lambda)\,R\,(\mu,A))^{-1}=\sum_{i=0}^{\infty}\left[(\mu-\lambda)\,R\,(\mu,A)\right]^{i}.$$

Thus,

$$\lambda-A=\left[\mathrm{Id}+(-\mu+\lambda)\,R\,(\mu,A)\right](\mu-A)$$

is surjective, and therefore also bijective. By the previous calculations the resolvent is given by

$$\begin{aligned}
R(\lambda,A)&=\left(\left[\mathrm{Id}+(-\mu+\lambda)\,R\,(\mu,A)\right](\mu-A)\right)^{-1}\\
&=(\mu-A)^{-1}\left(\left[\mathrm{Id}+(-\mu+\lambda)\,R\,(\mu,A)\right]\right)^{-1}\\
&=R(\mu,A)\left(\sum_{n=0}^{\infty}(\mu-\lambda)^{n}\,R\,(\mu,A)^{n}\right).
\end{aligned}$$

□

Related to the strongly continuous semigroup $(T(t))_{t\in\mathbb{R}_+}$, so far we have been dealing with the two objects *generator* $(A,\mathcal{D}(A))$ and *resolvent* $R(\lambda,A)$. Its definition relates the generator A to the semigroup $(T(t))_{t\in\mathbb{R}_+}$. Also, the definition of the resolvent $R(\lambda,A)$ provides a link between the resolvent and the generator A. The missing link is a direct relation between resolvent $R(\lambda,A)$ and semigroup $(T(t))_{t\in\mathbb{R}_+}$. As a preparation we observe the following:

LEMMA 1.4.28. *For a generator $(A,\mathcal{D}(A))$ of a strongly continuous semigroup*

$$(T(t))_{t\in\mathbb{R}_+}$$

on a Banach space X and for $\lambda \in \mathbb{C}$ and $t > 0$ the following equations hold

$$e^{-\lambda t}T(t)x - x = (A-\lambda)\int_0^t e^{-\lambda s}T(s)x\,ds \text{ for } x \in X$$
$$= \int_0^t e^{-\lambda s}T(s)\,(A-\lambda)\,x\,ds \text{ for } x \in \mathcal{D}(A).$$

PROOF. For the rescaled semigroup $S(t) = e^{-\lambda t}T(t)$ and its generator

$$(B, \mathcal{D}(B)) = (A-\lambda, \mathcal{D}(A))$$

from Lemma 1.4.16, Proposition 1.4.15 (iii) states that

$$S(t)x - x = (A-\lambda)\int_0^t S(t)x\,ds \text{ for } x \in X$$

and

$$S(t)x - x = \int_0^t S(t)\,(A-\lambda)\,x\,ds$$

which is the statement of the lemma. □

Heuristically, if in the Lemma above we choose λ large and then let $t \to \infty$, the left hand side converges to $-x$. This leads to the idea that

$$\lim_{t\to\infty}\int_0^t e^{-\lambda s}T(s)x = (A-\lambda)^{-1}\,x.$$

The idea is made precise in the *integral representation of the resolvent* below which provides the link between the semigroup and the resolvent of its generator. The theorem also shows that the spectrum of the generator is contained in some left half plane of $\mathbb{C}$. The proof follows the one presented in chapter II, section 1 of Engel, Nagel [**18**]

THEOREM 1.4.29. *Let*

$$(T(t))_{t\in\mathbb{R}_+}$$

be a strongly continuous semigroup on the Banach space X such that

$$\|T(t)\| \le Me^{\omega t}$$

for some constants $\omega \in \mathbb{R}$ and $M \ge 1$. For the generator $(A, \mathcal{D}(A))$ of

$$(T(t))_{t\in\mathbb{R}_+}$$

the following properties hold:
(i) If for $\lambda \in \mathbb{C}$ the expression

$$\begin{aligned} R(\lambda)x := & \int_0^\infty e^{-\lambda s} T(s)x\, ds \\ = & \lim_{t\to\infty} \int_0^t e^{-\lambda s} T(s)x\, ds \end{aligned}$$

exists for all $x \in X$, then $\lambda \in \rho(A)$ and

$$R(\lambda, A)x = \int_0^\infty e^{-\lambda s} T(s)x\, ds.$$

(ii) If $Re\lambda > \omega$, then $\lambda \in \rho(A)$ and

$$R(\lambda, A)x = \int_0^\infty e^{-\lambda s} T(s)x\, ds$$

and $\|R(\lambda, A)\| \le \frac{M}{Re\lambda - \omega}$.

Proof. We first show that ii) follows from i) .

ii) In order to show existence of

$$\lim_{t\to\infty} \int_0^t e^{-\lambda s} T(s)x\, ds$$

for all $x \in X$ we first need to show that

$$\left(\int_0^n e^{-\lambda s} T(s)x\, ds\right)_{n\in\mathbb{R}_+}$$

is a Cauchy sequence for all $x \in X$ if $\mathrm{Re}\lambda > \omega$. For any $u, v \in \mathbb{R}$ we have

$$\begin{aligned} \left\|\int_u^v e^{-\lambda s} T(s)x\, ds\right\| & \le \int_u^v \left\|e^{-\lambda s} T(s)x\right\|\, ds \\ & \le \int_u^v e^{-(\mathrm{Re}\lambda)s} M e^{\omega s} \|x\|\, ds \\ & = M \int_u^v e^{(\omega-\mathrm{Re}\lambda)s} \|x\|\, ds \\ & = M \left\{\frac{e^{(\omega-\mathrm{Re}\lambda)u}}{\mathrm{Re}\lambda-\omega} - \frac{e^{(\omega-\mathrm{Re}\lambda)v}}{\mathrm{Re}\lambda-\omega}\right\} \|x\| \end{aligned} \tag{1.4.3}$$

Hence, for all $x \in X$ and for any $\varepsilon > 0$ there exists some $N \in \mathbb{N}$ such that for all $n, m > N$ it holds

$$\left\|\int_n^m e^{-\lambda s} T(s)x\, ds\right\| \le \varepsilon.$$

Thus,

$$\left(\int_0^n e^{-\lambda s} T(s)x \, ds \right)_{n \in \mathbb{R}_+}$$

is a Cauchy sequence and

$$\lim_{n \to \infty} \left(\int_0^n e^{-\lambda s} T(s)x \, ds \right)$$

exists. By Equation 1.4.3 also

$$\lim_{t \to \infty} \left\| \left(\int_0^t e^{-\lambda s} T(s)x \, ds \right) - \lim_{n \to \infty} \left(\int_0^n e^{-\lambda s} T(s)x \, ds \right) \right\| = 0.$$

Hence also

$$\lim_{t \to \infty} \left(\int_0^t e^{-\lambda s} T(s)x \, ds \right)$$

exists. By i) $\lambda \in \rho(A)$ and

$$R(\lambda, A)x = \int_0^\infty e^{-\lambda s} T(s)x \, ds$$

thus by Equation 1.4.3

$$\|R(\lambda, A)x\| \leq \frac{M}{\text{Re}\lambda - \omega} \|x\| \text{ for all } x \in X.$$

i) We have to show that if for some $\lambda \in \mathbb{C}$ the integral

$$\int_0^\infty e^{-\lambda s} T(s)x \, ds$$

exists, then for all $x \in X$

$$\int_0^\infty e^{-\lambda s} T(s)x \, ds \in \mathcal{D}(A)$$

and

$$(\lambda - A) \left(\int_0^\infty e^{-\lambda s} T(s)x \, ds \right) = x,$$

and for all $x \in \mathcal{D}(A)$

$$\left(\int_0^\infty e^{-\lambda s} T(s) \left((\lambda - A) x \right) ds \right) = x.$$

The proof has two steps.
As a first step, we show the statement for $\lambda = 0$.
We show first that for all $x \in X$

$$\int_0^\infty T(s)x \, ds \in \mathcal{D}(A)$$

and

$$(-A)\left(\int_0^\infty T(s)x\,ds\right) = x. \tag{1.4.4}$$

Toward this end, we calculate

$$\begin{aligned}&\lim_{h\searrow 0}\frac{T(h)-T(0)}{h}\int_0^\infty T(s)x\,ds\\ &= \lim_{h\searrow 0}\frac{T(h)-\mathrm{Id}}{h}\lim_{t\to\infty}\int_0^t T(s)x\,ds.\end{aligned}$$

Since we can pull linear bounded operators in the integral (see Equation 1.3.2), this equals

$$\lim_{h\searrow 0}\frac{1}{h}\lim_{t\to\infty}\int_0^t T(s+h)x - T(s)x\,ds.$$

Rewriting the integration bounds in the integral this yields

$$\lim_{h\searrow 0}\frac{1}{h}\lim_{t\to\infty}\left(\int_0^{t+h} T(s)x\,ds - \int_0^h T(s)x\,ds - \int_0^t T(s)x\,ds\right).$$

Since $\lim_{t\to\infty}\int_0^t T(s)x\,ds$ converges by assumption we obtain

$$\lim_{t\to\infty}\int_0^{t+h} T(s)x\,ds - \int_0^t T(s)x\,ds = 0$$

for any $h > 0$ thus

$$\begin{aligned}\lim_{h\searrow 0}\lim_{t\to\infty}\frac{T(h)-\mathrm{Id}}{h}\int_0^t T(s)x\,ds &= \lim_{h\searrow 0}\frac{1}{h}\left(-\int_0^h T(s)x\,ds\right)\\ &= -x,\end{aligned}$$

the last step being possible thanks to the fundamental theorem of calculus (Equation 1.3.1) and strong continuity of

$$(T(t))_{t\in\mathbb{R}_+}.$$

Hence $\int_0^\infty T(s)x\,ds \in \mathcal{D}(A)$ and $A\left(\int_0^\infty T(s)x\,ds\right) = -x$.
Next, we fix some arbitrary $x \in \mathcal{D}(A)$ and show

$$\left(\int_0^\infty T(s)\,(-Ax)\;ds\right) = x.$$

By assumption we know

$$\lim_{t\to\infty}\int_0^t T(s)\,(-Ax)\;ds$$

converges in X. Proposition 1.4.15(iii) implies that for any $t \in \mathbb{R}_+$

$$\int_0^t T(s)(x)\ ds \in \mathcal{D}(A)$$

and

$$\lim_{t\to\infty} \int_0^t T(s)(-Ax)\ ds = \lim_{t\to\infty}\ -A\left(\int_0^t T(s)(x)\ ds\right).$$

Therefore it is sufficient to compute the right hand side. Since we know by assumption that

$$\lim_{t\to\infty} \int_0^t T(s)(x)\ ds$$

exists, closedness of the generator A (Proposition 1.4.24) yields

$$\lim_{t\to\infty}\ -A\left(\int_0^t T(s)(x)\ ds\right) = -A\left(\lim_{t\to\infty} \int_0^t T(s)(x)\ ds\right).$$

Thus,

$$\int_0^\infty T(s)(-Ax)\ ds = -A\left(\lim_{t\to\infty} \int_0^t T(s)(x)\ ds\right),$$

and Equation 1.4.4 states

$$-A\left(\lim_{t\to\infty} \int_0^t T(s)(x)\ ds\right) = x.$$

As a second step, we show the statement for any $\lambda \in \mathbb{C}$. So for a strongly continuous semigroup $(T(t))_{t\in\mathbb{R}_+}$ with generator $(A, \mathcal{D}(A))$ let

$$\lim_{t\to\infty} \int_0^t e^{-\lambda s} T(s)x\, ds$$

exist for all $x \in X$. We then use the rescaled semigroup

$$S(t) := e^{-\lambda t} T(t)$$

from Lemma 1.4.16 whose generator is given by

$$(B, \mathcal{D}(B)) = (A - \lambda, \mathcal{D}(A)).$$

Clearly

$$\lim_{t\to\infty} \int_0^t S(s)x\, ds$$

exists for all $x \in X$. Thus, by the first step of the proof for all $x \in X$

$$\int_0^\infty S(s)x\, ds \in \mathcal{D}(B)$$

and

$$(-B)\left(\int_0^\infty S(s)x\,ds\right) = x$$

and for all $x \in \mathcal{D}(B)$

$$\left(\int_0^\infty S(s)\left((-B)\,x\right)\,ds\right) = x.$$

We conclude by substituting back. □

This corollary is taken from [**20**].

COROLLARY 1.4.30. *Let*

$$(T(t))_{t\in\mathbb{R}_+}$$

be a strongly continuous semigroup on the Banach space X *such that*

$$\|T(t)\| \leq Me^{\omega t}$$

for some constants $\omega \in \mathbb{R}$ *and* $M \geq 1$. *Then for all* $n \in \mathbb{N}$ *and* $\lambda \in \mathbb{C}$ *such that* $Re\lambda > \omega$ *it holds*

$$\|(R(\lambda, A))^n\| \leq \frac{M}{(Re\lambda - \omega)^n}.$$

PROOF. By Theorem 1.4.29 and the property of the integral that we can pull linear operators inside (see Equation 1.3.2) we can write

$$\begin{aligned}
&(R(\lambda, A))^n\,x \\
&= \left(\int_0^\infty ... \int_0^\infty \int_0^\infty e^{-\lambda(s_1+s_2+...+s_n)}T(s_1+s_2+...+s_n)x\,ds_1ds_2...ds_n\right).
\end{aligned}$$

Using $\|T(t)\| \leq Me^{\omega t}$ we obtain

$$\begin{aligned}
&\|(R(\lambda, A))^n\,x\| \\
&\leq \left(\int_0^\infty ... \int_0^\infty \int_0^\infty \left|e^{-\lambda(s_1+s_2+...+s_n)}\right| Me^{\omega(s_1+s_2+...+s_n)}\,ds_1ds_2...ds_n\right)\|x\| \\
&= M\left(\int_0^\infty ... \int_0^\infty \int_0^\infty e^{(\omega-\mathrm{Re}\lambda)(s_1+s_2+...+s_n)}\,ds_1ds_2...ds_n\right)\|x\| \\
&\leq M\left(\frac{1}{\mathrm{Re}\lambda - \omega}\right)^n \|x\|.
\end{aligned}$$

□

1.4.3. Hille-Yosida Theorem. We turn our attention to the question which kind of linear operators $(A, \mathcal{D}(A))$ on X are generators of some strongly continuous semigroup. We have seen in Proposition 1.3.10 that for any *bounded* linear operator A the family

$$\left(e^{tA}\right)_{t\in\mathbb{R}_+}$$

of linear bounded operators is a uniformly continuous semigroup. However, generators of strongly continuous semigroup are *unbounded* if the semigroup is not uniformly continuous. We have already seen that not all unbounded linear operators are generators of semigroups and that some necessary conditions have to be satisfied. From Proposition 1.4.24 we remember that a generator $(A, \mathcal{D}(A))$ is closed and densely defined. Furthermore, from Proposition 1.4.29 (ii) we observe that there is $\omega \in \mathbb{R}$ such that $\mathrm{Re}\lambda > \omega$ implies $\lambda \in \rho(A)$ and from Corollary 1.4.30 we obtain that in this case the norm of the resolvent

$$R(\lambda, A) = (\lambda - A)^{-1}$$

is bounded by

$$\|(R(\lambda, A))^n\| \leq \frac{M}{(\mathrm{Re}\lambda - \omega)^n}$$

for some $M \geq 1$ and all $n \in \mathbb{N}$. We want to show that these necessary conditions are also sufficient.
For this purpose, we will approximate the unbounded operator $(A, \mathcal{D}(A))$ by bounded operators in order to apply the result for bounded operators from Proposition 1.3.10. For this approximation procedure, we need to find bounded operators that are closely connected to $(A, \mathcal{D}(A))$. One such operator is clearly the resolvent

$$R(\lambda, A) = (\lambda - A)^{-1}$$

which by definition is a bounded operator whenever $\lambda \in \rho(A)$ which is the case for $\mathrm{Re}\lambda > \omega$ according to Proposition 1.4.29, where $\omega \in \mathbb{R}$ is the exponent in the bound

$$\|T(t)\| \leq Me^{\omega t}$$

that holds for strongly continuous semigroups (see Proposition 1.4.9). Since we do not want the approximation of the generator to appear out of nowhere we choose not give present the most concise proof possible but instead to follow a more didactic approach in the next two lemmata.

This subsection is based on chapter II, section 3a of Nagel, Engel [**18**] and section 4 of Hairer [**20**].

In the following lemma we see that there are other operators that can be expressed in terms of the resolvent and that are bounded linear operators.

LEMMA 1.4.31. *Let $(A, \mathcal{D}(A))$ be linear operator on X and $\lambda \in \rho(A)$. For $x \in X$*

$$\lambda A R(\lambda, A)x = -\lambda x + \lambda^2 R(\lambda, A)x.$$

For $x \in \mathcal{D}(A)$

$$\lambda R(\lambda, A)Ax = -\lambda x + \lambda^2 R(\lambda, A)x.$$

PROOF. For $x \in X$

$$\lambda A R(\lambda, A)x = \lambda (A - \lambda) R(\lambda, A)x + \lambda^2 R(\lambda, A)x$$

and for $x \in \mathcal{D}(A)$

$$\lambda R(\lambda, A)Ax = \lambda R(\lambda, A) (A - \lambda) x + \lambda R(\lambda, A)\lambda x.$$

□

In order to approximate the generator $(A, \mathcal{D}(A))$ of a strongly continuous semigroup on X by bounded operators related to the resolvent we observe that for $\lambda > 0$ large enough by Theorem 1.4.29 $\lambda \in \rho(A)$. For $Ax \in \mathcal{D}(A)$ we have the identity

$$\begin{aligned} Ax &= R(\lambda, A) (\lambda - A) (Ax) \\ &= \lambda R(\lambda, A) (Ax) - R(\lambda, A) A (Ax) . \end{aligned} \tag{1.4.5}$$

By the previous lemma $\lambda R(\lambda, A)A$ is a linear bounded operator for λ large enough. However, $R(\lambda, A)A^2$ is in general not bounded. Therefore, we would like to show that $R(\lambda, A)A (Ax)$ approaches zero when λ approaches infinity. Following this idea we show the next lemma.

LEMMA 1.4.32. *Let $(A, \mathcal{D}(A))$ be a linear operator on a Banach space X for which there exists $M > 0$ and $\omega \in \mathbb{R}$ such that $[\omega, \infty) \subset \rho(A)$ and for all $\lambda \in [\omega, \infty)$*

$$\|\lambda R(\lambda, A)\| \leq M.$$

Then for all $y \in \mathcal{D}(A)$ it holds

$$\lim_{\lambda \to \infty} \lambda R(\lambda, A)y = y.$$

PROOF. We fix $y \in \mathcal{D}(A)$. $R(\lambda, A)$ exists and as in Equation 1.4.5 above we obtain

$$\lambda R(\lambda, A) (y) - y = R(\lambda, A)A (y) ,$$

which yields the bounds

$$\begin{aligned}\|\lambda R(\lambda, A)\,(y) - y\| &\leq \|R(\lambda, A)A\,(y)\| \\ &\leq \frac{1}{\lambda}\,\|\lambda R(\lambda, A)\|\,\|A\,(y)\| \\ &\leq \frac{1}{\lambda}M\,\|A\,(y)\|\,.\end{aligned}$$

Hence

$$\lim_{\lambda\to 0}\|\lambda R(\lambda, A)\,(y) - y\| = 0$$

for any $y \in \mathcal{D}(A)$. □

In the next step we would like to drop the assumption in Lemma 1.4.32 that $y \in \mathcal{D}(A)$. Since a generator $(A, \mathcal{D}(A))$ is densely defined, the equation

$$\lim_{\lambda\to\infty}\lambda R(\lambda, A)y = y$$

holds true on a dense subset of X. We can use the following lemma.

LEMMA 1.4.33. *Let X be a Banach space and let $(L_n)_{n\in\mathbb{N}} \in L(X)$ be a sequence of bounded linear operators. If there exists $C < \infty$ such that*

$$\sup_{n\in\mathbb{N}}\|L_n\| \leq C$$

then it holds: If for a dense subset D of X

$$\lim_{n\to\infty} L_n x = 0$$

for all $x \in D$, then

$$\lim_{n\to\infty} L_n x = 0$$

for all $x \in X$.

PROOF. For $x \in X$ fixed and arbitrary choose a sequence $(x_m)_{m\in\mathbb{N}}$ in D such that

$$\lim_{m\to\infty} x_m \to x.$$

Fix some arbitrary $\varepsilon > 0$ and choose m_ε such that $\|x_m - x\| < \varepsilon$ for all $m > m_\varepsilon$. We obtain

$$\begin{aligned}\left\|\lim_{n\to\infty} L_n x\right\| &\leq \left\|\lim_{n\to\infty} L_n\,(x_m)\right\| + \left\|\lim_{n\to\infty} L_n\,(x - x_m)\right\| \\ &\leq 0 + \lim_{n\to\infty}\|L_n\|\,\|x - x_m\| \\ &\leq 0 + C\varepsilon\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary this implies the statement of the lemma. □

After these preparations, we obtain the Yosida approximation.

Proposition 1.4.34. *(Yosida approximation) For a densely defined linear operator $(A, \mathcal{D}(A))$ on a Banach space X for which there exists $M > 0$ and $\omega \in \mathbb{R}$ such that $[\omega, \infty) \subset \rho(A)$ and for all $\lambda \in [\omega, \infty)$*

$$\|\lambda R(\lambda, A)\| \leq M$$

it holds true: For all $y \in X$ and $\lambda \in [\omega, \infty)$

$$\lim_{\lambda \to \infty} \lambda R(\lambda, A) y = y,$$

and for all $x \in \mathcal{D}(A)$

$$\lim_{\lambda \to \infty} \lambda R(\lambda, A) A x = A x,$$

where the linear operator $\lambda R(\lambda, A) A$ is bounded on $\mathcal{D}(A)$. Also for

$$A_\lambda := \lambda A R(\lambda, A) = -\lambda + \lambda^2 R(\lambda, A)$$

we obtain for all $x \in \mathcal{D}(A)$

$$\lim_{\lambda \to \infty} A_\lambda x = A x.$$

Proof. Since $(A, \mathcal{D}(A))$ is densely defined we apply Lemma 1.4.33 to Lemma 1.4.32 and obtain

$$\lim_{\lambda \to \infty} \lambda R(\lambda, A) y = y$$

for all $y \in X$. Setting $y = Ax$ yields the second limit. By Lemma 1.4.31 for all $x \in \mathcal{D}(A)$

$$\lambda A R(\lambda, A) x = \lambda R(\lambda, A) A x$$

which implies the third limit. By the identity in Lemma 1.4.31, the linear operator $\lambda R(\lambda, A) A$ is bounded on $\mathcal{D}(A)$ for all $\lambda \in [\omega, \infty)$. □

Using the approximations above, we can characterize the class of linear operators that are a generator of some strongly continuous semigroup. A first version of this theorem was proved by Hille and Yosida in 1948. The following -more general- version was shown by Feller, Miyadera and Phillips in 1952 (see Engel, Nagel, chapter III [**18**]).

Theorem 1.4.35. *(Hille-Yosida)*

A closed and densely defined operator $(A, \mathcal{D}(A))$ on the Banach space X is the generator of a strongly continuous semigroup

$$(T(t))_{t \in \mathbb{R}_+}$$

with

$$\|T(t)\| \leq M e^{\omega t}$$

for $M \geq 1$ and $\omega \in \mathbb{R}$ if and only if for all $\lambda \in \mathbb{C}$ the inequality $Re\lambda > \omega$ implies $\lambda \in \rho(A)$ and for all $\lambda \in \mathbb{C}$ such that $Re\lambda > \omega$ the bound

$$\|(R(\lambda, A))^n\| \leq \frac{M}{(Re\lambda - \omega)^n}$$

holds for every $n \geq 1$.

Proof. It has already been shown in Proposition 1.4.24, Proposition 1.4.29 (ii), and Corollary 1.4.30 that for a strongly continuous semigroup

$$(T(t))_{t\in\mathbb{R}_+}$$

with

$$\|T(t)\| \leq Me^{\omega t}$$

for $M \geq 1$ and $\omega \in \mathbb{R}$ the stated properties hold. We only need to show the other implication.

For this purpose, we use the Yosida approximation from Proposition 1.4.34 in order to approximate the operator $(A, \mathcal{D}(A))$ pointwise on $\mathcal{D}(A)$ by the sequence

$$(A_n)_{n\in\mathbb{N},\, n>\omega}$$

of bounded linear operators where each A_n is defined on all of X as

$$A_n := nAR(n, A).$$

By Proposition 1.3.10 we know that for each $n \in \mathbb{N}$, $n > \omega$ the family

$$(T_n(t))_{t\in\mathbb{R}_+}$$

of bounded linear operators on X defined by

$$T_n(t) := e^{tA_n} = \sum_{k=0}^{\infty} \frac{(tA_n)^k}{k!}$$

is a uniformly continuous semigroup. We want to show that

(i) for all $t \in \mathbb{R}_+$ and for each $x \in X$

$$T(t)x := \lim_{n\to\infty} T_n(t)x$$

exists and that

(ii)

$$\{T(t)\}_{t\in\mathbb{R}_+}$$

is a strongly continuous semigroup on X which

(iii) possesses A as a generator.

In order to show (i) it is enough to show that for all $t \in \mathbb{R}_+$ and for all $x \in X$

$$(T_n(t)x)_{n\in\mathbb{N},\, n>\omega}$$

is a Cauchy sequence in X. The proof has three steps.
First, we show that for all $x \in X$ and for each $n, m \in \mathbb{N}$, $n, m > \omega$

$$\|T_n(t)x - T_m(t)x\| \leq \int_0^t \|T_m(t-s)T_n(s)\| \, \|A_n x - A_m x\| \, ds. \tag{1.4.6}$$

In order to do so, we would like to use the fact that since for each $n \in \mathbb{N}$, $n > \omega$ A_n is a bounded operator with domain X, we know the derivative of $T_n(t)x$ with respect to t for any $x \in X$ and any $n \in \mathbb{N}$, $n > \omega$. Thus, for $x \in X$ and $n, m \in \mathbb{N}$, $n, m > \omega$ we write

$$T_n(t)x - T_m(t)x = \int_0^t \frac{d}{ds}\left(T_m(t-s)T_n(s)x\right) ds$$

using the fundamental theorem of calculus for Banach space valued integrals. We derive the integrand and obtain by the product rule for $x \in X$ and $n, m \in \mathbb{N}$, $n, m > \omega$:

$$T_n(t)x - T_m(t)x = \int_0^t \left(-A_m T_m(t-s)T_n(s)x + T_m(t-s)A_n T_n(s)x\right) ds.$$

A_n and T_n (and A_m and T_m) commute by definition of T_n (and T_m) for any $n, m \in \mathbb{N}$, $n, m > \omega$. A_n and A_m commute by their representations

$$A_n = -n\mathrm{Id} + n^2 R(n, A)$$

and

$$A_m = -m\mathrm{Id} + m^2 R(m, A)$$

for any $n, m \in \mathbb{N}$, $n, m > \omega$. Hence by definition of T_n also A_m and T_n commute for any $n, m \in \mathbb{N}$, $n, m > \omega$ and we obtain the following estimate:

$$\|T_n(t)x - T_m(t)x\| \leq \int_0^t \|T_m(t-s)T_n(s)\| \, \|A_n x - A_m x\| \, ds. \tag{1.4.7}$$

Second, since we know that

$$\lim_{n,m\to\infty} \|A_n x - A_m x\| = 0$$

for all $x \in \mathcal{D}(A)$, in order to show that for all $t \in \mathbb{R}_+$ and for all $x \in \mathcal{D}(A)$

$$(T_n(t)x)_{n\in\mathbb{N},\, n>\omega}$$

is a Cauchy sequence in X by the first step it suffices to show that there exists $C < \infty$ such that for all $s \in [0,t]$

$$\sup_{n\in\mathbb{N},\, n>\omega} \|T_n(s)\| \le C.$$

This is shown as follows. For $s \in [0,t]$ and $n \in \mathbb{N}$, $n > \omega$ by Lemma 1.3.9 we have for $T_n(s) = e^{sA_n}$:

$$\begin{aligned} T_n(s) &= e^{\left(-sn\mathrm{Id}+sn^2R(n,A)\right)} \\ &= e^{-sn}e^{\left(sn^2R(n,A)\right)}. \end{aligned}$$

We can estimate this and obtain for any $n \in \mathbb{N}$, $n > \omega$ and $s \in [0,t]$ and any $x \in X$

$$\begin{aligned} \|T_n(s)x\| &\le e^{-sn} \sum_{k=0}^{\infty} \frac{\left\| \left[sn^2\left(R(n,A)\right)\right]^k x \right\|}{k!} \\ &= e^{-sn} M \sum_{k=0}^{\infty} \frac{\left(\frac{sn^2}{(n-\omega)}\right)^k}{k!} \|x\| \\ &= e^{-sn} M e^{\frac{sn^2}{(n-\omega)}} \|x\| \\ &= M e^{\frac{sn\omega}{n-\omega}} \|x\|. \end{aligned}$$

Therefore $\|T_n(s)\| \le M e^{\frac{sn\omega}{n-\omega}}$ for any $n \in \mathbb{N}$, $n > \omega$ and $s \in [0,t]$. This implies that for all $s \in [0,t]$ and any $n \in \mathbb{N}$, $n > \omega$

$$\sup_{n\in\mathbb{N}} \|T_n(s)\| \le M \cdot \max\left\{ e^{\frac{sn_0\omega}{n_0-\omega}}, 1 \right\} \tag{1.4.8}$$

where

$$n_0 := \min\left\{n \in \mathbb{N} :\ n > \omega\right\}.$$

Hence by Inequality 1.4.6 and

$$\lim_{n,m\to\infty} \|A_n x - A_m x\| = 0$$

$$\|T_n(t)x - T_m(t)x\|$$

is a Cauchy sequence in X for all $x \in \mathcal{D}(A)$ and since X is a Banach space for all $x \in \mathcal{D}(A)$ and $t \in \mathbb{R}_+$ the Cauchy sequence converges to some

$$T(t)x := \lim_{n\to\infty} T_n(t)x.$$

We also see that the inequalities

$$\|T(t)x\| \le \lim_{n\to\infty} \|T_n(t)x\| \le M e^{t\omega} \|x\| \tag{1.4.9}$$

hold for all $x \in \mathcal{D}(A)$ and $t \in \mathbb{R}_+$.
Third, we want to show that

$$T(t)x := \lim_{n\to\infty} T_n(t)x.$$

exists also for any $x \in X\setminus\mathcal{D}(A)$ and $t \in \mathbb{R}_+$. We fix some $x \in X\setminus\mathcal{D}(A)$ and choose a sequence $(x_m)_{m\in\mathbb{N}}$ with $x_m \in \mathcal{D}(A)$ such that

$$x = \lim_{m\to\infty} x_m$$

for all $m \in \mathbb{N}$. Since $(x_m)_{m\in\mathbb{N}}$ is a Cauchy sequence in X we see because of Inequality 1.4.9

$$\begin{aligned}\|T(t)x_{m_1} - T(t)x_{m_2}\| &= \|T(t)\left(x_{m_1} - x_{m_2}\right)\| \\ &\le Me^{t\omega}\left\|x_{m_1} - x_{m_2}\right\|\end{aligned}$$

that also $(T(t)x_m)_{m\in\mathbb{N}}$ is a Cauchy sequence in X. Thus,

$$T(t)x := \lim_{m\to\infty} T(t)x_m$$

exists. This limit does not depend on the choice of the sequence which we see in the following way. We choose another sequence $(\tilde{x}_m)_{m\in\mathbb{N}}$ with $\tilde{x}_m \in \mathcal{D}(A)$ for all $m \in \mathbb{N}$ such that

$$x = \lim_{m\to\infty} \tilde{x}_m$$

and see that the inequality

$$\|T(t)\tilde{x}_m - T(t)x_m\| \le Me^{t\omega}\left\|\tilde{x}_m - x_m\right\|$$

which holds for all $m \in \mathbb{N}$ and $t \in \mathbb{R}_+$ implies that $(T(t)x_m)_{m\in\mathbb{N}}$ and $(T(t)\tilde{x}_m)_{m\in\mathbb{N}}$ converge to the same limit for all $t \in \mathbb{R}_+$. Furthermore, the estimate

$$\|T(t)x\| = \lim_{m\to\infty}\|T(t)x_m\| \le \lim_{m\to\infty} Me^{t\omega}\left\|x_m\right\| = Me^{t\omega}\left\|x\right\|$$

holds for any $t \in \mathbb{R}_+$. Combined with Inequality 1.4.9, this yields

$$\|T(t)\| \le Me^{t\omega}.$$

Since for any $t \in \mathbb{R}_+$

$$\left(T(t) - \lim_{n\to\infty} T_n(t)\right)x = 0$$

for $x \in \mathcal{D}(A)$, $\mathcal{D}(A)$ is dense in X and

$$\sup_{n\in\mathbb{N}}\|T(t) - T_n(t)\| \le Me^{t\omega} + M\cdot\max\left\{e^{\frac{tn_0\omega}{n_0-\omega}}, 1\right\}$$

by Lemma 1.4.33

$$T(t)x = \lim_{n\to\infty} T_n(t)x$$

holds for any $x \in X$ and any $t \in \mathbb{R}_+$.

In order to show (ii) we already know that the family

$$(T(t))_{t\in\mathbb{R}_+}$$

as defined above is a family of linear bounded operators on X. We first show that this family also satisfies the Functional Equation (1.1.1). This is the case since for any $x \in X$ and any $s, t \in \mathbb{R}_+$ we obtain

$$\begin{aligned}
&\|T(t+s)x - T(t)T(s)x\| \\
&= \left\| \lim_{n\to\infty} T_n(s+t)x - \lim_{n\to\infty} T_n(t)T(s)x \right\| \\
&= \left\| \lim_{n\to\infty} T_n(s)T_n(t)x - \lim_{n\to\infty} T_n(t)T(s)x \right\| \\
&\leq \left(M \cdot \max\left\{ e^{\frac{sn_0\omega}{n_0-\omega}}, 1 \right\} \right) \lim_{n\to\infty} \|(T_n(t)x - T(t)x)\| \\
&= 0.
\end{aligned}$$

For statement (ii) it remains to be shown that $(T(t))_{t\in\mathbb{R}_+}$ is strongly continuous. We fix $x \in X$ and by Proposition 1.4.8 it is enough to show that

$$t \to T(t)x$$

is continuous from the right at $t = 0$. We know that

$$t \to T_n(t)x$$

is continuous for any $n \in \mathbb{N}$. If we can show that on some interval $[0, t_0]$ $t \to T_n(t)x$ converges uniformly to $t \to T(t)x$ we know from calculus that then also $t \to T(t)x$ is continuous on this interval. In order to show this uniform convergence we fix some $\varepsilon > 0$ and $t_0 \in \mathbb{R}_+$ and by Inequality 1.4.7 and Inequality 1.4.8 there exists $n_1(t_0, x)$ such that for all $n, m > n_1(t_0, x)$ and any $s \in [0, t_0]$

$$\begin{aligned}
\|T_n(s)x - T_m(s)x\| &\leq t_0 \cdot \left(M \cdot \max\left\{ e^{\frac{t_0 n_0\omega}{n_0-\omega}}, 1 \right\} \right)^2 \|A_n x - A_m x\| \\
&\leq \varepsilon.
\end{aligned}$$

We can find $n_0(t, x) > n_1(t_0, x)$ such that for all $n \geq n_0(t, x)$

$$\begin{aligned}
\|T(t)x - T_n(t)x\| &= \left\| \lim_{n\to\infty} T_n(t)x - T_n(t)x \right\| \\
&< \varepsilon.
\end{aligned}$$

$n_0(t, x)$ depends on t but we can eliminate this dependence combining both inequalities where we observe for $n > n_1(t_0, x)$

$$\begin{aligned}
\|T(t)x - T_n(t)x\| &\leq \left\| T(t)x - T_{n_0(t,x)}(t)x \right\| + \left\| T(t)_{n_0(t,x)}x - T_n(t)x \right\| \\
&\leq 2\varepsilon.
\end{aligned}$$

Hence we have shown uniform convergence of $t \to T_n(t)x$ to $t \to T(t)x$ on $[0, t_0]$ and therefore strong continuity of the semigroup

$$(T(t))_{t \in \mathbb{R}_+}.$$

For the proof of (iii) we need to show that $(B, \mathcal{D}(B))$, the generator of $(T(t))_{t\in\mathbb{R}_+}$, coincides with $(A, \mathcal{D}(A))$.
We start by showing that for $x \in \mathcal{D}(A)$

$$\left\| \lim_{t \searrow 0} \frac{T(t)x - x}{t} - Ax \right\| = 0, \tag{1.4.10}$$

hence $\mathcal{D}(A) \subset \mathcal{D}(B)$ and $B = A$ on $\mathcal{D}(A)$. Later we prove $\mathcal{D}(A) = \mathcal{D}(B)$. The proof of Equation 1.4.10 has two steps.
As a first step, we show that for $y \in X$ on some interval $[0, t_0]$

$$\dot{\xi}_n : t \to T_n(t)A_n y$$

converges uniformly to

$$\eta : t \to T(t)Ax.$$

This is shown by the estimate

$$\begin{aligned} \|T_n(t)A_n x - T(t)Ax\| \leq& \|T_n(t)A_n x - T_n(t)Ax\| + \|T_n(t)Ax - T(t)Ax\| \\ \leq& M \cdot \max\left\{ e^{\frac{t_0 n_0 \omega}{n_0 - \omega}}, 1 \right\} \|A_n x - Ax\| \\ &+ \|T_n(t)Ax - T(t)Ax\| \end{aligned}$$

and the uniform convergence of

$$\xi_n : t \to T_n(t)y$$

to

$$\xi : t \to T(t)y$$

known from the proof of (ii).
As a second step we show

$$T(t)x - x = \int_0^t T(s)Axds.$$

For some $\varepsilon > 0$ we choose n_0 such that for $0 \leq t \leq t_0$ and all $n > n_0$

$$\|T_n(t)A_n x - T(t)Ax\| \leq \varepsilon$$

and

$$\|T_n(t)x - T(t)x\| \leq \varepsilon.$$

By the fundamental theorem of calculus

$$T_n(t)x - x = \int_0^t T_n(t)A_n xds.$$

Therefore

$$\begin{aligned}\left\|-\int_0^t T(s)Axds + T(t)x - x\right\| \leq &\left\|-\int_0^t T(s)Axds + \int_0^t T_n(t)A_nxds\right\| \\ &+ \left\|-\int_0^t T_n(t)A_nxds + T_n(t)x - x\right\| \\ &+ \|-T_n(t)x + T(t)x\| \\ \leq &\varepsilon t + \varepsilon\end{aligned}$$

and since $\varepsilon > 0$ was arbitrary

$$T(t)x - x = \int_0^t T(s)Axds.$$

We can use this and the fundamental theorem of calculus and the strong continuity of $(T(t))_{t\in\mathbb{R}_+}$ to directly compute

$$\begin{aligned}\left\|\lim_{t\searrow 0}\frac{T(t)x - x}{t} - Ax\right\| &= \left\|\lim_{t\searrow 0}\frac{\int_0^t T(s)Axds}{t} - Ax\right\| \\ &= 0.\end{aligned}$$

We still need to show that $\mathcal{D}(A) = \mathcal{D}(B)$. By assumption of the theorem, for $\mathrm{Re}\lambda > \omega$ one has $\lambda \in \rho(A)$. The bound $\|T(t)\| \leq Me^{t\omega}$ from Equation 1.4.9 yields that by Proposition 1.4.29 $\mathrm{Re}\lambda > \omega$ implies also $\lambda \in \rho(B)$. Therefore, for $\mathrm{Re}\lambda > \omega$ the resolvent $R(\lambda, A)$ is a bijection between X and $\mathcal{D}(A)$ and the resolvent $R(\lambda, B)$ is a bijection between X and $\mathcal{D}(B)$. Hence $\lambda - B$ is a bijection between $\mathcal{D}(B)$ and X and $\lambda - A$ is a bijection between $\mathcal{D}(A)$ and X. For $y \in X$ it holds $R(\lambda, A)y \in \mathcal{D}(A)$ and since $\mathcal{D}(A) \subset \mathcal{D}(B)$ and $B = A$ on $\mathcal{D}(A)$

$$\begin{aligned}y &= (\lambda - A)\, R(\lambda, A)y \\ &= (\lambda - B)\, R(\lambda, A)y.\end{aligned}$$

Hence by injectivity of $(\lambda - B)$

$$R(\lambda, A)y = R(\lambda, B)y$$

for any $y \in X$. Surjectivity of $R(\lambda, B)$ onto $\mathcal{D}(B)$ and $R(\lambda, A)$ onto $\mathcal{D}(A)$ yields that this is only possible if $\mathcal{D}(B) = \mathcal{D}(A)$.

□

1.4.4. Lumer-Phillips Theorem. In the case of so-called *dissipative* operators the conditions the operator needs to fulfill in order to generate a strongly continuous semigroup can be somewhat relaxed. The presentation here follows Chapter II, Section 3b in [**18**].

DEFINITION 1.4.36. A linear operator$(A, \mathcal{D}(A))$ on a Banach space X is called *dissipative* if for all $\lambda > 0$ and all $x \in \mathcal{D}(A)$

$$\|(\lambda - A)\, x\| \geq \lambda \, \|x\| \, .$$

PROPOSITION 1.4.37. *For a dissipative operator*$(A, \mathcal{D}(A))$ *the following statements hold true:*
(i) $\lambda - A$ *is injective for all* $\lambda > 0$ *and for all* $z \in rg\,(\lambda - A)$ *and for all* $\lambda > 0$

$$\left\|(\lambda - A)^{-1} z\right\| \leq \frac{1}{\lambda} \, \|z\| \, . \tag{1.4.11}$$

(ii) If $\lambda - A$ *is surjective for some* $\lambda > 0$ *then it is surjective for all* $\lambda > 0$ *and* $(0, \infty) \subset \rho(A)$*.*
(iii) A *is closed if* $rg\,(\lambda - A)$ *is closed for some* $\lambda > 0$ *and if* A *is closed then* $rg\,(\lambda - A)$ *is closed for all* $\lambda > 0$*.*
(iv) If $rg(A) \subset \overline{\mathcal{D}(A)}$*, then* A *is closable and its closure* $\bar{A}$ *is dissipative as well. In this case, for all* $\lambda > 0$

$$rg\left(\lambda - \bar{A}\right) = \overline{rg\,(\lambda - A)}.$$

PROOF. (i) If for some $x, y \in \mathcal{D}(A)$

$$(\lambda - A)\, x = (\lambda - A)\, y,$$

then by definition of dissipativity

$$0 = \|(\lambda - A)\,(x - y)\| \geq \lambda \, \|(x - y)\| \geq 0$$

hence $x = y$. Inequality 1.4.11 follows directly from the definition of dissipativity for $z = (\lambda - A)\, x$.
(ii) Let $\lambda_0 - A$ be surjective for some $\lambda_0 > 0$. By (i) $\lambda_0 - A$ is also bijective hence $\lambda_0 \in \rho(A)$. By (i)

$$\|R(\lambda_0, A)\| \leq \frac{1}{\lambda_0}$$

thus Proposition 1.4.27 yields that for $\lambda \in \mathbb{R}$ such that $|\lambda - \lambda_0| < \lambda_0$ $\lambda \in \rho(A)$. Therefore $(0, 2\lambda_0) \in \rho(A)$ and by (i)

$$\|R(\lambda, A)\| \leq \frac{1}{\lambda}$$

for any $\lambda \in (0, 2\lambda_0)$. By applying Proposition 1.4.27 again we obtain inductively $(0, \infty) \in \rho(A)$.
(iii) Let $\operatorname{rg}(\lambda_0 - A)$ be closed for some $\lambda_0 > 0$. By (i)

$$(\lambda_0 - A)^{-1} : \operatorname{rg}(\lambda_0 - A) \to D(A)$$

exists and is a bounded linear operator. Therefore, by closedness of $\operatorname{rg}(\lambda_0 - A)$ if a sequence $(x_n)_{n\in\mathbb{N}}$ in $\operatorname{rg}(\lambda_0 - A)$ converges to some x then $x \in \operatorname{rg}(\lambda_0 - A)$ and by continuity of $(\lambda_0 - A)^{-1}$

$$\lim_{n\to\infty} (\lambda_0 - A)^{-1} x_n = (\lambda_0 - A)^{-1} x.$$

Thus $(\lambda_0 - A)^{-1}$ is a closed operator. This implies that $\lambda_0 - A$ is a closed operator since their respective graphs are identical. We still need to show closedness of A. For this end, let $(z_n)_{n\in\mathbb{N}}$ be a sequence in $D(A)$ that converges to some $z \in X$ as n tends to infinity and let $(Az_n)_{n\in\mathbb{N}}$ converge to some y in X as n tends to infinity. Then,$z \in D(A)$ by closedness of $\lambda_0 - A$ and

$$\begin{aligned}\lim_{n\to\infty} Az_n &= -\lim_{n\to\infty} (\lambda_0 z_n - Az_n) + \lim_{n\to\infty} (\lambda_0 z_n) \\ &= -(\lambda_0 z - Az) + \lambda_0 z \\ &= Az.\end{aligned}$$

Hence A is closed.
For the opposite direction, closedness of A implies closedness of $\lambda - A$ for any $\lambda > 0$ just as in the calculation above. However, this yields closedness of $(\lambda - A)^{-1}$ for any $\lambda > 0$ where by (i)

$$(\lambda - A)^{-1} : \operatorname{rg}(\lambda - A) \to D(A)$$

is a well defined bounded linear operator. By continuity of $(\lambda - A)^{-1}$, for any sequence $(\tilde{y}_n)_{n\in\mathbb{N}}$ in $\operatorname{rg}(\lambda - A)$ converging to some $\tilde{y}$ in X as n tends to infinity also $(\lambda - A)^{-1}\tilde{y}_n$ converges to $(\lambda - A)^{-1}\tilde{y}$ in X. Closedness of $(\lambda - A)^{-1}$ yields $\tilde{y} \in \operatorname{rg}(\lambda - A)$ hence $\operatorname{rg}(\lambda - A)$ is closed for any $\lambda > 0$.
(iv) In order to show closability of A it suffices by Lemma 1.4.23 to show that if a sequence $(x_n)_{n\in\mathbb{N}} \subset D(A)$ satisfies $\lim_{n\to\infty} x_n = 0$ and

$$\lim_{n\to\infty} Ax_n = y$$

then $y = 0$. We need to use some inequality in order to find bounds for y. We can use the one appearing in the definition of dissipativity and obtain for any $\lambda > 0$ and $\omega \in D(A)$ and $n \in \mathbb{N}$

$$\|\lambda(\lambda - A)x_n + (\lambda - A)\omega\| \geq \lambda \|\lambda x_n + \omega\|.$$

The idea is now to send x_n on the right hand side to 0 and ω to y in order to obtain an upper bound for $\|y\|$ which turns out to be 0. We obtain for $n \to \infty$

$$\left\| -y + \left(\frac{\lambda - A}{\lambda} \right) \omega \right\| \geq \|\omega\|$$

and for $\lambda \to \infty$

$$\|-y + \omega\| \geq \|\omega\| .$$

Now the assumption $\mathrm{rg}(A) \subset \overline{\mathcal{D}(A)}$ enters and we choose a sequence $(\omega_n)_{n \in \mathbb{N}} \subset D(A)$ such that

$$\lim_{n \to \infty} \omega_n = y.$$

Then the closedness follows from

$$0 = \lim_{n \to \infty} \|-y + \omega_n\| \geq \lim_{n \to \infty} \|\omega_n\| = \|y\| .$$

In order to show that $\bar{A}$ is dissipative we have to show that for all $\lambda > 0$ and all $x \in \mathcal{D}(\bar{A})$

$$\left\| \left(\lambda - \bar{A} \right) x \right\| \geq \lambda \|x\| .$$

By the definition of the closure of an operator there is a sequence $(x_n)_{n \in \mathbb{N}} \subset D(A)$ such that $\lim_{n \to \infty} x_n = x$ and

$$\lim_{n \to \infty} A x_n = \bar{A} x$$

and the dissipativity of A yields

$$\|(\lambda - A) x_n\| \geq \lambda \|x_n\|$$

and taking the limit on both sides we have shown dissipativity of $\bar{A}$.

For the last assertion of (iv) let $y \in \mathrm{rg}\left(\lambda - \bar{A}\right)$. Then

$$y = \lambda x - \bar{A} x$$

for some $x \in \mathcal{D}(\bar{A})$ and again by the definition of the closure of an operator there is a sequence $(x_n)_{n \in \mathbb{N}} \subset D(A)$ such that $\lim_{n \to \infty} x_n = x$ and

$$\lim_{n \to \infty} A x_n = \bar{A} x.$$

Hence

$$y = \lim_{n \to \infty} \lambda x_n - A x_n$$

and $\mathrm{rg}\,(\lambda - A)$ is dense in $\mathrm{rg}\left(\lambda - \bar{A}\right)$. Since by (iii) $\mathrm{rg}\left(\lambda - \bar{A}\right)$ is closed we obtain

$$\overline{\mathrm{rg}\,(\lambda - A)} = \mathrm{rg}\left(\lambda - \bar{A}\right) .$$

□

THEOREM 1.4.38. *For a densely defined dissipative operator* $(A, \mathcal{D}(A))$ *on a Banach space* X *the following statements are equivalent:*

(i) The closure $\bar{A}$ *generates a contraction semigroup.*

(ii) The image of $\lambda - A$ *is dense in* X *for some* $\lambda > 0$ *and thus for all* $\lambda > 0$.

PROOF. (i) $\Rightarrow$ (ii). By Theorem 1.4.29(ii) for all $\lambda > 0$ it holds $\lambda \in \rho(\bar{A})$. Thus $\operatorname{rg}\left(\lambda - \bar{A}\right) = X$ for all $\lambda > 0$. By Proposition 1.4.37 (iv)

$$\overline{\operatorname{rg}(\lambda - A)} = \operatorname{rg}\left(\lambda - \bar{A}\right)$$

for all $\lambda > 0$. Thus the image of $\lambda - A$ is dense in X for all $\lambda > 0$.

(ii)$\Rightarrow$(i). Let $\lambda_0 > 0$ be such that the image of $\lambda_0 - A$ is dense in X. By Proposition 1.4.37 (iv) A is closable with dissipative closure $\bar{A}$ and

$$X = \overline{\operatorname{rg}(\lambda_0 - A)} = \operatorname{rg}\left(\lambda_0 - \bar{A}\right).$$

Hence $\lambda_0 - \bar{A}$ is surjective and by Proposition 1.4.37(ii) $\lambda - \bar{A}$ is surjective for any $\lambda > 0$ and $(0, \infty) \subset \rho(A)$. By Proposition 1.4.37 (i) for all $\lambda > 0$

$$\left\| R\left(\lambda, \bar{A}\right) \right\| \leq \frac{1}{\lambda}.$$

We may thus use Hille-Yosida Theorem (Theorem 1.4.35) and conclude. □

1.4.5. Adjoint Semigroups. We consider the semigroup

$$(T(t))_{t \in \mathbb{R}_+}$$

on a Banach space X and construct another semigroup on its so-called *dual space* that consists of the so-called *adjoint* operators. First, we introduce adjoint operators and some of their properties.

DEFINITION 1.4.39. The *dual space* X' of a normed vector space X is defined as

$$X' := \{x' : X \to \mathbb{K} | \text{ linear and continuous}\}.$$

Its elements are called *functionals*.

REMARK 1.4.40. If X is a normed vector space then by Proposition 1.2.5 X' is a Banach space. Its norm is the usual norm of the space of linear bounded maps from Definition 1.2.3

$$\|x'\| := \inf\{M_{x'} \in \mathbb{R}_+ : |x'(x)| \leq M_{x'} \|x\| \text{ for all } x \in X\}. \tag{1.4.12}$$

The dual space $(X')'$ of a dual space X' is called *bidual* and is written X''.

DEFINITION 1.4.41. Let X, Y be normed vector spaces. To a (in general unbounded) densely defined linear operator

$$L: \mathcal{D}(L) \subset X \to Y$$

we can associate a map

$$L': \mathcal{D}(L') \subset Y' \to X'$$

on

$$\mathcal{D}(L') := \left\{ y' \in Y' : \begin{array}{c} \text{there exists } x' \in X' \text{ such that} \\ y'(Lx) = x'(x) \text{ for all } x \in \mathcal{D}(L) \end{array} \right\}$$

in the following way:

$$L'(y')(x) := y'(L(x)).$$

The operator L' is called *adjoint* operator.

We see that for $y' \in \mathcal{D}(L')$ the key requirement is that $y' \circ L$ is bounded on $\mathcal{D}(L)$ even though L is in general unbounded.

The adjoint operator possesses the following properties:

LEMMA 1.4.42. *Let X, Y be normed vector spaces.*

(i) For a densely defined linear operator $(L, \mathcal{D}(L))$ and its adjoint operator $(L', \mathcal{D}(L'))$ $\mathcal{D}(L')$ is a vector space and L' is linear.

(ii) If L is bounded then $\mathcal{D}(L') = Y'$.

(iii) If L is bounded then also L' is bounded and $\|L'\| = \|L\|$.

PROOF.
(i) If $y_1', y_2' \in \mathcal{D}(L')$ and $\mu \in \mathbb{C}$ then for all $x \in \mathcal{D}(L)$

$$\begin{aligned}(\mu y_1' + \mu y_2')(L(x)) &= \mu y_1'(L(x)) + \mu y_2'(L(x)) \\ &= \mu L'(y_1')(x) + \mu L'(y_2')(x)\end{aligned}$$

and since

$$\mu L'(y_1') + \mu L'(y_2') \in X'$$

$\mathcal{D}(L')$ is a vector space and

$$L'(\mu y_1' + \mu y_2') = \mu L'(y_1') + \mu L'(y_2').$$

(ii) If L is bounded then $L'y' := y' \circ L \in X'$ for all $y' \in Y'$.
(iii) By Definition 1.2.3 of the operator norm

$$\|L\| = \sup_{\|x\| \leq 1} \|Lx\| .$$

Moreover by Corollary A.4.8

$$\sup_{\|x\| \leq 1} \|Lx\| = \sup_{\|x\| \leq 1} \sup_{\|y'\| \leq 1} |y'(Lx)| .$$

Thus,

$$\begin{aligned} \|L\| &= \sup_{\|x\| \leq 1} \sup_{\|y'\| \leq 1} |y'(Lx)| \\ &= \sup_{\|y'\| \leq 1} \sup_{\|x\| \leq 1} |y'(Lx)| \\ &= \sup_{\|y'\| \leq 1} \|y' \circ L\| \\ &= \sup_{\|y'\| \leq 1} \|L'y'\| \\ &= \|L'\| . \end{aligned}$$

□

We can now pose the question we would like to solve in this section. If we are given a strongly continuous semigroup $(T(t))_{t\in\mathbb{R}_+}$ on a Banach space X we would like to know whether or not $(T'(t))_{t\in\mathbb{R}_+}$ also forms a strongly continuous semigroup. A first result is the following lemma.

LEMMA 1.4.43. *If $(T(t))_{t\in\mathbb{R}_+}$ is a semigroup on a Banach space X then $(T'(t))_{t\in\mathbb{R}_+}$ is a semigroup on X'.*

PROOF. Since we know from Lemma 1.4.42 that $(T'(t))_{t\in\mathbb{R}_+}$ is a family of bounded linear operators on the Banach space X' we only need to show that it satisfies the Functional Equation (1.1.1). For $s, t \in \mathbb{R}_+$ we obtain for all $y' \in Y'$

$$\begin{aligned} T'(s+t)(y')(x) :&= y'\left(T(s+t)(x)\right) \\ &= y'\left(T(s)T(t)(x)\right) \\ &= \left\{T'(s)(y')\right\}\left(T(t)(x)\right) \\ &= T'(t)T'(s)(y')(x). \end{aligned}$$

□

Therefore the following definition is justified:

DEFINITION 1.4.44. If $(T(t))_{t\in\mathbb{R}_+}$ is a semigroup on a Banach space X, the family$(T'(t))_{t\in\mathbb{R}_+}$ of bounded linear operators on X' is called the *adjoint semigroup*.

However, in the following example we see that the adjoint semigroup of a strongly continuous semigroup is not necessarily strongly continuous.

EXAMPLE 1.4.45. In Example 1.4.11 we have seen that the left translation semigroup

$$(T_l(t))_{t\in\mathbb{R}_+}$$

on $L^1(\mathbb{R})$ is strongly continuous. We want to find the adjoint operators of the left translation. It is known from functional analysis (see Proposition A.4.3) that the dual space of $L^1(\mathbb{R})$ is the space $L^\infty(\mathbb{R})$ from Example A.4.2. For the right continuous semigroup

$$(T_r(t))_{t\in\mathbb{R}_+}$$

from Example 1.4.1 it holds for $f \in L^1(\mathbb{R})$ and $g \in L^\infty(\mathbb{R})$ and $t \in \mathbb{R}_+$

$$\begin{aligned}\int_{-\infty}^{\infty} [T_r(t)(g)](s)\cdot f(s)\,ds &= \int_{-\infty}^{\infty} g(s-t)\cdot f(s)\,ds\\ &= \int_{-\infty}^{\infty} g(s)\cdot f(s+t))\,ds\\ &= \int_{-\infty}^{\infty} g(s)\cdot [T_l(t)f(s)]\ ds.\end{aligned}$$

Hence,

$$(T_r(t))_{t\in\mathbb{R}_+}$$

on $L^\infty(\mathbb{R})$ is the family of adjoint operators of

$$(T_l(t))_{t\in\mathbb{R}_+}$$

on $L^1(\mathbb{R})$. Since we know that the latter is strongly continuous we would like to know whether this is also the case for its family of adjoint operators. However, we see that for $\text{sgn} \in L^\infty(\mathbb{R})$ defined as

$$\text{sgn}(x) := \begin{cases} 1 \text{ if } x > 0\\ 0 \text{ if } x = 0\\ -1 \text{ if } x < 0\end{cases}$$

we obtain for $t > 0$

$$\|T_r(t)\text{sgn} - \text{sgn}\|_{L^\infty(\mathbb{R})} = 1.$$

Thus, the adjoint semigroup is not strongly continuous.

Since the adjoint semigroup is in general not strongly continuous but we would still like to show some regularity a first idea is to ask for less than strong continuity. Therefore, we introduce*locally convex vector spaces* and *weak topologies* and hereby follow [**41**], chapter VIII .

DEFINITION 1.4.46. Let X be a $\mathbb{K}$-vector space. A map

$$p:\, X \to [0,\infty)$$

is called *seminorm* on X if
(i) for all $\lambda \in \mathbb{K}$ and $x \in X$ $p\,(\lambda x) = |\lambda|\, p(x)$
(ii) for all $x, y \in X p(x+y) \leq p(y) + p(x)$

REMARK 1.4.47. A seminorm for which holds that $p(x) = 0$ implies $x = 0$ is a norm.

Just as in example A.1.3 where a norm induced a topology also a family P of seminorms on X induces a topology. This is seen as follows. For a finite subset $F \subset P$ and $\varepsilon > 0$ we define

$$U_{F,\varepsilon} = \{x \in X \,:\, p(x) \leq \varepsilon \,\text{for all}\; p \in F\} \tag{1.4.13}$$

and the set of all such sets

$$\mathfrak{U} := \{U_{F,\varepsilon} \,:\, F \subset P \text{ finite}, \varepsilon > 0\}\,.$$

$\mathfrak{U}$ replaces the set of open balls in the case of normed spaces.

PROPOSITION 1.4.48. *Let X be a $\mathbb{K}$-vector space, let P be a set of seminorms on X, and let $\mathfrak{U}$ be defined as above. Then the family of subsets*

$$\tau := \{O \subset X \,:\, \textit{for any } x \in O \textit{ there is } U \in \mathfrak{U} \textit{ such that } x + U \subset O\}$$

is a topology on X.

PROOF. Concerning property (i) of a topology (in Definition A.1.1), clearly $\emptyset, X \in \tau$.
Concerning (ii), if $O_1, O_2 \in \tau$ then for $x \in O_1 \cap O_2$ there exists $F_1 \subset P$ finite, $\varepsilon_1 > 0$ such that $x + U_{F_1,\varepsilon_1} \subset O_1$ and there exists $F_2 \subset P$ finite, $\varepsilon_2 > 0$ such that $x + U_{F_2,\varepsilon_2} \subset O_2$. Then $x + U_{F_1 \cup F_2, \min(\varepsilon_1,\varepsilon_2)} \subset O_1 \cap O_2$.
Concerning (iii), if I is some index set and $O_i \in \tau$ for all $i \in I$ and $x \in \bigcup_{i\in I} O_i$ then $x \in O_j$ for some $j \in I$ and by definition there exists $F_j \subset P$ finite, $\varepsilon_j > 0$ such that $x + U_{F_j,\varepsilon_j} \subset O_j \subset \bigcup_{i \in I} O_i$. □

DEFINITION 1.4.49. Let X be a $\mathbb{K}$-vector space and P be a set of seminorms on X and τ the topology from Proposition 1.4.48. Then(X,τ) is called *locally convex topological vector space*.

Just as in the case of a normed vector space we can define dual spaces:

DEFINITION 1.4.50. The *dual space* X' of a locally convex topological vector space X is defined as

$$X' := \{L:\ X \to \mathbb{K}|\ \text{linear and continuous}\}.$$

DEFINITION 1.4.51. Let X, Y be $\mathbb{K}$-vector spaces and $\langle\cdot,\cdot\rangle$:

$$\begin{aligned} X \times Y &\to \mathbb{K} \\ (x,y) &\to \langle x,y\rangle \end{aligned}$$

a bilinear map. $(X, Y, \langle\cdot,\cdot\rangle)$ is called *dual pair* if for all $x \in X \setminus \{0\}$ there is $y \in Y$ such that $\langle x,y\rangle \neq 0$ and for all $y \in Y \setminus \{0\}$ there is $x \in X$ such that $\langle x,y\rangle \neq 0$.

LEMMA 1.4.52. *Let $(X, Y, \langle\cdot,\cdot\rangle)$ be a dual pair. For $y \in Y$ define* p_y :

$$\begin{aligned} X &\to [0,\infty) \\ x &\to |\langle x,y\rangle|. \end{aligned}$$

Then p_y is a seminorm on X.

PROOF. For $\lambda \in \mathbb{K}$ clearly $|\langle \lambda x,y\rangle| = |\lambda|\,|\langle x,y\rangle|$ and for $a,b \in X$ $|\langle a+b,y\rangle| = |\langle a,y\rangle + \langle b,y\rangle| \leq |\langle a,y\rangle| + |\langle b,y\rangle|$. □

DEFINITION 1.4.53. Let $(X, Y, \langle\cdot,\cdot\rangle)$ be a dual pair and

$$P := \{p_y:\ y \in Y\}$$

a family of seminorms on X. The topology that is induced by P on X via Proposition 1.4.48 is called $\sigma(X,Y)$-topology.

REMARK 1.4.54. Let $(X, Y, \langle\cdot,\cdot\rangle)$ be a dual pair and $(x_n)_{n\in\mathbb{N}} \subset X$ and $x \in X$. If for all $y \in Y$ $\lim\limits_{n\to\infty} \langle x_n,y\rangle = \langle x,y\rangle$ then $(x_n)_{n\in\mathbb{N}}$ converges to x in the $\sigma(X,Y)$-topology.

REMARK 1.4.55. For a locally convex topological vector space X and its dual space X' the map $\langle \cdot, \cdot \rangle$:

$$\begin{aligned} X \times X' &\to \mathbb{K} \\ (x, x') &\to x'(x) \end{aligned}$$

is bilinear.

PROPOSITION 1.4.56. *Let X be a Banach space. Then$(X, X', \langle \cdot, \cdot \rangle)$ is a dual pair.*

PROOF. If $x' \neq 0$ then there is some $x \in X$ such that $x'(x) \neq 0$. If $x \neq 0$ then by a corollary of the Hahn-Banach theorem (Corollary A.4.7) there exists $x' \in X'$ such that $x'(x) \neq 0$. □

Therefore, for $x \in X$ and $x' \in X'$ we can write $\langle x, x' \rangle = \langle x', x \rangle = x'(x)$.

REMARK 1.4.57. It can also be shown that$(X, X', \langle \cdot, \cdot \rangle)$ is a dual pair when X is only a locally convex topological vector space (see [**41**], Chapter VIII.3, Example (a)). For this purpose however one first needs to introduce the Hahn-Banach theorem for locally convex topological vector spaces (see [**41**], Chapter VIII.2), something we will omit.

DEFINITION 1.4.58. The topology $\sigma(X, X')$ is called *weak topology*, the topology $\sigma(X', X)$ is called *weak***- topology* .

Turning back to our question whether the adjoint semigroup is continuous in some sense we obtain the following result.

PROPOSITION 1.4.59. *If $(T(t))_{t \in \mathbb{R}_+}$ is a strongly continuous semigroup on X then for the adjoint semigroup $(T'(t))_{t \in \mathbb{R}_+}$*

$$t \to T'(t)y'(x)$$

is continuous for any $x \in X$ and any $y' \in X'$.

PROOF. Since

$$T'(t)y'(x) = (y' \circ T(t))(x)$$

it follows from continuity of y' and strong continuity of $t \to T(t)$ that

$$\begin{aligned} \lim_{t \to s} T'(t)y'(x) &= y' \circ \left(\lim_{t \to s} T(t)x \right) \\ &= y' \circ T(s)x \\ &= T'(s)y'(x). \end{aligned}$$

□

LEMMA 1.4.60. *The map*

$$t \to T'(t)y'(x)$$

is continuous for any $t \in \mathbb{R}_+$, $x \in X$ *and* $y' \in X'$ *if and only if for any* $y' \in X'$ *the map* $\xi'_{y'}$:

$$\begin{aligned} \mathbb{R}_+ &\to X' \\ t &\to T'(t)y' \end{aligned}$$

is continuous where X' *is equipped with the* $\sigma(X', X)$ *-topology. Therefore, this type of continuity is called weak***-continuity.*

PROOF. Fix $y' \in X'$. If $\mathbb{R}_+ \ni t \to T'(t)y'(x) \in \mathbb{K}$ is continuous for any $x \in X$ then also $t \to (T'(t)y' - z')(x)$ for any $z' \in X'$ and any $x \in X$. Thus, for any $\varepsilon > 0$ and $x \in X$ and the sets $U_{\{|\langle \cdot, x\rangle|\},\varepsilon}$ defined in 1.4.13 the set

$$\left(\xi'_{y'}\right)^{-1}\left(U_{\{|\langle \cdot, x\rangle|\},\varepsilon} + z'\right)$$

is open. Hence, for any finite set $F \subset X$ and any $\varepsilon > 0$ the set

$$\left(\xi'_{y'}\right)^{-1}\left(U_{\{|\langle \cdot, x\rangle| : x \in F\},\varepsilon} + z'\right) = \bigcap_{x \in F} \left(\xi'_{y'}\right)^{-1}\left(U_{\{|\langle \cdot, x\rangle|\},\varepsilon} + z'\right)$$

is open. Thus, $\xi'_{y'}$ is continuous when X' is equipped with the $\sigma(X', X)$ topology.

For the other direction, let $\varepsilon > 0$ and $x \in X$ be arbitrary. Then $\left(\xi'_{y'}\right)^{-1}\left(U_{\{|\langle \cdot, x\rangle|\},\varepsilon} + z'\right)$ is open for any $z' \in X'$ and for any $\varepsilon > 0$. For any $a \in \mathbb{K}$ and $x \in X \setminus \{0\}$ by Corollary A.4.7 there exists $a' \in X$ such that $a'(x) = a$. Thus, for any $\delta > 0$ and the δ-ball $B_\delta(a)$ around $a \in \mathbb{K}$ the sets

$$(T'(t)y' - a')(x) \in B_\delta(0)$$

and

$$T'(t)y'(x) \in B_\delta(a)$$

are the same and open. Hence, the map

$$t \to T'(t)y'(x)$$

is continuous for any $x \in X$. □

However, instead of using a weaker concept of continuity we can also restrict ourselves to a smaller space where we obtain even strong continuity. The exact result is the proposition below. The proof follows the one in **[20]**, Section 4.1.1.

PROPOSITION 1.4.61.
Let

$$(T(t))_{t\in\mathbb{R}_+}$$

be a strongly continuous semigroup on X *with generator* $(A, \mathcal{D}(A))$ *and* A' *the adjoint operator of* A *with domain* $\mathcal{D}(A')$. *Then*
i) for $X^\dagger := \overline{\mathcal{D}(A')} \subset X'$ *where the closure is to be taken with respect to the norm topology of* X' *and*

$$T^\dagger(t) := T'(t)|_{X^\dagger} \quad \textit{for all } t \in \mathbb{R}_+$$

the family

$$\left(T^\dagger(t)\right)_{t\in\mathbb{R}_+}$$

is a strongly continuous semigroup on $X^\dagger$ *and*
ii) its generator $A^\dagger$ *is given by the restriction of* A' *to the set*

$$\mathcal{D}(A^\dagger) := \left\{y' \in \mathcal{D}(A') :\, A'y' \in X^\dagger\right\}.$$

PROOF.
i) The proof has two steps.
First, we have to show that

$$\left(T^\dagger(t)\right)_{t\in\mathbb{R}_+}$$

is indeed a family of maps between the correct spaces, that is $T^\dagger(t) \in L(X^\dagger)$ for all $t \in \mathbb{R}_+$. In order to show this claim, we first let $y' \in \mathcal{D}(A')$, fix some $t \in \mathbb{R}_+$ and show that $T^\dagger(t)y' \in \mathcal{D}(A')$. In other words, we have to prove that there exits $A'\left(T^\dagger(t)y'\right) \in X'$ such that for all $x \in \mathcal{D}(A)$

$$\left(T^\dagger(t)y'\right)(Ax) = A'\left(T^\dagger(t)y'\right)x.$$

Since by Lemma 1.4.42 $T(t) \in L(X)$ implies $\mathcal{D}(T') = X'$ we obtain $T^\dagger(t)y' \in X'$ and $T^\dagger(t)y' = y' \circ T(t)$ and the equation we need to show reduces to

$$y' \circ T(t) \circ (Ax) = A' \circ y' \circ T(t)x$$

for all $x \in \mathcal{D}(A)$. Also $y' \in \mathcal{D}(A')$ implies $A'(y') \in X'$ and $A'(y')x = y' \circ Ax$ for all $x \in \mathcal{D}(A)$. By the invariance of $\mathcal{D}(A)$ under $T(t)$ (see Proposition 1.4.15) also $T(t)x \in \mathcal{D}(A)$ and the equation we need to show reduces to

$$y' \circ T(t) \circ (Ax) = y' \circ A \circ T(t)x.$$

This holds true due to Proposition 1.4.15. Furthermore, the equation

$$A'\left(T^{\dagger}(t)y'\right) = \underbrace{A' \circ y'}_{\in X'} \circ T(t)$$

yields $A'\left(T^{\dagger}(t)y'\right) \in X'$. Hence $T^{\dagger}(t)y' \in \mathcal{D}(A')$ for $y' \in \mathcal{D}(A')$.
We still need to show that $T^{\dagger}(t)y' \in X^{\dagger}$ for any $y' \in X^{\dagger} \setminus \mathcal{D}(A')$. We find a sequence $(y'_n)_{n\in\mathbb{N}} \subset \mathcal{D}(A')$ such that $\lim_{n\to\infty} y'_n = y'$ in X'. Since by Lemma 1.4.42 $T^{\dagger}$ is linear and bounded it is continuous on X'. Therefore

$$\begin{aligned} T^{\dagger}(t)\,(y') &= T^{\dagger}(t)\left(\lim_{n\to\infty} y'_n\right) \\ &= \lim_{n\to\infty} \underbrace{T^{\dagger}(t)\,(y'_n)}_{\in\mathcal{D}(A')} \end{aligned}$$

lies in $X^{\dagger}$.
In the second step, since we know from Lemma 1.4.43 that the functional equation holds for $\left(T^{\dagger}(t)\right)_{t\in\mathbb{R}_+}$, by the first part of the proof it is a semigroup on $X^{\dagger}$ and all we need to show is strong continuity on $X^{\dagger}$. Since by assumption $(T(t))_{t\in\mathbb{R}_+}$ is a strongly continuous semigroup, by Proposition 1.4.8 there is $\delta > 0$ and $M \geq 1$ such that $\|T(t)\| \leq M$ for $0 \leq t \leq \delta$. By Lemma 1.4.42 $\left\|T^{\dagger}(t)\right\| \leq \|T(t)\|$. Therefore, there is $\delta > 0$ and $M \geq 1$ such that $\left\|T^{\dagger}(t)\right\| \leq M$ for $0 \leq t \leq \delta$ and by Proposition 1.4.8 in order to show strong continuity of $\left(T^{\dagger}(t)\right)_{t\in\mathbb{R}_+}$ it is enough to show

$$\lim_{t\searrow 0} T^{\dagger}(t)y' = y'$$

on a dense subset of $X^{\dagger}$. As such a subset we take $\mathcal{D}(A')$. We need to find a suitable expression for $T^{\dagger}(t)y' = y' \circ T(t) \in X'$ and observe that it follows from Proposition 1.4.15 that for $x \in X$

$$T(t)x - x = \int_0^t AT(s)x ds.$$

We apply y' on both sides of the equation and remember from Equation 1.3.2 that we can pull linear maps in the Banach space valued integral. We obtain

$$(y' \circ T(t))\,(x) - y'\,(x) = \int_0^t (y' \circ A)\,T(s)(x)ds.$$

and since $y' \in \mathcal{D}(A')$ by the definition of adjoint operators $A'\,(y') \in X'$

$$\int_0^t (y' \circ A)\,T(s)(x)ds. = \int_0^t (A'\,(y')) \circ (T(s))\,(x)ds$$

Since this equation holds for all $x \in X$ we have shown

$$T^{\dagger}(t)y' - y' = \int_0^t \left(A'\left(y'\right)\right) \circ \left(T(s)\right) ds.$$

We can therefore bound for $0 \leq t \leq \delta$

$$\left\|T^{\dagger}(t)y' - y'\right\| \leq \int_0^t \left\|A'\left(y'\right)\right\| M\, ds.$$

With $t \to 0$ the right hand side converges to zero and we conclude.

ii) First, we show that the limit

$$\lim_{t \searrow 0} \frac{T^{\dagger}(t)y' - y'}{t} = \lim_{t \searrow 0} \frac{1}{t} \int_0^t \left(y' \circ A\right) T(s) ds.$$

exists for all $y' \in \mathcal{D}(A^{\dagger})$ and that this limit is $A'\left(y'\right)$. Since

$$\lim_{t \searrow 0} \frac{1}{t} \int_0^t \left(y' \circ A\right) T(s) ds = \lim_{t \searrow 0} \frac{1}{t} \int_0^t T^{\dagger}(s) \left(A'\left(y'\right)\right) ds$$

and the fact that

$$\left(T^{\dagger}(t)\right)_{t \in \mathbb{R}_+}$$

is a strongly continuous semigroup on $X^{\dagger}$ the right hand side converges by the fundamental theorem of calculus to $A'\left(y'\right)$.
Next, we want to show that the domain of the generator is not larger than

$$\left\{y' \in \mathcal{D}(A') :\ A'y' \in X^{\dagger}\right\}.$$

In order to determine this domain, we use the fact that for any $\lambda \in \rho(A^{\dagger})$ by its definition, $R(\lambda, A^{\dagger})$ is a bijection between the domain of the generator $A^{\dagger}$ and $X^{\dagger}$. Therefore, we would like to find an expression of $R(\lambda, A^{\dagger})$ in terms of A' without dependence on $A^{\dagger}$ in order to express the domain of $A^{\dagger}$ in terms of A'. We recall that by Lemma 1.4.42 and Proposition 1.4.9 for all $t \in \mathbb{R}_+$

$$\left\|T^{\dagger}(t)\right\| \leq \left\|T'(t)\right\| = \left\|T(t)\right\| \leq M e^{\omega t}$$

for some $M \geq 1$ and $\omega \in \mathbb{R}$. Hence, by Theorem 1.4.29 if $\mathrm{Re}\lambda > \omega$ then $\lambda \in \rho(A^{\dagger})$ and for all $x \in X^{\dagger}$

$$\begin{aligned} R(\lambda, A^{\dagger})x &= \int_0^{\infty} e^{-\lambda s} T^{\dagger}(s) x ds \\ &= \int_0^{\infty} e^{-\lambda s} T'(s) x ds. \end{aligned}$$

We want to express the right hand side by an operator depending on A'. We have

$$R(\lambda, A) = \int_0^\infty e^{-\lambda s} T(s) ds.$$

Since by definition (see Definition 1.4.25) it is a bounded operator for $\mathrm{Re}\lambda > \omega$ its transpose

$$R(\lambda, A)' = \left(\int_0^\infty e^{-\lambda s} T(s) ds \right)'$$

is defined on all of X' (by Lemma 1.4.42). The integral is defined as the limit of Riemann sums and transposing an operator is a continuous operation according to Lemma 1.4.42. Thus, we can pull the transpose in the integral and obtain

$$R(\lambda, A)' = \int_0^\infty e^{-\lambda s} T'(s) ds$$

and

$$R(\lambda, A^\dagger) = R(\lambda, A)'|_{X^\dagger} .$$

In order to simplify the right hand side further, we observe that

$$R(\lambda, A)' = (\lambda - A')^{-1}$$

holds since

$$x' \circ (\lambda - A)^{-1} (\lambda - A) = x' \in X'$$

for any $x' \in X'$, hence $x' \circ (\lambda - A)^{-1} \in \mathcal{D}(\lambda - A')$ and

$$\begin{aligned} (\lambda - A') (R(\lambda, A)'(x')) &= (\lambda - A') \left(x' \circ (\lambda - A)^{-1} \right) \\ &= x'. \end{aligned}$$

Moreover, for $x' \in \mathcal{D}(A')$

$$\begin{aligned} R(\lambda, A)'((\lambda - A') x') &= x' \circ (\lambda - A) \circ (\lambda - A)^{-1} \\ &= x'. \end{aligned}$$

Thus, we can rewrite the resolvent $R(\lambda, A^\dagger)$ as

$$\begin{aligned} R(\lambda, A^\dagger) &= \int_0^\infty e^{-\lambda s} T'(s) ds \Big|_{X^\dagger} \\ &= R(\lambda, A)'|_{X^\dagger} \\ &= R(\lambda, A')|_{X^\dagger} . \end{aligned}$$

$R(\lambda, A^\dagger)$ is a bijection between $X^\dagger$ and the space where the generator $A^\dagger$ is defined. Moreover,

$$R(\lambda, A')|_{X^\dagger} (X^\dagger) = \left\{ y' \in \mathcal{D}(A') : \, A'y' \in X^\dagger \right\}$$

by definition of $R(\lambda, A')$ and the operator $A^\dagger$ can be defined only on

$$\left\{y' \in \mathcal{D}(A') : A'y' \in X^\dagger\right\}.$$

. □

A result concerning the size of the space $X^\dagger$ from Proposition 1.4.61 is the following assertion. Like in the case of the previous proposition, we follow the proof of [**20**], section 4.1.1.

PROPOSITION 1.4.62. *Let*

$$\{T(t)\}_{t\in\mathbb{R}_+}$$

be a strongly continuous semigroup on X *with generator* $(A, \mathcal{D}(A))$ *and* A' *the adjoint operator of* A *with domain* $\mathcal{D}(A')$*. Then* $X^\dagger = \overline{\mathcal{D}(A')} \subset X'$ *is dense in* X' *in the weak-*- topology which means that for each* $y' \in X'$ *there exists a sequence* $(y'_n)_{n\in\mathbb{N}} \subset X^\dagger$ *such that*

$$\lim_{n\to\infty} y'_n(x) = y'(x)$$

for each $x \in X$.

PROOF. Fix $y' \in X'$. We have to find some good approximation for y' in $X^\dagger$. By Theorem 1.4.29 (ii) in conjunction with Proposition 1.4.9 $R(n, A)$ exists and is bounded (see Definition 1.4.25) for $n \in \mathbb{N}$ large enough, say $n > N$. Choosing an approximation of y' known from Proposition 1.4.34, we set

$$y'_n := ny'R(n, A)$$

for any $n \in \mathbb{N}$, $n > N$. By Proposition 1.4.34 we know that

$$\lim_{n\to\infty} nR(n, A)(x) \to x$$

for all $x \in X$ hence by continuity of y'

$$\begin{aligned}\lim_{n\to\infty} ny'\left(R(n, A)(x)\right) &= y'\left(\lim_{n\to\infty} nR(n, A)(x)\right)\\ &= y'\left(x\right).\end{aligned}$$

We would like to show that

$$(y'_n)_{n\in\mathbb{N}, n>N.} \subset \mathcal{D}(A') \subset X^\dagger$$

in order to conclude. All we need to show is that $y'_n \circ A \in X'$ for any $n \in \mathbb{N}$, $n > N$. This follows from

$$\begin{aligned}ny'R(n, A)A &= ny'R(n, A)\left(A - n\right) + ny'R(n, A)\left(n\right)\\ &= -ny' + n^2y'R(n, A)\end{aligned}$$

and the fact that $R(n, A)$ is bounded for $n \in \mathbb{N}$, $n > N$.

□

1.4.6. Weakly continuous semigroups. In this subsection we see that so-called *weakly continuous semigroups* (see Definition 1.4.64) are also strongly continuous. In the proof we follow [**18**], chapter I, Theorem 5.8 and need to cite several results from functional analysis.

REMARK 1.4.63. The map $f_{x,x'}$:

$$\begin{aligned} \mathbb{R}_+ &\to \mathbb{C} \\ t &\to \langle T(t)x, x' \rangle \end{aligned}$$

is continuous for any $x \in X$ and any $x' \in X'$ if and only if for any $x \in X$ the map ξ_x :

$$\begin{aligned} \mathbb{R}_+ &\to X \\ t &\to T(t)x \end{aligned}$$

is continuous where X is equipped with the $\sigma(X, X')$ topology.

PROOF. The proof is almost identical to the one of Lemma 1.4.60.
□

DEFINITION 1.4.64. In case the conditions of Lemma 1.4.63 holds true, we call the semigroup

$$(T(t))_{t \in \mathbb{R}_+}$$

weakly continuous.

LEMMA 1.4.65. *For a semigroup*

$$(T(t))_{t \in \mathbb{R}_+}$$

on a Banach space X such that for any $x \in X$ and $x' \in X'$ the map $f_{x,x'}$:

$$\begin{aligned} \mathbb{R}_+ &\to \mathbb{C} \\ t &\to \langle T(t)x, x' \rangle \end{aligned}$$

is continuous, it holds

$$\sup_{t \in [0,s]} \|T(t)\| < \infty$$

for any $s \in \mathbb{R}_+$.

