EVOLUTIONARY SEMIGROUPS ON PATH SPACES

ROBERT DENK, MARKUS KUNZE, AND MICHAEL KUPPER

Abstract. We introduce the concept of evolutionary semigroups on path spaces, generalizing the notion of transition semigroups to possibly non-Markovian stochastic processes. We study the basic properties of evolutionary semigroups and, in particular, prove that they always arise as the composition of the shift semigroup and a single operator called the expectation operator of the semigroup. We also prove that the transition semigroup of a Markov process can always be extended to an evolutionary semigroup on the path space whenever the Markov process can be realized with the appropriate path regularity. As first examples of evolutionary semigroups associated to non-Markovian processes, we discuss deterministic evolution equations and stochastic delay equations.

1. Introduction

An important object in the study of a Markov process (say with a Polish state space X) is its transition semigroup. Indeed, encoded in the transition semigroup are the transition probabilities of the process and, together with the initial distribution, these determine the finite dimensional marginals and thus the distribution of the process as a random variable with values in the space of all X-valued functions on $[0,\infty)$. In particular, in view of Kolmogorov's extension theorem, we can construct a Markov process with a prescribed transition semigroup. Depending on the Markov process in question, it is sometimes possible to replace the space of all X-valued functions on $[0,\infty)$ with a space of more regular functions. The most important examples are the space $C([0,\infty);X)$ of all continuous X-valued functions and $D([0,\infty);X)$, the space of all càdlàg X-valued functions. But in all cases we can say that all information about the stochastic process is encoded in the transition semigroup and vice versa.

In this article, we introduce the concept of *evolutionary semigroups on path spaces* to extend the notion of transition semigroups to possibly non-Markovian stochastic processes. Our main guidelines in establishing this concept are two statements which may be considered 'mathematical folk wisdom' – not proven facts but rather an intuitive understanding on how a sensible theory should look like. These are as follows:

If we enlarge the state space appropriately, every stochastic process becomes Markovian. This suggests that we should incorporate enough 'history' of the stochastic process into the state space to make it Markovian. Consequently, as an 'ultimate state space', we should consider a space of X-valued functions on the interval $(-\infty,0]$, where we consider the value at time $t=0$ as the present state of the process and the value at time $t<0$ as the position of the process at the past time $t<0$. Thus, if we seek to describe a stochastic process with continuous paths (and we will assume this throughout this introduction; however our general setup also allows for different 'path spaces', in particular for càdlàg paths), we should

Date: July 25, 2025.

1991 *Mathematics Subject Classification.* 47D03, 46G12, 60J35.

Key words and phrases. Transition semigroup, non-Markovian stochastic process, Koopman semigroup, (stochastic) delay equation.

use $\mathscr{X}^- = C((-\infty, 0]; X))$ as state space. This strategy of 'incorporating the past' into the state space is rather standard in the semigroup approach to (deterministic) delay equations, see [3] and the references therein, but it was also used for stochastic delay equations, see [43]. We should point out that often (in particular in the references just mentioned) only the history in a finite time horizon is considered, i.e. one uses $C([-h, 0]; X)$ instead of $C((-\infty, 0]; X)$.

However, there is a fundamental difference between semigroup theory for deterministic delay equations and the theory that we want to develop here. In the former, X itself is a Banach space and one constructs a semigroup on the space $C([-h, 0]; X)$ (or a similar space such as $L^p([-h, 0]; X)$) that describes the evolution of the solution of the delay equation. On the other hand, in this article, X is not a vector space in general and (similar to [43]) we want to use the 'path space' $\mathscr{X}^- = C((-\infty, 0]; X)$ as state space of a then Markovian process. Thus, the transition semigroup acts on the space $C_b(\mathscr{X}^-) = C_b\big(C((-\infty, 0]; X)\big)$. In a sense, the semigroup we seek to construct acts on a function space that is 'one level higher' than that considered in the classical theory. This is similar to considering the *Koopman operator* (or, in the time continuous setting, *Koopman semigroups*) in the study of dynamical systems, see [17]. Typically, the Koopman operator/semigroup acts either on an L^p-space or (closer to our situation) on the space $C(K)$ of continuous functions on a compact space K. Only recently, there was a generalization of this theory to the setting of completely regular spaces in [22].

As a second guideline we require that, *as time passes, the past of the process is merely shifted in time*. While this guideline does not need additional explanation, it raises an important question. What should happen with the past once it is shifted into the future? In the semigroup approach to delay equations mentioned above, this question is answered at an infinitesimal level. The generator of the shift semigroup is (roughly speaking) the first derivative. To obtain a semigroup governing the evolution of a delay equation, one considers a realization of the first derivative, defined on a subspace of $C^1([-h, 0]; X)$, where the delay equation itself enters the domain of the generator as a (Neumann-type) boundary condition at time $t = 0$. In our situation, the shift semigroup is a rather delicate object (see Remark 3.10) and this approach does not seem to work.

Therefore, in this article we follow a different approach. Setting $\mathscr{X} = C(\mathbb{R}; X)$, we identify the space $C_b(\mathscr{X}^-)$ with a subspace of $C_b(\mathscr{X})$ as follows: Given $\mathbb{x} \in \mathscr{X}$ and $\tilde{F} \in C_b(\mathscr{X}^-)$, we can extend $\tilde{F}$ to a function $F \in C_b(\mathscr{X})$, by setting $F(\mathbb{x}) = \tilde{F}(\mathbb{x}|_{(-\infty,0]})$. Note that $F(\mathbb{x}) = F(\mathbb{y})$ for any $\mathbb{x}, \mathbb{y} \in \mathscr{X}$ with $\mathbb{x}(t) = \mathbb{y}(t)$ for all $t \leq 0$ and this property actually characterizes functions that appear as extension of functions in $C_b(\mathscr{X}^-)$. As it turns out (see Lemma 2.5) this property is equivalent to F being measurable with respect to the σ-algebra $\mathscr{F}_0$ that is generated by the point evaluations $(\pi_t)_{t\leq 0}$. Similarly, functions $F \in C_b(\mathscr{X})$ whose value only depends on $\mathbb{x}|_{(-\infty,t]}$ are exactly those that are measurable with respect to $\mathscr{F}_t := \sigma(\pi_s : s \leq t)$. Let us write $C_b(\mathscr{X}, \mathscr{F}_t)$ for the space of bounded continuous functions $F : \mathscr{X} \to \mathbb{R}$ that are $\mathscr{F}_t$-measurable. With this identification at hand, we can work throughout on the space $C_b(\mathscr{X})$ and, in particular, avoid the central question posed above. If we shift a function $F \in C_b(\mathscr{X}, \mathscr{F}_0)$ (say by the time lapse $t > 0$), we again obtain a function in $C_b(\mathscr{X})$, which is $\mathscr{F}_t$-measurable rather than $\mathscr{F}_0$-measurable.

After this preparation, we now define an *evolutionary semigroup* as a semigroup $(\mathbb{T}(t))_{t\geq 0}$ on $C_b(\mathscr{X}, \mathscr{F}_0)$ such that for $0 \leq s \leq t$ and $F \in C_b(\mathscr{X}, \mathscr{F}_{-t})$ it holds $\mathbb{T}(s)F = \Theta_s F$, where $(\Theta_t)_{t\geq 0}$ is the *shift semigroup* on $C_b(\mathscr{X})$, see Section 3. This requirement is a rephrasement of our second guiding principle. In our first main result, Theorem 4.5, we prove that a semigroup on $C_b(\mathscr{X}, \mathscr{F}_0)$ is evolutionary if and only if it is given as $\mathbb{T}(t) = \mathbb{E}\Theta_t$ for some bounded linear operator $\mathbb{E}$ on $C_b(\mathscr{X})$ that takes values in $C_b(\mathscr{X}, \mathscr{F}_0)$ and satisfies $\mathbb{E}F = F$ for all $F \in C_b(\mathscr{X}, \mathscr{F}_0)$. This operator $\mathbb{E}$ determines the semigroup $(\mathbb{T}(t))_{t\geq 0}$ uniquely and is called the *expectation operator* of the semigroup. Proposition 4.1 shows that $\mathbb{E}$ indeed has properties that are characteristic for expectations. The expectation operator also connects (various notions of) the generator of $\mathbb{T}$ to that of the shift semigroup, see Propositions 4.10, 5.4 and 5.9. We point out that the expectation operator now settles the central question posed above: Once the shift operator Θ_t transports information into the future (i.e. producing a function that is not $\mathscr{F}_0$-measurable any more) the expectation operator $\mathbb{E}$ transforms this into an $\mathscr{F}_0$-measurable function (and thus into a function that we might view as a function on $\mathscr{X}^-$).

Let us now discuss some examples.

Delay equations. Our first example concerns deterministic delay equations. We fix $d \in \mathbb{N}$, $h > 0$ and define

$$\mathscr{C}_h := C([-h, 0]; \mathbb{R}^d).$$

As usual, for a function $y \in C([-h,\infty); \mathbb{R}^d)$ and $t \geq 0$, we define $y_t \in \mathscr{C}_h$, the *past at time* t, by setting $y_t(s) = y(t+s)$ for $s \in [-h, 0]$. Now, given a Lipschitz map $b : \mathscr{C}_h \to \mathbb{R}^d$ and an 'initial history' $\xi \in \mathscr{C}_h$, we may consider the delay equation

$$\begin{cases} y'(t) = b(y_t), & \text{for } t \geq 0, \\ \quad y_0 = \xi. \end{cases} \tag{1.1}$$

As b is assumed to be Lipschitz continuous, an argument based on the Banach fixed point theorem shows that (1.1) has a unique solution y^ξ, see [4, Theorem II.4.3.1]. In Section 6.1, we will see that, setting $X = \mathbb{R}^d$ and $\mathscr{X} = C(\mathbb{R}; \mathbb{R}^d)$ as above, we can associated an evolutionary semigroup $\mathbb{T}$ on $C_b(\mathscr{X}, \mathscr{F}_0)$ with such equations. To describe the expectation operator, given $\mathbbm{x} \in \mathscr{X}$, we define the function $\mathbbm{x}_- \in \mathscr{X}^- = C((-\infty, 0]; \mathbb{R}^d)$ as the restriction $\mathbbm{x}|_{(-\infty,0]}$. Note that $\mathbbm{x}_0 \in \mathscr{C}_h$ so that (1.1) has a unique solution $y^{\mathbbm{x}_0}$ for $\xi = \mathbbm{x}_0$. We set

$$[\mathbbm{x}_- \otimes_0 y^{\mathbbm{x}_0}](t) := \begin{cases} \mathbbm{x}(t), & \text{if } t \leq 0, \\ y^{\mathbbm{x}_0}(t), & \text{if } t > 0. \end{cases}$$

Then $\mathbbm{x}_- \otimes_0 y^{\mathbbm{x}_0} \in \mathscr{X}$ and we define for $F \in C_b(\mathscr{X})$,

$$[\mathbb{E}F](\mathbbm{x}) := F(\mathbbm{x}_- \otimes_0 y^{\mathbbm{x}_0}). \tag{1.2}$$

We will see in Section 6.1, that this is indeed an expectation operator that induces an evolutionary semigroup by setting $\mathbb{T}(t) := \mathbb{E}\Theta_t$.

We point out that there are several semigroup approaches to delay equations to be found in the literature, see, for example, [15], [3], [19, Section VI.6] or [4, Section II.4]. Typically, in these semigroup approaches, the map b is assumed to be linear and nonlinear maps can be handled via the variation of constants formula. On the other hand, we obtain an evolutionary semigroup $\mathbb{T}$ also for nonlinear b without relying on the variation of constants formula. The semigroup $\mathbb{T}$ can be seen as a Koopman approach to delay equations or as an extension of the results from [22] to equations that depend on the past.

Stochastic differential equations. Next, let $(B(t))_{t\geq 0}$ be an m-dimensional Brownian motion, defined on a probability space $(\Omega, \Sigma, \mathbf{P})$. Given functions $b : \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \to \mathbb{R}^{m\times d}$, we can consider the stochastic differential equation

$$\begin{cases} dY(t) = b(Y(t))\, dt + \sigma(Y(t))dB(t), & \text{for } t \geq 0, \\ \ Y(0) = y. \end{cases} \tag{1.3}$$

Under suitable assumptions on the coefficients b and σ, it is well known that for every initial datum $y \in \mathbb{R}^d$, Equation (1.3) has a unique solution $(Y^y(t))_{\geq 0}$. It turns out that also to this equation we can associate an evolutionary semigroup on the path space $\mathscr{X} = C(\mathbb{R}, \mathbb{R}^d)$. Motivated by (1.2), we set for $F \in C_b(\mathscr{X})$,

$$[\mathbb{E}F](\mathbb{x}) := \mathbf{E}\big[F(\mathbb{x}_- \otimes_0 Y^{\mathbb{x}(0)})\big],$$

where $\mathbf{E}$ refers to expectation with respect to the measure $\mathbf{P}$. Using the fact that the solutions to (1.3) are Markov processes, we prove in Section 6.2 that $\mathbb{E}$ is an expectation operator that induces an evolutionary semigroup. As a matter of fact, this is true not only for solutions of stochastic differential equations, but for all Markov processes that can be realized with continuous paths. Switching to the path space of càdlàg paths, the same result also applies to Markovian processes with càdlàg paths. In Theorem 6.7, we give a complete characterization of evolutionary semigroups that arise in this way.

Stochastic delay equations. For our last example, consider stochastic delay equations of the form

$$\begin{cases} dY(t) = b(Y_t)dt + \sigma(Y_t)dB(t), & \text{for } t \geq 0, \\ \quad Y_0 = \xi. \end{cases} \tag{1.4}$$

Here, once again, $(B(t))_{t\geq 0}$ is an m-dimensional Brownian motion on the probability space $(\Omega, \Sigma, \mathbf{P})$. However, compared to (1.3), the domain of the maps b and σ changes from $\mathbb{R}^d$ to the function space $\mathscr{C}_h$. Nevertheless, under Lipschitz assumptions on the coefficients, it is known that for any $\xi \in \mathscr{C}_h$, Equation (1.4) has a unique solution Y^ξ and we may define

$$[\mathbb{E}F](\mathbb{x}) := \mathbf{E}\big[F(\mathbb{x}_- \otimes_0 Y^{\mathbb{x}_0})\big].$$

We point out that the solutions Y^ξ are no longer Markov processes on $\mathbb{R}^d$, but they turn out to be Markov processes on the space $\mathscr{C}_h$. Using this fact, we prove in Section 6.3 that $\mathbb{E}$ is indeed the expectation operator of an evolutionary semigroup on the space $\mathscr{X} = C(\mathbb{R}, \mathbb{R}^d)$. In Section 6.4, we discuss some related results in the càdlàg setting.