PROOF. It is enough to show that there is some $\delta > 0$ such that

$$\sup_{t\in[0,\delta]} \|T(t)\| < \infty$$

because then by the semigroup property

$$\sup_{t\in[0,s]} \|T(t)\| < \infty$$

holds true for any $s \in \mathbb{R}_+$.

We show this by contradiction and assume that

$$\sup_{t\in[0,\delta]} \|T(t)\| = \infty$$

for any $\delta > 0$. Then there exists a sequence $(t_n)_{n\in\mathbb{N}}$ such that $t_n \searrow 0$ and

$$\lim_{t_n\to\infty} \|T(t_n)\| = \infty.$$

By the uniform boundedness principle (Theorem 1.4.6) there exists $y \in X$ such that

$$\limsup_{t_n\to 0} \|T(t_n)y\| = \infty.$$

Interpreting $(T(t_n)y)_{n\in\mathbb{N}}$ as a family of maps

$$\begin{aligned} X' &\to \mathbb{C} \\ x' &\to \langle T(t_n)y, x'\rangle \end{aligned}$$

we obtain that if for all $x' \in X'$ by continuity of $f_{y,x'}$ at $t = 0$

$$\limsup_{t_n\to 0} \langle T(t_n)y, x'\rangle < \infty$$

holds true. Then, by the uniform boundedness principle (Theorem 1.4.6)

$$\limsup_{t_n\to 0} \sup_{\substack{x' \in X \\ \|x'\| \le 1}}{}' \langle T(t_n)y, x'\rangle < \infty.$$

However, by a corollary of the Hahn-Banach theorem (Corollary A.4.8)

$$\sup_{\substack{x' \in X \\ \|x'\| \le 1}}{}' \langle T(t_n)y, x'\rangle = \|T(t_n)y\|,$$

which yields the contradiction

$$\infty > \limsup_{t_n\to 0} \sup_{\substack{x' \in X \\ \|x'\| \le 1}}{}' \langle T(t_n)y, x'\rangle = \limsup_{t_n\to 0} \|T(t_n)y\| = \infty.$$

□

In the proof of the next proposition we need to work with the so-called *convex hull* of a set.

DEFINITION 1.4.66. Let X be a $\mathbb{K}$ vector space and $M \subset X$ some subset. The *convex hull* of M is written co$M$ and defined as

$$\mathrm{co}M := \left\{ \sum_{i=0}^{n} \lambda_i m_i \middle| \begin{array}{c} n \in \mathbb{N},\, \lambda_i \in \mathbb{K}\, \sum\limits_{i=0}^{n} \lambda_i = 1,\, \lambda_i \geq 0, \\ m_i \in M \text{ for all } i \in \{0, ..., n\}\,. \end{array} \right\}$$

Its closure is written as $\overline{\mathrm{co}}M$.

In the following proof we need a different definition of an integral of Banach space valued functions than the one we have been using previously in Definition 1.4.67. We introduce the following definition from Rudin ([**38**], Definition 3.26):

DEFINITION 1.4.67. Let (Q, Σ, μ) be a measure space and let X be a Banach space. Let $f : Q \to X$ be a function such that for any $x' \in X'$ the function $x'(f) : Q \to \mathbb{K}$ is integrable with respect to μ. We write

$$\int_Q f d\mu = y$$

for $y \in X$ if for any $x' \in X'$

$$\int_Q x'(f)\, d\mu = x'(y)\,.$$

We are now able to show the main result of this subsection which is taken from Engel, Nagel ([**18**]), Chapter I, Theorem 5.8. Our proof follows the one presented there.

THEOREM 1.4.68. *A semigroup*

$$(T(t))_{t \in \mathbb{R}_+}$$

on a Banach space X is strongly continuous if and only if for any $x \in X$ and any $x' \in X'$ the map $f_{x,x'}$:

$$\begin{aligned} \mathbb{R}_+ &\to \mathbb{C} \\ t &\to \langle T(t)x, x' \rangle \end{aligned}$$

is continuous.

PROOF. We only have to show that weak continuity of

$$(T(t))_{t \in \mathbb{R}_+}$$

implies strong continuity of the semigroup. We use Proposition 1.4.8(iii). Condition (b) in Proposition 1.4.8(iii) has already been shown in Lemma 1.4.65. We still have to show that the set

$$E := \left\{ x : \lim_{t \to 0} \|T(t)x - x\| = 0 \right\}$$

is dense in X with respect to the topology induced by the norm. The outline of the proof is as follows. Since E is convex (see Definition A.3.71) it suffices by Proposition A.4.10 to show that some subset $D \subset E$ is dense in X with respect to the weak topology. As a first step, we find a set $D \subset X''$ such that X is contained in its weak closure. As a second step, we show $D \subset X$. This implies that D is weakly dense in X. Finally, we show that D is a subset of E and conclude.
In order to find such a weakly dense set, in the first step we fix $x \in X$ and $r > 0$ and we define for $x' \in X'$

$$\langle x_r, x' \rangle := \frac{1}{r} \int_0^r \langle T(s)x, x' \rangle \, ds.$$

The map $x' \to \langle x_r, x' \rangle$ is linear and due to

$$|\langle x_r, x' \rangle| \le \left(\sup_{s \in [0,r]} \|T(s)\| \right) \|x\| \, \|x'\| \, ds$$

by Proposition 1.2.7 continuous. Hence $x_r \in X''$. Because of continuity of $s \to \langle T(s)x - x, x' \rangle$ we obtain

$$\begin{aligned} \lim_{r \to 0} \frac{1}{r} \int_0^r \langle T(s)x, x' \rangle \, ds - \langle x, x' \rangle &= \lim_{r \to 0} \frac{1}{r} \int_0^r \langle T(s)x - x, x' \rangle \, ds \\ &= 0. \end{aligned}$$

Hence x_r converges in $\sigma(X'', X')$ to x as r tends to 0 and we define the set

$$D := \{ x_r \in X'' : \, x \in X, \, r > 0 \}.$$

As a second step, we show $D \subset X$. For this purpose, we can use Proposition A.4.11 on Banach space valued integration (according to Definition 1.4.67). Interpreting

$$\begin{aligned} [0, r] &\to X \\ s &\to T(s)x \end{aligned}$$

as a map that is continuous when X is equipped with the weak topology, this proposition yields that

$$x_r \in \overline{co} \{ T(s) \, x : \, s \in [0, r] \} \subset X$$

if

$$\overline{co}\left\{T\left(s\right)x:\ s\in\left[0,r\right]\right\}$$

is compact in X in the weak topology. In order to show such compactness we observe that since $s\to T(s)x$ is continuous, when X is equipped with the weak topology the image

$$\left\{T\left(s\right)x:\ s\in\left[0,r\right]\right\}$$

of the compact set $[0,r]$ is weakly compact. Then the Krein-Šmulian weak compactness theorem (Theorem A.4.12) states that the closed convex hull of a weakly compact set is also weakly compact. Hence

$$\overline{co}\left\{T\left(s\right)x:\ s\in\left[0,r\right]\right\}$$

is weakly compact and applying Proposition A.4.11 is justified and yields $D\subset X$ since

$$D\subset\overline{co}\left\{T\left(s\right)x:\ s\in\left[0,r\right]\right\}\subset X.$$

It remains to be shown that $D\subset E$. For any $x_r\in D$ by definition Corollary A.4.8 of Hahn-Banach

$$\lim_{t\to0}\|T(t)x_r-x_r\|=\lim_{t\to0}\sup_{\substack{x'\in X'\\ \|x'\|\le1}}\left|\langle T(t)x_r,x'\rangle-\langle x_r,x'\rangle\right|.$$

Thus, by definition of the adjoint operator (see Definition A.3.71) $(T(t))'$ of $T(t)$ and Lemma 1.4.42 (iii)

$$\lim_{t\to0}\|T(t)x_r-x_r\|=\lim_{t\to0}\sup_{\substack{x'\in X'\\ \|x'\|\le1}}\left|\left\langle x_r,\left(T(t)\right)'x'\right\rangle-\langle x_r,x'\rangle\right|.$$

Therefore, the definition of x_r yields

$$\begin{aligned}&\lim_{t\to0}\|T(t)x_r-x_r\|\\ &=\lim_{t\to0}\sup_{\substack{x'\in X'\\ \|x'\|\le1}}\left|\frac{1}{r}\int_0^r\left\langle T(s)x,\left(T(t)\right)'x'\right\rangle ds-\frac{1}{r}\int_0^r\langle T(s)x,x'\rangle\,ds\right|.\end{aligned}$$

Again by the definition of the adjoint operator and changing integration boundaries we obtain

$$\begin{aligned}&\lim_{t\to 0}\|T(t)x_r - x_r\|\\&= \lim_{t\to 0} \sup_{\substack{x' \in X' \\ \|x'\| \leq 1}} \left| \frac{1}{r}\int_t^{r+t} \langle T(s)x, x'\rangle\, ds - \frac{1}{r}\int_0^r \langle T(s)x, x'\rangle\, ds \right|.\end{aligned}$$

Since t converges to 0, we can rearrange the two boundaries of the integral and obtain

$$\begin{aligned}&\lim_{t\to 0}\|T(t)x_r - x_r\|\\&\leq \lim_{t\to 0} \sup_{\substack{x' \in X' \\ \|x'\| \leq 1}} \left(\frac{1}{r}\int_r^{r+t} |\langle T(s)x, x'\rangle|\, ds + \frac{1}{r}\int_0^t |\langle T(s)x, x'\rangle|\, ds \right).\end{aligned}$$

By definition of the norm of x' $|\langle T(s)x, x'\rangle| < \|x'\| \cdot \|T(s)x\|$, hence

$$\lim_{t\to 0}\|T(t)x_r - x_r\| \leq \lim_{t\to 0}\frac{t}{r}\|x\| \left(\sup_{r\leq s\leq r+t}\|T(s)\| + \sup_{0\leq s\leq t}\|T(s)\| \right)$$

which converges to 0 as $t \to 0$ using Lemma 1.4.65.

Hence, $x_r \in E$ for any $x_r \in D$ which yields $D \subset E$ and we conclude that E is dense in X in the weak topology, thus by Proposition A.4.10 and convexity in the norm topology. Therefore, the semigroup $(T(t))_{t\in\mathbb{R}_+}$ is strongly continuous by Lemma 1.4.65 and Proposition 1.4.8. □

COROLLARY 1.4.69. *A semigroup*

$$(T(t))_{t\in\mathbb{R}_+}$$

on a Banach space X is strongly continuous if for any $x \in X$ and $x' \in X'$ the map $f_{x,x'}$

$$\begin{aligned}\mathbb{R}_+ &\to \mathbb{C}\\ t &\to \langle T(t)x, x'\rangle\end{aligned}$$

is continuous at $t = 0$.

PROOF. Let $x \in X$ and $x' \in X'$ be arbitrary and fixed. We have to show continuity of $t \to \langle T(t)x, x'\rangle$ at any $t \in \mathbb{R}_+$ in order to use Proposition 1.4.68 which permits us to conclude. Let $s \in \mathbb{R}_+$ be arbitrary. Right continuity of $t \to \langle T(t)x, x'\rangle$ at $t = s$ follows immediately from

$$\langle T(h+s)x, x'\rangle = \left\langle T(h)\underbrace{(T(s)x)}_{:=y}, x'\right\rangle$$

and the continuity of $f_{y,x'}$ at $h = 0$. Regarding left continuity of $t \to \langle T(t)x, x'\rangle$ at $t = s$ we have to show that

$$\lim_{h \nearrow 0} \langle T(h+s)x, x'\rangle - \langle T(s)x, x'\rangle = 0.$$

It holds

$$\begin{aligned}\lim_{h \nearrow 0} \langle T(h+s)x, x'\rangle - \langle T(s)x, x'\rangle &= \lim_{h \nearrow 0} \langle T(h+s)\left(\mathrm{Id} - T(-h)\right)x, x'\rangle \\ &\leq \sup_{t \in [0,s]} \|T(t)\| \underbrace{\lim_{h \nearrow 0} \langle \left(\mathrm{Id} - T(-h)\right)x, x'\rangle}_{=0}\end{aligned}$$

and by Lemma 1.4.65

$$\sup_{t \in [0,s]} \|T(t)\| < \infty$$

for any $s \in \mathbb{R}_+$ thus $t \to \langle T(t)x, x'\rangle$ is left continuous for any $t \in \mathbb{R}_+$. □

Definition 1.4.70. Let X be a Banach space and $x \in X$ arbitrary. The map $i(x)$:

$$\begin{aligned} X' &\to \mathbb{K} \\ x' &\to x'(x) \end{aligned}$$

is clearly linear and bounded. Hence $i(x) \in X''$ for any $x \in X$. The map

$$\begin{aligned} X &\to X'' \\ x &\to i(x) \end{aligned}$$

is injective by Corollary A.4.7. If it is also surjective then X is called *reflexive*.

Corollary 1.4.71. *On a reflexive Banach space the adjoint semigroup of a strongly continuous semigroup is strongly continuous.*

Proof. Let X be a reflexive Banach space and let

$$(T(t))_{t \in \mathbb{R}_+}$$

be a strongly continuous semigroup on X. By Proposition 1.4.59 for the adjoint semigroup (see Definition 1.4.44)

$$(T'(t))_{t \in \mathbb{R}_+}$$

the map

$$t \to T'(t)x'(x)$$

is continuous for any $x \in X$ and any $x' \in X'$. Since X is reflexive this implies that also the map

$$t \to \langle T'(t)x', x'' \rangle$$

is continuous for any $x'' \in X''$ and any $x' \in X'$. Therefore, by Theorem 1.4.68 the adjoint semigroup

$$(T'(t))_{t \in \mathbb{R}_+}$$

is a strongly continuous semigroup on X'. $\square$

CHAPTER 2

Markov, Feller and Generalized Feller Semigroups

Certain semigroups are of particular interest in probability theory and can be used to define stochastic processes. For definitions and terminology of stochastic processes we refer the reader to Appendix A.3.3.

2.1. Markov Semigroups

There are several different ways to define Markov semigroups and processes. This section largely follows the presentation in Chapter 3.1 in Revuz-Yor [**35**]. $(E, \mathcal{E})$ will always denote a measurable space and

$$\left(\Omega, \mathcal{F}, (\mathcal{G}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$$

a filtered probability space (see Definition A.3.75).

2.1.1. Definition of Markov semigroups. Using transition probabilities that satisfy the Functional Equation (Equation 1.1.1) one obtains an important semigroup on the space of bounded measurable functions called *Markov semigroup*.

In order to introduce this semigroup, we first recall the definition of *transition kernels* and *transition probabilities*.

Definition 2.1.1. The map

$$\kappa : \Omega \times \mathcal{E} \to [0, \infty]$$

is called *transition kernel* from $(\Omega, \mathcal{F})$ to$(E, \mathcal{E})$ if
(i) for any $A \in \mathcal{E}$ the map

$$\kappa(\cdot, A) : \Omega \to [0, \infty]$$

is $\mathcal{F}$-measurable and
(ii) for any $\omega \in \Omega$ the map

$$\kappa(\omega, \cdot) : \mathcal{E} \to [0, \infty]$$

is a measure on $(E, \mathcal{E})$.
If $\kappa(\omega, E) = 1$ for all $\omega \in \Omega$, then κ is called *transition probability* from $(\Omega, \mathcal{F})$ to $(E, \mathcal{E})$. If $(\Omega, \mathcal{F}) = (E, \mathcal{E})$ then we speak of *transition kernels/probabilities on*$(E, \mathcal{E})$.

REMARK 2.1.2. If $\kappa(x, E) \leq 1$ for all $x \in E$ but $\kappa(y, E) < 1$ for some $y \in E$, then one can add a new element to the space E, the so-called *cemetery* $\{\triangle\}$, and on $E_\triangle := E \cup \{\triangle\}$ define a transition probability (the properties can easily be checked)

$$\kappa' : E_\triangle \times \sigma(\mathcal{E}, \{\triangle\}) \to [0, 1]$$

by

$$\kappa'|_{:E\times\mathcal{E}} = \kappa$$

and

$$\kappa'(\{\triangle\}, A) = 0 \text{ for any } A \in \mathcal{E}$$
$$\kappa'(x, A \cup \{\triangle\}) = \kappa'(x, A) + 1 - \kappa(x, E) \text{ for any } A \in \mathcal{E}$$

for any $x \in E_\triangle$. For any function f on E the convention is to extend it to $E_\triangle$ by setting $f(\triangle) = 0$. Usually, the precise distinction between κ' and κ will not be made and κ' will simply be called κ.

In the following we need two properties of transition kernels. The first is that the integral of a positive, jointly measurable function with respect to a transition kernel is measurable (see Lemma A.3.59). The second property is that by composing two kernels one obtains again a kernel (see Lemma A.3.60).

Above definitions and properties permit to define on the state space of transition probabilities a one-parameter family of mappings that fulfills the Functional Equation (Equation 1.3.3):

DEFINITION 2.1.3. A family $(p(t))_{t\in\mathbb{R}_+}$ of transition probabilities (kernels) on $(E, \mathcal{E})$ is called *semigroup of transition probabilities (kernels) on* $(E, \mathcal{E})$ if for all $x \in E$, for all $s, t \in \mathbb{R}_+$ and all $A \in \mathcal{E}$

$$p(s+t)(x, A) = \int_E p(s)(y, A)p(t)(x, dy) \tag{2.1.1}$$

and

$$p(0)(x, \cdot) = \delta_x$$

hold. Here δ_x denotes the Dirac measure (see ExampleA.3.21). This definition can be extended to the space $(E_\triangle, \sigma(\mathcal{E}, \{\triangle\}))$ if necessary.

REMARK 2.1.4. Equation 2.1.1 most authors call Chapman-Kolmogorov equation or Master Equation.

REMARK 2.1.5. If $(p(t))_{t\in\mathbb{R}_+}$ is a family of transition kernels on $(E,\mathcal{E})$ such that $p(t)(x,E)\le 1$ for all $x\in E$ and all $t\in\mathbb{R}_+$, then for the corresponding family of transition probabilities $(p'(t))_{t\in\mathbb{R}_+}$ on $(E_\triangle,\sigma(\mathcal{E},\{\triangle\}))$ defined as in Remark 2.1.2 the condition

$$p'(s+t)(x,A)=\int_{E_\triangle}p'(t)(y,A)p'(s)(x,dy) \text{ for all } A\in\sigma(\mathcal{E},\{\triangle\})$$

holds for any $x\in E_\triangle$ and $s,t\in\mathbb{R}_+$ if Equation 2.1.1 holds for $(p(t))_{t\in\mathbb{R}_+}$ for all $x\in E$, for all $s,t\in\mathbb{R}_+$ and all $A\in\mathcal{E}$. In order to simplify notation, the following statements will only be made for transition probabilities on $(E,\mathcal{E})$. They can be extended to transition probabilities on $(E_\triangle,\sigma(\mathcal{E},\{\triangle\}))$ when necessary.

A semigroup of transition probabilities leads to a one-parameter semigroup (see Definition 1.3.1) on the Banach space of measurable bounded, real-valued functions on $(E,\mathcal{E})$:

DEFINITION 2.1.6. For a semigroup of transition probabilities

$$(p(t))_{t\in\mathbb{R}_+}$$

on $(E,\mathcal{E})$ we define the *Markov semigroup*

$$(P(t))_{t\in\mathbb{R}_+}$$

on the space of bounded, real-valued, $\mathcal{E}$-measurable functions by

$$P(t)f(x):=\int_E f(y)p(t)(x,dy).$$

REMARK 2.1.7. For a Markov semigroup

$$(P(t))_{t\in\mathbb{R}_+}$$

on $(E,\mathcal{E})$ for $t\in\mathbb{R}_+$ $P(t)f$ is defined also if f is non-negative, real-valued an $\mathcal{E}$-measurable.

Such a semigroup permits to define a stochastic process called *Markov process*:

DEFINITION 2.1.8. Let $(p(t))_{t\in\mathbb{R}_+}$ be a semigroup of transition probabilities and $(P(t))_{t\in\mathbb{R}_+}$ be the respective Markov semigroup. An adapted process (see Definition A.3.87) $(\lambda_t)_{t\in\mathbb{R}_+}$ on

$$\left(\Omega,\mathcal{F},(\mathcal{G}_t)_{t\in\mathbb{R}_+},\mathbb{P}\right)$$

with state space $(E, \mathcal{E})$ is called *Markov process* with respect to $(\mathcal{G}_t)_{t\in\mathbb{R}_+}$ with semigroup of transition probabilities $(p(t))_{t\in\mathbb{R}_+}$ if for any $\mathcal{E}$-measurable non-negative function

$$f : E \to \mathbb{R}_+ \cup \{\infty\}$$

and any $0 \leq s < t$

$$\mathbb{E}\left[f(\lambda_t) | \, \mathcal{G}_s \right] = P(t-s) f(\lambda_s) \tag{2.1.2}$$

holds $\mathbb{P}$-almost surely. $\mathbb{P} \circ \lambda_0^{-1}$ is called *initial distribution* of $(\lambda_t)_{t\in\mathbb{R}_+}$.

Remark 2.1.9. By linearity, Equation 2.1.2 holds for $\mathcal{E}$-measurable bounded functions $f : E \to \mathbb{R}$ as well. The definition of Markov processes implies that f$p(t-s) = \kappa_{\lambda_t,\lambda_s}$ or any $t > s \geq 0$ where $\kappa_{\lambda_t,\lambda_s}$ is the regular conditional probability (see Definition A.3.67).

Remark 2.1.10. More generally, Markov processes can also be defined for families $(p_{s,t})_{s,t\in\mathbb{R}_+}$ of transition probabilities on $(E, \mathcal{E})$ such that for any $0 \leq s \leq r \leq t$

$$p_{s,t}(x, A) = \int_E p_{s,r}(y, A) p_{r,t}(x, dy).$$

In this case $P_{s,t}$ is not a semigroup but it is still possible to define a Markov process in the above way where we replace Equation 2.1.2 by

$$\mathbb{E}\left[f(\lambda_t) | \, \mathcal{G}_s \right] = P_{s,t} f(\lambda_s).$$

Such a family of transition probabilities and such Markov processes are called *inhomogeneous* whereas the ones introduced above are called *homogeneous*. In the following, we only consider homogeneous processes and when speaking of Markov processes we always intend homogeneous ones.

Remark 2.1.11. (Motivation)

By setting $f = 1_A$ for some $A \in \mathcal{E}$ in Equation 2.1.2 we see that a Markov process possesses the properties that
(i) for predictions about the future it is sufficient to know the present instead of the whole past and
(ii) that for such predictions the present time by itself is not important; what matters is only the difference between the future time for which a prediction is to be made and the present time.

It turns out that also the other direction is true which we want to show in the following. The assumption that for a stochastic process

$(\lambda_t)_{t\in\mathbb{R}_+}$ properties (i) and (ii) hold implies that the family of maps $(z_{\mathbb{P}})_{\mathbb{P}\in\mathcal{M}_1(\Omega,\mathcal{F})}$ defined for some $r\in\mathbb{R}_+$ as

$$\begin{aligned} z_{\mathbb{P}}: \mathbb{R}_+ &\to \mathcal{M}_1(E) \\ r &\to \mathbb{P}\circ\lambda_r^{-1} \end{aligned}$$

satisfies Assumption 1.1.3 and Assumption 1.1.4 from Section 1.1. Thus, for the family $(T(t))_{t\in\mathbb{R}_+}$ defined as

$$\begin{aligned} T(t):\ \mathcal{M}_1(E) &\to \mathcal{M}_1(E) \\ \mathbb{P}\circ\lambda_r^{-1} &\to \mathbb{P}\circ\lambda_{r+t}^{-1} \end{aligned}$$

$(T(t))_{t\in\mathbb{R}_+}$ does not depend on r and has to be a one-parameter family of mappings that satisfies the Functional Equation (Equation 1.3.3):

$$T(t)\circ T(s) = T(t+s) \tag{2.1.3}$$

for any $s,t\in\mathbb{R}_+$. The map

$$\begin{aligned} p(t):\ E\times\mathcal{E} &\to [0,\infty] \\ (x,A) &\to \mathbb{E}\left[1_{\lambda_{r+t}\in A}1_{\lambda_r=x}\right] \end{aligned}$$

(which by assumption does not depend on $r\in\mathbb{R}_+$) is a transition probability. By definition for any $r\in\mathbb{R}_+$

$$\begin{aligned} T(t)\left(\mathbb{P}\circ\lambda_r^{-1}(A)\right) &= \mathbb{P}\left(\lambda_{r+t}^{-1}(A)\right) \\ &= \int_E \mathbb{E}\left[1_{\lambda_{r+t}\in A}1_{\lambda_r=x}\right]\cdot\mathbb{P}\circ\lambda_r^{-1}(dx) \\ &= \int_E p(t)\,(x,A)\cdot\mathbb{P}\circ\lambda_r^{-1}(dx). \end{aligned}$$

Hence, Equation 2.1.3 yields that $(p(t))_{t\in\mathbb{R}_+}$ is a semigroup of transition probabilities and $(\lambda_t)_{t\in\mathbb{R}_+}$ is a Markov process with respect to its natural filtration

$$\left(\mathcal{F}_t^0\right)_{t\in\mathbb{R}_+} := \sigma\left((\lambda_s)_{0\le s\le t}\right)_{t\in\mathbb{R}_+}$$

which is defined as the smallest filtration on $(\Omega,\mathcal{F})$ such that $(\lambda_t)_{t\in\mathbb{R}_+}$ is adapted since for any $A\in\mathcal{E}$

$$\begin{aligned} \mathbb{E}\left[1_A(\lambda_t)\middle|\,\mathcal{F}_s^0\right] &= \mathbb{E}\left[1_A(\lambda_t)\middle|\,\lambda_s\right] \\ &= p(t-s)\,(\lambda_s,A) \\ &= P_{t-s}1_A(\lambda_s). \end{aligned}$$

2.1.2. Properties of Markov processes. If not stated otherwise, for a Markov process $(\lambda_t)_{t\in\mathbb{R}_+}$ with respect to a filtration as filtration we take the natural filtration $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ of the process $(\lambda_t)_{t\in\mathbb{R}_+}$.

The proofs of the next four propositions roughly follow Revuz-Yor [**35**], Chapter III.1 .

PROPOSITION 2.1.12. *([**35**] , Proposition 1.4)*
A stochastic process

$$(\lambda_t)_{t\in\mathbb{R}_+}$$

on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *with state space* $(E, \mathcal{E})$ *is a Markov process with respect to*

$$\left(\mathcal{F}_t^0\right)_{t\in\mathbb{R}_+}$$

with initial distribution v *and semigroup of transition probabilities*

$$(p(t))_{t\in\mathbb{R}_+}$$

if and only if for all $k \in \mathbb{N}$ *, all times* $0 = t_0 < ... < t_k$*, and all non-negative,* $\mathcal{E}$*-measurable functions* $f_0, ..., f_k$

$$\begin{aligned}\mathbb{E}\left[\prod_{i=0}^{k} f_i(\lambda_{t_i})\right] &\\ = \int_E \cdots \Bigg(\int_E \Bigg(\int_E & f_k(x_k)\cdot p\left(t_k - t_{k-1}\right)(x_{k-1}, dx_k)\Bigg) \\ &\cdot f_{k-1}(x_{k-1})\cdot p\left(t_{k-1}-t_{k-2}\right)(x_{k-2}, dx_{k-1})\Bigg)\cdots f_0(x_0)d\nu(x_0).\end{aligned} \tag{2.1.4}$$

PROOF. In order to show the first implication, assume $(\lambda_t)_{t\in\mathbb{R}_+}$ is such a Markov process. Then, the properties of conditional expectations yield

$$\begin{aligned}&\mathbb{E}\left[\prod_{i=0}^{k} f_i(\lambda_{t_i})\right]\\ &= \mathbb{E}\left[\mathbb{E}\left[\prod_{i=0}^{k} f_i(\lambda_{t_i})\middle|\, \sigma\left((\lambda_s)_{0\le s\le t_{k-1}}\right)\right]\right]\\ &= \mathbb{E}\left[\prod_{i=0}^{k-1} f_i(\lambda_{t_i})\cdot P\left(t_k - t_{k-1}\right) f(\lambda_{t_{k-1}})\right]\\ &= \mathbb{E}\left[\prod_{i=0}^{k-2} f_i(\lambda_{t_i})\cdot \mathbb{E}\left[\left(f_{k-1}\cdot P\left(t_k-t_{k-1}\right) f\right)(\lambda_{t_{k-1}})\middle|\, \sigma\left((\lambda_s)_{0\le s\le t_{k-2}}\right)\right]\right]\\ &= \mathbb{E}\left[\prod_{i=0}^{k-2} f_i(\lambda_{t_i})\cdot P\left(t_{k-2}-t_{k-1}\right)\left(f_{k-1}\cdot P\left(t_k-t_{k-1}\right) f\right)(\lambda_{t_{k-2}})\right]\end{aligned}$$

etc. This proves one implication of the proposition.

For the other implication, assuming Equation 2.1.4 holds we need to show Equation 2.1.2. It is enough to prove that for any measurable non-negative function $f : E \to \mathbb{R}_+ \cup \{\infty\}$, any $0 \le s < t$ and any $A \in \mathcal{F}_s^0$ the equation

$$\mathbb{E}\left[f(\lambda_t) \cdot 1_A\right] = \mathbb{E}\left[P\left(t-s\right) f(\lambda_s) \cdot 1_A\right]. \tag{2.1.5}$$

holds. The set

$$\mathcal{D} = \left\{ A \in \mathcal{F}_s^0 \middle| \text{Equation 2.1.5 holds} \right\}$$

clearly is a Dynkin system by the monotone convergence theorem. Applying Equation 2.1.4 to both sides of Equation 2.1.5, we observe that for $n \in \mathbb{N}$ and

$$0 = t_0 < t_1 < ... < t_n \le s$$

and

$$F_{0,} F_1, ..., F_n \in \mathcal{E}$$

the set

$$\bigcap_{i=0}^{n} \lambda_{t_i}^{-1}(F_i)$$

is contained in $\mathcal{D}$. Since the system of such sets is an intersection stable generator of the product σ-algebra $\mathcal{F}_s^0$, by Lemma A.3.15

$$\mathcal{D} = \mathcal{F}_s^0,$$

hence the assertion of the proposition follows. □

Proposition 2.1.13. *([**35**] , Theorem 1.5)*
Let E be a polish space (see Definition A.1.14), $\mathcal{E}$ its Borel σ-algebra and $\mathcal{E}^{\mathbb{R}_+}$ the product σ-algebra (see Definition A.3.7) of $E^{\mathbb{R}_+}$. For any semigroup of transition probabilities

$$\left(p(t)\right)_{t\in\mathbb{R}_+}$$

on $(E, \mathcal{E})$ and any probability measure v on $(E, \mathcal{E})$, there exists a unique probability measure $\mathbb{P}_\nu$ on

$$\left(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+}\right)$$

such that the coordinate process (see Definition A.3.5)

$$\left(\lambda_t\right)_{t\in\mathbb{R}_+}$$

on

$$\left(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+}, \mathbb{P}_\nu\right)$$

is a Markov process with respect to the filtration

$$\left(\mathcal{F}_t^0\right)_{t\in\mathbb{R}_+}$$

with semigroup of transition probabilities

$$(p(t))_{t\in\mathbb{R}_+}$$

and initial distribution v.

PROOF. We want to define a probability measure $\mathbb{P}_\nu$ on $\left(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+}\right)$ using Kolmogorov extension theorem. For $n \in \mathbb{N}$ and $0 = t_0 < ... < t_n$ let $J := \{t_0, ..., t_n\} \subset \mathbb{R}_+$. For $F_0, F_1, ..., F_n \in \mathcal{E}$ we set

$$\begin{aligned}&\mathbb{P}_J\left(\underset{i\in J}{\times} F_i\right)\\ &:= \int_{F_0} \cdots \left(\int_{F_{n-1}} \left(p\left(t_n - t_{n-1}\right)\left(x_{n-1}, F_n\right)\right) \cdot p\left(t_{n-1} - t_{n-2}\right)\left(x_{n-2}, dx_{n-1}\right)\right) \cdots d\nu(x_0),\end{aligned}$$

thereby obtaining a σ-additive map on the generator

$$\left\{ \underset{i\in J}{\times} F_i \,\middle|\, F_0, ..., F_n \in \mathcal{E} \right\}$$

of $\mathcal{E}^J$ which is a semi-ring (see Definition A.3.27). By applying Carathéodory extension theorem (Theorem A.3.29) we obtain a unique probability measure $\mathbb{P}_J$ on $\left(E^J, \mathcal{E}^J\right)$. Proceeding this way we obtain a projective family (see Definition A.3.3)

$$(\mathbb{P}_J)_{J\subset\mathbb{R}_+,\ \text{finite}}$$

of probability measures on

$$\left(E^J, \mathcal{E}^J\right)_{J\subset\mathbb{R}_+,\ \text{finite}}$$

and by Kolmogorov extension theorem (Theorem A.3.102) there exists a unique probability measure $\mathbb{P}_\nu$ on $\left(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+}\right)$ such that for all finite $J \subset \mathbb{R}_+$ and $F \in \mathcal{E}^J$

$$\mathbb{P}_\nu\left(\left(\Pi_J^{\mathbb{R}_+}\right)^{-1}(F)\right) = \mathbb{P}_J(F)$$

holds where $\Pi_J^{\mathbb{R}_+}$ is the projection from Definition A.3.5. By definition of $\mathbb{P}_\nu$, for the coordinate process $(\Pi_t)_{t\in\mathbb{R}_+}$ (see DefinitionA.3.5), on $\left(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+}, \mathbb{P}_\nu\right)$, subsequently denoted

$$(\lambda_t)_{t\in\mathbb{R}_+} := (\Pi_t)_{t\in\mathbb{R}_+},$$

Equation 2.1.4 holds for indicator functions $f_0, ..., f_n$ of sets in $\mathcal{E}$. By linearity of the integral and monotone convergence (Theorem A.3.57) and Proposition A.3.19 this implies that Equation 2.1.4 holds also for all non-negative, $\mathcal{E}$-measurable functions $f_0, ..., f_n$. Hence, by Proposition 2.1.12 $(\lambda_t)_{t\in\mathbb{R}_+}$ is a Markov process with the desired properties. $\square$

NOTATION 2.1.14. From now on, we always assume that for any initial distribution ν and semigroup of transition probabilities on $(E, \mathcal{E})$ there is a probability measure $\mathbb{P}_\nu$ on

$$\left(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+}\right)$$

such that the coordinate process is a Markov process with respect to $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$(e.g. because E is polish and we use Proposition 2.1.13). Unless specified otherwise, when we speak of a Markov process we always intend one obtained in such a way. We write $\mathbb{E}_\nu$ instead of $\mathbb{E}_{\mathbb{P}_\nu}$ and for $x \in E$ and Dirac measure (see Example A.3.21) δ_x we write $\mathbb{P}_x$ instead of $\mathbb{P}_{\delta_x}$.

PROPOSITION 2.1.15. *([**35**], Proposition 1.6)*
Let

$$(\lambda_t)_{t\in\mathbb{R}_+}$$

be a Markov process with the state space $(E, \mathcal{E})$ *on the measurable space*

$$\left(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+}\right)$$

with respect to

$$\left(\mathcal{F}_t^0\right)_{t\in\mathbb{R}_+}$$

with some semigroup of transition probabilities and for any initial distribution ν *let* $\mathbb{P}_\nu$ *be the corresponding probability measure. Let*

$$Z:\ E^{\mathbb{R}_+} \to \mathbb{R}_+ \cup \{\infty\}$$

be measurable with respect to $\mathcal{E}^{\mathbb{R}_+}$. *(Or let* $Z:\ E^{\mathbb{R}_+} \to \mathbb{R}$ *be measurable, bounded.) Then*

$$\begin{aligned} E &\to \mathbb{R} \cup \{\infty\} \\ x &\to \mathbb{E}_x\left[Z\right] \end{aligned}$$

is measurable with respect to $\mathcal{E}$ *and*

$$\mathbb{E}_\nu\left[Z\right] = \int_E \mathbb{E}_x\left[Z\right] d\nu(x).$$

PROOF. For some $n \in \mathbb{N}$ and $0 \leq t_1 < ... < t_n$ and arbitrary $F_{t_1}, ..., F_{t_n} \in \mathcal{E}$ we set

$$\Gamma = \left\{\lambda_{t_1} \in F_1, ..., \lambda_{t_n} \in F_n\right\}. \tag{2.1.6}$$

In the first step of the proof, we show the assertion of the proposition for maps $Z = 1_\Gamma$. By Proposition 2.1.12 for any $x \in E$

$$\mathbb{E}_x\left[1_\Gamma\right]$$
$$= \int_{F_1} \cdots \left(\int_{F_{n-1}} \left(P_{t_n - t_{n-1}}(x_{n-1}, F_n)\right) \cdot P_{t_{n-1}-t_{n-2}}(x_{n-2}, dx_{n-1})\right) \cdots P_{t_1 - t_0}(x, dx_1).$$

Since by Lemma A.3.59

$$x_{n-2} \to \int_{F_{n-1}} \left(P_{t_n - t_{n-1}}(x_{n-1}, F_n)\right) \cdot P_{t_{n-1}-t_{n-2}}(x_{n-2}, dx_{n-1})$$

is $\mathcal{E}$-measurable we can deduce inductively that

$$x \to \mathbb{E}_x \left[1_\Gamma\right]$$

is also $\mathcal{E}$-measurable. Additionally, by Proposition 2.1.12

$$\mathbb{E}_\nu \left[1_\Gamma\right] = \int_\Omega \mathbb{E}_x \left[1_\Gamma\right] d\nu(x) \tag{2.1.7}$$

holds.

In the second step of the proof, we prove the proposition for all positive measurable maps Z. We define the set $\mathcal{D} \subset \mathcal{E}^{\mathbb{R}_+}$ such that for all $F \in \mathcal{D}$ both the equation

$$\mathbb{E}_\nu \left[1_F\right] = \int_\Omega \mathbb{E}_x \left[1_F\right] d\nu(x)$$

holds and the map

$$x \to \mathbb{E}_x \left[1_F\right]$$

is $\mathcal{E}$-measurable. One shows easily, that $\mathcal{D}$ is a Dynkin system and by the previous step it contains the intersection stable generator of $\mathcal{E}^{\mathbb{R}_+}$, hence by Lemma A.3.15 $\mathcal{D} = \mathcal{E}^{\mathbb{R}_+}$. Since by Proposition A.3.19 any positive random variable Z can be written as limit of positive simple functions, the assertion of this proposition follows by Lemma A.3.17 and monotone convergence (Theorem A.3.57). □

PROPOSITION 2.1.16. *(Markov property,* **[35]** *, Proposition 1.7) Let*

$$(\lambda_t)_{t\in\mathbb{R}_+}$$

be a Markov process with the state space $(E, \mathcal{E})$ *on the measurable space*

$$\left(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+}\right)$$

with respect to

$$\left(\mathcal{F}_t^0\right)_{t\in\mathbb{R}_+}$$

with some semigroup of transition probabilities and for any initial distribution ν *let* $\mathbb{P}_\nu$ *be the corresponding probability measure. Let*

$$Z : E^{\mathbb{R}_+} \to \mathbb{R}_+ \cup \{\infty\}$$

be measurable with respect to $\mathcal{E}^{\mathbb{R}_+}$ *(Or let*

$$Z : E^{\mathbb{R}_+} \to \mathbb{R}$$

be measurable, bounded). Let $\mathbb{E}_{\lambda_t}[Z]$ be the composition of $x \to \mathbb{E}_x[Z]$ and

$$\begin{aligned} E^{\mathbb{R}_+} &\to E \\ \omega &\to \lambda_t(\omega). \end{aligned}$$

For $t \in \mathbb{R}_+$ let

$$\theta_t : E^{\mathbb{R}_+} \to E^{\mathbb{R}_+}$$

be the map

$$(\omega(s))_{s\in\mathbb{R}_+} \to (\omega(s+t))_{s\in\mathbb{R}_+}.$$

Then for any $t > 0$ and any initial distribution ν on $(E, \mathcal{E})$

$$\mathbb{E}_\nu\left[Z \circ \theta_t \middle| \mathcal{F}_t^0\right] = \mathbb{E}_{\lambda_t}[Z] \tag{2.1.8}$$

holds $\mathbb{P}_\nu$-almost surely.

Proof. As composition of two measurable maps $\mathbb{E}_{\lambda_t}[Z]$ is clearly measurable with respect to $\mathcal{F}_t^0$. As a first step, we want to show the proposition for $Z = 1_\Gamma$, where

$$\Gamma = \{\lambda_{t_1} \in F_1, ..., \lambda_{t_n} \in F_n\}$$

for some $n \in \mathbb{N}$ and $0 \le t_1 < ... < t_n$ and $F_1, ..., F_n \in \mathcal{E}$ arbitrary. We need to show that for any $B \in \mathcal{F}_t^0$ the equation

$$\mathbb{E}_\nu[(1_\Gamma \circ \theta_t) \cdot 1_B] = \mathbb{E}_\nu[\mathbb{E}_{\lambda_t}[1_\Gamma] \cdot 1_B] \tag{2.1.9}$$

holds. The system of sets $\mathcal{D} \subset \mathcal{E}^{\mathbb{R}_+}$ defined as set of sets $D \in \mathcal{D}$ such that the equation

$$\mathbb{E}_\nu[(1_\Gamma \circ \theta_t) \cdot 1_D] = \mathbb{E}_\nu[\mathbb{E}_{\lambda_t}[1_\Gamma] \cdot 1_D]$$

holds is a Dynkin system as one can easily show. For any $m \in \mathbb{N}$ and

$$B' = \{\lambda_{s_1} \in F'_1, ..., \lambda_{s_m} \in F'_m\} \tag{2.1.10}$$

where $0 \le s_1 < ... < s_m \le t$ and $F'_1, ..., F'_m \in \mathcal{E}$ are arbitrary the equation

$$\mathbb{E}_\nu[(1_\Gamma \circ \theta_t) \cdot 1_{B'}] = \mathbb{E}_\nu[\mathbb{E}_{\lambda_t}[1_\Gamma] \cdot 1_{B'}]$$

follows from applying Proposition 2.1.12 to both sides of the equation. Thus, $\mathcal{D}$ contains the (intersection stable) generator of the σ- algebra $\mathcal{E}^{\mathbb{R}_+}$, hence by Lemma A.3.15 $\mathcal{D} = \mathcal{E}^{\mathbb{R}_+}$ and the equation

$$\mathbb{E}_\nu\left[1_\Gamma \circ \theta_t \middle| \mathcal{F}_t^0\right] = \mathbb{E}_{\lambda_t}[1_\Gamma]$$

holds.

In a second step, we show the assertion of the proposition for all positive measurable maps Z. For this purpose, we observe that the system of sets $\mathcal{D}' \subset \mathcal{E}^{\mathbb{R}_+}$ defined as set of all sets $D' \in \mathcal{D}'$ such that the equation

$$\mathbb{E}_\nu \left[1_{D'} \circ \theta_t \middle| \mathcal{F}_t^0 \right] = \mathbb{E}_{\lambda_t} \left[1_{D'} \right]$$

holds is a Dynkin system that contains an intersection stable generator of $\mathcal{E}^{\mathbb{R}_+}$ hence by Lemma A.3.15 $\mathcal{D}' = \mathcal{E}^{\mathbb{R}_+}$. Since by Proposition A.3.19 any positive random variable Z can be written as limit of positive simple functions, the assertion of this proposition follows by monotone convergence (Theorem A.3.57). □

REMARK 2.1.17. If $p(t)(x, E) < 1$ for some $x \in E$ and some $t \in \mathbb{R}_+$ such that the construction from Remark 2.1.2 needs to be employed to obtain a semigroup of transition probabilities, Equation 2.1.8 in Proposition 2.1.16 is shown only on the set $\{\lambda_t \neq \Delta\}$ as by convention the right hand side of the equation is 0 if $\lambda_t = \Delta$.

Let $(\lambda_t)_{t\in\mathbb{R}_+}$ be a Markov process with the state space $(E, \mathcal{E})$ on the measurable space $\left(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+}\right)$ with respect to the natural filtration and for any initial distribution ν let $\mathbb{P}_\nu$ be the corresponding probability measure. The family $(P(t))_{t\in\mathbb{R}_+}$ of maps defined by

$$P(t)f(x) := \mathbb{E}_x \left[f(\lambda_t) \right]$$

for all $x \in E$ and $f \in \ell^\infty(E)$ is a one-parameter semigroup on $\ell^\infty(E)$ by the calculation

$$\begin{aligned} \mathbb{E}_x \left[f(\lambda_{s+t}) \right] &= \mathbb{E}_x \left[\mathbb{E}_x \left[f(\lambda_{s+t}) \middle| \mathcal{F}_t^0 \right] \right] \\ &= \mathbb{E}_x \left[\left(P(s) f \right) (\lambda_t) \right]. \end{aligned}$$

One can define the generator A of the semigroup, only if the restriction of the semigroup to some closed subspace $D \subset \ell^\infty(E)$ is strongly continuous. There are several ways to weaken the notion of a generator.

Following **[9]**, one can always define the *infinitesimal generator* of a Markov process:

DEFINITION 2.1.18. Let $(\lambda_t)_{t\in\mathbb{R}_+}$ be a Markov process with the state space $(E, \mathcal{E})$ on the measurable space $\left(E^{\mathbb{R}_+}, \mathcal{E}^{\mathbb{R}_+}\right)$ with respect to the natural filtration and for any initial distribution ν let $\mathbb{P}_\nu$ be the corresponding probability measure. Define

$$\mathcal{D}(\mathscr{A}) := \left\{ f \in \ell^\infty(E) : \lim_{t\searrow 0} \frac{\mathbb{E}_x \left[f(\lambda_t) \right] - f(x)}{t} \text{ exists for all } x \in E \right\},$$

and for any $f \in \mathcal{D}(\mathscr{A})$ and any $x \in E$ define

$$\mathscr{A} f(x) := \lim_{t\searrow 0} \frac{\mathbb{E}_x \left[f(\lambda_t) \right] - f(x)}{t}.$$

The linear map $\mathscr{A}$ is called *infinitesimal generator*.

The next proposition motivates a different way to generalize the notion of a generator for a Markov process:

PROPOSITION 2.1.19. *Let $(\lambda_t)_{t\in\mathbb{R}_+}$ be a Markov process with the state space $(E,\mathcal{E})$ on the measurable space $\left(E^{\mathbb{R}_+},\mathcal{E}^{\mathbb{R}_+}\right)$ with respect to the natural filtration $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ and for any initial distribution ν let $\mathbb{P}_\nu$ be the corresponding probability measure. Let $(P(t))_{t\in\mathbb{R}_+}$ be the family of maps defined by*

$$P(t)f(x) := \mathbb{E}_x\left[f(\lambda_t)\right]$$

for all $x\in E$ and $f\in\ell^\infty(E)$. Let $D\subset\ell^\infty(E)$ be a closed subspace and let the restriction of $(P(t))_{t\in\mathbb{R}_+}$ on D be strongly continuous with generator $(A,\mathcal{D}(A))$. Then for any $x\in E$ and $f\in\mathcal{D}(A)$ the process $\left(M_t^f\right)_{t\in\mathbb{R}_+}$ defined by

$$M_t^f := f(\lambda_t) - f(\lambda_0) - \int_0^t (Af)(\lambda_s)ds$$

is a martingale with respect to $\mathbb{P}_x$ and $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$.

PROOF. We observe that if $f\in\mathcal{D}(A)$ then clearly for all $x\in E$

$$\begin{aligned}\mathbb{E}_x\left[f(\lambda_t) - f(\lambda_0) - \int_0^t (Af)(\lambda_s)ds\right] &= P(t)f(x) - f(x) - \int_0^t \mathbb{E}_x\left[(Af)(\lambda_s)\right]ds \\ &= P(t)f(x) - f(x) - \int_0^t P(t)(Af)(x)ds \\ &= 0,\end{aligned}$$

where the last step is possible thanks to 1.4.15. Furthermore, the process $\left(M_t^f\right)_{t\in\mathbb{R}_+}$ defined by

$$M_t^f := f(\lambda_t) - f(\lambda_0) - \int_0^t (Af)(\lambda_s)ds$$

is clearly adapted with respect to its natural filtration $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ and bounded (since $Af\in\ell^\infty(E)$ by definition). By Proposition 2.1.16,

$\left(M_t^f\right)_{t\in\mathbb{R}_+}$ is a martingale with respect to $\mathbb{P}_x$ for any $x \in E$ since

$$\begin{aligned}
&\mathbb{E}_x\left[f(\lambda_t)-f(\lambda_0)-\int_0^t (Af)(\lambda_s)ds\middle|\mathcal{F}_s^0\right]\\
&= f(\lambda_s)-f(\lambda_0)-\int_0^s (Af)(\lambda_s)ds+\mathbb{E}_x\left[f(\lambda_t)-f(\lambda_s)-\int_s^t (Af)(\lambda_r)dr\middle|\mathcal{F}_s^0\right]\\
&= f(\lambda_s)-f(\lambda_0)-\int_0^s (Af)(\lambda_s)ds+\mathbb{E}_x\left[f(\lambda_{t-s}\circ\theta_s)-f(\lambda_0\circ\theta_s)-\int_0^{t-s} (Af)(\lambda_r\circ\theta_s)dr\middle|\mathcal{F}_s^0\right]\\
&= f(\lambda_s)-f(\lambda_0)-\int_0^s (Af)(\lambda_s)ds+\mathbb{E}_{\lambda_s}\left[f(\lambda_{t-s})-f(\lambda_0)-\int_0^{t-s} (Af)(\lambda_r)dr\right]\\
&= f(\lambda_s)-f(\lambda_0)-\int_0^s (Af)(\lambda_s)ds.
\end{aligned}$$

This motivates the following definition (see for Example [**35**], Definition VII.1.8): □

DEFINITION 2.1.20. Let $(\lambda_t)_{t\in\mathbb{R}_+}$ be a Markov process with the state space $(E,\mathcal{E})$ on the measurable space $\left(E^{\mathbb{R}_+},\mathcal{E}^{\mathbb{R}_+}\right)$ with respect to the natural filtration $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ and let $f: E\cup\{\Delta\}\to\mathbb{R}$ be measurable. If there exists a measurable map $g: E\cup\{\Delta\}\to\mathbb{R}$ such that for all $x\in E$ and for every $t\in\mathbb{R}_+$

$$\int_0^t |g(\lambda_s)|\, ds < \infty$$

$\mathbb{P}_x$-almost surely and

$$M_t^f := f(\lambda_t) - f(x) - \int_0^t g(\lambda_s)ds$$

is well defined and is a right continuous martingale with respect to $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ and probability measure $\mathbb{P}_x$ then one defines $\mathcal{G}f := g$ and $f\in\mathcal{D}(\mathcal{G})$ and calls $\mathcal{G}$ *extended infinitesimal generator*.

REMARK 2.1.21. The requirement that the martingale be right continuous becomes more clear when looking at Feller process and in particular at the existence of càdlàg modifications in this case (see Theorem 2.2.6).

This can be weakened further (see for example [**12**]):

DEFINITION 2.1.22. Let $(\lambda_t)_{t\in\mathbb{R}_+}$ be a Markov process with the state space $(E,\mathcal{E})$ on the measurable space $\left(E^{\mathbb{R}_+},\mathcal{E}^{\mathbb{R}_+}\right)$ with respect to the natural filtration $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ and let $f: E\cup\{\Delta\}\to\mathbb{R}$ be measurable. If there exists a measurable map $g: E\cup\{\Delta\}\to\mathbb{R}$ such that for all $x\in E$ and for every $\mathbb{P}_x$-almost surely and

$$M_t^f := f(\lambda_t) - f(x) - \int_0^t g(\lambda_s)ds$$

is well defined and is a local martingale (see Definition 3.0.2) with respect to $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ and probability measure $\mathbb{P}_x$ then one defines $\mathcal{G}f := g$ and $f\in\mathcal{D}(\mathcal{G})$ and calls $\mathcal{G}$ *extended generator*.

DEFINITION 2.1.23.

Let $(\lambda_t)_{t\in\mathbb{R}}$ be a Markov process on $\mathbb{R}^d$. Let the infinitesimal generator $\mathscr{A}$ be such that for any $f\in C_c^2(\mathbb{R}^d)$ it holds for $x\in\mathbb{R}^d$

$$\mathscr{A}f(x) = c(x)f(x) + \sum_{i\in\{1,\ldots,d\}} b_i(x)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j\in\{1,\ldots,d\}} a_{ij}(x)\frac{\partial^2 f}{\partial x_i\partial x_j}(x))$$

where for any $i,j\in\{1,...,d\}$ a_{ij}, b_i, and $c\leq 0$ are functions on $\mathbb{R}^d$ and the matrix

$$(a_{ij}(x))_{i,j\in\{1,\ldots,d\}}$$

is non-negative and symmetric for any $x\in\mathbb{R}^d$.
Then the vector

$$(b_i(x))_{i\in\{1,\ldots,d\}}$$

is called *drift* of $(\lambda_t)_{t\in\mathbb{R}}$, and the matrix

$$(a_{ij}(x))_{i,j\in\{1,\ldots,d\}}$$

is called *diffusion matrix* of $(\lambda_t)_{t\in\mathbb{R}}$ provided they are Borel-measurable and locally bounded. Furthermore, c of $(\lambda_t)_{t\in\mathbb{R}}$ is called *killing rate*.

REMARK 2.1.24. We want to heuristically explain the meaning of the functions c, b_i, and a_{ij} for $i,j\in\{1,...,d\}$ in Definition 2.1.23 and fix $x\in\mathbb{R}^d$.

Regarding c, we see by setting $f=1$ on some neighborhood around x that

$$\mathbb{E}_x\left[f(\lambda_h)\right] = hc(x) + 1 + o(h).$$

and the measure $\mathbb{P}_{\lambda_t}$ loses mass with rate c because the process is "killed" and moved to the cemetery, which explains its name.

Regarding b_i, by setting $c = 0$ and $f_i(y) = y_i$ we obtain

$$\mathbb{E}_x\left[\lambda_h^{(i)} - x_i\right] = hb_i(x) + o(h)$$

which shows that infinitesimally $(\lambda_t)_{t\in\mathbb{R}_+}$ moves by the vector

$$(b_i(x))_{i\in\{1,\dots,d\}})$$

which explains the term *drift*.

Regarding b_i, by setting $c = 0$ and

$$f_{ik}(y) = (y_i - x_i)(y_k - x_k)$$

we obtain

$$\mathbb{E}_x\left[\left(\lambda_h^{(i)} - x_i\right)\left(\lambda_h^{(k)} - x_k\right)\right] = ha_{ik}(x) + o(h).$$

Thus, the instantaneous rate of change at 0 of the covariance of the vector

$$\left(\lambda_h^{(i)}\right)_{i\in\{1,\dots,d\}}$$

is given by

$$(a_{ij})_{i,j\in\{1,\dots,d\}}\,.$$

EXAMPLE 2.1.25. (Brownian motion)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and $(W_t)_{t\in\mathbb{R}_+}$ be a Brownian motion on it (see Definition A.3.83 and Theorem A.3.84) and let $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ be its natural filtration. It is well known that$(W_t)_{t\in\mathbb{R}_+}$ is a Markov process which we will show in the following. Let f be a non-negative, measurable map. Set $g(x, y) := f(x + y)$. Then by Lemma A.3.64

$$\begin{aligned}\mathbb{E}\left[f(W_t)\right|\mathcal{F}_s] &= \mathbb{E}\left[f(W_t - W_s + W_s)\right|\mathcal{F}_s]\\ &= \mathbb{E}\left[g(W_t - W_s, W_s)\right|\mathcal{F}_s]\\ &= \mathbb{E}\left[g(W_t - W_s, W_s)\right|\sigma(W_s)]\\ &= \mathbb{E}\left[f(W_t)\right|\sigma(W_s)]\,.\end{aligned}$$

Furthermore, by Proposition A.3.69

$$\mathbb{E}\left[f(W_t)\right|\sigma(W_s)] = \int f(x)\kappa_{W_t,W_s}(W_s, dx).$$

By definition of Brownian motion for any $y \in \mathbb{R}$ and any Borel set $A \in \mathcal{B}(\mathbb{R})$

$$p(t-s)(y, A) := \kappa_{W_t,W_s}(y, A) = \int_A \frac{1}{\sqrt{2\pi(t-s)}} \exp\left(\frac{-(x-y)^2}{2(t-s)}\right) dx$$

and for any $t \in \mathbb{R}$ $y \to p(t)(y,A)$ is measurable by Lemma A.3.59. From the identity

$$\int_{\mathbb{R}} \exp\left(\frac{-x^2}{2\sigma^2}\right) dx = \sqrt{2\pi\sigma^2}$$

and completing the square it follows that$(p(t))_{t\in\mathbb{R}_+}$ is a semigroup of transition probabilities:

$$\begin{aligned}
&\int_{\mathbb{R}} p(s)(y,A)p(t)(x,dy)\\
&= \int_{\mathbb{R}} \left(\int_A \frac{1}{\sqrt{2\pi s}} \exp\left(\frac{-(z-y)^2}{2s}\right) dz\right) \frac{1}{\sqrt{2\pi t}} \exp\left(\frac{-(x-y)^2}{2t}\right) dy\\
&= \int_A \frac{1}{\sqrt{2\pi s}} \frac{1}{\sqrt{2\pi t}} \exp\left(\frac{-tz^2 - sx^2 + \left(\frac{tz+sx}{\sqrt{s+t}}\right)^2}{2st}\right) \left(\int_{\mathbb{R}} \exp\left(\frac{-\left(\frac{tz+sx}{s+t} - y\right)^2}{2st/(s+t)}\right) dy\right) dz\\
&= \int_A \frac{1}{\sqrt{2\pi(s+t)}} \exp\left(\frac{-(z-x)^2}{2(s+t)}\right) dz\\
&= p(s+t)(x,A).
\end{aligned}$$

For all $k \in \mathbb{N}$, all times $0 = t_0 < ... < t_k$, and all non-negative, measurable functions $f_0, ..., f_k$

$$\mathbb{E}\left[\prod_{i=0}^k f_i(W_{t_i})\right]$$

$$\mathbb{E}\left[\prod_{i=0}^k f_i(W_{t_i})\middle| \mathcal{F}_s\right]$$

$$\begin{aligned}
= \int_E \cdots \Bigg(\int_E \bigg(\int_E & f_k(x_k) \cdot p\,(t_k - t_{k-1})\,(x_{k-1}, dx_k)\bigg)\\
&\cdot f_{k-1}(x_{k-1}) \cdot p\,(t_{k-1} - t_{k-2})\,(x_{k-2}, dx_{k-1})\Bigg) \cdots f_0(x_0)\delta_0(x_0).
\end{aligned}$$

and by Proposition 2.1.12 $(W_t)_{t\in\mathbb{R}_+}$ is a Markov process with respect to $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$.

EXAMPLE 2.1.26. (Geometric Brownian motion on $E = \{x \in \mathbb{R} : x > 0\}$ is a Markov process)

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be some probability space and $(W_t)_{t\in\mathbb{R}_+}$ be a Brownian motion on it. For any $x \in E$ let

$$S_t^x = x \exp\left(\left(\alpha - \frac{\sigma^2}{2}\right) t + \sigma W_t\right).$$

With the substitution

$$y := \varphi^x(z) := x \exp\left(\sigma z + \left(\alpha - \frac{\sigma^2}{2}\right) t\right)$$

for $B, F \in \mathcal{B}(\mathbb{R}_+)$ the calculation

$$\begin{aligned}
\mathbb{E}\left[1_B(S_t^x)1_F(S_s^x)\right] &= \mathbb{E}\left[1_B\left(\varphi^x\left(W_t\right)\right)1_F\left(\varphi^x\left(W_s\right)\right)\right] \\
&= \mathbb{E}\left[1_{(\varphi^x)^{-1}(B)}\left(W_t\right)1_{(\varphi^x)^{-1}(F)}\left(W_s\right)\right] \\
&= \mathbb{E}\left[\kappa_{W_t,W_s}(W_s,(\varphi^x)^{-1}(B))1_{(\varphi^x)^{-1}(F)}\left(W_s\right)\right] \\
&= \mathbb{E}\left[\kappa_{W_t,W_s}(W_s,(\varphi^x)^{-1}(B))1_F(S_s^x)\right]
\end{aligned}$$

yields that the conditional regular probability (see Definition A.3.67)$\kappa_{S_t^x,S_s^x}$ is given by

$$\begin{aligned}
\kappa_{S_t^x,S_s^x}(w,B) &= \kappa_{W_t,W_s}(w,(\varphi^x)^{-1}(B)) \\
&= \int_{(\varphi^x)^{-1}(B)} \frac{1}{\sqrt{2\pi\left(t-s\right)}}\exp\left(\frac{-\left(z-w\right)^2}{2\left(t-s\right)}\right)dz \\
&= \int_B \frac{1}{\sqrt{2\pi\left(t-s\right)y^2\sigma^2}}\exp\left(\frac{-\left(\ln\left(\frac{y}{x}\right)-\left(\alpha-\frac{\sigma^2}{2}\right)t-w\right)^2}{2\left(t-s\right)\sigma^2}\right)dy
\end{aligned}$$

for any $w \in \mathbb{R}_+$, $B \in \mathcal{B}(\mathbb{R}_+)$. It follows then exactly as in the case of Brownian motion in Example 2.1.25 that $(S_t^x)_{t\in\mathbb{R}_+}$ is a Markov process with respect to its natural filtration.

Furthermore, by Proposition A.3.92

$$\left(\exp\left(\left(-\frac{\sigma^2}{2}\right)t+\sigma W_t\right)\right)_{t\in\mathbb{R}_+}$$

is a martingale with respect to the natural filtration. Hence,

$$\mathbb{E}\left[S_t^x\right] = xe^{\alpha t}\mathbb{E}\left[\exp\left(\left(-\frac{\sigma^2}{2}\right)t+\sigma W_t\right)\right] = xe^{\alpha t}.$$

REMARK 2.1.27. More generally, if $(\lambda_t)_{t\in\mathbb{R}_+}$ is the $\mathbb{R}^d$-valued solution of a stochastic differential equation (see Definition A.3.116)

$$d\lambda_t = \mu(\lambda_t)dt + \sigma(\lambda_t)dW_t$$

for a d-dimensional Brownian motion (see Definition A.3.83) $W = (W_t)_{t\in\mathbb{R}_+}$ on the filtered probability space

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

and for the measurable maps

$$\mu = \left(\mu^1,...,\mu^d\right) : \mathbb{R}^d \to \mathbb{R}^d$$

and

$$\sigma = \left(\sigma^{i,j}\right)_{i,j\in\{1,...,d\}} : \mathbb{R}^d \to \mathbb{R}^{d\times d}$$

and if there is a constant $C > 0$ such that for any $x, y \in \mathbb{R}$

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| < C\,|x - y|\,,$$

$(\lambda_t)_{t\in\mathbb{R}_+}$ is called *Ito diffusion*. One can show that $(\lambda_t)_{t\in\mathbb{R}_+}$ is a Markov process (see [**32**], Theorem 7.1.2). Then by Ito formula (A.3.115) μ is the drift and σ the diffusion matrix as defined in Definition 2.1.23.

2.2. Feller semigroups

In this section we introduce Feller processes as special class of Markov processes. Throughout this section, we make the assumption that the state space E of the processes is locally compact (see Definition A.1.12) and Hausdorff and has a base (see Definition A.1.2) with at most countably many elements. Among other things, these assumptions ensure that the space is polish (see Proposition A.1.15).

DEFINITION 2.2.1. A strongly continuous, positive (see Definition 2.3.39), contractive (see Definition 1.4.10) semigroup $(Q(t))_{t\in\mathbb{R}_+}$ on $C_0(E)$ is called *Feller semigroupon E.*

PROPOSITION 2.2.2. *([**35**], Proposition 2.2) Let $(P(t))_{t\in\mathbb{R}_+}$ be a Feller semigroup on E. Then there exists a semigroup of transition probabilities $(p(t))_{t\in\mathbb{R}_+}$ on $(E, \mathcal{B}(E))$ (or on $(E \cup \{\triangle\}, \mathcal{B}(E \cup \{\Delta\}))$) such that for any $t \in \mathbb{R}_+$ and any $f \in C_0(E)$*

$$P(t)f = \int_E f(y)p(t)(x, dy).$$

We call $(p(t))_{t\in\mathbb{R}_+}$ the associated Feller semigroup of transition probabilities.

REMARK 2.2.3. Either $p(t)(x,\cdot)$ is a probability measure on $(E, \mathcal{E})$ for any $x \in E$ or by Remark 2.1.2 $p'(t)(x,\cdot)$ is a probability measure on $(E \cup \{\triangle\}, \mathcal{B}(E \cup \{\Delta\}))$ for any $x \in E$. In order to simplify notation, in the following we assume without loss of generality the first case. We set the cemetery Δ (see Remark 2.1.2) to be ∞ obtained from the one-point-compactification of E (see Definition A.1.16).

PROOF. In the first step of the proof we construct a family of positive measures $(p(t)(x,\cdot))_{t\in\mathbb{R}_+}$ on $(E, \mathcal{B}(E))$ for any $x \in E$, and in the second we show that the family is a semigroup of transition probabilities on $(E, \mathcal{B}(E))$.

Fix some $x \in E$ and $t \in \mathbb{R}_+$. Then the map

$$\ell_x :\ f \to P(t)f(x)$$

is a bounded positive linear functional on $C_0(E)$ hence by Riesz representation theorem (A.4.13) there is a unique complex regular measure $P(x,\cdot)$ on $(E,\mathcal{B}(E))$ such that

$$P(t)f(x)=\int_E f(y)p(t)(x,dy)$$

holds for any $f\in C_0(E)$. We need to show that $p(t)(x,\cdot)$ is a positive measure. This is the case since due to metrizability of the topology of E (see Proposition A.1.13) and separability (choose one element in each of the sets of the countable base) any open set $O\subset E$ can be written as countable union of open balls

$$B_r(z):=\{y\in E:\, d(z,y)<r\}$$

for some $(r_k)_{k\in\mathbb{N}}\in\mathbb{R}_+$ and $(z_k)_{k\in\mathbb{N}}\in E$:

$$O=\bigcup_{k\in\mathbb{N}}B_{r_k}(z_k)$$

Thus, by approximating each $1_{B_{r_k}(z_k)}$ pointwise by a positive sequence $(f_{k_n})_{n\in\mathbb{N}}$ in $C_0(E)$ we obtain

$$p(t)(x,O)=\int_E 1_O(y)p(t)(x,dy)=\lim_{n\to\infty}\int_E\sup_{k\in\{1,\dots,n\}}\{f_{k_n}(y)\}\,p(t)(x,dy)\geq 0.$$

Since $p(t)(x,\cdot)$ is outer regular, this implies that $p(t)(x,\cdot)\geq 0$ on all of $\mathcal{B}(E)$.

As second step of the proof, we need to check the properties of a semigroup of transition probabilities. By Theorem A.4.13

$$p(t)(x,E)=\|\ell_x\|\leq 1$$

so either $p(t)(x,\cdot)$ is a probability measure on $(E,\mathcal{B}(E))$ for any $x\in E$ or by Remark 2.1.2 $p(t)(x,\cdot)$ is a probability measure on

$$(E\cup\{\triangle\},\sigma(\mathcal{B}(E)\cup\{\triangle\}))$$

for any $x\in E$.

Next, we show measurability of

$$x\to p(t)(x,A)$$

for any $A\in\mathcal{B}(E)$. Since by definition of $(P(t))_{t\in\mathbb{R}_+}$ the map

$$x\to\int_E f(y)p(t)(x,dy)$$

is in $C_0(E)$, it is in particular measurable with respect to $\mathcal{B}(E)$. As above, we can write any open set $O\subset E$ as countable union

$$O=\bigcup_{k\in\mathbb{N}}B_{r_k}(z_k)$$

for some $(r_k)_{k\in\mathbb{N}} \in \mathbb{R}_+$ and $(z_k)_{k\in\mathbb{N}} \in E$. We can approximate the function $1_{B_{r_k}(z_k)}$ pointwise with positive functions

$$(f_{k_n})_{n\in\mathbb{N}} \subset C_0(E)$$

which yields by dominated convergence (Theorem A.3.58)

$$x \to \int_E 1_O(y)p(t)(x,dy) = \lim_{n\to\infty} \int_E \sup_{k\in\{1,\dots,n\}} \{f_{k_n}(y)\}\, p(t)(x,dy),$$

hence $\mathcal{B}(E)$-measurability of $x \to p(t)(x,O)$. One easily shows that the family of sets $\mathcal{M} \subset 2^{\Omega}$ defined as

$$\mathcal{M} := \{M \in \mathcal{B}(E) :\, x \to p(t)(x,M) \text{ is measurable}\}$$

is a Dynkin system hence $\mathcal{M} = \mathcal{B}(E)$ by Lemma A.3.15.

Last, we see that indeed for any $s,t \in \mathbb{R}_+$, any $A \in \mathcal{B}(E)$ and any $x \in E$

$$p(s+t)(x,A) = \int_E p(s)(y,A)p(t)(x,dy)$$

since by assumption on $(P(t))_{t\in\mathbb{R}_+}$

$$P(s+t)f = P(t)P(s)f$$

holds for any $f \in C_0(E)$ and conclude by approximating 1_A by some sequence

$$(f_n)_{n\in\mathbb{N}} \subset C_0(E)$$

as before. □

PROPOSITION 2.2.4. *Let $(P(t))_{t\in\mathbb{R}_+}$ be a Feller semigroup on E and $(p(t))_{t\in\mathbb{R}_+}$ its associated Feller semigroup of transition probabilities. Then the family $(Q(t))_{t\in\mathbb{R}_+}$ of linear bounded maps*

$$Q(t)f(x) := \int_E f(y)p(t)(x,dy)$$

is a Markov semigroup and by Proposition 2.1.13 for any probability measure v on $(E,\mathcal{B}(E))$ there exists a Markov process $(\lambda_t)_{t\in\mathbb{R}_+}$ with respect to its natural filtration $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ with semigroup of transition probabilities $(p(t))_{t\in\mathbb{R}_+}$ and initial distribution v. Such a Markov process is called Feller process.

DEFINITION 2.2.5. A stochastic process $(\lambda_t)_{t\in\mathbb{R}_+}$ with a state space that is closed is called *càdlàg* process (from French: continue à droite, limite à gauche) if all paths are *càdlàg paths*, that is, they are right continuous (thus $\lambda_t = \lambda_{t+} := \lim_{s\searrow t} \lambda_s$ for any $t \in \mathbb{R}_+$) and possess left limits (hence $\lambda_{t-} := \lim_{s\nearrow t} \lambda_s$ exists in the state space for any $t > 0$). If all paths are left continuous and possess right limits, we call the process (and the paths) *càglàd*. For a càdlàg process $(\lambda_t)_{t\in\mathbb{R}_+}$ we can define the process $(\Delta\lambda_t)_{t\in\mathbb{R}_+}$as

$$\Delta\lambda_t := \begin{cases} \lambda_t - \lim\limits_{s\nearrow t}\lambda_s & \text{if } t>0 \\ 0 & \text{if } t=0. \end{cases}$$

Wherever $\Delta\lambda_t \neq 0$ for some $t>0$ we call $\Delta\lambda_t$ a *jump*.

One can show (see e.g. [**35**], Theorem 2.7)

THEOREM 2.2.6. *A Feller process possesses a version (see Definition A.3.86) that is a càdlàg process.*

NOTATION 2.2.7. We remind the reader of the naming convention for Markov processes from Notation 2.1.14 and add the convention that for Feller processes we always consider the version that is a càdlàg process.

For Markov processes we know that the Markov property holds (Proposition 2.1.16). For Feller processes we know more than that.

We define $\mathcal{F}^\nu_\infty$ as the completion (see Definition A.3.32) of $(\mathcal{B}(E))^{\mathbb{R}_+}$ with respect to $\mathbb{P}_\nu$ and set

$$\mathcal{F}_\infty = \bigcap_\nu \mathcal{F}^\nu_\infty,$$

where the intersection is taken over all probability measures on $(E,\mathcal{B}(E))$. Furthermore, we call $\mathcal{N}^\nu$the set of all $\mathbb{P}_\nu$-null sets on $\mathcal{F}^\nu_\infty$ and set
$\mathcal{F}^\nu_t := \sigma(\mathcal{N}^\nu \cup \mathcal{F}^0_t)$ and

$$\mathcal{F}_t = \bigcap_\nu \mathcal{F}^\nu_t,$$

where again the intersection is taken over all probability measures on $(E,\mathcal{B}(E))$.

THEOREM 2.2.8. *(Strong Markov property,* [**35**] *Theorem 3.1)*
Let $(\lambda_t)_{t\in\mathbb{R}_+}$ be a Feller process. Let

$$Z:\ E^{\mathbb{R}_+} \to \mathbb{R}_+ \cup \{\infty\}$$

be measurable with respect to $\mathcal{F}_\infty$ (Or let $Z:\ E^{\mathbb{R}_+}\to\mathbb{R}$ be measurable, bounded). Let $\mathbb{E}_{\lambda_t}[Z]$ be the composition of $x\to\mathbb{E}_x[Z]$ and

$$\begin{aligned} E^{\mathbb{R}_+} &\to E \\ \omega &\to \lambda_t(\omega). \end{aligned}$$

For $t \in \mathbb{R}_+$ let $\theta_t : E^{\mathbb{R}_+} \to E^{\mathbb{R}_+}$ be defined by

$$(\omega(s))_{s\in\mathbb{R}_+} \to (\omega(s+t))_{s\in\mathbb{R}_+} .$$

Let $\tau : E^{\mathbb{R}_+} \to \mathbb{R}_+$ be a $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$-stopping time (see Definition A.3.91) and $\mathcal{F}_\tau \subset \mathcal{F}_\infty$ the σ-algebra of events determined prior to the stopping time τ (see Definition A.3.91). Define

$$\lambda_\tau := \lambda_{\tau(\omega)}(\omega) \text{ on } \{\tau \neq \infty\}$$

and set

$$\lambda_\tau = \Delta \text{ on } \{\tau = \infty\} .$$

Furthermore, define

$$\theta_\tau := \theta_{\tau(\omega)}(\omega) \text{ on } \{\tau \neq \infty\}$$

and on $\{\tau = \infty\}$ set θ_τ as the map from $E^{\mathbb{R}_+}$ to Δ.

Then for any initial measure ν on $(E, \mathcal{B}(E))$:

$$\mathbb{E}_\nu [Z \circ \theta_\tau | \mathcal{F}_\tau] = \mathbb{E}_{\lambda_\tau} [Z]$$

on $\{\lambda_\tau \neq \Delta\}$.

Since a Feller semigroup $(P(t))_{t\in\mathbb{R}_+}$ on E is a strongly continuous semigroup on $C_0(E)$, all results on strongly continuous semigroups from Section 1.4 carry over. In particular, there exists a generator (see Definition 1.4.13):

Definition 2.2.9. The generator A of a Feller semigroup $(P(t))_{t\in\mathbb{R}_+}$ on E is given by

$$\begin{aligned} A:\ \ & \mathcal{D}(A) \subset C_0(E) \to C_0(E) \\ & f \to \lim_{h\searrow 0} \frac{P(h)f - f}{h} \end{aligned}$$

on the dense domain $\mathcal{D}(A)$ (see Proposition 1.4.24).

The limit is of course to be taken with respect to the norm of $C_0(E)$.

Regarding the form of the generator of a Feller semigroup, in case $E = \mathbb{R}^d$, $d \in \mathbb{N}$ one can show:

Theorem 2.2.10. *([**35**], Theorem VII.1.13)*
Let $(P(t))_{t\in\mathbb{R}_+}$ be a Feller semigroup on $\mathbb{R}^d$ and let

$$C_c^\infty(\mathbb{R}^d) \subset \mathcal{D}(A) .$$

Then

$$C_c^2(\mathbb{R}^d) \subset \mathcal{D}(A)$$

and for $f \in C_c^2(\mathbb{R}^d)$ and an open set U whose closure is compact it holds for $x \in U$:

$$Af(x) = c(x)f(x) + \sum_{i\in\{1,...,d\}} b_i(x)\frac{\partial f}{\partial x_i}(x) + \frac{1}{2}\sum_{i,j\in\{1,...,d\}} a_{ij}(x)\frac{\partial^2 f}{\partial x_i \partial x_j}(x))$$

(2.2.1)

$$+ \int_{\mathbb{R}^d\setminus\{x\}} \left[f(y) - f(x) - \sum_{i\in\{1,...,d\}} 1_U(y)\,(y_i - x_i)\,\frac{\partial f}{\partial x_i}(x)\right] N(x,dy)$$

for functions a_{ij}, $i,j \in \{1,...,d\}$ on U that do not depend on U such that for any $x \in U$ the matrix

$$(a_{ij})_{i,j\in\{1,...,d\}}$$

is non-negative and symmetric, for functions b_i $i \in \{1,...,d\}$ on U that may depend on U, for a function $c \leq 0$ on U that does not depend on U and for a kernel N such that $N(x,\cdot)$ is a Radon measure on $\mathbb{R}^d \setminus \{x\}$ that may depend on U.

Remark 2.2.11. According to Definition 2.1.23, $b = (b_i(x))_{i\in\{1,...,d\}}$ is the drift of the corresponding Feller process, $a = (a_{ij})_{i,j\in\{1,...,d\}}$ is the diffusion matrix, and c is the killing rate.

Remark 2.2.12. We want to heuristically explain the meaning of N. Setting $c = 0$, $U = \{x\}$and $f(y) = 1_A(y)$ for $x \notin A$

$$\mathbb{P}_x[\lambda_h \in A] = h \cdot N(x,A) + o(h)$$

which shows that $N(x,\cdot)$ measures the time derivative at 0 of the probability that the process $(\lambda_t)_{t\in\mathbb{R}_+}$ jumps from x into a certain set.

2.3. Generalized Feller semigroups

Generalized Feller semigroups have been introduced in a special setting by Röckner and Sobol [**36**] in 2006 and were defined and investigated more generally in [**15**] in 2010. They are defined on so-called $\mathcal{B}^\rho-$*spaces* which in turn are defined on completely regular spaces. Thus, in order to define such semigroups, in the first subsection we introduce completely regular spaces and other separation axioms. In the second subsection, we define admissible weight functions before introducing B-rho spaces in the third subsection. In this section (E,τ) will always denote a topological space that is completely regular. For additional terminology regarding topology, the reader is referred to Appendix A.1.

2.3.1. Separation axioms.

DEFINITION 2.3.1. A topological space (T, τ) is called *Hausdorff* if any two points $x, y \in T$ possess disjoint neighborhoods U_x and U_y.

DEFINITION 2.3.2. A topological space (E, τ) is called *completely regular* if it is Hausdorff and if for any closed set $A \subset Y$ and any point $x \in E \setminus A$ there exists a continuous function $f : E \to [0, 1]$ such that $f(x) = 1$ and $f(y) = 0$ for all $y \in A$.

DEFINITION 2.3.3. A topological space (N, τ) is called *normal* if it is Hausdorff and if for all disjoint closed sets A, $B \subset E$ there are disjoint neighborhoods U_A of A and U_B of B.

REMARK 2.3.4. There are different naming conventions in the literature. Some authors do not require completely regular spaces and normal spaces to be Hausdorff and call the space we call completely regular *Tychonoff* space. Others do not define a completely regular space or a normal space N to be Hausdorff but instead ask for less: They demand that for any two points $x, y \in N$ there must be neighborhoods U_x of x and U_y of y such that U_x does not contain y and U_y does not contain x. However, here we follow the conventions of Bourbaki **[7]** and use the definition in the sense stated above .

For normal spaces we have access to Urysohn's Lemma:

LEMMA 2.3.5. *(Urysohn's Lemma,* **[7]***, Chapter IX, §4, Theorem 1) Let N be a normal space and $A, B \subset N$ be nonempty closed sets. Then there is a continuous function $f : N \to [0, 1]$ with $f(A) = \{0\}$ and $f(B) = \{1\}$.*

Additionally, the Tietze-Urysohn extension theorem is known:

THEOREM 2.3.6. *(Tietze-Urysohn extension theorem) (***[7]***, Chapter IX, §4, Theorem 2 and* **[40]** *Lemma 7.9)*
Let (N, τ) be a topological space. N is normal if and only if for each closed subset $A \subset N$ and each continuous function $f : A \to \mathbb{R}$ there is a continuous extension $F : N \to \mathbb{R}$.
If $|f| : A \to \mathbb{R}$ is additionally bounded by $C < \infty$, then there is also a continuous extension $F : N \to \mathbb{R}$ such that $|F| : N \to \mathbb{R}$ is also bounded by $C < \infty$.

We note the following relationship between two of the separation axioms.

PROPOSITION 2.3.7. *A normal space is completely regular.*

PROOF. By Urysohn's Lemma (Lemma 2.3.5) the required function exists. □

For a completely regular space a statements similar to Tietze-Urysohn extension theorem and Urysohn's Lemma can be shown and will be used frequently in the following sections.

PROPOSITION 2.3.8. *Let E be completely regular and $K \subset E$ compact. Then a real-valued continuous function $f \in C(K, \mathbb{R})$ on K can be extended to a continuous function $F \in C(E, \mathbb{R})$ on all of E. If additionally $|f| < C < \infty$ then there is an extension $F \in C(E, \mathbb{R})$ such that $|F| < C < \infty$.*

PROOF. We would like to apply the Tietze-Urysohn extension theorem (Theorem 2.3.6). However, it allows the extension only on normal spaces. But we can use Proposition A.1.10 in order to embed E by an embedding i in a compact Hausdorff set N which by Proposition A.1.11 is normal. $i(K)$ is also compact on N with respect to the subspace topology $\tau(i(E))$ (see Lemma A.1.8) on N. Hence, by Lemma A.1.9 $i(K)$ is compact with respect to the topology of N. Since N is Hausdorff, the compact set $i(K)$ is closed (see Lemma A.1.5). We can apply the Tietze-Urysohn extension Theorem (Theorem 2.3.6) on N to extend the function $f \circ i^{-1} \in C(i(K), \mathbb{R})$ to a continuous function $G \in C(N, \mathbb{R})$ such

$$f \circ i^{-1}\big|_{i(K)} = G|_{i(K)}$$

and $|G| \leq C$ if $|f| \leq C$. Therefore, $F := G \circ i$ possesses the desired properties. □

PROPOSITION 2.3.9. *(Urysohn's Lemma in the completely regular case) Let E be completely regular, $K \subset E$ compact, $A \subset E$ closed and $A \cap K = \emptyset$. Then there is a continuous function $f : E \to [0,1]$ such that $f(K) = \{0\}$, $f(A) = \{1\}$.*

PROOF. As in Proposition 2.3.8, we embed E in a compact Hausdorff set N by an embedding i. $i(K)$ is compact, hence closed in the compact Hausdorff space N. Since $i(A)$ is closed in the subspace topology $\tau(i(E))$, there is a closed set $B \subset N$ such that $B \cap i(E) = i(A)$ and clearly $B \cap i(K) = \emptyset$. Applying Urysohn's Lemma in the normal space N we see that there is a continuous function $g : N \to [0,1]$ with $g(i(K)) = \{0\}$ and $g(B) = \{1\}$. Setting $f := g \circ i$, we conclude. □

COROLLARY 2.3.10. *Let E be a completely regular space, $\mathcal{B}(E)$ its Borel σ-algebra and μ a measure on $(E, \mathcal{B}(E))$ and $B \in \mathcal{B}(E)$. If there* *is a sequence of compact sets* $(K_n)_{n\in\mathbb{N}}$ *and open sets* $(O_n)_{n\in\mathbb{N}}$ *such that* $K_n \subset B \subset O_n$ *for any* $n \in \mathbb{N}$ *and*

$$\mu(O_n \setminus K_n) \to 0,$$

then there exists a sequence $(f_n)_{n\in\mathbb{N}}$ *of non-negative continuous functions with* $f_n \leq 1_{O_n}$ *for any* $n \in \mathbb{N}$ *such that*

$$\lim_{n\to\infty} f_n = 1_B$$

μ*-almost surely and in* $L^1(E,\mu)$*.*

PROOF. Thanks to Urysohn's Lemma in the complete regular case there is a sequence $(g_n)_{n\in\mathbb{N}}$ of non negative continuous functions with $1_{K_n} \leq g_n \leq 1_{O_n}$ for any $n \in \mathbb{N}$ such that $g_n \to 1_B$ in $L^1(E,\mu)$ and μ-probability. By Proposition A.3.53 there exists a subsequence $(g_{n_k})_{k\in\mathbb{N}}$ such that $g_{n_k} \to 1_B$ almost surely. □

2.3.2. Admissible weight functions. In the definition of B^ρ-spaces in subsection 2.3.3 *admissible weight functions* appear. For this purpose, we define and investigate them in this subsection.

DEFINITION 2.3.11. A function $\rho : E \to (0,\infty)$ is called *admissible weight function* if the sets

$$K_R := \{x \in E : \rho(x) \leq R\}$$

are compact for all $R \geq 0$. The pair (E, ρ) is called *weighted space*.

REMARK 2.3.12. The identity $E = \bigcup_{n\in\mathbb{N}} K_n$ yields that E is σ-compact which means that it is the countable union of compact sets.

In order to investigate admissible weight functions further, we recall the following definition from analysis (see also Figure 2.3.1):

DEFINITION 2.3.13. Let T be a topological space. A function

$$f : T \to \mathbb{R}$$

is called *lower* (*upper*) *semicontinuous* if for any $\varepsilon > 0$ and any $x \in T$ there exists a neighborhood U_x of x such that $f(y) > f(x) - \varepsilon$ $(f(y) < f(x) + \varepsilon)$ for all $y \in U_x$.

LEMMA 2.3.14. *An admissible weight function* $\rho : E \to (0,\infty)$ *is lower semicontinuous.*

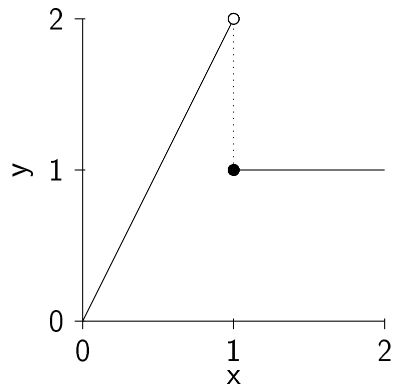


FIGURE 2.3.1. The function is lower semicontinuous at x=1, but not upper semicontinuous.

PROOF. For any $\varepsilon > 0$ and any $x \in E$ the set

$$K_{\rho(x)-\varepsilon} := \{y \in E : \rho(y) \leq \rho(x) - \varepsilon\}$$

is compact and $x \notin K_{\rho(x)-\varepsilon}$. Since E is Hausdorff, compact sets are closed by Lemma A.1.5. Thus, there exists a neighborhood U_x of x such that $U_x \cap K_{\rho(x)-\varepsilon} = \emptyset$. □

We know that on a compact set a continuous functions attains its maximum and minimum. For semicontinuous functions similar statements can be shown.

LEMMA 2.3.15. *A lower (upper) semicontinuous function f on a compact set K is bounded from below (above).*

PROOF. Let $f : K \to \mathbb{R}$ be lower semicontinuous and $\varepsilon > 0$. Then for any $x \in K$ there is a neighborhood U_x of x such that for any $y \in U_x$

$$f(y) > f(x) - \varepsilon.$$

Since K is compact, finitely many such neighborhoods suffice to cover K.

For upper semicontinuous functions the assertion follows in the same fashion. □

LEMMA 2.3.16. *On a compact set K a lower (upper) semicontinuous function $f : K \to \mathbb{R}$ attains its minimum (maximum).*

PROOF. Let $f : K \to \mathbb{R}$ be lower semicontinuous. By Lemma 2.3.15 f is bounded from below. By lower semicontinuity, for any

$x \in K$ there exists a neighborhood $U_x \ni x$ such that for all $y \in U_x$

$$f(y) > f(x) - \frac{1}{2}\left(f(x) - \inf_{x \in K} f(x)\right)$$

holds. Moreover,

$$\bigcup_{x \in K} U_x$$

is an open cover of the compact set K hence the neighborhoods of finitely many x_i, $i \in \{1, ..., n\}$ suffice to cover K :

$$K \subset \bigcup_{i \in \{1,...,n\}} U_{x_i}.$$

This implies that for all $y \in K$

$$f(y) > \frac{1}{2} \min_{i \in \{1,...,n\}} f(x_i) + \frac{1}{2}\left(\inf_{x \in K} f(x)\right)$$

holds. Taking the infimum on the left hand side yields

$$\inf_{x \in K} f(x) \geq \min_{i \in \{1,...,n\}} f(x_i) \geq \inf_{x \in K} f(x).$$

Thus, the infimum of $f : K \to \mathbb{R}$ is attained on K.

For upper semicontinuous functions the statement follows by the same reasoning. □

COROLLARY 2.3.17. *An admissible weight function $\rho : E \to (0, \infty)$ attains its minimum on E.*

The product space of weighted spaces is again a weighted space.

LEMMA 2.3.18. *Let (E_i, ρ_i), $i \in \{1, ..., n\}$ be weighted spaces. Then*

$$(E_1 \times ... \times E_n, \rho)$$

is a weighted space, where

$$\rho(x_1, ..., x_n) := \rho_1(x_1) \cdots \rho_n(x_n).$$

PROOF. We first show that $E_1 \times ... \times E_n$ is completely regular. It is clear that it is Hausdorff. Furthermore, for a closed set $A \subset E_1 \times ... \times E_n$ and $x \in E_1 \times ... \times E_n \setminus A$ by definition of the product topology (Definition A.1.7) we can find an open neighborhood U_x of x given by

$$U_x = \bigtimes_{i=1,...,n} U_x^i,$$

where each $U_x^i \subset E_i$, $i \in \{1, ..., n\}$ is a neighborhood of x. By definition of completely regular spaces, there exist continuous maps

$$f_i : E_i \to [0, 1]$$

such that $f_i(x) = 1$ and $f_i(y_i) = 0$ for all $y_i \in E_i \backslash U_x^i$ and the continuous map

$$f(x_1, ..., x_n) := f_1(x_1) \cdots f_n(x_n)$$

shows that $E_1 \times ... \times E_n$ is completely regular.

Next, we show that ρ is an admissible weight function. Let without loss of generality $\rho_i \geq 1$ for $i \in \{1, ..., n\}$. Let $R > 0$ be arbitrary. Then

$$\begin{aligned} &\{(x_1, ..., x_n) \in E_1 \times ... \times E_n : \rho_1(x_1) \cdots \rho_n(x_n) \leq R\} \\ &\subset \{x_1 \in E_1 : \rho_1(x_1) \leq R\} \times ... \times \{x_n \in E_n : \rho_n(x_n) \leq R\}. \end{aligned}$$

Since the right hand side is compact by Proposition A.1.17 we only need to show closedness of the left hand side. For $y = (y_1, ..., y_n)$ such that $\rho(y_1, ..., y_n) > R$, by lower semicontinuity of ρ_1,...ρ_n (see Lemma 2.3.14) for any $\varepsilon > 0$ there exist open neighborhoods $U_{y_1}^\varepsilon \subset E_1$ of y_1, ..., $U_{y_n}^\varepsilon \subset E_n$ of y_n such that for any $u_i \in U_{y_i}^\varepsilon$, $i \in \{1, ..., n\}$

$$\rho_i(u_i) > \rho_i(y_i) - \varepsilon.$$

Hence for $u \in U_{y_1}^\varepsilon \times ... \times U_{y_n}^\varepsilon$

$$\rho(u) > (\rho_1(y_1) - \varepsilon) \cdots (\rho_n(y_n) - \varepsilon)$$

and the right hand side is larger than R for ε small enough. Thus,

$$\{(x_1, ..., x_n) \in E_1 \times ... \times E_n : \rho_1(x_1) \cdots \rho_n(x_n) \leq R\}$$

is a closed subset of a compact set, hence compact. □

LEMMA 2.3.19. *A locally compact Hausdorff spaces E with countable base $\mathfrak{B}$ is a weighted space.*

PROOF. As a metrizable space (see Proposition A.1.15) , E is also completely regular. We show that E is σ-compact. By local compactness, for any $x \in E$ there is a compact neighborhood K_x. Thus, there exists an open neighborhood $O_x \subset K_x$ of x. By definition of the base, there is $B_x \subset O_x$, $B_x \in \mathfrak{B}$ such that $x \in B_x$. For any $x \in E$ the set $\overline{B_x} \subset K_x$ is compact and by assumption,

$$E = \bigcup_{x \in E} \overline{B_x}$$

is the union of countably many elements, hence E is σ-compact. Let

$$E = \bigcup_{n \in \mathbb{N}} K_n$$

be such a union. We define an admissible weight function in the following way:

$$\rho(x) := \min_{n \in \mathbb{N}} \{n : x \in K_n\}.$$

□

One example of a space on which it is always possible to find a continuous admissible weight function is of course $\mathbb{R}^n, n \in \mathbb{N}$.

2.3.3. $\mathscr{B}^\rho$-spaces. In this subsection (E, ρ) always denotes a weighted space and we introduce $\mathscr{B}^\rho$-spaces where we follow [**15**].

DEFINITION 2.3.20. For a Banach space Z and an admissible weight function ρ we define

$$B^\rho(E;Z) := \left\{ f : E \to Z : \sup_{x \in E} \rho(x)^{-1} \|f(x)\| < \infty \right\}.$$

REMARK 2.3.21. By Corollary 2.3.17 for the space $B^\rho(E;Z)$ we can assume that $\rho \geq 1$, if necessary.

2.3.20 In the following, we would like to show that, provided the norm is chosen well, this space is a Banach space. In order to do so, we recall the following fact:

Let Z be a Banach space and T be some set. The space of all bounded maps form T to Z

$$\ell^\infty(T;Z) := \left\{ f : T \to Z : \sup_{x \in T} \|f(x)\| < \infty \right\}$$

equipped with the norm

$$\|\cdot\|_\infty : f \to \sup_{x \in T} \|f(x)\|$$

is a Banach space (see Proposition A.4.4).

Turning back to the space $B^\rho(E;Z)$ we observe:

PROPOSITION 2.3.22. *$B^\rho(E;Z)$ is a vector space. On $B^\rho(E;Z)$ the map*

$$\|\cdot\|_\rho : f \to \sup_{x \in E} \frac{\|f(x)\|}{\rho(x)}$$

is a norm. $B^\rho(E;Z)$ endowed with the norm $\|\cdot\|_\rho$ is a Banach space.

PROOF. That $B^\rho(E;Z)$ is a vector space follows easily and that $\|\cdot\|_\rho$ possesses the properties of a seminorm is clear. Furthermore, $\|f\|_\rho = 0$ means that for all $x \in Y$

$$\frac{\|f(x)\|}{\rho(x)} \leq 0,$$

which by $\rho > 0$ implies $f = 0$. Hence, $\|\cdot\|_\rho$ is a norm.

Regarding completeness of $B^\rho(E;Z)$, let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $B^\rho(E;Z)$. Then for any $\varepsilon > 0$ there exists N_ε such that for all $n, m > N_\varepsilon$

$$\sup_{x\in E} \frac{\|f_n(x) - f_m(x)\|}{\rho(x)} < \varepsilon.$$

Thus, the sequence $(g_n)_{n\in\mathbb{N}}$ defined by

$$g_n(x) := \frac{f_n(x)}{\rho(x)}$$

is a Cauchy sequence in $\ell^\infty(E;Z)$, hence by Proposition A.4.4 converges in $\ell^\infty(E;Z)$ to a bounded function $g \in \ell^\infty(E;Z)$ as n tends to infinity. Defining

$$f(x) := \rho(x) \cdot g(x)$$

we obtain $\lim_{n\to\infty} f_n = f$ in $B^\rho(E;Z)$. □

REMARK 2.3.23. Similarly, we can define

$$\widetilde{B^\rho}(E;Z) := \left\{ f : E \to Z : \sup_{x\in E} \rho(x)^{-1} \|f(x)\| < \infty, f \text{ measurable} \right\}$$

and obtain, that this is a Banach space.

NOTATION 2.3.24. We denote the space of bounded continuous maps between a topological space T and a normed vector space N by $C_b(T,N)$.

We remark $C_b(E,Z) \subset B^\rho(E;Z)$ and define:

DEFINITION 2.3.25. The closure of $C_b(E,Z)$ in $B^\rho(E;Z)$ is denoted by $\mathscr{B}^\rho(E;Z)$.

$\mathscr{B}^\rho(E;Z)$ is a closed subspace of the Banach space $B^\rho(E;Z)$. It holds:

LEMMA 2.3.26. *A closed subset of a Banach space is itself complete.*

PROOF. Any Cauchy sequence in the closed subset converges to some limit in the Banach space. Since the subset is closed, the limit must also lie in the closed subset. □

Therefore we obtain:

PROPOSITION 2.3.27. *$\mathscr{B}^\rho(E;Z)$ is a Banach space.*

NOTATION 2.3.28. We write $\mathscr{B}^\rho(E) := \mathscr{B}^\rho(E;\mathbb{R})$, $\widetilde{B^\rho}(E) := \widetilde{B^\rho}(E;\mathbb{R})$, and $B^\rho(E) := B^\rho(E;Z)$.

In the following, we will study some important properties of this space. A first one concerns the nature of its dual space. For this purpose, we introduce Radon measures and signed measures. For additional remarks on measure theory we refer to Appendix A.3.

DEFINITION 2.3.29. Let $(\Omega, \mathcal{F})$ be a measurable space. A map $\mu : \mathcal{F} \to \mathbb{R}$ is called *signed measure* if for all pairwise disjoint sets $F_1, F_2, ... \in \mathcal{F}$

$$\mu\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mu(F_i).$$

If this identity holds for a map $\mu : \mathcal{F} \to \mathbb{C}$, then the map is called *complex measure*.

An important result for signed measures is the following:

THEOREM 2.3.30. *(Hahn-Jordan decomposition,* **[30]**, *Corollary 7.44)* *For a signed measure μ on the measurable space $(\Omega, \mathcal{F})$ there are unique positive finite measures μ^+ and μ^- such that*

$$\mu = \mu^+ - \mu^-$$

and there exists a set $A \in \mathcal{F}$ such that $\mu^+(A) = 0$ and $\mu^-(\Omega \setminus A) = 0$.

DEFINITION 2.3.31. For a signed or complex measure μ on a σ-algebra $\mathcal{F}$ we define the *total variation*

$$|\mu| : \mathcal{F} :\to \mathbb{R}_+ \cup \{\infty\}$$

as

$$|\mu|(F) = \sup \sum_{k=1}^{n} |\mu(F_k)|$$

where the supremum is taken over all finite partitions

$$F = \bigcup_{k=1}^{n} F_k,$$

such that the sets $(F_k)_{k \in \{1,...,n\}} \subset \mathcal{F}$ are pairwise disjoint.

An application of the Hahn-Jordan decomposition theorem immediately yields:

COROLLARY 2.3.32. *The total variation of a signed measure is given by*

$$|\mu| = \mu^+ + \mu^-.$$

DEFINITION 2.3.33. Let T be a Hausdorff topological space and $\mathcal{B}(T)$ its Borel σ-algebra. A *Radon measure* is a measure

$$\mu : \mathcal{B}(T) \to \mathbb{R}_+ \cup \{\infty\}$$

that is

(i) *locally finite*, which means that every point $x \in T$ has a neighborhood U_x such that $\mu(U_x) < \infty$, and

(ii) *inner regular*, which means that for every $B \in \mathcal{B}(T)$

$$\mu(B) = \sup \{\mu(K) : K \subset B, K \text{ compact}\}.$$

The space of Radon measures is denoted by $\mathcal{M}_+(T)$.

LEMMA 2.3.34. *A Radon measure μ on a Hausdorff topological space T is outer regular, that is for every $B \in \mathcal{B}(T)$*

$$\mu(B) = \inf \{\mu(O) : O \supset B, O \subset T, O \textit{ open}\}.$$

PROOF. For $B \in \mathcal{B}(T)$ and some $\varepsilon > 0$ we choose a compact set

$$K_\varepsilon \subset T \setminus B$$

such that

$$\mu((T \setminus B) \setminus K_\varepsilon) < \varepsilon.$$

K_ε is closed since T is Hausdorff. Thus, the set

$$(T \setminus K_\varepsilon) \supset B$$

is open and

$$\mu((T \setminus K_\varepsilon) \setminus B) < \varepsilon.$$

□

DEFINITION 2.3.35. A measure is called *regular* , if it is inner and outer regular. A signed or complex measure μ is called *regular* if $|\mu|$ is regular.

DEFINITION 2.3.36. Let T be a Hausdorff topological space. Let $\mathcal{B}(T)$ be its Borel σ-algebra. A *signed Radon measure* is a signed measure $\mu : \mathcal{B}(T) \to \mathbb{R}$ for which $|\mu| : \mathcal{B}(T) \to \mathbb{R}_+ \cup \{\infty\}$ is a Radon measure. Its space is denoted by $\mathcal{M}(T)$. In the same fashion, one can also define a *complex Radon measure* and call the space $\mathcal{M}_c(T)$.

Adding uniqueness to the statement already proved in [**15**], we can completely characterize the dual space of $\mathscr{B}^\rho(E)$ by the following theorem.

THEOREM 2.3.37. *(Riesz representation for $\mathscr{B}^\rho(E)$)*
Let $\ell : \mathscr{B}^\rho(E) \to \mathbb{R}$ be a continuous linear map. Then, there exists a unique signed Radon measure μ such that

$$\ell(f) = \int_E f(x)\mu(dx) \text{ for all } f \in \mathscr{B}^\rho(E). \tag{2.3.1}$$

Additionally,

$$\|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})} = \int_E \rho(x)\,|\mu|\,(dx).$$

On the other hand, for any signed Radon measure μ for which

$$\int_E \rho(x)\,|\mu|\,(dx)$$

is finite,

$$\begin{aligned} \mathscr{B}^\rho(E) &\to \mathbb{R} \\ f &\to \int_E f(x)\mu(dx) \end{aligned}$$

is a continuous linear map.

REMARK 2.3.38. We call the space of such signed Radon measures $\mathcal{M}^\rho(E)$. As dual space (see Definition 1.4.39) of the Banach space $\mathscr{B}^\rho(E)$ it is a Banach space itself (by Proposition 1.2.5) with the norm

$$\|\cdot\| : \mu \to \int_E \rho(x)\,|\mu|\,(dx).$$

In other words, the theorem states

$$(\mathscr{B}^\rho(E))' = \mathcal{M}^\rho(E).$$

PROOF. We start with the last part of the assertion which is much easier to show. We first note that by definition $f \in \mathscr{B}^\rho(E)$ is the pointwise limit of continuous, hence measurable functions and as such f is itself measurable (see Lemma A.3.17). Therefore, the integral

$$\int_E f(x)\mu(dx)$$

is defined. It is finite thanks to

$$\begin{aligned}\left|\int_E f(x)\mu(dx)\right| &= \left|\int_E f(x)\mu^+(dx) - \int_E f(x)\mu^-(dx)\right| \\ &\leq \int_E |f(x)|\,\mu^+(dx) + \int_E |f(x)|\,\mu^-(dx) \\ &= \int_E \frac{|f(x)|}{\rho(x)} \cdot \rho(x)\,|\mu|\,(dx) \\ &\leq \|f\|_\rho \cdot \left(\int_E \rho(x)\,|\mu|\,(dx)\right).\end{aligned}$$

This also implies that the functional is indeed continuous by Proposition 1.2.7.

As for the more difficult first part of the assertion, the proof has three steps. We first show the existence of a unique signed Radon measure μ such that for the map ℓ restricted to $C_b(E,\mathbb{R})$:

$$\ell|_{C_b(E,\mathbb{C})} : C_b(E,\mathbb{R}) \to \mathbb{R}$$

the equation

$$\ell|_{C_b(E,\mathbb{C})}(g) = \int_E g(x)\mu(dx) \text{ for all } g \in C_b(E,\mathbb{R})$$

holds true. In a second step, we prove

$$\|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})} = \int_E \rho(x)\,|\mu|\,(dx)$$

and in the third step we show

$$\ell(f) = \int_E f(x)\mu(dx) \text{ for all } f \in \mathscr{B}^\rho(E).$$

Regarding the first step, due to continuity of ℓ, by Proposition 1.2.7 there is a constant $M_\rho \geq 0$ such $|\ell(g)| \leq M_\rho \|g\|_\rho$ for all $g \in C_b(E,\mathbb{R})$. Due to Corollary 2.3.17

$$\begin{aligned}\|g\|_\rho &= \sup_{x\in E} \rho(x)^{-1}\,\|g(x)\| \\ &\leq \frac{1}{M_\rho} \cdot \underbrace{M_\rho \left(\frac{1}{\min\limits_{x\in E} \rho(x)}\right)}_{:=M_\infty} \left(\sup_{x\in E} \|g(x)\|\right).\end{aligned}$$

Hence, $|\ell(g)| \leq M_\infty \|g\|_\infty$ and again by Proposition 1.2.7 the map

$$\ell|_{C_b(E,\mathbb{C})} : C_b(E,\mathbb{R}) \to \mathbb{R}$$

is a continuous linear map as well. To this map, we would like to apply Proposition A.4.15. For any $\varepsilon > 0$, we therefore have to find a compact set K such that $|\ell(g)| \leq \varepsilon$ for all $g \in C_b(E, \mathbb{R})$, $|g| \leq 1$ for which $g = 0$ on K . Due to

$$|\ell(g)| \leq \|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})} \|g\|_\rho$$

it suffices to choose K such that the inequality

$$\|g\|_\rho \leq \frac{\varepsilon}{\|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})}} \tag{2.3.2}$$

holds for all $g \in C_b(E, \mathbb{R})$ for with $g|_K = 0$ and $|g| \leq 1$ hold. Since we know that for any such g the relation

$$\begin{aligned} \|g\|_\rho &= \sup_{x \in E} \frac{\|g(x)\|}{\rho(x)} \\ &\leq \sup_{x \in E \setminus K} \frac{1}{\rho(x)} \end{aligned}$$

holds, by definition of the admissible weight function ρ we can choose K to be

$$K := K_{\left(\frac{\|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})}}{\varepsilon}\right)} = \left\{ x \in E : \, \rho(x) \leq \frac{\|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})}}{\varepsilon} \right\}$$

in order to obtain Inequality 2.3.2 as desired. Thus, there exists a unique signed Radon measure μ such that

$$\ell(g) = \int_E g(x)\mu(dx) \text{ for all } g \in C_b(E, \mathbb{R}).$$

As second step of the proof, we show

$$\|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})} = \int_E \rho(x) \, |\mu| \, (dx).$$

Thanks to lower semicontinuity of the admissible weight function $\rho : E \to (0, \infty)$ (see Lemma 2.3.14), it is possible to apply Proposition A.4.17 in order to compute $\int_E \rho(x) \, |\mu| \, (dx)$. Application of Proposition A.4.17 yields

$$\int_E \rho(x) \, |\mu| \, (dx) = \sup_g \left| \int_E g(x)\mu(dx) \right|,$$

where the supremum is taken over all functions $g \in C_b(E, \mathbb{R})$ such that $|g| \leq \rho$, and such that g is $|\mu|$-integrable. For all these g it holds

$\|g\|_\rho \leq 1$. Hence

$$\sup_g \left| \int_E g(x)\mu(dx) \right| = \sup_g |\ell(g)| \leq \sup_g \|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})} \cdot \|g\|_\rho \leq \|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})} .$$

Regarding the other inequality, we observe that for

$$\|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})} = \sup_{f\in\mathscr{B}^\rho(E)} \|f\|_\rho^{-1} |\ell(f)|$$

it is sufficient to take the supremum over all $g \in C_b(E,\mathbb{R})$ since $C_b(E,\mathbb{R})$ is dense in $\mathscr{B}^\rho(E)$. This yields

$$\begin{aligned} \|\ell\|_{L(\mathscr{B}^\rho(E),\mathbb{R})} &= \sup_{g\in C_b(E,\mathbb{R})} \|g\|_\rho^{-1} |\ell(g)| \\ &\leq \sup_{g\in C_b(E,\mathbb{R})} \int_E \frac{|g(x)|}{\|g\|_\rho} |\mu| (dx) \\ &\leq \int_E \rho(x) |\mu| (dx) \end{aligned}$$

and we conclude the second step.

As third step, we have to show that for a continuous linear map

$$\ell : \mathscr{B}^\rho(E) \to \mathbb{R}$$

and the Radon measure μ found in the first step the equation

$$\ell(f) = \int_E f(x)\mu(dx)$$

holds for all $f \in \mathscr{B}^\rho(E)$. As shown in the beginning of this proof, for the Radon measure μ found in the first step of the proof the map

$$f \to \int_E f(x)\mu(dx)$$

is a continuous linear functional on $\mathscr{B}^\rho(E)$. By construction, it coincides with the continuous linear map ℓ on the dense subset $C_b(E,\mathbb{R})$. Hence they coincide on all of $\mathscr{B}^\rho(E)$.

□

In the case of positive linear maps $\mathscr{B}^\rho(E) \to \mathbb{R}$ we always obtain a statement like in the theorem above and do not need to check for continuity first. This is shown in the Proposition 2.3.41 which was proved in [**15**]. We first show that if the continuous functional in Theorem 2.3.37 is positive, then its Radon measure given by Theorem 2.3.37 is positive as well.

DEFINITION 2.3.39. Let A and B be some sets and call $\mathbb{R}^A$ the set of maps between A and $\mathbb{R}$. For $f \in \mathbb{R}^A$ we write $f \geq 0$ if $f(a) \geq 0$ for all $a \in A$. A map

$$T : \mathbb{R}^A \to \mathbb{R}^B$$

such that $f \geq 0$ implies $T(f) \geq 0$ is called *positive*.

COROLLARY 2.3.40. *If the continuous linear map $\ell : \mathscr{B}^\rho(E) \to \mathbb{R}$ is positive, then the unique finite Radon measure $\mu \in \mathcal{M}^\rho(E)$ from Theorem2.3.37 is positive.*

PROOF. By Theorem 2.3.37 we know that there exists a unique signed Radon measure μ such that Equation 2.3.1 and Equation 2.3.1 hold. Assume by contradiction that μ is not positive. If $\mu(O) \geq 0$ would hold for any open set $O \subset Y$, then by outer regularity of Radon measures (Lemma 2.3.34) μ were positive. Thus, there must be an open set $O \subset Y$ and $\varepsilon > 0$ such that $\mu(O) < -\varepsilon < 0$. By inner regularity of Radon measures choose $K_{\varepsilon/2} \subset O$ compact such that $|\mu|(O \setminus K_{\varepsilon/2}) < \frac{\varepsilon}{2}$. Then $-\frac{\varepsilon}{2} < \mu(O \setminus K_{\varepsilon/2})$ implies $\mu(K_{\varepsilon/2}) < -\frac{\varepsilon}{2}$. By Proposition 2.3.9 there exists a continuous function $g : E \to [0,1]$ such that $g(O) = 0$ and $g(K_{\varepsilon/2}) = 1$. We obtain

$$\begin{aligned}
\ell(g) &= \int_E g(x)\mu(dx) \\
&= \int_{O\setminus K_{\varepsilon/2}} g(x)\mu(dx) + \int_{K_{\varepsilon/2}} g(x)\mu(dx) \\
&\leq |\mu|(O \setminus K_{\varepsilon/2}) + \mu(K_{\varepsilon/2}) \\
&< \frac{\varepsilon}{2} - \frac{\varepsilon}{2} = 0
\end{aligned}$$

in contradiction to positivity of ℓ. Hence, the statement of the corollary holds. □

PROPOSITION 2.3.41.
Let $\ell : \mathscr{B}^\rho(E) \to \mathbb{R}$ be a positive linear map. Then there exists a unique finite Radon measure μ on E such that

$$\ell(f) = \int_E f(x)\mu(dx) \textit{ for all } f \in \mathscr{B}^\rho(E).$$

Additionally, μ is positive and $\mu \in \mathcal{M}^\rho(E)$.

PROOF. By Corollary 2.3.40 we only need to show that

$$\ell : \mathscr{B}^\rho(E) \to \mathbb{R}$$

is continuous. By contradiction, assume that this is not the case. Then by Proposition 1.2.7 there exists a sequence $(f_n)_{n\in\mathbb{N}}$ in $\mathscr{B}^\rho(E)$ such that $\|f_n\|_\rho = 1$ and $|\ell(f_n)| \geq n^3$ for all $n \in \mathbb{N}$. Hence either $\ell(-f_n) \geq n^3$ or $\ell(f_n) \geq n^3$ for all $n \in \mathbb{N}$ and positivity of ℓ implies that for all $n \in \mathbb{N}$ the inequalities

$$\ell(|f_n|) \geq \ell(-f_n)$$

and

$$\ell(|f_n|) \geq \ell(f_n)$$

hold. Thus, for all $n \in \mathbb{N}$ for $g_n := |f_n|$ we obtain $\|g_n\|_\rho = 1$ and

$$|\ell(g_n)| = \ell\left(|f_n|\right) \geq n^3.$$

In order to obtain a contradiction, we now construct a map $h \in \mathscr{B}^\rho(E)$ such that $\ell(h)$ is not defined anymore. This is done by setting

$$h := \sum_{n=1}^{\infty} \frac{g_n}{n^2}.$$

We see that $(h_m)_{m\in\mathbb{N}}$ given by

$$h_m := \sum_{n=1}^{m} \frac{g_n}{n^2}$$

converges to h in $\mathscr{B}^\rho(E)$. Furthermore $h_m, h > 0$ and $h - h_m > 0$ for all $m \in \mathbb{N}$ and positivity of ℓ yields $\ell(h) \geq \ell\left(h_m\right)$ for all $m \in \mathbb{N}$. The inequality

$$\ell\left(h_m\right) = \sum_{n=1}^{m} \frac{\ell\left(g_n\right)}{n^2} \geq \sum_{n=1}^{m} n$$

implies $\ell(h) \geq \lim_{m\to\infty} \ell\left(h_m\right) = \infty$ and we obtain that $\ell(h)$ is not defined in contradiction to our assumption. □

Investigating the space $\mathscr{B}^\rho(E)$ further, we obtain the following characterization. The proof is again based on [**15**].

THEOREM 2.3.42. *Let $f : E \to \mathbb{R}$. Then $f \in \mathscr{B}^\rho(E)$ if and only if (i) for all $R > 0$*

$$f|_{K_R} \in C_b(K_R, \mathbb{R}),$$

and
(ii)

$$\lim_{R\to\infty} \sup_{x\in E\setminus K_R} \frac{|f(x)|}{\rho(x)} = 0.$$

PROOF. We first show that $f \in \mathscr{B}^\rho(E)$ implies (i) and (ii).

We start with (i). So let $f \in \mathscr{B}^\rho(E)$ and fix $R > 0$. Then by density of $C_b(E,\mathbb{R})$ in $\mathscr{B}^\rho(E)$ there exists $(g_n)_{n\in\mathbb{N}} \subset C_b(E,\mathbb{R})$ such that

$$\|g_n - f\|_\rho < \frac{1}{n}$$

for any $n \in \mathbb{N}$. Therefore,

$$\begin{aligned}\sup_{x\in K_R} |g_n(x) - f(x)| &\leq R \cdot \sup_{x\in K_R} \frac{|g_n(x) - f(x)|}{\rho(x)} \\ &\leq R \cdot \frac{1}{n}.\end{aligned}$$

Hence on K_R

$$\left(g_n|_{K_R}\right)_{n\in\mathbb{N}} \subset C_b(K_R,\mathbb{R})$$

converges uniformly to f which implies that $f|_{K_R} \in C_b(K_R,\mathbb{R})$ which is a well known result from analysis.

As for $f \in \mathscr{B}^\rho(E)$ implying (ii), let $\delta > 0$ be arbitrary. We have to show that for R large enough the inequality

$$\sup_{x\in E\setminus K_R} \frac{|f(x)|}{\rho(x)} < \delta$$

holds. By density of $C_b(E,\mathbb{R})$ in $\mathscr{B}^\rho(E)$, choose $h \in C_b(E,\mathbb{R})$ such that

$$\|h - f\|_\rho < \frac{\delta}{2}.$$

Then

$$\begin{aligned}\sup_{x\in E\setminus K_R} \frac{|f(x)|}{\rho(x)} &\leq \sup_{x\in E\setminus K_R} \frac{|f(x) - h(x)|}{\rho(x)} + \sup_{x\in E\setminus K_R} \frac{|h(x)|}{\rho(x)} \\ &\leq \frac{\delta}{2} + \sup_{x\in E\setminus K_R} \frac{|h(x)|}{\rho(x)}.\end{aligned}$$

Since $h \in C_b(E,\mathbb{R})$ is bounded by $\|h\|_\infty$ is suffices to choose R such that

$$R > \frac{2\,\|h\|_\infty}{\delta}.$$

Next, we let $f : E \to \mathbb{R}$ and show that Properties (i) and (ii) imply $f \in \mathscr{B}^\rho(E)$. We will do this by constructing a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathscr{B}^\rho(E)$ that converges to f in $\mathscr{B}^\rho(E)$. For such a sequence $(f_n)_{n\in\mathbb{N}} \subset \mathscr{B}^\rho(E)$ that we have to construct, for any $n \in \mathbb{N}$ it has to be possible to approximate f_n in $\mathscr{B}^\rho(E)$ by continuous bounded functions. Therefore, it seems reasonable to investigate the candidate

$$g_n(\cdot) := \min\left(\left(\max\left(f(\cdot), -n\right)\right), n\right)$$

for any $n \in \mathbb{N}$, thereby already ensuring boundedness of all $(g_n)_{n\in\mathbb{N}}$. In order to show that the sequence $(g_n)_{n\in\mathbb{N}}$ possesses the desired properties, we have to prove $(g_n)_{n\in\mathbb{N}} \subset \mathscr{B}^\rho(E)$ and $\lim_{n\to\infty} g_n = f$ in $\mathscr{B}^\rho(E)$.

Concerning the proof of $(g_n)_{n\in\mathbb{N}} \subset \mathscr{B}^\rho(E)$, we fix $n \in \mathbb{N}$ and by Property (i) of our assumptions we obtain

$$g_n|_{K_R} \in C_b(K_R, \mathbb{R})$$

for any $R > 0$, $R \in \mathbb{N}$. By Proposition 2.3.8, for any $R > 0$, $R \in \mathbb{N}$ we can find a continuous extension $h_{n,R}$ of $g_n|_{K_R}$ on all of X such that $|h_{n,R}| \leq n$. $h_{n,R}$ approximates g_n in $\mathscr{B}^\rho(E)$ as $R > 0$, $R \in \mathbb{N}$ tends to infinity because

$$\begin{aligned} \lim_{R\to\infty} \|h_{n,R} - g_n\|_\rho &= \lim_{R\to\infty} \sup_{x\in E\setminus K_R} \frac{|h_{n,R}(x) - g_n(x)|}{\rho(x)} \\ &\leq \lim_{R\to\infty} \frac{2n}{R} \\ &= 0. \end{aligned}$$

Thus $g_n \in \mathscr{B}^\rho(E)$ for any $n \in \mathbb{N}$.

In order to show

$$\lim_{n\to\infty} g_n = f \text{ on } \mathscr{B}^\rho(E),$$

for any $\varepsilon > 0$ by Property (ii) we can choose $R > 0$, $R \in \mathbb{N}$ such that

$$\sup_{x\in E\setminus K_R} \frac{2\,|f(x)|}{\rho(x)} < \varepsilon.$$

We then choose $N \in \mathbb{N}$ such that

$$N > \max_{x\in K_R} |f(x)|\,,$$

which is possible because a continuous function attains its maximum on a compact set. Then, $g_n = f$ on K_R for any $n \in \mathbb{N}$, $n > N$ and

$$\begin{aligned} \|f - g_n\|_\rho &= \sup_{x\in E\setminus K_R} \frac{|f(x) - g_n(x)|}{\rho(x)} \\ &\leq \sup_{x\in E\setminus K_R} \frac{2\,|f(x)|}{\rho(x)} \\ &\leq \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this yields $\lim_{n\to\infty} g_n = f$ in $\mathscr{B}^\rho(E)$ and therefore $f \in \mathscr{B}^\rho(E)$. □

The following property is similar to the fact that on a compact set continuous functions attain their maximum. The proof follows **[15]**.

THEOREM 2.3.43. *Let $f \in \mathscr{B}^\rho(E)$. If*

$$\sup_{x\in E} f(x) > 0,$$

then there exists $z \in E$ such that for all $x \in E$

$$\frac{f(x)}{\rho(x)} \leq \frac{f(z)}{\rho(z)}.$$

PROOF. Due to

$$\sup_{x\in E} f(x) > 0$$

we choose some $y \in E$ such that $f(y) > 0$ and by Theorem 2.3.42 we can choose R large enough such that

$$\sup_{x\in E\setminus K_R} \frac{|f(x)|}{\rho(x)} < \frac{f(y)}{\rho(y)}.$$

In order to estimate the supremum of $\frac{f}{\rho}$ on the compact set K_R, we use that by Lemma 2.3.14 ρ is lower semicontinuous and by Lemma A.2.1 $\frac{1}{\rho}$ is upper semicontinuous. By Theorem 2.3.37 f is continuous on K_R. Thus, $\frac{1}{\rho} \cdot f$ is upper semicontinuous as well and attains its maximum on the compact set K_R at some point $z \in K_R$. This yields the statement of the theorem. □

COROLLARY 2.3.44. *Let $f \in \mathscr{B}^\rho(E)$. Then, there exists $z \in E$ such that $\|f\|_\rho = \frac{|f(z)|}{\rho(z)}$.*

PROOF. Set $g := |f| \in \mathcal{B}^\rho(E)$. By the previous Theorem there exists $z \in E$ such that $\|g\|_\rho = \frac{g(z)}{\rho(z)}$ and we conclude by substituting back in. □

Regarding maps defined on $\mathscr{B}^\rho(E)$, a composition with a bounded continuous map is continuous:

LEMMA 2.3.45. *Let $h \in C_b(\mathbb{R})$ and $f \in \mathscr{B}^\rho(E)$. Then*

$$\begin{aligned} \mathscr{B}^\rho(E) &\to \mathscr{B}^\rho(E) \\ f &\to h \circ f \end{aligned}$$

is a continuous map.

PROOF. Since $h \circ f|_{K_R}$ is continuous for any $R > 0$, by Theorem 2.3.42 $h \circ f \in \mathscr{B}^\rho(E)$. Thus, we only need to show continuity. Let

$g \in \mathscr{B}^\rho(E)$, $\varepsilon > 0$ and choose $R_\varepsilon > \frac{\|h\|_\infty}{\varepsilon}$. Then

$$\|h \circ f - h \circ g\|_\rho \leq \varepsilon + \left\| \frac{h \circ f - h \circ g}{\rho} \Big|_{K_{R_\varepsilon}} \right\|_\infty$$
$$\leq \varepsilon + \frac{1}{\inf\limits_{x \in E} \rho(x)} \cdot \left\| (h \circ f - h \circ g)|_{K_{R_\varepsilon}} \right\|_\infty .$$

Let $[a, b]$ be some interval such that $f(K_{R_\varepsilon}) \subset [a, b]$ (which is possible since on K_{R_ε} f is continuous and attains maximum and minimum.). Then h is uniformly continuous on $[a-1, b+1]$ and there is $\delta > 0$ such that for any $x_1, x_2 \in [a-1, b+1]$ such that $|x_1 - x_2| < \delta$ it holds $|h(x_1) - h(x_2)| < \varepsilon$. Thus, if $\|f - g\|_\rho < \frac{\delta}{R_\varepsilon}$ then

$$\left\| (f-g)|_{K_{R_\varepsilon}} \right\|_\infty \leq \|f - g\|_\rho \cdot R_\varepsilon < \delta,$$

and consequently

$$\left\| (h \circ f - h \circ g)|_{K_{R_\varepsilon}} \right\|_\infty < \varepsilon,$$

which shows continuity of $f \to h \circ f$. □

Theorem 2.3.42 shows that the space $\mathscr{B}^\rho(E)$ is closely related to the space of continuous maps on a compact space. For such spaces the Stone-Weierstrass theorem (Theorem A.2.6) holds. We show that a version of it also holds for $\mathscr{B}^\rho(E)$. For the definition of an algebra we refer the reader to Definition A.2.5.

PROPOSITION 2.3.46. *(Stone-Weierstraß for $\mathscr{B}^\rho(E)$)*
Let $A \subset C_b(E)$ be an algebra with respect to pointwise multiplication that contains 1_E and that separates points. Then A is dense in $\mathscr{B}^\rho(E)$ with respect to $\|\cdot\|_\rho$.

PROOF. The idea of the proof is to approximate elements of $\mathscr{B}^\rho(E)$ by continuous bounded maps on E which in turn can be approximated on K_R for any $R > 0$ via Stone-Weierstrass by elements in A that are restricted to K_R. However, such an element in A, albeit bounded, may have an arbitrary large bound that depends on R. Thus, it may not approximate with respect to $\|\cdot\|_\rho$ on all of E. Therefore, it is rescaled by a suitable polynomial such that an element in A is obtained whose bounds do not depend on R. This yields an approximation with respect to $\|\cdot\|_\rho$.

That idea will be carried out in the following. Let $h \in \mathscr{B}^\rho(E)$ and $\varepsilon > 0$. By definition of $\mathscr{B}^\rho(E)$ there exists $g_\varepsilon \in C_b(E)$ such that

$$\|g_\varepsilon - h\|_\rho < \varepsilon.$$

Set

$$R_\varepsilon := \max\left(\frac{\|g_\varepsilon\|_\infty}{\varepsilon}, 1\right).$$

The set $A_\varepsilon \subset C_b(K_{R_\varepsilon})$ defined as

$$A_\varepsilon := \left\{ f|_{K_{R_\varepsilon}} \,:\, f \in A \right\}$$

is an algebra that contains $1_{K_{R_\varepsilon}}$ and that separates points, hence by Stone-Weierstrass (Theorem A.2.6) A_ε is dense in $C(K_{R_\varepsilon})$. Thus, there is $f_\varepsilon \in A$ such that

$$\sup_{x \in K_{R_\varepsilon}} |f_\varepsilon(x) - g_\varepsilon(x)| < \varepsilon.$$

Clearly,

$$\alpha_\varepsilon := \sup_{x \in K_{R_\varepsilon}} |f_\varepsilon(x)| \leq \sup_{x \in E} |g_\varepsilon(x)| + \varepsilon =: \beta_\varepsilon.$$

Set

$$\gamma_\varepsilon := \sup_{x \in E} |f_\varepsilon(x)|.$$

By Tietze-Urysohn (Theorem 2.3.6) there exists a continuous map

$$\varphi_\varepsilon : [-\gamma_\varepsilon, \gamma_\varepsilon] \to [-\beta_\varepsilon, \beta_\varepsilon]$$

such that

$$\varphi_\varepsilon(y) = \begin{cases} y & \text{for } y \in [-\alpha_\varepsilon, \alpha_\varepsilon] \\ \beta_\varepsilon & \text{for } \ |y| \geq \beta_\varepsilon. \end{cases}$$

Again by Stone-Weierstrass, on a compact set the space of polynomials is dense in the space of continuous maps. This means that there is a polynomial p_ε on $[-\gamma_\varepsilon, \gamma_\varepsilon]$ such that

$$\sup_{y \in [-\gamma_\varepsilon, \gamma_\varepsilon]} |p_\varepsilon(y) - \varphi_\varepsilon(y)| < \varepsilon,$$

hence

$$\sup_{x \in E} \left| \frac{(p_\varepsilon \circ f_\varepsilon)(x) - (\varphi_\varepsilon \circ f_\varepsilon)(x)}{\rho(x)} \right| \leq \frac{\varepsilon}{\inf\limits_{x \in E} \rho(x)}.$$

Since A is an algebra $p_\varepsilon \circ f_\varepsilon \in A$ and

$$\begin{aligned}
\|h - p_\varepsilon \circ f_\varepsilon\|_\rho &\leq \|h - g_\varepsilon\|_\rho + \|g_\varepsilon - \varphi_\varepsilon \circ f_\varepsilon\|_\rho + \|\varphi_\varepsilon \circ f_\varepsilon - p_\varepsilon \circ f_\varepsilon\|_\rho \\
&\leq \varepsilon + \sup_{x \in K_{R_\varepsilon}} \left| \frac{g_\varepsilon(x) - (\varphi_\varepsilon \circ f_\varepsilon)(x)}{\rho(x)} \right| \\
&\quad + \sup_{x \in E \setminus K_{R_\varepsilon}} \left| \frac{g_\varepsilon(x) - (\varphi_\varepsilon \circ f_\varepsilon)(x)}{\rho(x)} \right| + \frac{\varepsilon}{\inf\limits_{x \in E} \rho(x)} \\
&\leq \varepsilon + \sup_{x \in K_{R_\varepsilon}} \left| \frac{g_\varepsilon(x) - f_\varepsilon(x)}{\rho(x)} \right| \\
&\quad + 2 \sup_{x \in E} \left| \frac{|g_\varepsilon(x)| + \varepsilon}{R_\varepsilon} \right| + \frac{\varepsilon}{\inf\limits_{x \in K_{R_\varepsilon}} \rho(x)} \\
&\leq \varepsilon + \frac{\varepsilon}{\inf\limits_{x \in K_{R_\varepsilon}} \rho(x)} + 2(\varepsilon + \varepsilon) + \frac{\varepsilon}{\inf\limits_{x \in K_{R_\varepsilon}} \rho(x)},
\end{aligned}$$

and A is dense in $\mathscr{B}^\rho(E)$. □

We recall the space $C_0(E, \mathbb{K})$. It is defined as the set of continuous $\mathbb{K}$-valued functions on E such that $\{x \in E : |f(x)| \geq \varepsilon\}$ is compact for any $\varepsilon > 0$. It is equipped with the norm (see Proposition A.4.4)

$$\|\cdot\|_\infty : f \to \sup_{x \in E} \|f(x)\| .$$

One can easily show that this is a Banach space (see Lemma A.4.5). We set $C_0(E) := C_0(E, \mathbb{R})$.

LEMMA 2.3.47. *If the admissible weight function ρ is continuous, then*
(i) $\mathscr{B}^\rho(E) \subset C(E)$,
(ii) $f \in C_0(E)$ implies $f \cdot \rho \in \mathscr{B}^\rho(E)$,
(iii) $f \in \mathscr{B}^\rho(E)$ implies $\frac{f}{\rho} \in C_0(E)$.

PROOF. (i) For $f \in \mathscr{B}^\rho(E)$ by definition of $\mathscr{B}^\rho(E)$ $\frac{f}{\rho}$ is the uniform limit of $\left(\frac{g_n}{\rho}\right)_{n \in \mathbb{N}}$ for some $(g_n)_{n \in \mathbb{N}} \subset C_b(E)$. Hence $\frac{f}{\rho}$ is continuous and therefore also f.

(ii) If $f \in C_0(E)$, then $f \cdot \rho$ is continuous and $\bigcup_{n \in \mathbb{N}} \{\rho < n\}$ is an open cover of E hence for any $\varepsilon > 0$ finitely many such sets suffice to cover the compact set $\{|f| \geq \varepsilon\}$. Thus, for any $\varepsilon > 0$ there exists $R_\varepsilon > 0$ such that $|f| < \varepsilon$ on $E \setminus K_{R_\varepsilon}$ and by Theorem 2.3.42 $f \cdot \rho \in \mathscr{B}^\rho(E)$.

(iii) By (i) $\frac{f}{\rho}$ is continuous. By Theorem 2.3.42 for any $\varepsilon > 0$ there is some $R'_\varepsilon > 0$ such that $\left\{\left|\frac{f}{\rho}\right| \geq \varepsilon\right\} \subset K_{R'_\varepsilon}$. Hence by closedness

$\left\{\left|\frac{f}{\rho}\right| \geq \varepsilon\right\}$ is compact and

$$\frac{f}{\rho} \in C_0(E).$$

□

In the next lemma, $C_c(E)$ denotes the continuous maps with compact support, which are the continuous maps f such that

$$\overline{\{x \in E : f(x) \neq 0\}}$$

is compact.

LEMMA 2.3.48. *Let (E, ρ) be a weighted space and E be locally compact. Then $C_c(E)$ is dense in $\mathscr{B}^\rho(E)$ with respect to $\|\cdot\|_\rho$.*

PROOF. By definition of $\mathscr{B}^\rho(E)$ we only need to show that $C_c(E)$ is dense in $C_b(E)$ with respect to $\|\cdot\|_\rho$. Let $f \in C_b(E)$ and $\varepsilon > 0$. Choose $R_\varepsilon := \frac{\|f\|_\infty}{\varepsilon}$. By local compactness each element in K_{R_ε} has a compact neighborhood hence by compactness of K_{R_ε} finitely many such compact neighborhood cover K_{R_ε}. The union of these finitely many neighborhoods is a compact neighborhood $U_{K_{R_\varepsilon}} \supset K_{R_\varepsilon}$. Let $V_{K_{R_\varepsilon}} \subset U_{K_{R_\varepsilon}}$ be an open neighborhood of K_{R_ε} and define the map $\tilde{g}_\varepsilon \in C_b(K_{R_\varepsilon} \cup U_{K_{R_\varepsilon}} \setminus V_{K_{R_\varepsilon}})$ as

$$\tilde{g}_\varepsilon := \begin{cases} f & \text{on } K_{R_\varepsilon} \\ 0 & \text{on } U_{K_{R_\varepsilon}} \setminus V_{K_{R_\varepsilon}}. \end{cases}$$

By normality of compact sets (Proposition A.1.11) and Tietze-Urysohn (Theorem 2.3.6) this map can be extended to $g'_\varepsilon \in C_b(U_{K_{R_\varepsilon}})$ such that $\|g'_\varepsilon\|_\infty = \|f\|_\infty$. Subsequently the map g'_ε can be extended to $g_\varepsilon \in C_c(E)$ with $\|g_\varepsilon\|_\infty = \|f\|_\infty$ by stetting $g_\varepsilon \equiv 0$ on $E \setminus U_{K_{R_\varepsilon}}$. Then

$$\begin{aligned} \|g_\varepsilon - f\|_\rho &\leq \sup_{x \in K_{R_\varepsilon}} \frac{|g_\varepsilon(x) - f(x)|}{\rho(x)} + \sup_{x \in E \setminus K_{R_\varepsilon}} \frac{|g_\varepsilon(x) - f(x)|}{\rho(x)} \\ &\leq 0 + \frac{2\|f\|_\infty}{R_\varepsilon} \\ &= 2\varepsilon, \end{aligned}$$

which proves the lemma since $\varepsilon > 0$ was arbitrary. □

2.3.4. Generalized Feller Semigroups. As before, in this subsection (E, ρ) always denotes a weighted space. Since we have seen that $X := \mathscr{B}^\rho(E)$ is a Banach space we can define one-parameter semigroups on it (see Definition 1.3.1). In a special setting this was done by

Röckner and Sobol [**36**] in 2006. Generalizing this idea, in 2010 Dörsek and Teichmann [**15**] introduced *generalized Feller semigroups*.

DEFINITION 2.3.49. Let $(P(t))_{t\in\mathbb{R}_+}$ be a family of bounded linear operators such that for any $t \in \mathbb{R}_+$

$$P(t) : \mathscr{B}^\rho(E) \to \mathscr{B}^\rho(E).$$

We call the family $(P(t))_{t\in\mathbb{R}_+}$ *generalized Feller semigroup* on $\mathscr{B}^\rho(E)$ if

P1 $P(0) = \mathrm{Id}$, where Id is the identity on $\mathscr{B}^\rho(E)$,
P2 $P(t+s) = P(s) \circ P(t)$ for all $s, t \in \mathbb{R}_+$,
P3 for all $f \in \mathscr{B}^\rho(E)$ and all $x \in E$

$$\lim_{t\searrow 0} P(t)f(x) = f(x),$$

P4 there exists $\varepsilon > 0$ and $C \in \mathbb{R}$ such that for all $t \in [0, \varepsilon]$

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))} \leq C,$$

and
P5 $P(t)$ is positive (in the sense of Definition 2.3.39) for all $t \in \mathbb{R}_+$.

REMARK 2.3.50. Compared with the definition of Feller semigroups (see Definition 2.2.1) in the definition of generalized Feller semigroups $C_0(E)$ is replaced by $\mathscr{B}^\rho(E)$ and instead of

$$\|Q(t)\|_{L(C_0(E))} \leq 1$$

in the case of Feller semigroups only **P4** is asked in the case of generalized Feller semigroups. Strong continuity and positivity are properties that Feller semigroups and generalized Feller semigroups have in common (for strong continuity and generalized Feller semigroups see Theorem 2.3.51).

Furthermore, if E is compact, then choosing$\rho = 1$

$$\mathscr{B}^\rho(E) = C_b(E) = C_0(E).$$

With Chapter 1 in mind, we are interested in the continuity properties of generalized Feller semigroups. It is proved in [**15**]:

THEOREM 2.3.51. *Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$. Then $(P(t))_{t\in\mathbb{R}_+}$ is strongly continuous on $\mathscr{B}^\rho(E)$.*

PROOF. The proof is possible thanks to the deep result from Chapter 1, Subsection 1.4.6 that strong continuity of a semigroup follows from weak continuity (see Corollary 1.4.69).

Therefore, it is enough to show that for any $f \in \mathscr{B}^\rho(E)$ and any $\mu \in (\mathscr{B}^\rho(E))'$ we obtain the right limit

$$\lim_{t\searrow 0} \langle P(t)f, \mu\rangle = \langle f, \mu\rangle .$$

We fix some $f \in \mathscr{B}^\rho(E)$ and $\mu \in (\mathscr{B}^\rho(E))'$. By Theorem 2.3.37

$$(\mathscr{B}^\rho(E))' = \mathcal{M}^\rho(E)$$

and

$$\langle P(t)f, \mu\rangle = \int_E P(t)f(x)\mu(dx).$$

Thus, the theorem follows immediately if we can show that

$$\lim_{t\to 0} \int_E \left(P(t)f(x) - f(x)\right)\mu(dx) = 0$$

holds true. By **P3,** for all $x \in E$ we are given the right limit

$$\lim_{t\searrow 0} P(t)f(x) = f(x).$$

Hence, we obtain the above limit by dominated convergence (Theorem A.3.58) if for some $t_0 > 0$ we can bound

$$(|P(t)f - f|)_{t\in\mathbb{R}_+,\, t<t_0}$$

by a μ- integrable function. By **P4** we obtain for $t < t_0 < \varepsilon$ the bounds

$$\begin{aligned}
|P(t)f(x) - f(x)| &\leq \sup_{x\in E} \frac{|P(t)f(x) - f(x)|}{\rho(x)}\rho(x) \\
&= \|P(t)f - f\|_\rho\, \rho(x) \\
&\leq \left(\|P(t)\|_{L(\mathscr{B}^\rho(E))} + \|\mathrm{Id}\|_{L(\mathscr{B}^\rho(E))}\right) \|f\|_\rho\, \rho(x) \\
&\leq (C+1)\, \|f\|_\rho\, \rho(x),
\end{aligned}$$

and by Theorem 2.3.37 $(C+1)\,\|f\|_\rho\,\rho(x)$ is indeed integrable with respect to μ. Hence, it is justified to apply the dominated convergence theorem which yields

$$\lim_{t\to 0} \int_E \left(P(t)f(x) - f(x)\right)\mu(dx) = 0.$$

Since $f \in \mathscr{B}^\rho(E)$ and $\mu \in (\mathscr{B}^\rho(E))'$ were arbitrary, we obtain the statement of the theorem. □

Since we know the dual space of $\mathscr{B}^\rho(E)$, we can connect generalized Feller semigroups to a family of positive finite Radon measures on $(E, \mathcal{B}(E))$. Furthermore,with respect to the Baire σ-algebra $\mathcal{B}_0(E)$ (see Definition A.3.38) we even obtain a semigroup of transition probabilities.

PROPOSITION 2.3.52. *Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ such that for any $t \in \mathbb{R}_+$*

$$\|P(t)\| \leq M e^{\omega t}$$

for some $M \geq 1$ and $\omega \in \mathbb{R}$.
(i) There exists a unique family of positive finite Radon measures

$$(p(t)(x,\cdot))_{t\in\mathbb{R}_+, x\in E}$$

on $(E, \mathcal{B}(E))$ such that for all $x \in E$, $t \in \mathbb{R}_+$ and $f \in \mathscr{B}^\rho(E)$

$$P(t)f(x) = \int_E f(y)p(t)(x,dy),$$

and $p(t)(x,\cdot) \in \mathcal{M}^\rho(E)$.
(ii) For all $x \in E$ and $t \in \mathbb{R}_+$ one defines for all positive measurable maps $f : E \to \mathbb{R}$ (or $f \in \widetilde{B^\rho}(E)$)

$$\tilde{P}(t)f(x) := \int_E f(y)p(t)(x,dy),$$

and obtains the bounds

$$\tilde{P}(t)\rho(x) = \sup_{\substack{f \in C_b(E) \\ |f| \leq \rho}} |(P(t)f)(x)| \leq \rho(x) \|P(t)\|_{L(\mathcal{B}^\rho(E))} \leq \rho(x) M e^{\omega t}.$$

Hence, for all $t \in \mathbb{R}_+$ $P(t)$ can also be interpreted as a bounded linear operator on

$$\widetilde{B^\rho}(E) := \left\{ f : E \to Z : \sup_{x\in E} \rho(x)^{-1} \|f(x)\| < \infty, f \text{ measurable} \right\},$$

which will be done frequently without further mention. Furthermore,

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))} = \|P(t)\|_{L\left(\widetilde{B^\rho}(E)\right)}.$$

(iii) If additionally E is locally compact (Definition A.1.12) with countable base (Definition A.1.2), then $(p(t))_{t\in\mathbb{R}_+}$ is a semigroup of transition kernels.
(iv) The family $(\hat{p}(t))_{t\in\mathbb{R}_+}$ defined as the restriction

$$\hat{p}(t) = p(t)|_{E\times\mathcal{B}_0(E)}$$

for any $t \in \mathbb{R}_+$ is a semigroup of transition kernels with respect to the Baire σ-algebra $\mathcal{B}_0(E)$.

REMARK 2.3.53. If the family of positive finite Radon measures $(p(t))_{t\in\mathbb{R}_+}$ on $(E,\mathcal{B}(E))$ permits to find a projective family of probability measures on $\left(E^J,\mathcal{B}(E)^J\right)_{J\subset\mathbb{R}_+,\,\text{finite}}$, then in a similar way as in Proposition 2.1.13 one obtains the existence of a stochastic process such that the conditional expectation is of a particular form. Results when this is the case will be presented later in Theorem 2.3.65 and Theorem 2.3.73.

PROOF. (i) By definition of generalized Feller semigroups, for any $t\in\mathbb{R}_+$ and $x\in E$ the map

$$\begin{aligned}\mathscr{B}^\rho(E)&\to\mathbb{R}\\ \ell_{t,x}:\, f&\to P(t)f(x)\end{aligned}$$

is positive and linear and by strong continuity of $(P(t))_{t\in\mathbb{R}_+}$ (see Theorem 2.3.51) and exponential boundedness of strongly continuous semigroups (see Proposition 1.4.9) there exist constants $\omega\in\mathbb{R}$ and $M\geq 1$ such that

$$|P(t)f(x)|\leq\rho(x)\,\|P(t)f\|_\rho\leq\rho(x)Me^{\omega t}\cdot\|f\|_\rho$$

holds true. Hence

$$\begin{aligned}\mathscr{B}^\rho(E)&\to\mathbb{R}\\ \ell_{t,x}:\, f&\to P(t)f(x)\end{aligned}$$

is also continuous. Thus, by Corollary 2.3.40 for any $t\in\mathbb{R}_+$ and any $x\in E$ there is a unique positive finite Radon measure $p(t)(x,\cdot)\in\mathcal{M}^\rho(E)$ such that

$$\left(P(t)f\right)(x)=\int_E f(y)p(t)(x,dy)$$

holds true.

(ii) $P(t)f$ is clearly well-defined. Since $p(t)(x,\cdot)\in\mathcal{M}^\rho(E)$ we observe

$$\int_E\rho(y)p(t)(x,dy)=\|\ell_{t,x}\|=\sup_{\substack{f\in C_b(E)\\ |f|\leq\rho}}\left|\left(P(t)f\right)(x)\right|.$$

The calculation

$$\begin{aligned}\sup_{\substack{f\in C_b(E)\\ |f|\le\rho}} |(P(t)f)(x)| &\le \rho(x) \sup_{\substack{f\in C_b(E)\\ |f|\le\rho}} \left(\sup_{x\in E}\frac{|(P(t)f)(x)|}{\rho(x)}\right)\\ &\le \rho(x) \sup_{\substack{f\in C_b(E)\\ |f|\le\rho}} \|P(t)f\|_\rho\\ &\le \rho(x)\,\|P(t)\|_\rho\end{aligned}$$

yields the statement of the Lemma.

(iii) In the first step, we show that $(p(t))_{t\in\mathbb{R}_+}$ is a family of transition kernels and only have to show that for any $A\in\mathcal{B}(E)$ and any $t\in\mathbb{R}_+$

$$x\to p(t)(x,A)$$

is measurable. Since E is locally compact with countable base, it is also metrizable (see Proposition A.1.13). Let d be a metric and for $x\in E$ and $\varepsilon>0$ let $B_\varepsilon(z):=\{y:\, d(z,y)<\varepsilon\}$ be an open ball.

We can approximate

$$1_{B_\varepsilon(z)}(y)=\lim_{n\to\infty}1\wedge\left(n\cdot\inf_{\widetilde{z}\notin B_\varepsilon(z)}d(y,\widetilde{z})\right).$$

(the infimum is strictly positive due to $E\setminus B_\varepsilon(z)$ being closed). Then by dominated convergence

$$\begin{aligned}p(t)(x,B_\varepsilon(z)) &= \lim_{n\to\infty}\int_E\left(1\wedge\left(n\cdot\inf_{\widetilde{z}\notin B_\varepsilon(z)}d(y,\widetilde{z})\right)\right)p(t)(x,dy)\\ &= \lim_{n\to\infty}P(t)\left(1\wedge\left(n\cdot\inf_{\widetilde{z}\notin B_\varepsilon(z)}d(y,\widetilde{z})\right)\right)(x),\end{aligned}$$

and $x\to p(t)(x,B_\varepsilon(z))$ is measurable as limit of measurable maps since

$$1\wedge\left(n\cdot\inf_{\widetilde{z}\notin B_\varepsilon(z)}d(y,\widetilde{z})\right)\in\mathscr{B}^\rho(E).$$

Since E is separable, there exists a sequence $(\widetilde{z}_n)_{n\in\mathbb{N}}\subset E$ such that $(B_\varepsilon(\widetilde{z}_i))_{i\in\mathbb{N},\,\varepsilon\in\mathbb{Q}}$ forms a countable base of $\mathcal{B}(E)$. As before

$$p(t)\left(x,\bigcup_{i\in\{1,\dots,m\}}B_\varepsilon(\widetilde{z}_i)\right)=\lim_{n\to\infty}P(t)\left(\sup_{i\in\{1,\dots,m\}}\left(1\wedge\left(n\cdot\inf_{\widetilde{z}_i\notin B_\varepsilon(\widetilde{z}_i)}d(y,\widetilde{z}_i)\right)\right)\right)(x),$$

and by taking the limit also $x\to p(t)(x,O)$ is measurable for any open set $O\subset E$. Since the system of sets $D\in\mathcal{B}(E)$ such that

$$x\to p(t)(x,D)$$

is measurable is a Dynkin system and it contains all open sets that are an intersection stable generator of $\mathcal{B}(E)$, it contains all of $\mathcal{B}(E)$ by Lemma A.3.15. Hence, $(p(t))_{t\in\mathbb{R}_+}$ is a family of transition kernels.

In the second step, we show that $(p(t))_{t\in\mathbb{R}_+}$ is a semigroup of transition kernels on $(E,\mathcal{B}(E))$, or in other words, that additionally for any $A\in\mathcal{B}(E)$ and any $s,t\in\mathbb{R}_+$

$$\int_E p(s)(y,A)p(t)(x,dy) = p(s+t)(x,A).$$

We know that for any $f\in\mathscr{B}^\rho(E)$

$$\begin{aligned}\int_E\int_E f(z)p(s)(y,dz)p(t)(x,dy) &= P(t)P(s)f\\ &= P(s+t)f\\ &= \int_E f(z)p(s+t)(x,dz)\end{aligned}$$

holds true. For any open set $O\subset E$ we can approach 1_O as in the first step of the proof by a sequence $\mathscr{B}^\rho(E)$, which yields by dominated convergence

$$\int_E p(s)(y,O)p(t)(x,dy) = p(s+t)(x,O).$$

As in the first step since the system of sets such that this equations holds is a Dynkin system it follows that

$$\int_E p(s)(y,A)p(t)(x,dy) = p(s+t)(x,A)$$

holds true for any $A\in\mathcal{B}(E)$.

(iv) For any $C_b(E)$-open set $A\in\mathcal{B}(E)$ (see Definition A.3.36) by definition there is a sequence $\left(f_n^A\right)_{n\in\mathbb{N}}$ such that $f_n^A\nearrow 1_A$ pointwise. Hence, for any $t\in\mathbb{R}_+$

$$x\to p(t)(x,A) = \lim_{n\to\infty} P(t)f_n^A(x)$$

is measurable with respect to Baire σ-algebra $\mathcal{B}_0(E)$ as limit of maps that are in $\mathscr{B}^\rho(E)$ and therefore Baire-measurable (by virtue of being pointwise limits of $C_b(E)$ functions). This property extends to all sets A in the Dynkin system generated by the $C_b(E)$-open sets. Since the system of $C_b(E)$-open sets is intersection stable by Lemma A.3.15 the property holds true also for the σ-algebra generated by the $C_b(E)$-open sets. By Lemma A.3.37 this is precisely $\mathcal{B}_0(E)$.

Furthermore, for any $A \in \mathcal{B}_0(E)$ and any $s, t \in \mathbb{R}_+$ by dominated convergence

$$\begin{aligned}\int_E p(s)(y,A)p(t)(x,dy) &= \lim_{n\to\infty} P(t)P(s)f_n^A \\ &= \lim_{n\to\infty} P(s+t)f_n^A \\ &= p(s+t)(y,A).\end{aligned}$$

$\square$

PROPOSITION 2.3.54. *Let $(\psi_t)_{t\in\mathbb{R}_+}$ be a family of maps such that for any $t \in \mathbb{R}_+$*

$$\psi_t : E \to E.$$

Then $(P(t))_{t\in\mathbb{R}_+}$ defined as

$$P(t)(f) := f \circ \psi_t$$

is a generalized Feller semigroup on $\mathscr{B}^\rho(E)$, called generalized Feller semigroup on $\mathscr{B}^\rho(E)$ of transport type, if and only if the following conditions hold:
(i) $\psi_0 = Id$.
(ii) For any $t_1, t_2 \in \mathbb{R}_+$

$$\psi_{t_1} \circ \psi_{t_2} = \psi_{t_1+t_2}.$$

(iii) For any $x \in E$

$$\lim_{t\searrow 0} \psi_t(x) = x.$$

(iv) For any $t \in \mathbb{R}_+$ and any $R > 0$

$$\psi_t|_{K_R} : K_R \to E$$

is continuous.
(v) For any $t \in \mathbb{R}_+$

$$\sup_{x\in E} \frac{\rho \circ \psi_t(x)}{\rho(x)} =: C_t < \infty.$$

(vi) For some $\delta > 0$ there is $C > 0$ such that for all $0 \leq t < \delta$

$$C_t < C.$$

Furthermore, for a generalized Feller semigroup of transport type the identity

$$P(t)\rho(x) = \sup_{\substack{f \in C_b(E) \\ |f| \leq \rho}} |f \circ \psi_t(x)| = \rho \circ \psi_t(x) \tag{2.3.3}$$

holds true.

PROOF. We first show that the conditions (i)-(vi) are sufficient in order to obtain a generalized Feller semigroup.

Fix $t \in \mathbb{R}_+$. We show that $P(t)$ is a bounded linear map from $\mathscr{B}^\rho(E)$ to $\mathscr{B}^\rho(E)$. For $f \in \mathscr{B}^\rho(E)$ and $n \in \mathbb{N}$, by definition of $\mathscr{B}^\rho(E)$, there is $f_n \in C_b(E)$ such that

$$\|f - f_n\|_\rho < \frac{1}{n}.$$

By Theorem 2.3.42 we obtain $f_n \circ \psi_t \in \mathscr{B}^\rho(E)$ for any $n \in \mathbb{N}$ since on the one hand $f_n \circ \psi_t|_{K_R} \in C_b(E)$ holds for any $R > 0$ and on the other hand

$$\lim_{R \to \infty} \sup_{x \in E \setminus K_R} \frac{|f_n \circ \psi_t(x)|}{\rho(x)} = 0.$$

The inequality

$$\begin{aligned} \sup_{x \in E} \frac{|f \circ \psi_t(x) - f_n \circ \psi_t(x)|}{\rho(x)} &= \sup_{x \in E} \frac{|f \circ \psi_t(x) - f_n \circ \psi_t(x)|}{\rho \circ \psi_t(x)} \cdot \frac{\rho \circ \psi_t(x)}{\rho(x)} \\ &\leq \frac{1}{n} \cdot C_t \end{aligned}$$

yields that $f \circ \psi_t \in \mathscr{B}^\rho(E)$ as a limit of functions in $\mathscr{B}^\rho(E)$. Moreover,

$$\begin{aligned} \|f \circ \psi_t\|_\rho &= \sup_{x \in E} \frac{|f \circ \psi_t(x)|}{\rho \circ \psi_t(x)} \cdot \frac{\rho \circ \psi_t(x)}{\rho(x)} \\ &\leq \|f\|_\rho \cdot C_t, \end{aligned}$$

hence $P(t)$ is a linear bounded operator on $\mathscr{B}^\rho(E)$.

The Properties **P1, P2,** and **P5** of generalized Feller semigroups are easy to check. For Property **P4** we see that for all $0 < t < \delta$

$$\|P(t)\| \leq C_t \leq C.$$

Regarding Property **P3,** we observe that for any $x \in E$ and any $0 \leq t < \delta$ the inequality

$$\rho \circ \psi_t(x) \leq C_\delta \cdot \rho(x) =: R_x$$

holds true. Therefore, $\psi_t(x) \in K_{R_x}$ for $t \in [0, \delta)$ and because of $f|_{K_{R_x}} \in C_b(E)$ for all $f \in \mathscr{B}^\rho(E)$ (see Theorem 2.3.42) we obtain

$$\lim_{t \searrow 0} f \circ \psi_t(x) = f(x)$$

for any $x \in E$.

Next, we show that if $(P(t))_{t \in \mathbb{R}_+}$ is a generalized Feller semigroup, then Properties (i) - (vi) and Equation 2.3.3 hold true.

Property (i) follows from $P(0) = \mathrm{Id}$ which yields

$$f \circ \psi_0 = f \text{ for all } f \in \mathscr{B}^\rho(E). \tag{2.3.4}$$

So by contradiction, if there was some $x \in E$ such that $\psi_0(x) \neq x$, then by definition of completely regular spaces, one could find some map $f_x \in C_b(E) \subset \mathscr{B}^\rho(E)$ such that $f_x(x) = 1$ and $f \circ \psi_0(x) = 0$. But this would contradict Equation 2.3.4.

Regarding Property (ii), as in the proof of Property (i) we obtain

$$f \circ (\psi_{t_1} \circ \psi_{t_2}) = f \circ (\psi_{t_1 + t_2}) \text{ for all } f \in \mathscr{B}^\rho(E). \tag{2.3.5}$$

and as above by contradiction, if Property (ii) did not hold, then one could find a map in $\mathscr{B}^\rho(E)$ that would contradict Equation 2.3.5.

Property (iii) can be shown in the same way, since by definition of generalized Feller semigroups

$$\lim_{t \searrow 0} f \circ \psi_t(x) = f(x)$$

holds for any $x \in E$ and any $f \in \mathscr{B}^\rho(E)$.

In order to show Property (iv), we fix some $R > 0$ and some arbitrary open set O in E. We have to show that $\psi_t^{-1}(O)$ is open in K_R which respect to the subspace topology. We know by Theorem 2.3.42 that $f \circ \psi_t|_{K_R}$ is continuous for any $f \in \mathscr{B}^\rho(E)$. For any $x \in O$, by definition of completely regular spaces, we know that we can find $f_x \in C_b(E)$ such that

$$|f_x| \leq 1,$$
$$f_x(x) = 1,$$

and

$$f_x(E \setminus O) \subset \{0\}.$$

Clearly,

$$\bigcup_{x \in O} \left(f_x \circ \psi_t|_{K_R} \right)^{-1} (0, 2)$$

is open in K_R with respect to the subspace topology. On the other hand

$$\bigcup_{x \in O} (f_x)^{-1} (0, 2) = O.$$

Thus,

$$\begin{aligned}\psi_t|_{K_R}^{-1}(O) &= \psi_t|_{K_R}^{-1}\left(\bigcup_{x\in O}(f_x)^{-1}(0,2)\right)\\ &= \bigcup_{x\in O}\left(f_x\circ\psi_t|_{K_R}\right)^{-1}(0,2)\end{aligned}$$

is open in K_R with respect to the subspace topology.

Regarding Equation 2.3.3, by Lemma 2.3.52 $P(t)\rho(x)$ is given for any $x\in E$ by

$$P(t)\rho(x) = \sup_{\substack{f\in C_b(E)\\ |f|\leq\rho}} |f\circ\psi_t(x)|\,.$$

We observe that for any $y\in E$ and any $n\in\mathbb{N}$ there is an open neighborhood $O_{n,y}$ of y such that

$$\rho(x) > \rho(y) - \frac{1}{n}$$

holds true for any $x\in O_{n,y}$. On $E\setminus O_{n,y}\cup\{y\}$ we define the function

$$g_{n,y}(x) := \begin{cases}\rho(y)-\frac{1}{n} & \text{for } x=y\\ 0 & \text{for } x\in E\setminus O_{n,y},\end{cases}$$

and by Proposition 2.3.8 we can extend $g_{n,y}$ to $f_{n,y}\in C_b(E)$ such that $|f_{n,y}|<\rho$ and $\rho(y)-f_{n,y}(y)=\frac{1}{n}$. Hence, for any $x\in E$

$$\sup_{\substack{f\in C_b(E)\\ |f|\leq\rho}} |f\circ\psi_s(x)| = \rho\circ\psi_s(x).$$

Finally, Property (v) and (vi) follow since for any $x\in E$

$$\rho\circ\psi_t(x) = P(t)\rho(x),$$

and by Theorem 2.3.51 and Proposition 1.4.9 the estimate

$$\|P(t)\|\leq Me^{\omega t}$$

holds true for some $M\geq 1$ and $\omega\in\mathbb{R}$. Thus, Lemma 2.3.52 implies that for any $x\in E$

$$\rho\circ\psi_t(x)\leq\rho(x)Me^{\omega t}.$$

□

EXAMPLE 2.3.55. For $\alpha>0$ let $(\psi_t^\alpha)_{t\in\mathbb{R}_+}$ be a family of maps such that for any $t\in\mathbb{R}_+$

$$\psi_t^\alpha:\mathbb{R}\to\mathbb{R}$$

is defined as

$$\psi_t^\alpha(x) = \alpha^t x.$$

Then we define $\rho(x) := \max\{\ln|x|, 1\}$ and

$$\begin{aligned}\sup_{x\in\mathbb{R}} \frac{\rho(\psi_t^\alpha(x))}{\rho(x)} &= \sup_{x\in\mathbb{R}} \frac{\rho(\alpha^t x)}{\rho(x)} \\ &= \max\left(\sup_{|x|>e} \frac{\rho(\alpha^t x)}{\ln|x|}, \sup_{|x|\le e} \frac{\rho(\alpha^t x)}{1}\right) \\ &\le t\ln(\alpha) + 1 \\ &=: C_t.\end{aligned}$$

Therefore, $(P_t^\alpha)_{t\in\mathbb{R}_+}$ defined as

$$P_t^\alpha(f) := f \circ \psi_t^\alpha$$

is a generalized Feller semigroup.

It is also a Feller semigroup as it is clear that it maps $C_0(\mathbb{R})$ to $C_0(\mathbb{R})$, is positive and contractive in the supremum norm. Moreover, strong continuity in the supremum norm follows from the fact that on compact sets in metric spaces continuous maps are uniformly continuous.

Example 2.3.56. (generalized Feller semigroup of transport type, but not Feller semigroup)

Consider $E = \mathbb{R}^2$ in polar coordinates, define

$$\begin{aligned}\rho :\ (0,\infty] \times (0, 2\pi] &\to (0,\infty] \\ (r,\varphi) &\to 1 + \frac{r}{\varphi},\end{aligned}$$

and $\rho(0,\varphi) = 1$. Then ρ is an admissible weight function. For the map

$$\psi_t(r,\varphi) := \left(re^{-\frac{t}{\varphi}}, \varphi\right)$$

by Proposition 2.3.54 $P(t)(f) := f\circ\psi_t$ is a generalized Feller semigroup by

$$\frac{\rho\left(\psi_t(r,\varphi)\right)}{\rho\left(r,\varphi\right)} = \frac{1 + \frac{re^{-\frac{t}{\varphi}}}{\varphi}}{1 + \frac{r}{\varphi}} \le 1.$$

But it is not a Feller semigroup because for $g(r) := e^{-r} \in C_0(\mathbb{R}^2)$

$$\left(g \circ \psi_t\right)(r,\varphi) = e^{-re^{-\frac{t}{\varphi}}} \notin C_0(\mathbb{R}^2)$$

as can be see by letting φ approach 0.

COROLLARY 2.3.57. *Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ of transport type and let $(\psi_t)_{t\in\mathbb{R}_+}$ be as in Proposition 2.3.54. For some $M \geq 1$ and $\omega \in \mathbb{R}$ let*

$$\|P(t)\| \leq Me^{\omega t},$$

and let A be the generator of $(P(t))_{t\in\mathbb{R}_+}$, and A' the adjoint of A (see Definition 1.4.41). Then

$$(Q(t))_{t\in\mathbb{R}_+}$$

defined as

$$Q(t)(\mu) := \mu \circ \psi_t^{-1}$$

is a semigroup on $\mathcal{M}^\rho(E)$ and

$$\left(Q(t)|_{\overline{\mathcal{D}(A')}} \right)_{t\in\mathbb{R}_+}$$

is a strongly continuous semigroup on $\overline{\mathcal{D}(A')} \subset \mathcal{M}^\rho(E)$.

PROOF. Fix $t \in \mathbb{R}_+$. By Theorem 2.3.37 we obtain for $\mu \in \mathscr{B}^\rho(E)' = \mathcal{M}^\rho(E)$ and $f \in \mathscr{B}^\rho(E)$ the identities

$$\begin{aligned} \mu\left(P(t)(f)\right) &= \int_E \left(f \circ \psi_t(x)\right) \mu(dx) \\ &= \int_E f(y) \left(\mu \circ \psi_t^{-1}\right)(dy) \\ &= \left(\mu \circ \psi_t^{-1}\right)(f). \end{aligned}$$

Furthermore, by Lemma 2.3.52 and Proposition 2.3.54

$$\begin{aligned} \int_E \rho(x)\mu \circ \psi_t^{-1}(dx) &= \int_E P(t)\rho(x)\mu(dx) \\ &= \int_E P(t)\rho(x)\mu(dx) \\ &\leq \int_E Me^{\omega t}\rho(x)\mu(dx). \end{aligned}$$

Hence,

$$\begin{aligned} Q(t): \quad \mathcal{M}^\rho(E) &\to \mathcal{M}^\rho(E) \\ \mu &\to \mu \circ \psi_t^{-1} \end{aligned}$$

is the adjoint operator of $P(t)$ and the statement of the Corollary follows from Lemma 1.4.43 and Proposition 1.4.61. □

On normed vector spaces, we can determine a subset of the domain of the generator and the generator on that subset. In the following, Df will denote the Fréchet derivative of f (see Definition A.2.2).