Organization of this article. In Section 2, we introduce the abstract concept of a 'path space' and establish some preliminary measurability results. Our main examples of path spaces are $C(\mathbb{R}; X)$ and $D(\mathbb{R}; X)$, as we will prove in Appendix B. In Section 3, we introduce the shift semigroup and present some results concerning its full generator and its C_b-generator. We should point out that the semigroups considered in this article are not strongly continuous, whence a different semigroup theory is used throughout. The main definitions and results of this theory are collected in Appendix A, where also references to the literature may be found. In Section 4, we take up our main line of study and introduce the central concepts of 'evolutionary semigroup' and 'expectation operator'. In Section 4 we do not work in $C_b(\mathscr{X})$ but rather in $B_b(\mathscr{X})$, the space of all bounded measurable functions on the path space $\mathscr{X}$. Continuity properties of evolutionary semigroups are discussed in Section 5. The choice of the path space plays an important role here. The case where $\mathscr{X} = C(\mathbb{R}; X)$ is the easier and all of the expected results hold true. The general case is much more involved and we actually impose additional assumptions on the 'path space' (which, however, are satisfied in the case of càdlàg paths). The basic problem in the case of càdlàg paths is that point evaluations are not continuous functions. Nevertheless, most results from the continuous case generalize. In the concluding Section 6, we present our examples.

Acknowledgement. We are grateful to Rainer Nagel for helpful comments.

2. Path spaces

Definition 2.1. Let (X, d) be a complete separable metric space. A *path space (with state space X)* is a pair $((\mathscr{X}, \mathbb{d}), \tau)$ consisting of a complete separable metric space $(\mathscr{X}, \mathbb{d})$ of right-continuous functions $\mathbb{x} : \mathbb{R} \to X$ and a map $\tau : \mathscr{X} \to \mathscr{X}$, such that the following conditions are satisfied:

(P1) The *evaluation maps* $\pi_t : \mathscr{X} \to X$, $\pi_t(\mathbb{x}) = \mathbb{x}(t)$ are Borel measurable and the Borel σ-algebra $\mathfrak{B}(\mathscr{X})$ is generated by these maps, i.e. $\mathfrak{B}(\mathscr{X}) = \sigma(\pi_t : t \in \mathbb{R})$. Every $\mathbb{x} \in \mathscr{X}$ is continuous at almost every $t \in \mathbb{R}$. Moroever, if $\mathbb{x}_n \to \mathbb{x}$ and $\mathbb{x}$ is continuous at t, then $\pi_t(\mathbb{x}_n) \to \pi_t(\mathbb{x})$.

We define some additional σ-algebras. Given an interval $I \subset \mathbb{R}$, we set

$$\mathscr{F}(I) := \sigma(\pi_t : t \in I).$$

Of particular importance is the case where $I = (-\infty, t)$ or $I = (-\infty, t]$ for some $t \in \mathbb{R}$. We define $\mathscr{F}_{t-} := \mathscr{F}((-\infty, t))$ and $\mathscr{F}_t := \mathscr{F}((-\infty, t])$.

(P2) The *stopping map* $\tau : \mathscr{X} \to \mathscr{X}$ is $\mathscr{F}_{0-}$-measurable and $\tau^{-1}(A) = A$ for every $A \in \mathscr{F}_{0-}$. Moreover, $\tau(\mathbb{x})$ is continuous at 0 for every $\mathbb{x} \in \mathscr{X}$ and the image $\mathscr{X}^- := \tau(\mathscr{X})$ is Polish, i.e. there exists a complete metric $\mathbb{d}^-$ on $\mathscr{X}^-$ that induces the same topology as $\mathbb{d}$ on $\mathscr{X}^-$.

(P3) For every $t \in \mathbb{R}$, the *shift* ϑ_t, defined by $[\vartheta_t \mathbb{x}](s) := \mathbb{x}(t + s)$, maps $\mathscr{X}$ to itself and the map $(t, \mathbb{x}) \mapsto \vartheta_t \mathbb{x}$ is continuous from $\mathbb{R} \times \mathscr{X} \to \mathscr{X}$.

Throughout, elements of the path space $\mathscr{X}$ will be denoted by lower case blackboard letters $\mathbb{x}, \mathbb{y}, \mathbb{z}$ whereas elements of the state space X will be denoted by lower case roman letters x, y, z. Likewise, scalar-valued functions on $\mathscr{X}$ will be denoted by upper case roman letters F, G, H, whereas scalar-valued functions on X will be denoted by lower case roman letters f, g, h.

Example 2.2. Our basic examples of path spaces are $\mathscr{X}_{\mathrm{C}} := C(\mathbb{R}; X)$, the space of all *continuous paths* from $\mathbb{R}$ to X, endowed with the metric that topologizes uniform convergence on compact subsets of $\mathbb{R}$ and $\mathscr{X}_{\mathrm{D}} := D(\mathbb{R}; X)$, the space of all *càdlàg paths* from $\mathbb{R}$ to X, endowed with Skorohod's J_1-metric. We point out that the stopping map is defined slightly different in these two examples: for $\mathscr{X}_{\mathrm{C}}$, we define $[\tau_{\mathrm{C}}(\mathbb{x})](t) = \mathbb{x}(0)$ for $t \geq 0$, whereas for càdlàg paths $\mathscr{X}_{\mathrm{D}}$ we put $[\tau_{\mathrm{D}}(\mathbb{x})](t) = \mathbb{x}(0-)$ for $t \geq 0$.

As our main results only use the abstract assumptions of Definition 2.1, we postpone the details and proofs to Appendix B.

Example 2.3. In the context of continuous paths, we can also consider the spaces

$$\mathscr{X}_{\mathrm{C},\ell} := \{\mathbb{x} \in C(\mathbb{R}; X) : \lim_{t \to -\infty} \mathbb{x}(t) \text{ exists}\} = C([-\infty, \infty); X)$$

and, if X is additionally a vector space,

$$\mathscr{X}_{\mathrm{C},0} := \{\mathbb{x} \in C(\mathbb{R}; X) : \lim_{x \to -\infty} \mathbb{x}(t) = 0\}$$

Both are path spaces with respect to the metric

$$\mathbb{d}(\mathbb{x}, \mathbb{y}) = \sum_{n=1}^{\infty} 2^{-n} \big[1 \wedge \sup_{t \in (-\infty, n]} d(\mathbb{x}(t), \mathbb{y}(t))\big].$$

This choice of path spaces is helpful, for example, in the context of delay equations with infinite delay, see Equation (3.2) on page 240 of [4].

Example 2.4. Another possible choice for a path space is motivated by adding a cemetary state to the space X. This is a standard procedure to turn a non-honest Markov process into an honest one. This construction can be replicated on the level of the path space. Indeed, given a locally compact X, we denote its one-point compactification of X by $X^\dagger$, where the added point $\dagger$ is used as a cemetery state. As X is a complete separable metric space, so is $X^\dagger$. Now set

$$\mathscr{X}_{\mathrm{C},\dagger} := \{\mathbb{x} \in C(\mathbb{R}; X^\dagger) : \mathbb{x}(t_0) = \dagger \text{ implies } \mathbb{x}(t) = \dagger \text{ for all } t \geq t_0\}.$$

We also introduce the following notation. Given an interval $I \subset \mathbb{R}$, we put

$$B_b(\mathscr{X}, \mathscr{F}(I)) := \{F \in B_b(\mathscr{X}) : F \text{ is } \mathscr{F}(I)\text{-measurable}\}$$

and $C_b(\mathscr{X}, \mathscr{F}(I)) := B_b(\mathscr{X}, \mathscr{F}(I)) \cap C_b(\mathscr{X})$. We often write $B_b(\mathscr{X}, \mathscr{F}_{t-})$ and $C_b(\mathscr{X}, \mathscr{F}_{t-})$ instead of $B_b(\mathscr{X}, \mathscr{F}((-\infty, t)))$ and $C_b(\mathscr{X}, \mathscr{F}((-\infty, t)))$. If $F : \mathscr{X} \to \mathbb{R}$ is $\mathscr{F}_{t-}$-measurable, we will say that it is *determined before* t, if F is $\mathscr{F}([t,\infty))$-measurable, we will say that it is *determined after* t. Thus, $B_b(\mathscr{X}, \mathscr{F}([t_1, t_2)))$ refers to the space of bounded, measurable functions that are determined before t_2 and after t_1. Endowed with the supremum norm $\|\cdot\|_\infty$, all of these spaces are Banach spaces.

Let us note some easy consequences concerning measurability.

Lemma 2.5. *Let $((\mathscr{X}, \mathbb{d}), \tau)$ be a path space and S be a Polish space with Borel σ-algebra $\mathfrak{B}(S)$.*

(a) *For every $t \in \mathbb{R}$ and interval $I \subset \mathbb{R}$, the map ϑ_t is $\mathscr{F}(I+t)/\mathscr{F}(I)$-measurable.*
(b) *For $t \in \mathbb{R}$, the map $\tau_t := \vartheta_{-t} \circ \tau \circ \vartheta_t$ is $\mathscr{F}_{t-}$-measurable and $(\tau_t)^{-1}(A) = A$ for all $A \in \mathscr{F}_{t-}$.*
(c) *A measurable function $\Phi : \mathscr{X} \to S$ is $\mathscr{F}_{t-}$-measurable if and only if $\Phi = \Phi \circ \tau_t$.*
(d) *It holds $\tau^2 = \tau$. Moreover, $\mathbb{x} \in \mathscr{X}^-$ if and only if $\mathbb{x} = \tau(\mathbb{x})$.*
(e) *It holds $\mathbb{x}(t) = [\tau(\mathbb{x})](t)$ for all $t < 0$ and $\mathbb{x} \in \mathscr{X}$.*

Proof. (a). This follows directly from the identity $\pi_r \circ \vartheta_t = \pi_{r+t}$ for all $r \in \mathbb{R}$.

(b). If $A \in \mathfrak{B}(\mathscr{X})$, then $B := \tau^{-1}((\vartheta_{-t})^{-1}(A)) \in \mathscr{F}_{0-}$ by (P2), so that $(\vartheta_t)^{-1}(B) \in \mathscr{F}_{t-}$ by (a), proving the $\mathscr{F}_{t-}$-measurability of τ_t. If $A \in \mathscr{F}_{t-}$, then $(\vartheta_{-t})^{-1}(A) \in \mathscr{F}_{0-}$ by (a) so that $B = (\vartheta_{-t})^{-1}(A)$ by (P2). At this point, (a) yields $A = (\vartheta_t)^{-1}(B) = (\tau_t)^{-1}(A)$.

(c). If $\Phi = \Phi \circ \tau_t$, then Φ is $\mathscr{F}_{t-}$-measurable as τ_t is. Conversely, if Φ is $\mathscr{F}_{t-}$-measurable, then for $A \in \Sigma$, $\Phi^{-1}(A) \in \mathscr{F}_{t-}$ so that $\Phi^{-1}(A) = (\tau_t)^{-1}(\Phi^{-1}(A))$ by (b). As A was arbitrary, it follows that $\Phi = \Phi \circ \tau_t$.

(d). Applying (c) to $\Phi = \tau$ yields $\tau^2 = \tau$. If $\mathbb{x} = \tau(\mathbb{x})$ then $\mathbb{x} \in \mathscr{X}^-$. Conversely, assume that $\mathbb{x} = \tau(\mathbb{y}) \in \mathscr{X}^-$. Then $\tau(\mathbb{x}) = \tau^2(\mathbb{y}) = \tau(\mathbb{y}) = \mathbb{x}$.

(e). Let $A := \{\mathbb{x} \in \mathscr{X} : \mathbb{x}(t) = [\tau(\mathbb{x})](t) \text{ for all } t < 0\}$. Then $A \in \mathscr{F}_{0-}$ and thus $A = \tau^{-1}(A)$ by (P2). On the other hand, $\tau^{-1}(A) = \{\mathbb{x} : [\tau(\mathbb{x})](t) = [\tau^2(\mathbb{x})](t) \text{ for all } t < 0\} = \mathscr{X}$ by (d). $\square$

As a consequence of Lemma 2.5(c), an $\mathscr{F}_{0-}$-measurable function $F \in B_b(\mathscr{X})$ is uniquely determined by its values on $\mathscr{X}^-$. We may thus use the map τ to identify functions on $\mathscr{X}^-$ with functions on $\mathscr{X}$ by means of the *extension map* $\Phi \mapsto \hat{\Phi}$, defined by $\hat{\Phi} = \Phi \circ \tau$. Concerning measurability, we have the following result.

Lemma 2.6. *Let $((\mathscr{X}, \mathbb{d}), \tau)$ be a path space, S be a Polish space with Borel σ-algebra $\mathfrak{B}(S)$ and $\Phi : \mathscr{X} \to S$.*

(a) *The Borel σ-algebra $\mathfrak{B}(\mathscr{X}^-)$ is the trace σ-algebra of $\mathscr{F}_{0-}$ on $\mathscr{X}^-$.*
(b) *Then the function Φ is $\mathfrak{B}(\mathscr{X}^-)$-measurable if and only if $\hat{\Phi}$ is $\mathscr{F}_{0-}$-measurable.*

Proof. (a). By [6, Lemma 6.2.4], $\mathfrak{B}(\mathscr{X}^-)$ is the trace of $\mathfrak{B}(\mathscr{X})$ on $\mathscr{X}^-$, i.e. $A \in \mathfrak{B}(\mathscr{X}^-)$ if and only if there exits $B \in \mathfrak{B}(\mathscr{X})$ with $A = B \cap \mathscr{X}^-$. For A of this form, it follows from Lemma 2.5(d) that $\tau(A) = A$ which, by (P2), implies $A \in \mathscr{F}_{0-}$. This shows that $\mathfrak{B}(\mathscr{X}^-) \subset \mathscr{F}_{0-}$ and thus $\mathfrak{B}(\mathscr{X}^-)$ is contained in the trace of $\mathscr{F}_{0-}$ on $\mathscr{X}^-$. The converse inclusion follows from considering $B \in \mathscr{F}_{0-}$.

(b). Assume that Φ is $\mathfrak{B}(\mathscr{X}^-)$-measurable. By (a), for every $A \in \Sigma$ it is $B := \Phi^{-1}(A) \in \mathfrak{B}(\mathscr{X}^-) \subset \mathscr{F}_{0-}$. By (P2), $\hat{\Phi}^{-1}(A) = \tau^{-1}(\Phi^{-1}(A)) = \tau^{-1}(B) = B$, proving that $\hat{\Phi}$ is $\mathscr{F}_{0-}$-measurable.

Conversely assume that $\hat{\Phi}$ is $\mathscr{F}_{0-}$-measurable. Then for every $A \in \Sigma$ we have $B := \hat{\Phi}^{-1}(A) \in \mathscr{F}_{0-}$, whence $\tau^{-1}(B) = B$. This implies $\tau(B) = B$ and thus $B \subset \mathscr{X}^-$ in view of Lemma 2.5(d). In particular, $B = B \cap \mathscr{X}^- \in \mathfrak{B}(\mathscr{X}^-)$. It follows that $\tau^{-1}(\Phi^{-1}(A)) = \tau^{-1}(B)$ and thus $\Phi^{-1}(A) = B$ proving that Φ is $\mathfrak{B}(\mathscr{X}^-)$-measurable. $\square$

3. The shift semigroup

Throughout, $((\mathscr{X}, \mathbb{d}), \tau)$ is a path space. We next introduce the *shift group* $(\Theta_t)_{t\in\mathbb{R}} \subset \mathscr{L}(B_b(\mathscr{X}), \sigma)$ by setting

$$(\Theta_t F)(\mathbb{x}) := F(\vartheta_t \mathbb{x})$$

for $t \in \mathbb{R}$ and $F \in B_b(\mathscr{X})$. Here, $\mathscr{L}(B_b(\mathscr{X}), \sigma)$ refers to the space of bounded kernel operators, see Section A.1 and, in particular, Lemma A.1.