PROPOSITION 2.3.58. *Let E be a normed vector space, let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ of transport type, and let $(\psi_t)_{t\in\mathbb{R}_+}$ be as in Proposition 2.3.54. Let A be the generator of $(P(t))_{t\in\mathbb{R}_+}$, and A' its adjoint. Let $t \to \psi_t(x)$ be continuously differentiable for any $x \in E$ and define the vector field*

$$v(x) := \lim_{t\searrow 0} \frac{\psi_t(x) - x}{t}.$$

Let $v \in C_c(E,E)$.
Then $C_c^1(E,\mathbb{R}) \subset \mathcal{D}(A)$ and for $f \in C_c^1(E,\mathbb{R})$

$$Af(x) = (Df(x))\,(v(x)) \text{ for } x \in E.$$

For $\mu \in \mathcal{D}(A')$ and $f \in C_c^1(E,\mathbb{R})$

$$A'\mu(f) = \int_E Df(x)(v(x))\mu(dx).$$

PROOF. First, we observe that for $f \in C_c^1(E,\mathbb{R})$ and any $x \in E$, the map

$$t \to (f \circ \psi_t)\, x$$

is continuously differentiable and for $s \geq 0$ the derivative is given by

$$\begin{aligned} \frac{d}{dt}\bigg|_{t=s} (f\circ\psi_t)\,(x) &= (Df(\psi_s\,(x)))\left(\lim_{h\searrow 0}\frac{\psi_h(\psi_s\,(x)) - \psi_s\,(x)}{h}\right) \\ &= (Df(\psi_s\,(x)))\, v(\psi_s\,(x)). \end{aligned}$$

We need to show that the difference quotient converges not only pointwise to the derivative but also in the $\|\cdot\|_\rho$-norm:

$$\lim_{h\to 0}\left\| \frac{f(\psi_h(x)) - f(x)}{h} - (Df(x))\,(v(x)) \right\|_\rho = 0.$$

For this purpose, we fix $\varepsilon > 0$ and we want to bound the left hand side by a suitable expression. By the mean value theorem for any $x \in E$ and any $h > 0$ there is $0 \leq s(x,h) \leq h$ such that

$$\frac{f(\psi_h(x)) - f(x)}{h} = \frac{d}{dt}\bigg|_{t=s(x,h)} (f\circ\psi_t)\,(x).$$

This yields the estimate

$$\begin{aligned}
&\left|\frac{f(\psi_h(x))-f(x)}{h\rho(x)}-\frac{(Df(x))\,(v(x))}{\rho(x)}\right| \\
&=\left|\frac{\frac{d}{dt}\big|_{t=s(x,h)}\,(f\circ\psi_t)\,x}{\rho(x)}-\frac{(Df(x))\,(v(x))}{\rho(x)}\right| \\
&\le\left|\frac{\big(Df(\psi_{s(x,h)}(x))\big)\,\big(v(\psi_{s(x,h)}(x))\big)}{\rho(x)}-\frac{\big(Df(\psi_{s(x,h)}(x))\big)\,(v(x))}{\rho(x)}\right| \\
&+\left|\frac{\big(Df(\psi_{s(x,h)}(x))\big)\,(v(x))}{\rho(x)}-\frac{(Df(x))\,(v(x))}{\rho(x)}\right|.
\end{aligned}$$

Furthermore, we note that $f\in C_c^1(E,\mathbb{R})$ and $v\in C_c(E,E)$ imply that there exists $\delta>0$ such that for any $x,y\in Y$ satisfying $\|x-y\|<\delta$ the inequalities

$$\|Df(x)-Df(y)\|_{L(E,\mathbb{R})}<\varepsilon,$$

and

$$\|v(x)-v(y)\|<\varepsilon$$

hold true. For any $0\le s\le h:=\frac{\delta}{\|v\|_\infty}$

$$\|\psi_s(x)-x\|<h\cdot\|v\|_\infty<\delta.$$

Thus,

$$\begin{aligned}
&\left|\frac{\big(Df(\psi_{s(x,h)}(x))\big)\,\big(v(\psi_{s(x,h)}(x))\big)}{\rho(x)}-\frac{\big(Df(\psi_{s(x,h)}(x))\big)\,(v(x))}{\rho(x)}\right| \\
&\le\left\|\frac{Df(\psi_{s(x,h)}(x))}{\rho(x)}\right\|_{L(E,\mathbb{R})}\cdot\underbrace{\big\|v(\psi_{s(x,h)}(x))-v(x)\big\|}_{<\varepsilon} \\
&\le\frac{\varepsilon}{\inf\limits_{x\in E}\rho(x)}\cdot\sup_{x\in E}\|Df(x)\|_{L(E,\mathbb{R})},
\end{aligned}$$

and

$$\begin{aligned}
&\left|\frac{\big(Df(\psi_{s(x,h)}(x))\big)\,(v(x))}{\rho(x)}-\frac{(Df(x))\,(v(x))}{\rho(x)}\right| \\
&\le\underbrace{\big\|Df(\psi_{s(x,h)}(x))-Df(x)\big\|_{L(E,\mathbb{R})}}_{<\varepsilon}\cdot\left\|\frac{v(x)}{\rho(x)}\right\|.
\end{aligned}$$

Combined, these results yield the bound

$$\left\| \frac{f(\psi_h(x)) - f(x)}{h} - (Df(x))\,(v(x)) \right\|_\rho$$
$$\leq \varepsilon \left(\underbrace{\frac{1}{\inf\limits_{x\in E} \rho(x)} \cdot \sup_{x\in E} \|Df(x)\|_{L(E,\mathbb{R})} + \sup_{x\in E} \left\| \frac{v(x)}{\rho(x)} \right\|}_{:=C} \right).$$

Since C does not depend on h and $\varepsilon > 0$ was arbitrary, the left hand side converges to 0 as $h \searrow 0$ which is what we had to show.

For $\mu \in \mathcal{D}(A') \subset \mathcal{M}^\rho(E)$ we obtain for $f \in C_c^1(E, \mathbb{R}) \subset \mathcal{D}(A)$

$$A'\mu(f) = \mu(Af) = \int_E (Df(x))\,(v(x))\, d\mu(x).$$

□

In order to determine the domain of the adjoint of the generator we have to make additional assumptions. In particular, we look at functions and measures defined on $\mathbb{R}^n$.

Proposition 2.3.59. *Let $E \subset \mathbb{R}^n, n \in \mathbb{N}$ and let the conditions of Proposition 2.3.58 be fulfilled and $v \in C_c^1(E, E)$. For the Lebesgue measure λ let $\mu \ll \lambda$ be given by the density g_μ:*

$$\mu(B) = \int_B g_\mu(x)\lambda(dx) \text{ for any } B \in \mathcal{B}(E), \tag{2.3.6}$$

and for $C > 0$ let $g_\mu \in C^1(E)$ fulfill

$$\int_E \rho(x) \left(\left| \frac{d}{dx_1} g_\mu + \dots \frac{d}{dx_n} g_\mu \right| \right)(x)\lambda(dx) < C \int_E \rho(x) \left| g_\mu(x) \right|(x)\lambda(dx).$$

Denote the space of such measures as $\mathcal{MC}^{\rho,1}(E) \subset \mathcal{M}^\rho(E)$. Then, $\mathcal{MC}^{\rho,1}(E) \subset \mathcal{D}(A')$ and for $\mu \in \mathcal{MC}^{\rho,1}(E)$

$$A'(\mu)(B) = -\int_B div\,(v \cdot g_\mu)\,(x)\lambda(dx) \text{ for any } B \in \mathcal{B}(Y). \tag{2.3.7}$$

$(Q(t))_{t\in\mathbb{R}_+}$ defined as

$$Q(t)(\mu) := \mu \circ \psi_t^{-1}$$

is a strongly continuous semigroup on $\overline{\mathcal{D}(A')} \subset \mathcal{M}^\rho(E)$ and its generator $A^\dagger$ is given by the restriction of A' to the set

$$\mathcal{D}(A^\dagger) := \left\{ y' \in \mathcal{D}(A') : \, A'y' \in \overline{\mathcal{D}(A')} \right\}.$$

PROOF. By

$$\int_E \rho(x) \left(\left| \frac{d}{dx_1} g_\mu + ... \frac{d}{dx_n} g_\mu \right| \right) (x) d\lambda(x) < C \int_E \rho(x) \left| g_\mu(x) \right| (x) d\lambda(x)$$

we obtain boundedness of the linear map

$$\begin{aligned} L: \quad & \mathcal{M}C^{\rho,1}(E) \subset \mathcal{M}^\rho(E) \to \mathcal{M}^\rho(E) \\ & \int_B g_\mu(x)\lambda(dx) \to -\int_B \operatorname{div}(v \cdot g_\mu)(x)\lambda(dx). \end{aligned}$$

In order to show $\mathcal{M}C^{\rho,1}(E) \subset \mathcal{D}(A')$, we have to show that for $\mu \in \mathcal{M}C^{\rho,1}(E)$ and $f \in \mathcal{D}(A)$

$$\mu(Af) = L\mu(f)$$

holds true. By definition and dominated convergence

$$\begin{aligned} \mu(Af) &= \int_E \left(\lim_{h \searrow 0} \frac{f(\psi_h(x)) - f(x)}{h} \cdot g_\mu(x) \right) \lambda(dx) \\ &= \lim_{h \searrow 0} \int_E \left(\frac{f(\psi_h(x)) - f(x)}{h} \cdot g_\mu(x) \right) \lambda(dx) \\ &= \lim_{h \searrow 0} \left(\int_E \left(\frac{f(\psi_h(x))}{h} \cdot g_\mu(x) \right) \lambda(dx) - \int_E \left(\frac{f(x)}{h} \cdot g_\mu(x) \right) \lambda(dx). \right) \end{aligned}$$

We want to substitute $z := \psi_h(x)$. By our assumptions and the Picard-Lindelöf theorem (Theorem A.2.3) also the inverse ψ_h^{-1}exists. We see that for some $s > 0$ and $x = \psi_s(y)$

$$\begin{aligned} \lim_{h \to 0} \frac{(\psi_h)^{-1}(x) - x}{h} &= \lim_{h \to 0} \frac{(\psi_{s-h}(y)) - \psi_s(y)}{h} = \\ &= -v(\psi_s(y)) \\ &= -v(x). \end{aligned}$$

Furthermore, by Theorem A.2.4 for any $t \geq 0$

$$x \to \psi_t(x)$$

is continuously differentiable and clearly $D\psi_0 = \mathrm{Id}$. We obtain

$$\begin{aligned} \mu(Af) = \lim_{h \searrow 0} \Bigg(& \int_E \left(\frac{f(z)}{h} \cdot g_\mu((\psi_h)^{-1}(z)) \right) \left((|D\psi_h|) ((\psi_h)^{-1}(z)) \right)^{-1} \lambda(dz) \\ & - \int_E \left(\frac{f(x)}{h} \cdot g_\mu(x) \right) \lambda(dx). \Bigg) \end{aligned}$$

Written in terms of difference quotients this yields (using our assumptions and dominated convergence):

$$\mu(Af) = \int_E \lim_{h\searrow 0} f(x) \left(\frac{g_\mu((\psi_h)^{-1}(x)) - g_\mu(x)}{h} \right.$$
$$\left. + \frac{g_\mu((\psi_h)^{-1}(x)) \left((|D\psi_h|)((\psi_h)^{-1}(x)) \right)^{-1} - g_\mu((\psi_h(x))^{-1})}{h} \right) \lambda(dx).$$

Calculating the derivatives using the chain rule we obtain:

$$\begin{aligned} \mu(Af) &= \int_E f(x) \left(-v(x) \cdot \nabla g_\mu(x) + g_\mu(x) \cdot \left. \frac{-\frac{d}{dh}\left((|D\psi_h|)((\psi_h)^{-1}(x))\right)}{(|\mathrm{Id}|(x))^2} \right|_{h=0} \right) \lambda(dx) \\ &= \int_E -f(x) v(x) \cdot \nabla g_\mu(x) \lambda(dx) \\ &\quad - \int_E f(x) g_\mu(x) \cdot \left. \left(\left(\frac{d}{dh}(|D\psi_h|) \right) ((\psi_h)^{-1}(x)) \right) \right|_{h=0} \lambda(dx) \\ &\quad + \int_E f(x) \left((D(|D\psi_h|))((\psi_h)^{-1}(x)) \right) \frac{d}{dh} \left((\psi_h)^{-1} \right)(x) \Big|_{h=0} \lambda(dx). \end{aligned}$$

Making use of the rules for the derivative of the determinant, we can show

$$\begin{aligned} \mu(Af) &= \int_E -f(x) v(x) \cdot \nabla g_\mu(x) \lambda(dx) \lambda(dx) \\ &\quad - \int_E f(x) g_\mu(x) \cdot \left(|D\psi_h| \cdot \mathrm{tr}\left((D\psi_h)^{-1} \frac{d}{dh}(D\psi_h) \right) ((\psi_h)^{-1}(x)) \Big|_{h=0} \right) \lambda(dx) \\ &\quad + \int_E f(x) \left(|D\psi_h| \cdot \mathrm{tr}\left((D\psi_h)^{-1} \left(D^2 \psi_h \right) \right) ((\psi_h)^{-1}(x)) \right) \frac{d}{dh} \left((\psi_h)^{-1} \right)(x) \Big|_{h=0} \lambda(dx). \end{aligned}$$

Finally, this simplifies to

$$\begin{aligned} \mu(Af) &= \int_E f(x) \left(-v(x) \cdot \nabla g_\mu(x) - g_\mu(x) \cdot (\mathrm{div}(v)(x)) \right) \lambda(dx) \\ &\quad + \int_E \underbrace{f(x) \left((1 \cdot \mathrm{tr}(0))((\psi_h)^{-1}(x)) \right) \frac{d}{dh} \left((\psi_h)^{-1} \right)(x) \Big|_{h=0}}_{=0} \lambda(dx) \\ &= L\mu(f). \end{aligned}$$

Thus, $\mathcal{M}C^{\rho,1}(E) \subset \mathcal{D}(A')$ and $L(\mu) = A'(\mu)$ for $\mu \in \mathcal{M}C^{\rho,1}(E)$.

□

We were able to characterize operators that generate strongly continuous semigroups by the Hille-Yosida theorem (Theorem 1.4.35) and would like to achieve a similar characterization of operators that generate generalized Feller semigroups. However, in this case we need to assume that for the generalized Feller semigroup $(P(t))_{t\in\mathbb{R}_+}$ with the usual norm bounds

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))} \leq M \exp(\omega t)$$

the condition $M = 1$ holds. Under this condition, operators that generate generalized Feller semigroups are characterized in the following theorem which was proved in [**15**].

THEOREM 2.3.60. *Let A be a linear operator on $\mathscr{B}^\rho(E)$ and $\mathcal{D}(A)$ its domain. Let $\omega \in \mathbb{R}$. A is closable and $\bar{A}$ generates a generalized Feller semigroup $(P(t))_{t\in\mathbb{R}_+}$ with*

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))} \leq \exp(\omega t)$$

for all $t \in \mathbb{R}_+$ if and only if
(i) $\mathcal{D}(A)$ is dense,
(ii) For some $\lambda_0 > \omega$ the linear operator $A - \lambda_0$ has a dense image (and hence for all $\lambda > \omega$ the linear operator $A - \lambda$ has a dense image) and,
(iii) A satisfies the generalized positive maximum principle, that is, for $f \in \mathcal{D}(A)$ with

$$\max\left(\frac{f}{\rho}, 0\right) \leq \frac{f(z)}{\rho(z)} \tag{2.3.8}$$

for some $z \in E$ the inequality

$$Af(z) \leq wf(z)$$

holds.

PROOF. First, we show the implication that for a closable linear operator A on $\mathscr{B}^\rho(E)$ with domain $\mathcal{D}(A)$ for which $\bar{A}$ generates a generalized Feller semigroup $(P(t))_{t\in\mathbb{R}_+}$ with

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))} \leq \exp(\omega t)$$

for all $t \in \mathbb{R}_+$, Properties (i) (ii) and (iii) hold.

(i) By Theorem 2.3.51 $(P(t))_{t\in\mathbb{R}_+}$ is a strongly continuous semigroup and by Proposition 1.4.24 its generator $\bar{A}$ has a dense domain. Hence also $\mathcal{D}(A)$ is dense.

(ii) By Theorem 1.4.29 for all $\lambda > \omega$ we have $\lambda \in \rho(\bar{A})$ and the operator $\bar{A} - \lambda$ is a bijection between its domain $\mathcal{D}(\bar{A})$ and $\mathscr{B}^\rho(E)$. Since the graph of $\bar{A} - \lambda$ is the closure of the graph of $A - \lambda$, the range of $A - \lambda$ is dense in $\mathscr{B}^\rho(E)$.

(iii) For $f \in \mathcal{D}(A)$ and $z \in E$ such that

$$\max\left(\frac{f}{\rho}, 0\right) \leq \frac{f(z)}{\rho(z)}$$

we want to find an estimate for $\frac{P(t)f(z)-f(z)}{t}$ and want to take the limit as $t \to 0$. For this purpose, we can use assumption

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))} \le \exp(\omega t)$$

and would like to work with functions in $\mathcal{D}(A)$ that take values only in $\mathbb{R}_+$ in order to make better use of assumption 2.3.8. Therefore, for $f \in \mathcal{D}(A)$ we look at $\max(f,0)$ (which is clearly in $\mathscr{B}^\rho(E)$) and we obtain by positivity of the generalized Feller semigroup **(P5)**

$$P(t)f(z) \le P(t)(\max(f,0))(z).$$

Then, our estimate on $\|P(t)\|_{L(\mathcal{B}^\rho(E))}$ yields

$$\begin{aligned} P(t)(\max(f,0))(z) &\le \rho(z) \|P(t)(\max(f,0))\|_\rho \\ &\le \rho(z)\exp(\omega t)\|\max(f,0)\|_\rho \end{aligned}$$

and the inequality

$$\max\left(\frac{f}{\rho},0\right) \le \frac{f(z)}{\rho(z)}$$

for some $z \in E$ implies

$$\rho(z)\exp(\omega t)\|\max(f,0)\|_\rho \le \rho(z)\exp(\omega t)\frac{f(z)}{\rho(z)}.$$

Hence, we obtain

$$P(t)f(z) \le \exp(\omega t) f(z)$$

and

$$\frac{P(t)f(z)-f(z)}{t} \le \frac{\exp(\omega t) f(z) - f(z)}{t}.$$

Taking the limit as $t \to 0$ yields Property (iii).

Next, we show the opposite implication. In a first step, we use Lumer-Phillips theorem (Theorem 1.4.38) in order to find a strongly continuous semigroup generated by $\bar{A} - \omega$. Later, we show that its rescaled semigroup indeed is a generalized Feller semigroup with the desired properties.

For the application of Lumer-Phillips theorem we need to show that $A - \omega$ is densely defined, that there is some $\lambda > 0$ such that $\lambda - (A - \omega)$ has a dense image, and that $A - \omega$ is dissipative. The first two conditions follow immediately from Properties (i) and (ii).

As for the dissipativity of $A-\omega$ let $f \in \mathcal{D}(A)$ and $\lambda > 0$ be arbitrary. To prove dissipativity of $A - \omega$ we have to show that

$$\|(\lambda - (A - \omega)) f\|_\rho \ge \lambda \|f\|_\rho.$$

We want to use Property (iii) for bounding $\|f\|_\rho$. More precisely, by Corollary 2.3.44 there exists $z \in E$ such that

$$\|f\|_\rho = \frac{|f(z)|}{\rho(z)}.$$

Defining $g := (\mathrm{sgn} f(z)) \cdot f$ implies

$$\|f\|_\rho = \frac{g(z)}{\rho(z)}.$$

We see that $g \in \mathcal{D}(A)$ since $\mathcal{D}(A)$ has to be a vector space and we can apply Property (iii) to g. Therefore,

$$\begin{aligned}
\lambda \|f\|_\rho &= \lambda \frac{g(z)}{\rho(z)} \\
&\leq \lambda \frac{g(z)}{\rho(z)} + \omega \frac{g(z)}{\rho(z)} - A \frac{g(z)}{\rho(z)} \\
&\leq \sup_{x \in E} \frac{|(\lambda + \omega - A)\, g(x)|}{\rho(x)} \\
&= \|(\lambda + \omega - A)\, f\|_\rho
\end{aligned}$$

and $A - \omega$ is dissipative.

Therefore, the the Lumer-Phillips theorem (Theorem 1.4.38) can be applied and yields that $\bar{A} - \omega$ generates a strongly continuous semigroup $\{S(t)\}_{t \in \mathbb{R}_+}$ on $\mathscr{B}^\rho(E)$ such that for all $f \in \mathscr{B}^\rho(E)$ and for all $t \geq 0$

$$\|S(t)f\|_\rho \leq \|f\|_\rho .$$

For the family of linear operators $\{P(t)\}_{t \in \mathbb{R}_+}$ on $\mathscr{B}^\rho(E)$ defined for any $t \in \mathbb{R}_+$ as

$$P(t) = S(t)e^{\omega t}$$

Lemma 1.4.16 yields that $\{P(t)\}_{t \in \mathbb{R}_+}$ is a strongly continuous semigroup as well and that its generator is given by $\bar{A}$. Furthermore, for all $t \geq 0$

$$\|P(t)f\|_\rho \leq e^{\omega t} \|f\|_\rho .$$

Having found the strongly continuous semigroup $\{P(t)\}_{t \in \mathbb{R}_+}$, in a next step we show that this is indeed a generalized Feller semigroup with the desired properties. We need to check Properties **P3, P4**, and **P5**.

Regarding Property **P4**, this follows directly form

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))} \leq \exp(\omega t) .$$

Property **P3** follows from strong continuity of $\{P(t)\}_{t\in\mathbb{R}_+}$and

$$\lim_{t\to 0}\left(P(t)f(x)-f(x)\right) \le \rho(x)\lim_{t\to 0}\left(\sup_{x\in E}\frac{|P(t)f(x)-f(x)|}{\rho(x)}\right)$$
$$= \rho(x)\lim_{t\to 0}\left\|P(t)f-f\right\|_{\rho}.$$

In order to show property **P5** (positivity) of $\{P(t)\}_{t\in\mathbb{R}_+}$we first observe that by Post-Widder Inversion Formula (Theorem A.5.1) positivity of $\left(\lambda-\bar{A}\right)^{-1}$ for any $\lambda>\omega$ implies positivity of $\{P(t)\}_{t\in\mathbb{R}_+}$.
In order to show positivity of $\left(\lambda-\bar{A}\right)^{-1}$ for any $\lambda>\omega$, we fix $\lambda_0>\omega$ and functions $f,g\in\mathscr{B}^{\rho}(E)$ that satisfy $f=\left(\lambda_0-\bar{A}\right)^{-1}g$, such that f is not positive. Clearly $f\in\mathcal{D}(\bar{A})$. We show that under these assumptions g is not positive either. An equivalent way of stating that f is not positive is

$$\alpha := \inf_{x\in E}\frac{f(x)}{\rho(x)} < 0.$$

We would like to show that

$$\beta := \inf_{x\in E}\frac{g(x)}{\rho(x)} < 0 \tag{2.3.9}$$

holds as well and for this purpose we would like to bound β by a suitable expression depending on α. However, if we substitute

$$\left(\lambda_0-\bar{A}\right)f = g$$

in Inequality 2.3.9 we encounter the expression $\bar{A}f$. We would like to apply Property (iii) of this theorem and since this is only possible for functions in $\mathcal{D}(A)$ we choose a sequence $(f_n)_{n\in\mathbb{N}}\subset\mathcal{D}(A)$ such that $Af_n\to\bar{A}f$ and $f_n\to f$ in $\mathscr{B}^{\rho}(Y)$ as $n\to\infty$. By uniform convergence, limit and infimum can be interchanged, hence

$$\inf_{x\in E}\frac{g(x)}{\rho(x)} = \inf_{x\in E}\lim_{n\to\infty}\frac{\left(\lambda_0-\bar{A}\right)f_n(x)}{\rho(x)}$$
$$= \lim_{n\to\infty}\inf_{x\in E}\frac{\left(\lambda_0-A\right)f_n(x)}{\rho(x)}.$$

Similarly,

$$\inf_{x\in E}\frac{f(x)}{\rho(x)} = \lim_{n\to\infty}\inf_{x\in E}\frac{f_n(x)}{\rho(x)}.$$

Since $f_n \to f$ in $\mathscr{B}^\rho(E)$ as n tends to infinity and f is not positive there is some $M \in \mathbb{N}$ such that for all $n > M$

$$\sup_{x\in E} - f_n(x) > 0$$

and Theorem 2.3.43 yields that for any $n \in \mathbb{N}$, $n > M$ there is $z_n \in E$ such that for all $x \in E$

$$\frac{-f_n(x)}{\rho(x)} \leq \frac{-f_n(z_n)}{\rho(z_n)}$$

hence

$$\inf_{x\in E} \frac{f_n(x)}{\rho(x)} = \frac{f_n(z_n)}{\rho(z_n)}.$$

We can apply the generalized positive maximum principle to $-f_n$ at z_n for $n \in \mathbb{N}$, $n > M$ and obtain

$$A\left(-f_n(z_n)\right) \leq \omega\left(-f_n(z_n)\right).$$

Therefore, the following estimates hold:

$$\begin{aligned}
\inf_{x\in E} \frac{g(x)}{\rho(x)} &= \lim_{n\to\infty} \left(\inf_{x\in E} \frac{\left(\lambda_0 - A\right) f_n(x)}{\rho(x)} \right) \\
&\leq \lim_{n\to\infty} \frac{\left(\lambda_0 - A\right) f_n(z_n)}{\rho(z_n)} \\
&\leq \lim_{n\to\infty} \frac{\lambda_0 f_n(z_n) - \omega\left(f_n(z_n)\right)}{\rho(z_n)} \\
&= \left(\lambda_0 - \omega\right) \lim_{n\to\infty} \left(\inf_{x\in E} \frac{f_n(x)}{\rho(x)} \right) \\
&= \left(\lambda_0 - \omega\right) \inf_{x\in E} \frac{f(x)}{\rho(x)} \\
&< 0
\end{aligned}$$

To sum up, we have shown that for $\lambda > \omega$ $g = \left(\lambda - \bar{A}\right) f$ cannot be positive when $f \in \mathscr{B}^\rho(E)$ is not positive which proves that the linear operator $\left(\lambda - \bar{A}\right)^{-1}$ is positive. Thus $\{P(t)\}_{t\in\mathbb{R}_+}$is positive and property **P5** of generalized Feller semigroup holds. In conjunction with the previous parts of the proof this shows that$\{P(t)\}_{t\in\mathbb{R}_+}$ is indeed a generalized Feller semigroup with generator $\bar{A}$ such that

$$\left\|P(t)\right\|_{L(\mathscr{B}^\rho(E))} \leq \exp\left(\omega t\right).$$

□

If a generalized Feller semigroup $(P(t))_{t\in\mathbb{R}_+}$ with

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))} \leq M \exp(\omega t)$$

does not satisfy $M = 1$, then Property (iii) in Theorem 2.3.60 does not have to hold anymore. This is shown in the following counterexample.

EXAMPLE 2.3.61. Let $X = \mathbb{R}$ and

$$\rho(s) := \begin{cases} 1 & \text{if } |s| \leq 1 \\ |s|+1 & \text{else.} \end{cases}$$

(Compare with Figure 2.3.2.) For $t \in \mathbb{R}_+$ let the maps

$$P_t : \mathscr{B}^\rho(\mathbb{R}) \to \mathscr{B}^\rho(\mathbb{R})$$

be defined as

$$P_t(f)(s) := f(s+t).$$

The calculation

$$\begin{aligned} \sup_{s\in\mathbb{R}} \frac{|f(s+t)|}{\rho(s)} &= \sup_{s\in\mathbb{R}} \frac{|f(s+t)|}{\rho(s+t)} \frac{\rho(s+t)}{\rho(s)} \\ &\leq \|f\|_\rho (t+2) \end{aligned}$$

shows by Proposition 2.3.54 that $(P_t)_{t\in\mathbb{R}_+}$ is a generalized Feller semigroup.

There is indeed no $\omega \in \mathbb{R}$ such that

$$\|P_t\|_{L(\mathscr{B}^\rho(E))} \leq \exp(\omega t)$$

holds true for all $t \in \mathbb{R}$. In order to see this, we note that for any $t > 0$ there is $f_t \in \mathscr{B}^\rho(\mathbb{R})$ such that

$$\|f_t\|_\rho = 1$$

and

$$f_t(1+t) = 2+t.$$

holds. Then

$$\frac{P_t(f_t)(1)}{\rho(1)} = 2+t.$$

Hence $\|P_t\|_\rho \geq 2+t$ for all $t > 0$ and if there was $\omega \in \mathbb{R}$ such that

$$\|P_t\|_{L(\mathscr{B}^\rho(E))} \leq \exp(\omega t)$$

then

$$2 = \lim_{t\searrow 0} (2+t) \leq \lim_{t\searrow 0} \|P_t\|_{L(\mathscr{B}^\rho(E))} \leq \lim_{t\searrow 0} \exp(\omega t) = 1$$

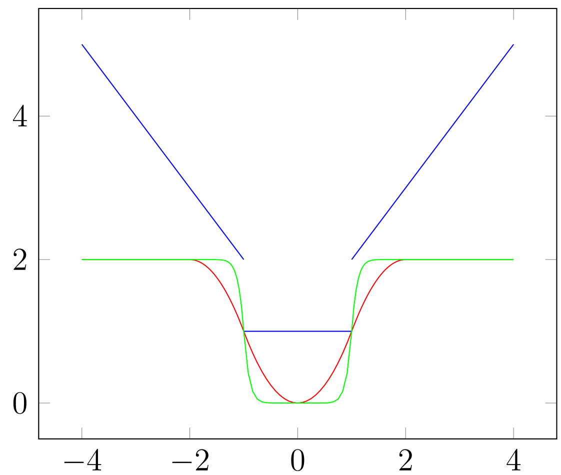


FIGURE 2.3.2. ρ is blue, f_2 is red, f_{10} is green

would yield a contradiction. Moreover, for all $f \in \mathscr{B}^\rho(\mathbb{R})$ such that $\|f\|_\rho \leq 1$ we obtain for all $s, t \geq 0$

$$\frac{|P_t(f)(s)|}{\rho(s)} = \frac{|f(s+t)|}{\rho(s)} \leq \frac{\rho(s+t)}{\rho(s)} \leq 2 + t.$$

Thus,

$$\|P_t\|_\rho = 2 + t \leq 2 \exp\left(\frac{t}{2}\right)$$

for all $t \geq 0$.

Fix same arbitrary $n \in \mathbb{N}$, $n \geq 2$. Define

$$f_n(s) := \begin{cases} |s|^n & \text{if } |s| \leq 1 \\ 2 - (2 - |s|)^n & \text{if } 1 < |s| < 2 \\ 2 & \text{else.} \end{cases}$$

$f_n \in \mathscr{B}^\rho(\mathbb{R})$ and $\|f_n\|_\rho = 1$ hold for all $n \in \mathbb{N}$, $n \geq 1$. Moreover,

$$\sup_{s \in \mathbb{R}} \frac{|f_n(s)|}{\rho(s)} \leq \frac{f_n(1)}{\rho(1)}.$$

For all $s \in \mathbb{R}$

$$\lim_{t \searrow 0} \frac{f_n(s+t) - f_n(s)}{t} = g_n(s)$$

holds true with $g_n \in \mathscr{B}^\rho(\mathbb{R})$ defined as

$$g_n(s) := \begin{cases} -n\,|s|^{n-1} & \text{if } -1 \leq s < 0 \\ ns^{n-1} & \text{if } 0 \leq s \leq 1 \\ n\,(2-s)^{n-1} & \text{if } 1 < s < 2 \\ -n\,(2+s)^{n-1} & \text{if } -2 < s < -1 \\ 0 & \text{else.} \end{cases}$$

In order to show that $f_n \in \mathscr{B}^\rho(\mathbb{R})$ lies in the domain of the generator A of the semigroup $(P_t)_{t\in\mathbb{R}_+}$ we have to prove convergence of the difference quotient with respect to $\|\|_\rho$:

$$\limsup_{t\searrow 0} \sup_{s\in\mathbb{R}} \left| \frac{1}{\rho(s)} \left(\frac{f_n(s+t) - f_n(s)}{t} - g_n(s) \right) \right| = 0.$$

For any $s \in \mathbb{R}$

$$\begin{aligned} |f_n(s+t) - f_n(s) - tg_n(s)| &= \left| \int_s^{s+t} g_n(r)dr - tg_n(s) \right| \\ &\leq \left| \int_s^{s+t} \left(g_n(s) + \sup_{s\in\mathbb{R}} \left| \frac{d}{ds} g_n(s) \right| \right) dr - tg_n(s) \right| \\ &\leq t \left(\sup_{s\in\mathbb{R}} \left| \frac{d}{ds} g_n(s) \right| \right) \\ &\leq tn\,(n-1)\,. \end{aligned}$$

This inequality shows the convergence of the difference quotient with respect to $\|\|_\rho$ and

$$Af_n = g_n.$$

In particular, at $s = 1$

$$Af_n(1) = n.$$

Since $n \in \mathbb{N}$, $n \geq 2$ was arbitrary, this shows that Property (iii) of Theorem 2.3.60 does not hold.

In the next lemma we see that for generalized Feller semigroups of transport type, the problems encountered in Example 2.3.61 can be overcome by choosing a more appropriate weight function as was done by **[14]**.

LEMMA 2.3.62. *Let (E, ρ) be a weighted space and let $(\psi_t)_{t\in\mathbb{R}_+}$ be a family of maps such that for any $t \in \mathbb{R}_+$*

$$\psi_t : E \to E$$

and such that the conditions from Lemma 2.3.54 are satisfied. Let $(P(t))_{t\in\mathbb{R}_+}$ *be the associated generalized Feller semigroup on* $\mathscr{B}^\rho(E)$ *of transport type, and let* $\omega \in \mathbb{R}$ *and* $M_\omega \geq 1$ *be such that*

$$P(t)\rho \leq M_\omega \exp(\omega t)\rho.$$

Then

$$\tilde{\rho}_\omega(x) := \sup_{t\in\mathbb{R}_+} \exp\left(-\omega t\right)\left(P(t)\rho\right)(x)$$

is an admissible weight function on E *such that the norms* $\|\cdot\|_{\tilde{\rho}_\omega}$ *and* $\|\cdot\|_\rho$ *are equivalent and the bound*

$$\|P(t)\|_{L(\mathscr{B}^{\tilde{\rho}_\omega}(E))} \leq \exp\left(\omega t\right)$$

holds for any $t \in \mathbb{R}_+$.

Proof. By Lemma 2.3.52 there exists $\omega \in \mathbb{R}$ and $M_\omega \geq 1$ such that

$$P(t)\rho \leq M_\omega \exp(\omega t)\rho. \tag{2.3.10}$$

We fix such $\omega \in \mathbb{R}$ and want to find an admissible weight function $\tilde{\rho}_\omega$ such that for any $t \in \mathbb{R}_+$

$$\tilde{\rho}_\omega \geq \exp\left(-\omega t\right)\cdot\left(\tilde{\rho}_\omega \circ \psi_t\right) \tag{2.3.11}$$

holds true as it would permit the estimate

$$\begin{aligned}\|P(t)f\|_{\tilde{\rho}_\omega} &= \sup_{x\in E}\frac{|f\left(\psi_t(x)\right)|}{\tilde{\rho}_\omega(x)}\\ &\leq \exp\left(\omega t\right)\sup_{x\in E}\frac{|f\left(\psi_t(x)\right)|}{\tilde{\rho}_\omega\left(\psi_t(x)\right)}\\ &= \exp\left(\omega t\right)\|f\|_{\tilde{\rho}_\omega}.\end{aligned}$$

Therefore, we define

$$\tilde{\rho}_\omega(x) := \sup_{t\in\mathbb{R}_+} \exp\left(-\omega t\right)\left(P(t)\rho\right)(x)$$

which implies Inequality 2.3.11 and show that this is indeed an admissible weight function. For this purpose, we fix $R > 0$ and we have to show that

$$\{\tilde{\rho}_\omega \leq R\} = \bigcap_{t\in\mathbb{R}_+}\{P(t)\rho \leq \exp\left(\omega t\right)R\}$$

is compact. By Lemma 2.3.54

$$P(t)\rho(x) = \rho(\psi_t(x))$$

holds for any $t \in \mathbb{R}_+$and any $x \in E$. The set

$$\{P(0)\rho \leq \exp\left(\omega\cdot 0\right)R\} = \{\rho \leq R\}$$

is compact hence closed since E is Hausdorff. For any $t \in \mathbb{R}_+$, by continuity of $\psi_t|_{K_R}$,

$$\{P_t\rho \leq \exp(\omega t) R\} \cap \{\rho \leq R\} = \left\{\rho \circ \psi_t|_{K_R} \leq \exp(\omega t) R\right\} \cap \{\rho \leq R\}$$
$$= \left(\psi_t|_{K_R}\right)^{-1} \{\rho \leq \exp(\omega t) R\} \cap \{\rho \leq R\}$$

is closed. Thus, by

$$\{\tilde{\rho}_\omega \leq R\} = \bigcap_{t\in\mathbb{R}_+} \{P(t)\rho \leq \exp(\omega t) R\}$$
$$= \bigcap_{t\in\mathbb{R}_+} \{P_t\rho \leq \exp(\omega t) R\} \cap \{\rho \leq R\}$$

$\{\tilde{\rho}_\omega \leq R\}$ is a closed subset of the compact set $\{\rho \leq R\}$, hence compact. Thus, $\tilde{\rho}_\omega$ is an admissible weight function.

By definition $\rho \leq \tilde{\rho}_\omega$ and by Inequality 2.3.10

$$\tilde{\rho}_\omega(x) \leq M_\omega \rho(x).$$

This shows that the norms $\|\cdot\|_\rho$and $\|\cdot\|_{\tilde{\rho}_\omega}$ are equivalent. □

EXAMPLE 2.3.63. Continuing Example 2.3.61, we take $\omega = \frac{1}{2}$ and obtain the admissible weight function

$$\tilde{\rho}_{\frac{1}{2}}(x) := \begin{cases} \exp\left(\frac{s-1}{2}\right) & \text{if } |s| \leq 1 \\ |s| + 1 & \text{else.} \end{cases}$$

We see that for the family of maps $(f_n)_{n\in\mathbb{N}} \subset \mathscr{B}^{\tilde{\rho}_{\frac{1}{2}}}(\mathbb{R})$ from Example 2.3.61 1 is not a maximum of $\frac{f_n}{\tilde{\rho}_{\frac{1}{2}}}$ anymore, or in other words

$$\sup_{s\in\mathbb{R}} \frac{|f_n(s)|}{\tilde{\rho}_{\frac{1}{2}}(s)} > \frac{f_n(1)}{\tilde{\rho}_{\frac{1}{2}}(1)}.$$

Therefore $(f_n)_{n\in\mathbb{N}} \subset \mathscr{B}^{\tilde{\rho}_{\frac{1}{2}}}(\mathbb{R})$ is not a counterexample to Property (iii) of Theorem 2.3.60. On the contrary, by Theorem 2.3.60 for $h \in \mathcal{D}(A)$ with

$$\max\left(\frac{h}{\tilde{\rho}_{\frac{1}{2}}}, 0\right) \leq \frac{h(z)}{\tilde{\rho}_{\frac{1}{2}}(z)}$$

for some $z \in E$

$$Ah(z) \leq \frac{1}{2} h(z)$$

holds true.

2.3.5. Generalized and Extended Feller Processes. As before, in this subsection (E,ρ) always denotes a weighted space. We let I be some index set and let $J \subset I$ be a finite subset.

For a finite index set $J = \left\{j_1, ..., j_{n(J)}\right\}$ we deal with the product space

$$E^J := E_{j_1} \times ... \times E_{j_{n(J)}},$$

where $E_j = E$ for any $j \in J$. We write any element $x_J \in E^J$ as

$$x_J := \left(x_{j_1}, ..., x_{j_{n(J)}}\right).$$

We recall that by Lemma 2.3.18 ,

$$\left(E^J, \rho^{\otimes J}\right)$$

is a weighted space where

$$\rho^{\otimes J}(x_J) := \rho_{j_1}\left(x_{j_1}\right) \cdots \rho_{j_{n(J)}}\left(x_{j_{n(J)}}\right),$$

with

$$\begin{aligned} \rho_j :\ E_j &\to \mathbb{R} \\ x &\to \rho(x). \end{aligned}$$

We rigorously prove a statement made in Theorem 2.11 in [**14**] and show that for a generalized Feller semigroup $(P(t))_{t\in\mathbb{R}_+}$ with $P(t)1 = 1$ for any $t \in \mathbb{R}_+$, we can define a probability space on which for the coordinate process $(\lambda_t)_{t\in\mathbb{R}_+}$ and $f \in \mathscr{B}^\rho(E)$ the conditional expectation of $f(\lambda_t)$ can be expressed in terms of the generalized Feller semigroup (see Equation 2.3.12).

Definition 2.3.64. Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$, let $\nu \in \mathcal{M}^\rho(E)$ be a probability measure and let $(\lambda_t)_{t\in\mathbb{R}_+}$ be an adapted stochastic process on the filtered probability space

$$\left(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\mathbb{R}_+}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}_\nu\right).$$

If for any $t \geq s \geq 0$ and any $f \in \mathscr{B}^\rho(E)$

$$\mathbb{E}_\nu\left[f(\lambda_t)\middle|\,\mathcal{F}_s\right] = P\left(t-s\right) f(\lambda_s) \tag{2.3.12}$$

holds true $\mathbb{P}_\nu$ -almost surely and

$$\mathbb{P}_\nu \circ \lambda_0^{-1} = \nu$$

then $(\lambda_t)_{t\in\mathbb{R}_+}$ is called *generalized Feller process* with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ and $(P(t))_{t\in\mathbb{R}_+}$ and with initial distribution ν.

We can show the existence of generalized Feller processes:

THEOREM 2.3.65. *Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ such that for all $t \in \mathbb{R}_+$*

$$P(t)1 = 1,$$

and

$$\|P(t)\| \leq Me^{\omega t}$$

for $\omega \in \mathbb{R}$ and $M \geq 1$. Then on the measurable space

$$\left(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\mathbb{R}_+}\right)$$

for any $x_0 \in E$ there exists a measure $\mathbb{P}_{x_0}$ and a right continuous filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ (see Definition A.3.75) such that for any $t \geq s \geq 0$ and any $f \in \mathscr{B}^\rho(E)$ the canonical process $(\lambda_t)_{t\in\mathbb{R}_+}$ is adapted with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$,

$$\mathbb{E}_{x_0}\left[f(\lambda_t)|\,\mathcal{F}_s\right] = P\left(t-s\right)f(\lambda_s) \tag{2.3.13}$$

holds true $\mathbb{P}_{x_0}$ -almost surely, and

$$\mathbb{P}_{x_0} \circ \lambda_0^{-1} = \delta_{x_0}.$$

REMARK 2.3.66. In general, a generalized Feller process $(\lambda_t)_{t\in\mathbb{R}_+}$ is strictly speaking not a Markov process, since Equation 2.3.13 does not necessarily hold true for any positive Borel- measurable map f. This is due to the fact, that indicator functions of Borel sets can be approximated with continuous bounded functions by Corollary 2.3.10 only almost everywhere with respect to one (or finitely many) measure(s), but not necessarily simultaneously with respect to the entire family of measures $\left(\left(p(t-s)\right)(x,\cdot)\right)_{x\in E}$ on $(E, \mathcal{B}(E))$ obtained by Proposition 2.3.52. However, for indicator functions of $C_b(E)$-open sets (see Definition A.3.36) Equation 2.3.13 holds true by dominated convergence. Since by Lemma A.3.37 the $C_b(E)$-open sets generate the Baire σ-algebra $\mathcal{B}_0(E)$ (see Definition A.3.38) we conclude again by dominated convergence that Equation 2.3.13 holds true for any indicator function of sets in $\mathcal{B}_0(E)$. Thus, a generalized Feller process $(\lambda_t)_{t\in\mathbb{R}_+}$ is a Markov process with respect to the measurable space

$$\left(E^{\mathbb{R}_+}, \left(\mathcal{B}_0(E)\right)^{\mathbb{R}_+}\right),$$

its natural filtration, and the probability measure $\mathbb{P}_{x_0}$ restricted to this space.

Note that, for locally compact separable spaces E a generalized Feller process is a Markov process in the classical sense, meaning that Equation 2.3.13 holds for all non-negative Borel-measurable maps. This is the case since thanks to Urysohn's Lemma in the completely regular case (see Proposition 2.3.9) and separability an open set can be approximated pointwise by continuous bounded functions. Therefore, Equation 2.3.13 holds also for any map f that is the indicator function of a set in the Dynkin system generated by the open sets. By Lemma A.3.15 such a Dynkin system is the entire σ-algebra and by Proposition A.3.19 the generalized Feller process is a Markov process.

When we speak of generalized Feller processes usually we mean those obtained via Theorem 2.3.65. As for Markov processes $\mathbb{P}_x := \mathbb{P}_{\delta_x}$ for any $x \in E$.

We remind the reader that while $\mathcal{B}(E^{\mathbb{R}_+}) \supset \mathcal{B}(E)^{\mathbb{R}_+}$ holds true because on $\mathcal{B}(E^{\mathbb{R}_+})$ every projection is continuous, hence measurable with respect to $\mathcal{B}(E^{\mathbb{R}_+})$, the inclusion $\mathcal{B}(E^{\mathbb{R}_+}) \subset \mathcal{B}(E)^{\mathbb{R}_+}$ is in general not true when the topology of E does not have a countable base (see Definition A.1.2).

Proof. The proof has three steps. In the first step, we construct a projective family of probability measures (see Definition 2.1.3) on

$$\left(E^J, \mathcal{B}\left(E^J\right)\right)_{J \subset \mathbb{R}_+,\ \text{finite}}.$$

In the second step we use Theorem A.3.104 and obtain a probability measure on $\left(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\mathbb{R}_+}\right)$. The coordinate process $(\lambda_t)_{t\in\mathbb{R}_+}$ on this space then yields for any $t \geq s \geq 0$ and any $f \in \mathscr{B}^\rho(E)$

$$\mathbb{E}_{x_0}\left[f(\lambda_t) \middle| \mathcal{F}_s^0\right] = P\left(t-s\right) f(\lambda_s) \tag{2.3.14}$$

where $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ is the natural filtration of the coordinate process. In the third step, we take the right continuous extension of this filtration and show Equation 2.3.13.

For the first step, we let $R > 0$ and fix some $x_0 \in E_0$. For any $r \geq 0$ let $p(r)(x_0, \cdot)$ be the Radon measure given by the Riesz representation theorem (Proposition 2.3.52) via

$$P(r)f(x_0) = \int_{E_t} f(y)p(r)(x_0, dy).$$

For $0 \leq r_1 < r_2$ we define $\mu_{x_0}^{R,\{r_1,r_2\}} \in \mathcal{M}^{\rho_{r_1}\otimes\rho_{r_2}}(E_{r_1} \times E_{r_2})$ by Riesz representation (Proposition 2.3.52) on $E_{r_1} \times E_{r_2}$, and Lemma 2.3.67 and the continuous functional

$$f_{r_1} \cdot f_{r_2} \to \int_{E_{r_1}} \left(1_{\left\{\rho_{r_1}(y)<R\right\}} \cdot f_{r_1}(y) \cdot P(r_2 - r_1)f_{r_2}(y)\right) p(r_1)(x_0, dy)$$

for $f_{r_1} \in \mathscr{B}^{\rho_{r_1}}(E_{r_1})$ and $f_{r_2} \in \mathscr{B}^{\rho_{r_2}}(E_{r_2})$ as the unique measure in $\mathcal{M}^{\rho_{r_1}\otimes\rho_{r_2}}(E_{r_1} \times E_{r_2})$ such that

$$\begin{aligned}&\int_{E_{r_1}} \left(1_{\{\rho_{r_1}(y)<R\}} \cdot f_{r_1}(y) \cdot P(r_2 - r_1) f_{r_2}(y)\right) p(r_1)(x_0, dy)\\ &= \int_{E_{r_1}\times E_{r_2}} f_{r_1}(y) f_{r_2}(z) \mu_{x_0}^{R,\{r_1,r_2\}}(dy, dz).\end{aligned}$$

By $P(r_2 - r_1)1 = 1$ we obtain

$$\begin{aligned}\mu_{x_0}^{R,\{r_1,r_2\}}(E_{r_1} \times E_{r_2}) &= p(r_1)\left(x_0, \{\rho_{r_1}(y) < R\}\right)\\ &\le p(r_1)(x_0, E_{r_1})\\ &= P(r_1)1(x_0)\\ &= 1.\end{aligned}$$

Then for any $A \in \mathcal{B}(E_{r_1} \times E_{r_2})$ by monotonicity and boundedness, we can define

$$p_{x_0}^{\{r_1,r_2\}}(A) := \lim_{R\to\infty} \mu_{x_0}^{R,\{r_1,r_2\}}(A).$$

One can easily show that $p_{x_0}^{\{r_1,r_2\}}$ is a measure and by dominated convergence $p_{x_0}^{\{r_1,r_2\}}$ is a probability measure on

$$\left(E_{r_1} \times E_{r_2}, \mathcal{B}\left(E_{r_1} \times E_{r_2}\right)\right).$$

Furthermore, for any $r_3 > r_2$ by Riesz representation (Proposition 2.3.52) on $E_{r_1} \times E_{r_2} \times E_{r_3}$, Lemma 2.3.67 and the continuous functional

$$f_{r_1} \cdot f_{r_2} \cdot f_{r_3} \to \int_{E_{r_1}\times E_{r_2}} 1_{\{\rho_{r_1}(y)<R\}} \cdot 1_{\{\rho_{r_2}(z)<R\}} \cdot f_{r_1}(y) \cdot f_{r_2}(z) \cdot P(r_3 - r_2) f_{r_3}(z) p_{x_0}^{\{r_1,r_2\}}(dy, dz)$$

we define $\mu_{x_0}^{R,\{r_1,r_2,r_3\}}$ as the unique measure in $\mathcal{M}^{\rho_{r_1}\otimes\rho_{r_2}\otimes\rho_{r_3}}(E_{r_1} \times E_{r_2} \times E_{r_3})$ and for any $A \in \mathcal{B}(E_{r_1} \times E_{r_2} \times E_{r_3})$ again by monotonicity and boundedness

$$p_{x_0}^{\{r_1,r_2,r_3\}}(A) := \lim_{R\to\infty} \mu_{x_0}^{R,\{r_1,r_2,r_3\}}(A).$$

Inductively, in this way we can define a family of probability measures

$$\left(p_{x_0}^J\right)_{J\subset\mathbb{R}_+,\,\text{finite}}$$

on the respective measurable spaces

$$\left(E^J, \mathcal{B}\left(E^J\right)\right)_{J\subset\mathbb{R}_+,\,\text{finite}}.$$

By an $\varepsilon/3$ -argument and monotonicity it follows from $\mu_{x_0}^{R,J} \in \mathcal{M}^{\rho^{\otimes J}}(E^J)$ for each finite $J \subset \mathbb{R}_+$ and each $R > 0$ that for each finite $J \subset \mathbb{R}_+$ the measure $p_{x_0}^J$ is inner regular, hence a Radon measure.

In order to apply the Generalized Kolmogorov Extension Theorem (Theorem A.3.104), we need to show that this family is projective, i.e. for any finite J and $i \in J$ and any $A \in (\mathcal{B}(E))^{J\setminus\{i\}}$

$$p_{x_0}^J(A \times E_i) = p_{x_0}^{J\setminus\{i\}}(A).$$

We show this property by induction and start with the case $J = \{r_1, r_2\}$. For $f \in \mathscr{B}^{\rho_{r_1}}(E_{r_1})$

$$\int_{E_{r_1}} f(y)\mu_{x_0}^{R,\{r_1,r_2\}}(dy, E_{r_2}) = \int_{E_{r_1}} f(y)1_{K_R}(y)p(r_1)(x_0, dy)$$

which implies by uniqueness of the Radon measure (see Proposition A.4.15) for any $A \in \mathcal{B}(E_{r_1})$

$$\mu_{x_0}^{R,\{r_1,r_2\}}(A, E_{r_2}) = p(r_1)(x_0, A \cap K_R).$$

Thus,

$$\begin{aligned} p_{x_0}^{\{r_1,r_2\}}(A \times E_{r_2}) &= \lim_{R\to\infty} \mu_{x_0}^{R,\{r_1,r_2\}}(A \times E_{r_2}) \\ &= p(r_1)(x, A). \end{aligned}$$

Furthermore, for any $f \in C_b(E_{r_2})$ it holds true by dominated convergence that

$$\begin{aligned} &\int_{E_{r_1}\times E_{r_2}} 1(y)f(z)p_{x_0}^{\{r_1,r_2\}}(dy, dz) \\ &= \lim_{R\to\infty} \int_{E_{r_1}\times E_{r_2}} 1(y)f(z)\mu_{x_0}^{R,\{r_1,r_2\}}(dy, dz) \\ &= \lim_{R\to\infty} \int_{E_{r_2}} 1_{\{\rho_{r_1}(y)<R\}} P(r_2 - r_1)f(y)p(r_1)(x_0, dy) \\ &= P(r_1)P(r_2 - r_1)f(x) \\ &= P(r_2)f(x) \\ &= \int_{E_{r_2}} f(z)p(r_2)(x_0, dz). \end{aligned}$$

Since the functionals

$$f \to \int_{E_{r_2}} f(z)p_{x_0}^{\{r_1,r_2\}}(E, dz)$$

and

$$f \to \int_{E_{r_2}} f(z)p(r_2)(x_0, dz)$$

coincide on $C_b(E_{r_2})$ and satisfy the conditions of Proposition A.4.15, by uniqueness of the Radon measure in Proposition A.4.15 for any $A \in \mathcal{B}(E_{r_2})$

$$p_{x_0}^{\{r_1,r_2\}}(E_{r_1} \times A) = p(r_2)(x_0, A).$$

This implies in particular that for any $f_{r_2} \in \mathscr{B}^{\rho_{r_2}}(E_{r_2})$

$$\int_{E_{r_1}\times E_{r_2}} 1(y) f_{r_2}(z) p_{x_0}^{\{r_1,r_2\}}(dy, dz) < \infty.$$

Next, we assume that there is $N \in \mathbb{N}$ such that for any $n \leq N$ and any arbitrary finite index set $J_n := \{r_1, ..., r_n\} \in \mathbb{R}_+^n$, $0 \leq r_1 < ... < r_n$ for any $i \in \{1, ..., n\}$ and any $A \in \mathcal{B}\left(E^{J_n \setminus \{r_i\}}\right)$

$$p_{x_0}^{J_n}(A \times E_{r_i}) = p_{x_0}^{J_n \setminus \{r_i\}}(A),$$

and for any $f_{r_n} \in \mathscr{B}^{\rho_{r_n}}(E_{r_n})$

$$\int_{E_{r_1}\times ... \times E_{r_n}} 1(y_1) \cdot ... 1(y_{n-1}) \cdot f_{r_n}(y_n) p_{x_0}^{J_n}(dy_1, ..., dy_n) < \infty.$$

We want to show that for $J_{N+1} := \{r_1, ..., r_{N+1}\} \in \mathbb{R}_+^{N+1}$, where $r_{N+1} > ... > r_1 \geq 0$ are arbitrary, for any $i \in \{1, ..., N+1\}$ and any $A \in \mathcal{B}\left(E^{J_{N+1} \setminus \{r_i\}}\right)$

$$p_{x_0}^{J_{N+1}}(A \times E_{r_i}) = p_{x_0}^{J_{N+1} \setminus \{r_i\}}(A),$$

and for any $f_{r_{N+1}} \in \mathscr{B}^{\rho_{r_{N+1}}}(E_{r_{N+1}})$

$$\int_{E_{r_1}\times ... \times E_{r_{N+1}}} 1(y_1) \cdot ... 1(y_N) \cdot f_{r_{N+1}}(y_{N+1}) p_{x_0}^{J_{N+1}}(dy_1, ..., dy_{N+1}) < \infty.$$

In case $i = N+1$ and for bounded $f_{r_1} \in \mathscr{B}^{\rho_{r_1}}(E_{r_1})$,...,$f_{r_N} \in \mathscr{B}^{\rho_{r_N}}(E_{r_N})$

$$\begin{aligned}
&\int_{E_{r_1}\times ... \times E_{r_{N+1}}} f_{r_1}(y_1) \cdot ... \cdot f_{r_N}(y_N) 1(y_{N+1}) p_{x_0}^{J_{N+1}}(dy_1, ..., dy_{N+1}) \\
&= \lim_{R\to\infty} \int_{E_{r_1}\times ... \times E_{r_N}} f_{r_1}(y_1) \cdot ... \cdot f_{r_N}(y_N) \underbrace{P(r_{N+1} - r_N)1}_{=1}(y_N) \mu_{x_0}^{R,J_N}(dy_1, ..., dy_N) \\
&= \int_{E_{r_1}\times ... \times E_{r_N}} f_{r_1}(y_1) \cdot ... \cdot f_{r_N}(y_N) p_{x_0}^{J_{N+1}}(dy_1, ..., dy_N),
\end{aligned}$$

and we conclude by uniqueness of the Radon measure and Lemma 2.3.67 that for any $A \in \mathcal{B}\left(E^{J_{N+1} \setminus \{r_{N+1}\}}\right)$

$$p_{x_0}^{J_{N+1}}(A \times E_{r_{N+1}}) = p_{x_0}^{J_{N+1} \setminus \{r_{N+1}\}}(A),$$

since the corresponding functionals coincide on $C_b(E^{J_{N+1}\setminus\{r_{N+1}\}})$ and satisfy the conditions of Proposition A.4.15. Furthermore,

$$\begin{aligned}&\int_{E_{r_1}\times\ldots\times E_{r_{N+1}}} 1(y_1)\cdot\ldots 1(y_N)\cdot f_{r_{N+1}}(y_{N+1})p_{x_0}^{J_{N+1}}(dy_1,\ldots,dy_{N+1})\\ &=\int_{E_{r_1}\times\ldots\times E_{r_N}} 1(y_1)\cdot\ldots 1(y_N)\cdot P(r_{N+1}-r_N)f_{r_{N+1}}(y_N)p_{x_0}^{J_N}(dy_1,\ldots,dy_N)\\ &<\infty,\end{aligned}$$

by assumption.
In case $i=N$ and for $f_{r_1}\in C_b(E_{r_1})$,...,$f_{r_{N-1}}\in C_b(E_{r_{N-1}})$ and $f_{r_{N+1}}\in C_b(E_{r_{N+1}})$ by dominated convergence

$$\begin{aligned}&\int_{E_{r_1}\times\ldots\times E_{r_{N+1}}} f_{r_1}(y_1)\cdot\ldots\cdot f_{r_{N-1}}(y_{N-1})1(y_N)f_{r_{N+1}}(y_{N+1})p_{x_0}^{J_{N+1}}(dy_1,\ldots,dy_{N+1})\\ &=\lim_{R\to\infty}\int_{E_{r_1}\times\ldots\times E_{r_N}} f_{r_1}(y_1)\cdot\ldots\cdot f_{r_{N-1}}(y_{N-1})\cdot P(r_{N+1}-r_N)f_{r_{N+1}}(y_N)\mu_{x_0}^{R,J_N}(dy_1,\ldots,dy_N)\\ &=\int_{E_{r_1}\times\ldots\times E_{r_N}} f_{r_1}(y_1)\cdot\ldots f_{r_{N-1}}(y_{N-1})\cdot P(r_{N+1}-r_N)f_{r_{N+1}}(y_N)p_{x_0}^{J_N}(dy_1,\ldots,dy_N)\\ &=\int_{E_{r_1}\times\ldots\times E_{r_{N-1}}} f_{r_1}(y_1)\cdot\ldots\cdot f_{r_{N-1}}(y_{N-1})\cdot P(r_{N+1}-r_{N-1})f_{r_{N+1}}(y_{N-1})p_{x_0}^{J_{N-1}}(dy_1,\ldots,dy_{N-1})\\ &=\int_{E_{r_1}\times\ldots\times E_{r_{N-1}}\times E_{r_{N+1}}} f_{r_1}(y_1)\cdot\ldots\cdot f_{r_{N-1}}(y_{N-1})\cdot f_{r_{N+1}}(y_{N+1})p_{x_0}^{J_{N+1}\setminus\{r_N\}}(dy_1,\ldots,dy_{N-1},dy_{N+1}),\end{aligned}$$

and we can conclude as before.

For $i\in\{1,\ldots,N-1\}$ the desired properties follow in the same way by definition of $p_{x_0}^{J_{N+1}}$ and integration of continuous bounded functions and from the assumption that the properties hold true for any $n\leq N$. Thus, by induction it follows that for any $m\in\mathbb{N}$ and any arbitrary finite index set $J_m:=\{r_1,\ldots,r_m\}\in\mathbb{R}_+^m$, $0\leq r_1<\ldots<r_m$ for any $i\in\{1,\ldots,m\}$ and any $A\in\mathcal{B}\left(E^{J_m\setminus\{r_i\}}\right)$

$$p_{x_0}^{J_m}(A\times E_{r_i})=p_{x_0}^{J_m\setminus\{r_i\}}(A).$$

Therefore, the family

$$\left(p_{x_0}^J\right)_{J\subset\mathbb{R}_+,\,\text{finite}}$$

is projective.

In the second step of the proof, in order to construct a measure $\mathbb{P}_{x_0}$ for any $x_0\in E$ on $\left(E^{\mathbb{R}_+},\mathcal{B}(E)^{\mathbb{R}_+}\right)$ we want to use Theorem A.3.104. For this purpose since $\mathcal{B}(E)^J\subset\mathcal{B}(E^J)$ for any finite $J\subset\mathbb{R}_+$, we can define the measure

$$\hat{p}_{x_0}^J:=p_{x_0}^J\big|_{\mathcal{B}(E)^J},$$

and we have to find a compact class (see Definition A.3.103) $\mathcal{C}$ in E such that for each $t \in \mathbb{R}_+$ and $A \in \mathcal{B}(E_t)$

$$p_{x_0}^{\{t\}}(A) = \sup\left\{p_{x_0}^{\{t\}}(C) : C \subset A \,\text{and}\; C \in \mathcal{C}\right\}. \tag{2.3.15}$$

We show that

$$\mathcal{C} := \{C : C \text{ compact}, \, C \subset K_R \text{ for some } R \geq 0\}$$

is such a compact class and start by showing that $\mathcal{C}$ is a compact class at all. We choose some arbitrary sequence $\{C_l\}_{l\in\mathbb{N}} \subset \mathcal{C}$ such that

$$\bigcap_{l\in\mathbb{N}} C_l = \emptyset.$$

For C_1 we choose $R_1 \geq 0$ such that $C_1 \subset K_{R_1}$. Then

$$\bigcup_{l\in\mathbb{N}} E \setminus C_l \supset K_{R_1}$$

is an open cover of the compact set K_{R_1} hence finitely many sets, say without loss of generality $\{E \setminus C_1, E \setminus C_2, ..., E \setminus C_m\}$, suffice to cover it. Thus,

$$\bigcap_{l\in\{1,...,m\}} C_l \cap K_{R_1} = \emptyset$$

and $C_1 \subset K_{R_1}$ yields that

$$\bigcap_{l\in\{1,...,m\}} C_l = \emptyset,$$

and $\mathcal{C}$ is a compact class. By the identity

$$\bigcup_{R\geq 0} K_R = E,$$

for any $\varepsilon > 0$ and $t \in \mathbb{R}_+$ there is some $R_\varepsilon \geq 0$ such that

$$p_{x_0}^{\{t\}}(E) - p_{x_0}^{\{t\}}(K_{R_\varepsilon}) < \varepsilon,$$

and by inner regularity of the Radon measure $p_{x_0}^{\{t\}}$ for any $A \in \mathcal{B}(E)$ there is a compact set $A_\varepsilon \subset A$ such that

$$p_{x_0}^{\{t\}}(A) - p_{x_0}^{\{t\}}(A_\varepsilon) < \varepsilon.$$

Hence

$$p_{x_0}^{\{t\}}(A) - p_{x_0}^{\{t\}}(A_\varepsilon \cap K_{R_\varepsilon}) < 2\varepsilon,$$

and Equation 2.3.15 holds true. Thus, the conditions of Theorem A.3.104 are satisfied.

By applying that theorem, we obtain a probability measure $\mathbb{P}_{x_0}$ on the measurable space $\left(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\mathbb{R}_+}\right)$ such that for any finite $J \subset \mathbb{R}_+$ and

$$A \in \mathcal{B}(E)^J \subset \mathcal{B}(E^J)$$

the probability is given by

$$\mathbb{P}_{x_0}\left(\left(\Pi_J^{\mathbb{R}_+}\right)^{-1}(A)\right) = \hat{p}_{x_0}^J(A) = p_{x_0}^J(A) \tag{2.3.16}$$

with $\Pi_J^{\mathbb{R}_+}$ being the projection from $E^{\mathbb{R}_+}$ on E^J as defined in Definition A.1.6. Moreover, by definition

$$\mathbb{P}_{x_0} \circ (\lambda_0)^{-1} = p_{x_0}^{\{0\}} = q(0)(x_0, \cdot) = \delta_{x_0},$$

and for any $f \in \mathscr{B}^\rho(E)$

$$\mathbb{E}_{x_0}\left[f(\lambda_t)\right] = \int_{E_t} f(y)p(t)(x_0, dy) = P(t)f(x_0) < \infty.$$

Let $(\lambda_t)_{t\in\mathbb{R}_+} := (\Pi_t)_{t\in\mathbb{R}_+}$ be the the coordinate process (see Definition A.3.5) on $\left(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\mathbb{R}_+}\right)$ and let $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ be its natural filtration (see Definition A.3.80). We next show Equation 2.3.14 and to that end that for any $f \in \mathscr{B}^\rho(E)$, any $0 \le s < t$, any $F \in \mathcal{F}_s^0$, and any $x_0 \in E$ the equation

$$\mathbb{E}_{x_0}\left[f(\lambda_t) \cdot 1_F\right] = \mathbb{E}_{x_0}\left[P\left(t-s\right)f(\lambda_s) \cdot 1_F\right]$$

holds true. Since $\mathbb{E}_{x_0}\left[f(\lambda_t)\right] < \infty$ by Proposition A.3.31 it is enough to check

$$\mathbb{E}_{x_0}\left[f(\lambda_t) \cdot 1_G\right] = \mathbb{E}_{x_0}\left[P\left(t-s\right)f(\lambda_s) \cdot 1_G\right]$$

for all $G \in \mathcal{G}$ of an intersection stable generator $\mathcal{G} \subset \mathcal{F}_s^0$. The set

$$\left\{\bigcap_{j\in J} (\lambda_j)^{-1}(O_j) : \ J \subset \mathbb{R}_+, \ \text{finite}, O_j \subset E_j \text{ open for all } j \in J\right\}$$

is such an intersection stable generator. For any $x_0 \in E$ we fix $k \in \mathbb{N}$ and $0 \le r_1 \le r_2 \le ... \le r_k \le s$ and set $J^{'} := \{r_1, r_2, ..., r_k, s, t\}$. For $O_{r_1} \in E_{r_1}$,..., $O_{r_k} \in E_{r_k}$ open by definition

$$\begin{aligned}&\mathbb{E}_{x_0}\left[f(\lambda_t) \cdot 1_{O_{r_1}}(\lambda_{r_1}) \cdot ... \cdot 1_{O_{r_k}}(\lambda_{r_k})\right]\\ &= \int_{E^{J^{'}}} \left(f(t) \cdot 1_{O_{r_1}}(r_1) \cdot ... \cdot 1_{O_{r_k}}(r_k)\right) p_{x_0}^{J^{'}}(dr_1, ..., dr_k, ds, dt)\end{aligned}$$

and

$$\mathbb{E}_{x_0}\left[P(t-s)f(\lambda_s)\cdot 1_{O_{r_1}}(\lambda_{r_1})\cdot ...\cdot 1_{O_{r_k}}(\lambda_{r_k})\right]$$
$$=\int_{E^{J'\setminus\{t\}}}\left(P(t-s)f(s)\cdot 1_{O_{r_1}}(r_1)\cdot ...\cdot 1_{O_{r_k}}(r_k)\right)p_{x_0}^{J'\setminus\{t\}}(dr_1,...,dr_k,ds).$$

By arguing as before by Proposition A.4.15 it is therefore enough to show that the functionals

$$f\to\int_{E^{J'}}\left(f(t)\cdot 1_{O_{r_1}}(r_1)\cdot ...\cdot 1_{O_{r_k}}(r_k)\right)p_{x_0}^{J'}(dr_1,...,dr_k,ds,dt)$$

and

$$f\to\int_{E^{J'\setminus\{t\}}}\left(P(t-s)f(s)\cdot 1_{O_{r_1}}(r_1)\cdot ...\cdot 1_{O_{r_k}}(r_k)\right)p_{x_0}^{J'\setminus\{t\}}(dr_1,...,dr_k,ds)$$

coincide of $C_b(E_s)$.

By Corollary 2.3.10 there exists sequences of maps $(b_l^i)_{l\in\mathbb{N},i\in\{1,...,k\}}$ where $b_l^i\in C_b(E_{r_i})$ for any $l\in\mathbb{N}$ and $i\in\{1,...,k\}$ such that

$$\prod_{i\in\{1,...,k\}}b_l^i\to\prod_{i\in\{1,...,k\}}1_{O_{r_i}\times...\times O_{r_k}\times E_s\times E_t}$$

$p_{x_0}^{J'}$-almost surely and

$$\prod_{i\in\{1,...,k\}}b_l^i\to\prod_{i\in\{1,...,k\}}1_{O_{r_i}\times...\times O_{r_k}\times E_s}$$

$p_{x_0}^{J'\setminus\{t\}}$-almost surely. For $f\in C_b(E_s)$ by the assumption $P(t-s)1=1$ the map $P(t-s)f$ remains bounded and

$$\begin{aligned}
&\int_{E^{J'}}\left(f(t)\cdot 1_{O_{r_1}}(r_1)\cdot ...\cdot 1_{O_{r_k}}(r_k)\right)p_{x_0}^{J'}(dr_1,...,dr_k,ds,dt)\\
&=\lim_{l\to\infty}\int_{E^{J'}}\left(f(t)\cdot b_l^1(r_1)\cdot ...\cdot b_l^k(r_k)\right)p_{x_0}^{J'}(dr_1,...,dr_k,ds,dt)\\
&=\lim_{l\to\infty}\lim_{R\to\infty}\int_{E^{J'}}\left(f(t)\cdot b_l^1(r_1)\cdot ...\cdot b_l^k(r_k)\right)\mu_{x_0}^{R,J'}(dr_1,...,dr_k,ds,dt)\\
&=\lim_{l\to\infty}\lim_{R\to\infty}\int_{E^{J'\setminus\{t\}}}1_{\{\rho_{r_1}(y)<R\}}\cdot ...\cdot 1_{\{\rho_{r_k}(y)<R\}}\cdot ...\\
&\cdot\left(P(t-s)f(s)\cdot b_l^1(r_1)\cdot ...\cdot b_l^k(r_k)\right)p_{x_0}^{J'\setminus\{t\}}(dr_1,...,dr_k,ds)\\
&=\lim_{l\to\infty}\int_{E^{J'\setminus\{t\}}}\left(P(t-s)f(s)\cdot b_l^1(r_1)\cdot ...\cdot b_l^k(r_k)\right)p_{x_0}^{J'\setminus\{t\}}(dr_1,...,dr_k,ds)\\
&=\int_{E^{J'\setminus\{t\}}}\left(P(t-s)f(s)\cdot 1_{O_{r_1}}(r_1)\cdot ...\cdot 1_{O_{r_k}}(r_k)\right)p_{x_0}^{J'\setminus\{t\}}(dr_1,...,dr_k,ds).
\end{aligned}$$

This shows Equation 2.3.14.

In the third step of the proof, we show that for the right continuous enlargement (see Definition A.3.76)

$$(\mathcal{F}_t)_{t\in\mathbb{R}_+} := \left(\mathcal{F}^0_{t+}\right)_{t\in\mathbb{R}_+}$$

the equation

$$\mathbb{E}_{x_0}\left[f(\lambda_t)\middle|\,\mathcal{F}_s\right] = P\left(t-s\right)f(\lambda_s)$$

holds as well $\mathbb{P}_{x_0}$- almost surely for $f \in \mathscr{B}^\rho(E)$ and $t \geq s \geq 0$ and any $x_0 \in E$. We fix such x_0 and $f \in \mathscr{B}^\rho(E)$. By Proposition A.3.95

$$\mathbb{E}_{x_0}\left[f(\lambda_t)\middle|\,\mathcal{F}_s\right] = \lim_{r\searrow s}\mathbb{E}_{x_0}\left[f(\lambda_t)\middle|\,\mathcal{F}^0_r\right] \tag{2.3.17}$$

holds true $\mathbb{P}_{x_0}$- almost surely for any $t \geq s \geq 0$. Thus, it is sufficient to show

$$P\left(t-s\right)f(\lambda_s) = \lim_{r\searrow s} P\left(t-r\right)f(\lambda_r)$$

$\mathbb{P}_{x_0}$- almost surely for any $t \geq s \geq 0$. Since we know by Equation 2.3.17 that the $\mathbb{P}_{x_0}$- almost sure limit exists, by Proposition A.3.53 it is enough to show that

$$P\left(t-s\right)f(\lambda_s) = \lim_{r\searrow s} P\left(t-r\right)f(\lambda_r)$$

holds in $\mathbb{P}_{x_0}$- probability for any $t \geq s \geq 0$. We show this first for $s = 0$. In this case, the left hand side is deterministic by $\mathbb{P}_{x_0}\circ(\lambda_0)^{-1} = \delta_{x_0}$, hence by Proposition A.3.52 it is sufficient to show convergence in law (see Definition A.3.51). Therefore, we have to show that for any $h \in C_b(\mathbb{R})$

$$\lim_{r\searrow 0}\mathbb{E}_{x_0}\left[h\left(P\left(t-r\right)f(\lambda_r)\right)\right] = h\left(P\left(t\right)f(x_0)\right). \tag{2.3.18}$$

The map

$$\begin{aligned}[0,t] &\to \mathscr{B}^\rho(E)\\ r &\to P\left(t-r\right)f\end{aligned}$$

is continuous by Theorem 2.3.51 and

$$\begin{aligned}\mathscr{B}^\rho(E) &\to \mathscr{B}^\rho(E)\\ P\left(t-r\right)f &\to h\circ\left(P\left(t-r\right)f\right)\end{aligned}$$

is continuous by Lemma 2.3.45. Thus, since $[0,t]$ is compact and images of compact sets under continuous mappings are compact, by strong continuity of $(P(t))_{t\in\mathbb{R}_+}$ Lemma 1.4.7 yields

$$\begin{aligned}
&\lim_{r\searrow 0}\left(\mathbb{E}_{x_0}\left[h\left(P\left(t-r\right)f(\lambda_r)\right)\right]-\left(h\circ P\left(t\right)f\right)\left(x_0\right)\right)\\
&=\lim_{r\searrow 0}\left(P\left(r\right)\left(h\circ P\left(t-r\right)f\right)\left(x_0\right)-\left(h\circ P\left(t\right)f\right)\left(x_0\right)\right)\\
&=\left(\lim_{r\searrow 0}P\left(r\right)\left(\lim_{r\searrow 0}h\circ P\left(t-r\right)f\right)\left(x_0\right)-\left(h\circ P\left(t\right)f\right)\left(x_0\right)\right)\\
&=0.
\end{aligned}$$

Thus, Equation 2.3.18 holds and

$$P\left(t\right)f(\lambda_0)=\lim_{r\searrow 0}P\left(t-r\right)f(\lambda_r)$$

in $\mathbb{P}_{x_0}$- probability for $t\geq 0$. We still need to show

$$P\left(t-s\right)f(\lambda_s)=\lim_{r\searrow s}P\left(t-r\right)f(\lambda_r)$$

$\mathbb{P}_{x_0}$- almost surely for any $t\geq s\geq 0$. For $\varepsilon>0$, by definition of $\mathscr{B}^{\rho_0\cdot\rho_{r-s}}\left(E_0\times E_{r-s}\right)$ there exists $\left(f_n\right)_{n\in\mathbb{N}}\subset C_b\left(E_0\times E_{r-s}\right)$ such that $f_n\left(x,y\right)\to P\left(t-s\right)f\left(x\right)-P\left(t-r\right)f\left(y\right)$ for any $(x,y)\in E_0\times E_{r-s}$. Then by dominated convergence

$$\begin{aligned}
&\lim_{r\searrow s}\mathbb{E}_{x_0}\left[1_{|P(t-s)f(\lambda_0)-P(t-r)f(\lambda_{r-s})|>\varepsilon}\circ\theta_s\right]\\
&=\lim_{r\searrow s}\lim_{n\to\infty}\mathbb{E}_{x_0}\left[1_{|f_n(\lambda_0,\lambda_{r-s})|>\varepsilon}\circ\theta_s\right].
\end{aligned}$$

The set

$$O_n:=|f_n(\lambda_0,\lambda_{r-s})|>\varepsilon$$

is open, hence by Corollary 2.3.10 there is a sequence $\left(h_{n,m}\right)_{m\in\mathbb{N}}\subset C_b\left(E_0\times E_{r-s}\right)$ such that $\mathbb{P}_{x_0}$-almost surely

$$1_{O_n}=\lim_{m\to\infty}h_{n,m},$$

and $0\leq h_{n,m}\leq 1_{O_n}$. By Lemma 2.3.67 we can approximate $h_{n,m}$ by cylinder functions and by Proposition 2.3.52 (iv) we obtain

$$\begin{aligned}
&\lim_{r\searrow s}\mathbb{E}_{x_0}\left[1_{|P(t-s)f(\lambda_0)-P(t-r)f(\lambda_{r-s})|>\varepsilon}\circ\theta_s\right]\\
&=\lim_{r\searrow s}\lim_{n\to\infty}\lim_{m\to\infty}\mathbb{E}_{x_0}\left[h_{n,m}\circ\theta_s\right]\\
&=\lim_{r\searrow s}\lim_{n\to\infty}\lim_{m\to\infty}\mathbb{E}_{x_0}\left[\mathbb{E}_{\lambda_s}\left[h_{n,m}\right]\right]\\
&\leq\lim_{r\searrow s}\lim_{n\to\infty}\mathbb{E}_{x_0}\left[\mathbb{E}_{\lambda_s}\left[1_{O_n}\right]\right]\\
&=\lim_{r\searrow s}\mathbb{E}_{x_0}\left[\mathbb{E}_{\lambda_s}\left[1_{|P(t-s)f(\lambda_0)-P(t-r)f(\lambda_{r-s})|>\varepsilon}\right]\right]\\
&=0.
\end{aligned}$$

This yields

$$P\,(t-s)\,f(\lambda_s) = \lim_{r \searrow s} P\,(t-r)\,f(\lambda_r)$$

in $\mathbb{P}_{x_0}$- probability hence $\mathbb{P}_{x_0}$- almost surely since we know by Equation 2.3.17 that the $\mathbb{P}_{x_0}$- almost sure limit exists. Thus,

$$\mathbb{E}_{x_0}\left[\, f(\lambda_t) \middle|\, \mathcal{F}_s\right] = P\,(t-s)\,f(\lambda_s).$$

□

Adapting a proof in [**4**], we can show:

LEMMA 2.3.67. *Let* (E_i, ρ_i) $i \in \{1, ..., n\}$ *be weighted spaces and*

$$\rho\,(x_1, ..., x_n) := \rho_1\,(x_1) \cdots \rho_n\,(x_n)\,.$$

Then the linear map

$$\begin{aligned} \Psi: \quad \mathscr{B}^{\rho_1}(E_1) \otimes ... \otimes \mathscr{B}^{\rho_n}(E_n) &\to \mathscr{B}^{\rho}(E_1 \times ... \times E_n) \\ f_1 \otimes ... \otimes f_n &\to f_1 \cdots f_n \end{aligned}$$

is injective and its image is a dense linear subspace of $\mathscr{B}^{\rho}(E_1 \times ... \times E_n)$.

PROOF. First, we observe that by Lemma 2.3.18

$$(E_1 \times ... \times E_n, \rho)$$

is indeed a weighted space. Furthermore, for $f_i \in \mathscr{B}^{\rho_i}(E_i)$, $i \in \{1, ..., n\}$ the map

$$(x_1, ..., x_n) \to f_1(x_1) \cdots f_n(x_n)$$

is in $\mathscr{B}^{\rho}(E_1 \times ... \times E_n)$. In order to see this, observe that clearly

$$\sup_{(x_1,...,x_n) \in E_1 \times ... \times E_n} \left| \frac{f_1(x_1) \cdots f_n(x_n)}{\rho_1\,(x_1) \cdots \rho_n\,(x_n)} \right| < \infty.$$

Let $\varepsilon > 0$. For sequences $(g_i^m)_{m \in \mathbb{N}} \subset C_b(E_i)$, $i \in \{1, ..., n\}$ such that

$$\lim_{m \to \infty} \|g_i^m - f_i\|_{\rho_i} = 0$$

we obtain for $(g^m)_{m \in \mathbb{N}} \subset C_b(E_1 \times ... \times E_n)$ defined as

$$g^m := g_1^m \cdots g_n^m,$$

for m large enough the inequality

$$\begin{aligned}
&\left\| f_1 \cdots f_n - g^m \right\|_\rho \\
&= \sup_{(x_1,\dots,x_n)\in E_1\times\dots\times E_n} \left| \frac{f_1(x_1)\cdots f_n(x_n)}{\rho_1(x_1)\cdots\rho_n(x_n)} - \frac{g_1^m(x_1)\cdot f_2(x_2)\cdot\cdot f_n(x_n)}{\rho_1(x_1)\cdots\rho_n(x_n)} \right. \\
&\quad \left. + \frac{g_1^m(x_1)\cdot f_2(x_2)\cdot\cdot f_n(x_n)}{\rho_1(x_1)\cdots\rho_n(x_n)} - \frac{g_1^m(x_1)\cdots g_n^m(x_n)}{\rho_1(x_1)\cdots\rho_n(x_n)} \right| \\
&\leq \sup_{(x_1,\dots,x_n)\in E_1\times\dots\times E_n} \left(\underbrace{\left| \frac{f_1(x_1)-g_1^m(x_1)}{\rho_1(x_1)} \right|}_{<\varepsilon} \left| \frac{f_2(x_2)\cdots f_n(x_n)}{\rho_2(x_2)\cdots\rho_n(x_n)} \right| \right. \\
&\quad \left. + \underbrace{\left| \frac{g_1^m(x_1)}{\rho_1(x_1)} \right|}_{<\left\| f_i \right\|_{\rho_i}+\varepsilon} \left| \frac{f_2(x_2)\cdot\cdot f_n(x_n)}{\rho_2(x_2)\cdots\rho_n(x_n)} - \frac{g_2^m(x_2)\cdots g_n^m(x_n)}{\rho_2(x_2)\cdots\rho_n(x_n)} \right| \right),
\end{aligned}$$

and hence inductively

$$\lim_{m\to\infty} \left\| g^m - f \right\|_\rho = 0.$$

Thus,

$$(x_1, ..., x_n) \to f_1(x_1)\cdots f_n(x_n)$$

is in $\mathscr{B}^\rho(E_1 \times ... \times E_n)$.

The map

$$\begin{gathered}
\mathscr{B}^{\rho_1}(E_1) \times ...\mathscr{B}^{\rho_n}(E_n) \to \mathscr{B}^\rho(E_1 \times ... \times E_n) \\
(f_1, ..., f_n) \to f_1 \cdot ... \cdot f_n
\end{gathered}$$

is multilinear. Therefore, by definition of the tensor product (Proposition A.2.7) there exists a linear map

$$\begin{gathered}
\mathscr{B}^{\rho_1}(E_1) \otimes ... \otimes \mathscr{B}^{\rho_1}(E_1) \to \mathscr{B}^\rho(E_1 \times ... \times E_n) \\
f_1 \otimes ... \otimes f_n \to f_1 \cdots f_n.
\end{gathered}$$

In order to show injectivity of this map, for

$$0 \neq u \in \mathscr{B}^{\rho_1}(E_1) \otimes ... \otimes \mathscr{B}^{\rho_n}(E_n)$$

according to Lemma A.2.8 we choose a representation

$$u = \sum_{j=1}^{m} f_1^j \otimes ... \otimes f_n^j,$$

with $\left\{f_i^j\right\}_{j\in\{1,...,m\}} \subset \mathscr{B}^{\rho_i}(E_i)$ for any $i \in \{1,...,n\}$ and

$$\left\{f_1^j\right\}_{j\in\{1,...,m\}}, ..., \left\{f_n^j\right\}_{j\in\{1,...,m\}}$$

linearly independent. We need to show that

$$\sum_{j=1}^{m} f_1^j \cdot ... \cdot f_n^j \neq 0.$$

It is enough to observe that by linear independence of

$$\left\{f_1^j\right\}_{j\in\{1,...,m\}}, ..., \left\{f_{n-1}^j\right\}_{j\in\{1,...,m\}}$$

there are $z_i \in E_i$ such that $f_i^1(z_i) \neq 0$ for any $i \in \{1, ..., m-1\}$, hence

$$f_1^1(z_1) \cdots f_{n-1}^1(z_{n-1}) \neq 0,$$

and by linear independence of $\{f_n^j\}_{j\in\{1,...,m\}}$

$$\sum_{j=1}^{m} f_1^j(z_1) \cdots f_{n-1}^j(z_{n-1}) f_n^j \neq 0.$$

Thus, there is some $z_n \in E_n$ such that

$$\sum_{j=1}^{m} f_1^j(z_1) \cdots f_{n-1}^j(z_{n-1}) f_n^j(z_n) \neq 0,$$

and Ψ is injective.

Density of the image of Ψ follows directly from Stone-Weierstrass for $\mathscr{B}^{\rho}$-spaces (Proposition 2.3.46) as the image contains an algebra that separates points and contains $1_{E_1\times...\times E_n}$. □

COROLLARY 2.3.68. *Let* (E_i, ρ_i) $i \in \{1, ..., n\}$ *be weighted spaces and*

$$\rho(x_1, ..., x_n) := \rho_1(x_1) \cdots \rho_n(x_n).$$

Then the linear map

$$\begin{aligned} \Psi: \quad & \widetilde{B^{\rho_1}}(E_1) \otimes ... \otimes \widetilde{B^{\rho_n}}(E_n) \to \widetilde{B^{\rho}}(E_1 \times ... \times E_n) \\ & f_1 \otimes ... \otimes f_n \to f_1 \cdots f_n \end{aligned}$$

is injective and its image is a linear subspace of $D \subset \widetilde{B^{\rho}}(E_1 \times ... \times E_n)$ *given by*

$$D := \left\{ \sum_{j=1}^{m} f_1^j \cdot ... \cdot f_n^j,\, m \in \mathbb{N},\, f_i^j \in \widetilde{B^{\rho_i}}(E_i), i \in \{1, ..., n\}, j \in \{1, ..., m\} \right\}.$$

PROOF. The proof is a simplified version of the proof of Lemma 2.3.67. □

In Theorem 2.3.65 due to problems with the measurability of

$$x \to \mathbb{P}_x[A]$$

for $A \in \mathcal{B}(E)^{\mathbb{R}_+}$ as initial distributions of the generalized Feller process we could only use Dirac measures. However, when the admissible weight function is Baire measurable it is possible to use initial distributions in $\mathcal{M}^\rho(E)$:

PROPOSITION 2.3.69. *Let ρ be Baire measurable. Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ such that for all $t \in \mathbb{R}_+$*

$$P(t)1 = 1,$$

and

$$\|P(t)\| \leq Me^{\omega t}$$

for $\omega \in \mathbb{R}$ and $M \geq 1$. Then on the measurable space

$$\left(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\mathbb{R}_+}\right)$$

for any probability measure $\nu \in \mathcal{M}^\rho(E)$ there exists a measure $\mathbb{P}_\nu$ with mass $\nu(E)$ and a right continuous filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$(see Definition A.3.75) such that for any $t \geq s \geq 0$ and any $f \in \mathscr{B}^\rho(E)$ the canonical process $(\lambda_t)_{t\in\mathbb{R}_+}$ is adapted with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$,

$$\mathbb{E}_\nu\left[f(\lambda_t)|\,\mathcal{F}_s\right] = P\left(t-s\right)f(\lambda_s) \tag{2.3.19}$$

holds true $\mathbb{P}_\nu$ -almost surely, and

$$\mathbb{P}_\nu \circ \lambda_0^{-1} = \nu.$$

PROOF. By Theorem 2.3.65 for any $x_0 \in E$ for the Dirac distribution $\nu = \delta_{x_0}$ the statement of the proposition holds true. Moreover, we observe that by definition for any $R > 0$ the map $1_{\{y\in E:\, \rho(y)<R\}}$ is Baire measurable. By Proposition 2.3.52 $(p(t))$ is a semigroup of transition probabilities with respect to the Baire σ-algebra $\mathcal{B}_0(E)$. Thus, we can show inductively that for any $R > 0$, any finite $J \subset \mathbb{R}_+$ and any $A \in (\mathcal{B}_0(E))^J$for

$$\left(\mu_{x_0}^{R,J}\right)_{J\subset\mathbb{R}_+,\,\text{finite}}$$

and for

$$\left(p_{x_0}^{J}\right)_{J\subset\mathbb{R}_+,\,\text{finite}}$$

as defined in Theorem 2.3.65 the maps

$$x_0 \to \mu_{x_0}^{R,J}(A)$$

and

$$x_0 \to p_{x_0}^{J}(A)$$

are measurable with respect to $\mathcal{B}_0(E)$. Therefore, for $0 \leq r_1 < r_2$ we can define $\mu_\nu^{R,\{r_1,r_2\}} \in \mathcal{M}^{\rho_{r_1}\otimes\rho_{r_2}}(E_{r_1} \times E_{r_2})$ by Riesz representation (Proposition 2.3.52) on $E_{r_1}\times E_{r_2}$, and Lemma 2.3.67 and the continuous functional

$$f_{r_1} \cdot f_{r_2} \to \int_E \left(\int_{E_{r_1}} \left(1_{\left\{\rho_{r_1}(y)<R\right\}} \cdot f_{r_1}(y) \cdot P(r_2 - r_1) f_{r_2}(y) \right) p(r_1)(x, dy) \right) \nu(dx)$$

for $f_{r_1} \in \mathscr{B}^{\rho_{r_1}}(E_{r_1})$ and $f_{r_2} \in \mathscr{B}^{\rho_{r_2}}(E_{r_2})$ as the unique measure in $\mathcal{M}^{\rho_{r_1}\otimes\rho_{r_2}}(E_{r_1} \times E_{r_2})$ such that

$$\begin{aligned} &\int_E \left(\int_{E_{r_1}} \left(1_{\left\{\rho_{r_1}(y)<R\right\}} \cdot f_{r_1}(y) \cdot P(r_2 - r_1) f_{r_2}(y) \right) p(r_1)(x, dy) \right) \nu(dx) \\ &= \int_{E_{r_1}\times E_{r_2}} f_{r_1}(y) f_{r_2}(z) \mu_\nu^{R,\{r_1,r_2\}}(dy, dz). \end{aligned}$$

Inductively, as in the proof of Theorem 2.3.65 we can define for any probability measure $\nu \in \mathcal{M}^\rho(E)$ the family

$$\left(p_\nu^J\right)_{J\subset\mathbb{R}_+,\,\text{finite}}$$

of probability measures on the respective measurable spaces

$$\left(E^J, \mathcal{B}\left(E^J\right)\right)_{J\subset\mathbb{R}_+,\,\text{finite}} .$$

Moreover, for any $n \in \mathbb{N}$ and any arbitrary finite index set $J := \{r_1, ..., r_n\} \in \mathbb{R}_+^n$, $0 \leq r_1 < ... < r_n$ for $f_{r_1} \in C_b(E_{r_1})$,..., $f_{r_n} \in C_b(E_{r_n})$

$$\begin{aligned} &\int_{E_{r_1}\times\ldots\times E_{r_n}} f_{r_1}(y_1)\cdot\ldots\cdot f_{r_n}(y_n) p_\nu^J(dy_1, ..., dy_n). \\ &= \int_E \left(\lim_{R\to\infty} \int_{E_{r_1}\times\ldots\times E_{r_{n-1}}} 1_{\left\{\rho_{r_1}(y_1)<R\right\}} \cdot\ldots\cdot 1_{\left\{\rho_{r_n}(y_n)<R\right\}} \right. \\ &\left. \left(f_{r_1}(y_1)\cdot\ldots\cdot f_{r_{n-1}}(y_{n-1}) P(r_n - r_{n-1}) f_{r_n}(y_{n-1})\right) p_x^{J\setminus\{r_n\}}(dy_1, ..., dy_{n-1}) \right) \nu(dx) \\ &= \int_E \left(\int_{E_{r_1}\times\ldots\times E_{r_n}} f_{r_1}(y_1)\cdot\ldots\cdot f_{r_n}(y_n) p_x^J(dy_1, ..., dy_n) \right) \nu(dx). \end{aligned}$$

Furthermore, by dominated convergence it follows that for any bounded Baire-measurable maps $g_{r_1} \in \ell^\infty(E_{r_1})$,..., $g_{r_n} \in \ell^\infty(E_{r_n})$

$$\begin{aligned} &\int_{E_{r_1}\times\ldots\times E_{r_n}} g_{r_1}(y_1)\cdot\ldots\cdot g_{r_n}(y_n) p_\nu^J(dy_1, ..., dy_n). \\ &= \int_E \left(\int_{E_{r_1}\times\ldots\times E_{r_n}} g_{r_1}(y_1)\cdot\ldots\cdot g_{r_n}(y_n) p_x^J(dy_1, ..., dy_n) \right) \nu(dx). \end{aligned} \tag{2.3.20}$$

This shows with the results of the proof of Theorem 2.3.65 that the family

$$\left(p_\nu^J\right)_{J\subset\mathbb{R}_+,\,\text{finite}}$$

is projective. By generalized Kolmogorov extension theorem (Theorem A.3.104) this yields a probability measure $\mathbb{P}_\nu$ on $\left(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\mathbb{R}_+}\right)$. It follows that

$$\mathbb{P}_\nu\left(\left(\Pi_J^{\mathbb{R}_+}\right)^{-1}(A)\right) = \int_E p_x^J(A)\nu(dx),$$

and

$$\mathbb{P}_\nu \circ (\lambda_0)^{-1} = \nu,$$

and for any $f \in \mathscr{B}^\rho(E)$ and any $r \geq 0$

$$\mathbb{E}_\nu\left[f(\lambda_r)\right] = \int_E P(r)f(x)\nu(dx) < \infty.$$

As in Theorem 2.3.65 one can show that for any $t \geq s \geq 0$ and any $f \in \mathscr{B}^\rho(E)$ for the natural filtration$(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$

$$\mathbb{E}_\nu\left[f(\lambda_t)|\,\mathcal{F}_s^0\right] = P\left(t-s\right)f(\lambda_s)$$

holds true $\mathbb{P}_\nu$ -almost surely.

In the last part of the proof, we show that for the right continuous enlargement $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ of the natural filtration

$$\mathbb{E}_{\mathbb{P}_\nu}\left[f(\lambda_t)|\,\mathcal{F}_s\right] = P\left(t-s\right)f(\lambda_s)$$

holds as well $\mathbb{P}_\nu$- almost surely for $f \in \mathscr{B}^\rho(E)$ and $t \geq s \geq 0$ and any probability measure $\nu \in \mathcal{M}^\rho(E)$. In Theorem 2.3.65 this was shown for any $x_0 \in E$ for the Dirac distribution $\nu = \delta_{x_0}$. As in Theorem 2.3.65

$$P\left(t\right)f(\lambda_0) = \lim_{r\searrow 0} P\left(t-r\right)f(\lambda_r)$$

in $\mathbb{P}_{x_0}$- almost surely for any $t \geq 0$. By definition of the Baire σ-algebra the map

$$x \to 1_{|P(t)f(\lambda_0)-P(t-r)f(\lambda_r)|>\varepsilon}(x)$$

is Baire measurable. Then by Equation 2.3.20 and dominated convergence

$$\begin{aligned}&\lim_{r\searrow 0}\mathbb{E}_\nu\left[1_{|P(t)f(\lambda_0)-P(t-r)f(\lambda_r)|>\varepsilon}\right]\\&= \int_E\left(\lim_{r\searrow 0}\mathbb{E}_{x_0}\left[1_{|P(t)f(\lambda_0)-P(t-r)f(\lambda_r)|>\varepsilon}\right]\right)d\nu(x_0)\\&= 0\end{aligned}$$

and

$$P\left(t\right)f(\lambda_0) = \lim_{r\searrow 0} P\left(t-r\right)f(\lambda_r)$$

in $\mathbb{P}_\nu$- probability hence $\mathbb{P}_\nu$- almost surely since we know by Proposition A.3.95 that

$$\mathbb{E}_{\mathbb{P}_\nu}\left[f(\lambda_t)|\,\mathcal{F}_s\right] = \lim_{r\searrow s}\mathbb{E}_{\mathbb{P}_\nu}\left[f(\lambda_t)|\,\mathcal{F}_r^0\right]$$

and that the $\mathbb{P}_\nu$- almost sure limit exists. The rest follows as in Theorem 2.3.65 $\square$

PROPOSITION 2.3.70. *Let $(\lambda_t)_{t\in\mathbb{R}_+}$ be a generalized Feller process on (E,ρ). For $t \in \mathbb{R}_+$ let*

$$\theta_t : E^{\mathbb{R}_+} \to E^{\mathbb{R}_+}$$

be the map

$$(\omega(s))_{s\in\mathbb{R}_+} \to (\omega(s+t))_{s\in\mathbb{R}_+}.$$

For any finite $J \subset \mathbb{R}_+$ and $i \in J$ let $f \in C_b\left(E^J\right)$ or let $f \in \mathscr{B}^{\rho^{\otimes J}}\left(E^J\right)$ such that

$$f := f_i \cdot g$$

with $g \in C_b(E^{J\setminus\{i\}})$ and $f_i \in \mathscr{B}^{\rho_i}(E_i)$. For any such f let $\mathbb{E}_{\lambda_t}\left[f \circ \Pi_J^{\mathbb{R}_+}\right]$ be the composition of $x \to \mathbb{E}_x\left[f \circ \Pi_J^{\mathbb{R}_+}\right]$ and

$$E^{\mathbb{R}_+} \to E$$

$$\omega \to \lambda_t(\omega).$$

Then if ρ is Baire measurable then for any $t > 0$ and any $\nu \in \mathcal{M}^\rho(E)$

$$\mathbb{E}_\nu\left[f \circ \Pi_J^{\mathbb{R}_+} \circ \theta_t \middle| \mathcal{F}_t^0\right] = \mathbb{E}_{\lambda_t}\left[f \circ \Pi_J^{\mathbb{R}_+}\right] \tag{2.3.21}$$

holds true $\mathbb{P}_\nu$-almost surely.

PROOF. For any $i, j \in J$ denote by $e_{j,i}$ the map

$$E^{\{j\}} \to E^{\{i\}}$$

$$x \to x.$$

For $t \in \mathbb{R}_+$ denote

$$J + t := \bigcup_{j\in J} \{j + t\}.$$

For any f such that $\mathbb{E}_x[f]$ is defined for any $x \in E$ we can approximate f by cylinder functions according to Lemma 2.3.67. Then by Proposition 2.3.52 $\mathbb{E}_{\lambda_t}[f]$ is $\mathcal{F}_t^0$- measurable. We need to show that for any such f and any $A \in \mathcal{F}_t^0$

$$\mathbb{E}_\nu\left[(f \circ \theta_t)\, 1_A\right] = \mathbb{E}_\nu\left[\mathbb{E}_{\lambda_t}[f]\, 1_A\right]$$

holds true. Since one can easily show that the system of sets $D \in \mathcal{F}_t^0$ for which

$$\mathbb{E}_\nu\left[(f \circ \theta_t)\, 1_D\right] = \mathbb{E}_\nu\left[\mathbb{E}_{\lambda_t}[f]\, 1_D\right]$$

holds is a Dynkin system, by Lemma A.3.15 it is enough to show

$$\mathbb{E}_\nu\left[(f \circ \theta_t) 1_G\right] = \mathbb{E}_\nu\left[\mathbb{E}_{\lambda_t}[f] 1_G\right]$$

for any $G \in \mathcal{G}$, where $\mathcal{G}$ is an intersection stable generator of $\mathcal{F}_t^0$. Therefore, choosing $n \in \mathbb{N}$, $0 \leq s_1 < ... < s_n \leq t$ and $F_{s_i} \in \mathcal{B}(E_{s_i})$, $i \in \{1, ..., n\}$ arbitrary and denoting

$$J' := \{s_1, ..., s_n\}$$

it is enough to show that

$$\mathbb{E}_\nu\left[(f \circ \theta_t) 1_{B'}\right] = \mathbb{E}_\nu\left[\mathbb{E}_{\lambda_t}[f] 1_{B'}\right]$$

holds true for

$$B' = \{\lambda_{s_1} \in F_{s_1}, ..., \lambda_{s_n} \in F_{s_n}\}.$$

Let $g := \prod_{j \in J} g_j$ with $g_j \in C_b(E_j)$. We observe that for any $x_0 \in E$

$$\mathbb{E}_{x_0}\left[g \circ \Pi_J^{\mathbb{R}_+}\right] = P(s_1)\left(g_{s_1} \cdot ...P(s_{n-1} - s_{n-2})\left(g_{s_{n-1}} \cdot P(s_n - s_{n-1}) g_{s_n}\right)\right)(x_0). \tag{2.3.22}$$

Denote

$$\tilde{J} := (J + t) \cup \{t\} \cup J'.$$

By definition

$$\begin{aligned}&\mathbb{E}_\nu\left[\left(g \circ \Pi_J^{\mathbb{R}_+} \circ \theta_t\right) 1_{B'}\right] \\ &= \left(\int_{E^{\tilde{J}}} \left(\prod_{j \in J} g_j\left(e_{t+j,j}\left(x_{t+j}\right)\right)\right) 1_{\left\{x_{s_1} \in F_{s_1}, ..., x_{s_n} \in F_{s_n}\right\}} p_\nu^{\tilde{J}}\left(dx_{\tilde{J}}\right)\right).\end{aligned}$$

Since $p_\nu^{\tilde{J}}$ is a Radon measure for any probability measure $\nu \in \mathcal{M}^\rho(E)$, by Corollary 2.3.10 for any such ν and $i \in \{1, ..., n\}$ there exists a sequence $\left(h_{i,x_0}^m\right)_{m \in \mathbb{N}} \subset C_b\left(E_{s_i}\right)$ such that

$$h_{i,x_0}^m 1_{E^{\tilde{J}} \setminus \{s_i\}} \to 1_{\left\{\Pi_{s_i}^{\tilde{J}} \in F_{s_i}\right\}}$$

$p_\nu^{\tilde{J}}$-almost surely and

$$h_{i,x_0}^m 1_{E^{\{t\} \cup J' \setminus \{s_i\}}} \to 1_{\left\{\Pi_{s_i}^{\{t\} \cup J'} \in F_{s_i}\right\}}$$

$p_\nu^{\{t\} \cup J'}$-almost surely (where $\Pi_{s_i}^{\tilde{J}}$ and $\Pi_{s_i}^{\{t\} \cup J'}$ are the projections from Definition A.1.6). Then by Equation2.3.20

$$\begin{aligned}&\left(\int_{E^{\tilde{J}}} \left(\prod_{j \in J} g_j\left(e_{t+j,j}\left(x_{t+j}\right)\right)\right) 1_{\left\{x_{s_1} \in F_{s_1}, ..., x_{s_n} \in F_{s_n}\right\}} p_\nu^{\tilde{J}}\left(dx_{\tilde{J}}\right)\right) \\ &= \int_{E_0} \lim_{m \to \infty} \left(\int_{E^{\tilde{J}}} \left(\prod_{j \in J} g_j\left(e_{t+j,j}\left(x_{t+j}\right)\right)\right) \left(\prod_{i \in \{1, ..., n\}} h_{i,x_0}^m\left(x_{s_i}\right)\right) p_{x_0}^{\tilde{J}}(dx_{\tilde{J}})\right) d\nu(x_0).\end{aligned}$$

Applying the definition of $p_{x_0}^{\tilde{J}}$ multiple times and dominated convergence a comparison with $\mathbb{E}_x\left[g \circ \Pi_J^{\mathbb{R}_+}\right]$ yields

$$\int_{E_0} \lim_{m\to\infty} \left(\int_{E^{\tilde{J}}} \left(\prod_{j\in J} g_j\left(e_{t+j,j}\left(x_{t+j}\right)\right)\right) \left(\prod_{i\in\{1,\ldots,n\}} h_{i,x_0}^m\left(x_{s_i}\right)\right) p_{x_0}^{\tilde{J}}(dx_{\tilde{J}})\right) d\nu(x_0)$$

$$= \int_{E_0} \lim_{m\to\infty} \left(\int_{E^{\tilde{J}}} \left(\mathbb{E}_{x_t}\left[g \circ \Pi_J^{\mathbb{R}_+}\right]\right) \left(\prod_{i\in\{1,\ldots,n\}} h_{i,x_0}^m\left(x_{s_i}\right)\right) p_{x_0}^{\{t\}\cup J'}(dx_{\tilde{J}})\right) d\nu(x_0)$$

$$= \left(\int_{E^{\tilde{J}}} \left(\mathbb{E}_{x_t}\left[g \circ \Pi_J^{\mathbb{R}_+}\right]\right) 1_{\left\{x_{s_1}\in F_{s_1},\ldots,x_{s_n}\in F_{s_n}\right\}} p_{\nu}^{\{t\}\cup J'}\left(dx_{\tilde{J}}\right)\right),$$

where we again used Equation2.3.20 in the last step. Thus,

$$\mathbb{E}_\nu\left[\left(g \circ \Pi_J^{\mathbb{R}_+} \circ \theta_t\right) 1_{B'}\right] = \mathbb{E}_\nu\left[\mathbb{E}_{\lambda_t}\left[g \circ \Pi_J^{\mathbb{R}_+}\right] 1_{B'}\right].$$

For $f \in \mathscr{B}^{\rho^{\otimes J}}\left(E^J\right)$ such that

$$f := f_i \cdot g$$

with $g \in C_b(E^{J\setminus\{i\}})$ and $f_i \in \mathscr{B}^{\rho_i}\left(E_i\right)$

$$\mathbb{E}_\nu\left[\left(f \circ \Pi_J^{\mathbb{R}_+} \circ \theta_t\right) 1_{B'}\right] = \mathbb{E}_\nu\left[\mathbb{E}_{\lambda_t}\left[f \circ \Pi_J^{\mathbb{R}_+}\right] 1_{B'}\right]$$

follows from dominated convergence, which yields the assertion of the proposition. □

In Theorem 2.3.65 for the generalized Feller semigroup $(P(t))_{t\in\mathbb{R}_+}$ we required

$$P(t)1 = 1 \text{ for all } t \in \mathbb{R}_+,$$

which means that for any $t \in \mathbb{R}_+$ $P(t)$ is an isometry with respect to the supremum norm. We would like to replace this condition by one that instead depends on the $\|\cdot\|_\rho$-norm.

We recall the cemetery Δ from Remark 2.1.2 and equip $E \cup \{\Delta\}$ with a topology such that

$$\mathcal{B}(E \cup \{\Delta\}) = \sigma\left(\mathcal{B}(E), \{\Delta\}\right).$$

Consistent with the convention in Remark 2.1.2, we define $\mathscr{B}^\rho(E\cup\{\triangle\}$ as the space of maps f such that $f|_E \in \mathscr{B}^\rho(E)$ and $f(\Delta) = 0$. The space $C_0\left(E \cup \{\triangle\}\right)$ is defined in the same way.

Definition 2.3.71. Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$, let ν be a probability measure on

$$\left(E \cup \{\Delta\}, \mathcal{B}(E \cup \{\Delta\})\right)$$

and let $(\gamma_t)_{t\in\mathbb{R}_+}$ be an adapted stochastic process on the filtered probability space

$$\left((E\cup\{\triangle\})^{\mathbb{R}_+},\mathcal{B}(E\cup\{\Delta\})^{\mathbb{R}_+},(\mathcal{F}_t)_{t\in\mathbb{R}_+},\mathbb{P}'_\nu\right).$$

If for any $t\geq s\geq 0$ and any real-valued map f on $E\cup\{\Delta\}$ that is bounded and Baire-measurable

$$\mathbb{E}_{\mathbb{P}'_\nu}\left[f(\gamma_t)|\,\mathcal{F}_s\right]=\frac{P\,(t-s)\,(f\cdot\rho)}{\rho}(\gamma_s) \tag{2.3.23}$$

holds true $\mathbb{P}'_\nu$ -almost surely and

$$\mathbb{P}'_\nu\circ\gamma_0^{-1}=\nu,$$

then $(\gamma_t)_{t\in\mathbb{R}_+}$ is called *extended Feller process* with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ and with respect to $(P(t))_{t\in\mathbb{R}_+}$ with initial distribution υ.

The reason why these processes a named in such a way will become clear in Theorem 2.3.93.

REMARK 2.3.72. As for generalized Feller processes, due to the subtle measurability conditions extended Feller processes are in general not Markov processes in the classical sense but on separable locally compact spaces this is the case.

For contractive generalized Feller semigroups we obtain existence of extended Feller processes as can be seen in the next theorem. We remind the reader of the convention in Proposition 2.3.52(ii) that for all positive measurable maps $f:\,E\to\mathbb{R}$ (or $f\in\widetilde{B^\rho}(E)$)

$$\tilde{P}(t)f(x):=\int_E f(y)p(t)(x,dy)$$

will simply be written as $P(t)f(x)$.

THEOREM 2.3.73. *Let ρ be measurable with respect to the Baire σ-algebra $\mathcal{B}_0(E)$. Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ such that for all $t\in\mathbb{R}_+$*

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))}\leq 1.$$

Then for any probability measure ν on

$$(E\cup\{\Delta\},\mathcal{B}(E\cup\{\Delta\}))$$

there exists a probability measure $\mathbb{P}'_\nu$ on the measurable space

$$\left((E\cup\{\triangle\})^{\mathbb{R}_+},\mathcal{B}(E\cup\{\Delta\})^{\mathbb{R}_+}\right)$$

such that for the canonical process $(\gamma_t)_{t\in\mathbb{R}_+}$ and the natural filtration $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ for any $t \geq s \geq 0$ and any real-valued map f on $E\cup\{\Delta\}$ that is bounded and Baire-measurable

$$\mathbb{E}_{\mathbb{P}'_\nu}\left[f(\gamma_t)\middle|\,\mathcal{F}_s^0\right] = \frac{P\left(t-s\right)\left(f\cdot\rho\right)}{\rho}(\gamma_s) \tag{2.3.24}$$

holds true $\mathbb{P}'_\nu$ - almost surely and

$$\mathbb{P}'_\nu \circ \gamma_0^{-1} = \nu.$$

If f is such that $f \cdot \rho \in \mathscr{B}^\rho(E\cup\{\triangle\})$ then Equation 2.3.24 holds true also for the right continuous extension of the filtration.

Proof. We first define a family of sub-probability measures on the space

$$\left((E\cup\{\Delta\})^J, (\mathcal{B}\left(E\cup\{\Delta\}\right))^J\right)_{J\subset\mathbb{R}_+,\text{ finite}}.$$

After showing that this family of probability measures is projective, we can apply the generalized Kolmogorov extension theorem (Theorem A.3.104) and obtain the statement of this theorem. This proof is based on the one of Theorem 2.3.65.

We fix some probability measure ν on $(E\cup\{\Delta\}, \mathcal{B}(E\cup\{\Delta\}))$. In the first step of the proof, we define a family of probability measures

$$\left((p_\nu^J)\right)_{J\subset\mathbb{R}_+,\text{ finite}}$$

on

$$\left((E\cup\{\Delta\})^J, (\mathcal{B}\left(E\cup\{\Delta\}\right))^J\right)_{J\subset\mathbb{R}_+,\text{ finite}}.$$

We fix some $s\in\mathbb{R}_+$ and by Theorem 2.3.37 we find $p(s)(x,\cdot)\in\mathcal{M}^\rho(E)$ such that

$$P(s)f(x) = \int_E f(y)p(s)(x,dy)\,\text{for all } x\in Y.$$

By Proposition 2.3.52 for all $x\in E$

$$P(s)\rho(x) \leq \rho(x),$$

and we define for all $x\in E$ the measures $q(s)(x,\cdot)$

$$q(s)(x,A) := \int_E 1_A(y)\frac{\rho(y)}{\rho(x)}p(s)(x,dy) \text{ for } A\in\mathcal{B}(E).$$

Consequently, $q(s)(x,E)\leq 1$. For any $s\in\mathbb{R}_+$ for any $x\in E$ we define the measures $\tilde{q}(s)(x,\cdot)$ on $E\cup\{\Delta\}$ by

$$\tilde{q}(s)(x,\cdot)|_{\mathcal{B}(E)} := q(s)(x,\cdot)$$

and

$$\tilde{q}(s)(x,\{\Delta\}) := 1 - q(s)(x,E).$$

Furthermore

$$\tilde{q}(s)(\Delta, \{\Delta\}) := 1$$

for any $s \in \mathbb{R}_+$. Thanks to Proposition 2.3.52 (iv), on the space

$$\tilde{\ell}^\infty(E \cup \{\Delta\})$$

of bounded Baire measurable maps we can define the semigroup $(Q(t))_{t\in\mathbb{R}_+}$ by

$$Q(t)f(x) = \int_E f(y)\tilde{q}(t)(x, dy).$$

For any finite $J := \{r_1, ..., r_n\} \subset \mathbb{R}_+$ by Lemma 2.3.67 there is a unique continuous map $j_{J,\nu} : \mathscr{B}^{\rho^{\otimes J}}((E \cup \{\Delta\})^J) \to \mathbb{R}$ such that

$$f_{r_1} \cdot ... \cdot f_{r_n} \to \int_E Q(r_1)\left(\left(\frac{f_{r_1}}{\rho_{r_1}}\right) \cdot ... \cdot \left(Q(r_{n-1} - r_{n-2})\left(\frac{f_{r_{n-1}}}{\rho_{r_{n-1}}}\right) \cdot \left(Q(r_n - r_{n-1})\left(\frac{f_{r_n}}{\rho_{r_n}}\right)\right)\right)\right)(x_0)\nu(dx_0)$$

for any $f \in \mathscr{B}^{\rho^{\otimes J}}\left((E \cup \{\Delta\})^J\right)$ given by

$$f(x_J) := \prod_{j\in J} f_j(x_j).$$

By Theorem 2.3.37 there exists a unique finite positive Radon measure

$$\mu_\nu^J \in \mathcal{M}^{\rho^{\otimes J}}\left((E \cup \{\Delta\})^J\right),$$

such that for any $f \in \mathscr{B}^{\rho^{\otimes J}}\left((E \cup \{\Delta\})^J\right)$.

$$j_{J,\nu}(f) = \int_{E^J} f(x_J)\mu_\nu^J(dx_J),$$

and

$$\int_{E^J} \rho^{\otimes J}(x_J)\mu_\nu^J(dx_J) = 1.$$

We define the family of finite measures

$$\left(q_\nu^J\right)_{J\subset\mathbb{R}_+,\text{ finite}}$$

on

$$\left((E \cup \{\Delta\})^J, \mathcal{B}((E \cup \{\Delta\})^J)\right)_{J\subset\mathbb{R}_+,\text{ finite}}$$

by

$$\begin{aligned} \mathcal{B}(E^J) &\to [0,1] \\ A &\to \int_A \rho^{\otimes J}(x_J)\mu_\nu^J(dx_J). \end{aligned}$$

We first observe the non-obvious fact that for any finite $J \subset \mathbb{R}_+$ the measure p_ν^J is a Radon measure (since the space E^J is non necessarily polish). Fix a finite $\tilde{J} \subset \mathbb{R}_+$. Let $A \in \mathcal{B}\left(E^{\tilde{J}}\right)$ and $\varepsilon > 0$ be arbitrary. Then by $\bigcup_{R>0} K_R = E$ there exists $R_\varepsilon > 0$ such that

$$q_\nu^{\tilde{J}}\left(E^{\tilde{J}} \setminus (K_{R_\varepsilon})^{\tilde{J}}\right) < \frac{\varepsilon}{2}.$$

Since $\mu_\nu^{\tilde{J}}$ is a Radon measure there exists $K \subset A \cap (K_{R_\varepsilon})^{\tilde{J}}$ such that

$$\mu_\nu^{\tilde{J}}\left(A \cap (K_{R_\varepsilon})^{\tilde{J}} \setminus K\right) < \frac{\varepsilon}{2R_\varepsilon}.$$

Thus,

$$\begin{aligned} q_\nu^{\tilde{J}}(A \setminus K) &\leq q_\nu^{\tilde{J}}\left(E^{\tilde{J}} \setminus (K_{R_\varepsilon})^{\tilde{J}}\right) + q_\nu^{\tilde{J}}\left(A \cap (K_{R_\varepsilon})^{\tilde{J}} \setminus K\right) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2}, \end{aligned}$$

and the probability measure $q_\nu^{\tilde{J}}$ is inner regular, hence a Radon measure.

We need to show that the family

$$\left(q_\nu^J\right)_{J \subset \mathbb{R}_+,\ \text{finite}}$$

is projective. To this end, it is sufficient to show for any finite $J := \{r_1, ..., r_n\} \subset \mathbb{R}_+$ and $j \in \{1, ..., n\}$ for any $A_i \in \mathcal{B}(E \cup \{\Delta\})^{r_i}$, $i \in \{1, ..., n\} \setminus \{j\}$

$$q_\nu^J(A_1 \times ... \times A_{j-1} \times E_j \times A_{j+1} ... \times A_n) = q_\nu^{J \setminus \{r_j\}}(A_1 \times ... \times A_{j-1} \times A_{j+1} ... \times A_n).$$

We observe that by Corollary 2.3.10 indicator functions of open sets can be approximated almost surely by continuous bounded maps. Hence, any set in the Borel σ-algebra can be approximated almost surely by continuous bounded maps. With such approximations and dominated convergence projectivity of the family

$$\left(q_\nu^J\right)_{J \subset \mathbb{R}_+,\ \text{finite}}$$

follows from the definition of the family

$$\left(\mu_\nu^J\right)_{J \subset \mathbb{R}_+,\ \text{finite}}$$

on the cylinder functions.

In the last step of the proof, as in the proof of Theorem 2.3.65 on can easily show that

$$\begin{aligned} \mathcal{C} :&= \{C : C \text{ compact}, C \subset K_R \text{ for some } R \geq 0\} \\ &\quad \bigcup \{C \cup \{\Delta\} : C \text{ compact}, C \subset K_R \text{ for some } R \geq 0\} \end{aligned}$$

is a compact class in $E \cup \{\Delta\}$ and that for each $t \in \mathbb{R}_+$ and $A \in \mathcal{B}(E \cup \{\Delta\})$

$$\left(q_\nu^{\{t\}}\right)(A) = \sup\left\{\left(q_\nu^{\{t\}}\right)(C) : C \subset A \text{ and } C \in \mathcal{C}\right\},$$

such that we can apply Theorem A.3.104. This yields a measure $\mathbb{P}'_\nu$ on

$$\left((E \cup \{\Delta\})^{\mathbb{R}_+}, (\mathcal{B}(E \cup \{\Delta\}))^{\mathbb{R}_+}\right).$$

Furthermore,

$$\mathbb{P}'_\nu \circ \gamma_0^{-1} = \nu,$$

by definition of $\mathbb{P}'_\nu$ via the functional $j_{\{0\},\nu}$.

Equation 2.3.24 follows from the fact that we can approximate bounded Borel-measurable functions almost surely by continuous bounded function according to Corollary 2.3.10 and the same reasoning as in the proof of Theorem 2.3.65.

Finally, for f such that $f \cdot \rho \in \mathscr{B}^\rho(E \cup \{\triangle\})$ right continuity of the filtration follows as in the proof of Theorem 2.3.65 for

$$(\mathcal{F}_t)_{t \in \mathbb{R}_+} := \left(\mathcal{F}^0_{t+}\right)_{t \in \mathbb{R}_+}.$$

□

COROLLARY 2.3.74. *Let ρ be Baire measurable and let $(P(t))_{t \in \mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ such that for some $\omega \in \mathbb{R}$ and all $t \in \mathbb{R}_+$*

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))} \leq e^{\omega t}.$$

Then for any probability measure ν on $(E \cup \{\triangle\}, \mathcal{B}(E \cup \{\Delta\}))$ there exists a probability measure $\mathbb{P}'_\nu$ on

$$\left((E \cup \{\triangle\})^{\mathbb{R}_+}, \mathcal{B}(E \cup \{\Delta\})^{\mathbb{R}_+}\right)$$

such that for the canonical process $(\gamma_t)_{t \in \mathbb{R}_+}$ for any $t \geq s \geq 0$ and any real-valued map f on $E \cup \{\Delta\}$ that is bounded and Baire-measurable

$$\mathbb{E}_{\mathbb{P}'_\nu}\left[f(\gamma_t) \middle| \mathcal{F}^0_s\right] = \frac{e^{-\omega t} P(t-s)(f \cdot \rho)}{\rho}(\gamma_s) \tag{2.3.25}$$

holds true $\mathbb{P}'_\nu$ - almost surely (where $(\mathcal{F}^0_t)_{t \in \mathbb{R}_+}$ is the natural filtration) and

$$\mathbb{P}'_\nu \circ \gamma_0^{-1} = \nu.$$

If f is such that $f \cdot \rho \in \mathscr{B}^\rho(E \cup \{\triangle\})$, then Equation 2.3.25 holds true also for the right continuous extension of the filtration.

PROOF. Define the rescaled semigroup (see Lemma 1.4.16) $(S(t))_{t\in\mathbb{R}_+}$ for any $t \in \mathbb{R}_+$ by

$$S(t) := e^{-\omega t} P(t).$$

Then clearly $(S(t))_{t\in\mathbb{R}_+}$ is also a generalized Feller semigroup and satisfies the conditions of Theorem 2.3.73. This directly yields the statement of this corollary. □

REMARK 2.3.75. $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup and let $(p(t)(x,\cdot))_{t\in\mathbb{R}_+,\, x\in E}$ be the family of positive finite Radon measures from Proposition 2.3.52. Comparing

$$q(s)(x,dy) := \frac{\rho(y)}{\rho(x)} p(s)(x,dy)$$

and $p(s)(x,dy)$ we see that since $p(s)(x,\cdot)$ integrates ρ which becomes arbitrarily large outside of compact sets , $p(s)(x,\cdot)$ must have a very small mass in the periphery (the set, where ρ is large). In the center (the set, where ρ is small) however, the mass of $p(s)(x,\cdot)$ may be very large (but finite). On the other hand, by definition $q(s)(x,\cdot)$ has less mass than $p(s)(x,\cdot)$ in the center but more in the periphery. Under the right condition (namely $\|P(s)\|_{L(\mathscr{B}^\rho(E))} = 1$) such rescaling of mass leads to $q(s)(x,\cdot)$ being a sub-probability measure even when the mass of $p(s)(x,\cdot)$ is greater than one (but finite).

If the conditions of both of Theorem 2.3.73 and of Theorem 2.3.65 are satisfied, then in comparison to the generalized Feller process $(\lambda_t)_{t\in\mathbb{R}_+}$ the (sub-)probability measure $q(s)(x,dy)$ drives the extended Feller process $(\gamma_t)_{t\in\mathbb{R}_+}$ with higher probability to areas where ρ is large and reduces the probability for areas where ρ is small to be entered by the process. So relatively speaking, we can say that the generalized Feller process $(\lambda_t)_{t\in\mathbb{R}_+}$ "lives more in the center" whereas the extended Feller process $(\gamma_t)_{t\in\mathbb{R}_+}$ "lives more in the periphery". Thus, if the process $(\gamma_t)_{t\in\mathbb{R}_+}$ starts in the periphery at $x \in E$ and map the f is small in the periphery then over time $\mathbb{E}_x(f(\gamma_t))$ remains small. The precise result can be found in Corollary 2.3.92.

REMARK 2.3.76. By Jensen's inequality (Theorem A.3.73) for a monotone concave (Definition A.3.72) function ρ and a supermartingale $(\lambda_t)_{t\in\mathbb{R}_+}$ the inequality

$$\mathbb{E}_x\left[\rho\left(\lambda_t\right)\right] \leq \rho\left(\mathbb{E}_x\left[(\lambda_t)\right]\right) \leq \rho\left(\mathbb{E}_x\left[(\lambda_0)\right]\right) = \rho(x),$$

holds true, hence the condition

$$\|P(t)\|_{L(\mathcal{B}^{\rho}(Y))} \leq 1$$

holds true for $(P(t))_{t\in\mathbb{R}_+}$ defined by $P(t)\,:\, f \to \mathbb{E}_x\left[f\left(\lambda_t\right)\right]$ for any $t \in \mathbb{R}_+$.

REMARK 2.3.77. Of course in Theorem 2.3.65 and in Theorem 2.3.73 for an interval $I \subset \mathbb{R}_+$ that contains 0 we can also work on the product spaces

$$\left(E^I, \mathcal{B}(E)^I\right),$$

and

$$\left(E \cup \{\Delta\}^I, \left(\mathcal{B}(E \cup \{\Delta\})\right)^I\right)$$

respectively.

REMARK 2.3.78. Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup such that there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a family of stochastic processes $(\lambda_t^x)_{t\in I,\, x\in E}$ on$(\Omega, \mathcal{F}, \mathbb{P})$ where $I = \mathbb{R}_+$ or $I \supset \{0\}$ is an interval, such that $\mathbb{P}\left(\lambda_0^x = x\right) = 1$ and Equation 2.3.13 holds true. Then for $\lambda^x := (\lambda_t^x)_{t\in I,}$ on $\left(E^I, \mathcal{B}(E)^I\right)$, and the distribution $\mathbb{P}_{\lambda^x}$ (as in the sense of Definition A.3.25) $\mathbb{P}_x := \mathbb{P}_{\lambda^x}$ is a probability measure as in Theorem 2.3.65. Therefore, also the family of stochastic processes $(\lambda_t^x)_{t\in I,\, x\in E}$ will be called generalized Feller process. An equivalent statement can be made about a family of stochastic processes $(\gamma_t^x)_{t\in I,\, x\in E}$, Theorem 2.3.73, and extended Feller processes.

Next we want to compare the measures and thus the corresponding canonical processes in Theorem 2.3.65 and in Theorem 2.3.73. It is important to remember, that on the space $\left(E^I, \mathcal{B}(E)^I\right)$ the canonical processes $(\lambda_t)_{t\in I}$ and $(\gamma_t)_{t\in I}$ are the same. However, we choose to denote them differently in order to point out that the probability measures on the spaces $\left(E^I, \mathcal{B}(E)^I\right)$ and

$$\left((E \cup \{\Delta\})^I, \left(\mathcal{B}(E \cup \{\Delta\})\right)^I\right)$$

are different.

PROPOSITION 2.3.79. *Let $T > 0$ and let $I \subset \mathbb{R}_+$ be an interval that contains 0 or let $I = \mathbb{R}_+$ and let ρ be Baire measurable. Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^{\rho}(E)$ such that both* *the conditions of Theorem 2.3.65 and of Theorem 2.3.73 are fulfilled. Let*

$$(p(t)\,(x,\cdot))_{t\in I,;x\in E}$$

be the family of probability measures from Proposition 2.3.52 such that for all $x \in E$, $t \in \mathbb{R}_+$ *and* $f \in \mathscr{B}^\rho(E)$

$$P(t)f(x) = \int_E f(y)p(t)(x,dy).$$

For any initial distribution ν *denote by* $\mathbb{P}_\nu$ *the measure on*

$$\left(E^I, \mathcal{B}(E)^I\right)$$

such that for the canonical process $(\lambda_t)_{t\in I}$ *Equation 2.3.13 holds true for any* $t \geq s,\ s,t \in I$. *Let*

$$(q(t)\,(x,\cdot))_{t\in[0,T];x\in E}$$

be the family of (sub-) probability measures defined as

$$q(s)(x,A) := \int_E 1_A(y)\frac{\rho(y)}{\rho(x)}p(s)(x,dy) \text{ for } A \in \mathcal{B}(E),$$

and for any initial distribution ν *denote by* $\mathbb{P}'_\nu$ *the measure on*

$$\left((E\cup\{\Delta\})^I, (\mathcal{B}(E\cup\{\Delta\}))^I\right)$$

such that the canonical process $(\gamma_t)_{t\in I}$ *fulfills Equation 2.3.24 for any* $t \geq s,\ s,t \in I$.

Then the following assertions hold true:

(i) For all $t \in I$

$$\|P(t)\|_{L(\mathcal{B}^\rho(E))} = 1,$$

and for $x'_0 \in \underset{x\in E}{\arg\min}\,\rho(x)$

$$\mathbb{P}'_{x'_0}\Big|_{\mathcal{B}(E)^I}$$

is a probability measure.

(ii) Let $I = [0,T]$ *and let* $A \in \mathcal{B}(E)^{[0,T]}$. *Then*

$$\mathbb{P}'_\nu[A] = \mathbb{E}_{\mathbb{P}'_\nu}[1_A] = \mathbb{E}_{\mathbb{P}_\nu}\left[1_A\cdot\frac{\rho(\lambda_T)}{\rho(\lambda_0)}\right],$$

and

$$\mathbb{E}_{\mathbb{P}'_\nu}\left[1_A\cdot\frac{\rho(\gamma_0)}{\rho(\gamma_T)}\right] = \mathbb{E}_{\mathbb{P}_\nu}[1_A] = \mathbb{P}_\nu[A]$$

hold true, hence $\mathbb{P}'_\nu|_{\mathcal{B}(E)^{[0,T]}}$ *and* $\mathbb{P}_\nu$ *are equivalent measures (see Definition A.3.32).*

PROOF. (i) By Corollary 2.3.17

$$\arg\min_{x\in E}\rho(x)$$

is non-empty. Let $C := \rho(x_0')$ and let $t \in I$ be arbitrary. Then by positivity of generalized Feller semigroups

$$(P(t)\rho) \geq (P(t)\,(C\cdot 1))$$

and by assumption of Theorem 2.3.65

$$(P(t)\,(C\cdot 1)) = (C\cdot 1)\,.$$

Hence

$$(P(t)\rho)\,(x_0') \geq C = \rho(x_0')$$

and by Proposition 2.3.52

$$(P(t)\rho)\,(x_0') \leq \rho(x_0'),$$

which proves

$$(P(t)\rho)\,(x_0') = \rho(x_0').$$

Since $t \in I$ was arbitrary by Proposition 2.3.52

$$\|P(t)\|_{L(\mathcal{B}^\rho(E))} = 1,$$

holds true for any $t \in I$. The equation

$$\mathbb{E}'_{x_0'}(1_E(\gamma_t)) = \frac{(P(t)\rho)\,(x_0')}{\rho(x_0')} = 1$$

for any $t \in I$ yields by definition of the probabilities $\mathbb{P}'_{x_0'}$ in Theorem 2.3.65 for any $t \in I$

$$\mathbb{P}'_{x_0'}(E^{[0,t]}) = 1.$$

(ii) Just like $\mathbb{P}'_\nu|_{\mathcal{B}(E)^{[0,T]}}$ the map $\mathbb{Q}_\nu$:

$$A \to \mathbb{E}_{\mathbb{P}_\nu}\left[1_A \cdot \frac{\rho(\lambda_T)}{\rho(\lambda_0)}\right]$$

is a measure on

$$\left(E^{[0,T]}, \mathcal{B}(E)^{[0,T]}\right).$$

Its mass is given by

$$\begin{aligned}
\mathbb{E}_\nu\left[1_{E^{[0,T]}}\frac{\rho(\lambda_T)}{\rho(\lambda_0)}\right] &= \int_E\left(\int_E \rho(x_T)p(T)(x_0,dx_T)\right)\frac{1}{\rho(x_0)}d\nu(x_0)\\
&= \mathbb{E}'_\nu(1_E(\gamma_T))\\
&= \mathbb{P}'_\nu(E^{[0,T]}).
\end{aligned}$$

By Proposition A.3.31, it is enough to show that $\mathbb{Q}_\nu$ and $\mathbb{P}'_\nu|_{\mathcal{B}(E)^{[0,T]}}$ coincide on an intersection stable generator of $\mathcal{B}(E)^{[0,T]}$. This is indeed the case as for any $x_0 \in E$, $n \in \mathbb{N}$,$\{t_1, ..., t_n\} \subset [0,T]$, and $A_{t_1}, ..., A_{t_n} \in \mathcal{B}(E)$ one can approximate the indicator functions $1_{A_{t_1}}$, ...,$1_{A_{t_n}}$ and ρ using an adaption of Corollary 2.3.10 by non-negative continuous bounded functions that converge almost surely with respect to $p_{x_0}^{\{0,t_1,...,t_n,T\}}$, as defined in the proof of Theorem 2.3.65 and $q_{x_0}^{\{0,t_1,...,t_n,T\}}$, as defined in the proof of Theorem 2.3.73. Then one obtains by dominated convergence, and the definition of the measures$\mathbb{P}'_\nu$ and $\mathbb{P}_\nu$

$$
\begin{aligned}
&\mathbb{E}'_\nu \left[1_{E^{[0,T]}} \cdot 1_{A_{t_1}}(\gamma_{t_1}) \cdot ... \cdot 1_{A_{t_n}}(\gamma_{t_n})\right] \\
&= \mathbb{E}'_\nu \left[1_E(\gamma_T) \cdot 1_{A_{t_1}}(\gamma_{t_1}) \cdot ... \cdot 1_{A_{t_n}}(\gamma_{t_n})\right] \\
&= \left(\int_E \left(\int_{A_{t_1}} ... \left(\int_{A_{t_n}} \left(\int_E \rho(x_T) p(T-t_n)(x_{t_n}, dx_T)\right) p(t_n - t_{n-1})(x_{t_{n-1}}, dx_{t_n})\right) ...\right) \frac{1}{\rho(x_0)} d\nu(x_0)\right) \\
&= \mathbb{E}_{\mathbb{P}_\nu} \left[1_E(\gamma_T) \cdot 1_{A_{t_1}}(\gamma_{t_1}) \cdot ... \cdot 1_{A_{t_n}}(\gamma_{t_n}) \frac{\rho \circ \lambda_T}{\rho \circ \lambda_0}\right].
\end{aligned}
$$

□

PROPOSITION 2.3.80. *Let $E = \mathbb{R}^d$ for some $d \in \mathbb{N}$, let $I \subset \mathbb{R}_+$ be an interval containing 0 or all of $\mathbb{R}_+$ and let $(\lambda_t^x)_{t\in I}$ be an Ito diffusion (see Remark 2.1.27) with state space E with drift μ and diffusion matrix σ (see Remark 2.1.27), i.e., let $(\lambda_t^x)_{t\in I, x\in E}$ with $\lambda_0^x = x$ $\mathbb{P}$- a.s. satisfy the stochastic differential equation*

$$d\lambda_t^x = \mu(\lambda_t^x)dt + \sigma(\lambda_t^x)dW_t$$

with $(W_t)_{t\in I}$ the d-dimensional Brownian motion (see Definition A.3.83) on the filtered probability space

$$\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in I}, \mathbb{P}\right).$$

Let $(P(t))_{t\in I}$ be a semigroup of linear bounded operators on $\mathscr{B}^\rho(E)$ defined by

$$P(t)f(x) := \mathbb{E}\left[f(\lambda_t^x)\right]$$

for $f \in \mathscr{B}^\rho(E)$ and let

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))} \le 1.$$

Let$\rho \in C^2(E)$, and let $\mathbb{P}'$ be another probability measure on

$$\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in I}\right)$$

such that for the family of stochastic processes $(\gamma_t^x)_{t\in I, x\in E}$ with $\gamma_0^x = x$ $\mathbb{P}'$-a.s. and any real-valued map f on $E \cup \{\Delta\}$ that is bounded and

Baire-measurable

$$\mathbb{E}'\left[f(\gamma_t^x)\right] = \frac{P\left(t\right)\left(f\cdot\rho\right)}{\rho}(x)$$

holds true.
Then the drift $\mu' = (\mu'_1, ..., \mu'_d)$ *of* $(\gamma_t)_{t\in I}$ *with respect to* $\mathbb{P}'$ *is given by*

$$\mu'_i = \mu_i + \sum_{j=1}^{d} \frac{d\rho}{dx_j}(x)\frac{\sigma_{ij}^2(x)}{\rho(x)},$$

the diffusion matrix is $\sigma' = \sigma$, *and the killing rate* $c' < 0$ *is*

$$c'(x) = \left(\sum_{i=1}^{d} \frac{d\rho}{dx_i}(x)\mu_i(x) + \frac{1}{2}\sum_{j=1}^{d}\sum_{i=1}^{d} \frac{d^2\rho}{dx_i dx_j}(x)\sigma_{ij}^2(x)\right)\frac{1}{\rho(x)}.$$

(see Definition 2.1.23).

Proof. By Ito formula (see Theorem A.3.115) and the assumption $\|P(t)\|_{L(\mathscr{B}^\rho(E))} \le 1$ for any $x \in E$ and $t \in I$

$$\begin{aligned}\mathbb{E}\left[\rho(\lambda_t^x)\right] &= \rho(x) + \int_0^t \left(\sum_{i=1}^{d} \frac{d\rho}{dx_i}(x)\mu_i(x) + \frac{1}{2}\sum_{j=1}^{d}\sum_{i=1}^{d} \frac{d^2\rho}{dx_i dx_j}(x)\sigma_{ij}^2(x)\right) ds \\ &\le \rho(x).\end{aligned}$$

Hence,

$$\sum_{i=1}^{d} \frac{d\rho}{dx_i}(x)\mu_i(x) + \frac{1}{2}\sum_{j=1}^{d}\sum_{i=1}^{d} \frac{d^2\rho}{dx_i dx_j}(x)\sigma_{ij}^2(x) \le 0.$$

Furthermore, for any $x \in E$ and $f \in C_c^2(E)$ the infinitesimal generator $\mathscr{A}'$ of $(\gamma_t)_{t\in I}$ is given by

$$\begin{aligned}\mathscr{A}'f(x) &= \lim_{t\searrow 0}\frac{\mathbb{E}'\left[f(\gamma_t^x)\right] - f(x)}{t} \\ &= \lim_{t\searrow 0}\frac{1}{t}\left(\frac{P\left(t\right)\left(f\cdot\rho\right)}{\rho}(x) - f(x)\right) \\ &= \lim_{t\searrow 0}\frac{1}{t}\left(\frac{\mathbb{E}\left[(f\cdot\rho)(\lambda_t^x)\right]}{\rho(x)} - f(x)\right) \\ &= \lim_{t\searrow 0}\frac{1}{t}\left(\frac{(f\cdot\rho)(x) + \int_0^t \sum_{i=1}^{d}\left(\frac{d(f\cdot\rho)}{dx_i}\mu_i(\lambda_s)\right)ds + \frac{1}{2}\int_0^t \sum_{j=1}^{d}\sum_{i=1}^{d}\left(\frac{d^2(f\cdot\rho)}{dx_i dx_j}\sigma_{ij}^2(\lambda_s)\right)ds}{\rho(x)} - f(x)\right) \\ &= \sum_{i=1}^{d}\frac{d(f\cdot\rho)}{dx_i}(x)\frac{\mu_i(x)}{\rho(x)} + \frac{1}{2}\sum_{j=1}^{d}\sum_{i=1}^{d}\frac{d^2(f\cdot\rho)}{dx_i dx_j}(x)\frac{\sigma_{ij}^2(x)}{\rho(x)}.\end{aligned}$$

Applying the product rule yields the assertion of the Proposition. $\square$

REMARK 2.3.81. If in Proposition 2.3.80

$$\sum_{i=1}^{d} \frac{d\rho}{dx_i}(x)\mu_i(x) + \frac{1}{2}\sum_{j=1}^{d}\sum_{i=1}^{d} \frac{d^2\rho}{dx_i dx_j}(x)\sigma_{ij}^2(x) = 0$$

holds true, the change of measure from $\mathbb{P}$ to $\mathbb{P}'$ only produces a drift as extra term. Let $0 \leq T < \infty$. Let $\tilde{\mathbb{P}}_T$ be the probability measure in Girsanov (Theorem A.3.118) and let $\tilde{W}$ be the d-dimensional Brownian motion with respect to $\tilde{\mathbb{P}}_T$ defined by $W_t = \tilde{W}_t + \int_0^t a(\lambda_s)ds$ for $a = \left(a^1, ..., a^d\right) : \mathbb{R}^d \to \mathbb{R}^d$. Then by

$$\begin{aligned} d\lambda_t &= \mu(\lambda_t)dt + \sigma(\lambda_t)d\tilde{W}_t + \sigma(\lambda_t)a(\lambda_t)dt \\ &= \left(\mu(\lambda_t) + \sigma(\lambda_t)a(\lambda_t)\right)dt + \sigma(\lambda_t)d\tilde{W}_t, \end{aligned}$$

the drift of $(\lambda_t)_{t\in[0,T]}$ with respect to $\tilde{\mathbb{P}}_T$ is $\mu(\lambda_t) + \sigma(\lambda_t)a(\lambda_t)$. Thus, if a is such that for any $x \in E$ and any $1 \leq i \leq d$

$$\sum_{j=1}^{d} \frac{d\rho}{dx_j}(x)\frac{\sigma_{ij}^2(x)}{\rho(x)} = \sum_{j=1}^{d} \sigma_{ij}(x)a^j(x)$$

holds true, then on $[0, T]$ the probability measure $\tilde{\mathbb{P}}_T$ appearing in the theorem of Girsanov is the probability measure $\mathbb{P}'$ from Proposition 2.3.80.

EXAMPLE 2.3.82. Let (E, ρ) be some weighted space and let $(\psi_t)_{t\in\mathbb{R}_+}$ be a family of maps such that $(P(t))_{t\in\mathbb{R}_+}$ defined as

$$P(t)(f) := f \circ \psi_t$$

is a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ of transport type as defined in Proposition 2.3.54. We first determine a generalized Feller process and later choose a specific $(\psi_t)_{t\in\mathbb{R}_+}$ and admissible weight function ρ such that the conditions of Theorem 2.3.73 are satisfied, which permits us to construct a extended Feller process as in the proof of Theorem 2.3.73.

Regarding the generalized Feller process, for any $x \in E$ define

$$\beta(x) := (\psi_t(x))_{t\in\mathbb{R}_+} \in E^{\mathbb{R}_+}$$

and let

$$\mathbb{P}_x := \delta_{\beta(x)} \in \mathcal{M}_1\left(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\mathbb{R}_+}\right).$$

Then for the canonical process $(\lambda_t)_{t\in\mathbb{R}_+}$ and any $t \in \mathbb{R}_+$ and any $x \in E$

$$\mathbb{P}_x\left(\lambda_t = \psi_t(x)\right) = 1,$$

hence for the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ where for any $t \in \mathbb{R}_+$ $\mathcal{F}_t$ is simply the σ-algebra generated by the $\mathcal{B}(E)^{\mathbb{R}_+}$-null sets, for $t > s \geq 0$ and

$f \in \mathscr{B}^\rho(E)$ the stochastic process $f(\lambda_t)$ is measurable with respect to $\mathcal{F}_s$ and for any $x \in E$ $\mathbb{P}_x$-almost surely

$$\begin{aligned}\mathbb{E}_{\mathbb{P}_x}\left[f(\lambda_t)\right|\mathcal{F}_s] &= f(\lambda_t)\\ &= f \circ \psi_t(x)\\ &= (f \circ \psi_{t-s})\,(\psi_s(x))\\ &= P\,(t-s)\,f(\psi_s(x))\\ &= P\,(t-s)\,f(\lambda_s).\end{aligned}$$

Furthermore, by definition

$$\mathbb{P}_x \circ \lambda_0^{-1} = \delta_x,$$

thus $(\lambda_t)_{t\in\mathbb{R}_+}$ is a generalized Feller process.

Next, let in particular $E = \mathbb{R}$ and let $\rho(x) := x^2 + 1$ be the admissible weight function on E and let

$$\begin{aligned}\psi_t :\ E &\to E\\ x &\to e^{-t}x.\end{aligned}$$

Then $\|P(t)\|_{L(\mathscr{B}^\rho(E))} \leq 1$ for any $t \in \mathbb{R}_+$ since for $|f| \leq \rho$, $f \in \mathscr{B}^\rho(E)$ and any $t \in \mathbb{R}_+$

$$\begin{aligned}\frac{|P(t)f(x)|}{\rho(x)} &= \frac{|f\,(\psi_t(x))|}{\rho(x)}\\ &\leq \frac{\rho\,(\psi_t(x))}{\rho(x)}\\ &= \frac{\rho\,(e^{-t}x)}{\rho(x)}\\ &= \frac{(e^{-t}x)^2 + 1}{x^2+1}\\ &\leq 1.\end{aligned}$$

Thus, the conditions of Theorem 2.3.73 are satisfied. By definition of $\mathbb{P}_x$ for any $x \in E$ the semigroup of transition probabilities $(q(t))_{t\in\mathbb{R}_+}$ of the generalized Feller process $(\lambda_t)_{t\in\mathbb{R}_+}$ is given by $p(t)(x, dy) = \delta_{e^{-t}x}(dy)$ for any $t \in \mathbb{R}_+$, $x \in E$. Thus, by construction in Theorem 2.3.73, the semigroup of transition probabilities $(q(t))_{t\in\mathbb{R}_+}$ of the process $(\gamma_t)_{t\in\mathbb{R}_+}$ from Theorem 2.3.73 is given by

$$\begin{aligned}q(t)(x, dy) &= \frac{\rho(e^{-t}x)}{\rho(x)}\delta_{e^{-t}x}(dy)\\ &= \frac{e^{-2t}x^2+1}{x^2+1}\delta_{e^{-t}x}(dy)\end{aligned}$$

for any $t \in \mathbb{R}_+$, $x \in E$.

EXAMPLE 2.3.83. (Geometric Brownian motion on $E = \{x \in \mathbb{R} : x > 0\}$ with $\rho(x) = \sqrt{x}$)

Continuing Example 2.1.26, for any $x \in E$ and $\frac{1}{2}\mu - \frac{1}{8}\sigma^2 \leq 0$ let

$$S_t^x = x \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t\right),$$

where $(W_t)_{t\in\mathbb{R}_+}$ is the Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$. We show that $(S_t^x)_{t\in\mathbb{R}_+}$ is a generalized Feller process, that the conditions of Theorem 2.3.73 are fulfilled and construct the corresponding extended Feller process $\left((S_t^x)'\right)_{t\in\mathbb{R}_+}$ explicitly.

As seen in Example 2.1.26, for any $t \in \mathbb{R}_+$ $\mathbb{P}_{S_t^x}$ is given by $\mathbb{P}_{S_t^x}(A) = \int_A \kappa(t)(x,y)dy$ for any $A \in \mathcal{B}(E)$ and,

$$\kappa(t)(x,y) = \frac{1}{\sqrt{2\pi}}\frac{1}{y\sigma\sqrt{t}} \exp\left(\frac{-\left(\ln y - \ln x - \left(\mu - \frac{1}{2}\sigma^2\right)t\right)^2}{2\sigma^2 t}\right).$$

It was also shown that $(S_t^x)_{t\in\mathbb{R}_+}$ is a Markov process, hence for any real-valued measurable map f on E that is bounded or positive and any $t \in \mathbb{R}_+$

$$P(t)f(x) := \mathbb{E}\left[f(S_t^x)\right],$$

given by

$$P(t)f(x) = \int_E f(y)\frac{1}{\sqrt{2\pi}}\frac{1}{y\sigma\sqrt{t}} \exp\left(\frac{-\left(\ln y - \ln x - \left(\mu - \frac{1}{2}\sigma^2\right)t\right)^2}{2\sigma^2 t}\right) dy$$

is a Markov semigroup. We need to show that for any $t \in \mathbb{R}_+$ $P(t)$ is a linear map form $\mathscr{B}^\rho(E)$ to $\mathscr{B}^\rho(E)$ that is bounded. We calculate by Ito formula (Theorem A.3.115)

$$\begin{aligned}\mathbb{E}\left[(S_t^x)^{1/2}\right] &= x^{1/2} + \mathbb{E}\left[\int_0^t \frac{1}{2}(S_s^x)^{-1/2}\left(S_s^x\mu\right)ds - \int_0^t \frac{1}{8}(S_s^x)^{-3/2}\left(S_s^x\sigma\right)^2 ds\right] \\ &= x^{1/2} + \left(\frac{1}{2}\mu - \frac{1}{8}\sigma^2\right)\int_0^t \mathbb{E}\left[(S_s^x)^{1/2}\right]ds,\end{aligned}$$

and we see from

$$\frac{d}{dt}\left(\mathbb{E}\left[(S_t^x)^{1/2}\right]\right) = \left(\frac{1}{2}\mu - \frac{1}{8}\sigma^2\right)\mathbb{E}\left[(S_t^x)^{1/2}\right]$$

that

$$P(t)\rho(x) = \mathbb{E}\left[(S_t^x)^{1/2}\right] = \rho(x)\exp\left(\left(\frac{1}{2}\mu - \frac{1}{8}\sigma^2\right)t\right).$$

Hence, by positivity of $P(t)$ on its domain and assumption $\frac{1}{2}\mu-\frac{1}{8}\sigma^2 \leq 0$ we obtain $\|P(t)\|_{L(\mathcal{B}^\rho(E))} \leq 1$ and

$$P(t)f \in B^\rho(E) = \left\{ f : E \to E : \sup_{x\in E} \rho(x)^{-1} \|f(x)\| < \infty \right\}.$$

In order to show $P(t)f \in \mathscr{B}^\rho(E)$ for any $f \in \mathscr{B}^\rho(E)$, by density of $C_b(E)$ in $\mathscr{B}^\rho(E)$ and continuity of $P(t)$ it is sufficient to show that $f \in C_b(E)$ implies $P(t)f \in \mathscr{B}^\rho(E)$. We show below that in this case even $P(t)f \in C_b(E)$ holds true.

Let $(x_n)_{n\in\mathbb{N}} \subset E$ be a sequence such that $\lim_{n\to\infty} x_n = x$. Let $a := \inf_{n\in\mathbb{N}} x_n$ and $b := \sup_{n\in\mathbb{N}} x_n$. Then for any $n \in \mathbb{N}$

$$\exp\left(\frac{-\left(\ln y - \ln x_n - \left(\alpha - \frac{1}{2}\sigma^2\right) t\right)^2}{2\sigma^2 t} \right) \leq \exp\left(\frac{-\left(\ln y - \ln a - \left(\alpha - \frac{1}{2}\sigma^2\right) t\right)^2}{2\sigma^2 t} \right)$$

if $y \leq ae^{\left(\alpha-\frac{1}{2}\sigma^2\right)t}$. Furthermore,

$$\exp\left(\frac{-\left(\ln y - \ln x_n - \left(\alpha - \frac{1}{2}\sigma^2\right) t\right)^2}{2\sigma^2 t} \right) \leq \exp\left(\frac{-\left(\ln y - \ln b - \left(\alpha - \frac{1}{2}\sigma^2\right) t\right)^2}{2\sigma^2 t} \right)$$

if $y \geq be^{\left(\alpha-\frac{1}{2}\sigma^2\right)t}$. Hence, for all $y > 0$

$$\begin{aligned} \frac{1}{y\sigma\sqrt{t}} \exp\left(\frac{-\left(\ln y - \ln x_n - \left(\alpha - \frac{1}{2}\sigma^2\right) t\right)^2}{2\sigma^2 t} \right) &\leq \frac{1}{y\sigma\sqrt{t}} \exp\left(\frac{-\left(\ln y - \ln a - \left(\alpha - \frac{1}{2}\sigma^2\right) t\right)^2}{2\sigma^2 t} \right). \\ &+ \frac{1}{y\sigma\sqrt{t}} \exp\left(\frac{-\left(\ln y - \ln b - \left(\alpha - \frac{1}{2}\sigma^2\right) t\right)^2}{2\sigma^2 t} \right) \\ &+ \frac{1}{y\sigma\sqrt{t}} 1_{\left\{ae^{\left(\alpha-\frac{1}{2}\sigma^2\right)t} \leq y \leq be^{\left(\alpha-\frac{1}{2}\sigma^2\right)t}\right\}}(y). \end{aligned}$$

The right hand side is integrable, thus by dominated convergence

$$\begin{aligned} \lim_{n\to\infty} P(t)f(x_n) &= \int_E \lim_{n\to\infty} f(y) \frac{1}{\sqrt{2\pi}} \frac{1}{y\sigma\sqrt{t}} \exp\left(\frac{-\left(\ln y - \ln x_n - \left(\alpha - \frac{1}{2}\sigma^2\right) t\right)^2}{2\sigma^2 t} \right) dy \\ &= \int_E f(y) \frac{1}{\sqrt{2\pi}} \frac{1}{y\sigma\sqrt{t}} \exp\left(\frac{-\left(\ln y - \ln x - \left(\alpha - \frac{1}{2}\sigma^2\right) t\right)^2}{2\sigma^2 t} \right) dy \\ &= P(t)f(x). \end{aligned}$$

Since boundedness of $P(t)f$ is clear by definition of $P(t)$, we obtained that for any $t \in \mathbb{R}_+$ $P(t)$ maps $C_b(E)$ to $C_b(E)$.

Hence $(P(t))_{t\in\mathbb{R}_+}$ is a family of bounded linear operators on $\mathscr{B}^\rho(E)$ that fulfills properties **P1**, **P2** and **P5** of generalized Feller semigroups (see Definition 2.3.49) by virtue of being a Markov semigroup. Furthermore **P4** holds true thanks to $\|P(t)\|_{L(\mathscr{B}^\rho(E))} \leq 1$ and regarding

P3 for all continuous bounded maps f by dominated convergence and continuity of the Brownian motion

$$\lim_{t\to 0} P(t)f(x) = \mathbb{E}_x\left[\lim_{t\to 0} f(S_t^x)\right] = f(x).$$

By density of $C_b(E)$ in $\mathscr{B}^\rho(E)$ this convergence extends to any $f \in \mathscr{B}^\rho(E)$ (just like in the proof of Proposition 2.3.89). Therefore $(P(t))_{t\in\mathbb{R}_+}$ is a contractive generalized Feller semigroup.

Since clearly for any $t \in \mathbb{R}_+$

$$(P(t)1)(x) = 1$$

$(P(t))_{t\in\mathbb{R}_+}$ fulfills the conditions of Theorem 2.3.73 and of Theorem 2.3.65. Hence, we can define $(Q(t))_{t\in\mathbb{R}_+}$ by

$$Q(t)(f) := \frac{P(t)(f\cdot\rho)}{\rho}$$

and by Theorem 2.3.73 for any $x \in E$ there exists a probability measure $\mathbb{P}'_x$ and a Markov process $\left((S_t^x)'\right)_{t\in\mathbb{R}_+}$ such that $(S_0^x)' = x$.

We next find the process $\left((S_t^x)'\right)_{t\in\mathbb{R}_+}$ explicitly. Thanks to Proposition 2.3.80 we already know what drift and killing rate are. Let $\left(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}\right)$ be a different probability space and let $\tilde{\tau} : \tilde{\Omega} \to \mathbb{R}_+$ be measurable and distributed as $\mathbb{P}(\tilde{\tau} \le t) = 1 - \exp\left(\left(\frac{1}{2}\mu - \frac{1}{8}\sigma^2\right)t\right)$. Define $\Omega' := \Omega \times \tilde{\Omega}$, $\mathcal{F}' := \mathcal{F} \otimes \tilde{\mathcal{F}}$, and $\mathbb{P}' := \mathbb{P} \otimes \tilde{\mathbb{P}}$. Then on the product space $(\Omega', \mathcal{F}', \mathbb{P}')$ one can define the Brownian motion $(W'_t)_{t\in\mathbb{R}_+}$ as

$$W'_t(\omega, \tilde{\omega}) := W_t(\omega)$$

and $\tau' : \tilde{\Omega} \to \mathbb{R}_+$ as

$$\tau'(\omega, \tilde{\omega}) := \tilde{\tau}(\tilde{\omega}).$$

τ' and W'_t are independent for any $t \in \mathbb{R}_+$ and with

$$(S_t^x)' = \begin{cases} x\exp\left(\mu t + \sigma W'_t\right) & \text{for } t < \tau' \\ \Delta & \text{for } t \ge \tau' \end{cases}$$

we obtain that also $\left((S_t^x)'\right)_{t\in\mathbb{R}_+}$ is a Markov process and that

$$\frac{1}{\rho}\mathbb{E}\left[(f\cdot\rho)(S_t^x)\right] = \mathbb{E}'\left[f((S_t^x)')\right]$$

as can be seen by the following calculation:

$$
\begin{aligned}
&\frac{1}{\rho}\mathbb{E}\left[(f\cdot\rho)\left(S_t^x\right)\right]\\
&=\int_{\mathbb{R}_+} f(y)\frac{\sqrt{y}}{\sqrt{x}}\frac{1}{\sqrt{2\pi t}}\frac{1}{y\sigma}\exp\left(\frac{-\left(\ln y-\ln x-\left(\mu-\frac{1}{2}\sigma^2\right)t\right)^2}{2\sigma^2 t}\right)dy\\
&=\int_{\mathbb{R}_+} f(y)\frac{1}{\sqrt{2\pi t}}\frac{1}{y\sigma}\exp\left(\frac{-\left(\ln y-\ln x-\left(\mu-\frac{1}{2}\sigma^2\right)t\right)^2+2\sigma^2 t\ln\sqrt{y}-2\sigma^2 t\ln\sqrt{x}}{2\sigma^2 t}\right)dy\\
&=\int_{\mathbb{R}_+} f(y)\frac{1}{\sqrt{2\pi t}}\frac{1}{y\sigma}\exp\left(\frac{-\left((\ln y-\ln x-\mu t)+\frac{1}{2}\sigma^2 t\right)^2+\sigma^2 t\ln y-\sigma^2 t\ln x}{2\sigma^2 t}\right)dy\\
&=\int_{\mathbb{R}_+} f(y)\frac{1}{\sqrt{2\pi t}}\frac{1}{y\sigma}\exp\left(\frac{-(\ln y-\ln x-\mu t)^2+\sigma^2 t^2\mu-\left(\frac{1}{2}\sigma^2 t\right)^2}{2\sigma^2 t}\right)dy\\
&=\int_{\mathbb{R}_+} f(y)\frac{1}{\sqrt{2\pi t}}\frac{1}{y\sigma}\exp\left(\frac{-(\ln y-\ln x-\mu t)^2}{2\sigma^2 t}\right)\exp\left(\frac{1}{2}\mu t-\frac{1}{8}\sigma^2 t\right)dy\\
&=\mathbb{E}'\left[f\left(\left(S_t^x\right)'\right)\right].
\end{aligned}
$$

If instead of $\left(S_t^x\right)' = \left(\left(S_t^x\right)'\right)_{t\in\mathbb{R}_+}$ one wants to obtain a stochastic process defined on

$$\left(\left(E\cup\{\triangle\}\right)^{\mathbb{R}_+}, \mathcal{B}(E\cup\{\Delta\})^{\mathbb{R}_+}\right)$$

as in Theorem 2.3.73 one can choose the distribution $\mathbb{P}_{(S^x)'}$ and consider the canonical process.

2.3.6. Relationship between extended Feller processes and generalized Feller processes and Feller processes. Next we would like to investigate the relationship between on the one hand the extended Feller process $(\gamma_t)_{t\in\mathbb{R}_+}$ from Theorem 2.3.73 and Feller processes and on the other hand between generalized Feller processes $(\lambda_t)_{t\in\mathbb{R}_+}$ from Theorem 2.3.65 and Feller processes. Also in this subsection, $(E,\,\rho)$ will always denote a weighted space and as σ-algebra on this space we always take the Borel σ-algebra. We start with a first result:

PROPOSITION 2.3.84. *Let (E,ρ) be a weighted space and E be locally compact. Let $(\lambda_t)_{t\in\mathbb{R}_+}$ be a Feller process on E with semigroup of transition probabilities $(p(t))_{t\in\mathbb{R}_+}$ on $(E,\mathcal{B}(E))$ with initial distribution $\nu\in\mathcal{M}^\rho(E)$. Let there be $t_0>0$ and $C>0$ such that for all $x\in E$ and $0\le t\le t_0$*

$$\mathbb{E}_x\left[\rho(\lambda_t)\right]\le C\rho(x).$$

Then $(\lambda_t)_{t\in\mathbb{R}_+}$ is a generalized Feller process with respect to a right continuous filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ and with initial distribution ν.

PROOF. We need to show that $\left(\tilde{P}(t)\right)_{t\in\mathbb{R}_+}$ given by

$$\begin{aligned}\tilde{P}(t):\ \mathscr{B}^\rho(E) &\to \mathscr{B}^\rho(E)\\ f &\to \int_E f(y)p(t)(\cdot,dy)\end{aligned}$$

is a generalized Feller semigroup and that there is a right continuous filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ such that $(\lambda_t)_{t\in\mathbb{R}_+}$ is adapted with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ and for initial distribution $\nu \in \mathcal{M}^\rho(E)$ and any $f \in \mathscr{B}^\rho(E)$ and $0 \leq s \leq t$

$$\mathbb{E}_{\mathbb{P}_\nu}\left[f(\lambda_t)\middle| \mathcal{F}_s\right] = \tilde{P}\left(t-s\right) f(\lambda_s)$$

holds true $\mathbb{P}_\nu$ -almost surely. We know that $(P(t))_{t\in\mathbb{R}_+}$ given by

$$\begin{aligned}P(t):\ C_0(E) &\to C_0(E)\\ f &\to \int_E f(y)p(t)(\cdot,dy)\end{aligned}$$

is a Feller semigroup.

First, we show that $\tilde{P}(t)\left(\mathscr{B}^\rho(E)\right) = \mathscr{B}^\rho(E)$ and that

$$\tilde{P}(t):\ \mathscr{B}^\rho(E) \to \mathscr{B}^\rho(E)$$

is a linear bounded map. For any $f \in \mathscr{B}^\rho(E)$ and $0 \leq t \leq t_0$

$$\begin{aligned}\tilde{P}(t)f(x) &= \int_E f(y)p(t)(x,dy)\\ &= \int_E \frac{f(y)}{\rho(y)}\rho(y)p(t)(x,dy)\\ &\leq \|f\|_\rho C\rho(x).\end{aligned}$$

Hence, for $0 \leq t \leq t_0$

$$\tilde{P}(t):\mathscr{B}^\rho(E) \to B^\rho(E)$$

is a linear bounded map with

$$\left\|\tilde{P}(t)\right\|_{L(\mathcal{B}^\rho(E))} \leq C.$$

By Lemma 2.3.48 for any $\varepsilon > 0$ and any $f \in \mathcal{B}^\rho(E)$ there is $g_\varepsilon \in C_0(E)$ such that $\|f - g_\varepsilon\|_\rho < \varepsilon$. Hence for $0 \leq t \leq t_0$

$$\left\|\tilde{P}(t)f - \tilde{P}(t)g_\varepsilon\right\|_\rho < C\varepsilon$$

and since $\tilde{P}(t)g_\varepsilon = P(t)g_\varepsilon \in C_0(E)$ it follows that that $\tilde{P}(t)(\mathscr{B}^\rho(E)) = \mathscr{B}^\rho(E)$. For any $0 < s$ there is $n \in \mathbb{N}$ such that $s/n < t_0$ and since

$(p(t))_{t\in\mathbb{R}_+}$ is a semigroup of transition probabilities on $(E,\mathcal{B}(E))$

$$\begin{aligned}\tilde{P}(s)f(x) &= \int_E f(y)p(s)(x,dy)\\ &= \int_E f(y)p(\frac{s}{n}+...+\frac{s}{n})(x,dy)\\ &= \left(\tilde{P}(\frac{s}{n})...\left(\tilde{P}(\frac{s}{n})f\right)\right)(x).\end{aligned}$$

Hence,

$$\tilde{P}(t):\mathscr{B}^\rho(E)\to\mathscr{B}^\rho(E)$$

is a linear bounded map for any $t>0$ and

$$\left\|\tilde{P}(t)\right\|_{L(\mathcal{B}^\rho(E))}\leq C^{\lceil t/t_0\rceil}.$$

In order to show that

$$\tilde{P}(t):\mathscr{B}^\rho(E)\to\mathscr{B}^\rho(E)$$

is indeed a generalized Feller semigroup we have to show the properties **P1**,...,**P5** from Definition 2.3.49 hold. **P1** and **P2** follow immediately from the fact $(p(t))_{t\in\mathbb{R}_+}$ is a semigroup of transition probabilities. **P4** follows by assumption and positivity (**P5**) is obvious. It remains to be shown that for all $f\in\mathscr{B}^\rho(E)$ and all $x\in E$

$$\lim_{t\searrow 0}\tilde{P}(t)f(x)=f(x).$$

Fix $f\in\mathscr{B}^\rho(E)$ and $x\in E$. By Lemma 2.3.48 for any $\varepsilon>0$ there is $g_\varepsilon\in C_0(E)$ such that $\|f-g_\varepsilon\|_\rho<\varepsilon$. By strong continuity of Feller semigroups

$$\lim_{t\searrow 0}\left\|\tilde{P}(t)g_\varepsilon-g_\varepsilon\right\|_\infty=0.$$

Thus,

$$\begin{aligned}\lim_{t\searrow 0}\left|\tilde{P}(t)f(x)-f(x)\right| &= \lim_{t\searrow 0}\left|\tilde{P}(t)f(x)-\tilde{P}(t)g_\varepsilon(x)\right|\\ &\quad+\left|\tilde{P}(t)g_\varepsilon(x)-g_\varepsilon(x)\right|\\ &\quad+|g_\varepsilon(x)-f(x)|\\ &\leq \lim_{t\searrow 0}\left\|\tilde{P}(t)\right\|_{L(\mathcal{B}^\rho(E))}\|f(x)-g_\varepsilon(x)\|_\rho\,\rho(x)\\ &\quad+\lim_{t\searrow 0}\left|\tilde{P}(t)g_\varepsilon(x)-g_\varepsilon(x)\right|\\ &\quad+|g_\varepsilon(x)-f(x)|\\ &\leq C\varepsilon\rho(x)+\varepsilon\rho(x).\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, this implies

$$\lim_{t \searrow 0} \tilde{P}(t) f(x) = f(x).$$

Hence, $\left(\tilde{P}(t)\right)_{t\in\mathbb{R}_+}$ is a generalized Feller semigroup.