Lemma 2.5(a) immediately yields

Corollary 3.1. *Let $I \subset \mathbb{R}$ be an interval and $t \in \mathbb{R}$. If $F \in B_b(\mathscr{X}, \mathscr{F}(I))$, then $\Theta_t F \in B_b(\mathscr{X}, \mathscr{F}(I + t))$.*

Let us briefly recall some notions concerning semigroups. A C_b*-semigroup* is a family $\mathbb{S} = (\mathbb{S}_t)_{t\geq 0}$ of Markovian kernel operators on $\mathscr{X}$ that leave the space $C_b(\mathscr{X})$ invariant and such that the orbit of every bounded, continuous function is jointly continuous in t and $\mathbb{x}$. The C_b*-generator* $\mathbb{A}$ of a C_b-semigroup $\mathbb{S}$ is defined as follows. We have $F \in D(\mathbb{A})$ and $\mathbb{A}F = G$ if and only if $\sup_{t\in(0,1)} t^{-1}\|\mathbb{S}(t)F - F\|_\infty < \infty$ and $t^{-1}(\mathbb{S}(t)F(\mathbb{x}) - F(\mathbb{x})) \to G(\mathbb{x})$ for all $\mathbb{x} \in \mathscr{X}$ as $t \to 0$. See Theorem A.12 for equivalent descriptions of the C_b-generator.

Proposition 3.2. *Given (P1) and (P2), condition (P3) is fulfilled if and only if $(\Theta_t)_{t\in\mathbb{R}}$ is a C_b-group. In this case, the C_b-generator of $(\Theta_t)_{t\in R}$ is denoted by $\mathbb{D}$ and the following hold true:*

(a) *$\mathbb{D}$ is a* derivation, *i.e. for $F, G \in D(\mathbb{D})$ also the product $FG \in D(\mathbb{D})$ and $\mathbb{D}(FG) = (\mathbb{D}F)G + F(\mathbb{D}G)$;*
(b) *If $F \in D(\mathbb{D})$ is determined before (after) time t, then so is $\mathbb{D}F$.*

Proof. If (P3) is satisfied then for every $F \in C_b(\mathscr{X})$ the map $(t, \mathbb{x}) \mapsto (\Theta_t F)(\mathbb{x}) = F(\vartheta_t \mathbb{x})$ is continuous which shows that $(\Theta_t)_{t\in\mathbb{R}}$ is a C_b-group.

To see the converse, assume that there exist sequences $t_n \to t$ and $\mathbb{x}_n \to \mathbb{x}$ such that $\mathbb{d}(\vartheta_{t_n}\mathbb{x}_n, \vartheta_t\mathbb{x}) \geq \varepsilon > 0$ for all $n \in \mathbb{N}$. As $\mathscr{X}$ is a metric space, we find a bounded continuous function F on $\mathscr{X}$ such that $F(\vartheta_t\mathbb{x}) = 1$ whereas $F(\mathbb{y}) = 0$ whenever $\mathbb{d}(\mathbb{y}, \vartheta_t\mathbb{x}) \geq \varepsilon$. For this function F we have

$$0 \equiv (\Theta_{t_n} F)(\mathbb{x}_n) \not\to (\Theta_t F)(\mathbb{x}) = 1.$$

Thus, for this particular F the map $(t, \mathbb{x}) \mapsto [\Theta_t F](\mathbb{x})$ is not continuous, whence $(\Theta_t)_{t\in\mathbb{R}}$ is not a C_b-group.

(a). Let $F, G \in D(\mathbb{D})$. Using the characterization of the C_b-generator from Theorem A.12(iv), we find

$$\frac{\Theta_t(FG) - FG}{t}(\mathbb{x}) = (\Theta_t F)(\mathbb{x})\frac{(\Theta_t G)(\mathbb{x}) - G(\mathbb{x})}{t} + \frac{(\Theta_t F)(\mathbb{x}) - F(\mathbb{x})}{t}G(\mathbb{x})$$

which is uniformly bounded and converges to $F(\mathbb{x})[\mathbb{D}G](\mathbb{x}) + [\mathbb{D}F](\mathbb{x})G(\mathbb{x})$ as $t \to 0$.

(b). Let $F \in D(\mathbb{D})$ be $\mathscr{F}_{t-}$-measurable. By Corollary 3.1 $\Theta_{-s}F$ is $\mathscr{F}_{t-}$-measurable for every $s \geq 0$. Thus, for every $s > 0$ the difference quotient $(\Theta_{-s}F - F)/(-s)$ is $\mathscr{F}_{t-}$-measurable hence so is the (pointwise) limit $\mathbb{D}F$.

In the case where F is determined after time t we can proceed similarly, considering the difference quotients $(\Theta_s F - F)/s$ for $s > 0$ instead. $\square$

Besides the C_b-generator $\mathbb{D}$ also the *full generator* $\mathbb{D}_{\mathsf{full}}$ of $(\Theta_t)_{t\in\mathbb{R}}$ is of interest. This operator is typically multivalued and is defined via the Laplace transform of the semigroup on $B_b(\mathscr{X})$. Consequently, it may contain functions that are not continous, which will be important in what follows. For more information about the full generator, we refer to Section A.2.

Corollary 3.3. *The full generator $\mathbb{D}_{\mathsf{full}}$ of $(\Theta_t)_{t\in\mathbb{R}}$ is a derivation in the sense that if $(F_j, G_j) \in \mathbb{D}_{\mathsf{full}}$ for $j = 1, 2$, then also $(F_1F_2, F_1G_2 + F_2G_1) \in \mathbb{D}_{\mathsf{full}}$.*

Proof. Using the characterization of the full generator from Proposition A.4(iii), this follows from the multiplicativity of $(\Theta_t)_{t\in\mathbb{R}}$ and the product rule for Sobolev functions. $\square$

Next, we introduce some specific elements of $\mathbb{D}_{\mathsf{full}}$. Given $f \in B_b(X)$ and $a < b$ and $t \in \mathbb{R}$, we define $F_a^b(f)$ and $F_t(f)$ by setting

$$[F_a^b(f)](\mathbb{x}) := \int_a^b f(\mathbb{x}(s))\, ds \quad \text{and} \tag{3.1}$$

$$[F_t(f)](\mathbb{x}) := f(\mathbb{x}(t)) \tag{3.2}$$

for every $\mathbb{x} \in \mathscr{X}$. Obviously, $F_a^b(f)$ and $F_t(f)$ are bounded measurable functions and $F_a^b(f) \in B_b(\mathscr{X}, \mathscr{F}([a,b)))$. On the other hand, in general $F_t(f)$ is not $\mathscr{F}_{t-}$-measurable. As this leads to technical problems in what follows, we will often replace $F_t(f)$ by the function $F_t^\star(f)$, defined by

$$[F_t^\star(f)](\mathbb{x}) := \limsup_{n\to\infty} n \int_{t-2/n}^{t-1/n} f(\mathbb{x}(s))\, ds. \tag{3.3}$$

In the case where $\mathscr{X} \in \{\mathscr{X}_{\mathrm{C}}, \mathscr{X}_{\mathrm{D}}\}$, we set

$$t^\star := \begin{cases} t, & \text{if } \mathscr{X} = \mathscr{X}_{\mathrm{C}}, \\ t-, & \text{if } \mathscr{X} = \mathscr{X}_{\mathrm{D}}. \end{cases}$$

For future reference, we collect some easy properties of $F_t^\star(f)$.

Lemma 3.4. *Given $t \in \mathbb{R}$ and $f \in B_b(X)$, the following hold true:*
(a) *$F_t^\star(f)$ is $\mathscr{F}_{t-}$-measurable;*
(b) *If $f \in C_b(X)$ and t is a continuity point of $\mathbb{x}$, then $[F_t^\star(f)](\mathbb{x}) = [F_t(f)](\mathbb{x})$;*
(c) *If $\mathscr{X} \in \{\mathscr{X}_{\mathrm{C}}, \mathscr{X}_{\mathrm{D}}\}$, then $F_t^\star(f) = F_{t^\star}(f)$ for all $t \in \mathbb{R}$ and $f \in C_b(X)$;*
(d) *For $s \in \mathbb{R}$, it is $\Theta_s F_t^\star(f) = F_{t+s}^\star(f)$.*

Lemma 3.5. *Let $f \in C_b(X)$ and $a < b$ then $(F_a^b(f), F_b^\star(f) - F_a^\star(f)) \in \mathbb{D}_{\mathsf{full}}$.*

Proof. For $t > 0$, we have

$$\begin{aligned}
[\Theta_t F_a^b(f) - F_a^b(f)](\mathbb{x}) &= \int_a^b f(\mathbb{x}(t+s))\, ds - \int_a^b f(\mathbb{x}(s))\, ds \\
&= \int_b^{b+t} f(\mathbb{x}(s))\, ds - \int_a^{a+t} f(\mathbb{x}(s))\, ds \\
&= \int_0^t f(\mathbb{x}(b+s))\, ds - \int_0^t f(\mathbb{x}(a+s))\, ds \\
&= \int_0^t [\Theta_s(F_b(f) - F_a(f))](\mathbb{x})\, ds \\
&= \int_0^t [\Theta_s(F_b^\star(f) - F_a^\star(f))](\mathbb{x})\, ds.
\end{aligned}$$

Here, we have used Lemma 3.4(b) and the continuity assumption in (P1) in the last equality. Now Proposition A.4 yields $(F_a^b(f), F_b^\star(f) - F_a^\star(f)) \in \mathbb{D}_{\mathsf{full}}$. □

Definition 3.6. We define the operator $\mathbb{D}_0$ as follows: $D(\mathbb{D}_0)$ is the algebra generated by all functions of the form $F_a^b(f)$, where $a < b$ and $f \in C_b(X)$ and

$$\mathbb{D}_0 \prod_{j=1}^n F_{a_j}^{b_j}(f_j) = \sum_{k=1}^n (F_{b_k}^\star(f_k) - F_{a_k}^\star(f_k)) \prod_{j\neq k} F_{a_j}^{b_j}(f_j).$$

Then $\mathbb{D}_0$ is a slice (see Definition A.7) of $\mathbb{D}_{\mathsf{full}}$ by Lemma 3.5 and Proposition 3.2(b).

Remark 3.7. (a) If $f \in C_b(X)$, then $F_a^b(f) \in C_b(\mathscr{X})$ as a consequence of the continuity requirement in (P1). Note, however, that in general $F_b(f) - F_a(f) \notin C_b(\mathscr{X})$. However, if $\mathscr{X} = \mathscr{X}_{\mathrm{C}}$, see Section B.1, then $F_a^b(f) \in D(\mathbb{D})$ for $f \in C_b(X)$ and it follows that $D(\mathbb{D}_0) \subset D(\mathbb{D})$, i.e. it is a subset of the domain of the C_b-generator $\mathbb{D}$.

(b) In the general case, there are several choices of functions $G \in B_b(\mathscr{X})$ such that $(F_a^b(f), G) \in \mathbb{D}_{\mathsf{full}}$. For the slice $\mathbb{D}_0$, we choose $G = F_b^\star(f) - F_a^\star(f)$ and the latter is $\mathscr{F}_{b-}$-measurable. Thus, Proposition 3.2(b) generalizes to the slice $\mathbb{D}_0$. The proof of Lemma 3.5 shows that another possible choice is $G = F_b(f) - F_a(f)$ which is general not $\mathscr{F}_{b-}$-measurable.

If $\mathscr{X} = \mathscr{X}_{\mathrm{D}}$, see Subsection B.2, then for every $\lambda \in [0,1]$ we may choose

$$G(\mathbb{x}) = \lambda f(\mathbb{x}(b)) + (1-\lambda) f(\mathbb{x}(b-)) - [\lambda f(\mathbb{x}(a)) + (1-\lambda) f(\mathbb{x}(a-))].$$

Recall that a sequence $(F_n)_{n\in\mathbb{N}} \subset B_b(\mathscr{X})$ *bp-converges* to $F \in B_b(\mathscr{X})$ if it is uniformly bounded and and converges pointwise to F. Kernel operators are well-behaved with respect to bp-convergence, see Lemma A.1. A subset $M \subset B_b(\mathscr{X})$ is called *bp-closed* if for every sequence $(F_n)_{n\in\mathbb{N}} \subset M$ that bp-converges to F it follows that $F \in M$. The *bp-closure* of a set is the smallest bp-closed set that contains it. If the bp-closure of a set is all of $B_b(\mathscr{X})$, it is called *bp-dense* in $B_b(\mathscr{X})$.

Lemma 3.8. *$D(\mathbb{D}_0)$ is bp-dense in $B_b(\mathscr{X})$.*

Proof. Denote by M the bp-closure of the algebra $D(\mathbb{D}_0)$. Standard arguments show that M is also an algebra. By the right-continuity of the paths in $\mathscr{X}$ it follows that for every $f \in C_b(X)$, it is $nF_t^{t+1/n}(f) \to F_t(f)$ pointwise. Consequently, $F_t(f) \in M$ for all $f \in C_b(X)$ and $t \in \mathbb{R}$. Using that $C_b(X)$ is bp-dense in $B_b(X)$, it follows that $F_t(f) \in M$ for all $f \in B_b(X)$ and $t \in \mathbb{R}$. As M is an algebra, for all choices of $t_1 < \ldots < t_n \in \mathbb{R}$ and $A_1, \ldots, A_n \in \mathfrak{B}(X)$ we have

$$\mathbb{1}_{\{\mathbb{x} \in \mathscr{X} \,:\, \mathbb{x}(t_j) \in A_j \text{ for } j=1,\ldots,n\}} = \prod_{j=1}^n F_{t_j}(\mathbb{1}_{A_j}) \in M,$$

i.e. indicator functions of cylinder sets belong to M. As the cylinder sets form a generator of the Borel σ-algebra that is stable under intersections, measure theoretic induction yields $M = B_b(\mathscr{X})$. □

We end this section by establishing that the shift group is uniquely determined by the slice $\mathbb{D}_0$. This is done by generalizing the concept of a core of an operator, see [19, Definition II.1.6]. As we are using both the concept of the C_b-generator and the concept of the full generator, there are two different concepts of a core. For the former, the appropriate concept is that of a β_0-core (see Lemma A.14), for the latter the notion of a bp-core (see Lemma A.6) is used instead.

Corollary 3.9. *$D(\mathbb{D}_0)$ a bp-core for $\mathbb{D}_{\mathsf{full}}$. If $\mathscr{X} = \mathscr{X}_{\mathrm{C}}$, then $D(\mathbb{D}_0)$ is a β_0-core for $\mathbb{D}$.*

Proof. This follows from Corollary A.8. Indeed, noting that $\Theta_t F_a^b(f) = F_{a+t}^{b+t}(f)$ and $\Theta_t F_s(f) = F_{t+s}(f)$ condition (i) is satisfied, whereas (ii) follows from Lemma 3.8. Condition (iii) is an immediate consequence of the fact that for $f \in C_b(X)$ the map $t \mapsto f \circ \pi_t$ is continuous in almost every point as a consequence of (P1).

Now consider the case of continuous paths. As $D(\mathbb{D}_0)$ is an algebra that separates the points of $\mathscr{X}$, the Stone–Weierstraß Theorem, see [26, Theorem 11], yields that $D(\mathbb{D}_0)$ is dense in $C_b(\mathscr{X})$ with respect to β_0. As $\Theta_t D(\mathbb{D}_0) \subset D(\mathbb{D}_0)$ for all $t \in \mathbb{R}$, it follows from Lemma A.14 that $D(\mathbb{D}_0)$ is a β_0-core for $\mathbb{D}$. □

Remark 3.10. Even in the case where $\mathscr{X} = \mathscr{X}_{\mathrm{C}}$, we cannot expect any function of the form $F_0(f)$ (or, more generally, $F_t(f)$) which is not constant to belong to $D(\mathbb{D})$. To see this, we consider the case $X = \mathbb{R}^d$. Assume that $F_0(f)$ belongs to $D(\mathbb{D})$. Fixing $x_0, v \in \mathbb{R}^d$, we consider $\varkappa(t) = x_0 + |t|v$. It follows that

$$[\mathbb{D}F_0(f)](\varkappa) = \lim_{t\downarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} = \lim_{t\uparrow 0} \frac{f(x_0 - tv) - f(x_0)}{t} = -[\mathbb{D}F_0(f)](\varkappa).$$

This implies that f has in the point x_0 directional derivative 0 in direction v. As x_0 and v were arbitrary, f must be constant.

4. Evolutionary semigroups on spaces of measurable functions

We are now prepared to introduce the central notions of this article, namely *evolutionary semigroups* and their associated *expectation operators*. We start with the latter and first prove that several possible defining properties are in fact equivalent. Throughout this section $((\mathscr{X}, \mathbb{d}), \tau)$ is a path space.

Proposition 4.1. *Let $\mathbb{E} \in \mathscr{L}(B_b(\mathscr{X}), \sigma)$ be a Markovian kernel operator with associated kernel $\Bbbk$. The following are equivalent:*

(i) *For every $F \in B_b(\mathscr{X})$, the function $\mathbb{E}F$ is $\mathscr{F}_{0-}$-measurable and if F is $\mathscr{F}_{0-}$-measurable, then $\mathbb{E}F = F$.*
(ii) *For every $F \in C_b(\mathscr{X})$, the function $\mathbb{E}F$ is $\mathscr{F}_{0-}$-measurable and if F is $\mathscr{F}_{0-}$-measurable, then $\mathbb{E}F = F$.*
(iii) *Given $A_- \in \mathscr{F}_{0-}$, $A \in \mathfrak{B}(\mathscr{X})$ and $\varkappa \in \mathscr{X}$,*

$$\Bbbk(\varkappa, A_- \cap A) = \delta_{\tau(\varkappa)}(A_-)\Bbbk(\tau(\varkappa), A).$$

(iv) *For every $F \in B_b(\mathscr{X})$, the function $\mathbb{E}F$ is $\mathscr{F}_{0-}$-measurable. Moreover,*

$$\mathbb{E}(FG) = F\mathbb{E}G,$$

for all $F, G \in B_b(\mathscr{X})$ where F is $\mathscr{F}_{0-}$-measurable.

Proof. (i) ⇒ (ii). This is trivial.

(ii) ⇒ (i). Set

$$M = \{F \in B_b(\mathscr{X}) : \mathbb{E}F \text{ is } \mathscr{F}_{0-}\text{-measurable}\}.$$

By (ii), $C_b(\mathscr{X}) \subset M$. Moreover, using that $\mathbb{E}$ is σ-continuous (see Lemma A.1), it is easy to see that if $(F_n)_{n\in\mathbb{N}}$ is a bounded sequence in M that converges pointwise to F, then also $F \in M$. This proves that M is bp-closed. As $\mathscr{X}$ is Polish, $C_b(\mathscr{X})$ is bp-dense in $B_b(\mathscr{X})$ (see [20, Proposition 3.4.2]) whence $M = B_b(\mathscr{X})$.

Similarly, one also sees that the second property extends to $B_b(\mathscr{X})$.

(i) ⇒ (iii). As $\mathbb{E}F$ is always $\mathscr{F}_{0-}$-measurable it follows from Lemma 2.5 that

$$\Bbbk(\varkappa, A) = (\mathbb{E}\mathbb{1}_A)(\varkappa) = (\mathbb{E}\mathbb{1}_A)(\tau(\varkappa)) = \Bbbk(\tau(\varkappa), A)$$

for all $A \in \mathscr{B}(\mathscr{X})$ and $\varkappa \in \mathscr{X}$.

Now let $A_- \in \mathscr{F}_{0-}$, $A \in \mathfrak{B}(\mathscr{X})$ and $\varkappa \in \mathscr{X}$ be given. By (i), $\Bbbk(\varkappa, A_-) = \mathbb{E}\mathbb{1}_{A_-}(\varkappa) = \mathbb{1}_{A_-}(\varkappa)$. It follow that for $\varkappa \notin A_-$

$$0 \le \Bbbk(\varkappa, A_- \cap A) \le \Bbbk(\varkappa, A_-) = 0.$$

On the other hand, if $\mathbb{x} \in A_-$, the same argument shows $\Bbbk(\mathbb{x}, A_-^c \cap A) = 0$ and we obtain

$$\Bbbk(\mathbb{x}, A) = \Bbbk(\mathbb{x}, A_- \cap A) + \Bbbk(\mathbb{x}, A_-^c \cap A) = \Bbbk(\mathbb{x}, A_- \cap A).$$

Altogether,

$$\Bbbk(\mathbb{x}, A_- \cap A) = \Bbbk(\tau(\mathbb{x}), A_- \cap A) = \delta_{\tau(\mathbb{x})}(A_-)\Bbbk(\tau(\mathbb{x}), A).$$

(iii) ⇒ (iv). Condition (iii) yields $\mathbb{E}(FG) = F\mathbb{E}G$ whenever $F = \mathbb{1}_{A_-}$ for some $A_- \in \mathscr{F}_{0-}$ and $G = \mathbb{1}_A$ for some $A \in \mathfrak{B}(\mathscr{X})$. Using measure theoretic induction twice, this easily extends to arbitrary functions.

(iv) ⇒ (i). Let $F \in B_b(\mathscr{X})$ be $\mathscr{F}_{0-}$-measurable. As $\mathbb{1}$ is $\mathfrak{B}(\mathscr{X})$-measurable and $\mathbb{E}\mathbb{1} = \mathbb{1}$ by Markovianity, (iv) yields $\mathbb{E}F = \mathbb{E}(F\mathbb{1}) = F\mathbb{E}\mathbb{1} = F\mathbb{1} = F$. □

Note that condition (iv) of Proposition 4.1 implies that $\mathbb{E}$ is local in the sense that $\mathbb{E}(\mathbb{1}_A G) = \mathbb{1}_A \mathbb{E}G$ for all $A \in \mathscr{F}_{0-}$ and $G \in B_b(\mathscr{X})$, and can therefore be viewed as a conditional function/expectation, see [16] and the references therein.

Definition 4.2. An operator $\mathbb{E}$ satisfying the equivalent conditions of Proposition 4.1 is called *expectation operator*. Given an expectation operator $\mathbb{E}$, we define $\mathbb{E}_t := \Theta_t \mathbb{E} \Theta_{-t}$ for $t \geq 0$. We say that $\mathbb{E}$ is *homogeneous* if $\mathbb{E} = \mathbb{E}\mathbb{E}_t$ for all $t \geq 0$.

Remark 4.3. In the case of continuous paths, i.e. $((\mathscr{X}, \mathbb{d}), \tau) = ((\mathscr{X}_{\mathrm{C}}, \mathbb{d}_{\mathrm{C}}), \tau_{\mathrm{C}})$, we give an alternative representation of $\mathbb{E}$ that follows from Proposition 4.1(iii). To that end, we identify $\mathscr{X}^- := \tau(\mathscr{X})$ with $C((-\infty, 0]; X)$ and define $\mathscr{X}^+ := C([0, \infty); X)$. Then the map

$$\mathscr{X}^- \times \mathscr{X}^+ \to \mathscr{X}, \ (\mathbb{x}_-, \mathbb{x}_+) \mapsto \mathbb{x}_- \otimes_0 \mathbb{x}_+$$

with

$$(\mathbb{x}_- \otimes_0 \mathbb{x}_+)(t) := \begin{cases} \mathbb{x}_-(t), & \text{if } t \leq 0, \\ \mathbb{x}_+(t) - \mathbb{x}_+(0) + \mathbb{x}_-(0), & \text{if } t \geq 0, \end{cases}$$

is well-defined, and its restriction to the closed subspace of all compatible pairs

$$\mathscr{X}_{\mathrm{comp}} := \{(\mathbb{x}_-, \mathbb{x}_+) \in \mathscr{X}^- \times \mathscr{X}^+ : \mathbb{x}_-(0) = \mathbb{x}_+(0)\}$$

is an isomorphism. Let $\mathbb{E} \in \mathscr{L}(B_b(\mathscr{X}), \sigma)$ be a Markovian kernel operator with associated kernel $\Bbbk$. Identifying $\mathscr{X}$ and $\mathscr{X}_{\mathrm{comp}}$, Proposition 4.1(iii) yields

$$\Bbbk((\mathbb{x}_-, \mathbb{x}_+), A_- \times A_+) = \delta_{\mathbb{x}_-}(A_-)\Bbbk_+(\mathbb{x}_-, A_+) \tag{4.1}$$

for $(\mathbb{x}_-, \mathbb{x}_+) \in \mathscr{X}_{\mathrm{comp}}$ and $A_\pm \in \mathfrak{B}(\mathscr{X}^\pm)$, where we define

$$\begin{aligned} \Bbbk_+(\mathbb{x}_-, A_+) &:= \Bbbk\big(\mathbb{x}_-, \{\mathbb{x}_- \otimes_0 \mathbb{y}_+ : \mathbb{y}_+ \in A_+, \, \mathbb{x}_-(0) = \mathbb{y}_+(0)\}\big) \\ &= \Bbbk\big(\mathbb{x}_-, (\{\mathbb{x}_-\} \times A_+) \cap \mathscr{X}_{\mathrm{comp}}\big). \end{aligned}$$

Note that by (4.1), the measure $\Bbbk(\mathbb{x}, \cdot)$ is extended by zero to $(\mathscr{X}^- \times \mathscr{X}^+) \setminus \mathscr{X}_{\mathrm{comp}}$. For $F \in B_b(\mathscr{X})$, (4.1) and a Fubini argument yield

$$\begin{aligned} (\mathbb{E}F)(\mathbb{x}) &= \int_{\mathscr{X}} F(\mathbb{y})\Bbbk(\mathbb{x}, d\mathbb{y}) = \int_{\mathscr{X}^- \times \mathscr{X}^+} F(\mathbb{y}_- \otimes_0 \mathbb{y}_+)\Bbbk((\mathbb{x}_-, \mathbb{x}_+), d(\mathbb{y}_-, \mathbb{y}_+)) \\ &= \int_{\mathscr{X}^-} \int_{\mathscr{X}^+} F(\mathbb{y}_- \otimes_0 \mathbb{y}_+)\Bbbk_+(\mathbb{x}_-, d\mathbb{y}_+)\delta_{\mathbb{x}_-}(d\mathbb{y}_-) \\ &= \int_{\mathscr{X}^+} F(\mathbb{x}_- \otimes_0 \mathbb{y}_+)\Bbbk_+(\mathbb{x}_-, d\mathbb{y}_+). \end{aligned} \tag{4.2}$$

If we want to extend this construction to càdlàg paths, it is a natural question how to glue two functions from $D([0, \infty); X)$ and $D((-\infty, 0]; X)$ together at 0, i.e. if there might be a jump at 0 or not. If the evolutionary semigroup induced by $\mathbb{E}$ is a C_b-semigroup, then Corollary 5.11 below yields that

$$\mathbb{P}^{\mathbb{x}}(\{\mathbb{y} : \mathbb{y} \text{ is continuous at } 0\}) = 1,$$

which suggests that we should glue paths $\mathbb{x}_- \in D((-\infty,0];X)$ and $\mathbb{x}_+ \in D([0,\infty);X)$ continuously together in 0. Thus, mutatis mutandis, the same construction as above can be extended to càdlàg paths in this situation.

The above description of $\mathbb{E}$ will be used for stochastic (delay) differential equations below, where we specify $\mathbb{k}_+$ in terms of the unique solution of the equation, see Section 6 for details.