Finally, since $(\lambda_t)_{t\in\mathbb{R}_+}$ is a Markov process with respect to its natural filtration $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ for any initial distribution ν and $f \in \mathscr{B}^\rho(E)$ and $0 \le s \le t$ it holds

$$\mathbb{E}_{\mathbb{P}_\nu}\left[f(\lambda_t) \middle| \mathcal{F}_s^0 \right] = \int_E f(y) p(t-s)(\lambda_s, dy) = \tilde{P}(t-s) f(\lambda_s)$$

$\mathbb{P}_\nu$ -almost surely. As in the last step of the proof in Proposition 2.3.69, this equation can be extended to the right continuous extension of $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$, which yields the statement of the proposition. □

For an investigation where E is not locally compact we introduce the following spaces:

DEFINITION 2.3.85. Set $\widetilde{\ell^\infty}(E) \subset \ell^\infty(E)$ as

$$\widetilde{\ell^\infty}(E) := \{ f : f \in \ell^\infty(E),\, f \text{ measurable} \},$$

and $\ell^\rho(E) \subset \widetilde{\ell^\infty}(E)$ as

$$\ell^\rho(E) := \left\{ \frac{f}{\rho} : f \in \mathscr{B}^\rho(E) \right\}.$$

REMARK 2.3.86. As closed subspaces of the Banach space $\ell^\infty(E)$, both $\widetilde{\ell^\infty}(E)$ and $\ell^\rho(E)$ are Banach spaces with respect to $\|\cdot\|_\infty$.

For the next proposition we recall the definition of

$$\widetilde{B^\rho}(E) := \widetilde{B^\rho}(E, \mathbb{R})$$

in Remark 2.3.23.

PROPOSITION 2.3.87. *Let ρ be an admissible weight function. Define*

$$\begin{aligned} \Phi : L(\widetilde{B^\rho}(E)) &\to L(\widetilde{\ell^\infty}(E)) \\ P &\to \frac{P\left((\cdot)\cdot\rho\right)}{\rho}. \end{aligned}$$

Then the following assertions hold true:
(i) Φ is an isometric isomorphism between $L(\widetilde{B^\rho}(E))$ and $L(\widetilde{\ell^\infty}(E))$,
(ii) $\Phi|_{L(\mathscr{B}^\rho(E)))}$ is an isometric isomorphism between $L(\mathscr{B}^\rho(E))$ and $L(\ell^\rho(E))$.

PROOF. Clearly, $\frac{P((\cdot)\cdot\rho)}{\rho} \in L(\widetilde{\ell^\infty}(E))$ is well defined and Φ is linear.
(i) We show first that Φ is an isometry.
We calculate for any $f \in \widetilde{B^\rho}(E)$

$$\begin{aligned}\|(\Phi P)\, f\|_\infty &= \|P\,(f\cdot\rho)\|_\rho \\ &\leq \|P\|_{L\left(\widetilde{B^\rho}(E)\right)} \cdot \|f\|_\infty .\end{aligned}$$

Hence

$$\|(\Phi P)\|_{L(\widetilde{\ell^\infty}(E))} \leq \|P\|_{L\left(\widetilde{B^\rho}(E)\right)} .$$

Furthermore, for $\varepsilon > 0$ let $g_\varepsilon \in \widetilde{B^\rho}(E)$ be such that

$$\|Pg_\varepsilon\|_\rho \geq \left(\|P\|_{L\left(\widetilde{B^\rho}(E)\right)} - \varepsilon\right)\|g_\varepsilon\|_\rho .$$

Then $\frac{g_\varepsilon}{\rho} \in \widetilde{\ell^\infty}(E)$ and

$$\begin{aligned}\left\|(\Phi P)\left(\frac{g_\varepsilon}{\rho}\right)\right\|_\infty &= \left\|\frac{P\left(\frac{g_\varepsilon}{\rho}\cdot\rho\right)}{\rho}\right\|_\infty \\ &= \|Pg_\varepsilon\|_\rho \\ &\geq \left(\|P\|_{L(\mathscr{B}^\rho(E))} - \varepsilon\right)\left\|\frac{g_\varepsilon}{\rho}\right\|_\infty ,\end{aligned}$$

which shows that

$$\|(\Phi P)\|_{L(\widetilde{\ell^\infty}(E))} \geq \|P\|_{L\left(\widetilde{B^\rho}(E)\right)} - \varepsilon.$$

Thus, Φ is an isometry.

Regarding Φ being an isometric isomorphism between $L(\widetilde{B^\rho}(E))$ and $L(\widetilde{\ell^\infty}(E))$ by injectivity of isomorphisms we only need to show

$$\Phi\left(L(\widetilde{B^\rho}(E))\right) = L(\widetilde{\ell^\infty}(E)).$$

However, this is clear since for any $Q \in L(\widetilde{\ell^\infty}(E))$

$$Q'(\cdot) := Q\left(\frac{(\cdot)}{\rho}\right)\cdot\rho$$

is a linear map from $\widetilde{B^\rho}(E)$ to $\widetilde{B^\rho}(E)$ and the calculation

$$\|Q'f\|_\rho = \left\|Q\left(\frac{f}{\rho}\right)\cdot\rho\right\|_\rho \leq \|Q\|_{L(\widetilde{\ell^\infty}(E))}\|f\|_\rho$$

shows that $Q' \in L(\widetilde{B^\rho}(E))$. Then $\Phi\,(Q') = Q$ yields surjectivity of Φ.

(ii) Follows just like (i). $\square$

COROLLARY 2.3.88. *There is an isometric isomorphism between contractive generalized Feller semigroups on $\mathscr{B}^\rho(E)$ and strongly continuous, contractive, positive (see Definition 2.3.39) semigroups on $\ell^\rho(E)$.*

PROOF. Use Proposition 2.3.87 above and strong continuity of generalized Feller semigroups (see Theorem 2.3.51). The respective required semigroup properties follow immediately. □

Next, we want to characterize positive semigroups on $\ell^\rho(E)$, for which it is possible to obtain a transformation to a contractive generalized Feller semigroup.

PROPOSITION 2.3.89. *Let $(Q(t))_{t\in\mathbb{R}_+}$ be a positive semigroup on $\ell^\rho(E)$ such that there is some $\omega \in \mathbb{R}$ such that for any $t \in \mathbb{R}_+$*

$$\|Q(t)\|_{L(\ell^\rho(E))} \leq e^{\omega t}.$$

Then $(P(t))_{t\in\mathbb{R}_+}$ defined by

$$P(t) : \mathscr{B}^\rho(E) \to \mathscr{B}^\rho(E)$$
$$f \to e^{-\omega t} Q(t)\left(\frac{f}{\rho}\right) \cdot \rho$$

is a contractive generalized Feller semigroup on E with generator A if and only if
(i) for any $g \in \ell^\rho(E)$ such that $g \cdot \rho \in C_b(E)$ and any $x \in E$

$$\lim_{t\searrow 0} \left(Q(t)g\right)(x) = g(x)$$

holds true.
In this case, $\left(Q(t)|_{\ell^\rho(E)}\right)_{t\in\mathbb{R}_+}$ is strongly continuous on $\ell^\rho(E)$ and its generator $\tilde{A}$ is given by $\tilde{A}f = \frac{A(f\cdot\rho)}{\rho}$.

PROOF. If $(P(t))_{t\in\mathbb{R}_+}$ is a generalized Feller semigroup on E, then

$$f \to e^{\omega t}\frac{P(t)\left(f\cdot\rho\right)}{\rho}$$

is a positive semigroup on $\ell^\rho(E)$ such that for any $t \in \mathbb{R}_+$

$$\|Q(t)\|_{L(\ell^\rho(E))} \leq e^{\omega t}.$$

P3 of the definition of generalized Feller semigroups states that for any $f \in \mathscr{B}^\rho(E)$ and any $x \in E$

$$\lim_{t\searrow 0} Q(t)\left(\frac{f}{\rho}\right)(x) = \left(\frac{f}{\rho}\right)(x),$$

hence in particular the equation holds true for any $f \in C_b(E)$. (ii) follows form Theorem 2.3.42.

On the other hand, let $(Q(t))_{t\in\mathbb{R}_+}$ fulfill (i) and (ii). $f \in \mathscr{B}^\rho(E)$ implies $\frac{f}{\rho} \in \ell^\rho(E)$, hence

$$P(t)f = Q(t)\left(\frac{f}{\rho}\right) \cdot \rho \in \mathscr{B}^\rho(E).$$

By

$$\|P(t)f\|_\rho = e^{-\omega t}\left\|Q(t)\left(\frac{f}{\rho}\right)\right\|_\infty \leq \|f\|_\rho,$$

$(P(t))_{t\in\mathbb{R}_+}$ is contractive.

Thus, $(P(t))_{t\in\mathbb{R}_+}$ is a contractive semigroup on $\mathscr{B}^\rho(E)$. It is a generalized Feller semigroup because **P1** and **P2** obviously hold true, and **P4** and **P5** follow from contractivity and positivity in the assumption. In order to show **P3** we observe that (i) implies that for any $f \in C_b(E)$ and any $x \in E$

$$\lim_{t\searrow 0}\left(P(t)f\right)(x) = f(x),$$

and by density of $C_b(E)$ in $\mathscr{B}^\rho(E)$ for any $\varepsilon > 0$ and $g \in \mathscr{B}^\rho(E)$ there is $f_\varepsilon \in C_b(E)$ such that $\|g - f_\varepsilon\|_\rho < \varepsilon$ and

$$\begin{aligned}
&\lim_{t\to 0} P(t)g(x) - g(x)\\
&= \lim_{t\to 0} P(t)g(x) - \lim_{t\to 0} P(t)f_\varepsilon(x) + \lim_{t\to 0} P(t)f_\varepsilon(x) - f_\varepsilon(x) + f_\varepsilon(x) - g(x)\\
&\leq \lim_{t\to 0}\|P(t)\left(g(x) - f_\varepsilon(x)\right)\|_\rho \rho(x) + \underbrace{\lim_{t\to 0} P(t)f_\varepsilon(x) - f_\varepsilon(x)}_{=0} + \underbrace{\|f_\varepsilon - g\|}_{\leq\varepsilon}\rho(x)\\
&\leq \lim_{t\to 0}\underbrace{\|P(t)\|_\rho}_{\leq 1}\varepsilon\rho(x) + \varepsilon\rho(x)\\
&\leq 2\varepsilon\rho(x).
\end{aligned}$$

Since $\varepsilon > 0$ was arbitrary

$$\lim_{t\to 0} P(t)g(x) = g(x)$$

for any $g \in \mathcal{B}^\rho(E)$ and any $x \in E$ which yields **P3**. Hence, $(P(t))_{t\in\mathbb{R}_+}$ is a contractive generalized Feller semigroup.

Regarding strong continuity of $\left(Q(t)|_{\ell^\rho(E)}\right)_{t\in\mathbb{R}_+}$ we use Proposition 2.3.87(ii) and for the generator $\tilde{A}$ observe that for any $f \in \ell^\rho(E)$

$$\lim_{t\searrow 0}\left\|Q(t)f - f - \frac{A\left(f\cdot\rho\right)}{\rho}\right\|_\infty = \lim_{t\searrow 0}\|P\left(f\cdot\rho\right) - f\cdot\rho - A\left(f\cdot\rho\right)\|_\rho = 0.$$

□

We now would like to take a closer look at the results in the important case when the admissible weight function is continuous.

COROLLARY 2.3.90. *Let* $(Q(t))_{t\in\mathbb{R}_+}$ *be a positive semigroup on* $\ell^\rho(E)$ *that fulfills the conditions of Proposition 2.3.89 for some* $\omega \in \mathbb{R}$ *and a continuous admissible weight function* ρ. *Then on the measurable space*

$$\left((E \cup \{\triangle\})^{\mathbb{R}_+}, \mathcal{B}(E \cup \{\Delta\})^{\mathbb{R}_+}\right)$$

for any probability measure ν *on* $(E, \mathcal{B}(E))$ *there exists a probability measure* $\mathbb{P}'_\nu$ *on*

$$\left((E \cup \{\triangle\})^{\mathbb{R}_+}, \mathcal{B}(E \cup \{\Delta\})^{\mathbb{R}_+}\right),$$

such that for the canonical process $(\gamma_t)_{t\in\mathbb{R}_+}$ *and the natural filtration* $(\mathcal{F}_t^0)_{t\in\mathbb{R}_+}$ *for any* $t \geq s \geq 0$ *and any* $f \in \mathscr{B}^\rho(E \cup \{\triangle\})$ *that is bounded or positive*

$$\mathbb{E}_{\mathbb{P}'_\nu}\left[f(\gamma_t)\middle|\, \mathcal{F}_s^0\right] = e^{-\omega t}\left(Q\left(t-s\right)(f)\right)(\gamma_s) \tag{2.3.26}$$

holds true $\mathbb{P}'_\nu$ *- almost surely and*

$$\mathbb{P}'_\nu \circ \gamma_0^{-1} = \nu.$$

Here, for any $t \in \mathbb{R}_+$ *and positive* $f \in \mathscr{B}^\rho(E\cup\{\triangle\})$ $Q(t)f$ *is interpreted as*

$$Q(t)f := \frac{P(t)\left(f \cdot \rho\right)}{\rho}$$

and the convention in Proposition 2.3.52(ii).

PROOF. Apply Proposition 2.3.89 and Theorem 2.3.73. □

LEMMA 2.3.91. *If the admissible weight function* ρ *is continuous, then* $C_0(E) = \ell^\rho(E)$.

PROOF. Follows from Lemma 2.3.47(ii) and (iii). □

COROLLARY 2.3.92. *If the admissible weight function* ρ *is continuous, then there is an isometric isomorphism between contractive generalized Feller semigroups on* $\mathscr{B}^\rho(E)$ *and strongly continuous, contractive, positive semigroups on* $C_0(E)$.

The following theorem is the reason why $(\gamma_t)_{t\in\mathbb{R}_+}$ was name *extended Feller process*. As stated before, we remind the reader that on general weighted spaces an extended Feller process is not automatically a Markov process due to the subtle measurability issues. However, this is true if the weighted space is locally compact with countable base.

THEOREM 2.3.93. *Let (E,ρ) be a weighted space and let ρ be continuous. Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ and let $\omega\in\mathbb{R}$ be such that for any $t\in\mathbb{R}_+$*

$$\|P(t)\|_{L(\mathscr{B}^\rho(E))}\le e^{\omega t}.$$

Then $(Q(t))_{t\in\mathbb{R}_+}$ defined as

$$Q(t)f:=e^{-\omega t}\frac{P(t)(f\cdot\rho)}{\rho}$$

is a strongly continuous, positive, contractive semigroup on $C_0(E)$ and for any probability measure ν on $(E,\mathcal{B}(E))$ there exists a probability measure $\mathbb{P}'_\nu$ on

$$\left(E^{\mathbb{R}_+},\mathcal{B}(E)^{\mathbb{R}_+}\right)$$

and a right continuous filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ such that and for any $t\ge s\ge 0$ and for any $f\in C_0(E)$ for the canonical process $(\gamma_t)_{t\in\mathbb{R}_+}$

$$\mathbb{E}_{\mathbb{P}'_\nu}\left[f(\gamma_t)\middle|\mathcal{F}_s\right]=Q(t-s)f(\gamma_s) \tag{2.3.27}$$

holds true $\mathbb{P}'_\nu$ - almost surely and

$$\mathbb{P}'_\nu\circ\lambda_0^{-1}=\nu$$

holds true.

PROOF. This follows immediately from Corollary 2.3.74 and Proposition 2.3.87. □

2.3.7. Path properties. In the main theorems of this subsection, path properties of generalized Feller processes and of extended Feller processes are shown. In the case of generalized Feller processes, this already was proved in [**14**]. However, we correct their statement in one point. Also in this subsection, $(E,\,\rho)$ always denotes a weighted space.

In order to prove path properties of generalized Feller processes, we first need the following result regarding regularity of submartingales. We base the proof on the one in [**37**].

PROPOSITION 2.3.94. *Let $\left(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in\mathbb{R}_+},\mathbb{P}\right)$ be a filtered probability space and let $(\lambda_t)_{t\in\mathbb{R}_+}$ be a submartingale with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. Let $\Omega'\subset\Omega$ in $\mathcal{F}$ be the set where*

$$\lambda_{t+}:=\lim_{r\searrow t,\,r\in\mathbb{Q}}\lambda_r$$

exists for any $t\in\mathbb{R}_+$. Define $\left(\bar\lambda_t\right)_{t\in\mathbb{R}_+}$ by

$$\bar\lambda_t:=\begin{cases}\lambda_{t+} & \text{on } \Omega'\\ 0 & \text{elsewhere.}\end{cases}$$

Then $\mathbb{P}(\Omega') = 1$ *and there is a set* $\tilde{\Omega} \subset \Omega$ *in* $\mathcal{F}$ *with* $\mathbb{P}\left(\tilde{\Omega}\right) = 1$ *such that* $\left(\bar{\lambda}_t\right)_{t\in\mathbb{R}_+}$ *has càdlàg paths on* $\tilde{\Omega}$. *Moreover,* $\left(\bar{\lambda}_t\right)_{t\in\mathbb{R}_+}$ *is a modification of* $(\lambda_t)_{t\in\mathbb{R}_+}$ *if*

$$t \to \lambda_t$$

is right-continuous in $L^1(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. By Theorem A.3.96 $\mathbb{P}(\Omega') = 1$. By Proposition A.3.98 $(\lambda_{t+})_{t\in\mathbb{R}_+}$ is a submartingale with respect to $(\mathcal{F}_{t+})_{t\in\mathbb{R}_+}$ and by Theorem A.3.96 there is a set $\hat{\Omega} \subset \Omega$ in $\mathcal{F}$ with $\mathbb{P}\left(\hat{\Omega}\right) = 1$ on which

$$\lim_{r\nearrow t,\, r\in\mathbb{Q}} \lambda_{r+}$$

exists. One can then show easily that $(\lambda_{t+})_{t\in\mathbb{R}_+}$ has càdlàg paths on $\tilde{\Omega} := \hat{\Omega} \cap \Omega'$. Thus, by definition, this is true also for $\left(\bar{\lambda}_t\right)_{t\in\mathbb{R}_+}$.

Fix some arbitrary $t_0 \in \mathbb{R}_+$. By definition of $\bar{\lambda}_{t_0}$ and $\mathbb{P}(\Omega') = 1$,

$$\lim_{r\searrow t_0,\, r\in\mathbb{Q}} \lambda_r = \bar{\lambda}_{t_0}$$

$\mathbb{P}$-almost surely. By Theorem A.3.93 (where in a neighbourhood $O(t_0)$ of t_0 boundedness of $\sup_{r\in O(t_0)\cap\mathbb{Q}} \mathbb{E}\left[|\lambda_r|\right]$ follows from right-continuity in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ of $t \to \lambda_t$)

$$\lim_{r\searrow t_0,\, r\in\mathbb{Q}} \lambda_r$$

converges in $L^1(\Omega, \mathcal{F}, \mathbb{P})$. By uniqueness of the limits (which follows from Proposition A.3.53, Proposition A.3.56 and Remark A.3.50)

$$\lim_{r\searrow t_0,\, r\in\mathbb{Q}} \lambda_r = \bar{\lambda}_{t_0}$$

in $L^1(\Omega, \mathcal{F}, \mathbb{P})$. Thus, right-continuity in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ of $t \to \lambda_t$ implies

$$\left\|\bar{\lambda}_{t_0} - \lambda_{t_0}\right\|_{L^1(\Omega,\mathcal{F},\mathbb{P})} = 0,$$

or, what is equivalent, $\bar{\lambda}_{t_0} = \lambda_{t_0}$ $\mathbb{P}$-almost surely. As $t_0 \in \mathbb{R}_+$ was arbitrary, we conclude. □

Moreover, [**37**] mentions that a similar result holds true also for left continuity, which we prove below:

Proposition 2.3.95. *Let* $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ *be a filtered probability space and let* $(\lambda_t)_{t\in\mathbb{R}_+}$ *be a submartingale with respect to* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. *Let* $\Omega' \subset \Omega$ *be the set in* $\mathcal{F}$ *where*

$$\lambda_{t-} := \lim_{r\nearrow t,\, r\in\mathbb{Q}} \lambda_r$$

exists for any $t \in \mathbb{R}_+$. Define $\left(\hat{\lambda}_t\right)_{t\in\mathbb{R}_+}$ by

$$\hat{\lambda}_t := \begin{cases} \lambda_{t-} & \text{on } \Omega' \\ 0 & \text{elsewhere.} \end{cases}$$

Then $\mathbb{P}(\Omega') = 1$ and there is a set $\tilde{\Omega} \subset \Omega$ in $\mathcal{F}$ with $\mathbb{P}\left(\tilde{\Omega}\right) = 1$ such that $\left(\hat{\lambda}_t\right)_{t\in\mathbb{R}_+}$ has càglàd paths on $\tilde{\Omega}$. Moreover, $\left(\hat{\lambda}_t\right)_{t\in\mathbb{R}_+}$ is a modification of $(\lambda_t)_{t\in\mathbb{R}_+}$ if

$$t \to \lambda_t$$

is left-continuous in $L^1(\Omega, \mathcal{F}, \mathbb{P})$.

Proof. By Theorem A.3.96 $\mathbb{P}(\Omega') = 1$. By Proposition A.3.99 $(\lambda_{t-})_{t\in\mathbb{R}_+}$ is a submartingale with respect to $(\mathcal{F}_{t-})_{t\in\mathbb{R}_+}$ and by Theorem A.3.96 there is a set $\hat{\Omega} \subset \Omega$ in $\mathcal{F}$ with $\mathbb{P}\left(\hat{\Omega}\right) = 1$ on which

$$\lim_{r\searrow t,\, r\in\mathbb{Q}} \lambda_{r-}$$

exists. One can then show easily that $(\lambda_{t-})_{t\in\mathbb{R}_+}$ has càglàd paths on $\tilde{\Omega} := \hat{\Omega} \cap \Omega'$. Thus, by definition, also $\left(\hat{\lambda}_t\right)_{t\in\mathbb{R}_+}$.

Fix some arbitrary $s_0, t_0 \in \mathbb{R}_+$ such that $s_0 < t_0$. By definition of $\left(\hat{\lambda}_t\right)_{t\in\mathbb{R}_+}$ and $\mathbb{P}(\Omega') = 1$,

$$\lim_{r\nearrow t_0,\, r\in\mathbb{Q}} \lambda_r = \hat{\lambda}_{t_0}$$

$\mathbb{P}$-almost surely. Fix an arbitrary $c \in \mathbb{R}_-$. We next show uniform integrability (see Definition A.3.54) of the family of random variables $(\max(\lambda_r, c))_{r\in\mathbb{Q}\cap[s_0,t_0]}$. Let $a \in \mathbb{R}_+$. Then

$$\begin{aligned}
&\int_{\{|\max(\lambda_r,c)|>a\}} |\max(\lambda_r, c)|\, d\mathbb{P} \\
&= \int_{\{\max(\lambda_r,c)>a\}} \max(\lambda_r, c)\, d\mathbb{P} - \int_{\{\max(\lambda_r,c)<-a\}} \max(\lambda_r, c)\, d\mathbb{P} \\
&\leq \int_{\{\lambda_r>a\}} \left(\lambda_r 1_{\{\lambda_r>c\}} + c 1_{\{\lambda_r\leq c\}}\right) d\mathbb{P} - \int_{\{\max(\lambda_r,c)<-a\}} c d\mathbb{P} \\
&\leq \int_{\{\lambda_r>a\}} \lambda_r d\mathbb{P} - \int_{\{\max(\lambda_r,c)<-a\}} c d\mathbb{P} \\
&\leq \int_{\{\lambda_r>a\}} \lambda_{t_0} d\mathbb{P} - \int_{\{\max(\lambda_r,c)<-a\}} c d\mathbb{P}
\end{aligned}$$

Moreover,

$$\begin{aligned}\mathbb{P}\left(\{\lambda_r > a\}\right) &\leq \frac{1}{a}\mathbb{E}\left[\lambda_r 1_{\{\lambda_r > a\}}\right] \\ &\leq \frac{1}{a}\mathbb{E}\left[|\lambda_r|\right].\end{aligned}$$

Hence,

$$\lim_{a\to\infty}\mathbb{P}\left(\{\lambda_r > a\}\right) = 0,$$

and clearly also

$$\lim_{a\to\infty}\mathbb{P}\left(\{\max\left(\lambda_r, c\right) < -a\}\right) = 0.$$

Therefore,

$$\lim_{a\to\infty}\int_{\{|\max(\lambda_r,c)|>a\}}\left|\max\left(\lambda_r, c\right)\right| d\mathbb{P} = 0$$

and the family of random variables $\left(\max\left(\lambda_r, c\right)\right)_{r\in\mathbb{Q}\cap[s_0,t_0]}$ is uniformly integrable (see Proposition A.3.55).

By Proposition A.3.56 (and Proposition A.3.53)

$$\lim_{r\nearrow t_0,\, r\in\mathbb{Q}}\max\left(\lambda_r, c\right)$$

converges in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ to $\max\left(\hat{\lambda}_{t_0}, c\right)$. Left-continuity in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ of $t \to \lambda_t$ yields

$$\begin{aligned}&\lim_{r\nearrow t_0}\left(\int_\Omega\left|\max\left(\lambda_r, c\right) - \max\left(\lambda_{t_0}, c\right)\right| d\mathbb{P}\right) \\ &= \lim_{r\nearrow t_0}\left(\int_\Omega |\lambda_r - \lambda_{t_0}|\, 1_{\{\lambda_r>c,\, \lambda_{t_0}>c\}} + |c - \lambda_{t_0}|\, 1_{\{\lambda_r\leq c,\, \lambda_{t_0}>c\}} + |\lambda_r - c|\, 1_{\{\lambda_r>c,\, \lambda_{t_0}\leq c\}} d\mathbb{P}\right) \\ &\leq \lim_{r\nearrow t_0}\left(\int_\Omega |\lambda_r - \lambda_{t_0}|\, 1_{\{\lambda_r>c,\, \lambda_{t_0}>c\}} + |\lambda_r - \lambda_{t_0}|\, 1_{\{\lambda_r\leq c,\, \lambda_{t_0}>c\}} + |\lambda_r - \lambda_{t_0}|\, 1_{\{\lambda_r>c,\, \lambda_{t_0}\leq c\}} d\mathbb{P}\right) \\ &= 0.\end{aligned}$$

This implies

$$\left\|\max\left(\hat{\lambda}_{t_0}, c\right) - \max\left(\lambda_{t_0}, c\right)\right\|_{L^1(\Omega,\mathcal{F},\mathbb{P})} = 0,$$

or, what is equivalent, $\max\left(\hat{\lambda}_{t_0}, c\right) = \max\left(\lambda_{t_0}, c\right)$ $\mathbb{P}$-almost surely. Since $c \in \mathbb{R}_-$ was arbitrary and the countable union of null sets is a null set

$$\hat{\lambda}_{t_0} = \lim_{n\to\infty}\max\left(\hat{\lambda}_{t_0}, -n\right) = \lim_{n\to\infty}\max\left(\lambda_{t_0}, -n\right) = \lambda_{t_0}$$

$\mathbb{P}$-almost surely.

As $t_0 \in \mathbb{R}_+$ was arbitrary, we conclude. $\square$

In order to state the result regarding path properties of generalized Feller processes, we remind the reader of the definition of the resolvent in Definition 1.4.25. Since by Theorem 2.3.51 a generalized Feller semigroup is strongly continuous, we can define the generator of the semigroup (see Definition 1.4.13)

THEOREM 2.3.96. *Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ such that for any $t \in \mathbb{R}_+$ $P(t)1 = 1$ and*

$$\|P(t)\| \leq Me^{\omega t}$$

for some $M \geq 1$ and $\omega \in \mathbb{R}$. Let A be the generator of $(P(t))_{t\in\mathbb{R}_+}$. Let $x_0 \in E$ and let $(\lambda_t)_{t\in\mathbb{R}_+}$ be the generalized Feller process from Theorem 2.3.65 on the measurable space

$$\left(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\mathbb{R}_+}\right)$$

with probability measure $\mathbb{P}_{x_0}$ and right continuous filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ as in Theorem 2.3.65.

(i) For every countable family $(f_n)_{n\in\mathbb{N}} \subset \mathscr{B}^\rho(E)$ and $\beta > \omega$, $\beta \in \mathbb{N}$ for the family of stochastic processes $\left(Z_t^{\beta,n}\right)_{t\in\mathbb{R}_+}$ defined as

$$Z_t^{\beta,n} := \beta R(\beta, A) f_n(\lambda_t)$$

there exists a family of stochastic processes

$$\left(\bar{Z}_t^{\beta,n}\right)_{t\in\mathbb{R}_+}$$

with càdlàg paths (and one with càglàd paths), such that for all $t \in \mathbb{R}_+$ there is a $\mathbb{P}_{x_0}$- null set $\mathcal{N}_t \in \sigma\left(\bigcup_{t\in\mathbb{R}_+} \mathcal{F}_t\right)$ for which

$$Z_t^{\beta,n} = \bar{Z}_t^{\beta,n} \text{ on } E^{\mathbb{R}_+} \setminus \mathcal{N}_t$$

for all $n \in \mathbb{N}$ and all $\beta > \omega$, $\beta \in \mathbb{N}$.

(ii) Let ρ be Baire measurable and let $\nu \in \mathcal{M}^\rho(E)$ be the initial distribution. If additionally to the assumptions in (i) $M = 1$ holds true, then

$$(\exp(-\omega t)\,\rho(\lambda_t))_{t\in\mathbb{R}_+}$$

is a supermartingale and if $t \to P(t)\rho(x)$ is continuous for ν-almost any $x \in E$, then the supermartingale has a version such that the paths are càdlàg or càglàd. In this case, there exists a family of stochastic processes with càdlàg paths (and one with càglàd paths)

$$\left(\left(\overline{f_n(\lambda_t)}\right)_{t\in\mathbb{R}_+}\right)_{n\in\mathbb{N}}$$

such that for all $t \in \mathbb{R}_+$ there is a null set $\mathcal{N}_t' \in \mathcal{B}(E)^{\mathbb{R}_+}$ for which

$$f_n(\lambda_t) = \overline{f_n(\lambda_t)} \text{ on } E^{\mathbb{R}_+} \setminus \mathcal{N}_t'$$

for all $n \in \mathbb{N}$.

(iii) If additionally to the assumptions in (i) and (ii) there exists a countable family $(f_n)_{n\in\mathbb{N}} \subset \mathscr{B}^\rho(E)$ of sequentially continuous functions, i.e. for any $(x_m)_{m\in\mathbb{N}} \subset E$ with $x_m \to x \in E$ for any $n \in \mathbb{N}$

$$f_n(x_m) \to f_n(x),$$

and if this family separates points, i.e for any $y, z \in E$,$y \neq z$ there exists $l \in \mathbb{N}$ such that

$$f_l(y) \neq f_l(z),$$

then $(\lambda_t)_{t\in\mathbb{R}_+}$ has a version with càdlàg paths (and one with càglàd paths).

Proof. (i) We only treat the càdlàg case since the càglàd case follows along the same lines. By Theorem 1.4.29 if $\alpha > \omega$, then $\alpha \in \rho(A)$ and for all $f \in \mathscr{B}^\rho(E)$

$$(\alpha - A)^{-1} f := R(\alpha, A)f = \int_0^\infty e^{-\alpha s} P(s) f \, ds.$$

In order to find a càdlàg version, we would like to use Proposition 2.3.94. For this purpose, we fix $f \in \mathscr{B}^\rho(E)$, $f \geq 0$ and $\alpha > \omega$ we define the stochastic process $\left(Y_t^{\alpha,f}\right)_{t\in\mathbb{R}_+}$ by

$$Y_t^{\alpha,f} := \exp(-\alpha t) \, R(\alpha, A) f(\lambda_t).$$

We show that $\left(Y_t^{\alpha,f}\right)_{t\in\mathbb{R}_+}$ is a supermartingale with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. Let $0 \leq s \leq t$ and calculate

$$\mathbb{E}_{x_0}\left[Y_t^{\alpha,f} \middle| \mathcal{F}_s \right] = \exp(-\alpha t) \, \mathbb{E}_{x_0}\left[\int_0^\infty \exp(-\alpha u) \, P(u) f(\lambda_t) du \middle| \mathcal{F}_s \right].$$

By the definition of the Riemann integral (see Definition 1.3.5), positivity of the semigroup $(P(t))_{t\in\mathbb{R}_+}$, and monotone convergence for conditional expectations (Proposition A.3.65) and thanks to $(\lambda_t)_{t\in\mathbb{R}_+}$ being a generalized Feller process

$$\begin{aligned}
&\mathbb{E}_{x_0}\left[\int_0^\infty \exp\left(-\alpha u\right) P(u) f(\lambda_t) du \middle| \mathcal{F}_s\right] \\
&= \mathbb{E}_{x_0}\left[\lim_{n\to\infty}\lim_{m\to\infty} (n/m)\cdot \sum_{i=0}^{m-1} \exp\left(-\alpha in/m\right) P(in/m) f(\lambda_t) \middle| \mathcal{F}_s\right] \\
&= \lim_{n\to\infty}\lim_{m\to\infty} (n/m)\cdot \sum_{i=0}^{m-1} \exp\left(-\alpha in/m\right) \mathbb{E}_{x_0}\left[P(in/m) f(\lambda_t) \middle| \mathcal{F}_s\right] \\
&= \lim_{n\to\infty}\lim_{m\to\infty} (n/m)\cdot \sum_{i=0}^{m-1} \exp\left(-\alpha in/m\right) P(in/m + t - s) f(\lambda_s) \\
&= \exp\left(\alpha\left(t-s\right)\right) \lim_{n\to\infty}\lim_{m\to\infty} (n/m)\cdot \sum_{i=0}^{m-1} \exp\left(-\alpha\left(in/m + t - s\right)\right) P(in/m + t - s) f(\lambda_s).
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathbb{E}_{x_0}\left[Y_t^{\alpha,f} \middle| \mathcal{F}_s\right] &= \exp\left(-\alpha s\right) \int_{t-s}^\infty \left(\exp\left(-\alpha r\right) P(r) f(\lambda_s)\right) dr \\
&\leq Y_s^{\alpha,f}.
\end{aligned}$$

The last inequality followed from $f \geq 0$ and positivity of the semigroup $(P(t))_{t\in\mathbb{R}_+}$. Furthermore,

$$\mathbb{E}_{x_0}\left[Y_t^{\alpha,f}\right] = \exp\left(-\alpha s\right) \int_{t-s}^\infty \left(\exp\left(-\alpha r\right) \mathbb{E}_{x_0}\left[P(r) f(\lambda_s)\right]\right) dr,$$

and due to absolute continuity of the integral (Theorem A.3.43)

$$t \to \int_{t-s}^\infty \left(\exp\left(-\alpha r\right) \mathbb{E}_{x_0}\left[P(r) f(\lambda_s)\right]\right) dr,$$

and therefore

$$t \to \mathbb{E}_{x_0}\left[Y_t^{\alpha,f}\right]$$

is continuous.

We can apply Proposition 2.3.94 to $\left(-Y_t^{\alpha,f}\right)_{t\in\mathbb{R}_+}$ and obtain that there is a set $\Omega'_{\alpha,f} \in \mathcal{B}(E)^{\mathbb{R}_+}$ with $\mathbb{P}_{x_0}\left(\Omega'_{\alpha,f}\right) = 1$ on which

$$\lim_{r\searrow t,\, r\in\mathbb{Q}} - Y_s^{\alpha,f}$$

exists and there is a set $\tilde{\Omega}_{\alpha,f} \subset \Omega'_{\alpha,f}$ in $\mathcal{F}$ with $\mathbb{P}_{x_0}\left(\tilde{\Omega}_{\alpha,f}\right) = 1$ such that $\left(-\bar{Y}_t^{\alpha,f}\right)_{t\in\mathbb{R}_+}$ defined as

$$-\bar{Y}_t^{\alpha,f} := \begin{cases} \lim_{r\searrow t,\, r\in\mathbb{Q}} - Y_s^{\alpha,f} & \text{on } \Omega'_{\alpha,f} \\ 0 & \text{elsewhere,} \end{cases}$$

has càdlàg paths on $\tilde{\Omega}_{\alpha,f}$. Moreover,

$$\left(-\bar{Y}_t^{\alpha,f}\right)_{t\in\mathbb{R}_+}$$

is a version of

$$\left(-Y_t^{\alpha,f}\right)_{t\in\mathbb{R}_+}.$$

Therefore, for any $\beta > \omega\ \beta, n \in \mathbb{N}$ and $f_n \in \mathscr{B}^\rho(E)$, $f_n \geq 0$ for the stochastic process $\left(Z_t^{\beta,n}\right)_{t\in\mathbb{R}_+}$ defined as

$$Z_t^{\beta,n} := \beta R(\beta, A) f_n(\lambda_t)$$

the limit

$$\lim_{r\searrow t,\, r\in\mathbb{Q}} Z_t^{\beta,n}$$

exists on a set $\Omega'_{\beta,n} \subset \Omega$ in $\mathcal{F}$ with $\mathbb{P}_{x_0}\left(\Omega'_{\beta,n}\right) = 1$. $\left(Z_t^{\beta,n}\right)_{t\in\mathbb{R}_+}$ has a version $\left(\bar{Z}_t^{\beta,n}\right)_{t\in\mathbb{R}_+}$ defined as

$$\bar{Z}_t^{\beta,n} := \begin{cases} \lim\limits_{r\searrow t,\, r\in\mathbb{Q}} Z_t^{\beta,n} & \text{where it exists} \\ 0 & \text{elsewhere,} \end{cases}$$

with càdlàg paths on $\tilde{\Omega}_{\beta,n}$ in $\mathcal{F}$ with $\mathbb{P}_{x_0}\left(\tilde{\Omega}_{\beta,n}\right) = 1$. Since the countable union of null sets is a null set, for any $t \in \mathbb{R}_+$ there exists a null set $\mathcal{N}_t \in \mathcal{B}(E)^{\mathbb{R}_+}$ such that $Z_t^{\beta,n} = \bar{Z}_t^{\beta,n}$ on $E^{\mathbb{R}_+} \setminus \mathcal{N}_t$ for any $\beta, n \in \mathbb{N}$, $\beta > \omega$. In particular, for any $\beta > \omega\ \beta \in \mathbb{N}$ and any $g \in \mathscr{B}^\rho(E)$, clearly $g^+, g^- \in \mathscr{B}^\rho(E)$, and $g^+, g^- \geq 0$ and the process $\left(Z_t^{\beta,g}\right)_{t\in\mathbb{R}_+}$ defined as

$$\beta R(\beta, A) g(\lambda_t) = \beta R(\beta, A) g^+(\lambda_t) - \beta R(\beta, A) g^-(\lambda_t)$$

has a version with càdlàg paths.

(ii) As before, we only show the càdlàg case, as the càglàd case follows along the same lines.

If $P(t)\rho \leq \exp(\omega t)\,\rho$ holds for some $\omega \in \mathbb{R}$, then

$$\begin{aligned} \mathbb{E}_\nu\left[\exp(-\omega t)\,\rho(\lambda_t)\middle|\, \mathcal{F}_s\right] &= \exp(-\omega t)\, P(t-s)\rho(\lambda_s) \\ &\leq \exp(-\omega s)\,\rho(\lambda_s) \end{aligned}$$

and $\left(\exp(-\omega t)\,\rho(\lambda_t)\right)_{t\in\mathbb{R}_+}$ is a non-negative supermartingale. By Proposition 2.3.94 there exists a set $\Omega'_\rho \subset \Omega$ in $\mathcal{F}$ with $\mathbb{P}_\nu(\Omega'_\rho) = 1$ such that

$$\lim_{r\searrow t,\, r\in\mathbb{Q}} \rho(\lambda_r)$$

exist and a stochastic process $\left(\overline{\rho(\lambda_t)}\right)_{t\in\mathbb{R}_+}$ with càdlàg paths that is a version of $(\rho(\lambda_t))_{t\in\mathbb{R}_+}$.

Furthermore, by Yosida approximation (Proposition 1.4.34)

$$\lim_{\beta\to\infty}\|\beta R(\beta,A)f_n - f_n\|_\rho = 0. \tag{2.3.28}$$

Hence uniformly in $t\in\mathbb{R}_+$

$$\lim_{\beta\to\infty}\left|\limsup_{r\searrow t,\, r\in\mathbb{Q}}\frac{\beta R(\beta,A)f_n(\lambda_r)}{\rho(\lambda_r)} - \limsup_{r\searrow t,\, r\in\mathbb{Q}}\frac{f_n(\lambda_r)}{\rho(\lambda_r)}\right| = 0,$$

and

$$\lim_{\beta\to\infty}\left|\liminf_{r\searrow t,\, r\in\mathbb{Q}}\frac{\beta R(\beta,A)f_n(\lambda_r)}{\rho(\lambda_r)} - \liminf_{r\searrow t,\, r\in\mathbb{Q}}\frac{f_n(\lambda_r)}{\rho(\lambda_r)}\right| = 0.$$

Thus, when

$$\lim_{r\searrow t,\, r\in\mathbb{Q}}\frac{Z_r^{\beta,n}}{\rho(\lambda_r)}$$

exists for any large $\beta\in\mathbb{N}$, then does

$$\lim_{r\searrow t,\, r\in\mathbb{Q}}\frac{f_n(\lambda_r)}{\rho(\lambda_r)}.$$

As seen in (i) for any $\beta>\omega$, $\beta\in\mathbb{N}$ there is a set $\Omega'_{\beta,n}\subset\Omega$ in $\mathcal{F}$ with $\mathbb{P}_\nu\left(\Omega'_{\beta,n}\right)=1$ on which

$$\lim_{r\searrow t,\, r\in\mathbb{Q}} Z_r^{\beta,n}$$

exists for any $t\in\mathbb{R}_+$. Since $\rho>0$ attains its minimum on E (see Corollary 2.3.17) on $\Omega'_{\beta,n}\cap\Omega'_\rho$ also

$$\lim_{r\searrow t,\, r\in\mathbb{Q}}\frac{Z_r^{\beta,n}}{\rho(\lambda_r)}$$

exists and for any $t\in\mathbb{R}_+$ we define

$$\overline{\frac{f_n}{\rho}(\lambda_t)} := \begin{cases}\lim_{r\searrow t,\, r\in\mathbb{Q}}\frac{f_n(\lambda_r)}{\rho(\lambda_r)} & \text{on } \bigcap_{\beta\in\mathbb{N},\,\beta>\omega}\Omega'_{\beta,n}\cap\Omega'_\rho\\ 0 & \text{elsewhere.}\end{cases}$$

$\left(\overline{\frac{f_n}{\rho}(\lambda_t)}\right)_{t\in\mathbb{R}_+}$ is a version of $\left(\frac{f_n(\lambda_t)}{\rho(\lambda_t)}\right)_{t\in\mathbb{R}_+}$ since for any $\delta > 0$ and β large enough on $\bigcap_{\beta\in\mathbb{N},\,\beta>\omega} \Omega'_{\beta,n} \cap \Omega'_\rho$

$$\begin{aligned}\left|\overline{\frac{f_n}{\rho}(\lambda_t)} - \frac{f_n(\lambda_t)}{\rho(\lambda_t)}\right| &\leq \left|\lim_{r\searrow t,\, r\in\mathbb{Q}} \frac{f_n(\lambda_r)}{\rho(\lambda_r)} - \lim_{r\searrow t,\, r\in\mathbb{Q}} \frac{Z_r^{\beta,n}}{\rho(\lambda_r)}\right| \\ &\quad + \left|\lim_{r\searrow t,\, r\in\mathbb{Q}} \frac{Z_r^{\beta,n}}{\rho(\lambda_r)} - \frac{Z_t^{\beta,n}}{\rho(\lambda_t)}\right| + \left|\frac{Z_t^{\beta,n}}{\rho(\lambda_t)} - \frac{f_n(\lambda_t)}{\rho(\lambda_t)}\right| \\ &\leq \left|\lim_{r\searrow t,\, r\in\mathbb{Q}} \frac{Z_r^{\beta,n}}{\rho(\lambda_r)} - \frac{Z_t^{\beta,n}}{\rho(\lambda_t)}\right| + 2\delta \\ &= \left|\overline{\overline{\frac{\bar{Z}_t^{\beta,n}}{\rho(\lambda_t)}}} - \frac{Z_t^{\beta,n}}{\rho(\lambda_t)}\right| + 2\delta,\end{aligned}$$

and we know that for all $\beta, n \in \mathbb{N}$, $\beta > \omega$

$$\left(\overline{\overline{\frac{\bar{Z}_t^{\beta,n}}{\rho(\lambda_t)}}}\right)_{t\in\mathbb{R}_+}$$

and

$$\left(\frac{Z_t^{\beta,n}}{\rho(\lambda_t)}\right)_{t\in\mathbb{R}_+}$$

are versions of each other. We obtain for any $\varepsilon > 0$ and for β large enough and any $s, t \in \mathbb{R}_+$ on $\bigcap_{\beta\in\mathbb{N},\,\beta>\omega} \Omega'_{\beta,n} \cap \Omega'_\rho$

$$\left|\overline{\frac{f_n}{\rho}(\lambda_t)} - \overline{\frac{f_n}{\rho}(\lambda_s)}\right| \leq 2\varepsilon + \left|\overline{\overline{\frac{\bar{Z}_t^{\beta,n}}{\rho(\lambda_t)}}} - \overline{\overline{\frac{\bar{Z}_s^{\beta,n}}{\rho(\lambda_s)}}}\right|.$$

Therefore, also $\left(\overline{\frac{f_n}{\rho}(\lambda_t)}\right)_{t\in\mathbb{R}_+}$ has càdlàg paths.

Thus, for any $n \in \mathbb{N}$

$$\left(\overline{\frac{f_n}{\rho}(\lambda_t)} \cdot \overline{\rho(\lambda_t)}\right)_{t\in\mathbb{R}_+}$$

has càdlàg paths and is a version of

$$(f_n(\lambda_t))_{t\in\mathbb{R}_+}\,.$$

The statement of the theorem then follows from the fact that the countable union of null sets is a null set.

(iii) follows directly. □

EXAMPLE 2.3.97. Unlike stated in [**14**], Theorem 2.13 it is in general not true that for $f \subset \mathscr{B}^\rho(E)$ and generalized Feller process $(\lambda_t)_{t\in\mathbb{R}_+}$ the stochastic process

$$\left(\frac{f(\lambda_t)}{\rho(\lambda_t)}\right)_{t\in\mathbb{R}_+}$$

has a version with left-continuous paths. As a counterexample take $E = \mathbb{R}$,

$$\rho(x) := \begin{cases} -x & \text{if } x < -1 \\ x+2 & \text{if } -1 \leq x < 0 \\ x+1 & \text{if } x \geq 0, \end{cases}$$

and $f(x) := 1$ for $x \in \mathbb{R}$ with $\lambda_t(x) = x + t$. According to Proposition 2.3.54

$$P(t)(f) := f \circ \lambda_t$$

is a generalized Feller semigroup of transport type and as shown in Example 2.3.82 $(\lambda_t)_{t\in\mathbb{R}_+}$ is a generalized Feller process. But for $x < 0$ and $\mathbb{P}_x$ as in Example 2.3.82 clearly

$$\frac{f(\lambda_t)}{\rho(\lambda_t)} = \frac{1}{\rho(x+t)}$$

is $\mathbb{P}_x$-almost surely not left-continuous at $t = -x$.

COROLLARY 2.3.98. *Let E be separable and locally compact and let $\nu \in \mathcal{M}^\rho(E)$ be a probability measure . If $(P(t))_{t\in\mathbb{R}_+}$ is a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ such that for any $t \in \mathbb{R}_+$*

$$P(t)1 = 1$$

and $P(t)\rho \leq \exp(\omega t)\,\rho$ holds true for some $\omega \in \mathbb{R}$, and

$$t \to P(t)\rho(x)$$

is continuous for ν-almost any $x \in E$, then the generalized Feller process $(\lambda_t)_{t\in\mathbb{R}_+}$ associated to $(P(t))_{t\in\mathbb{R}_+}$ via Theorem 2.3.65 has a càdlàg or càglàd version. In particular, this is the case for a generalized Feller semigroup on $\mathscr{B}^\rho(E)$ of transport type with an appropriately chosen weight function (see Lemma 2.3.62).

PROOF. One can easily find a countable sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(E)$ that separates points and apply Theorem 2.3.96 (iii). □

For the generalized Feller semigroup $(P(t))_{t\in\mathbb{R}_+}$ on $\mathscr{B}^\rho(E)$ let the conditions of Theorem 2.3.73 be satisfied and let $(\gamma_t)_{t\in\mathbb{R}_+}$ be the corresponding stochastic process. Then according to Proposition 2.3.87 the semigroup $(Q(t))_{t\in\mathbb{R}_+}$ on $\ell^\rho(E)$ (see Definition 2.3.85) defined by

$$Q(t)f = \frac{P(t)\,(f\cdot\rho)}{\rho}$$

is strongly continuous, contractive and positive. In order to show regularity of the paths of $f(\gamma_t)$ for any $f\in\ell^\rho(E)$ one can proceed as in the proof of Theorem 2.3.96 but for the Yosida approximation in Equation 2.3.28 one obtains an approximation with respect to the norm $\|\cdot\|_\infty$. This yields the following result:

THEOREM 2.3.99. *Let $(P(t))_{t\in\mathbb{R}_+}$ be a generalized Feller semigroup on $\mathscr{B}^\rho(E)$, let the conditions of Theorem 2.3.73 be satisfied and let $(\gamma_t)_{t\in\mathbb{R}_+}$ be the corresponding stochastic process on*

$$\left((E\cup\{\triangle\})^{\mathbb{R}_+}, \mathcal{B}(E\cup\{\Delta\})^{\mathbb{R}_+}\right).$$

(i) For every countable family $(f_n)_{n\in\mathbb{N}}\subset\ell^\rho(E\cup\{\triangle\})$ there exists a family of stochastic processes with càdlàg or càglàd paths

$$\left(\left(\overline{f_n(\gamma_t)}\right)_{t\in\mathbb{R}_+}\right)_{n\in\mathbb{N}}$$

such that for all $t\in\mathbb{R}_+$ there is a null set $\mathcal{N}_t\in\mathcal{B}(E\cup\{\Delta\})^{\mathbb{R}_+}$ for which

$$f_n(\gamma_t) = \overline{f_n(\gamma_t)} \text{ on } (E\cup\{\triangle\})^{\mathbb{R}_+}\setminus\mathcal{N}_t$$

for all $n\in\mathbb{N}$.

(ii) If additionally to the assumption in (i) there exists a countable family $(f_n)_{n\in\mathbb{N}}\subset\ell^\rho(E\cup\{\triangle\})$ of sequentially continuous functions that separates points, then $(\gamma_t)_{t\in\mathbb{R}_+}$ has a version with càdlàg or càglàd paths.

CHAPTER 3

Affine and Polynomial Processes

Affine processes and polynomial processes are a special classes of Markov processes.

Let V be a d-dimensional vector space with scalar product $\langle \cdot, \cdot \rangle$. On $V + iV$ the scalar product is defined as

$$\langle a + ib, c + id \rangle = \langle a, c \rangle + i \langle b, c \rangle + i \langle a, d \rangle - \langle b, d \rangle .$$

Let $\|\cdot\|$ denote the norm induced by scalar product. Let $E \subset V$ be a subset and let $\mathcal{B}(E)$ and $\mathcal{B}(V)$ be the respective Borel σ-algebras. $S(V)$ ($S_+(V)$) denotes the set of (positive semidefinite) symmetric matrices on V. We recall the definition of semigroups of transition probabilities (Definition 2.1.3) and Markov processes (Definition 2.1.8) and will use Notation 2.1.14. Moreover, we remind the reader of the definition of the cemetery state Δ in Remark 2.1.2 and the convention $f(\Delta) = 0$ for any map f and add the convention $\|\Delta\| = \infty$. We write $E_\Delta := E \cup \{\Delta\}$ and as in Remark 2.1.5 for simplicity of notation the statements on Markov processes will be made only for the state space E. They are valid also for the augmented state space E_Δ.

Let $(p(t))_{t\in\mathbb{R}_+}$ be a semigroup of transition probabilities on $(E, \mathcal{B}(E))$, $x \in E$ and let $\mathbb{P}_x$ be the probability measure on

$$\left(E^{\mathbb{R}_+}, \mathcal{B}(E)^{\mathbb{R}_+}\right)$$

given by 2.1.13 such that the coordinate process $(\lambda_t)_{t\in\mathbb{R}_+}$ is a Markov process with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ starting at x with semigroup of transition probabilities $(p(t))_{t\in\mathbb{R}_+}$. Let$(P(t))_{t\in\mathbb{R}_+}$ be its Markov semigroup (see Definition 2.1.6). The natural filtration of $(\lambda_t)_{t\in\mathbb{R}_+}$ is called $(\mathcal{F}^0_t)_{t\in\mathbb{R}_+}$. We set

$$\mathcal{F}^0 := \sigma\left(\bigcup_{t\in\mathbb{R}_+} \mathcal{F}^0_t\right) \subset \mathcal{B}(E)^{\mathbb{R}_+}.$$

DEFINITION 3.0.1. A family of Markov processes

$$\left((\lambda_t)_{t\in\mathbb{R}_+}, (\mathbb{P}_x)_{x\in E}\right)$$

is called *stochastically continuous* if for any $t \in \mathbb{R}_+$ and any $x \in E$

$$\lim_{s \to t} p(s)(x, \cdot) = p(t)(x, \cdot)$$

weakly (see Definition A.3.74) on E.

One can show that affine processes are semimartingales with characteristics that are of a special affine form. The precise statement is made in Theorem 3.1.12. In order to be able to state this result, we next define semimartingales. For a complete introduction into semimartingales and their characteristics, the reader is referred to [**24**] on which the brief introduction in this thesis is based. In order to be as self contained as possible, all necessary definitions can be found in Appendix A.6.

Definition 3.0.2. Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space. A stochastic process $(\lambda_t)_{t\in\mathbb{R}_+}$ adapted to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ is called *local martingale* if there exists a sequence $(\tau_n)_{n\in\mathbb{N}}$ of stopping times (see Definition A.3.101) such that $\mathbb{P}$- almost surely $\tau_n < \tau_{n+1}$ and $\lim_{n\to\infty} \tau_n \to \infty$ hold true and such that the stopped process

$$\left(\lambda_{\min(t,\tau_n)}\right)_{t\in\mathbb{R}_+}$$

is a martingale (see Definition A.3.91) for any $n \in \mathbb{N}$.

Definition 3.0.3. A continuous function $f : \mathbb{R}_+ \to \mathbb{R}$ is called of *finite variation* if for all $t \geq 0$

$$V_t(f) = \sup\left\{ \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)| \,\middle|\, 0 = t_0 \leq t_1 \leq ... \leq t_n \leq t,\, n \in \mathbb{N} \right\}$$

is finite.
A real-valued stochastic process $(\lambda_t)_{t\in\mathbb{R}_+}$ is said to be of finite variation if all paths are of finite variation.

Definition 3.0.4. Let

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

be a filtered probability space. A real-valued stochastic process $(Y_t)_{t\in\mathbb{R}_+}$ is called *semimartingale* if it can be written $\mathbb{P}$-almost surely as

$$Y_t = Y_0 + M_t + A_t \text{ for all } t \in \mathbb{R}_+,$$

where Y_0 is $\mathscr{F}_0$-measurable, $(M_t)_{t\in\mathbb{R}_+}$ is a local martingale starting at $M_0 = 0$ $\mathbb{P}$-almost surely and $(A_t)_{t\in\mathbb{R}_+}$ is a càdlàg, adapted process of finite variation starting at $A_0 = 0$ $\mathbb{P}$-almost surely. An $\mathbb{R}^d$-valued stochastic process $\left(Y_t^1, ..., Y_t^d\right)_{t\in\mathbb{R}_+}$ is called $d-$*dimensional semimartingale* if all components are real-valued semimartingales.

We recall, that on a set K whose closure is compact the generator of a Feller process on $\mathbb{R}^d$ can be written as the sum of four summands that depend on drift, diffusion matrix, killing rate and a Radon measure, respectively (see Theorem 2.2.10). Semimartingales can similarly be characterized by so-called *semimartingale characteristics* (for a precise definition see Definition A.6.21). For a given truncation function (see Definition A.6.11), this is a triple consisting of generalizations of drift, covariance matrix (see Theorem 2.2.10 and Definition 2.1.23) and the compensator of a jump measure (see Theorem A.6.16). In particular, the semimartingale characteristics lead to the canonical decomposition of a semimartingale (see Theorem A.6.25). This is the decomposition of a semimartingale as the sum of four summands: generalized drift, a continuous local martingale, characterized by its covariance matrices, compensated small jumps, and large jumps.

3.1. Affine Processes

Affine processes have been introduced in 1971 by Kawazu and Watanabe [**28**], and in 2003 were characterized on the canonical state space $\mathbb{R}_+^n \times \mathbb{R}^m$ by Duffie, Filipovic, Schachermayer [**16**] who have also shown that on $\mathbb{R}_+^n \times \mathbb{R}^m$ affine processes are Feller processes. Regarding more general state spaces, in 2013, Cuchiero and Teichmann [**13**] showed regularity and path properties if the state space is a certain non-empty subset of a finite dimensional real vector space.

3.1.1. Affine processes on general state spaces. For the definition of affine processes on general state spaces $E \subset V$ we follow [**13**].We define

$$\mathcal{U} := \left\{ u \in V + iV |\, x \to e^{\langle u,x\rangle} \text{ is bounded on } E \right\},$$

and for any $p \geq 1$

$$\mathcal{U}^p := \left\{ u \in V + iV |\, x \to e^{\langle u,x\rangle} \leq p \text{ on } E \right\}.$$

Furthermore, we assume that E contains $d+1$ elements $x_1, ..., x_{d+1}$ such that for every $j \in \{1, ..., d+1\}$ the set

$$(x_1 - x_j, ..., x_{j-1} - x_j, x_{j+1} - x_j, ..., x_{d+1} - x_j)$$

is linearly independent.

Definition 3.1.1. The Markov process $(\lambda_t)_{t\in\mathbb{R}_+}$ together with the family of probability measures $(\mathbb{P}_x)_{x\in E}$ on the filtered measurable space

$$\left(E^{\mathbb{R}_+}, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}\right)$$

and its semigroup of transition probabilities $(p(t))_{t\in\mathbb{R}_+}$ are called *affine* if for every $(t,u)\in\mathbb{R}_+\times\mathcal{U}$ there exist $\Phi(t,u)\in\mathbb{C}$ and $\Psi(t,u)\in V+iV$ such that for all $x\in E$ and $p\geq 1$ on the subset of $\mathbb{R}_+\times\mathcal{U}^p$ where $\Phi\neq 0$ the map $(t,u)\to\langle\Psi(t,u),x\rangle$ is locally continuous and

$$\int_E e^{\langle u,\xi\rangle}p(t)(x,d\xi)=\Phi(t,u)\,e^{\langle\Psi(t,u),x\rangle}, \tag{3.1.1}$$

and if $\left((\lambda_t)_{t\in\mathbb{R}_+},(\mathbb{P}_x)_{x\in E}\right)$ is stochastically continuous (see Definition 3.0.3).

In Theorem 3.6 in [**13**], it it proved that there exists a càdlàg modification of an affine process:

Theorem 3.1.2. *Let* $\left((\lambda_t)_{t\in\mathbb{R}_+},(\mathbb{P}_x)_{x\in E}\right)$ *be an affine process on*

$$\left(E^{\mathbb{R}_+}, \mathcal{F}^0, (\mathcal{F}^0_t)_{t\in\mathbb{R}_+}\right).$$

Then for each $x\in E$ *there exists a* $\mathbb{P}_x$*-version* $\left(\tilde{\lambda}_t\right)_{t\in\mathbb{R}_+}$ *of* $(\lambda_t)_{t\in\mathbb{R}_+}$ *such that* $\left(\tilde{\lambda}_t\right)_{t\in\mathbb{R}_+}$ *has càdlàg paths and is adapted with respect to the completion* $(\mathcal{F}^x_t)_{t\in\mathbb{R}_+}$ *of* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ *with respect to the probability measure* $\mathbb{P}_x$ *(see Definition A.3.77).*

In the following, when we talk about an affine process, we always mean the version with càdlàg paths.

The affine process

$$\left((\lambda_t)_{t\in\mathbb{R}_+},(\mathbb{P}_x)_{x\in E}\right)$$

may take the value $\Delta\notin\mathbb{R}^d$. Thus, Definition 3.0.4 cannot be applied directly. Therefore, if $V\setminus\bar{E}$ is nonempty, then we identify Δ as one point in $V\setminus\bar{E}$. If $V=\bar{E}$ we extend V to $V\times\mathbb{R}$, set $E_1:=E\times\{0\}$, and choose Δ as one point in $(V\times\mathbb{R})\setminus\overline{E_1}$. We introduce the stopping times

$$T_\Delta(\omega):=\inf\{t\in\mathbb{R}_+|\,\lambda_t(\omega)=\Delta\}$$

and

$$T_{expl}:=\begin{cases}T_\Delta, & \text{if } T_k^{'}<T_\Delta \text{ for all } k\\ \infty, & \text{if } T_k^{'}=T_\Delta \text{ for some } k\end{cases}$$

with

$$T_k' := \inf\left\{ t \in \mathbb{R}_+ | \lim_{s \nearrow t} \|\lambda_t(\omega)\| \geq k \text{ or } \|\lambda_t(\omega)\| \geq k \right\}.$$

Affine processes are semimartingales as was shown in Theorem 5.8 in [**13**]:

THEOREM 3.1.3. *Let* $\left((\lambda_t)_{t\in\mathbb{R}_+}, (\mathbb{P}_x)_{x\in E}\right)$ *be an affine process on*

$$\left(E^{\mathbb{R}_+}, \mathcal{F}^0, \left(\mathcal{F}_t^0\right)_{t\in\mathbb{R}_+}\right)$$

and let τ *be a stopping time such that* $\tau < T_{expl}$. *Then* $\left(\lambda_t 1_{[0,T_\Delta)}\right)_{t\in\mathbb{R}_+}$ *is a semimartingale with state space* $E \cup \{0\}$ *and* $\left(\lambda_{\min(t,\tau)}\right)_{t\in\mathbb{R}_+}$ *is a semimartingale with state space* $E \cup \Delta$. *There is a version* (B, C, ν) *of the characteristics of* $\left(\lambda_{\min(t,\tau)}\right)_{t\in\mathbb{R}_+}$ *relative to a truncation function* χ *(see Definition A.6.11) such that*

$$B_{t,i} = \int_0^{\min(t,\tau)} b_i(\lambda_{s-})ds,$$
$$C_{t,ij} = \int_0^{\min(t,\tau)} c_{ij}(\lambda_{s-})ds,$$
$$\nu(\omega, dt, d\xi) = K\left(\lambda_t(\omega), d\xi\right) 1_{[0,\tau]} dt,$$

with measurable functions $b : E \to V$ *and* $c : E \to S_+(V)$. *Furthermore,* $K(x, d\xi)$ *is a transition kernel (see Definition 2.1.1) from* $(E, \mathcal{B}(E))$ *into* $(V, \mathcal{B}(V))$ *and for all* $x \in E$

$$\int_V \min\left(\|\xi\|^2, 1\right) K(x, d\xi) < \infty,$$
$$K(x, \{0\}) = 0,$$

and $x + supp\left(K(x, \cdot)\right) \subset E \cup \{\Delta\}$.

Furthermore, the differential characteristics of affine processes depend in an affine way on the process itself (see Theorem3.1.5). One can also show that affine processes are regular:

DEFINITION 3.1.4. The affine process $\left((\lambda_t)_{t\in\mathbb{R}_+}, (\mathbb{P}_x)_{x\in E}\right)$ is called *regular* if

$$F(u) := \partial_t^+ \left(\Phi(t, u)\right)\Big|_{t=0}$$

and

$$R(u) := \partial_t^+ \left(\Psi(t, u)\right)\Big|_{t=0}$$

exist for all $(x, u) \in E \times \mathcal{U}$ and are continuous functions on $\mathcal{U}^p$ for any $p \geq 1$.

The following statement was proofed in [**13**] (Theorem 6.4). Among other things, it states that the Fourier-Laplace transform of an affine process is given by the solution of an ordinary differential equation.

THEOREM 3.1.5. *Every affine process in regular. Furthermore, for* $u \in \mathcal{U}$

$$\begin{aligned} F(u) &= \frac{1}{2}\langle au, u\rangle + \langle b, u\rangle - c \\ &\quad + \int_V \left(e^{\langle u,\xi\rangle} - 1 - \langle u, \chi(\xi)\rangle\right) m(d\xi) \\ \langle R(u), x\rangle &= \frac{1}{2}\langle \alpha(x)u, u\rangle + \langle \beta(x), u\rangle - \langle \gamma, x\rangle \\ &\quad + \int_V \left(e^{\langle u,\xi\rangle} - 1 - \langle u, \chi(\xi)\rangle\right) \mu(x, d\xi) \end{aligned} \tag{3.1.2}$$

where $\chi : V \to V$ *is a truncation function (see Definition A.6.11) such that* $\chi(\Delta - x) = 0$ *for all* $x \in E$, $b \in V$, $a \in S(V)$, m *is a signed measure,* $c \in \mathbb{R}$, $\gamma \in V$ *and*

$$\begin{aligned} &x \to \beta(x), \\ &x \to \alpha(x), \\ &x \to \mu(x, d\xi), \end{aligned}$$

are restrictions of linear maps on V *such that*

$$\begin{aligned} b(x) &= b + \beta(x) \\ c(x) &= a + \alpha(x) \\ K(x, d\xi) &= m(d\xi) + \mu(x, d\xi) + (c + \langle\gamma, x\rangle)\delta_{\{\Delta - x\}}(d\xi), \end{aligned}$$

where the expressions on the left hand side are known from Theorem 3.1.3.

Moreover, on the set $\mathcal{Q} := \{(t, u) \in \mathbb{R}_+ \times \mathcal{U} |\, \Phi(s, u) \neq 0$ *for all* $s \in [0, t]\}$

$$\begin{aligned} &\partial_t \Psi(t, u) = R(\Psi(t, u)) && \Psi(0, u) = u \\ &\partial_t \Phi(t, u) = \Phi(t, u)\, F(\Psi(s, u))\, ds && \Phi(0, u) = 1. \end{aligned} \tag{3.1.3}$$

3.1.2. Affine processes on the canonical state space. On the canonical state space $E = \mathbb{R}_+^m \times \mathbb{R}^n$ affine processes were characterized in [**16**]. We mostly follow their notation.

We set $d = m + n$ and $V = \mathbb{R}^d$ and for $x \in E$ we use the convention $x = (y, z)$ with $y \in \mathbb{R}_+^m$ and $z \in \mathbb{R}^n$. The first m indices are collected in the index set $\ell := \{1, ..., m\}$ and the following ones in the index set $\mathscr{J} := \{m + 1, ..., d\}$. We observe that for any $p \geq 1$

$$\begin{aligned}\mathcal{U} &= \left\{u \in V + iV \,|\, x \to e^{\langle u,x\rangle} \text{ is bounded on } E\right\}\\ &= \mathbb{C}^m_- \times i\mathbb{R}^n\\ &= \left\{u \in V + iV \,|\, x \to e^{\langle u,x\rangle} \le p \text{ on } E\right\}\\ &= \mathcal{U}^p.\end{aligned}$$

On the canonical state space the definition of affine processes can be simplified in the following way:

DEFINITION 3.1.6. The Markov process $(\lambda_t)_{t\in\mathbb{R}_+}$ together with the family of probability measures $(\mathbb{P}_x)_{x\in E}$ on the filtered measurable space

$$\left(E^{\mathbb{R}_+}, \mathcal{F}^0, \left(\mathcal{F}^0_t\right)_{t\in\mathbb{R}_+}\right)$$

and its semigroup of transition probabilities $(p(t))_{t\in\mathbb{R}_+}$ is called *affine* if for every $(t,u) \in \mathbb{R}_+ \times i\mathbb{R}^d$ there exist $\Phi(t,u) \in \mathbb{C}$ and $\Psi(t,u) \in \mathbb{C}^d$ such that for all $x \in E$

$$\int_E e^{\langle u,\xi\rangle} p(t)(x,d\xi) = e^{\varphi(t,u)+\langle\Psi(t,u),x\rangle}$$

and if $\left((\lambda_t)_{t\in\mathbb{R}_+}, (\mathbb{P}_x)_{x\in E}\right)$ is stochastically continuous (see Definition 3.0.3).

We will see in Theorem 3.1.8 that an affine process is a Feller process with (see Theorem 2.2.10 and Definition 2.1.23) drift $b + \beta x$, diffusion matrix

$$(a_{kl} + \langle \alpha_{\ell,kl}, y\rangle)_{k,l\in\{1,\dots,d\}},$$

Radon measure

$$N(x,d\xi) = m(d\xi) + y_1\mu_1(d\xi) + \dots + y_m\mu_m(d\xi),$$

and killing rate $c + \langle\gamma, y\rangle$. Since the affine process may not exit the set E, this implies that the parameters

$$(a, \alpha, b, \beta, c, \gamma, m, \mu)$$

must be of a particular form. This is formulated in the definition of admissible parameters in Definition 3.1.7.

In order to introduce admissible parameters, we first introduce more notation. For a d-tuple $\beta = (\beta_1, \dots, \beta_d)$ we define $\beta_I := (\beta_i)_{i\in I}$ and for a $d \times d$ matrix $\alpha = (\alpha_{ij})$ we set $\alpha_{IJ} := (\alpha_{ij})_{i\in I, j\in J}$. Also $\Psi^{\mathcal{Y}} := \Psi_\ell$ and $\Psi^{\mathcal{Z}} := \Psi_{\mathscr{J}}$. We write $\mathbf{1} := (1, \dots, 1)$ for the dimension that makes sense in a given situation. For $i \in \ell$ we define $\ell(i) := \ell \setminus \{i\}$ and for

$i \in \ell$ we set $\mathscr{J}(i) := \mathscr{J} \cup \{i\}$. For $i \in \ell$ we define the $m \times m$-matrix $\mathrm{Id}(i)$ by

$$\mathrm{Id}(i)_{kl} := \begin{cases} 1 & \text{if } i = k = l \\ 0 & \text{else.} \end{cases}$$

The continuous truncation function χ (see also Definition A.6.11) is defined by

$$\chi = (\chi_1, ..., \chi_d) : \mathbb{R}^d \to [-1,1]^d$$

$$\chi_k(\xi) := \left\{ \begin{array}{cc} 0 & \text{if } \xi_k = 0, \\ (1 \wedge |\xi_k|) \frac{\xi_k}{|\xi_k|}, & \text{otherwise.} \end{array} \right. \tag{3.1.4}$$

It is an important tool in dissecting small jumps of which there may be infinitely many from large jumps of which there are only finitely many (see Lemma A.6.5). We can now define:

Definition 3.1.7. (*admissible parameters*) The parameters

$$(a, \alpha, b, \beta, c, \gamma, m, \mu)$$

are called *admissible* if

$$a \in S_+(V) \text{ with } a_{\ell\ell} = 0, \tag{3.1.5}$$

(3.1.6)

$$\alpha = (\alpha_1, ..., \alpha_m) \text{ with } \alpha_i \in S_+(V) \text{ and } a_{i,\ell\ell} = \alpha_{i,ii}\mathrm{Id}(i) \text{ for all } i \in \ell,$$

$$b \in E, \tag{3.1.7}$$

$$\beta \in \mathbb{R}^{d\times d} \text{ such that } \beta_{\ell\mathscr{J}} = 0 \text{ and } \beta_{i\ell(i)} \in \mathbb{R}_+^{m-1} \text{ for all } i \in \ell, \tag{3.1.8}$$

$$c \in \mathbb{R}_+, \tag{3.1.9}$$

$$\gamma \in \mathbb{R}_+^m, \tag{3.1.10}$$

$$m \text{ is a Borel measure on } E \setminus \{0\} \text{ satisfying} \tag{3.1.11}$$

$$M := \int_{E\setminus\{0\}} \left(\langle \chi_\ell(\xi), \mathbf{1} \rangle + \left\| \chi_{\mathscr{J}}(\xi) \right\|^2 \right) m(d\xi) < \infty, \tag{3.1.12}$$

(3.1.13)

$\mu = (\mu_1, ..., \mu_m)$; $(\mu_i)_{i\in 1,..,m}$ are Borel measure on $E \setminus \{0\}$ satisfying

(3.1.14)

$$\mathcal{M}_i := \int_{E\setminus\{0\}} \left(\left\langle \chi_{\ell(i)}(\xi), \mathbf{1} \right\rangle + \left\| \chi_{\mathscr{J}(i)}(\xi) \right\|^2 \right) \mu_i(d\xi) < \infty.$$

The following theorem was proved by Duffie, Filipovic and Schachermayer ([**16**], Theorem 2.7). It shows that affine processes are Feller processes and characterizes them by means of their admissible parameters. Furthermore, it shows that the functions Ψ and Φ are given as the solution of an ordinary differential equation.

THEOREM 3.1.8. *Let* $\left((\lambda_t)_{t\in\mathbb{R}_+}, (\mathbb{P}_x)_{x\in E} \right)$ *be an affine process on*

$$\left(E^{\mathbb{R}_+}, \mathcal{F}^0, \left(\mathcal{F}_t^0\right)_{t\in\mathbb{R}_+} \right).$$

Then $(\lambda_t)_{t\in\mathbb{R}_+}$ *is a Feller process on*

$$\left(E^{\mathbb{R}_+}, \mathcal{F}^0, \left(\mathcal{F}_t^0\right)_{t\in\mathbb{R}_+}, \mathbb{P}_x \right)$$

for any $x \in E$. *Let* A *be the generator of* $(\lambda_t)_{t\in\mathbb{R}_+}$ *(see Definition 2.2.9) Then* $C_c^\infty(E)$ *is a core of* A *(see Definition 1.4.20) and* $C_c^2(E) \subset \mathcal{D}(A)$. *Furthermore, there exist admissible parameters*

$$(a, \alpha, b, \beta, c, \gamma, m, \mu)$$

such that for $f \in C_c^2(E)$

$$\begin{aligned} Af(x) = &\frac{1}{2} \sum_{k,l=1}^{d} (a_{kl} + \langle \alpha_{\ell,kl}, y \rangle) \frac{\partial^2 f(x)}{\partial x_k \partial x_l} \\ &+ \langle b + \beta x, \nabla f(x) \rangle - (c + \langle \gamma, y \rangle) f(x) \\ &+ \int_{E\setminus\{0\}} (f(x+\xi) - f(x) - \langle \nabla_{\mathcal{J}} f(x), \chi_{\mathcal{J}}(\xi) \rangle m(d\xi) \qquad (3.1.15) \\ &+ \sum_{i=1}^{m} \int_{E\setminus\{0\}} (f(x+\xi) - f(x) - \left\langle \nabla_{\mathcal{J}(i)} f(x), \chi_{\mathcal{J}(i)}(\xi) \right\rangle y_i \mu_i(d\xi). \end{aligned}$$

Additionally,

$$\int_E e^{\langle u, \xi \rangle} p(t)(x, d\xi) = e^{\varphi(t,u) + \langle \Psi(t,u), x \rangle}$$

holds for $(t, u) \in \mathbb{R}_+ \times \mathcal{U}$ *and* $\varphi(t, u)$ *and* $\Psi(t, u)$ *are given by*

$$\varphi(t, u) = \int_0^t F(\Psi(s, u))\, ds \qquad (3.1.16)$$

and

$$\partial_t \Psi^{\mathcal{Y}}(t,u) = R^{\mathcal{Y}}\left(\Psi^{\mathcal{Y}}(t,u), e^{t\cdot\beta^{\mathcal{Z}}} w\right)$$

$$\Psi^{\mathcal{Z}}(t,u) = e^{t\cdot\beta^{\mathcal{Z}}} w \tag{3.1.17}$$

with $\Psi(0,u) = u$. *The vector fields* $R^{\mathcal{Y}}$ *and* F *are given by*

$$\begin{aligned} F(u) = &\frac{1}{2}\langle au, u\rangle + \langle b, u\rangle - c \\ &+ \int_{E\setminus\{0\}} e^{\langle u,\xi\rangle} - 1 - \langle u_{\mathcal{J}}, \chi_{\mathcal{J}}(\xi)\rangle\, m(d\xi) \\ R_i^{\mathcal{Y}}(u) = &\frac{1}{2}\langle \alpha_i u, u\rangle + \left\langle \beta_i^{\mathcal{Y}}, u\right\rangle - \gamma_i \\ &+ \int_{E\setminus\{0\}} e^{\langle u,\xi\rangle} - 1 - \left\langle u_{\mathcal{J}(i)}, \chi_{\mathcal{J}(i)}(\xi)\right\rangle \mu_i(d\xi) \quad \textit{for } i \in \{1,...m\} \end{aligned} \tag{3.1.18}$$

with $\beta_i^{\mathcal{Y}} = \left(\beta^T\right)_{i\{1,\dots,d\}} \in \mathbb{R}^d$ *for* $i \in \{1,...,m\}$
and $\beta^{\mathcal{Z}} = \left(\beta^T\right)_{\mathcal{J}\mathcal{J}}$.

On the other hand, for admissible parameters $(a,\alpha,b,\beta,c,\gamma,m,\mu)$ *there exists a unique affine semigroup* $(P(t))_{t\in\mathbb{R}_+}$ *on* $(E,\mathcal{B}(E))$ *with generator given by Equation 3.1.15 such that*

$$\int_E e^{\langle u,\xi\rangle} p(t)(x,d\xi) = e^{\varphi(t,u)+\langle\Psi(t,u),x\rangle}$$

holds for all $(t,u) \in \mathbb{R}_+ \times \mathcal{U}$ *with* $\varphi(t,u)$ *and* $\Psi(t,u)$ *given by Equations 3.1.16 and 3.1.17.*

Remark 3.1.9. Affine processes with the convex cone of symmetric positive semidefinite matrices as state space are Feller processes as well. See [**11**] for details.

Remark 3.1.10. The fact that a regular affine process on E is a Feller process implies in particular by Theorem 2.2.6 that there exists a modification of the affine process, that is càdlàg.

Remark 3.1.11. The differential equation

$$\partial_t \Psi^{\mathcal{Y}}(t,u) = R^{\mathcal{Y}}\left(\Psi^{\mathcal{Y}}(t,u), e^{t\cdot\beta^{\mathcal{Z}}} w\right)$$

is called *generalized Riccati equation*.

The following theorem was proved in [**16**]. It shows that regular affine processes are semimartingales with characteristics that depend in an affine way on the process itself. Furthermore, it shows, that by assuming a semimartingale with certain characteristics one obtains a process which is distributed like an affine process.

THEOREM 3.1.12. *An affine process*

$$\left((\lambda_t)_{t\in\mathbb{R}_+}, (\mathbb{P}_x)_{x\in E}\right) = \left(\left((Y_t)_{t\in\mathbb{R}_+}, (Z_t)_{t\in\mathbb{R}_+}\right), (\mathbb{P}_x)_{x\in E}\right)$$

with admissible parameters

$$(a, \alpha, b, \beta, c, \gamma, m, \mu)$$

is a semimartingale. If $p(t)(x, E) = 1$ for all $(t, x) \in \mathbb{R}_+ \times E$ then for any $x \in E$

$$\left(\lambda_t 1_{\{t<T_\Delta\}}\right)_{t\in\mathbb{R}_+}$$

is a semimartingale on

$$\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}_x\right)$$

and has the characteristics (B, C, ν) associated to truncation function χ from Equation 3.1.4 given by

$$B_t = \int_0^t \left(\tilde{b} + \tilde{\beta}\lambda_s\right) ds, \tag{3.1.19}$$

$$C_t = \int_0^t \left(a + \sum_{i=1}^m \alpha_i Y_s^i\right) ds, \tag{3.1.20}$$

and

$$v(dt, d\xi) = \left(m(d\xi) + \sum_{i=1}^m Y_t^i \mu_i(d\xi)\right) dt \tag{3.1.21}$$

with the definitions

$$E \ni \tilde{b} := b + \int_{E\setminus\{0\}} (\chi_\ell(\xi), 0)\, m(d\xi),$$

$$\mathbb{R}^{d\times d} \ni \tilde{\beta}_{kl} := \begin{cases} \beta_{kl} + (1-\delta_{kl}) \int_{E\setminus\{0\}} \chi_k(\xi)\mu_\ell(d\xi), & \text{if } l \in \ell, \\ \beta_{kl} & \text{if } l \in \mathscr{J}, \text{for } 1 \le k \le d. \end{cases}$$

If $\lambda' = (Y', Z')$ is such that

$$\left(\lambda_t' 1_{\{t<T_\Delta\}}\right)_{t\in\mathbb{R}_+}$$

is an E-valued semimartingale on some filtered probability space

$$\left(\Omega^{'}, \mathcal{F}^{'}, \left(\mathcal{F}^{'}{}_{t}\right)_{t\in\mathbb{R}_+}, \mathbb{P}^{'}\right)$$

with $\mathbb{P}^{'}\left(\lambda_0^{'} = x\right) = 1$ *and its characteristics* $\left(B^{'}, C^{'}, \nu^{'}\right)$ *are given by Formulas 3.1.19, 3.1.20, and 3.1.21 where* $(\lambda_t)_{t\in\mathbb{R}_+}$ *is replaced by* $\left(\lambda^{'}{}_{t}\right)_{t\in\mathbb{R}_+}$ *then*

$$\mathbb{P}^{'} \circ \left(\lambda^{'}\right)^{-1} = \mathbb{P}_x.$$

Corollary 3.1.13. *Let* $f \in B^{\rho}(\mathcal{U})$ *and* $\rho : \mathcal{U} \to \mathbb{R}_+$ *given by*

$$\rho(u) := \|u\|^2 + 1.$$

Assume that for any $t \in \mathbb{R}_+$

$$\sup_{u\in\mathcal{U}} \frac{\rho \circ \Psi(t,u)}{\rho(u)} =: C_t < \infty,$$

where $\Psi(t,u)$ *is given by 3.1.17 and for some* $\delta > 0$ *there is* $C > 0$ *such that for all* $0 \leq t < \delta$

$$C_t < C.$$

Then $(f \circ \Psi(t,\cdot))_{t\in\mathbb{R}_+}$ *is a generalized Feller semigroup on* $\mathscr{B}^{\rho}(\mathcal{U})$.