Lemma 4.4. *Let $\mathbb{E}$ be an expectation operator with associated kernel $\mathbb{k}$. The following are equivalent:*

(i) *$\mathbb{E}$ is homogeneous.*
(ii) *For every $\mathbb{x} \in \mathscr{X}$, $t \geq 0$ and $A \in \mathscr{B}(\mathscr{X})$,*

$$\int_{\mathscr{X}} \mathbb{k}(\vartheta_t \mathbb{y}, \vartheta_t A)\mathbb{k}(\mathbb{x}, d\mathbb{y}) = \mathbb{k}(\mathbb{x}, A).$$

Proof. Let $F \in B_b(\mathscr{X})$. Note that

$$(\mathbb{E}\Theta_{-t}F)(\mathbb{x}) = \int_{\mathscr{X}} F(\vartheta_{-t}\mathbb{y})\mathbb{k}(\mathbb{x}, d\mathbb{y}) = \int_{\mathscr{X}} F(\mathbb{z})\mathbb{k}(\mathbb{x}, d\mathbb{z} \circ (\vartheta_{-t})^{-1}),$$

where we write $\mathbb{k}(\mathbb{x}, d\mathbb{z} \circ (\vartheta_{-t})^{-1})$ for the push-forward measure of $\mathbb{k}(\mathbb{x}, \cdot)$ under the map ϑ_{-t}. Therefore

$$(\mathbb{E}_t F)(\mathbb{x}) = (\Theta_t \mathbb{E}\Theta_{-t}F)(\mathbb{x}) = (\mathbb{E}\Theta_{-t}F)(\vartheta_t \mathbb{x}) = \int_{\mathscr{X}} F(\mathbb{z})\mathbb{k}(\vartheta_t \mathbb{x}, d\mathbb{z} \circ (\vartheta_{-t})^{-1}).$$

It follows that

$$\begin{aligned}(\mathbb{E}\mathbb{E}_t F)(\mathbb{x}) &= \int_{\mathscr{X}} (\mathbb{E}_t F)(\mathbb{y})\mathbb{k}(\mathbb{x}, d\mathbb{y}) \\ &= \int_{\mathscr{X}}\int_{\mathscr{X}} F(\mathbb{z})\mathbb{k}(\vartheta_t \mathbb{y}, d\mathbb{z} \circ (\vartheta_{-t})^{-1})\mathbb{k}(\mathbb{x}, d\mathbb{y}) \\ &= \int_{\mathscr{X}} F(\mathbb{z}) \int_{\mathscr{X}} \mathbb{k}(\vartheta_t \mathbb{y}, d\mathbb{z} \circ (\vartheta_{-t})^{-1})\mathbb{k}(\mathbb{x}, d\mathbb{y}).\end{aligned}$$

This is the same as $(\mathbb{E}F)(\mathbb{x})$ for all $F \in B_b(\mathscr{X})$ if and only if

$$\int_{\mathscr{X}} \mathbb{1}_A(\mathbb{z}) \int_{\mathscr{X}} \mathbb{k}(\vartheta_t \mathbb{y}, d\mathbb{z} \circ (\vartheta_{-t})^{-1})\mathbb{k}(\mathbb{x}, d\mathbb{y}) = \int_{\mathscr{X}} \mathbb{k}(\vartheta_t \mathbb{y}, \vartheta_t A)\mathbb{k}(x, d\mathbb{y}) = \mathbb{k}(\mathbb{x}, A)$$

for all $\mathbb{x} \in \mathscr{X}$ and $A \in \mathscr{B}(\mathscr{X})$. □

Theorem 4.5. *Let $(\mathbb{T}(t))_{t\geq 0}$ be a semigroup of Markovian kernel operators on $B_b(\mathscr{X}, \mathscr{F}_{0-})$. The following are equivalent:*

(i) *For every $F \in B_b(\mathscr{X}, \mathscr{F}_{0-})$ and $t \geq 0$, it is $\mathbb{T}(t)\Theta_{-t}F = F$. In particular, if F is determined before $-t_0 < 0$, then $\mathbb{T}(t)F = \Theta_t F$ for all $0 \leq t \leq t_0$.*
(ii) *There exists a homogeneous expectation operator $\mathbb{E}$ such that*

$$\mathbb{T}(t) = \mathbb{E}\Theta_t \quad \textit{for all } t \geq 0. \tag{4.3}$$

Conversely, if $\mathbb{E}$ is a homogeneous expectation operator, then Equation (4.3) defines a semigroup of Markovian kernel operators on $B_b(\mathscr{X}, \mathscr{F}_{0-})$, which satisfies condition (i).

Proof. (i) ⇒ (ii). Fix $s \geq 0$. If F is determined before s, Lemma 2.5 implies that $\Theta_{-s}F$ is determined before 0. Thus, the expression $\mathbb{T}(s)\Theta_{-s}F$ is well-defined and $F \mapsto \mathbb{T}(s)\Theta_{-s}F$ defines a σ-continuous mapping from $B_b(\mathscr{X}, \mathscr{F}_{s-})$ to $B_b(\mathscr{X}, \mathscr{F}_{0-})$. Consequently, we find a kernel $\mathbb{k}_s$ such that

$$\mathbb{T}(s)\Theta_{-s}F(\mathbb{x}) = \int_{\mathscr{X}} F(\mathbb{y})\mathbb{k}_s(\mathbb{x}, d\mathbb{y})$$

for all $F \in B_b(\mathscr{X}, \mathscr{F}_{s-})$. It follows from Lemma 2.5 that for $F \in B_b(\mathscr{X}, \mathscr{F}_{s-})$ the function $\Theta_{-(r+s)}F$ is determined before $-r < 0$. Consequently, (i) implies that $\Theta_r\Theta_{-(r+s)}F = \mathbb{T}(r)\Theta_{-(r+s)}F$ and thus

$$\mathbb{T}(s)\Theta_{-s}F = \mathbb{T}(s)\Theta_r\Theta_{-(r+s)}F = \mathbb{T}(s)\mathbb{T}(r)\Theta_{-(r+s)}F = \mathbb{T}(s+r)\Theta_{-(r+s)}F.$$

This shows that $\mathbb{k}_s(\mathbb{x}, A) = \mathbb{k}_{s+r}(\mathbb{x}, A)$ for all $r \geq 0$ and $A \in \mathscr{F}_{s-}$. We may thus define

$$\mathbb{k}(\mathbb{x}, A) := \mathbb{k}_s(\mathbb{x}, A)$$

whenever $A \in \mathscr{F}_{s-}$. In this way, $\mathbb{k}(\mathbb{x}, \cdot)$ defines a finitely additive measure on $\mathscr{D} = \bigcup_{s\geq 0} \mathscr{F}_{s-}$.

Note that $\mathbb{k}_s(\mathbb{x}, \cdot)$ is even σ-additive on $\mathscr{F}_{s-}$ and thus can be viewed as a Radon measure on the Polish space $\mathscr{X}_s = \tau_s(\mathscr{X})$. Consequently, given $\varepsilon > 0$ and $A \in \mathscr{F}_{s-}$, we find a compact subset $K \subset \mathscr{X}_s$ such that $\mathbb{k}_s(\mathbb{x}, A \setminus K) \leq \varepsilon$. Note that K is also a compact subset of $\mathscr{X}$. This shows that given $A \in \mathscr{D}$ and $\varepsilon > 0$ we find a compact subset K of $\mathscr{X}$ that also belongs to $\mathscr{D}$ such that $\mathbb{k}(\mathbb{x}, A \setminus K) \leq \varepsilon$. At this point [6, Theorem 1.4.3] implies that $\mathbb{k}(\mathbb{x}, \cdot)$ is σ-additive on $\mathscr{D}$, whence it can be extended to a measure on $\mathfrak{B}(\mathscr{X})$ by means of the Carathéodory theorem (see [6, Cor. 1.11.9]).

We note that the map $\mathbb{x} \mapsto \mathbb{k}(\mathbb{x}, A)$ is measurable whenever $A \in \mathscr{D}$. By a monotone class argument, this extends to arbitrary $A \in \mathfrak{B}(\mathscr{X})$, proving that $\mathbb{k}$ is a kernel. We may thus define

$$[\mathbb{E}F](\mathbb{x}) := \int_{\mathscr{X}} F(\mathbb{y})\mathbb{k}(\mathbb{x}, d\mathbb{y})$$

for all $F \in B_b(\mathscr{X})$. By construction, $\mathbb{E}F = \mathbb{T}(s)\Theta_{-s}F$ whenever F is determined before $s \geq 0$. Applying this to $F = \Theta_s G$ for some $G \in B_b(\mathscr{X}, \mathscr{F}_{0-})$ yields $\mathbb{T}(s)G = \mathbb{E}\Theta_s G$ for all $G \in B_b(\mathscr{X}, \mathscr{F}_{0-})$. In particular, $\mathbb{E}G = G$ for $G \in B_b(\mathscr{X}, \mathscr{F}_{0-})$.

On the other hand, as $\mathbb{T}(s)\Theta_{-s}F$ is $\mathscr{F}_{0-}$-measurable for all $F \in B_b(\mathscr{X}, \mathscr{F}_s)$ it follows that $\mathbb{k}_s(\mathbb{x}, \cdot) = \mathbb{k}_s(\tau(\mathbb{x}), \cdot)$. As s is arbitrary, $\mathbb{k}(\mathbb{x}, A) = \mathbb{k}(\tau(\mathbb{x}), A)$ whenever $A \in \mathscr{D}$. A monotone class argument extends this to $A \in \mathfrak{B}(\mathscr{X})$ and it follows that $\mathbb{E}F$ is $\mathscr{F}_{0-}$-measurable for all $F \in B_b(\mathscr{X})$. Thus, condition (i) of Proposition 4.1 is satisfied, proving that $\mathbb{E}$ is an expectation operator.

It remains to prove that $\mathbb{E}$ is homogeneous. To that end, note that if $F \in B_b(\mathscr{X}, \mathscr{F}_{0-})$, then

$$\mathbb{E}\Theta_t\mathbb{E}\Theta_s F = \mathbb{T}(t)\mathbb{T}(s)F = \mathbb{T}(t+s)F = \mathbb{E}\Theta_t\Theta_s F$$

for all $t, s \geq 0$. Putting $G = \Theta_s F$, it follows from Lemma 2.5 that G is determined before s and that any G determined before s can be written in this form. This implies that $\mathbb{E}\Theta_t\mathbb{E}G = \mathbb{E}\Theta_t G$ for every G that is determined before some time $s \geq 0$ and thus $\mathbb{E}\Theta_t\mathbb{E}\mathbb{1}_A = \mathbb{E}\Theta_t\mathbb{1}_A$ for every $A \in \mathscr{D} := \bigcup_{s>0} \mathscr{F}_{s-}$. As this is a generator for $\mathfrak{B}(\mathscr{X})$, measure theoretic induction yields $\mathbb{E}\Theta_t\mathbb{E} = \mathbb{E}\Theta_t$ proving that $\mathbb{E}$ is homogeneous.

(ii) $\Rightarrow$ (i). If $\mathbb{T}(t)$ is defined by Equation (4.3), then for $F \in B_b(\mathscr{X}, \mathscr{F}_{0-})$

$$\mathbb{T}(t)\Theta_{-t}F = \mathbb{E}\Theta_t\Theta_{-t}F = \mathbb{E}F = F$$

since $\mathbb{E}$ is an expectation operator. This proves (i). For the addendum observe that if $\mathbb{T}$ is not a priori assumed to be a semigroup but given by Equation (4.3), then the semigroup law follows from the homogeneity of $\mathbb{E}$. Indeed, as $\mathbb{E}$ is homogeneous, $\mathbb{E}\Theta_t\mathbb{E} = \mathbb{E}\Theta_t$ for every $t \geq 0$ and hence

$$\mathbb{T}(t)\mathbb{T}(s) = \mathbb{E}\Theta_t\mathbb{E}\Theta_s = \mathbb{E}\Theta_t\Theta_s = \mathbb{E}\Theta_{t+s} = \mathbb{T}(t+s)$$

for all $t, s \geq 0$. □

Definition 4.6. A semigroup $\mathbb{T}$ of Markovian kernel operators on $B_b(\mathscr{X}, \mathscr{F}_{0-})$ is called *evolutionary* if it satisfies the equivalent conditions of Theorem 4.5. In this case, the operator $\mathbb{E}$ is called the *associated expectation operator* of $\mathbb{T}$.

We next show that the operator $\mathbb{E}_t$ can be interpreted as 'conditional expectations given $\mathscr{F}_{t-}$'. Moreover, we obtain a property similar to the Markov property for evolutionary semigroups.

Proposition 4.7. *Let $(\mathbb{T}(t))_{t\geq 0}$ be an evolutionary semigroup with expectation operator $\mathbb{E}$. We denote the kernel of $\mathbb{E}$ by $\Bbbk$ and write $\mathbb{P}^{\mathbb{x}}$ for the probability measure $\Bbbk(\mathbb{x},\cdot)$ and $\mathbb{E}^{\mathbb{x}}$ for the (conditional) expectation with respect to $\mathbb{P}^{\mathbb{x}}$.*

(a) *For every $F \in \mathbb{B}_b(\mathscr{X})$ and $t \geq 0$, it is*

$$\big[\mathbb{E}^{\mathbb{x}}[F|\mathscr{F}_{t-}]\big](\mathbb{y}) = [\mathbb{E}_t F](\mathbb{y}) \quad \text{for } \mathbb{P}^{\mathbb{x}}\text{-a.e. } \mathbb{y}.$$

(b) *For every $s,t \geq 0$, $F \in B_b(\mathscr{X}, \mathscr{F}_{0-})$ and $\mathbb{x} \in \mathscr{X}$, we have*

$$\big[\mathbb{E}^{\mathbb{x}}[\Theta_{t+s}F\big|\mathscr{F}_{t-}]\big](\mathbb{y}) = \big[\mathbb{T}(s)F\big](\vartheta_t\mathbb{y}) \quad \text{for } \mathbb{P}^{\mathbb{x}}\text{-a.e. } \mathbb{y}.$$

Proof. (a). Fix $F \in B_b(\mathscr{X})$ and $t \geq 0$. If $A \in \mathscr{F}_{t-}$, then $\mathbb{1}_A \in B_b(\mathscr{X}, \mathscr{F}_{t-})$ and thus $\Theta_{-t}\mathbb{1}_A \in B_b(\mathscr{X}, \mathscr{F}_{0-})$. By Proposition 4.1(iv), we have

$$\mathbb{E}[\Theta_{-t}\mathbb{1}_A F] = \mathbb{E}[(\Theta_{-t}\mathbb{1}_A)(\Theta_{-t}F)] = (\Theta_{-t}\mathbb{1}_A)\mathbb{E}[\Theta_{-t}F].$$

Applying Θ_t on both sides, it follows that $\mathbb{E}_t(\mathbb{1}_A F) = \mathbb{1}_A\mathbb{E}_t F$. Using this, we see that for $\mathbb{x} \in \mathscr{X}$

$$\mathbb{E}^{\mathbb{x}}[\mathbb{1}_A F] = [\mathbb{E}(\mathbb{1}_A F)](\mathbb{x}) = [\mathbb{E}\mathbb{E}_t(\mathbb{1}_A F)](\mathbb{x}) = \mathbb{E}[\mathbb{1}_A\mathbb{E}_t F](\mathbb{x}) = \mathbb{E}^{\mathbb{x}}[\mathbb{1}_A\mathbb{E}_t F].$$

Here, we have used homogenity of $\mathbb{E}$ in the second equality. As $A \in \mathscr{F}_{t-}$ was arbitrary, (a) is proved.

(b). Fix $s,t \geq 0$ and $F \in B_b(\mathscr{X}, \mathscr{F}_{0-})$. It follows from (a) that for $\mathbb{P}^{\mathbb{x}}$-a.e. $\mathbb{y}$,

$$\big[\mathbb{E}^{\mathbb{x}}[\Theta_{t+s}F|\mathscr{F}_{t-}]\big](\mathbb{y}) = [\mathbb{E}_t(\Theta_{t+s}F)](\mathbb{y}) = [\Theta_t\mathbb{E}\Theta_s F](\mathbb{y}) = [\mathbb{T}(s)F](\vartheta_t\mathbb{y}). \qquad \square$$

As already mentioned in the introduction, in (stochastic) delay equations the case of finite delay $h > 0$ is of particular interest and in this case, one often uses a space of functions on the interval $[-h, 0]$ as a state space. In our more general framework we identify functions on the interval $[-h, 0]$ with functions on $\mathscr{X}^-$ that are measurable with respect to $\mathscr{F}([-h, 0))$.

Lemma 4.8. *Let $\mathbb{T}$ be an evolutionary semigroup with expectation operator $\mathbb{E}$ and $h > 0$. The following are equivalent:*

(i) $\mathbb{T}(t)B_b(\mathscr{X}, \mathscr{F}([-h, 0))) \subset B_b(\mathscr{X}, \mathscr{F}([-h, 0)))$ *for all $t > 0$.*
(ii) $\mathbb{E}B_b(\mathscr{X}, \mathscr{F}([0, \infty))) \subset B_b(\mathscr{X}, \mathscr{F}([-h, 0)))$.

Proof. (i) $\Rightarrow$ (ii). We prove $\mathbb{E}F \in B_b(\mathscr{X}, \mathscr{F}([-h, 0)))$ for $F \in B_b(\mathscr{X}, \mathscr{F}([0, \infty)))$ in several steps.

Step 1. We prove the assertion for $F = F_t(f)$ with $t \geq 0$ and $f \in C_b(X)$.

To that end, fix $t \geq 0$ and $f \in C_b(X)$. For every $s > 0$ we have

$$\mathbb{E}F^\star_{t+s}(f) = \mathbb{E}\Theta_{t+s}F^\star_0(f) = \mathbb{T}(t+s)F^\star_0(f) \in B_b(\mathscr{X}, \mathscr{F}([-h, 0)))$$

by (i), as $F^\star_0 f \in B_b(\mathscr{X}, \mathscr{F}([-h, 0)))$. For every $\mathbb{x} \in \mathscr{X}$ it is $[F^\star_{t+s}(f)](\mathbb{x}) = [F_{t+s}(f)](\mathbb{x})$ for almost every $s > 0$. By right continuity of the paths, $F^\star_{t+s}(f)$ converges pointwise to $F_t(f)$ as $s \to 0$. Since $\mathbb{E}$ is bp-continuous and $B_b(\mathscr{X}, \mathscr{F}([-h, 0)))$ is bp-closed, it follows that $\mathbb{E}F_t(f) \in B_b(\mathscr{X}, \mathscr{F}([-h, 0)))$ as claimed. Note that this shows in particular that $F^\star_0(f) = \mathbb{E}F^\star_0(f) = \mathbb{E}F_0(f)$.

Step 2. We prove the assertion for $F = \prod_{j=1}^n F_{t_j}(f)$ with $0 \leq t_1 < \ldots < t_n$ and $f_1, \ldots, f_n \in C_b(X)$.

This is proved by induction over n. The case $n = 1$ is exactly Step 1. Assuming the claim to be proved for the product of n functions, let $0 \leq t_1 < \ldots < t_n <$

t_{n+1} and $f_1, \ldots f_{n+1} \in C_b(X)$ be given. We put $G = \prod_{j=2}^{n+1} F_{t_j}(f_j)$, so that $F = \prod_{j=1}^{n+1} F_{t_j}(f_j) = F_{t_1}(f_1)G$. It follows that

$$\begin{aligned} \mathbb{E}F &= \mathbb{E}F_{t_1}(f_1)G = \mathbb{E}\Theta_{t_1}F_0(f_1)\Theta_{-t_1}G = \mathbb{E}\Theta_{t_1}\mathbb{E}F_0(f_1)\Theta_{-t_1}G \\ &= \mathbb{E}\Theta_{t_1}\mathbb{E}F_0^\star(f_1)\Theta_{-t_1}G = \mathbb{E}\Theta_{t_1}F_0^\star(f_1)\mathbb{E}\Theta_{-t_1}G \\ &= \mathbb{T}(t_1)\big[F_0^\star(f_1)\mathbb{E}\Theta_{-t_1}G\big] \in B_b(\mathscr{X}, \mathscr{F}([-h,0))). \end{aligned}$$

In the above calculation, the third equality uses that $\mathbb{E}$ is homogeneous (whence $\mathbb{E}\Theta_{t_1}\mathbb{E} = \mathbb{E}\Theta_{t_1}$), the fourth follows from Step 1 and the fifth from Lemma 4.4(iv). By induction hypothesis, $\mathbb{E}\Theta_{-t_1}G \in B_b(\mathscr{X}, \mathscr{F}([-h,0)))$ so that also $F_0^\star(f_1)\mathbb{E}\Theta_{-t_1}G \in B_b(\mathscr{X}, \mathscr{F}([-h,0)))$. Using our assumption (i) again, it follows that indeed $\mathbb{E}F \in B_b(\mathscr{X}, \mathscr{F}([-h,0)))$.

Step 3. We prove the general case.

By a bp-closedness argument, it follows that the assertion of Step 2 is still valid for $f_1, \ldots, f_n \in B_b(X)$. In particular, $\mathbb{E}F \in B_b(\mathscr{X}, \mathscr{F}([-h,0)))$ whenever $F = \mathbb{1}_A$ where A is a cylinder set of the form

$$A = \{\varkappa : \varkappa(t_j) \in A_j \text{ for } j = 1, \ldots, n\}$$

for some $0 \le t_1 < \ldots < t_n$ and $A_1, \ldots, A_n \in \mathfrak{B}(X)$. As these sets are a generator of $\mathscr{F}([0,\infty))$ which is stable under intersections, another bp-closedness argument shows that $\mathbb{E}\mathbb{1}_A \in B_b(\mathscr{X}, \mathscr{F}([-h,0)))$ for all $A \in \mathscr{F}([0,\infty))$ and the general case follows from linearity and yet another bp-closedness argument.

(ii) $\Rightarrow$ (i). Let $F = \prod_{j=1}^n F_{t_j}(f_j)$ where $-r \le t_1 < \ldots < t_n < 0$ and $f_1, \ldots f_n \in B_b(X)$. Note that $\Theta_t F = \prod_{j=1}^n F_{t_j+t}(f_j)$. We pick the index k such that $t_k + t < 0 \le t_{k+1} + t$. It follows that

$$\mathbb{T}(t)F = \mathbb{E}\Theta_t F = \mathbb{E}\prod_{j=1}^n F_{t_j+t}(f) = \prod_{j=1}^k F_{t_j+t}(f)\mathbb{E}\prod_{j=k+1}^n F_{t_j+t}(f)$$

by Proposition 4.1(iv). At this point, (ii) implies $\mathbb{T}(t)F \in B_b(\mathscr{X}, \mathscr{F}([-r,0)))$. As $\mathscr{F}([-h,0)) = \sigma(\pi_t : -h \le t < 0)$, measure theoretic induction shows $\mathbb{T}(t)F \in B_b(\mathscr{X}, \mathscr{F}([-h,0)))$ for all $F \in B_b(\mathscr{X}, \mathscr{F}([-h,0)))$. ☐

So far, we have not imposed any measurability assumption on the orbits of an evolutionary semigroup. As it turns out, it is automatically satisfied.

Lemma 4.9. *Let $\mathbb{T}$ be an evolutionary semigroup with expectation operator $\mathbb{E}$. Then for every $F \in B_b(\mathscr{X}, \mathscr{F}_{0-})$, the map $(t,\varkappa) \mapsto [\mathbb{T}(t)F](\varkappa)$ is measurable.*

Proof. First consider $F \in C_b(\mathscr{X}, \mathscr{F}_{0-})$. As $(\Theta_t)_{t\ge 0}$ is a C_b-semigroup by Proposition 3.2, if $t_n \to t$, then $\Theta_{t_n}F \to \Theta_t F$ pointwise. As $\mathbb{E}$ is a kernel operator, it follows that

$$[\mathbb{T}(t_n)F](\varkappa) = [\mathbb{E}\Theta_{t_n}F](\varkappa) \to [\mathbb{E}\Theta_t F](\varkappa) = [\mathbb{T}(t)F](\varkappa)$$

for every $\varkappa \in \mathscr{X}$. This shows that for fixed $\varkappa \in \mathscr{X}$ the map $t \mapsto \psi(t,\varkappa) := [\mathbb{T}(t)F](\varkappa)$ is continuous. At the same time, for fixed $t \in [0,\infty)$, the map $\varkappa \mapsto \psi(t,\varkappa)$ is measurable. These two facts imply that ψ is product measurable. Indeed, setting

$$\psi_n(t,\varkappa) = \sum_{j=0}^{n2^n} \mathbb{1}_{[\frac{j}{2^n}, \frac{j+1}{2^n})}(t)\psi\Big(\frac{j}{2^n}, \varkappa\Big),$$

we see that ψ_n is product measurable and converges pointwise to ψ.

Next, consider the set

$$M := \{F \in B_b(\mathscr{X}) : (t,\varkappa) \mapsto [\mathbb{T}(t)(F\circ\tau)](\varkappa) \text{ is measurable }\}.$$

By the above, $C_b(\mathscr{X}) \subset M$. A moments thought shows that M is bp-closed. As $C_b(\mathscr{X})$ is bp-dense in $B_b(\mathscr{X})$ (see [20, Proposition 3.4.2]), it follows that $M =$

$B_b(\mathscr{X})$. As $F \circ \tau \in B_b(\mathscr{X}, \mathscr{F}_{0-})$ for every $F \in B_b(\mathscr{X})$ and every element of $B_b(\mathscr{X}, \mathscr{F}_{0-})$ is of this form (by Lemma 2.5), the claim follows. ☐

It follows from Lemma 4.9 that we are very close to the notion of transition semigroup, see Definition A.3. However, our setting is slightly different from that considered there, as $\mathscr{F}_{0-}$ is not the Borel σ-algebra of $\mathscr{X}$. It is also clear, at least on an intuitive level, that $B_b(\mathscr{X}, \mathscr{F}_{0-})$ does not separate the points in $\mathscr{M}(\mathscr{X})$, the space of all bounded measures on $\mathscr{X}$. Note that the duality between measurable functions and measures lies at the heart of the theory of transition semigroups. Thus, at first glance, we cannot use the theory presented in the appendix.

However, by means of the extension map $F \mapsto \hat{F} = F \circ \tau$, we can identify $B_b(\mathscr{X}, \mathscr{F}_{0-})$ with the space $B_b(\mathscr{X}^-, \mathfrak{B}(\mathscr{X}^-))$, see Lemma 2.6. By means of this identification, we are in the situation considered in the appendix, namely we work on the norming dual pair $(B_b(\mathscr{X}^-), \mathscr{M}(\mathscr{X}^-))$. To switch from $\mathscr{X}$ to $\mathscr{X}^-$, we formally also have to change the semigroup. To wit, if $\mathbb{T}$ is an evolutionary semigroup, we define

$$[\mathbb{T}^-(t)F](\mathbb{x}) := [\mathbb{T}(t)\hat{F}](\mathbb{x}) \quad \text{for all } F \in B_b(\mathscr{X}^-) \text{ and } x \in \mathscr{X}^-. \tag{4.4}$$

It then follows, that $\mathbb{T}^-$ defines a transition semigroup on $B_b(\mathscr{X}^-)$ and we may talk of its Laplace transform, its full generator, continuity properties etc. However, in order not to overburden notation, we will not distinguish between the semigroup $\mathbb{T}$ and the modified semigroup $\mathbb{T}^-$.

We may now describe the full generator $\mathbb{A}_{\mathsf{full}}$ of an evolutionary semigroup in terms of the expectation operator $\mathbb{E}$ and the full generator $\mathbb{D}_{\mathsf{full}}$ of the shift semigroup. It turns out that $\mathbb{A}_{\mathsf{full}}$ is already uniquely determined by the expectation operator $\mathbb{E}$ and the slice $\mathbb{D}_0$ from Definition 3.6. We use the following notation:

$$D_-(\mathbb{D}_0) = D(\mathbb{D}_0) \cap B_b(\mathscr{X}, \mathscr{F}_{0-}) \text{ and } D_+(\mathbb{D}_0) = D(\mathbb{D}_0) \cap B_b(\mathscr{X}, \mathscr{F}([0,\infty))).$$

Proposition 4.10. *Let $\mathbb{T}$ be an evolutionary semigroup with expectation operator $\mathbb{E}$ and full generator $\mathbb{A}_{\mathsf{full}}$. As usual, we denote the full generator of the shift group by $\mathbb{D}_{\mathsf{full}}$. Then:*

(a) *It is $(F,G) \in \mathbb{A}_{\mathsf{full}}$ if and only if there exits $(U,V) \in \mathbb{D}_{\mathsf{full}}$ with $F = \mathbb{E}U$ and $G = \mathbb{E}V$.*
(b) *Define $\mathbb{A}_0$ by*
$$D(\mathbb{A}_0) := \Big\{ \sum_{j=1}^n U_j \mathbb{E} V_j : n \in \mathbb{N}, U_j \in D_-(\mathbb{D}_0), V_j \in D_+(\mathbb{D}_0) \text{ for } j = 1, \ldots, n \Big\}$$
and
$$\mathbb{A}_0 \sum_{j=1}^n U_j \mathbb{E} V_j = \sum_{j=1}^n \big[(\mathbb{D}_0 U_j) \mathbb{E} V_j + U_j \mathbb{E} \mathbb{D}_0 V_j \big].$$
Then $\mathbb{A}_0$ is a slice of $\mathbb{A}_{\mathsf{full}}$ and $D(\mathbb{A}_0)$ is a bp-core for $\mathbb{A}_{\mathsf{full}}$.
(c) *$D(\mathbb{A}_0)$ (and thus $D(\mathbb{A}_{\mathsf{full}})$) is bp-dense in $B_b(\mathscr{X}, \mathscr{F}_{0-})$. In view of Lemma A.5, $\mathbb{T}$ is uniquely determined by $\mathbb{A}_{\mathsf{full}}$.*

Proof. (a) As $\mathbb{E}$ is homogeneous, it is $\mathbb{E}\Theta_t = \mathbb{E}\Theta_t\mathbb{E}$ for all $t \geq 0$. By Proposition A.4, we have $(U,V) \in \mathbb{D}_{\mathsf{full}}$ if and only if

$$\Theta_t U - U = \int_0^t \Theta_s V \, ds$$

for all $t \geq 0$. Applying $\mathbb{E}$ to this equality and using the homogeneity yields

$$\mathbb{E}\Theta_t \mathbb{E} U - \mathbb{E} U = \mathbb{T}(t)\mathbb{E}U - \mathbb{E}U = \int_0^t \mathbb{E}\Theta_s \mathbb{E} V \, ds = \int_0^t \mathbb{T}(t) \mathbb{E} V \, ds.$$

By Proposition A.4, this shows that $(F,G) = (\mathbb{E}U, \mathbb{E}V) \in \mathbb{A}_{\mathsf{full}}$.

To see the converse, first observe that the resolvents of $\mathbb{A}_{\mathsf{full}}$ and $\mathbb{D}_{\mathsf{full}}$ are related via

$$R(\lambda, \mathbb{A}_{\mathsf{full}})F = \int_0^\infty e^{-\lambda t}\mathbb{E}\Theta_t F\, dt = \mathbb{E}\int_0^\infty e^{-\lambda t}\Theta_t F\, dt = \mathbb{E}R(\lambda, \mathbb{D}_{\mathsf{full}})F.$$

Consequently, $(F, G) \in \mathbb{A}_{\mathsf{full}}$ if and only if

$$F = R(\lambda, \mathbb{A}_{\mathsf{full}})(\lambda F - G) = \mathbb{E}R(\lambda, \mathbb{D}_{\mathsf{full}})(\lambda F - G) = \mathbb{E}U,$$

for $U = R(\lambda, \mathbb{D}_{\mathsf{full}})(\lambda F - G) \in D(\mathbb{D}_{\mathsf{full}})$. Note that in this case for $V = \lambda U - \lambda F + G$ we have $(U, V) \in \mathbb{D}_{\mathsf{full}}$. Moreover, $\mathbb{E}V = G$ as $\mathbb{E}F = \mathbb{E}^2 U = \mathbb{E}U$ by construction and $\mathbb{E}G = G$ since G is $\mathscr{F}_{0-}$-measurable.