Proof. $(\Psi(t,\cdot))_{t\in\mathbb{R}_+}$ fulfills the conditions of Proposition 2.3.54.

Proposition 7.4 (ii) in **[16]** states that$\Psi(s,\cdot) \circ \Psi(t,\cdot) = \Psi(s+t,\cdot)$ and from **[16]** Proposition 6.1 and Proposition 6.4 it follows that

$$(t,u) \to \Psi(t,u)$$

is continuous on $\mathbb{R}_+ \times \mathcal{U}$ which implies that Property (iii) and (iv) in Proposition 2.3.54 hold true. □

Corollary 3.1.14. *Let* $n = 0$ *and let* $f \in \mathscr{B}^{\rho}(\mathcal{U})$ *and* $\rho : \mathcal{U} \to \mathbb{R}_+$ *given by*

$$\rho(u) := \|u\|^2 + 1.$$

Then $(f \circ \Psi(t,\cdot))_{t\in\mathbb{R}_+}$ *is a generalized Feller semigroup on* $\mathscr{B}^{\rho}(\mathcal{U})$.

$C_c^1(\mathcal{U}) \subset \mathcal{D}(A)$ *and for* $f \in C_c^1(\mathcal{U})$ *the generator is given by*

$$Af(u) = (Df)(u) \cdot R(u).$$

Let $\mu \in \mathcal{M}^{\rho}(\mathcal{U})$ *be given by*

$$\mu(B) = \int_B g_\mu(u) d\lambda(u) \textit{ for any } B \in \mathcal{B}(\mathcal{U})$$

for the Lebesgue measure λ *and* $g_\mu \in C^1(\mathcal{U})$ *and let*

$$\int_{\mathcal{U}} \left(\|u\|^2+1\right)\left(\frac{d}{dx_1}g_\mu + ...\frac{d}{dx_n}g_\mu\right)(u)d\lambda(u)$$
$$\leq C\int_{\mathcal{U}} \left(\|u\|^2+1\right) g_\mu(u)d\lambda(u)$$

for some $C>0$. *Then,* $\mu \in \mathcal{D}(A')$ *and*

$$A'(\mu)(B) = -\int_B div\left(R\cdot g_\mu\right)(u)d\lambda(u) \text{ for any } B\in\mathcal{B}(\mathcal{U}).$$

$(Q(t))_{t\in\mathbb{R}_+}$ *defined on* $\overline{\mathcal{D}(A')}\subset\mathcal{M}^\rho(\mathcal{U})$ *as*

$$Q(t)(\mu) := \mu\circ\psi_t^{-1}$$

is a strongly continuous semigroup. Define

$$\mathcal{D}(A^\dagger) := \left\{\mu\in\mathcal{D}(A') :\ A'\mu\in\overline{\mathcal{D}(A')}\right\}.$$

If $div\left(R\cdot g_\mu\right)\in C^1(\mathcal{U})$ *and*

$$\int_{\mathcal{U}} \left(\|u\|^2+1\right)\left(\frac{d}{dx_1}\left(R\cdot g_\mu\right) + ...\frac{d}{dx_n}\left(R\cdot g_\mu\right)\right)(u)d\lambda(u)$$
$$\leq C'\int_{\mathcal{U}} \left(\|u\|^2+1\right)\left(R\cdot g_\mu\right)(u)d\lambda(u)$$

for some $C'>0$*, then* $\mu\in\mathcal{D}(A^\dagger)$ *and the generator* $A^\dagger$ *of* $(Q(t))_{t\in\mathbb{R}_+}$ *is given by the restriction of* A' *to* $\mathcal{D}(A^\dagger)$.

PROOF. By **[16]**, Inequality 6.16 for all $t\in\mathbb{R}_+$

$$\|\Psi(t,u)\|^2 \leq \|u\|^2,$$

which implies that the conditions of Corollary 3.1.13 are satisfied. The results for the generator follow from Proposition 2.3.58 and Corollary 2.3.57 and Proposition 2.3.59 yields the results for the adjoint semigroup and its generator. □

3.2. Polynomial Processes

Polynomial processes were introduced in **[12]**. This section mainly follows their presentation.

Throughout this section, let E be a closed subspace of $V=\mathbb{R}^d$ for some $d\in\mathbb{N}$ and let the filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ be right continuous.

Definition 3.2.1. Let

$$\mathcal{P}_m := \left\{ E \ni x \to \sum_{\boldsymbol{j}, |\boldsymbol{j}| \le m} \alpha_{\boldsymbol{j}} x^{\boldsymbol{j}},\, \Delta \to 0 \,\middle|\, \alpha_{\boldsymbol{j}} \in \mathbb{R}^d \right\}$$

be the finite dimensional vector space of polynomials on $E \cup \{\Delta\}$ of degree m. Here, $\boldsymbol{j} = (j_1, ..., j_d) \in \mathbb{N}_0^d$, $|\boldsymbol{j}| = j_1 + ... + j_d$ and $x^{\boldsymbol{j}} := x_1^{j_1} \cdot ... \cdot x_d^{j_d}$ for $x = (x_1, ..., x_d) \in E$.

Definition 3.2.2. Let $m \in \mathbb{N}$. If for any $k \in \{0, ..., m\}$ and any $t \in \mathbb{R}_+$

$$P_t(\mathcal{P}_k) \subset \mathcal{P}_k$$

holds true, and if for any $x \in E$ and any $f \in \mathcal{P}_m$

$$\lim_{t \searrow 0} P_t f(x) \to f(x),$$

then $(\lambda_t)_{t \in \mathbb{R}_+}$ is called $m-$ *polynomial process*. If $(\lambda_t)_{t \in \mathbb{R}_+}$ is an m-polynomial process for any $m \in \mathbb{N}$, then it is called *polynomial process*.

We recall the definition of the extended generator (Definition 2.1.20). Using semigroup theory, the following theorem is proved in [**12**]:

Theorem 3.2.3. *The following three statements are equivalent:*
(i) $(\lambda_t)_{t \in \mathbb{R}_+}$ *is an* m*-polynomial process for some* $m \in \mathbb{N}$.
(ii) For every $k \in \{0, 1, ...m\}$ *there is a bounded linear map* A_k *on* $\mathcal{P}_k$ *such that for any* $t \in \mathbb{R}_+$

$$P(t)|_{\mathcal{P}_k} = e^{tA_k}.$$

(iii) For any $x \in E$, $t \in \mathbb{R}_+$ *and* $f \in \mathcal{P}_m$

$$P_t \,|f|\, (x) < \infty,$$

$\mathcal{P}_m$ *is in the domain of the extended generator* $\mathcal{G}$, *and*

$$M_t^f := f(\lambda_t) - f(x) - \int_0^t \mathcal{G} f(\lambda_s) ds$$

is a martingale with respect to $(\mathcal{F}_t)_{t \in \mathbb{R}_+}$ *and the probability measure* $\mathbb{P}_x$. *Furthermore, for any* $k \in \{0, ..., m\}$

$$\mathcal{G}(\mathcal{P}_k) \subset \mathcal{P}_k.$$

REMARK 3.2.4. For an m-polynomial process $(\lambda_t)_{t\in\mathbb{R}_+}$, by Theorem 3.2.3 and Proposition 2.3.94 for any $f \in \mathcal{P}_m$ there is a version of the stochastic process $\left(M_t^f\right)_{t\in\mathbb{R}_+}$ that has càdlàg paths (see Definition 2.2.5) and is a martingale with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ and the probability measure $\mathbb{P}_x$. Since $\mathcal{P}_m$ is finite dimensional this implies that for any $t \in \mathbb{R}_+$ $\mathbb{P}_x$-almost surely

$$M_t^f = \lim_{s\searrow t,\, s\in\mathbb{Q}} M_s^f$$

holds true for any $f \in \mathcal{P}_m$. Hence, for any $t \in \mathbb{R}_+$ $\mathbb{P}_x$-almost surely

$$f(\lambda_t) = \lim_{s\searrow t,\, s\in\mathbb{Q}} f(\lambda_s) \tag{3.2.1}$$

for any $f \in \mathcal{P}_m$. Since $\mathcal{P}_m$ separates points, also $\mathbb{P}_x$-almost surely for any $t \in \mathbb{R}_+$

$$\lambda_t = \lim_{s\searrow t,\, s\in\mathbb{Q}} \lambda_s$$

(if not, then on some set with non-zero probability there would be some $i \in \{1, ..., d\}$and $t \in \mathbb{R}_+$ such that

$$\lambda_t^i \neq \limsup_{s\searrow t,\, s\in\mathbb{Q}} \lambda_s^i$$

or

$$\lambda_t^i \neq \liminf_{s\searrow t,\, s\in\mathbb{Q}} \lambda_s^i$$

where λ_t^i is the i-th component of λ_t . But then some point-separating $g \in \mathcal{P}_m$ would yield a contradiction to Equation 3.2.1). One obtains such a result also for the existence of left limits. Therefore, an m-polynomial process $(\lambda_t)_{t\in\mathbb{R}_+}$ has a càdlàg modification and in the following when talking about m-polynomial processes we will always mean one whose paths are càdlàg.

Writing

$$\mu(x, d\xi) = x_1\mu_1(d\xi) + ... + x_d\mu_d(d\xi)$$

from Theorem 3.1.5 one can show as in Example 3.1 in [**12**]:

PROPOSITION 3.2.5. *On a state space $E \subset \mathbb{R}^d$, $d \geq 2$ that contains $d+1$ elements $x_1, ..., x_{d+1}$ such that for every $j \in \{1, ..., d+1\}$ the set*

$$(x_1 - x_j, ..., x_{j-1} - x_j, x_{j+1} - x_j, ..., x_{d+1} - x_j)$$

is linearly independent an affine process $(\lambda_t)_{t\in\mathbb{R}_+}$ is r -polynomial if $\gamma = 0$,

$$\int_{\|\xi\|>1} \|\xi\|^r \, m\,(d\xi) < \infty,$$

and for any $i \in \{1, ..., m\}$

$$\int_{\|\xi\|>1} \|\xi\|^r \mu_i(d\xi) < \infty.$$

Proof. From Theorem 3.1.5 it follows that there is $C > 0$ such that

$$\int_{\mathbb{R}^d} \|\xi\|^r K(\lambda_t, d\xi) \leq C\left(1 + \left\|\lambda_t 1_{\{t<T_\Delta\}}\right\|^r\right).$$

The Proposition follows then from Theorem 2.15 in [**12**]. □

Lemma 3.2.6. *If* $(\lambda_t)_{t\in\mathbb{R}_+}$ *is an* m*-polynomial process and* $\rho \in \mathcal{P}_k$ $k \in \{0, ..., m\}$ *then there is a bounded linear map* A_k *on* $\mathcal{P}_k$ *such that for all* $x \in E$ *and* $t \in \mathbb{R}_+$

$$P(t)\rho(x) = \mathbb{E}_x[\rho(\lambda_t)] = \left(e^{tA_m}\rho\right)(x) \leq e^{t\|A_m\|}\rho(x)$$

holds true and $\mathbb{E}_x[\rho(\lambda_t)] < \infty$ *for all* $t \in \mathbb{R}_+$ *and for all* $x \in E$.

Proof. Follows directly from Theorem 3.2.3 (ii). □

Proposition 3.2.7. *Let* $(\lambda_t)_{t\in\mathbb{R}_+}$ *be a polynomial process and for some* $m \in \mathbb{N}$ *let* $\rho \in \mathcal{P}_m$ *be an admissible weight function on* E. *For any* $f \in C_b(E)$ *and any* $t \in \mathbb{R}_+$ *let* $P(t)f|_{K_R}$ *be continuous for any* $R > 0$. *Then* $(\lambda_t)_{t\in\mathbb{R}_+}$ *is a generalized Feller process on* (E, ρ).

Proof. We have to show that $(\lambda_t)_{t\in\mathbb{R}_+}$ is the stochastic process in Theorem 2.3.65. By definition of the Markov process $(\lambda_t)_{t\in\mathbb{R}_+}$ for any $t \geq s \geq 0$ and any measurable map $f : E \to \mathbb{R}_+$

$$\mathbb{E}_{\mathbb{P}_\nu}[f(\lambda_t)|\mathcal{F}_s] = P(t-s) f(\lambda_s) \tag{3.2.2}$$

holds true $\mathbb{P}_\nu$ -almost surely and

$$\mathbb{P}_\nu \circ \lambda_0^{-1} = \nu.$$

By Lemma 3.2.6

$$\mathbb{E}_{\mathbb{P}_\nu}[f(\lambda_t)|\mathcal{F}_s] = P(t-s) f(\lambda_s)$$

holds true $\mathbb{P}_\nu$ -almost surely for all $f \in \mathscr{B}^\rho(E)$ as well. In order to show that $(\lambda_t)_{t\in\mathbb{R}_+}$ is a generalized Feller process we still have to prove that $(P(t))_{t\in\mathbb{R}_+}$ is a generalized Feller semigroup. We fix some $t \in \mathbb{R}_+$ and first show that $f \in \mathscr{B}^\rho(E)$ implies $P(t) f \in \mathscr{B}^\rho(E)$. By Lemma 3.2.6 for any

$$f \in \widetilde{B^\rho}(E) := \left\{f : E \to \mathbb{R} : \sup_{x\in E} \rho(x)^{-1} \|f(x)\| < \infty, f \text{ measurable}\right\}$$

the map

$$x \to \int_E f(y)p(t)(x, dy)$$

is well defined and

$$\begin{aligned} P(t) f(x) &= \int_E f(y)p(t)(x, dy) \\ &\leq e^{t\|A_m\|} \|f\|_\rho \rho(x). \end{aligned}$$

Hence, $P(t)$ is a linear bounded map from $\widetilde{B^\rho}(E)$ to $\widetilde{B^\rho}(E)$. In order to show $P(t) f \in \mathscr{B}^\rho(E)$ for any $f \in \mathscr{B}^\rho(E)$, by continuity of $P(t)$ with respect to $\|\cdot\|_\rho$ and density of $C_b(E)$ in $\mathscr{B}^\rho(E)$ it is sufficient to show that $P(t) f \in \mathscr{B}^\rho(E)$ holds true for any $f \in C_b(E)$. By Theorem 2.3.42 this is the case since $P(t)f$ is clearly bounded and by assumption $P(t)f|_{K_R}$ is continuous for any $R > 0$.

Regarding the properties of generalized Feller semigroups in Definition 2.3.49, the properties **P1** and **P2** and positivity **(P5)** are clearly satisfied for $(P(t))_{t\in\mathbb{R}_+}$. **P4** holds true due to Lemma 3.2.6. Finally, **P3** is fulfilled since the paths of $(\lambda_t)_{t\in\mathbb{R}_+}$ are càdlàg for all $f \in \mathscr{B}^\rho(E)$ hence for all $x \in E$ we obtain by dominated convergence

$$\lim_{t\searrow 0} P(t)f(x) = \lim_{t\searrow 0} \mathbb{E}_x [f(\lambda_t)] = f(x).$$

Thus $(P(t))_{t\in\mathbb{R}_+}$ is a generalized Feller semigroup and $(\lambda_t)_{t\in\mathbb{R}_+}$ is a generalized Feller process on (E, ρ). □

COROLLARY 3.2.8. *On a state space $E \subset \mathbb{R}^d$, $d \geq 2$ that contains $d + 1$ elements $x_1, ..., x_{d+1}$ such that for every $j \in \{1, ..., d + 1\}$ let $(\lambda_t)_{t\in\mathbb{R}_+}$ be an affine process with such that $\gamma = 0$,*

$$\int_{\|\xi\|>1} \|\xi\|^r m(d\xi) < \infty,$$

and for any $i \in \{1, ..., m\}$

$$\int_{\|\xi\|>1} \|\xi\|^r \mu_i(d\xi) < \infty.$$

Let $\rho \in \mathcal{P}_r$ be an admissible weight function on E. If for any $f \in C_b(E)$ and any $t \in \mathbb{R}_+$ $P(t)f|_{K_R}$ is continuous for any $R > 0$, then $(\lambda_t)_{t\in\mathbb{R}_+}$ is a generalized Feller process on (E, ρ).

PROOF. Combine Proposition 3.2.7 and Proposition 3.2.5. □

3.3. Stochastic Representation of ODEs

We have seen in Theorem 3.1.8 and in Theorem 3.1.5 that for an affine process

$$\left((\lambda_t)_{t\in\mathbb{R}_+}, (\mathbb{P}_x)_{x\in E}\right)$$

with state space E the Fourier-Laplace transform is given by a solution of the ordinary differential equation

$$\partial_t \Psi^{\mathcal{Y}}(t,u) = R^{\mathcal{Y}}\left(\Psi^{\mathcal{Y}}(t,u), e^{t\cdot\beta^{\mathcal{Z}}} w\right), \qquad \Psi(0,u) = u,$$
$$\Psi^{\mathcal{Z}}(t,u) = e^{t\cdot\beta^{\mathcal{Z}}} w.$$

In the following we turn this idea around and ask for which ordinary differential equations we obtain a stochastic representation via affine processes.

Definition 3.3.1. Let
(i) $E \subset \mathbb{R}^d$, $d \in \mathbb{N}$ contain $d+1$ elements $x_1, ..., x_{d+1}$ such that for every $j \in \{1, ..., d+1\}$ the set

$$(x_1 - x_j, ..., x_{j-1} - x_j, x_{j+1} - x_j, ..., x_{d+1} - x_j)$$

is linearly independent,
(ii) μ be a signed d-dimensional vector valued measure on E such that for all $i \in \{1, ..., d\}$

$$\int_E \min\left(\|(x,y)\|, 1\right)\left(\mu_i^+(dx,dy) + \mu_i^-(dx,dy)\right) < \infty,$$

(iii) λ be the affine process on E with respect to the characteristics $(B, 0, \nu)$ with truncation function χ and with

$$K(x, d\xi) = x_1\mu_1(d\xi) + ... + x_d\mu_d(d\xi)$$

and

$$b(x) = \int_V \chi(\xi)\left(x_1\mu_1(d\xi) + ... + x_d\mu_d(d\xi)\right)$$

as in Theorem 3.1.3,
(iv) $\hat{\mathcal{U}} \subset \mathbb{R}_+ \times \mathbb{C}^d$ be such that for $(t,u) \subset \hat{\mathcal{U}}$ for any $x \in E$

$$\mathbb{E}_x\left[e^{\langle u, \lambda_t\rangle}\right] = \Phi(t,u)\, e^{\langle \Psi(t,u), x\rangle},$$

and such that on $\hat{\mathcal{U}}$

$$\partial_t \Psi(t,u) = R(\Psi(t,u)) \qquad \Psi(0,u) = u$$
$$\partial_t \Phi(t,u) = \Phi(t,u)\, F(\Psi(s,u))\, ds \qquad \Phi(0,u) = 1,$$

where Φ, Ψ, R, F are given by Equation 3.1.1 and Equation 3.1.2. Then we call $\left(E, \hat{\mathcal{U}}, \lambda, \mu\right)$ *admissible setting*.

THEOREM 3.3.2. *Let $E_1 \subset \mathbb{R}^d$ and let ν_+^{re} be an $\mathbb{R}_+^d$-valued σ-finite measure on E_1 and let ν_-^{re} , ν_+^{im} and ν_-^{im} be $\mathbb{R}_+^d$-valued finite measures on E_1. Let*

$$\mu := \nu_+^{re} - \nu_-^{re} + i\nu_+^{im} - i\nu_-^{im}.$$

and

$$\nu := \nu_+^{re} + \nu_-^{re} + \nu_+^{im} + \nu_-^{im}.$$

On $E := E_1 \times \mathbb{Z} \times \mathbb{Z}$ define the measures

$$\begin{aligned}\tilde{\nu}_+^{re} &:= \nu_+^{re} \circ \left(j_+^{re}\right)^{-1},\\ \tilde{\nu}_-^{re} &:= \nu_-^{re} \circ \left(j_-^{re}\right)^{-1},\\ \tilde{\nu}_+^{im} &:= \nu_+^{im} \circ \left(j_+^{im}\right)^{-1},\\ \tilde{\nu}_-^{im} &:= \nu_-^{im} \circ \left(j_-^{im}\right)^{-1},\end{aligned}$$

and

$$\tilde{\nu} := \tilde{\nu}_+^{re} + \tilde{\nu}_-^{re} + \tilde{\nu}_+^{im} + \tilde{\nu}_-^{im},$$

via the maps

$$\begin{aligned}j_+^{re} &: E_1 \ni x \to (x, 0, 0) \in E,\\ j_-^{re} &: E_1 \ni x \to (x, 1, 0) \in E,\\ j_+^{im} &: E_1 \ni x \to (x, 0, 1) \in E,\\ j_-^{im} &: E_1 \ni x \to (x, 1, 1) \in E.\end{aligned}$$

Assume that $\left(E, \hat{\mathcal{U}}, \tilde{N}, (\tilde{\nu}, 0, 0)\right)$ is an admissible setting. Let

$$\left(0, \log f, i\pi, \frac{i}{2}\pi\right) \in \hat{\mathcal{U}}.$$

If $e_h \in E_1$ for some $h \in \{1, ..., d\}$, then on $\exp\left(\hat{\mathcal{U}}\right)$ the ordinary differential equation

$$\begin{aligned}\partial_t u_h(t) &= u_h(t) \int_{\mathbb{R}^d} \left(u(t)^{\boldsymbol{n}} - 1\right) \mu_h(d\boldsymbol{n})\\ u(0) &= f,\end{aligned} \tag{3.3.1}$$

with $u^{\boldsymbol{n}} = u_1^{n_1} \cdot ... \cdot u_d^{n_d}$ permits the stochastic representation

$$u_h(t) = \mathbb{E}_{(e_h,0,0)}\left[e^{\langle \log(f), N_t\rangle + \langle i\pi\mathbf{1}, Z_{1,t}\rangle + \frac{i}{2}\pi\langle \mathbf{1}, Z_{2,t}\rangle}\right]. \tag{3.3.2}$$

PROOF. By Theorem 3.1.5 for the affine process $\tilde{N} = (N, Z_1, Z_2)$

$$\begin{aligned}&\mathbb{E}_{(e_h,0,0)}\left[e^{\langle \log(f), N_t\rangle + i\pi, Z_{1,t} + \frac{i}{2}\pi Z_{2,t}}\right]\\&= \tilde{\Phi}\left(t, (\log(f), i\pi, \frac{i}{2}\pi)\right) e^{\left\langle \tilde{\Psi}\left(t,(\log(f), i\pi, \frac{i}{2}\pi)\right), (e_h,0,0)\right\rangle}\end{aligned}$$

where $\tilde{\Phi}$ and $\tilde{\Psi} = \left(\tilde{\Psi}^{\mathcal{Y}}, \tilde{\Psi}^{\mathcal{Z}}\right)$, are given by Equation 3.1.3 and $\tilde{\Psi}^{\mathcal{Y}} := \left(\tilde{\Psi}_1, ..., \tilde{\Psi}_d\right)$ and $\tilde{\Psi}^{\mathcal{Z}} := \left(\tilde{\Psi}_{d+1}, \tilde{\Psi}_{d+2}\right)$. This yields $\tilde{\Phi} \equiv 1$ and

$$\tilde{\Psi}^{\mathcal{Z}}\left(t, (\log(f), i\pi, \frac{i}{2}\pi)\right) = \left(i\pi, \frac{i}{2}\pi\right).$$

Moreover, with $\tilde{\nu} = (\tilde{\nu}_1, ..., \tilde{\nu}_d, 0, 0)$, $\mu = (\mu_1, ..., \mu_d)$ and $R = (R_1, .., R_{d+2})$ for any $h \in \{1, ..., d\}$

$$\begin{aligned}&\partial_t \tilde{\Psi}_h\left(t, (\log(f), i\pi, \frac{i}{2}\pi)\right)\\&= R_h\ \left(\tilde{\Psi}^{\mathcal{Y}}\left(t, \left(\log(f), i\pi, \frac{i}{2}\pi\right), i\pi, \frac{i}{2}\pi\right)\right)\\&= \int_{\mathbb{R}^{d+2}} \left(e^{\left\langle \tilde{\Psi}^{\mathcal{Y}}, n\right\rangle + \langle i\pi, z_1\rangle + \left\langle \frac{i}{2}\pi, z_2\right\rangle} - 1\right) \tilde{\nu}_h(dn, dz_1, dz_2)\\&= \int_{\mathbb{R}^d} \left(e^{\left\langle \tilde{\Psi}^{\mathcal{Y}}, n\right\rangle} - 1\right) \mu_h(dn).\end{aligned}$$

We substitute,

$$u_h(t) := \exp\left(\tilde{\Psi}_h\left(t, (\log(f), 1, i\pi)\right)\right).$$

and conclude. □

REMARK 3.3.3. The vector field on the right hand side in Equation 3.3.1 is in general not locally Lipschitz continuous. Therefore, standard theory involving the Picard-Lindelöf theorem can in general not be applied. Furthermore, the right hand side in Equation 3.3.2 can easily be simulated numerically.

COROLLARY 3.3.4. *On the state space $E = \mathbb{Z}^d \times \mathbb{R} \times \mathbb{R}$ with*

$$\left(\tilde{N}\right)_{t\in\mathbb{R}_+} = (N_t, Z_{1,t}, 0)_{t\in\mathbb{R}_+}$$

let $\left(E,\hat{\mathcal{U}},\tilde{N},(\mu,0,0)\right)$ *be an admissible setting, where with multi index* $\mathbf{k}=(k_1,...,k_d)$ *for any* $h\in\{1,...,d\}$

$$\sum_{|\mathbf{k}|=0}^{\infty}\left|a_{\mathbf{k}}^{h}\right|<\infty,$$

and

$$\mu_h\left(dn,dz_1,dz_2\right)=\sum_{|\mathbf{k}|=0}^{\infty}\left|a_{\mathbf{k}}^{h}\right|\delta_{\left\{\mathbf{k}-e_h,1_{\left\{a_{\mathbf{k}}^{h}<0\right\}},0\right\}}\left(dn,dz_1,dz_2\right).$$

Let

$$\left(0,\log f,i\pi,\frac{i}{2}\pi\right)\in\hat{\mathcal{U}}.$$

Then for any $h\in\{1,...,d\}$ *on* $\exp\left(\hat{\mathcal{U}}\right)$

$$u_h(t)=\mathbb{E}_{(e_h,0,0)}\left[\prod_{j=1}^{d}g_j^{N_{j,t}}e^{i\pi Z_{1,t}}\right]$$

is a stochastic representation of

$$\partial_t u_h(t)=\left(\sum_{|\mathbf{k}|=0}^{\infty}a_{\mathbf{k}}^{h}u^{\mathbf{k}}(t)-1\right), \tag{3.3.3}$$

$$u_h(0)=g_h. \tag{3.3.4}$$

with $u^{\mathbf{k}}=u_1^{k_1}\cdot...\cdot u_d^{k_d}$.

PROOF. Choose

$$\nu_{+}^{\text{re}}\left(dn\right)=\sum_{|\mathbf{k}|=0}^{\infty}a_{\mathbf{k}}^{h}1_{\left\{a_{\mathbf{k}}^{h}>0\right\}}\delta_{\{\mathbf{k}-e_h\}}\left(dn\right)$$

and

$$\nu_{-}^{\text{re}}\left(dn\right)=\sum_{|\mathbf{k}|=0}^{\infty}a_{\mathbf{k}}^{h}1_{\left\{a_{\mathbf{k}}^{h}<0\right\}}\delta_{\{\mathbf{k}-e_h\}}\left(dn\right)$$

in Theorem 3.3.2. □

APPENDIX A

Appendix

A.1. Topology

DEFINITION A.1.1. Let X be a set and τ a set of subsets of X. Then τ is called *topology* if
(i) $\emptyset, X \in \tau$,
(ii) an intersection of finitely many elements of τ is in τ,
(iii) any union of elements of τ is in τ.

A *topological space* is a pair (X, τ). For $x \in X$ a set $U_x \subset X$ that contains x is called *neighborhood* of x if there is $O \in \tau$ such that $x \in O \subset U_x$.

DEFINITION A.1.2. A *base* $\mathfrak{B} \subset \tau$ of a topological space (X, τ) is a family of subsets of X such that any element in τ can be written as the union of some elements in $\mathfrak{B}$. A *local base* of $x \in X$ is a family of neighborhoods $\mathfrak{U}_x$ of x such that for any neighborhood V of x there is some $U \in \mathfrak{U}_x$ such that $U \subset V$.

EXAMPLE A.1.3. If X is a vector space with norm $\|\cdot\|$ (or a metric space with metric d) then define for $r > 0$ the open balls

$$B_r(x) := \{y \in X : \|y - x\| < r\} \;\; (B_r(x) := \{y \in X : d(x, y) < r\})$$

and

$$\mathcal{O} : = \left\{O \subset X : \text{for any } x \in O \text{ there is } r(x) > 0 \text{ such that } B_{r(x)}(x) \subset O\right\}$$

Then$\mathcal{O}$ is a topology.

DEFINITION A.1.4. A topological space (X, τ) is called *metrizable* if there is a metric d on X such that the topology generated by d (see Example A.1.3) coincides with τ.

The proof of the following Lemma roughly follows [**40**], Chapter 8.

LEMMA A.1.5. *In a Hausdorff topological space compact sets are closed.*

Proof. Let X be a Hausdorff topological space and $K \subset X$ be compact. Without loss of generality assume $X \setminus K \neq \emptyset$. We have to show that $X \setminus K$ is open. Since X is Hausdorff for any $x \in X \setminus K$ and any $y \in K$ there exist disjoint neighborhoods $U_{x,y} \ni x$ and $U_{y,x} \ni y$. Moreover,

$$\bigcup_{y \in K} U_{y,x}$$

is an open cover of the compact set K hence a finite number of neighborhoods suffices to cover K. Calling these points whose neighborhoods suffice y_i for $i \in \{1, ..., n\}$ we obtain that the sets

$$\bigcup_{i \in \{1,...,n\}} U_{y_i,x} \supset K$$

and

$$\bigcap_{i \in \{1,...,n\}} U_{x,y_i} \ni x$$

are open and disjoint. Therefore

$$\bigcap_{i \in \{1,...,n\}} U_{x,y_i}$$

is an open neighborhood of x in $X \setminus K$. As $x \in X \setminus K$ was arbitrary such a neighborhood exists for all $x \in X \setminus K$ and it follows that $X \setminus K$ is open. $\square$

Definition A.1.6. Let I be an index set and $(\Omega_i)_{i \in I}$ be a family of sets. For $J \subset J' \subset I$ the map

$$\Pi_J^{J'} : \underset{i \in J'}{\times} \Omega_i \to \underset{i \in J}{\times} \Omega_i$$
$$\omega' \to \omega'|_J$$

is called *projection*. For $i \in I$ the map $\Pi_{\{i\}}^{J'}$ is written as $\Pi_i^{J'}$ and for $J' = I$ it is called *coordinate map* and simply written as Π_i.

Definition A.1.7. (product topology)
Let $(X_i, \tau_i)_{i \in I}$ be topological spaces,

$$X := \prod_{i \in I} X_i,$$

and let $(\Pi_i)_{i \in I}$ be coordinate maps from Definition A.1.6. Then the basis

$$\mathcal{B} := \left\{ \bigcap_{j \in J} \Pi_j^{-1}(O_j) \subset X : O_j \in \tau_j, \, J \subset I \text{ finite} \right\}$$

defines the *product topology* τ on X. By definition it is the coarsest topology (the one with least open sets) such that all coordinate maps are continuous.

LEMMA A.1.8. *If* (X, τ) *is a topological space and* $U \subset X$ *then* (U, τ_U) *is a topological space as well where* τ_U *is defined as*

$$\tau_U := \{O \cap U : O \in \tau\}.$$

τ_U *is called subspace topology.*

PROOF. This is shown by a simple verification of the three properties of a topology. □

LEMMA A.1.9. *Let* (X, τ) *be a topological space and let* $S \subset X$ *be equipped with the subspace topology* τ_S. *Then a set* $K \subset S$ *is compact in*(S, τ_S) *if and only in it is compact in* (X, τ).

PROOF. Let $K \subset S$ be compact in (X, τ) and let

$$\bigcup_{i \in I} O_i^S$$

be a cover of K of sets that are open in (X, τ_S). Then for each O_i^S there is a set $O_i \in \tau$ such that $O_i^S = O_i \cap S$ and by assumption there is a finite cover

$$K \subset \bigcup_{j \in \{1,\dots,n\}} O_{i_j}$$

of sets open in (X, τ) . Hence,

$$K \subset \bigcup_{j \in \{1,\dots,n\}} O_{i_j}^S$$

is a finite cover of sets open in τ_S and K is compact in (X, τ_S).

Let $K \subset S$ be compact in (X, τ_S) . Then for any cover

$$\bigcup_{i \in I} O_i$$

of sets open in (X, τ) the cover

$$\bigcup_{i \in I} O_i \cap S$$

of sets open in (X, τ_S) has a finite subcover, say

$$K \subset \bigcup_{j \in \{1,\dots,n\}} O_{i_j} \cap S.$$

Hence

$$K \subset \bigcup_{j \in \{1,\dots,n\}} O_{i_j}$$

is a finite cover of K of sets open in (X, τ) and K is compact in (X, τ).
$\square$

Proposition A.1.10. *([**7**], Chapter IX, §1, Proposition 3 and Corollary)*

A topological space X is completely regular if and only if it is homeomorphic to a subspace of a compact Hausdorff space K. This means, there exists a map $i : X \to K$ called embedding that is continuous and that possesses an inverse on $i(X)$ that is continuous when we equip $i(X) \subset K$ with the subspace topology form Lemma A.1.8 .

Proposition A.1.11. *A compact Hausdorff space is normal.*

Proof. Let X be a compact Hausdorff space and $A, B \subset X$ be closed, disjoint sets. Fix some $x \in A$. Since X is Hausdorff for any $y \in B$ there exist disjoint neighborhoods $U_{x,y}$ of x and $U_{y,x}$ of y.

$$\bigcup_{y \in B} U_{y,x} \supset B$$

is an open cover of B hence by compactness of B finitely many neighborhoods suffice to cover B. Thus, there are $y_i \in B$, $i \in \{1, \dots, n\}$such that:

$$V_x := \bigcup_{i \in \{1,\dots n\}} U_{y_i,x} \supset B.$$

The intersection

$$\tilde{U}_x := \bigcap_{i \in \{1,\dots n\}} U_{x,y_i}$$

is a neighborhood of x and by construction V_x and $\tilde{U}_x$ are disjoint.

$$\bigcup_{x \in A} \tilde{U}_x \supset A$$

is an open cover of A and by compactness finitely many neighborhoods suffice to cover A. Thus, there are $x_j \in A$, $j \in \{1, \dots, m\}$ such that:

$$U_{A,B} := \bigcup_{j \in \{1,\dots m\}} \tilde{U}_{x_j} \supset A.$$

By construction, $U_{A,B}$ and

$$U_{B,A} := \bigcap_{j \in \{1,\dots m\}} V_{x_j}$$

are disjoint open neighborhoods. □

DEFINITION A.1.12. A topological space (X, τ) is called *locally compact* if each point $x \in X$ has a compact neighbourhood.

PROPOSITION A.1.13. *([**40**], Proposition 10.15) A topological space that is locally compact and Hausdorff and possesses a base with just countably many elements is metrizable.*

DEFINITION A.1.14. A separable, topological space whose topology is generated by a complete metric is called *polish space*.

PROPOSITION A.1.15. *A topological space that is locally compact and Hausdorff and possesses a base with just countably many elements is polish.*

PROOF. Combine [**40**], Proposition 13.17 and [**40**], Proposition 10.15. □

DEFINITION A.1.16. (one-point-compactification) Let (X, τ) be a topological space and $\infty \notin X$. Then $X^* := X \cup \{\infty\}$ equipped with the topology

$$\tau^* := \tau \cup \{ X^* \setminus A | \, A \text{ is closed and compact in } (X, \tau)\}$$

is compact.

PROPOSITION A.1.17. *([**40**], Proposition 8.12) (Tychonoff)*
A non-empty product space

$$X = \prod_{i \in I} X_i$$

is compact if and only if all X_i are compact.

A.2. Analysis

LEMMA A.2.1. *Let Y be a topological space. The following properties hold true:*
(i) A continuous function $f : Y \to \mathbb{R}$ is upper and lower semicontinuous.
(ii) If $f : Y \to (0, \infty)$ is lower (upper) semicontinuous then $g := \frac{1}{f}$ is upper (lower) semicontinuous.

(iii) If $f : Y \to (0, \infty)$ and $g : Y \to (0, \infty)$ are lower (upper) semicontinuous then also $h := f \cdot g$ is lower (upper) semicontinuous.

Proof.
(i) Clear.
(ii) Let $f : Y \to (0, \infty)$ be lower semicontinuous and $x \in Y$ and $\varepsilon > 0$ arbitrary. Then for

$$f(x) > \delta := \frac{\varepsilon \left(f(x)\right)^2}{1 + \varepsilon f(x)} > 0$$

there exists a neighborhood U_x of x such that $f(y) > f(x) - \delta > 0$ for all $y \in U_x$. Therefore,

$$\begin{aligned} \frac{1}{f(y)} &< \frac{1}{f(x) - \delta} \\ &= \left(1 + \frac{\delta}{f(x) - \delta}\right) \frac{1}{f(x)} \\ &= \frac{1}{f(x)} + \varepsilon \end{aligned}$$

for all $y \in U_x$. For f being upper semicontinuous, the assertion is proved in the same way.
(iii) Let $f : Y \to (0, \infty)$ and $g : Y \to (0, \infty)$ be lower semicontinuous and let $x \in Y$ and $\varepsilon > 0$ be arbitrary. Then for $\delta = \frac{\varepsilon}{f(x)+g(x)} > 0$ there exist neighborhoods U_x and V_x of x such that

$$f(y) > f(x) - \delta$$

for all $y \in U_x$ and

$$g(y) > g(x) - \delta$$

for all $y \in V_x$. Then for $y \in U_x \cap V_x$

$$\begin{aligned} h(y) &= f(y) \cdot g(y) \\ &> \left(f(x) - \delta\right) \left(g(x) - \delta\right) \\ &= h(x) - \frac{\varepsilon}{f(x) + g(x)} \left(f(x) + g(x)\right) + \delta^2 \\ &> h(x) - \varepsilon. \end{aligned}$$

For f and g being upper semicontinuous, the assertion is proved in the same way. □

Definition A.2.2. Let X, Y be normed vector spaces and

$$f : \ X \to Y.$$

For $x \in X$, f is said to be *Fréchet differentiable at* x if there is a linear bounded map

$$Df(x): \ X \to Y$$

such that

$$\lim_{\|h\|_X \searrow 0} \frac{\|f(x+h) - f(x) - Df(x)h\|_Y}{\|h\|_X} = 0.$$

If f is Fréchet differentiable at x for any $x \in X$ f it is simply called *Fréchet differentiable*.

THEOREM A.2.3. *([**2**], Theorem 7.6) Let E be a finite dimensional Banach space, $J \subset \mathbb{R}$ and $D \subset E$ open. Assume that*

$$f: \ J \times D \to E$$

is continuous and that for any $x \in D$ there is a neighbourhood in $J \times D$ such that f is Lipschitz continuous on this neighbourhood. Then for every $(t_0, x_0) \in J \times D$ there exists a unique solution

$$u(\cdot, t_0, x_0) : J(t_0, x_0) \to D$$

of

$$\frac{dx}{dt} = f(t, x) \ \text{and} \ x(t_0) = x_0,$$

for which $J(t_0, x_0)$ is the maximal interval of existence. $J(t_0, x_0)$ is open:

$$J(t_0, x_0) = \left(t^-(t_0, x_0), t^+(t_0, x_0)\right),$$

and we either have

$$t^- := t^-(t_0, x_0) = \inf J,$$

and

$$t^+ := t^+(t_0, x_0) = \inf J,$$

or

$$\lim_{t \to t^{\pm}} \min \left\{ dist\left(u(t, t_0, x_0), \partial D\right), \frac{1}{\|u(t, t_0, x_0)\|} \right\} = 0.$$

THEOREM A.2.4. *([**21**] Theorem 3.1) Let $t \in \mathbb{R}$, $y, y_0 \in \mathbb{R}^d$, $d \in \mathbb{N}$ and f, η be $\mathbb{R}^d$-valued. Let $f(t, y, z)$ be continuous on an open (t, y, z)-set E and possess continuous first order partials $\frac{\partial f}{\partial y^k}$, $\frac{\partial f}{\partial z^i}$ with respect to the components y and z. Then the unique solution*

$$y = \eta(t, t_0, y_0, z)$$

of

$$y' = f(t, y, z) \ \text{and} \ y(t_0) = y_0,$$

where $z = (z_1, ..., z_e)$ is a set of parameters, is of C^1-class on its open domain of definition.

Definition A.2.5. Let A be a $\mathbb{K}$-vector space with a bilinear map

$$\cdot: \ A \times A \to A$$

such that for any $a, b, c \in A$ and any $\mu, \lambda \in \mathbb{K}$

$$a \cdot (b + c) = a \cdot b + a \cdot c,$$

$$(b + c) \cdot a = b \cdot a + c \cdot a$$

and

$$(\lambda a) \cdot (\mu b) = (\lambda \mu)\,(b \cdot a)\,.$$

Then A is called *algebra*.

Theorem A.2.6. *(Stone-Weierstrass on $\mathbb{R}$, [**30**], VIII.4.7)*
Let K be compact and $A \subset C(K)$ be an algebra with respect to pointwise multiplication . If
(i)$1_K \in A$, and
(ii)A separates points (which means that for any $x, y \in K$, $x \neq y$ there is $f \in A$ such that $f(x) \neq f(y)$),
then A is dense in $C(K)$ with respect to $\|\cdot\|_\infty$.

Proposition A.2.7. *([**4**], Definition 1.1) Let U, V be $\mathbb{K}$- vector spaces. There exists a $\mathbb{K}$- vector space $U \otimes V$ denoted as tensor product space and a map*

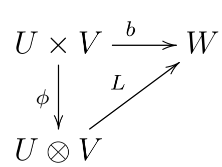
$$\phi: \ U \times V \to U \otimes V,$$

denoted as canonical bilinear map such that for any $\mathbb{K}$- vector space W and any bilinear map

$$b: \ U \times V \to W$$

there is a linear map $L: \ U \otimes V \to W$ such that

$$L \circ \phi = b.$$



$U \otimes V$ and ϕ are unique up to a bijective linear map and $u \otimes v := \phi(u, v)$ for any $u \in U$, $v \in V$.

Lemma A.2.8. *([**4**], Remark 1.2 (5))Let U, V be $\mathbb{K}$- vector spaces. Every $x \in U \otimes V$ is given by $x = \sum_{i=1}^{n} u_i \otimes v_i$ for some $n \in \mathbb{N}$ and linear independent $(u_i)_{i \in \{1,\dots,n\}} \subset U$ and $(v_i)_{i \in \{1,\dots,n\}} \subset V$.*

LEMMA A.2.9. *([4], Remark 1.2 (6))Let* U, V, W *be* $\mathbb{K}$*- vector spaces. Then*

$$U \otimes V \otimes W := (U \otimes V) \otimes W = U \otimes (V \otimes W).$$

A.3. Probability Theory

A.3.1. σ-Algebras and Measures.

DEFINITION A.3.1. Let Ω be some set and 2^Ω its power set which is the set of all subsets of Ω. A family of sets $\Sigma \subset 2^\Omega$ is called *σ- algebra* on Ω if

(i) $\Omega \in \Sigma$,

(ii) if $E \in \Sigma$ then $\Omega \setminus E \in \Sigma$, and

(iii) if $E_i \in \Sigma$ for any $i \in \mathbb{N}$ then $\bigcup_{i \in \mathbb{N}} E_i \in \Sigma$.

If Σ is a σ-algebra on Ω, then the pair (Ω, Σ) is called *measurable space*.

DEFINITION A.3.2. Let Ω be some set and 2^Ω its power set and $\mathcal{G} \subset 2^\Omega$ a family of subsets of Ω. $\sigma(\mathcal{G})$ is denotes the smallest σ-algebra such that $\mathcal{G} \subset \sigma(\Omega)$ and $\mathcal{G}$ is called *generator* of $\sigma(\mathcal{G})$.

LEMMA A.3.3. *For* $g : (\Omega_1, \Sigma_1) \to (\Omega_2, \Sigma_2)$ *and a generator* $\mathcal{G}$ *of* Σ_2 *g is measurable if and only if* $g^{-1}(A) \in \Sigma_1$ *for all* $A \in \mathcal{G}$.

PROOF. Define

$$\mathcal{M} := \left\{ A \in \Sigma_2 : \; g^{-1}(A) \in \Sigma_1 \right\}.$$

One shows easily that this is a σ-algebra. Since $\mathcal{G} \subset \mathcal{M}$ this implies $\Sigma_2 \subset \mathcal{M} \subset \Sigma_2$. $\square$

DEFINITION A.3.4. Let T be a topological space and $\mathcal{O}$ the set of all open sets. Then $\mathcal{B}(T) := \sigma(\mathcal{O})$ is called *Borel σ- algebra*. If not explicitly stated otherwise, on a topological space the σ-algebra considered is the Borel σ-algebra.

Definition A.3.5. Let I be an index set, let $(\Omega_i)_{i\in I}$ be a family of sets and let $J \subset J' \subset I$. The family $(\Pi_i)_{i\in I}$ of coordinate maps (see Definition A.1.6) is called *coordinate process* if $\Omega_i = \Omega$ for all $i \in I$ and if $(\Omega, \mathcal{F})$ is a measurable space.

Definition A.3.6. Let Σ_1 be a σ-algebra on Ω_1 and Σ_2 be a σ-algebra on Ω_2. A function

$$f : \Omega_1 \to \Omega_2$$

is called Σ_1-Σ_2- *measurable* if for all $E \in \Omega_2$ also $f^{-1}(E) \in \Omega_1$. If it is unambiguous which σ-algebra is meant, the map is often just called measurable . The smallest σ-algebra on Ω_1 with respect to which f is measurable is denoted $\sigma(f)$. If I is a non-empty index and $f_i : \Omega_1 \to \Omega_2$ for all$i \in I$ then $\sigma(f_i : i \in I)$ is the smallest σ-algebra on Ω_1 with respect to which all f_i, $i \in I$ are measurable.

Definition A.3.7. Let I be an index set and $(\Omega_i, \mathcal{F}_i)$ be measurable spaces for all $i \in I$. Then

$$\bigotimes_{i\in I} \mathcal{F}_i := \sigma\left(\Pi_j^I : j \in I\right)$$

is defined as the smallest σ-algebra on $\times_{i\in I} \Omega_i$ such that the coordinate maps $\Pi_j^I : \times_{i\in I} \Omega_i \to \Omega_j$ are measurable for all $j \in I$. Written differently,

$$\bigotimes_{i\in I} \mathcal{F}_i = \sigma\left(\left(\Pi_i^I\right)^{-1}(E_i) : E_i \in \mathcal{F}_i,\, i \in I\right).$$

Corollary A.3.8. *For any $i \in I$ let $\mathcal{G}_i \subset \mathcal{F}_i$ be a generator of $\mathcal{F}_i$. Then*

$$\mathcal{G} := \left(\left(\Pi_i^I\right)^{-1}(E_i) : E_i \in \mathcal{G}_i,\, i \in I\right)$$

is a generator of $\bigotimes_{i\in I} \mathcal{F}_i$.

Proof. For any $i \in I$ by Lemma A.3.3 the projection

$$\Pi_i^I : \left(\times_{i\in I} \Omega_i, \sigma(\mathcal{G})\right) \to (\Omega_i, \mathcal{F}_i)$$

is measurable. Hence

$$\bigotimes_{i\in I} \mathcal{F}_i \subset \sigma(\mathcal{G}) \subset \bigotimes_{i\in I} \mathcal{F}_i.$$

□

LEMMA A.3.9. *([**30**],Corollary 1.97) (Doob-Dynkin lemma) Let* $X : \Omega_1 \to \Omega_2$ *be a map between some non-empty set* Ω_1 *and a measurable space* $(\Omega_2, \mathcal{F})$*. Then* $f : \Omega_1 \to \mathbb{R} \cup \{-\infty, \infty\}$ *is measurable with respect to* $\sigma(X)$ *if and only if there is a* $\mathcal{F}$*-measurable function*

$$g : \Omega_2 \to \mathbb{R} \cup \{-\infty, \infty\}$$

such that $f = g(X)$*.*

LEMMA A.3.10. *([**30**], Corollary 1.82) Let* I *be a non-empty index set and* $(E_1, \mathcal{E}_1)$*,* $(E_2, \mathcal{E}_2)$ *and for all* $i \in I$ *also* $(\Omega_i, \mathcal{F}_i)$ *be measurable spaces. For each* $i \in I$ *let* Z_i *be a map from* E_2 *to* Ω_i *such that* $\mathcal{E}_2 = \sigma(Z_i : i \in I)$*. Then a map* $Y : E_1 \to E_2$ *is* $\mathcal{E}_1$*-* $\mathcal{E}_2$ *measurable if and only if* $Z_i \circ Y$ *is* $\mathcal{E}_1 - \mathcal{F}_i$ *measurable for all* $i \in I$*.*

PROPOSITION A.3.11. *([**30**], Proposition 1.23)* $\mathcal{B}(\mathbb{R})$ *is identical to the* σ*-algebra generated by the intervals* $(-\infty, a)$ *,* $a \in \mathbb{Q}$*, or by the intervals* $(-\infty, a]$ *,* $a \in \mathbb{Q}$*, or by the intervals* (a, ∞) *,* $a \in \mathbb{Q}$ *or by the intervals* $[a, \infty)$ *,* $a \in \mathbb{Q}$*.*

DEFINITION A.3.12. Let Ω be some set and 2^Ω its power set which is the set of all subsets of Ω. A family of sets $\mathcal{D} \subset 2^\Omega$ is called *Dynkin-system* on Ω if

(i) $\Omega \in \mathcal{D}$,

(ii) if $A, B \in \mathcal{D}$ and $A \supset B$, then $A \setminus B \in \mathcal{D}$, and

(iii) if $E_i \in \mathcal{D}$ for any $i \in \mathbb{N}$ and if the sets $(E_i)_{i\in\mathbb{N}}$ are pairwise disjoint, then $\bigcup_{i\in\mathbb{N}} E_i \in \mathcal{D}$.

REMARK A.3.13. For $\mathcal{G} \subset 2^\Omega$ the smallest Dynkin system $\mathcal{D}$ on Ω such that $\mathcal{G} \subset \mathcal{D}$ is denoted $\delta(\mathcal{G})$.

DEFINITION A.3.14. Let Ω be some set. $\mathcal{P} \subset 2^\Omega$ is called *intersection stable* if A, $B \in \mathcal{P}$ implies $A \cap B \in \mathcal{P}$.

LEMMA A.3.15. *([**30**], Proposition 1.19) (Dynkin's* π*-*λ *theorem) If* $\mathcal{E} \subset 2^\Omega$ *is intersection stable, then*

$$\delta(\mathcal{E}) = \sigma(\mathcal{E}).$$

LEMMA A.3.16. *Let* $(\Omega_1, \mathcal{F}_1)$, $(\Omega_2, \mathcal{F}_2)$, $(E, \mathcal{E})$ *be measurable spaces. Let* $f : \Omega_1 \times \Omega_2 \to E$ *be* $\mathcal{F}_1 \otimes \mathcal{F}_2$*-*$\mathcal{E}$*-measurable. For* $\omega_1 \in \Omega_1$ *fixed*

$$\begin{aligned} &\Omega_2 \to E \\ &\omega_2 \to f(\omega_1, \omega_2) \end{aligned}$$

is $\mathcal{F}_2$*-measurable.*

PROOF. For the map

$$\begin{aligned} i_{\omega_1} : \ &\Omega_2 \to \Omega_1 \times \Omega_2 \\ &\omega_2 \to (\omega_1, \omega_2) \end{aligned}$$

$i_{\omega_1}^{-1}(A_1 \times A_2) \in \mathcal{F}_2$ holds for any $A_1 \in \mathcal{F}_1$, $A_2 \in \mathcal{F}_2$, hence by Lemma A.3.3 i_{ω_1} is measurable, thus $f \circ i_{\omega_1}$. □

LEMMA A.3.17. *Let* (Ω, Σ) *be a measurable space and let*

$$f_n : (\Omega, \Sigma) \to (\mathbb{R} \cup \{-\infty, \infty\}, \mathcal{B}(\mathbb{R} \cup \{-\infty, \infty\})), \ n \in \mathbb{N}$$

be a sequence of measurable functions. Then

$$\inf_{n\in\mathbb{N}} f_n, \ \sup_{n\in\mathbb{N}} f_n, \ \liminf_{n\to\infty} f_n, \ \limsup_{n\to\infty} f_n,$$

and , if it exists, $\lim_{n\to\infty} f_n$ *are measurable.*

PROOF. We apply Lemma A.3.3 and obtain for any $a \in \mathbb{R}$

$$\left(\inf_{n\in\mathbb{N}} f_n\right)^{-1} (-\infty, a) = \bigcup_{n\in\mathbb{N}} f_n^{-1} (-\infty, a) \in \Sigma_1,$$

hence, by Proposition A.3.11 indeed $\mathcal{B}(\mathbb{R} \cup \{-\infty, \infty\}) \subset \mathcal{M}$ and

$$\inf_{n\in\mathbb{N}} f_n$$

is measurable. The same holds true for

$$\sup_{n\in\mathbb{N}} f_n.$$

By writing

$$\liminf_{n\to\infty} f_n = \sup_{N\in\mathbb{N}} \left(\inf_{\substack{n \in \mathbb{N} \\ n > N}} f_n \right)$$

we obtain measurability of

$$\liminf_{n\to\infty} f_n$$

and analogously of

$$\limsup_{n\to\infty} f_n.$$

Finally, we observe, that if it exists,

$$\lim_{n\to\infty} f_n = \limsup_{n\to\infty} f_n = \liminf_{n\to\infty} f_n.$$

□

DEFINITION A.3.18. Let (Ω, Σ) be a measurable space. A function

$$f : \Omega \to \mathbb{R}$$

is called *simple function* , if it can be written as

$$f = \sum_{i=1}^{n} \alpha_i 1_{A_i}$$

for some $n \in \mathbb{N}$, where $\alpha_1, ..., \alpha_n \in \mathbb{R}$ and $A_1, ..., A_n \in \Sigma$ are pairwise disjoint.

PROPOSITION A.3.19. *Let* (Ω, Σ) *be a measurable space and*

$$f : \Omega \to [0, \infty]$$

measurable. Then there exists a sequence of non-negative simple functions $(f_n)_{n\in\mathbb{N}}$ *such that*

$$f_n \nearrow f.$$

PROOF. For $n \in \mathbb{N}$ define

$$f_n = \min\left(2^{-n} \lfloor 2^n f \rfloor, n\right).$$

□

DEFINITION A.3.20. Let (Ω, Σ) be a measurable space . A map $\mu : \Sigma \to \mathbb{R}_+ \cup \{\infty\}$ is called *measure* if it is σ-*additive* which means, that for all pairwise disjoint sets $E_1, E_2, ... \in \Sigma$

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i).$$

If additionally $\mu(\Omega) = 1$ then μ is called *probability measure*. If $\mu(\Omega) < \infty$ then μ is called *finite*. If there is a sequence $(E_i)_{i\in\mathbb{N}} \subset \Sigma^{\mathbb{N}}$such that $\Omega = \bigcup_{i\in\mathbb{N}} E_i$ and $\mu(E_i) < \infty$ the map is called σ- *finite*. The set of all probability measures on (Ω, Σ) is denoted by $\mathcal{M}_1(\Omega, \Sigma)$. The set of all σ-finite measures on (Ω, Σ) is called $\mathcal{M}_\sigma(\Omega, \Sigma)$.

Example A.3.21. On a measurable space (Ω, Σ) for $\omega \in \Omega$ the map

$$\begin{aligned} \delta_\omega : \Sigma &\to [0,1] \\ A &\to \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{else} \end{cases} \end{aligned}$$

is a probability measure and called *Dirac measure* in ω.

Definition A.3.22. For a set Ω, a σ-algebra $\Sigma \subset 2^\Omega$ and a measure $\mu : \Sigma \to \mathbb{R}_+ \cup \{\infty\}$ we call the triple (Ω, Σ, μ) *measure space*. If μ is a probability measure, then (Ω, Σ, μ) is called *probability space*.

Definition A.3.23. Two measurable spaces (Ω, Σ) and (Ω', Σ') are called *isomorphic* if there is a measurable bijective map $\varphi : \Omega \to \Omega'$, called *isomorphism*, such that φ^{-1} is measurable. Two measure spaces (Ω, Σ, μ) and (Ω', Σ', μ') are called if (Ω, Σ) and (Ω', Σ') are isomorphic and for their isomorphism φ the equation $\mu' = \mu \circ \varphi^{-1}$ holds true.

Definition A.3.24. A measurable map between a probability space $(\Omega, \Sigma, \mathbb{P})$ and a measurable space $(E, \mathcal{E})$ is called *random variable*. If we speak of random variables on $(\Omega, \Sigma, \mathbb{P})$ without specifying the space E, then $E = \mathbb{R}$ is meant.

Definition A.3.25. For a probability space $(\Omega, \Sigma, \mathbb{P})$ and a real-valued random variable X the probability measure $P_X := \mathbb{P} \circ X^{-1}$ is called *distribution*. For $\mu = \mathbb{P} \circ X^{-1}$ we write

$$X \sim \mu.$$

Example A.3.26. If $\mu \in \mathbb{R}$, $\sigma^2 > 0$ and X is a real-valued random variable such that for any $x \in \mathbb{R}$

$$\mathbb{P}(X \leq x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt,$$

then $\mathcal{N}_{\mu,\sigma^2} := \mathbb{P} \circ X^{-1}$ is called *normal distribution with parameters μ and σ^2*.

If X is $\mathbb{R}^d$-valued, $\mu \in \mathbb{R}^d$, Σ is a positive definite $d \times d$ matrix and for any $x = (x_1, ..., x_d) \in \mathbb{R}^d$

$$\mathbb{P}(X \leq x) = \det(2\pi\Sigma)^{-1/2} \int_{-\infty}^{x_d} ... \left(\int_{-\infty}^{x_1} \exp\left(-\frac{1}{2} \left\langle t - \mu, \Sigma^{-1}(t - \mu) \right\rangle \right) dt_1 \right) dt_d,$$

with $t = (t_1, ..., t_d) \in \mathbb{R}^d$, then $\mathcal{N}_{\mu,\Sigma} := \mathbb{P} \circ X^{-1}$ is called *d-dimensional normal distribution with parameters μ and Σ.*

Definition A.3.27. On Ω a *semiring* is a system of sets $\mathcal{S} \subset 2^\Omega$such that

(i)$\emptyset \in \mathcal{S}$

(ii) if $A, B \in \mathcal{S}$, then $A \setminus B$ is the finite union of pairwise disjoint sets in S

(iii) if $A, B \in \mathcal{S}$, then $A \cap B \in \mathcal{S}$

Definition A.3.28. Let $\mathcal{S}$ be a semiring on Ω and $\mu : \mathcal{S} \to [0, \infty]$. μ is called σ-subadditive if for any $A, A_1, ..., A_n \in \mathcal{S}$ such that $A \subset \bigcup_{i=1}^{n} A_i$

$$\mu(A) \leq \sum_{i=1}^{n} \mu(A_i).$$

Theorem A.3.29. *(*[**30**]*, Proposition 1.53) (Caratheodory extension theorem)*

Let $\mathcal{S}$ be a semiring on Ω and $\mu : \mathcal{S} \to [0, \infty]$ be an additive, σ-subadditive, σ-finite map with $\mu(\emptyset) = 0$. Then there exists a unique extension of μ to a measure $\bar{\mu}$ on $\sigma(\mathcal{S})$ and $\bar{\mu}$ is σ-finite.

Definition A.3.30. A measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, $n \in \mathbb{N}$ obtained by extending the map

$$\mu(\underset{i \in \{1,...,n\}}{\times} [a_i, b_i]) = (b_1 - a_1) \cdot ... \cdot (b_n - a_n)$$

for $a_i \leq b_i$, $i \in \{1, ..., n\}$ by Theorem A.3.29 (and Proposition A.3.11 and Definition A.3.7) is called *Lebesgue measure*. If not stated otherwise, as a measure on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ we always choose the Lebesgue measure.

Proposition A.3.31. *Let μ_1 and μ_2 be two (signed) measures on the measurable space (Ω, Σ) that coincide on an intersection stable generator $\mathcal{E}$ of Σ for which there are sets $E_1, E_2, \ldots \in \mathcal{E}$, $E_n \subset E_{n+1}$ for all $n \in \mathbb{N}$, such that*

$$\Omega = \bigcup_{n \in \mathbb{N}} E_n$$

and $\mu_1(E_n) = \mu_2(E_n) < \infty$. Then μ_1 and μ_2 coincide everywhere.

Proof. For $E \in \mathcal{E}$ such that $\mu_1(E) = \mu_2(E) < \infty$ define

$$\mathcal{D}_E := \{A \in \Sigma : \mu_1(E \cap A) = \mu_2(E \cap A)\}.$$

One can easily show that this is a Dynkin system and clearly $\mathcal{E} \subset \mathcal{D}_E$, hence $\delta(\mathcal{E}) \subset \mathcal{D}_E$. Therefore, by Proposition A.3.15

$$\delta(\mathcal{E}) = \sigma(\mathcal{E}) = \Sigma.$$

Thus, $\mathcal{D}_E = \Sigma$ and for any $A \in \Sigma$ and any $E \in \mathcal{E}$ such that $\mu_1(E) = \mu_2(E) < \infty$ we obtain:

$$\mu_1(E \cap A) = \mu_2(E \cap A).$$

In particular, for any $A \in \Sigma$

$$\mu_1(A) = \lim_{n \to \infty} \mu_1(E_n \cap A) = \lim_{n \to \infty} \mu_2(E_n \cap A) = \mu_2(A).$$

□

Definition A.3.32. Let (Ω, Σ, μ) be a measure space. A set $N \in \Sigma$ such that $\mu(N) = 0$ is called *null set*. If for a second measure ν on (Ω, Σ) any null set with respect to μ is also a null set with respect to ν then ν is said to be *absolutely continuous* with respect to μ which is denoted as $\nu \ll \mu$. If both $\nu \ll \mu$ and $\nu \gg \mu$ then μ and ν are called *equivalent*. If for any null set N also $A \in \Sigma$ for any $A \subset N$ then the measure space (Ω, Σ, μ) is called *complete*. If $\Sigma' \supset \Sigma$ is the smallest σ-algebra such that for a measure μ' on Σ' with $\mu'|_{\Sigma} = \mu$ the measure space (Ω, Σ', μ') is complete then (Ω, Σ', μ') is called *completion of* (Ω, Σ, μ). If a property holds on the set $\Omega \setminus N$ where N is a null set, it is said to hold *almost everywhere* or *almost surely* in case if μ is a probability measure.

Lemma A.3.33. *The sum of two signed Radon measures is a signed Radon measure.*

PROOF. Let μ_1, μ_2 be signed Radon measures on (Ω, Σ). Then $\mu_1 + \mu_2$ is clearly a signed measure on (Ω, Σ) and $|\mu_1 + \mu_2| \leq |\mu_1| + |\mu_2|$ is locally finite. For any open set $O \subset \Omega$and $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset O$ such that

$$\varepsilon \geq |\mu_1| (O \setminus K_\varepsilon) = \mu_1^+ (O \setminus K_\varepsilon) + \mu_1^- (O \setminus K_\varepsilon)$$

and

$$\varepsilon \geq |\mu_2| (O \setminus K_\varepsilon) = \mu_2^+ (O \setminus K_\varepsilon) + \mu_2^- (O \setminus K_\varepsilon) .$$

Thus,

$$2\varepsilon \geq \left(\mu_1^+ + \mu_2^+\right) (O \setminus K_\varepsilon) + \left(\mu_1^- + \mu_2^-\right) (O \setminus K_\varepsilon) ,$$

and the inequalities

$$\left(\mu_1^+ + \mu_2^+\right) (O \setminus K_\varepsilon) \geq (\mu_1 + \mu_2)^+ (O \setminus K_\varepsilon) ,$$

and

$$\left(\mu_1^- + \mu_2^-\right) (O \setminus K_\varepsilon) \geq (\mu_1 + \mu_2)^- (O \setminus K_\varepsilon) ,$$

imply

$$\begin{aligned} |\mu_1 + \mu_2| (O \setminus K_\varepsilon) &= (\mu_1 + \mu_2)^+ (O \setminus K_\varepsilon) + (\mu_1 + \mu_2)^- (O \setminus K_\varepsilon) \\ &\leq 2\varepsilon. \end{aligned}$$

Therefore, $|\mu_1 + \mu_2|$ is inner regular and $\mu_1 + \mu_2$ is a signed Radon measure. □

PROPOSITION A.3.34. *([**17**], Chapter VIII, §1, Proposition 1.5) Let (X, τ) be a topological space . A finite measure on the measurable space $(X, \mathcal{B}(X))$ is a Radon measure if for all open sets $O \subset X$*

$$\mu (O) = \sup \{\mu (K) : \ K \subset O, \ K \textit{ compact}\} .$$

DEFINITION A.3.35. Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and I an index set. A family of sets $(A_i)_{i \in I} \subset \Sigma$ is called *independent* if for any finite $J \subset I$

$$\mathbb{P} \left(\bigcap_{i \in J} A_i \right) = \prod_{i \in J} \mathbb{P} (A_i) .$$

A family $(\mathcal{B}_i)_{i \in I}$ such that $\mathcal{B}_i \subset \Sigma$ for any $i \in I$ is called *independent* if for any finite $J \subset I$ and any family of sets $(B_i)_{i \in J} \subset \Sigma$ such that $B_i \in \mathcal{B}_i$ for any $i \in J$

$$\mathbb{P} \left(\bigcap_{i \in J} B_i \right) = \prod_{i \in J} \mathbb{P} (B_i)$$

holds true.

A family of random variables $(X_i)_{i \in I}$ on $(\Omega, \Sigma, \mathbb{P})$ is called *independent* if the family $(\sigma (X_i))_{i \in I}$ is independent.

Definition A.3.36. Let $\Omega \neq \emptyset$. A set $O \subset \Omega$ is called $C_b(\Omega)$-*open*, if there exists a sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(\Omega)$ such that pointwise $f_n \nearrow 1_O$. The system of sets that are $C_b(\Omega)$- open is called $\mathcal{G}(C_b(\Omega))$.

In M.Schweizer's lecture notes "Measure and Integration" (version July 22, 2017) it is shown in Lemma IV.1.11:

Lemma A.3.37. *Let $\Omega \neq \emptyset$ and let $\mathcal{G}(C_b(\Omega))$ be the system of sets that are $C_b(\Omega)$- open. Then*

$$\sigma(\mathcal{G}(C_b(\Omega))) = \sigma(f|\ f \in C_b(\Omega)),$$

the smallest σ-algebra such that all maps in $C_b(\Omega)$ are measurable.

Definition A.3.38. The smallest σ-algebra such that all maps in $C(\Omega)$ are measurable i.e. $\sigma(C(\Omega))$ is called *Baire σ-algebra* and is written $\mathcal{B}_0(\Omega)$.

Remark A.3.39. Clearly, $\mathcal{B}_0(\Omega) = \sigma(C_b(\Omega))$.

A.3.2. Integration and Conditional Expectation.

Definition A.3.40. Let (Ω, Σ, μ) be a measure space. For a measurable function $f : \Omega \to [0, \infty]$ the integral

$$\int_\Omega f d\mu$$

is defined as the supremum of

$$\sum_{i=1}^{n} \alpha_i \mu(A_i)$$

over all positive simple functions

$$g = \sum_{i=1}^{n} \alpha_i 1_{A_i} \leq f.$$

Definition A.3.41. Let (Ω, Σ, μ) be a measure space. For a measurable function $f : \Omega \to \mathbb{R}$ and $f^+ := \max(f, 0)$ and $f^- := \max(-f, 0)$ the integral is defined as

$$\int_\Omega f d\mu := \int_\Omega f^+ d\mu - \int_\Omega f^- d\mu,$$

provided not both $\int_\Omega f^+ d\mu$ and $\int_\Omega f^- d\mu$ are infinity.

For a signed measure μ on (Ω, Σ) and a measurable function $f : \Omega \to \mathbb{R}$ such that $\int_\Omega f d\,|\mu| < \infty$ the integral is defined as

$$\int_\Omega f d\mu := \int_\Omega f d\mu^+ - \int_\Omega f d\mu^-$$

If μ is a probability measure, then

$$\mathbb{E}\left[f\right] := \int_\Omega f d\mu$$

is called *expected value*.

DEFINITION A.3.42. Let (Ω, Σ, μ) be a measure space. The set of all measurable functions $f : \Omega \to \mathbb{R} \cup \{-\infty, \infty\}$ such that

$$\int_\Omega |f|\, d\mu < \infty$$

is called $\mathcal{L}^1(\Omega, \Sigma, \mu)$. Any element $f \in \mathcal{L}^1(\Omega, \Sigma, \mu)$ is called *integrable*. If $f^2 \in \mathcal{L}^1(\Omega, \Sigma, \mu)$, then f is called *square − integrable*.

THEOREM A.3.43. *(**[23]**, Theorem 12.34) Let (Ω, Σ, μ) be a measure space and $f \in L^1(\Omega, \Sigma, \mu)$. Then for every $\varepsilon > 0$ there exists $\delta > 0$ such that for any $A \in \Sigma$ with $\mu(A) < \delta$*

$$\int_A |f|\, d\mu < \varepsilon.$$

LEMMA A.3.44. *Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space and $(E_1, \mathcal{E}_1)$, $(E_2, \mathcal{E}_2)$ measurable. Let $X, Y \in \mathcal{L}^1(\Omega, \Sigma, \mathbb{P})$ be independent random variable that take values in $(E_1, \mathcal{E}_1)$ and $(E_2, \mathcal{E}_2)$. If $f : (E_1, E_2) \to \mathbb{R}_+$ is measurable then*

$$\mathbb{E}\left[f(X, Y)\right] = \int_{E_1} \int_{E_2} f(x, y) P_X(dx) P_Y(dy).$$

PROOF. For $F_1 \in \mathcal{E}_1$ and $F_2 \in \mathcal{E}_2$ by Proposition A.3.45

$$\begin{aligned}\mathbb{E}\left[1_{F_1\times F_2}(X,Y)\right] &= \mathbb{E}\left[1_{F_1}(X)1_{F_2}(Y)\right]\\ &= \mathbb{E}\left[1_{F_1}(X)\right]\mathbb{E}\left[1_{F_2}(Y)\right]\\ &= \int_{E_2}\left(\int_{E_1} 1_{F_1}(x)P_X(dx)\right)1_{F_2}(y)P_Y(dy)\\ &= \int_{E_1}\int_{E_2} 1_{F_1\times F_2}(x,y)P_X(dx)P_Y(dy).\end{aligned}$$

holds true. Hence the family $\mathcal{D}$ of sets$\{B \in \mathcal{E}_1 \otimes \mathcal{E}_2\}$ such that

$$\mathbb{E}\left[1_B(X,Y)\right] = \int_{E_1}\int_{E_2} 1_B(x,y)P_X(dx)P_Y(dy)$$

holds true contains the (intersection stable) generator of $\mathcal{E}_1 \otimes \mathcal{E}_2$ (see Corollary A.3.8) and is a Dynkin system as one can easily show. Thus, by Lemma A.3.15 $\mathcal{D} = \mathcal{E}_1 \otimes \mathcal{E}_2$. Linearity of the integral and the expected value and Proposition A.3.19 and monotone convergence (Theorem A.3.57) yield the statement of the Lemma. □

PROPOSITION A.3.45. *([**30**],Proposition 5.4) Let* $(\Omega, \Sigma, \mathbb{P})$ *be a probability space. If* $X, Y \in \mathcal{L}^1(\Omega, \Sigma, \mathbb{P})$ *are independent then* $X \cdot Y \in \mathcal{L}^1(\Omega, \Sigma, \mathbb{P})$ *and*

$$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y].$$

DEFINITION A.3.46. If a measure μ on the measurable space (Ω, Σ) is given by the measure ν and the measurable map $f : \Omega \to \mathbb{R}_+$ by the relation

$$\mu(A) = \int_\Omega f(x)\nu(dx) \text{ for any } A \in \Sigma$$

then f is called *density of* μ with respect to ν.

THEOREM A.3.47. *(Radon-Nikodym, [**30**],Corollary 7.34)*
For σ*-finite measures* μ *and* ν *on the measurable space* (Ω, Σ) ν *has a density with respect to* μ *if and only if* v *is absolutely continuous with respect to* μ.

DEFINITION A.3.48. Let (Ω, Σ, μ) be a σ -finite measure space and E a topological space with Borel σ- algebra $\mathcal{B}(E)$. Let $f_1, f_2, ... : \Omega \to E$ be a sequence of measurable maps. We say that

$$f_n \to f$$

converges μ-almost everywhere if there is a null set $N \in \Sigma$ such that for any $\omega \in \Omega \setminus N$

$$f_n(\omega) \to f(\omega).$$

DEFINITION A.3.49. Let (Ω, Σ, μ) be a σ -finite measure space and (E, d) a metric space with Borel σ- algebra $\mathcal{B}(E)$. Let $f_1, f_2, ... : \Omega \to E$ be a sequence of measurable maps. We say that

$$f_n \to f$$

converges in μ-probability if for any $\varepsilon > 0$ and $A \in \Sigma$ such that $\mu(A) < \infty$

$$\lim_{n\to\infty} \mu\left(\{d\left(f_n, f\right) > \varepsilon\} \cap A\right) = 0.$$

REMARK A.3.50. Let (Ω, Σ, μ) be a σ -finite measure space and (E, d) a metric space with Borel σ- algebra $\mathcal{B}(E)$. Let $f_1, f_2, ... : \Omega \to E$ be a sequence of measurable maps and let $f, g : \Omega \to E$ be measurable. If $f_n \to f$ and $f_n \to g$ converge in μ-probability then for any $A \in \Sigma$ such that $\mu(A) < \infty$ and any $n \in \mathbb{N}$ and any $\varepsilon > 0$

$$\mu\left(\{d\left(g, f\right) > \varepsilon\} \cap A\right) \leq \mu\left(\{d\left(g, f_n\right) > \varepsilon/2\} \cap A\right) + \mu\left(\{d\left(f_n, f\right) > \varepsilon/2\} \cap A\right).$$

The right hand side converges to 0. Choosing a sequence $(A_n)_{n\in\mathbb{N}} \subset \Sigma$ of sets of finite mass such that $A_n \to \Omega$ we conclude that the limit with respect to convergence in μ-probability is μ-almost everywhere unique.

DEFINITION A.3.51. Let E be a metric space and $\mathcal{B}(E)$ its Borel σ-algebra. Let $X, X_1, X_2, ...$ be random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that take values in $(E, \mathcal{B}(E))$. Then $X_n \to X$ *converges in law* if for all $h \in C_b(E)$

$$\mathbb{E}\left[h(X_n)\right] \to \mathbb{E}\left[h(X)\right].$$

PROPOSITION A.3.52. *(*[**26**]*, Lemma 3.7)*

Let E be a metric space and $\mathcal{B}(E)$ its Borel σ-algebra. Let $X, X_1, X_2, ...$ be random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that take values in $(E, \mathcal{B}(E))$. Then

$$X_n \to X$$

in $\mathbb{P}$-probability implies

$$X_n \to X$$

in law. If X is $\mathbb{P}$-almost surely constant, then also the converse is true.

Proposition A.3.53. *([**30**], Corollary 6.13) Let (Ω, Σ, μ) be a measure space and (E, d) a separable metric space with Borel σ- algebra $\mathcal{B}(E)$. Let $f_1, f_2, \ldots : \Omega \to E$ be a sequence of measurable maps. Then*

$$f_n \to f$$

convergences in μ-probability if and only if there exists a subsequence $f_{n_1}, f_{n_2}, \ldots : \Omega \to E$ that converges μ-almost everywhere to f.

Definition A.3.54. ([**30**], Definition 6.16) Let (Ω, Σ, μ) be a measure space. The family $\mathcal{F} \subset L^1(\Omega, \Sigma, \mu)$ is called *uniformly integrable* if

$$\inf_{0 \leq g \leq L^1(\Omega,\Sigma,\mu)} \sup_{f \in \mathcal{F}} \int_\Omega (|f| - g)^+ \, d\mu = 0$$

Proposition A.3.55. *([**30**], Proposition 6.17) Let (Ω, Σ, μ) be a measure space and let μ be finite. Then the family $\mathcal{F} \subset L^1(\Omega, \Sigma, \mu)$ is uniformly integrable*
(i) if

$$\inf_{0 \leq a < \infty} \sup_{f \in \mathcal{F}} \int_\Omega (|f| - a)^+ \, d\mu = 0,$$

(ii) or if

$$\inf_{0 \leq a < \infty} \sup_{f \in \mathcal{F}} \int_{\{|f| > a\}} |f| \, d\mu = 0.$$

Proposition A.3.56. *([**30**], Proposition 6.25) Let (Ω, Σ, μ) be a measure space and $(f_n)_{n \in \mathbb{N}} \subset L^1(\Omega, \Sigma, \mu)$. Then the following statements are equivalent:*

(i) There is $f \in L^1(\Omega, \Sigma, \mu)$ such that

$$\lim_{n \longrightarrow \infty} \|f_n - f\|_{L^1(\Omega,\Sigma,\mu)} = 0$$

.

(ii) $(f_n)_{n \in \mathbb{N}}$ is uniformly integrable and there is a measurable map f such that

$$\lim_{n \longrightarrow \infty} f_n = f$$

in μ-probability.

The limits in (i) and (ii) coincide.

Theorem A.3.57. *(Monotone Convergence Theorem, [**30**], Proposition 4.20)*

Let (Ω, Σ, μ) *be a measure space, and*

$$(f_n)_{n\in\mathbb{N}} \subset \mathcal{L}^1(\Omega, \Sigma, \mu)$$

and

$$f : \Omega \to [-\infty, \infty]$$

be measurable. Let $f_n \nearrow f$ *almost everywhere for* $n \to \infty$. *Then*

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

THEOREM A.3.58. *(Dominated Convergence Theorem,* **[30]***, Corollary 6.26)*

Let (Ω, Σ, μ) *be a measure space,* f *measurable and*

$$(f_n)_{n\in\mathbb{N}} \subset \mathcal{L}^1(\Omega, \Sigma, \mu)$$

such that almost everywhere $f_n \to f$. *If there exists* $0 \le g \in \mathcal{L}^1(\Omega, \Sigma, \mu)$ *such that* $|f_n| \le g$ *almost everywhere for all* $n \in \mathbb{N}$ *then*

$$\lim_{n\to\infty} \int_\Omega |f_n - f| \, d\mu = 0$$

and $f \in \mathcal{L}^1(\Omega, \Sigma, \mu)$.

The proof of the next Lemma follows **[30]**, Lemma 14.20.

LEMMA A.3.59. *Let* κ *be a transition kernel on* $(E, \mathcal{E})$ *and*

$$f : E \times E \to [0, \infty]$$

be measurable with respect to $\mathcal{E} \otimes \mathcal{E}$. *Then*

$$x \to \int\limits_E f(x, y)\kappa(x, dy)$$

is $\mathcal{E}$*-measurable .*

PROOF. The integral is defined because by Lemma A.3.16 for all $x \in E$ the map $y \to f(x, y)$ is measurable. For $A \in \mathcal{E}$ and $B \in \mathcal{E}$ and $f = 1_{A\times B}$ we obtain measurability for

$$x \to \int\limits_E f(x, y)\kappa(x, dy) = 1_A \kappa(x, B).$$

Defining $\mathcal{D}$ as the set of sets $A \in \mathcal{E} \otimes \mathcal{E}$ such that $x \to \int\limits_E 1_A \kappa(x, dy)$ is measurable we show easily that $\mathcal{D}$ is a Dynkin system. Since $\mathcal{D}$ contains the intersection stable generator of $\mathcal{E} \otimes \mathcal{E}$ we obtain by Lemma A.3.15 that $\mathcal{D} = \mathcal{E} \otimes \mathcal{E}$. Therefore

$$x \to \int_E f(x,y)\kappa(x,dy)$$

is measurable for all simple functions (see Definition A.3.18). Since any jointly measurable function

$$f: \ E \times E \to [0,\infty]$$

can be written as the limit of simple functions (Proposition A.3.19), the statement of the lemma follows from the fact, that the limit of measurable functions is measurable (Lemma A.3.17). □

LEMMA A.3.60. *If*

$$\kappa_1 : E \times \mathcal{E} \to [0,\infty]$$

and

$$\kappa_2 : E \times \mathcal{E} \to [0,\infty]$$

are transition kernels on $(E,\mathcal{E})$, *then*

$$\kappa_3(x,A) := \int_E \kappa_2(y,A) \cdot \kappa_1(x,dy)$$

is a transition kernel on $(E,\mathcal{E})$.