(b) Noting that if $a < 0 < c$, then $F_a^b(f) = F_a^0(f) + F_0^b(f)$, where $F_a^0(f)$ is $\mathscr{F}_0$-measurable and $F_0^b(f)$ is $\mathscr{T}_0$-measurable, it is easy to see that any element U of $D(\mathbb{D}_0)$ can be written in the form $U = \sum_{j=1}^n U_j V_j$, where $U_j \in D_-(\mathbb{D}_0)$ and $V_j \in D_+(\mathbb{D}_0)$ for $j = 1, \dots, n$. Using part (a) and Proposition 4.1 (taking Remark 3.7 into account), it follows that $\mathbb{A}_0$ is a slice of $\mathbb{A}_{\mathsf{full}}$.

Using that $D(\mathbb{D}_0)$ is a bp-core for $\mathbb{D}_{\mathsf{full}}$ by Corollary 3.9, it follows from part (a) and the σ-continuity of $\mathbb{E}$ that $\mathbb{A}_0$ is a bp-core for $\mathbb{A}$.

(c) It follows from (b) that $D_-(\mathbb{D}_0) \subset D(\mathbb{A}_0)$. The proof now follows along the lines of that of Lemma 3.8. ☐

5. Continuity properties of evolutionary semigroups

In this section, we study the question, whether an evolutionary semigroup actually is a C_b-semigroup. The answer depends on the path space $\mathscr{X}$. The case where $\mathscr{X} = \mathscr{X}_{\mathrm{C}}$ is significantly easier than the general case and also yields better results. This is due to the fact that both the evaluation maps π_t and the stopping map τ are continuous in this case. The former yields that the operator $\mathbb{D}_0$ from Definition 3.6 is not only a slice of $\mathbb{D}_{\mathsf{full}}$, but actually a restriction of the C_b-generator $\mathbb{D}$, since in this case $\mathbb{D}_0 F_a^b(f) = F_b(f) - F_a(f) \in C_b(\mathscr{X})$ whenever $f \in C_b(X)$. That τ is continuous implies that the extension map $F \mapsto \hat{F}$ does not only establish an isomorphism between $B_b(\mathscr{X}^-)$ and $B_b(\mathscr{X}, \mathscr{F}_{0-})$, but also an isomorphism between $C_b(\mathscr{X}^-)$ and $C_b(\mathscr{X}, \mathscr{F}_{0-})$.

In the case $\mathscr{X} = \mathscr{X}_{\mathrm{D}}$ of càdlàg paths, functions of the form $F = F_t(f)$ for $f \in C_b(X)$ are not continuous on $\mathscr{X}$, whence $\mathbb{D}_0$ is not a restriction of the C_b-generator in this case. Moreover, for $t = 0$ the function $F = F_0(f)$ belongs to $C_b(\mathscr{X}^-)$ but its extension $\hat{F}$ is not continuous. To overcome this problem, we will impose additional assumptions on the path space (which are satisfied for càdlàg paths in the J_1-topology) which allow us to generalize at least part of our results.

5.1. **The case of continuous paths.** Throughout this subsection, we consider the case $((\mathscr{X}, \mathbb{d}), \tau) = ((\mathscr{X}_{\mathrm{C}}, \mathbb{d}_{\mathrm{C}}), \tau_{\mathrm{C}})$ from Subsection B.1. However, in order to not overburden notation, we will use our generic notation. Note that in this case $\mathscr{F}_{t-} = \sigma(\pi_s : s < t) = \sigma(\pi_s : s \le t) = \mathscr{F}_t$.

Before stating and proving our first result, we need a little preparation. Let us briefly recall the notion of a *convergence determining set*, see, e.g. [20, p. 112]. If (S, d) is a metric space, and μ_n, μ are Borel measures on S, then μ_n converges weakly to μ if $\int_S f\, d\mu_n \to \int_S f\, d\mu$ for every $f \in C_b(S)$; we write $\mu_n \rightharpoonup \mu$ in this case. A subset $M \subset C_b(S)$ is called *convergence determining* if convergence $\int_S f\, d\mu_n \to \int_S f\, d\mu$ for all $f \in M$ implies $\mu_n \rightharpoonup \mu$.

A set $M \subset C_b(S)$ is said to *strongly separate points* in S if, given $x \in S$ and $\delta > 0$, there exist $\{f_1, \dots, f_n\} \subset M$ such that

$$\inf_{y : d(x,y) \ge \delta} \max_{k=1,\dots,n} |f_k(x) - f_k(y)| > 0.$$

It follows from [20, Theorem 3.4.5] that any algebra that strongly separates points in S is convergence determining.

Lemma 5.1. *The sets*

$$\mathscr{D} := \bigcup_{t>0} C_b(\mathscr{X}, \mathscr{F}_t)$$

and

$$\mathscr{P} := \Big\{ \sum_{j=1}^n F_j G_j : n \in \mathbb{N}, F_j \in C_b(\mathscr{X}, \mathscr{F}_0), G_j \in C_b(\mathscr{X}, \mathscr{F}([0,\infty))) \Big\}.$$

are convergence determining for $\mathscr{X}$.

Proof. Fix $\mathbb{x} \in \mathscr{X}$ and $\delta > 0$. Pick n_0 so large, that $2^{-n_0} < \delta/2$ and put

$$F(\mathbb{y}) := \sup_{t\in[-n_0,n_0]} d(\mathbb{x}(t), \mathbb{y}(t)).$$

Obviously, $F \in C_b(\mathscr{X}, \mathscr{F}_{n_0}) \subset \mathscr{D}$ and $F(\mathbb{x}) = 0$. On the other hand, if $\mathbb{d}(\mathbb{x}, \mathbb{y}) > \delta$, we must have $F(\mathbb{y}) \geq \delta/2$ as otherwise $\mathbb{d}(\mathbb{x}, \mathbb{y}) < \delta$ by choice of n_0. This shows that $\mathscr{D}$ strongly separates points. As $\mathscr{D}$ is also an algebra, [20, Theorem 3.4.5] yields the claim. The proof for $\mathscr{P}$ is similar, considering the functions

$$F(\mathbb{y}) := \sup_{t\in[-n_0,0]} d(\mathbb{x}(t), \mathbb{y}(t)) \quad \text{and} \quad G(\mathbb{y}) := \sup_{t\in[0,n_0]} d(\mathbb{x}(t), \mathbb{y}(t)).$$ ∎

We are now ready to present the main result of this subsection. In its formulation (and also in the rest of this subsection), we once again do not distinguish between an evolutionary semigroup $\mathbb{T}$ and its modification $\mathbb{T}^-$ on $B_b(\mathscr{X}^-)$ defined by (4.4). This in particular concerns parts (a) and (b) of the theorem.

Theorem 5.2. *Let $\mathbb{T}$ be an evolutionary semigroup with expectation operator $\mathbb{E}$ and C_b-generator $\mathbb{A}$. The following are equivalent:*

(i) $\mathbb{T}(t)C_b(\mathscr{X}, \mathscr{F}_0) \subset C_b(\mathscr{X}, \mathscr{F}_0)$ *for all* $t > 0$*;*
(ii) $\mathbb{E}C_b(\mathscr{X}) \subset C_b(\mathscr{X}, \mathscr{F}_0)$*;*
(iii) $\mathbb{E}C_b(\mathscr{X}, \mathscr{F}([0,\infty))) \subset C_b(\mathscr{X}, \mathscr{F}_0)$.

If these equivalent conditions are satisfied, the following hold:

(a) $\mathbb{T}$ *induces (in the sense of (4.4)) a C_b-semigroup on $C_b(\mathscr{X}^-)$;*
(b) *The operator $\mathbb{A}_0$ from Proposition 4.10 is a restriction of $\mathbb{A}$. Moreover, $D(\mathbb{A}_0)$ is a β_0-core for $\mathbb{A}$.*

Proof. Throughout the proof, we denote the kernel of $\mathbb{E}$ by $\mathbb{k}$ and write $\mathbb{P}^{\mathbb{x}}$ for the probability measure $\mathbb{k}(\mathbb{x}, \cdot)$.

(i) ⇒ (ii). As $\mathbb{E}F = \mathbb{T}(t)\Theta_{-t}F$ whenever F is determined before t, it follows from (i) that $\mathbb{E}F \in C_b(\mathscr{X}, \mathscr{F}_0)$ whenever F belongs to the set $\mathscr{D}$ from Lemma 5.1. This means that if $(\mathbb{x}_n) \subset \mathscr{X}^-$ converges to $\mathbb{x} \in \mathscr{X}^-$, then

$$\int_{\mathscr{X}} F(\mathbb{y})\mathbb{P}^{\mathbb{x}_n}(d\mathbb{y}) = [\mathbb{E}F](\mathbb{x}_n) \to [\mathbb{E}F](\mathbb{x}) = \int_{\mathscr{X}} F(y)\mathbb{P}^{\mathbb{x}}(d\mathbb{y}) \tag{5.1}$$

for every $F \in \mathscr{D}$. As $\mathscr{D}$ is convergence determining by Lemma 5.1, Equation (5.1) holds true for every $F \in C_b(\mathscr{X})$, i.e. $\mathbb{E}F$ is continuous on $\mathscr{X}^-$ for every $F \in C_b(\mathscr{X})$. As τ is a continuous map taking values in $\mathscr{X}^-$ and $[\mathbb{E}F](\mathbb{x}) = [\mathbb{E}F](\tau(\mathbb{x}))$ for all $\mathbb{x} \in \mathscr{X}$, (ii) follows.

(ii) ⇒ (iii). This implication is trivial.

(iii) ⇒ (i). If $F \in C_b(\mathscr{X}, \mathscr{F}_0)$ and $G \in C_b(\mathscr{X}, \mathscr{F}([0,\infty))$, then it follows from Proposition 4.1, that $\mathbb{E}(FG) = F\mathbb{E}G$ and (iii) shows that $\mathbb{E}(FG) \in C_b(\mathscr{X}, \mathscr{F}_0)$. It follows that $\mathbb{E}H \in C_b(\mathscr{X}, \mathscr{F}_0)$ for every Element H in the set $\mathscr{P}$ from Lemma 5.1. As this set is convergence determining, similar arguments as in the proof of (i) ⇒

(ii) show that $\mathbb{E}F \in C_b(\mathscr{X}, \mathscr{F}_0)$ for every $F \in C_b(\mathscr{X})$, i.e. (ii) holds true. That (ii) implies (i) follows immediately from the representation $\mathbb{T}(t)F = \mathbb{E}\Theta_t F$.

Now assume that the equivalent conditions are satisfied.

(a). Let $F \in C_b(\mathscr{X}, \mathscr{F}_0)$, $(t_n) \subset [0,\infty)$ converge to t and $(\mathbb{x}_n) \subset \mathscr{X}^-$ converge to $\mathbb{x}$. By Proposition 3.2 and Theorem A.10, $\Theta_{-t_n}F \to \Theta_{-t}F$ with respect to β_0. By (ii), $\mathbb{P}^{\mathbb{x}_n} \to \mathbb{P}^{\mathbb{x}}$ with respect to $\sigma(\mathscr{M}(\mathscr{X}), C_b(\mathscr{X}))$. It follows that

$$[\mathbb{T}(t_n)F](\mathbb{x}_n) = \int_{\mathscr{X}} [\Theta_{-t_n}F](\mathbb{y})\mathbb{P}^{\mathbb{x}_n}(d\mathbb{y}) \to \int_{\mathscr{X}} [\Theta_{-t}F](\mathbb{y})\mathbb{P}^{\mathbb{x}}(d\mathbb{y}) = [\mathbb{T}(t)F](\mathbb{x}).$$

(b). Note that, in the case of continuous paths, we actually have $\mathbb{D}_0 \subset \mathbb{D}$. By the construction in Proposition 4.10 this fact together with (ii) shows that $D(\mathbb{A}_0) \subset C_b(\mathscr{X}, \mathscr{F}_0)$ and $\mathbb{A}_0 F \in C_b(\mathscr{X}, \mathscr{F}_0)$ for every $F \in D(\mathbb{A}_0)$. Consequently, $\mathbb{A}_0 \subset \mathbb{A}$.

To prove that $D(\mathbb{A}_0)$ is a β_0-core for $\mathbb{A}$, let $F \in D(\mathbb{A})$ be given. By Proposition 4.10(a), there is some $(U,V) \in D(\mathbb{D}_{\mathsf{full}})$ with $F = \mathbb{E}U$ and $\mathbb{A}F = \mathbb{E}V$. Inspecting the proof of that proposition and noting that $R(\lambda, \mathbb{D}_{\mathsf{full}})$ maps $C_b(\mathscr{X})$ to itself, we see that we may choose continuous U, V. This implies $U \in D(\mathbb{D})$ and $V = \mathbb{D}U$. By the above, there is a net $(U_\alpha) \subset D(\mathbb{D}_0)$ with $U_\alpha \to U$ and $\mathbb{D}U_\alpha \to \mathbb{D}U = V$ with respect to β_0. It follows from (ii) that $\mathbb{E}$ is β_0-continuous and thus $F_\alpha := \mathbb{E}U_\alpha \to F$ and $\mathbb{A}F_\alpha = \mathbb{E}\mathbb{D}U_\alpha \to \mathbb{E}V = \mathbb{A}F$ with respect to β_0. ☐

Definition 5.3. An *evolutionary C_b-semigroup* is an evolutionary semigroup that satisfies the equivalent conditions of Theorem 5.2.

Proposition 5.4. *Let $\mathscr{X} = \mathscr{X}_{\mathrm{C}}$ and $(\mathbb{T}(t))_{t\geq 0}$ be a C_b-semigroup on $C_b(\mathscr{X}^-)$ with C_b-generator $\mathbb{A}$. Then $\mathbb{T}$ is induced by an evolutionary semigroup if and only if $D_-(\mathbb{D}_0) \subset D(\mathbb{A})$ and $\mathbb{A}U = \mathbb{D}U$ for $U \in D_-(\mathbb{D}_0)$.*

Proof. Assume that $D_-(\mathbb{D}_0) \subset D(\mathbb{A})$ and $\mathbb{A}U = \mathbb{D}U$ for $U \in D_-(\mathbb{D}_0)$. Fix $U \in D_-(\mathbb{D}_0)$. As $D(\mathbb{D}_0)$ is invariant under the shift semigroup, $\Theta_{-t}U \in D_-(\mathbb{D}_0) \subset D(\mathbb{A})$ for all $t > 0$. Define $\varphi : [0,\infty) \to C_b(\mathscr{X}^-)$ by setting $\varphi(t) = \mathbb{T}(t)\Theta_{-t}U$.