PROOF. By Lemma A.3.59 for every $A \in \mathcal{E}$ the map

$$x \to \int_E \kappa_2(y,A) \cdot \kappa_1(x,dy)$$

is measurable and by monotone convergence (Theorem A.3.57)

$$A \to \int_E \kappa_2(y,A) \cdot \kappa_1(x,dy)$$

is a measure on $(E,\mathcal{E})$. □

DEFINITION A.3.61. Let $(\Omega,\Sigma,\mathbb{P})$ be a probability space, let X, Y be two random variables and let $\mathcal{F} \subset \pm$be a sub σ-algebra. Let $X \in \mathcal{L}^1(\Omega,\Sigma,\mathbb{P})$ or $X \geq 0$. Then Y is called *conditional expectation of* X *with respect to* $\mathcal{F}$ if Y is $\mathcal{F}$- measurable and for each $A \in \mathcal{F}$

$$\mathbb{E}[X \cdot 1_A] = \mathbb{E}[Y \cdot 1_A].$$

In this case, it is written

$$Y = \mathbb{E}[X|\mathcal{F}].$$

If $X = 1_E$ for some $E \in \Sigma$, then Y ,the conditional expectation of X with respect to $\mathcal{F}$, is called *conditional probability of E with respect to $\mathcal{F}$* and is written

$$Y = \mathbb{P}\left[E | \mathcal{F}\right].$$

Proposition A.3.62. *([30], Proposition 8.12) On a measure space $(\Omega, \Sigma, \mathbb{P})$, where $X \in \mathcal{L}^1(\Omega, \Sigma, \mathbb{P})$ or $X \geq 0$. and $\mathcal{F} \subset \pm$is a sub σ-algebra $Y := \mathbb{E}\left[X | \mathcal{F}\right]$ exists and if $Y' = \mathbb{E}\left[X | \mathcal{F}\right]$ holds as well then $Y = Y'\mathbb{P}$-almost surely.*

Proposition A.3.63. *([30], Proposition 8.14) Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space, let $X, Y \in \mathcal{L}^1(\Omega, \Sigma, \mathbb{P})$ or $X \geq 0$ and let $\mathcal{G} \subset F \subset \Sigma$ be σ-algebras. Then*
(i)

$$\mathbb{E}\left[\mathbb{E}\left[X | \mathcal{F}\right] | \mathcal{G}\right] = \mathbb{E}\left[X | \mathcal{G}\right],$$

(ii) for $\sigma(X)$ independent of $\mathcal{F}$

$$\mathbb{E}\left[X | \mathcal{F}\right] = \mathbb{E}\left[X\right],$$

(iii) for Y measurable with respect to $\mathcal{F}$ and $\mathbb{E}\left[|XY|\right] < \infty$

$$\mathbb{E}\left[XY | \mathcal{F}\right] = Y\mathbb{E}\left[X | \mathcal{F}\right],$$

Proof. (i) Let $A \in \mathcal{G}$. Then $A \in \mathcal{F}$ and by definition

$$\mathbb{E}\left[1_A \cdot \mathbb{E}\left[X | \mathcal{F}\right]\right] = \mathbb{E}\left[1_A \cdot X\right] = \mathbb{E}\left[1_A \cdot \mathbb{E}\left[X | \mathcal{G}\right]\right].$$

(ii) Clearly $\mathbb{E}\left[X\right]$ is measurable with respect to $\mathcal{F}$ and X and 1_A are independent. Thus, by Proposition A.3.45 for any $A \in \mathcal{F}$

$$\mathbb{E}\left[X 1_A\right] = \mathbb{E}\left[X\right] \mathbb{E}\left[1_A\right] = \mathbb{E}\left[\mathbb{E}\left[X\right] 1_A\right].$$

(iii) (sketch) Assume $Y, X \geq 0$ and approximate Y by

$$Y_n = 2^{-n} \lfloor 2^n Y \rfloor .$$

Then by monotone convergence (Theorem A.3.57) for any $A \in \mathcal{F}$

$$\lim_{n \to \infty} \mathbb{E}\left[1_A Y_n \mathbb{E}\left[X | \mathcal{F}\right]\right] = \mathbb{E}\left[1_A Y \mathbb{E}\left[X | \mathcal{F}\right]\right]$$

and

$$\lim_{n \to \infty} \mathbb{E}\left[1_A Y_n \mathbb{E}\left[X | \mathcal{F}\right]\right] = \mathbb{E}\left[1_A XY\right].$$

For the general case set $Y := Y^+ - Y^-$ and $X := X^+ - X^-$. $\square$

Lemma A.3.64. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X, Y be random variables with measurable state space $(E, \mathcal{E})$. Let X be independent of $\mathcal{F}$ and Y be measurable with respect to $\mathcal{F}$. Then for the measurable positive or bounded map $g: E \times E \to \mathbb{R}$*

$$\mathbb{E}\left[g(X,Y)\middle|\,\mathcal{F}\right] = \mathbb{E}\left[g(X,Y)\middle|\,\sigma(Y)\right].$$

Proof. For $A, B \in \mathcal{E}$ clearly

$$\begin{aligned}\mathbb{E}\left[1_A(X)1_B(Y)\middle|\,\mathcal{F}\right] &= \mathbb{E}\left[1_A(X)\right]1_B(Y)\\ &= \mathbb{E}\left[1_A(X)1_B(Y)\middle|\,\sigma(Y)\right].\end{aligned}$$

Furthermore, the set $\mathcal{D} \subset \mathcal{E} \otimes \mathcal{E}$ such that for $D \in \mathcal{D}$ the equation

$$\mathbb{E}\left[1_D(X,Y)\middle|\,\mathcal{F}\right] = \mathbb{E}\left[1_D(X,Y)\middle|\,\sigma(Y)\right]$$

holds true is a Dynkin system. Since by the first step, it contains the generator of the σ-algebra $\mathcal{E} \otimes \mathcal{E}$, by Lemma A.3.15 it contains all of $\mathcal{E} \otimes \mathcal{E}$. Then the assertion of this Lemma follows from Proposition A.3.19 and monotone convergence for conditional expectations (see Proposition A.3.65). $\square$

Proposition A.3.65. *(monotone convergence) On a probability space $(\Omega, \Sigma, \mathbb{P})$ let $(X_n)_{n\in\mathbb{N}}$ be a monotonically increasing sequence of non-negative random variables such that*

$$\lim_{n\to\infty} X_n = X$$

converges $\mathbb{P}$-almost surely. Then

$$\lim_{n\to\infty} \mathbb{E}\left[X_n\middle|\,\mathcal{F}\right] = \mathbb{E}\left[X\middle|\,\mathcal{F}\right]$$

$\mathbb{P}$-almost surely.

Proof. For any $A \in \mathcal{F}$ by monotonically of the conditional expectation and monotone convergence (see Theorem A.3.57)

$$\begin{aligned}\mathbb{E}\left[\lim_{n\to\infty}\mathbb{E}\left[X_n\middle|\,\mathcal{F}\right]1_A\right] &= \lim_{n\to\infty}\mathbb{E}\left[\mathbb{E}\left[X_n\middle|\,\mathcal{F}\right]1_A\right]\\ &= \lim_{n\to\infty}\mathbb{E}\left[X_n 1_A\right]\\ &= \mathbb{E}\left[X 1_A\right].\end{aligned}$$

$\square$

PROPOSITION A.3.66. *(dominated convergence,* **[30]**, *Proposition 8.14) On a probability space* $(\Omega, \Sigma, \mathbb{P})$ *let* $Y \in \mathcal{L}^1(\Omega, \Sigma, \mathbb{P})$ *be a positive random variable and* $(X_n)_{n\in\mathbb{N}}$ *a sequence of random variables such that* $|X_n| < Y$ *for all* $n \in \mathbb{N}$ *. If*

$$\lim_{n\to\infty} X_n = X,$$

$\mathbb{P}$*-almost surely then for sub* σ*-algebra* $\mathcal{F} \subset \Sigma$

$$\lim_{n\to\infty} \mathbb{E}[X_n | \mathcal{F}] = \mathbb{E}[X | \mathcal{F}]$$

holds true almost surely and in $\mathcal{L}^1(\Omega, \Sigma, \mathbb{P})$.

DEFINITION A.3.67. (Regular conditional probability) Let $(E, \mathcal{E})$ be a measurable space, $(\Omega, \mathcal{A}, \mathbb{P})$ a probability space, and $\mathcal{F} \subset \mathcal{A}$ a sub-σ-algebra. Let X, Y be a random variables on $(\Omega, \mathcal{A}, \mathbb{P})$ with values in $(E, \mathcal{E})$. If $\kappa_{X,\mathcal{F}}$ is a transition probability (see 2.1.1) from $(\Omega, \mathcal{F})$ to $(E, \mathcal{E})$ such that for any $F \in \mathcal{F}$ and any $B \in \mathcal{E}$

$$\mathbb{E}[1_B(X)1_F] = \int_\Omega \kappa_{X,\mathcal{F}}(\omega, B)1_F(\omega)\mathbb{P}(d\omega),$$

then $\kappa_{X,\mathcal{F}}$ is called *regular conditional probability*. Furthermore, with Lemma A.3.9 for any $B \in \mathcal{E}$ define the $\sigma(Y)$- measurable map

$$y \to \kappa_{X,Y}(y, B)$$

such that

$$\kappa_{X,Y}(Y(\omega), B) = \kappa_{X,\sigma(Y)}(\omega, B)$$

for any $\omega \in \Omega$.

PROPOSITION A.3.68. *(***[30]**, *Proposition 8.36) Let* $B \in \mathcal{B}(\mathbb{R})$ *be a Borel set and let* $(E, \mathcal{E})$ *be a measurable space that is isomorphic to* $(B, \mathcal{B}(B))$ *(see Definition A.3.23). Furthermore, let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space, and* $\mathcal{F} \subset \mathcal{A}$ *a sub-*σ*-algebra and let* X *be a random variable on* $(\Omega, \mathcal{A}, \mathbb{P})$ *with values in* $(E, \mathcal{E})$*.Then the Regular conditional probability* $\kappa_{X,\mathcal{F}}$ *exists.*

PROPOSITION A.3.69. *(***[30]**, *Proposition 8.37) Let* $B \in \mathcal{B}(\mathbb{R})$ *be a Borel set, let* $(E, \mathcal{E})$ *be a polish space, let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space, and* $\mathcal{F} \subset \mathcal{A}$ *a sub-*σ*-algebra and let* X *be a random variable on* $(\Omega, \mathcal{A}, \mathbb{P})$ *with values in* $(E, \mathcal{E})$*. Let* $f : E \to \mathbb{R}$ *be measurable and* $\mathbb{E}[|f(X)|] < \infty$*. Then, for* $\mathbb{P}$*-almost all* $\omega \in \Omega$

$$\mathbb{E}[f(X) | \mathcal{F}](\omega) = \int f(x)\kappa_{X,\mathcal{F}}(\omega, dx).$$

LEMMA A.3.70. *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space,* $(E, \mathcal{E})$ *be a measurable space and* X, Y *be independent* $(E, \mathcal{E})$*-valued random variables on* $(\Omega, \mathcal{A}, \mathbb{P})$*. Then for* $Z := X + Y$ *and any* $B \in \mathcal{E}$

$$\kappa_{Z,X}(x, B) = P_Y(B - x) = P_{Y+x}(B).$$

PROOF. Due to independence of X and Y, by Lemma A.3.44 for any $F \in \mathcal{E}$ and any $B \in \mathcal{E}$

$$\begin{aligned}\mathbb{E}\left[1_B(X+Y)1_F(X)\right] &= \int_E \left(\int_E 1_B(x+y)1_F(x)P_X(dx)\right) P_Y(dy) \\ &= \int_E \left(\int_E 1_{B-x}(y)P_Y(dy)\right) 1_F(x)P_X(dx)) \\ &= \int_E \left(P_Y(B-x)\right) 1_F(x)P_X(dx))\end{aligned}$$

□

DEFINITION A.3.71. (convex set)

In a $\mathbb{K}$-vector space V a subset $C \subset V$ is called *convex* if for any $\lambda \in [0, 1]$ and any $c_1, c_2 \in C$

$$\lambda c_1 + (1 - \lambda)\, c_2 \in C.$$

DEFINITION A.3.72. (convex map)

Let C be a convex set. Then a map $f : C \to \mathbb{R}$ is called *convex map* if for any $\lambda \in [0, 1]$ and any $c_1, c_2 \in C$

$$f\left(\lambda c_1 + (1 - \lambda)\, c_2\right) \leq \lambda f\left(c_1\right) + (1 - \lambda)\, f\left(c_2\right).$$

A map $f : C \to \mathbb{R}$ is called concave if $-f$ is convex.

THEOREM A.3.73. *(Jensen's Inequality) (*[**30**]*, Proposition 7.9)*

Let $I \subset \mathbb{R}$ *be an interval let* X *be a random variable that takes values in* I *and let* $\mathbb{E}[|X|] < \infty$*. If* φ *is convex, then* $\mathbb{E}[\varphi(X)^-] < \infty$ *and*

$$\varphi\left(\mathbb{E}[X]\right) \leq \mathbb{E}\left[\varphi(X)\right].$$

DEFINITION A.3.74. Let E be a metric space, $\mathcal{E}$ its Borel σ-algebra and $\mu, \mu_1, \mu_2, ...$ finite measures on $(E, \mathcal{E})$. The sequence $(\mu_n)_{n \in \mathbb{N}}$ is said to *converge weakly* to μ if for all continuous bounded functions $f \in C_b(E)$

$$\lim_{n\to\infty} \int_E f(x) d\mu_n(x) = \int_E f(x) d\mu(x).$$

A.3.3. Stochastic Processes.

DEFINITION A.3.75. Let $(\Omega, \mathcal{F})$ be a measurable space and let $I \subset \mathbb{R}$. Let $(\mathcal{F}_t)_{t\in I}$ be a family of σ-algebras such that for all $s, t \in I$ with $s \leq t$

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}.$$

Then $(\mathcal{F}_t)_{t\in I}$ is called *filtration*. If $(\Omega, \mathcal{F}, \mu)$ is a measure (probability) space, then $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in I}, \mu\right)$ is called *filtered measure (probability) space*. If

$$\mathcal{F}_s = \bigcap_{t>s,\, t\in I} \mathcal{F}_t$$

for all $s \in I$ then the filtration is called *right continuous*. If

$$\mathcal{F}_t = \sigma\left(\bigcup_{s<t,\, s\in I} \mathcal{F}_s\right)$$

for all $t \in I$ then the filtration is called *left continuous*.

DEFINITION A.3.76. Given a filtration $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ on a measurable space $(\Omega, \mathcal{F})$ by setting $\mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$ one can define the *right continuous enlargement* $(\mathcal{F}_{t+})_{t\in\mathbb{R}_+}$ of $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

DEFINITION A.3.77. (Completion of a filtration with respect to a family of measures)

Let $\left(\Omega, (\mathcal{G}_t)_{t\in\mathbb{R}_+}\right)$ be a filtered measurable space and

$$\mathcal{G}_\infty = \sigma\left(\bigcup_{t\in\mathbb{R}_+} \mathcal{G}_t\right).$$

Let $(\mu_\nu)_{\nu\in\mathcal{M}}$ be a family of measures on $(\Omega, \mathcal{G}_\infty)$. For every $\nu \in \mathcal{M}$ define $(\Omega, \mathcal{F}^\nu_\infty, \mu'_\nu)$ as the completion (see Definition A.3.32) of $(\Omega, \mathcal{G}_\infty, \mu_\nu)$ and set

$$\mathcal{F}_\infty := \bigcap_{\nu\in\mathcal{M}} \mathcal{F}^\nu_\infty.$$

Furthermore, call $\mathcal{N}^\nu$ the set of all μ'_ν-null sets on $\mathcal{F}^\nu_\infty$ and set

$$\mathcal{F}^\nu_t := \sigma\left(\mathcal{N}^\nu \cup \mathcal{G}_t\right)$$

$$\mathcal{F}_t := \bigcap_{\nu\in\mathcal{M}} \mathcal{F}^\nu_t.$$

Then we call $(\mathcal{F}_t^\nu)_{t\in\mathbb{R}_+}$ the *completion of the filtration* $(\mathcal{G}_t)_{t\in\mathbb{R}_+}$ *with respect to the measure* μ_ν and $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ the *completion of the filtration* $(\mathcal{G}_t)_{t\in\mathbb{R}_+}$ *with respect to the family of measures* $(\mu_\nu)_{\nu\in\mathcal{M}}$.

LEMMA A.3.78. *(*[**26**]*, Lemma 6.8) Let* $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mu\right)$ *be a filtered measure space. Then for* $(\mathcal{F}_t^\mu)_{t\in\mathbb{R}_+}$*, the completion of* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ *with respect to* μ*, and for* $((\mathcal{F}_{t+})^\mu)_{t\in\mathbb{R}_+}$*, the completion of* $(\mathcal{F}_{t+})_{t\in\mathbb{R}_+}$ *with respect to* μ*, for any* $t\in\mathbb{R}_+$

$$(\mathcal{F}_{t+})^\mu = \bigcap_{s>t} (\mathcal{F}_s^\mu)\,.$$

DEFINITION A.3.79. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $I \subset \mathbb{R}$ and $X = (X_t)_{t\in I}$ a family of random variables that take values in a measure space $(E, \mathcal{E})$ called *state space*. Then X is called *stochastic process*. For each fixed $\omega \in \Omega$ the map

$$\begin{aligned} I &\to E \\ t &\to X_t(\omega) \end{aligned}$$

is called *path*.

DEFINITION A.3.80. Let $I \subset \mathbb{R}$ and let $(X_t)_{t\in I}$ be a stochastic process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then the filtration $(\mathcal{F}_t)_{t\in I}$ defined as

$$\mathcal{F}_t := \sigma\left((X_s)_{s\le t,\, s\in I}\right)$$

is called *natural filtration*.

DEFINITION A.3.81.

A stochastic process $X = (X_t)_{t\in I}$ is said to have *independent increments* if for any $n\in\mathbb{N}$ and $0 < t_1 < t_2 < \ldots < t_n$ the family

$$\left(X_{t_i} - X_{t_{i-1}}\right)_{i\in\{1,\ldots,n\}}$$

is independent (see Definition A.3.35).

DEFINITION A.3.82. A real-valued stochastic process $X = (X_t)_{t\in I}$ is said to have *stationary increments* if for any $r, s_1, s_2 \in I$

$$X_{s_2+r} - X_{s_2} \sim X_{s_1+r} - X_{s_1}.$$

DEFINITION A.3.83. (Brownian motion, [**30**], Definition 21.8)

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space. A real-valued, adapted stochastic process $W = (W_t)_{t\in\mathbb{R}_+}$ is called *Brownian motion* if
(i) $W_0 = 0$,
(ii) W has stationary and independent increments (See Definition A.3.82 and A.3.81),
(iii) for any $t > 0$ $W_t \sim \mathcal{N}_{0,t}$ (see Example A.3.26),
(iv) the paths are continuous $\mathbb{P}$- almost surely.

An $\mathbb{R}^d$-valued adapted stochastic process $W = (W_t)_{t\in\mathbb{R}_+}$ is called *d-dimensional Browninan motion with initial distribution* μ if for any $B \in \mathcal{B}(\mathbb{R}^d)$

$$\mathbb{P}(W_0 \in B) = \mu(B),$$

for any $t > s > 0$ $W_t - W_s \sim \mathcal{N}_{0,t-s}$, and the other properties of the one-dimensional case hold accordingly.

THEOREM A.3.84. *(existence Brownian motion,* [**30**]*, Proposition 21.9)*

There exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *and a Brownian motion* $W = (W_t)_{t\in\mathbb{R}_+}$ *on it.*

LEMMA A.3.85. *For* $Z \sim \mathcal{N}(a, b)$

$$\mathbb{E}\left[\exp(Z)\right] = e^{b^2/2-a}$$

PROOF.

$$\begin{aligned}\mathbb{E}\left[\exp(Z)\right] &= \int_{-\infty}^{\infty} e^x e^{-\frac{(x-a)^2}{2b^2}}\,dx \\ &= \int_{-\infty}^{\infty} e^{-\frac{\left(x+b^2-a\right)^2-\left(b^2-a\right)^2+a^2}{2b^2}}\,dx \\ &= e^{b^2/2-a}.\end{aligned}$$

□

DEFINITION A.3.86. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $X = (X_t)_{t\in\mathbb{R}_+}$ and $Y = (Y_t)_{t\in\mathbb{R}_+}$ be stochastic processes. X and Y are called *modifications* if $\mathbb{P}(X_t = Y_t) = 1$ for any $t \in \mathbb{R}_+$ and X and Y are called *indistinguishable* if on a set $\Omega' \subset \Omega$ with $\mathbb{P}(\Omega') = 1$ the paths $t \to X_t(\omega)$ and $t \to Y_t(\omega)$ are equal.

Definition A.3.87. Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space. A stochastic process X is called *adapted* with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ if X_t is $\mathcal{F}_t$-measurable for any $t \in \mathbb{R}_+$.

Definition A.3.88. A stochastic process $(X_t)_{t\in\mathbb{R}_+}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ is called *measurable* if the map

$$\begin{aligned} \mathbb{R}_+ \times \Omega &\to E \\ (t, \omega) &\to X_t(\omega) \end{aligned}$$

is measurable with respect to the product σ-algebra $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{F}$.

Definition A.3.89. A stochastic process $(X_t)_{t\in\mathbb{R}_+}$ on the filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ is called *progressively measurable* with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ if for any $t \geq 0$

$$\begin{aligned} [0, t] \times \Omega &\to E \\ (s, \omega) &\to X_s(\omega) \end{aligned}$$

is $\mathcal{B}([0,t]) \otimes \mathcal{F}_t$-measurable.

Clearly, any progressively measurable stochastic process is measurable.

Proposition A.3.90. *(***[27]***Proposition 1.13) If a stochastic process X on the filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ is adapted and every path is left continuous or every path is right continuous, then X is progressively measurable with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.*

Definition A.3.91. Let $I \subset \mathbb{R}$ and let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in I}, \mathbb{P}\right)$ be a filtered probability space. A stochastic process $(X_t)_{t\in I}$ on $(\Omega, \mathcal{F})$ is called *submartingale with respect to* $(\mathcal{F}_t)_{t\in I}$ if it is adapted to $(\mathcal{F}_t)_{t\in I}$, if $X_t \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ for all $t \in I$ and if for any $s, t \in I$ such that $s > t$

$$\mathbb{E}\left[X_s \middle| \mathcal{F}_t\right] \geq X_t.$$

It is called *supermartingale with respect to* $(\mathcal{F}_t)_{t\in I}$ if instead the inequality

$$\mathbb{E}\left[X_s \middle| \mathcal{F}_t\right] \leq X_t,$$

holds and *martingale with respect to* $(\mathcal{F}_t)_{t\in I}$ if it is both supermartingale and submartingale with respect to $(\mathcal{F}_t)_{t\in I}$.

PROPOSITION A.3.92. *Let* $(W_t)_{t\in\mathbb{R}_+}$ *be the Brownian motion on the probability space* $(\Omega,\mathcal{F},\mathbb{P})$ *and* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ *be its natural filtration defined as* $\mathcal{F}_t := \sigma\left((W_s)_{0\le s\le t}\right)$. *Then*

(i)

$$(W_t)_{t\in\mathbb{R}_+}$$

is a martingale with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

(ii)

$$\left(\exp\left(\left(-\frac{\sigma^2}{2}\right)t+\sigma W_t\right)\right)_{t\in\mathbb{R}_+}$$

is a martingale with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

PROOF. (i) By construction $(W_t)_{t\in\mathbb{R}_+}$ is adapted, $(W_t)^2\in\mathcal{L}^1(\mathbb{P})$ implies $W_t\in\mathcal{L}^1(\mathbb{P})$ and for any $0\le s\le t$ independent increments of the Brownian motion imply that $\sigma(W_t-W_s)$ is independent of $\mathcal{F}_s$. Hence by Proposition A.3.63

$$\begin{aligned}\mathbb{E}\left[W_t\middle|\,\mathcal{F}_s\right] &= \mathbb{E}\left[W_t-W_s+W_s\middle|\,\mathcal{F}_s\right]\\ &= \mathbb{E}\left[W_t-W_s\right]+W_s\\ &= W_s.\end{aligned}$$

(ii) By Lemma A.3.85 for any $t\in\mathbb{R}_+$

$$\mathbb{E}\left[\left|\exp\left(\left(-\frac{\sigma^2}{2}\right)t+\sigma W_t\right)\right|\right]<\infty.$$

As argued in (i) for any $0\le s\le t$ $\sigma(W_t-W_s)$ is independent of $\mathcal{F}_s$ and by Proposition A.3.63

$$\mathbb{E}\left[\exp\left(\left(-\frac{\sigma^2}{2}\right)t+\sigma W_t\right)\middle|\mathcal{F}_s\right]=\mathbb{E}\left[\exp\left(\sigma\left(W_t-W_s\right)\right)\right]\exp\left(\left(-\frac{\sigma^2}{2}\right)t+\sigma W_s\right).$$

Thus, by Lemma A.3.85

$$\mathbb{E}\left[\exp\left(\sigma\left(W_t-W_s\right)\right)\right]=e^{\sigma^2(t-s)/2}$$

and

$$\left(\exp\left(\left(-\frac{\sigma^2}{2}\right)t+\sigma W_t\right)\right)_{t\in\mathbb{R}_+}$$

is a martingale with respect to $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. □

Theorem A.3.93. *([**35**], Theorem II.2.3) Let $I = \{0, -1, -2, ...\}$ and let $\left(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in I}, \mathbb{P}\right)$ be a filtered probability space. Let $(X_n)_{n\in I}$ be a submartingale with respect to $(\mathcal{F}_n)_{n\in I}$. Then*

$$\lim_{n\searrow -\infty} X_n$$

converges $\mathbb{P}$-almost surely. If additionally $\sup_{n\in I}\mathbb{E}\left[|X_n|\right] < \infty$, then $(X_n)_{n\in I}$ is uniformly integrable,

$$\lim_{n\searrow -\infty} X_n$$

converges in $L^1(\Omega, \mathcal{F}, \mathbb{P})$ and for every $m \in I$

$$\lim_{n\searrow -\infty} X_n \leq \mathbb{E}\left[X_m \middle| \mathcal{F}_{-\infty}\right]$$

with

$$\mathcal{F}_{-\infty} := \bigcap_{n\in I} \mathcal{F}_n.$$

Proposition A.3.94. *([**30**], Proposition 11.7) Let*

$$\left(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P}\right)$$

be a filtered probability space and let $(X_n)_{n\in\mathbb{N}}$ be a uniformly integrable (see Definition A.3.54) submartingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}}$. Then there exists a random variable X_∞ that is measurable with respect to $\mathcal{F}_\infty := \sigma\left(\bigcup_{n\in\mathbb{N}} \mathcal{F}_n\right)$ and

$$X_n \to X_\infty$$

$\mathbb{P}$-almost surely and with respect to $L^1(\Omega, \mathcal{F}, \mathbb{P})$. Furthermore,

$$X_n \leq \mathbb{E}\left[X_\infty \middle| \mathcal{F}_n\right]$$

for any $n \in \mathbb{N}$.

Proposition A.3.95. *([**35**], Corollary II.2.4) Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of random variables on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\lim_{t\to\infty} X_n$ converges $\mathbb{P}$-almost surely to the random variable X. Let Y be a random variable such that $\mathbb{E}\left[|Y|\right] < \infty$ and $|X_n| \leq Y$ for all $n \in \mathbb{N}$. Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a decreasing sequence of σ-algebras such that $\mathcal{F}_n \subset \mathcal{F}$ for any $n \in \mathbb{N}$. Then*

$$\mathbb{E}\left[X \middle| \bigcap_{n\in\mathbb{N}} \mathcal{F}_n\right] = \lim_{n\to\infty}\left[X_n \middle| \mathcal{F}_n\right]$$

$\mathbb{P}$-almost surely.

THEOREM A.3.96. *([**35**], Theorem II.2.5) Let* $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ *be a filtered probability space and let* $(X_t)_{t\in\mathbb{R}_+}$ *be a submartingale with respect to* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$*. Then* $\mathbb{P}$*-almost everywhere for any* $t \in \mathbb{R}_+$

$$\lim_{r\searrow t,\, r\in\mathbb{Q}} X_r$$

exists and for any $t > 0$

$$\lim_{r\nearrow t,\, r\in\mathbb{Q}} X_r$$

exists.

DEFINITION A.3.97. Let $(\Omega, \mathcal{F})$ be a measurable space and let $(X_t)_{t\in\mathbb{R}_+}$ be a stochastic process. For any $t \in \mathbb{R}_+$ set

$$X_{t+} := \limsup_{r\searrow t,\, r\in\mathbb{Q}} X_r,$$

and for any $t > 0$ set

$$X_{t-} := \lim_{r\nearrow t,\, r\in\mathbb{Q}} X_r.$$

PROPOSITION A.3.98. *([**35**], Proposition II.2.6) Let*

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

be a filtered probability space and let $(X_t)_{t\in\mathbb{R}_+}$ *be a submartingale with respect to* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$*. If* $\mathbb{E}\left[|X_t|\right] < \infty$ *for any* $t \in \mathbb{R}_+$*, then* $\mathbb{E}\left[|X_{t+}|\right] < \infty$ *for any* $t \in \mathbb{R}_+$*, and almost surely*

$$X_t \leq \mathbb{E}\left[X_{t+} \middle| \mathcal{F}_t\right].$$

Moreover, $(X_{t+})_{t\in\mathbb{R}_+}$ *is a submartingale with respect to* $(\mathcal{F}_{t+})_{t\in\mathbb{R}_+}$ *(see Definition A.3.76).*

If additionally, the map

$$t \to \mathbb{E}\left[X_t\right]$$

is right continuous, then

$$X_t = \mathbb{E}\left[X_{t+} \middle| \mathcal{F}_t\right].$$

If $(X_t)_{t\in\mathbb{R}_+}$ *is a martingale with respect to* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$*, then* $(X_{t+})_{t\in\mathbb{R}_+}$ *is a martingale with respect to* $(\mathcal{F}_{t+})_{t\in\mathbb{R}_+}$*.*

Proposition A.3.99. *([**35**],Proposition II.2.7) Let*

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

be a filtered probability space and let $(X_t)_{t\in\mathbb{R}_+}$ *be a submartingale with respect to* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. *If* $\mathbb{E}\left[|X_t|\right]<\infty$ *for any* $t\in\mathbb{R}_+$, *then* $\mathbb{E}\left[|X_{t-}|\right]<\infty$ *for any* $t\in\mathbb{R}_+$, *and almost surely*

$$X_{t-}\leq\mathbb{E}\left[X_t|\,\mathcal{F}_{t-}\right].$$

$(X_{t-})_{t\in\mathbb{R}_+}$ *is a submartingale with respect to* $(\mathcal{F}_{t-})_{t\in\mathbb{R}_+}$ *(see Definition A.3.76).*

If additionally, the map

$$t\to\mathbb{E}\left[X_t\right]$$

is left continuous, then

$$X_{t-}=\mathbb{E}\left[X_t|\,\mathcal{F}_{t-}\right].$$

If $(X_t)_{t\in\mathbb{R}_+}$ *is a martingale with respect to* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$, *then* $(X_{t-})_{t\in\mathbb{R}_+}$ *is a martingale with respect to* $(\mathcal{F}_{t-})_{t\in\mathbb{R}_+}$.

Proposition A.3.100. *([**35**],Theorem II.2.9) Let* $(\Omega,\mathcal{F},\mathbb{P})$ *be a filtered probability space and let* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ *be a right continuous and complete filtration on it. Let* $(X_t)_{t\in\mathbb{R}}$ *be a submartingale with respect to* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$. *If*

$$t\to\mathbb{E}\left[X_t\right]$$

is right-continuous, then $(X_t)_{t\in\mathbb{R}}$ *has a càdlàg version that is a submartingale with respect to* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$.

Definition A.3.101. Let $\left(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in\mathbb{R}_+},\mathbb{P}\right)$ be a filtered probability space. A random variable $\tau:\,\Omega\to\mathbb{R}_+$ is called *stopping time* if for any $t\in\mathbb{R}_+$

$$\{\tau\leq t\}\in\mathcal{F}_t.$$

The σ-algebra

$$\mathcal{F}_\tau:=\{A\in\mathcal{F}:\;A\cap\{\tau\leq t\}\in\mathcal{F}_t\,\text{for all }t\in\mathbb{R}_+\}$$

is called the *σ-algebra of events determined prior to the stopping time* τ.

DEFINITION. Let T be an index set, and $(\Omega_t, \mathcal{F}_t)$ be measurable spaces for all $t \in T$. A family of probability measures $(P_F)_{F\subset T,\text{ finite}}$ such that any P_F for $F \subset T$ finite is defined on the measurable space

$$\left(\bigtimes_{t\in F}\Omega_t, \bigotimes_{t\in F}\mathcal{F}_t\right)$$

is called *projective family* if for all finite $K \subset L \subset T$

$$P_L \circ \left(\Pi_K^L\right)^{-1} = P_K,$$

where

$$\Pi_K^L : \left(\bigtimes_{t\in L}\Omega_t, \bigotimes_{t\in L}\mathcal{F}_t\right) \to \left(\bigtimes_{t\in K}\Omega_t, \bigotimes_{t\in K}\mathcal{F}_t\right)$$

is the projection from Definition A.3.5.

THEOREM A.3.102. *(Kolmogorov extension theorem,* **[34]***, Theorem 2.19)*

Let $T \neq \emptyset$ be an index set, for any $t \in T$ let Ω_t be a polish space, and let $(P_F)_{F\subset T,\text{ finite}}$ be a projective family of probability measures on

$$\left(\bigtimes_{t\in F}\Omega_t, \bigotimes_{t\in F}\mathcal{B}\left(\Omega_t\right)\right).$$

Then there exists a unique probability measure $\mathbb{P}$on

$$\left(\bigtimes_{t\in T}\Omega_i, \bigotimes_{t\in T}\mathcal{B}\left(\Omega_t\right)\right)$$

such that for all $F \subset T,$ finite and

$$A \in \bigotimes_{t\in F}\mathcal{B}\left(\Omega_t\right)$$

$$\mathbb{P}\left(\left(\Pi_F^T\right)^{-1}(A)\right) = P_F(A),$$

where Π_F^T is the projection from Definition A.3.5.

Theorem A.3.102 can be generalized using *compact classes*.

Definition A.3.103. A family $\mathcal{C}$ of subsets of a space X is called *compact class* if for any sequence $(C_n)_{n\in\mathbb{N}} \subset \mathcal{C}$ such that the intersection $\bigcap_{n\in\mathbb{N}} C_n$ is empty, already some finite intersection $\bigcap_{i\in I,\text{ finite}} C_i$ is empty.

Theorem A.3.104. *(Generalized Kolmogorov Extension Theorem,* [**1**]*, Theorem 15.26)*

Let $T \neq \emptyset$ be an index set and let

$$(\Omega_t, \Sigma_t)_{t\in T}$$

be a family of measurable spaces and for each finite subset $F \subset T$ let P_F be a probability measure on

$$\Omega_F = \bigtimes_{t\in F} \Omega_t$$

equipped with the product σ-algebra

$$\Sigma_F = \bigotimes_{t\in F} \Sigma_t.$$

If $\{P_F\}_{F\subset T}$ is a projective family of probability measures and if for each $t \in T$ there is a compact class (see Definition A.3.103) $\mathcal{C}_t \in \Sigma_t$ such that for each $A\in\Sigma_t$

$$P_t(A) = \sup\{P_t(C) : C \subset A \text{ and } C \in \mathcal{C}_t\},$$

then there is a unique probability measure $\mathbb{P}$ on

$$\Omega_T = \bigtimes_{t\in T} \Omega_t$$

and

$$\Sigma_T = \bigotimes_{t\in T} \Sigma_t.$$

such that for all $F \subset T$, finite and $A \in \Sigma_T$

$$\mathbb{P}\left(\left(\Pi_F^T\right)^{-1}(A)\right) = P_F(A),$$

where Π_J^T is the projection from Definition A.3.5.

A.3.4. Stochastic integration. Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space.

Definition A.3.105. The space of $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$-martingales $(M_t)_{t\in\mathbb{R}_+}$ at $M_0 = 0$ such that $M_t \in L^2$ for any $t \in \mathbb{R}_+$ is denoted by $\mathscr{M}$. If additionally all paths are continuous the set is denoted by$\mathscr{M}$. For $M \in \mathscr{M}_2$ we set $\|M\|_{\mathscr{M}_2} := \sum_{n=1}^{\infty} 2^{-n}\left(\|M_n\|_{L^2} \wedge 1\right)$ and define the metric (when identifying indistinguishable processes) $d(M,N) := \|M-N\|_{\mathscr{M}_2}$ on $\mathscr{M}$.

THEOREM A.3.106. *([27], Chapter I, Theorem 5.23)$\mathscr{M}$ is a complete metric space with respect to $\|\cdot\|_{\mathscr{M}_2}$ when identifying indistinguishable processes and $\mathscr{M}$ is a closed subspace of $\mathscr{M}$.*

DEFINITION A.3.107. A stochastic process $X = (X_t)_{t\in\mathbb{R}_+}$ is called *simple* if there is a sequence of real numbers $(t_n)_{n\in\mathbb{N}}$, $t_0 = 0$ such that $\lim_{n\to\infty} t_n$ and a sequence $(\xi_n)_{n\in\mathbb{N}}$ of random variables such that for any $n \in \mathbb{N}$ ξ_n is $\mathcal{F}_{t_n}$-measurable and such that there is a constant $C > 0$ and for any $\omega \in \Omega$

$$\sup_{n\in\mathbb{N}} |\xi_n(\omega)| \le C.$$

The set of simple process is denoted by $\mathscr{L}_0$.

DEFINITION A.3.108. For $M \in \mathscr{M}$ and $X \in \mathscr{L}_0$ and $0 \le t < \infty$ define

$$\int_0^t X_s dM_s := \sum_{i=0}^{\infty} X_{t_i} \left(M_{t_{i+1}\wedge t} - M_{t_i\wedge t} \right).$$

THEOREM A.3.109. *([27], Chapter IV, Theorem 1.8) If M is a local martingale (see Definition 3.0.2) there is a continuous adapted, increasing process$\langle M\rangle$ starting at $\langle M\rangle_0 = 0$ such that $M^2 - \langle M\rangle$ is a martingale. It is unique up to indistinguishability.*

DEFINITION A.3.110. Let $M \in \mathscr{M}$. $\mathscr{L}$ denotes the set of all $\mathcal{F}_t$ -adapted measurable stochastic processes such that for all $t > 0$

$$\mathbb{E}\left[\int_0^t X_s^2 d\langle M\rangle_s\right] < \infty.$$

The set of all elements of $\mathscr{L}$ that are progressively measurable (see Definition A.3.89) is denoted by $\mathscr{L}^*$. On $\mathscr{L}$

$$d(X,Y) := \sum_{n=1}^{\infty} 2^{-n} \left(\mathbb{E}\left[\int_0^n (X_s - Y_s)^2 d\langle M\rangle_s\right] \wedge 1 \right)$$

defines a metric.

THEOREM A.3.111. *([27], Chapter IV, Proposition 1.22) Denote by $\mathscr{L}_\infty^*$ the set of elements of $\mathscr{L}^*$ such that*

$$\mathbb{E}\left[\int_0^\infty X_s^2 d\langle M\rangle_s\right] < \infty.$$

Then $\mathscr{L}_\infty^*$ is a Hilbert space with respect to the scalar product $\langle X, Y\rangle := \mathbb{E}\left[\int_0^\infty X_s Y_s d\langle M\rangle_s\right]$.

Proposition A.3.112. *(**[27]**, Chapter IV, Proposition 2.8) $\mathscr{L}_0$ is dense in $\mathscr{L}^*$ with respect to the metric from Definition A.3.110.*

Proposition A.3.113. *(**[27]**, Chapter IV, Equation 2.14) (Ito isometry) For $X \in \mathscr{L}_0$ and $M \in \mathscr{M}$*

$$\left(\int_0^t X_s dM_s\right)^2 = \mathbb{E}\left[\int_0^t X_s^2 d\left\langle M\right\rangle_s\right].$$

Using density of $\mathscr{L}_0$ in $\mathscr{L}^*$ and Ito isometry, we can define the stochastic integral for all elements in $\mathscr{L}^*$:

Definition A.3.114. (**[27]**,Chapter IV, Definition 2.9)

For $X \in \mathscr{L}^*$ the *stochastic integral of X with respect to $M \in \mathscr{M}$* is the unique square integrable martingale N such that for every sequence $(X_n)_{n\in\mathbb{N}} \subset \mathscr{L}_0$ such that

$$\lim_{n\to\infty} d(X_n, X) = 0$$

also

$$\lim_{n\to\infty} \left\|\left(\int_0^t (X_n)_s\, dM_s\right)_{t\in\mathbb{R}_+} - N\right\|_{\mathscr{M}} = 0.$$

Such $N \in M_2^c$ is denoted by $\left(\int_0^t X_s dM_s\right)_{t\in\mathbb{R}_+}$.

Theorem A.3.115. *(**[27]**,Chapter IV, Definition 3.6) (Ito formula)*

Let X be a real-valued stochastic process such that it has $\mathbb{P}$-almost surely the decomposition

$$X_t = X_0 + M_t + B_t,$$

where $M = (M_t)_{t\in\mathbb{R}_+} \in \mathscr{M}$, and $B = (B_t)_{t\in\mathbb{R}_+}$ is the difference of continuous non decreasing adapted processes starting at 0. Let $f : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ be continuously differentiable in the first, and twice continuously differentiable in the second variable. Then $\mathbb{P}$-almost surely

$$\begin{aligned} f(t, X_t) = f(0, X_0) &+ \int_0^t \frac{\partial f}{\partial t}(s, X_s)ds + \int_0^t \frac{\partial f}{\partial x^i}(s, X_s)dB_s \\ &+ \int_0^t \frac{\partial f}{\partial x^i}(s, X_s)dM_s \\ &+ \frac{1}{2}\int_0^t \frac{\partial^2 f}{(\partial x)^2}(s, X_s)d\left\langle M\right\rangle_s. \end{aligned}$$

Definition A.3.116. Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space. Let $W = \left(W_t^1, .., W_t^d\right)_{t\in\mathbb{R}_+}$ be the d-dimensional Brownian motion (see Definition A.3.83) and $\mu = \left(\mu^1, ..., \mu^d\right) : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma = \left(\sigma^{i,j}\right)_{i,j\in\{1,...,d\}} : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$ be measurable maps such that stochastic integrals below exist. A $\mathbb{R}^d$-valued stochastic process $X = \left(X_t^1, ..., X_t^d\right)_{t\in\mathbb{R}_+}$ that satisfies the integral equations

$$X_t^1 = X_0^1 + \int_0^t \mu^1(s, X_s)ds + \int_0^t \sigma^{1,1}(s, X_s)dW_s^1 + ... + \int_0^t \sigma^{1,d}(s, X_s)dW_s^d$$
$$\vdots$$
$$X_t^d = X_0^d + \int_0^t \mu^d(s, X_s)ds + \int_0^t \sigma^{d,1}(s, X_s)dW_s^1 + ... + \int_0^t \sigma^{d,d}(s, X_s)dW_s^d$$

for any $t \in \mathbb{R}_+$ is said to satisfy the *stochastic differential equation*

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t.$$

Proposition A.3.117. *([**5**], Proposition 5.12) (Kolmogorov forward equation/ Fokker-Planck equation)*

Let $X = (X_t)_{t\in\mathbb{R}_+}$ *be the solution of the stochastic differential equation*

$$\begin{aligned} dX_t &= \mu(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_s &= y, \end{aligned}$$

and let $\mu : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ *and* $\sigma : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ *be sufficiently smooth. If for any* $t \in \mathbb{R}_+$ *the distribution of* $\mathbb{P}(X_t \leq x)$ *is given by a density* $f(s, y; t, x)$ *then*

$$\begin{aligned} \frac{\partial}{\partial t} f(s, y; t, x) &= \left(\mathcal{A}^* f\right)(s, y; t, x) \text{ for all } (t, x) \in (s, T) \times \mathbb{R} \\ f(s, y; t, x) &\to \delta_y \text{ as } t \searrow s. \end{aligned}$$

$\mathcal{A}^*$ *is defined by*

$$\left(\mathcal{A}^* f\right)(t, x) = -\frac{\partial}{\partial x}\left(\mu(t, x)f(t, x)\right) + \frac{1}{2}\frac{\partial^2}{(\partial x)^2}\left((\sigma(t, x))^2 f(t, x)\right).$$

Theorem A.3.118. *(Girsanov, [**27**], Theorem 5.1)*

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ *be a filtered probability space and let* $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$ *be right continuous (see Definition A.3.75) and let* $\mathcal{F}_0$ *be complete (see Definition A.3.32). Let* $W = \left(W_t^1, .., W_t^d\right)_{t\in\mathbb{R}_+}$ *be the d-dimensional Brownian motion such that* $\mathbb{P}(W_0 = 0) = 1$*. Let* $a = \left(a_t^1, .., a_t^d\right)_{t\in\mathbb{R}_+}$ *be* *a vector of adapted stochastic processes such that for any* $1 \leq i \leq d$ *and any* $0 \leq T < \infty$

$$\mathbb{P}\left(\int_0^T \left(a_t^i\right)^2 dt < \infty\right) = 1.$$

If $(Z_t)_{t\in\mathbb{R}_+}$ *defined by*

$$Z_t := \exp\left(\sum_{i=1}^d \int_0^t a_s^i dW_s^i - \frac{1}{2}\int_0^t \left\|a_s^i\right\|^2 ds\right)$$

is a martingale, then for each $0 \leq T < \infty$ *the* d*-dimensional stochastic process* $\widetilde{W} = \left(\widetilde{W_t^1}, .., \widetilde{W_t^d},\right)_{t\in[0,T]}$ *defined for any* $1 \leq i \leq d$ *by*

$$\widetilde{W_t^i} = W_t^i - \int_0^t a_t^i ds$$

is a d-dimensional Brownian motion on $\left(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, \tilde{\mathbb{P}}_T\right)$ where $\tilde{\mathbb{P}}_T$ is defined by

$$\tilde{\mathbb{P}}_T(A) := \mathbb{E}\left[1_A Z_T\right]$$

for any $A \in \mathcal{F}_T$.

A.4. Functional Analysis

Example A.4.1. Let (Ω, Σ, μ) be a measure space. For $1 \leq p < \infty$ we define the space

$$\mathcal{L}^p(\Omega, \Sigma, \mu) := \left\{ f : \Omega \to \mathbb{C} : \begin{array}{c} \text{measureable,} \\ \int_\Omega |f(x)|^p d\mu < \infty. \end{array} \right\}$$

It can be shown (see for example **[39]**, Chapter 3 or **[41]**, Chapter I) that $\mathcal{L}^p(\Omega, \Sigma, \mu)$ is a vector space and that the map

$$\|\cdot\|^*_{L^p(\Omega,\Sigma,\mu)} : \ f \to \left(\int_\Omega |f(s)|^p d\mu\right)^{1/p}$$

is a seminorm on $\mathcal{L}^p$ (see 1.4.46). For any $f \in \mathcal{L}^p(\mathbb{R})$ we define the sets

$$[f] := \left\{ g \in \mathcal{L}^p(\Omega, \Sigma, \mu) : \ \|g - f\|^*_{L^p(\Omega,\Sigma,\mu)} = 0 \right\}$$

and the set of these sets

$$L^p(\Omega, \Sigma, \mu) := \{[f] : \ f \in \mathcal{L}^p(\Omega, \Sigma, \mu)\}.$$

Then the map $\|\cdot\|_{L^p(\mathbb{R})}$ defined as

$$\|[f]\|_{L^p(\mathbb{R})} := \|f\|^*_{L^p(\mathbb{R})}$$

is a norm on the space $L^p(\Omega, \Sigma, \mu)$ which is a vector space. It can be shown (see [**38**], Chapter 3 or [**41**], Chapter I) that with respect to this norm $L^p(\Omega, \Sigma, \mu)$ is also complete, thus a Banach space.

EXAMPLE A.4.2. Let (Ω, Σ, μ) be a measure space. Define

$$\mathcal{L}^\infty(\Omega, \Sigma, \mu) := \left\{ f : \Omega \to \mathbb{C} : \begin{array}{c} f \text{ is measurable,} \\ \text{there is } C < \infty \text{ such that} \\ |f| < C \text{ almost surely.} \end{array} \right\}$$

(for the definition of almost surely see Definition A.3.32). On $\mathcal{L}^\infty(\Omega, \Sigma, \mu)$ we define the map $\|\cdot\|^*_{L^\infty(\Omega,\Sigma,\mu)}$ via

$$\|f\|^*_{L^\infty(\Omega,\Sigma,\mu))} = \inf\{C \in \mathbb{R}_+ : \ |f| < C \text{ almost surely}\}.$$

It can be shown (see for example [**41**], Chapter I) that $\mathcal{L}^\infty(\Omega, \Sigma, \mu)$ is a vector space and that the map $\|\cdot\|^*_{L^\infty(\Omega,\Sigma,\mu)}$ is a seminorm on this vector space. If we introduce the sets

$$[f] := \left\{ g \in \mathcal{L}^\infty(\Omega, \Sigma, \mu) : \ \|g - f\|^*_{\mathcal{L}^\infty(\Omega,\Sigma,\mu)} = 0 \right\}$$

and define

$$L^\infty(\Omega, \Sigma, \mu) := \{[f] : \ f \in \mathcal{L}^\infty(\Omega, \Sigma, \mu)\}$$

then it can be proved that the map $\|\cdot\|_{L^\infty(\Omega,\Sigma,\mu)}$ defined as

$$\|[f]\|_{L^\infty(\Omega,\Sigma,\mu)} := \|f\|^*_{L^\infty(\Omega,\Sigma,\mu)}$$

is a norm on $L^\infty(\Omega, \Sigma, \mu)$ and that $L^\infty(\Omega, \Sigma, \mu)$ is complete, hence a Banach space.

PROPOSITION A.4.3. *([**41**], Proposition II.2.4) Let $1 \le p < \infty$ and let q be defined by $\frac{1}{p} + \frac{1}{q} = 1$. Let (Ω, Σ, μ) be a σ-finite measure space (see Definition A.3.20). Then the map*

$$\begin{aligned} L^q &\to (L^p)' \\ f &\to Tf, \end{aligned}$$

defined by

$$Tf(g) = \int_\Omega g f d\mu$$

is an isometric isomorphism.

Following Chapter I in [**41**], and recalling that for a Banach space Z and some set T

$$\ell^\infty(T;Z) := \left\{ f : T \to Z : \sup_{x\in T} \|f(x)\| < \infty \right\}$$

we can show:

PROPOSITION A.4.4. *$\ell^\infty(T;Z)$ is a vector space, the map*

$$\|\cdot\|_\infty : f \to \sup_{x\in T} \|f(x)\|$$

is a norm on $\ell^\infty(T;Z)$, and with respect to this norm $\ell^\infty(T;Z)$ is a Banach space.

PROOF. That $\ell^\infty(T;Z)$ is a vector space follows easily because for $f, g \in \ell^\infty(T;Z)$

$$\sup_{x\in X} \|f(x) + g(x)\|_\infty \leq \sup_{x\in X} \|f(x)\|_\infty + \sup_{x\in X} \|g(x)\|_\infty < \infty$$

and for $\lambda \in \mathbb{K}$

$$\sup_{x\in X} \|\lambda f(x)\|_\infty = \lambda \sup_{x\in X} \|f(x)\|_\infty .$$

These expressions imply also that $\|\cdot\|_\infty$ is a norm since $\|0\|_\infty = 0$ and $\|f\|_\infty = 0$ clearly yields that $f = 0$.

In order to show that $\ell^\infty(T;Z)$ is a Banach space we need to show that any Cauchy sequence $(f_n)_{n\in\mathbb{N}} \subset \ell^\infty(T;Z)$ converges to some $f \in \ell^\infty(T;Z)$ as n tends to infinity. In the proof we will use that pointwise each Cauchy sequence $(f_n(x))_{n\in\mathbb{N}} \subset Z$ converges.

Let $\varepsilon > 0$ be arbitrary. For $\varepsilon/2$ there is $N_{\varepsilon/2}$ such that for all $m, n > N_{\varepsilon/2}$

$$\|f_n - f_m\|_\infty < \varepsilon/2.$$

Hence for any $x \in T$ and for $m, n > N_{\varepsilon/2}$

$$\|f_n(x) - f_m(x)\| < \varepsilon/2.$$

Thus $(f_n(x))_{n\in\mathbb{N}} \subset Z$ is a Cauchy sequence for all $x \in Z$ and by assumption of Z being a Banach space there exists $f(x) \in Z$ such that $\lim_{n\to\infty} f_n(x) = f(x)$ in Z. Having shown pointwise convergence we still need to show that $f_n \to f$ in $\ell^\infty(T;Z)$. Since for all $m, n > N_{\varepsilon/2}$

$$\|f_n(x) - f_m(x)\| < \varepsilon/2.$$

and for any x there is $N_{x,\varepsilon/2}$ such that for all $m > N_{x,\varepsilon/2}$

$$\|f_m(x) - f(x)\| < \varepsilon/2$$

we conclude that

$$\begin{aligned}\|f_n(x) - f(x)\| &\leq \|f_m(x) - f(x)\| + \|f_n(x) - f_m(x)\| \\ &< \varepsilon/2 + \varepsilon/2\end{aligned}$$

for all $x \in X$. Hence $\ell^\infty(T; Z)$ is a Banach space. □

LEMMA A.4.5. *Let X be a topological space and $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$. The space $C_0(X, \mathbb{K})$ equipped with the norm*

$$\|\cdot\|_\infty : f \to \sup_{x \in X} |f(x)|.$$

is a Banach space.

PROOF. For the sequence $(f_n)_{n\in\mathbb{N}} \subset C_0(X, \mathbb{K})$ its limit

$$\lim_{n\to\infty} f_n$$

is continuous as the uniform limit of continuous functions and for any $\varepsilon > 0$ there is $n_0 \in \mathbb{N}$ such that

$$\left\{x \in X : \left|\lim_{n\to\infty} f_n(x)\right| \geq \varepsilon\right\} \subset \{x \in X : |f_{n_0}(x)| \geq \varepsilon/2\}.$$

Therefore, for any $\varepsilon > 0$

$$\left\{x \in X : \left|\lim_{n\to\infty} f_n(x)\right| \geq \varepsilon\right\}$$

is a closed subset of a compact set, hence compact and $C_0(X, \mathbb{K})$ is a closed subspace of

$$\ell^\infty(X; \mathbb{K}) := \left\{f : X \to \mathbb{K} : \sup_{x\in X} |f(x)| < \infty\right\}$$

which is a Banach space by Proposition A.4.4. Thus, $C_0(X, \mathbb{K})$ is a Banach space as well (by Lemma 2.3.26). □

THEOREM A.4.6. *(Hahn-Banach) ([**41**], Theorem III.1.5) Let X be a normed vector space and $U \subset X$ a vector subspace. To any continuous linear map $u' : U \to \mathbb{K}$ there is a continuous linear map $x' : X \to \mathbb{K}$ such that*

$$x'|_U = u'$$

and

$$\|x'\| = \|u'\|.$$

COROLLARY A.4.7. *Let X be a normed vector space and $x \in X$, $x \neq 0$. Then there is $x' \in X'$ such that $\|x'\| = 1$ and $x'(x) = \|x\|$.*

PROOF. The linear span

$$\operatorname{lin}\{x\} := \{\lambda x \mid \lambda \in \mathbb{K}\}$$

is vector subspace of X and $u' : \operatorname{lin}\{x\} \to \mathbb{K}$ defined as

$$u'(\lambda x) := \lambda \|x\| \text{ for } \lambda \in \mathbb{K}$$

is a continuous linear map and $\|u'\| = 1$. By the Hahn-Banach theorem there is a continuous linear extension $x' : X \to \mathbb{K}$ of u' such that $\|x'\| = 1$ and $x'(x) = \|x\|$. □

COROLLARY A.4.8. *Let X be a normed vector space. Then for all $x \in X$*

$$\|x\| = \sup_{\substack{x' \in X' \\ \|x'\| \leq 1}} |x'(x)|$$

PROOF. By Corollary A.4.7

$$\|x\| \leq \sup_{\substack{x' \in X' \\ \|x'\| \leq 1}} |x'(x)| .$$

On the other hand, by the definition of the norm of linear maps

$$\sup_{\substack{x' \in X' \\ \|x'\| \leq 1}} |x'(x)| \leq \|x'\| \|x\| = \|x\| .$$

□

THEOREM A.4.9. *(Open mapping theorem) ([**41**], Theorem IV.3.3) Let X and Y be Banach spaces and $L : X \to Y$ be a linear bounded surjective operator. Then L maps open sets to open sets.*

PROPOSITION A.4.10. *([**41**], Proposition III.3.8) Let X be a Banach space. For convex sets in X the weak and the norm closure coincide.*

THEOREM A.4.11. *([**38**] Theorem 3.27) Let X be a Banach space, K be a compact Hausdorff space and μ be a probability measure on the*

Borel σ-algebra of K. If $f : K \to X$ is continuous and $\overline{co} f(K)$ is compact in X then

$$\int_K f d\mu = y$$

exists in the sense of Definition 1.4.67 and $y \in \overline{co} f(K)$.

THEOREM A.4.12. *(Krein-Šmulian weak compactness theorem,* **[31]** *Theorem 2.8.14) The closed convex hull of a weakly compact subset of a Banach space is itself weakly compact.*

THEOREM A.4.13. *(Riesz representation theorem,* **[39]**, *Theorem 6.19) Let X be a locally compact Hausdorff space and let Φ be a complex valued bounded linear operator on $C_0(X, \mathbb{C})$. Then there is a unique regular (Definition 2.3.35) complex measure μ on $(X, \mathcal{B}(X))$ such that*

$$\Phi f = \int_X f d\mu.$$

Additionally, for the total variation $|\mu|$ (see Definition 2.3.31) of the complex measure μ

$$\|\Phi\| = |\mu|(X).$$

PROPOSITION A.4.14. *(***[6]**, *§5, Proposition 5) Let X be a completely regular space and $\ell : C_b(X, \mathbb{C}) \to \mathbb{C}$ be a continuous linear map . There exists a complex Radon measure μ on X such that for all $f \in C_b(X, \mathbb{C})$*

$$\ell(f) = \int_X f(x)\mu(dx),$$

if and only if for each $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset X$ such that for any function $f \in C_b(X, \mathbb{C})$ with $|f| \leq 1$ and $f|_{K_\varepsilon} = 0$

$$|\ell(f)| < \varepsilon$$

holds. The complex Radon measure is unique.

Making slight adjustments in the proof in (**[6]**, §5, Proposition 5) one obtains also a version of the above proposition, that holds on $C_b(X, \mathbb{R})$:

Proposition A.4.15. *Let X be a completely regular space and $\ell : C_b(X,\mathbb{R}) \to \mathbb{R}$ be a continuous linear map . There exists a signed Radon measure μ on X such that for all $f \in C_b(X,\mathbb{R})$*

$$\ell(f) = \int_X f(x)\mu(dx),$$

if and only if for each $\varepsilon > 0$ there exists a compact set $K_\varepsilon \subset X$ such that for any function $f \in C_b(X,\mathbb{R})$ with $|f| \le 1$ and $f|_{K_\varepsilon} = 0$

$$|\ell(f)| < \varepsilon$$

holds. The signed Radon measure is unique.

Proposition A.4.16. *([**6**], §5, Proposition 1b) Let X be a completely regular Hausdorff space. Let μ be a complex Radon measure on X and $f : X \to \mathbb{R}_+$ a lower semicontinuous map. Then*

$$\int_X f(x)\,|\mu|\,(dx) = \sup_g \left| \int_X g(x)\mu(dx) \right|,$$

where the supremum is taken over all functions $g \in C_b(X,\mathbb{C})$ such that $|g| \le f$, and g is $|\mu|$-integrable.

Noting that for a signed measure μ, a compact set K and $f \in C(K,\mathbb{R}_+)$ the identity

$$\int_K f(x)\,|\mu|\,(dx) = \sup_{\substack{|g| \le f, \\ g \in C(K,\mathbb{R})}} \int_K g(x)\mu(dx)$$

holds (see[**8**], Chapter III, §1, n.6, Equation 9), making slight adjustments in the proof of [**6**], §5, Proposition 1b one obtains the real version of the preceding proposition:

Proposition A.4.17. *Let X be a completely regular Hausdorff space. Let μ be a signed Radon measure on X and $f : X \to \mathbb{R}_+$ a lower semicontinuous map. Then*

$$\int_X f(x)\,|\mu|\,(dx) = \sup_g \int_X g(x)\mu(dx),$$

where the supremum is taken over all functions $g \in C_b(X,\mathbb{R})$ such that $|g| \le f$, and g is $|\mu|$-integrable.

Definition A.4.18. A Banach algebra A is a Banach space that is an algebra (see Definition A.2.5) such that for any $x, y \in A$

$$\|x \cdot y\| \leq \|x\| \cdot \|y\| .$$

If there is $e \in A$ such that for any $x \in A$

$$ex = xe = x$$

and $\|e\| = 1$, then e is called *unity*.

Theorem A.4.19. *([10], Theorem VII.2.2) For a Banach algebra with unity the set of invertible elements G is open and the map*

$$\begin{aligned} G &\to G \\ x &\to x^{-1} \end{aligned}$$

is continuous.

A.5. More Semigroups

Theorem A.5.1. *(Post-Widder Inversion Formula)([18], Corollary III.5.5)*

For every strongly continuous semigroup $(T(t))_{t\in\mathbb{R}_+}$ on X with generator $(A, \mathcal{D}(A))$ one has for all $x \in X$

$$T(t)x = \lim_{n\to\infty} \left[\frac{n}{t} R\left(\frac{n}{t}, A\right)\right]^n x = \lim_{n\to\infty} \left[I - \frac{t}{n} A\right]^{-n} x$$

uniformly for t in compact intervals.

A.6. Semimartingales

This subsection of the appendix is entirely taken from [24].

Definition A.6.1. Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space with right continuous filtration and let $d \in \mathbb{N}$. The σ-algebra $\mathscr{P}$ on $\Omega \times \mathbb{R}_+$ that is generated by all $\mathbb{R}^d$-valued adapted stochastic processes (as mappings on $\Omega \times \mathbb{R}_+$) with left continuous paths is called *predictableσ-algebra*. A $\mathbb{R}^d$-valued stochastic process $(X_t)_{t\in\mathbb{R}_+}$ is also called *predictable* if the map

$$\begin{aligned} \Omega \times \mathbb{R}_+ : &\to \mathbb{R}^d \\ (\omega, t) &\to X_t(\omega) \end{aligned}$$

is measurable with respect to $\mathscr{P}$.

Definition A.6.2. Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space with right continuous filtration . The σ-algebra $\mathscr{O}$ on $\Omega \times \mathbb{R}_+$ that is generated by all $\mathbb{R}^d$-valued adapted càdlàg (see Definition 2.2.5) stochastic processes (as mappings on $\Omega \times \mathbb{R}_+$) is called *optional$\sigma$-algebra*. A $\mathbb{R}^d$-valued stochastic process $(X_t)_{t\in\mathbb{R}_+}$ is called *optional* if the map

$$\begin{aligned}\Omega \times \mathbb{R}_+ :&\to \mathbb{R}^d\\ (\omega, t) &\to X_t(\omega)\end{aligned}$$

is measurable with respect to $\mathscr{O}$.

Proposition A.6.3. *([**24**], Proposition I.1.24)*

$$\mathscr{P} \subset \mathscr{O}.$$

Corollary A.6.4. *Due to Proposition A.3.90*

$$\mathscr{O} \subset \mathscr{F} \otimes \mathcal{B}(\mathbb{R}_+).$$

Lemma A.6.5. *([**33**], Lemma 13.12) Let E be a Polish space and for $T > 0$*

$$f : [0,T] :\to E$$

be a càdlàg function. Then, for $n \in \mathbb{N}$

$$\#\left\{t \in (0,T] |\, \|\Delta f(t)\| > \frac{1}{n}\right\} < \infty.$$

Hence, a càdlàg function

$$g : \mathbb{R}_+ :\to E$$

has at most countably many jumps.

Proof. By contradiction, if there were infinitely many jumps of f greater than $\frac{1}{n}$ one would find a strictly increasing sequence

$$(t_k)_{k\in\mathbb{N}} \subset (0,T]$$

where these jumps occur and by compactness of the interval there would be a subsequence converging to some $t \in (0,T]$ in contradiction to left continuity of f at t. □

Definition A.6.6. Two local martingales M, N are called *orthogonal* if $M \cdot N$ is a local martingale.

Definition A.6.7. A local martingale M is called *purely discontinuous local martingale* if $M_0 = 0$ and if M is orthogonal to all continuous local martingales.

Theorem A.6.8. *(*[**24**]*, Theorem I.4.18) Any local martingale M admits a unique (up to indistinguishability) decomposition*

$$M = M_0 + M^c + M^d,$$

where $M_0^c = M_0^d = 0$, and $M^c$is a continuous local martingale, and $M^d$is a purely discontinuous local martingale. M^c is called the continuous part of M, M^d is called purely discontinuous part of M.

Theorem A.6.9. *(* [**24**]*, Theorem I.4.2)*

To each pair (M, N) of square integrable local martingales there is a predictable unique (up to indistinguishability) process with finite variation $\langle M, N\rangle$ such that

$$MN - \langle M, N\rangle$$

is a local martingale.

Proposition A.6.10. *(*[**24**]*, Proposition I.4.27)*

Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space with right continuous filtration. Let X be a semimartingale. There is a unique (up to indistinguishability) continuous local martingale X^c starting at 0 *such that for any decomposition of X according to Definition 3.0.4 given by*

$$X = X_0 + M + A$$

up to indistinguishability $X^c = M^c$ holds (where M^c is the continuous local martingale from Theorem A.6.8). X^c is called continuous martingale part of X.

Definition A.6.11. A truncation function on $\mathbb{R}^d$ is a bounded function

$$h : \mathbb{R}^d \to \mathbb{R}^d$$

such that $h(x) = x$ in a neighborhood of 0.

Definition A.6.12. Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space with right continuous filtration. Let E be a polish space and $\mathcal{E}$ its Borel σ-algebra. A *random measure* on $\mathbb{R}_+ \times E$ is a family

$$(\mu(\omega, dt, dx))_{\omega\in\Omega}$$

of measures on

$$(\mathbb{R}_+ \times E,\, \mathcal{B}(\mathbb{R}_+) \otimes E)$$

satisfying $\mu(\omega, \{0\} \times E) = 0$ for all $\omega \in \Omega$.

Definition A.6.13. Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space with right continuous filtration, let $E \subset \mathbb{R}^d$, and let

$$(\mu(\omega, dt, dx))_{\omega\in\Omega}$$

be a random measure on $\mathbb{R}_+ \times E$. Let $W : \Omega \times \mathbb{R}_+ \times E \to E$ be a map that is measurable with respect to $\mathscr{O} \otimes \mathcal{E}$. Then due to Corollary A.6.4 and Lemma A.3.16 for any $\omega \in \Omega$ the map

$$(t, x) \to W(\omega, t, x)$$

is $\mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$-measurable and we can define

$$W * \mu_t(\omega) := \int\limits_{[0,t]\times E} W(\omega, s, x)\mu(\omega, ds, dx)$$

if $\int\limits_{[0,t]\times E} |W(\omega, s, x)|\, \mu(\omega, ds, dx) < \infty$ and $W * \mu_t(\omega) := \infty$ otherwise.

Definition A.6.14. Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space with right continuous filtration, let $E \subset \mathbb{R}^d$, and let

$$(\mu(\omega, dt, dx))_{\omega\in\Omega}$$

be a random measure on $\mathbb{R}_+ \times E$. We call a random measure

$$(\mu(\omega, dt, dx))_{\omega\in\Omega}$$

optional if

$$(t, \omega) \to W * \mu_t(\omega)$$

is optional for any optional map $W : \Omega \times \mathbb{R}_+ \times E \to E$

We call a random measure

$$(\mu(\omega, dt, dx))_{\omega\in\Omega}$$

predictable if

$$(t, \omega) \to W * \mu_t(\omega)$$

is predictable for any predictable map $W : \Omega \times \mathbb{R}_+ \times E \to E$

We call an random measure

$$(\mu(\omega, dt, dx))_{\omega\in\Omega}$$

that is optional $\tilde{\mathscr{P}}$-σ $finite$ if there exists a strictly positive map V on $\Omega \times \mathbb{R}_+ \times E$, measurable with respect to $\tilde{\mathscr{P}} := \mathscr{P} \otimes \mathcal{E}$ such that

$$\omega \to \int\limits_{\mathbb{R}_+\times E} V(\omega, s, x)\mu(\omega, ds, dx)$$

is integrable.

Definition A.6.15. Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space with right continuous filtration, let $E \subset \mathbb{R}^d$ and let $\mathcal{E}$ be its Borel σ-algebra. An *integer − valued random measure*

$$(\mu(\omega, dt, dx))_{\omega\in\Omega}$$

on $\mathbb{R}_+ \times E$ is a random measure on $\mathbb{R}_+ \times E$ such that

(i)$\mu(\omega, \{t\} \times E) \leq 1$ for all $\omega \in \Omega$ and all $t \in \mathbb{R}_+$

(ii) for all $A \in \mathcal{B}(\mathbb{R}_+) \otimes \mathcal{E}$ and for all $\omega \in \Omega$:$\mu(\omega, A) \in \mathbb{N} \cup \{\infty\}$

(iii) $(\mu(\omega, dt, dx))_{\omega\in\Omega}$ is $\tilde{\mathscr{P}}$-σ-finite.

Theorem A.6.16. *(*[**24**] *Theorem II.1.8)*

Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ *be a filtered probability space with right continuous filtration , let* $E \subset \mathbb{R}^d$ *and let* $\mathcal{E}$ *be its Borel* σ*-algebra. Let*

$$(\mu(\omega, dt, dx))_{\omega\in\Omega}$$

be an $\tilde{\mathscr{P}}$-σ *finite random measure on* $\mathbb{R}_+ \times E$*. Then there exists a random measure*

$$(\mu^p(\omega, dt, dx))_{\omega\in\Omega}$$

on $\mathbb{R}_+ \times E$ *called compensator of* μ *which is unique up to a* $\mathbb{P}$*-null set and which is characterized as being a predictable random measure on* $\mathbb{R}_+ \times E$ *such that either :*

(i) For each non-negative $\mathscr{P} \otimes \mathcal{E}$*-measurable function* W *on* $\Omega \times \mathbb{R}_+ \times E$

$$\int\limits_{\Omega} |W * \mu_\infty(\omega)|\, d\mathbb{P}(\omega) = \int\limits_{\Omega} |W * \mu^p_\infty(\omega)|\, d\mathbb{P}(\omega).$$

or

(ii) Let W be a non-negative $\mathscr{P} \oplus \mathcal{E}$-measurable function on $\Omega \times \mathbb{R}_+ \times E$. Let $(\tau_n)_{n\in\mathbb{N}}$ be an increasing sequence of $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-stopping times with $\lim_{n\to\infty} \tau_n \to \infty$ $\mathbb{P}$- almost surely. Let

$$(t,\omega) \to W * \mu_{\min(\tau_n,t)}(\omega)$$

be càdlàg, adapted, starting at 0, with non-decreasing paths for any $n \in \mathbb{N}$ and

$$\int_\Omega \left| W * \mu_{\min(\tau_n,\infty)}(\omega) \right| d\mathbb{P}(\omega) < \infty,$$

for any $n \in \mathbb{N}$. Then for any $n \in \mathbb{N}$ also

$$(t,\omega) \to W * \mu^p_{\min(\tau_n,t)}(\omega)$$

is càdlàg, adapted, starting at 0, with non-decreasing paths and

$$\int_\Omega \left| W * \mu^p_{\min\,(\tau_n,\,\infty)}(\omega) \right| d\mathbb{P}(\omega) < \infty.$$

for any $n \in \mathbb{N}$. Additionally,

$$(W * \mu_t - W * \mu^p_t)_{t\in\mathbb{R}_+}$$

is a local martingale.

Definition A.6.17. Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space with right continuous filtration. A subset A of $\Omega \times \mathbb{R}_+$ is called *thin* if it can be written as

$$A = \bigcup_{n\in\mathbb{N}} \{(\omega, T_n(\omega))\}$$

for a sequence $(T_n)_{n\in\mathbb{N}}$ of stopping times.

Proposition A.6.18. *([**24**], Proposition II.1.14)*
Let

$$\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$$

be a filtered probability space with right continuous filtration. Let $E \subset \mathbb{R}^d$ and let $\mathcal{E}$ be its Borel σ-algebra and let $(\mu(\omega, dt, dx))_{\omega\in\Omega}$ be an integer-valued random measure on $\mathbb{R}_+ \times E$. Let δ_x be the Dirac measure in x (see Example A.3.21). Then there exists a thin set $D \subset \Omega \times \mathbb{R}_+$ and an optional E-valued optional process β such that

$$\mu(\omega, dt, dx) = \sum_{s\geq 0} 1_D(\omega, s)\delta_{(s,\beta_s(\omega))}(dt, dx).$$

PROPOSITION A.6.19. *([**24**], Proposition II.1.16) Let*

$$\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$$

be a filtered probability space with right continuous filtration. Let X be a $\mathbb{R}^d$-valued càdlàg process. Then

$$\mu^X(\omega, dt, dx) = \sum_{s\geq 0} 1_{\{\Delta X_s(\omega)\neq 0\}} \delta_{(s,\Delta X_s(\omega))}(dt, dx)$$

is an integer-valued measure on $\mathbb{R}_+ \times \mathbb{R}^d$.

LEMMA A.6.20. *([**24**], Equation II .2.5) Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space with right continuous filtration. Let X be a semimartingale and h a truncation function. Then with*

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

the process

$$X(h) := X - \tilde{X}(h)$$

admits the decomposition

$$X(h) = X_0 + M(h) + B(h)$$

with $\mathscr{F}_0$- measurable random variableX_0, a d-dimensional local martingale $M(h)$ starting at 0 and $B(h)$ a predictable d-dimensional process with finite variation.

DEFINITION A.6.21. Let $\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$ be a filtered probability space with right continuous filtration. Let $(X_t)_{t\in\mathbb{R}_+}$ be a d-dimensional semimartingale. For a fixed truncation function h, the triple (B, C, ν) is called *semimartingale characteristics associated with* h where

(i) $B = B(h)$ is the predictable d-dimensional process with finite variation appearing in Lemma A.6.20,

(ii) $C = (C^{ij})_{1\leq i,j\leq d}$ for $C^{ij} = \left\langle (X^c)^i, (X^c)^j \right\rangle$ is a continuous process with finite variation (for the covariation process see TheoremA.6.9) for the continuous martingale part X^c(see Proposition A.6.10)

(iii) $\nu = (\nu(\omega, dt, dx))_{\omega\in\Omega}$ is the compensator (see Theorem A.6.16) of the random measure $\mu^X = \left(\mu^X(\omega, dt, dx)\right)_{\omega\in\Omega}$ on $\mathbb{R}_+ \times \mathbb{R}^d$ from

Proposition A.6.19 and as such ν is a predictable random measure on $\mathbb{R}_+ \times \mathbb{R}^d$.

Definition A.6.22. Let
$$\left(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, \mathbb{P}\right)$$
be a filtered probability space with right continuous filtration. Let $E \subset \mathbb{R}^d$ and let μ be an integer-valued random measure on $\mathbb{R}_+ \times E$. Let
$$W : \Omega \times \mathbb{R}_+ \times E \to \mathbb{R}$$
be a process that is measurable with respect to $\mathscr{P} \otimes \mathcal{E}$. Let ν be the compensator of μ and define
$$\widehat{W}_t(\omega) := \begin{cases} \int\limits_E W(\omega, t, x)\nu(\omega; \{t\} \times dx) & \text{if } \int\limits_E |W(\omega, t, x)|\, \nu(\omega; \{t\} \times dx) < \infty \\ \infty & \text{else.} \end{cases}$$
Furthermore, let D and β_t be the thin set and optional process from Proposition A.6.18 and define
$$\tilde{W} := W(\omega, t, \beta_t(\omega))1_D(\omega, t) - \widehat{W}_t(\omega).$$
Then W is said to belong to $G_{loc}(\mu)$ if for the process $(Q_t)_{t\in\mathbb{R}_+}$ defined by
$$Q_t := \left(\sum_{s\le t} \left(\tilde{W}_s\right)^2\right)^{1/2}$$
there is a increasing sequence $(\tau_n)_{n\in\mathbb{N}}$ of $(\mathscr{F}_t)_{t\in\mathbb{R}_+}$-stopping times with
$$\lim_{n\to\infty} \tau_n \to \infty$$
$\mathbb{P}$- almost surely such that for any $n \in \mathbb{N}$ the stopped process
$$\left(Q_{\min(t,\tau_n)}\right)_{t\in\mathbb{R}_+}$$
has càdlàg, adapted, non-decreasing paths starting at 0 and
$$\mathbb{E}\left[Q_{\tau_n}\right] < \infty$$
for any $n \in \mathbb{N}$.

Definition A.6.23. If $W \in G_{loc}(\mu)$ then $W * (\mu - \nu)$ is defined as any purely discontinuous local martingale M such that ΔM and $\tilde{W}$ are indistinguishable.

In **[24]** below Definition I.1.27 it is shown that:

PROPOSITION A.6.24. *If* $W \in G_{loc}(\mu)$ *then* $W * (\mu - \nu)$ *exists and is unique up to indistinguishability.*

THEOREM A.6.25. *(*[**24**]*, Theorem II.2.34) (Canonical representation of a semimartingale) Let* $(X_t)_{t\in\mathbb{R}_+}$ *be a* d*-dimensional semimartingale and let* (B, C, ν) *be its characteristics relative to a truncation function (see Definition A.6.11)* h*. Furthermore, let* μ^X *be the random measure associated to* X *via Proposition A.6.19 and* X^c *its continuous martingale part (see Proposition A.6.10). Then for* $W^i(\omega, t, x) := h^i(x)$ $W \in G_{loc}(\mu^X)$ *for* $1 \leq i \leq d$ *and*

$$X = X_0 + X^c + h * \left(\mu^X - \nu\right) + (x - h(x)) * \mu^X + B,$$

where the d-dimensional integral $h * \left(\mu^X - \nu\right)$ is defined componentwise.

Bibliography

[1] Charalambos D. Aliprantis and Kim C. Border. *Infinite Dimensional Analysis.* Springer, 2005.

[2] Herbert Amann. *Ordinary Differential Equations.* Walter De Gruyter, 1990.

[3] 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: Series B (Statistical Methodology)*, 63(2):167–241, 2001.

[4] Klaus D. Bierstedt. Introduction to topological tensor products. Lecture Notes University of Paderborn 2007.

[5] Tomas Björk. *Arbitrage Theory in Continuous Time.* Oxford University Press, 2009.

[6] Nicolas Bourbaki. *Éléments de mathématique - Livre VI - Intégration - Chapitre IX - Intégration sur les éspaces topologiques séparés.* DIFFUSION C.C.L.S., 1969.

[7] Nicolas Bourbaki. *Éléments de mathématique - Topologie générale Chapitres 5 à 10.* Springer, 1971.

[8] Nicolas Bourbaki. *Intégration Chapitres 1 à 4.* Springer, 2007.

[9] Leo Breiman. *Probability.* Society for Industrial and Applied Mathematics, 1992.

[10] John B. Conway. *A Course in Functional Analysis.* Springer, 1985.

[11] Christa Cuchiero, Damir Filipović, Eberhard Mayerhofer, and Josef Teichmann. Affine processes on positive semidefinite matrices. *The Annals of Applied Probability*, 2009.

[12] Christa Cuchiero, Martin Keller-Ressel, and Josef Teichmann. Polynomial processes and their applications to mathematical finance. *Finance and Stochastics*, 16(4):711–740, 2012.

[13] Christa Cuchiero and Josef Teichmann. *Séminaire de Probabilités XLV. Lecture Notes in Mathematics, vol 2078.*, chapter Path Properties and Regularity of Affine Processes on General State Spaces. Springer, Heidelberg, 2013.

[14] Christa Cuchiero and Josef Teichmann. Generalized feller processes and markovian lifts of stochastic volterrra processes: The affine case. *Journal of Evolution Equations*, 2020.

[15] Philipp Doersek and Josef Teichmann. A semigroup point of view on splitting schemes for stochastic (partial) equations. *ArXiv e-Prints*, 2010.

[16] D. Duffie, D. Filipovic, and W. Schachermayer. afffine processes and applications in finance. *Annals of Applied Probability*, 13:984–1053, 2003.

[17] Jürgen Elstrodt. *Maß- und Integrationstheorie*. Springer, 2010.

[18] Klaus-Jochen Engel and Rainer Nagel. *One-Parameter Semigroups for Linear Evolution Equations*. Springer, 2000.

[19] Lawrence C. Evans. *Partial Differential Equations*. American Mathematical Society, 2010.

[20] Martin Hairer. An introduction to stochastic pdes. http://www.hairer.org/notes/SPDEs.pdf.

[21] Philip Hartman. *Ordinary Differential Equations*. SIAM, 2002.

[22] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 1993.

[23] E. Hewitt and K.R. Stromberg. *Real and abstract analysis*. Springer, 1965.

[24] Jean Jacod and Albert N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer, 2003.

[25] Ying Jiao, Chunhua Ma, and Simone Scotti. Alpha-cir model with branching processes in sovereign interest rate modeling. *Finance and Stochastics*, 21(3):789–813, 2017.

[26] Olav Kallenberg. *Foundations of Modern Probability*. Springer, 1997.

[27] Ioannis Karatzas and Steven E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 2000.

[28] K. Kawazu and S.Watanabe. Branching processes with immigration and related limit theorems. *Teor. Veroyatnost. i Primenen., 16:1 (1971), 34â51*, 1971.

[29] Martin Keller-Ressel, Thorsten Schmidt, and Robert Wardenga. Affine processes beyond stochastic continuity. *The Annals of Applied Probability*, 29(6):3387 – 3437, 2019.

[30] Achim Klenke. *Wahrscheinlichkeitstheorie*. Springer, 2008.

[31] Robert E. Megginson. *An Introduction to Banach Space Theory*. Springer, 1998.

[32] Bernt Oksendal. *Stochastic Differential Equations*. Springer, 2000.

[33] Andrea Pascucci. *PDE and Martingale Methods in Option Pricing*. Springer-Verlag, 2011.

[34] Marcus Reiss. Stochastic processes (lecture notes 2016/17). https://www.math.hu-berlin.de/ mreiss/GliederungWS16.pdf, January 2017.

[35] D. Revuz and M. Yor. *Continuous martingales and Brownian motion,*. 1994.

[36] Michael Röckner and Zeev Sobol. Kolmogorov equations in infinite dimensions: Well-posedness and regularity of solutions, with applications to stochastic generalized burgers equations. *Annals of Probability*, 2006.

[37] L.C.G. Rogers and David Williams. *Diffusions, Markov Processes, and Martingales*. Cambridge University Press, 1994.

[38] Walter Rudin. *Functional Analysis*. McGraw-Hill Book Company, 1973.

[39] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Company, 1987.

[40] Boto von Querenburg. *Mengentheoretische Topologie*. Springer, 2000.

[41] Dirk Werner. *Funktionalanalysis*. Springer, 1995.