Now, fix $t \geq 0$ and note that

$$\frac{\varphi(t+h) - \varphi(t)}{h} = \mathbb{T}(t+h)\frac{\Theta_{-(t+h)}U - \Theta_{-t}U}{h} + \frac{\mathbb{T}(t+h) - \mathbb{T}(t)}{h}\Theta_{-t}U.$$

Since $\Theta_{-t}U \in D(\mathbb{A})$ the last term β_0-converges to $\mathbb{T}(t)\mathbb{A}\Theta_{-t}U = \mathbb{T}(t)\mathbb{D}\Theta_{-t}U$ as $h \to 0$.

Next, observe that

$$\Delta_h U := \frac{\Theta_{-(t+h)}U - \Theta_{-t}U}{h} \to -\mathbb{D}\Theta_{-t}U$$

with respect to β_0. Now, let $\mu \in \mathscr{M}(\mathscr{X}^-)$ be given. As the map $r \mapsto \mathbb{T}(r)'\mu$ is $\sigma(\mathscr{M}(\mathscr{X}^-), C_b(\mathscr{X}^-))$-continuous, the set

$$S := \{\mathbb{T}(r)'\mu : r \in [t-1, t+1] \cap [0,\infty)\}$$

is $\sigma(\mathscr{M}(\mathscr{X}^-), C_b(\mathscr{X}^-))$-compact. By Theorems 5.8 and 5.4 of [50], β_0 is not only the Mackey topology of the dual pair $(C_b(\mathscr{X}^-), \mathscr{M}(\mathscr{X}^-))$ but also the topology of uniform convergence on the $\sigma(\mathscr{M}(\mathscr{X}^-), C_b(\mathscr{X}^-))$-compact subsets of $\mathscr{M}(\mathscr{X}^-)$. Thus, $p_S(F) := \sup_{\nu\in S} |\langle F, \nu\rangle|$ is a β_0-continuous seminorm. It follows that for $|h| < \max\{t, 1\}$

$$\begin{aligned}|\langle \mathbb{T}(t+h)(\Delta_h U + \mathbb{D}\Theta_{-t}U), \mu\rangle| &= |\langle \Delta_h U + \mathbb{D}\Theta_{-t}U, \mathbb{T}(t+h)'\mu\rangle| \\ &\leq p_S(\Delta_h U + \mathbb{D}\Theta_{-t}U) \to 0\end{aligned}$$

as $h \to 0$. Noting that $\mathbb{T}(t+h)\mathbb{D}\Theta_{-t}U \to \mathbb{T}(t)\mathbb{D}\Theta_{-t}U$ as $h \to 0$ by the continuity of $\mathbb{T}$, it follows that altogether

$$\left\langle \frac{\varphi(t+h) - \varphi(t)}{h}, \mu \right\rangle \to 0 - \langle \mathbb{T}(t)\mathbb{D}\Theta_{-t}U, \mu\rangle + \langle \mathbb{T}(t)\mathbb{D}\Theta_{-t}U, \mu\rangle = 0.$$

This implies that the map $t \mapsto \langle \varphi(t), \mu \rangle$ is constant, whence $\langle \mathbb{T}(t)\Theta_{-t}U, \mu\rangle = \langle U, \mu\rangle$ for all $t \geq 0$ and $\mu \in \mathscr{M}(\mathscr{X}^-)$. As μ was arbitrary, the Hahn–Banach theorem implies $\mathbb{T}(t)\Theta_{-t}U = U$ for all $t \geq 0$.

We note that $D_-(\mathbb{D}_0)$ is an algebra that separates the points in $\mathscr{X}^-$ and is thus dense in $C_b(\mathscr{X}^-)$ by the Stone–Weierstraß Theorem, see [26, Theorem 11]. It follows that, given $F \in C_b(\mathscr{X}^-)$, we find a net $(U_\alpha)_\alpha \subset D_-(\mathbb{D}_0)$ that converges to F with respect to β_0. By the continuity properties of $\mathbb{T}(t)$ and Θ_{-t}, it follows that the equality $\mathbb{T}(t)\Theta_{-t}U_\alpha = U_\alpha$ implies $\mathbb{T}(t)\Theta_{-t}F = F$. By a bp-closure argument, this equality extends to arbitrary $F \in B_b(\mathscr{X}, \mathscr{F}_0)$, proving that $\mathbb{T}$ is induced by an evolutionary semigroup.

The converse implication follows immediately from Theorem 5.2. □

5.2. **The general case.** We now turn to more general path spaces, where our main interest lies in the space of all càdlàg paths. If $\mathscr{X} = \mathscr{X}_\mathrm{D}$ then for $f \in C_b(X)$ the function $F_t(f)$ is not a continuous function on $\mathscr{X}_\mathrm{D}$. More precisely, $F_t(f)$ is continuous at the point $\mathbb{x} \in \mathscr{X}_\mathrm{D}$ if and only if $\mathbb{x}$ is continuous at t. On the other hand, $F_0(f)$ is continuous on $\mathscr{X}_\mathrm{D}^-$, as every $\mathbb{x} \in \mathscr{X}_\mathrm{D}^-$ is continuous at 0. This is an example of a function $F \in C_b(\mathscr{X}_\mathrm{D}^-)$ whose extension $\hat{F}$ is not continuous on $\mathscr{X}$.

To study the subtle interaction of the extension map with continuous functions, we introduce the following spaces:

$$\hat{C}_b(\mathscr{X}^-) := \{\hat{F} : F \in C_b(\mathscr{X}^-)\} = \{F \in B_b(\mathscr{X}, \mathscr{F}_{0-}) : F|_{\mathscr{X}^-} \text{ is continuous }\}$$

is the space of all extensions to $\mathscr{X}$ of continuous functions on $\mathscr{X}^-$.

$$C_{b,\mathrm{ext}}(\mathscr{X}^-) := \{F \in C_b(\mathscr{X}^-) : \hat{F} \in C_b(\mathscr{X}, \mathscr{F}_{0-})\} = \{F|_{\mathscr{X}^-} : F \in C_b(\mathscr{X}, \mathscr{F}_{0-})\}$$

is the space of all continuous functions on $\mathscr{X}^-$ whose extensions are continuous on all of $\mathscr{X}$.

To ensure that these spaces are rich enough, we impose some additional assumptions on path spaces.

Definition 5.5. Let $((\mathscr{X}, \mathbb{d}), \tau)$ be a path space. We say that $((\mathscr{X}, \mathbb{d}), \tau)$ is *proper* if

(P4) If $\mathbb{x}_n \to \mathbb{x}$, then $\tau(\vartheta_t \mathbb{x}_n) \to \tau(\vartheta_t \mathbb{x})$ in $\mathscr{X}$ for almost every $t \in \mathbb{R}$.

(P5) Given a compact subset $K \subset \mathscr{X}^-$ and $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\mathbb{d}(\mathbb{x}, \tau(\vartheta_{-t}\mathbb{x})) \leq \varepsilon$$

for all $0 \leq t \leq \delta$ and $\mathbb{x} \in K$.

(P6) There is a constant $c \in (0,1)$ such that given $\delta > 0$ we find $t_0 > 0$ such that $\mathbb{d}(\tau_{t_0}(\mathbb{x}), \tau_{t_0}(\mathbb{y})) \geq c\delta$ for all $\mathbb{x}, \mathbb{y} \in \mathscr{X}$ with $\mathbb{d}(\mathbb{x}, \mathbb{y}) \geq \delta$.

In Section B.3, we will prove that $\mathscr{X}_\mathrm{D}$ is a proper path space, so that all results of this subsection are applicable in this case.

Proposition 5.6. *Let $((\mathscr{X}, \mathbb{d}), \tau)$ be a proper path space.*

(a) *Given $F \in \hat{C}_b(\mathscr{X}^-)$ and $r > 0$ we define*

$$F_r(\mathbb{x}) := \frac{1}{r}\int_0^r F(\vartheta_{-t}\mathbb{x})\, dt.$$

Then $F_r \in C_b(\mathscr{X}, \mathscr{F}_{0-})$ and $F_r(\mathbb{x}) \to F(\mathbb{x})$ as $r \to 0$ for every $\mathbb{x} \in \mathscr{X}^-$, uniformly on compact subsets of $\mathscr{X}^-$.

(b) *Given $\mathbb{x} \in \mathscr{X}^-$ and $\delta > 0$ there is $F \in C_{b,\mathrm{ext}}(\mathscr{X}^-)$ with*

$$\inf\{|F(\mathbb{x}) - F(\mathbb{y})| : \mathbb{y} \in \mathscr{X}^-, \mathbb{d}(\mathbb{x}, \mathbb{y}) \geq \delta\} > 0$$

In other words, $C_{b,\mathrm{ext}}(\mathscr{X}^-)$ strongly separates the points in $\mathscr{X}^-$.

(c) *$C_{b,\mathrm{ext}}(\mathscr{X}^-)$ is convergence determining for $\mathscr{X}^-$.*

Proof. (a). First observe that as F is determined before 0, $F \circ \vartheta_{-t}$ is determined before $-t$ and thus also before 0. Integrating, it follows that F_r is $\mathscr{F}_{0-}$-measurable. Now let $\mathbb{x}_n \to \mathbb{x}$. As F is $\mathscr{F}_{0-}$-measurable, $F(\vartheta_{-t}\mathbb{x}_n) = F(\tau(\vartheta_{-t}\mathbb{x}_n))$ for all $t > 0$ and $n \in \mathbb{N}$ by Lemma 2.5(c). It follows from (P4) that $\tau(\vartheta_{-t}\mathbb{x}_n) \to \tau(\vartheta_{-t}\mathbb{x})$ for almost all $t \in (0, r)$. By continuity of F on $\mathscr{X}^-$, $F(\tau(\vartheta_{-t}\mathbb{x}_n)) \to F(\tau(\vartheta_{-t}\mathbb{x}))$ for almost all $t \in (0, r)$ and the dominated convergence theorem yields $F_r(\mathbb{x}_n) \to F_r(\mathbb{x})$, proving that $F_r \in C_b(\mathscr{X}, \mathscr{F}_{0-})$.

Now let $K \subset \mathscr{X}^-$ be compact and $\varepsilon > 0$. As F is uniformly continuous on K we find $\delta > 0$ such that $|F(\mathbb{x}) - F(\mathbb{y})| \leq \varepsilon$ for all $\mathbb{x}, \mathbb{y} \in K$ with $\mathbb{d}(\mathbb{x}, \mathbb{y}) \leq \delta$. By (P5), there is $\rho > 0$ such that $\mathbb{d}(\mathbb{x}, \tau(\vartheta_{-t}\mathbb{x})) \leq \delta$ for all $\mathbb{x} \in K$ and $0 \leq t \leq \rho$. It follows for $0 < r < \rho$ and $\mathbb{x} \in K$ that

$$|F(\mathbb{x}) - F_r(\mathbb{x})| \leq \frac{1}{r}\int_0^r |F(\mathbb{x}) - F(\tau(\vartheta_{-t}\mathbb{x}))|\, dr \leq \varepsilon.$$

(b). Given $\mathbb{x} \in \mathscr{X}^-$ and $\delta > 0$, pick r_0 such that $\mathbb{d}(\mathbb{x}, \tau(\vartheta_{-t}\mathbb{x})) \leq \delta/4$ for all $0 \leq t \leq r_0$. This is possible by (P5). As $\mathscr{X}^-$ is a metric space, we find $F \in C_b(\mathscr{X}^-)$ with $0 \leq F \leq 1$ such that $F(\mathbb{y}) = 0$ for all $\mathbb{y} \in \mathscr{X}^- \cap B(\mathbb{x}, \delta/4)$ and $F(\mathbb{y}) = 1$ for all $\mathbb{y} \in \mathscr{X}^- \setminus B(\mathbb{x}, \delta/2)$.

Consider the function F_{r_0} defined as in part (a). Then $F_{r_0} \in C_b(\mathscr{X}, \mathscr{F}_{0-})$ and $F(\mathbb{x}) = 0$. If $\mathbb{y} \in \mathscr{X}^-$ satisfies $\mathbb{d}(\mathbb{x}, \mathbb{y}) \geq \delta$, we find by (P5) $0 < \varepsilon < r_0$ such that $\mathbb{d}(\mathbb{y}, \tau(-\vartheta_{-t}\mathbb{y})) \leq \delta/4$ for all $0 \leq t \leq \varepsilon$. It follows that $\mathbb{d}(\mathbb{x}, \tau(\vartheta_{-t}\mathbb{y})) \geq \delta/2$ for all $0 \leq t \leq \varepsilon$ and thus

$$|F(\mathbb{x}) - F(\mathbb{y})| = F(\mathbb{y}) = \frac{1}{r_0}\int_0^{r_0} F(\tau(\vartheta_{-t}\mathbb{y})\, dt \geq \frac{1}{r_0}\int_0^{\varepsilon} 1\, dt = \frac{\varepsilon}{r_0} > 0.$$

(c). This follows from part (b) and [20, Theorem 3.4.5]. ☐

Corollary 5.7. *Let $((\mathscr{X}, \mathbb{d}), \tau)$ be a proper path space. Define*

$$\mathscr{D} := \bigcup_{t\in\mathbb{R}} C_b(\mathscr{X}, \mathscr{F}_{t-})$$

and

$$\mathscr{P} := \Big\{ \sum_{j=1}^n F_j G_j : n \in \mathbb{N}, F_j \in C_b(\mathscr{X}, \mathscr{F}_{0-}), G_j \in C_b(\mathscr{X}, \mathscr{F}([0,\infty))) \Big\}.$$

Then both $\mathscr{D}$ and $\mathscr{P}$ are convergence determining for $\mathscr{X}$ and β_0-dense in $C_b(\mathscr{X})$.

Proof. We note that $\mathscr{X}^-$ is homeomorphic to $\mathscr{X}^t := \tau_t(\mathscr{X})$ via the map ϑ_{-t}. Thus, it follows from from Proposition 5.6(b) that $C_b(\mathscr{X}, \mathscr{F}_{t-})$ strongly separates points in $\mathscr{X}^t$. Using condition (P6) it follows that $\mathscr{D}$ strongly separates the points in $\mathscr{X}$. As $\mathscr{D}$ is an algebra, [20, Theorem 3.4.5] yields that $\mathscr{D}$ is convergence determining and the Stone–Weierstraß Theorem [26, Theorem 11] yields the density.

The proof for $\mathscr{P}$ is similar. We note that to prove that $\mathscr{P}$ strongly separates points, we can consider a product FG, where $F \in C_b(\mathscr{X}, \mathscr{F}_{0-})$, i.e. F is an extension of an element of $C_{b,\mathrm{ext}}(\mathscr{X}^-)$, and G is of the form $G(\mathbb{y}) = \mathbb{d}_0^t(\mathbb{y}, \mathbb{x})$ (cf. Equation (B.1)) for suitably chosen t and $\mathbb{x}$. ☐

We can now generalize Theorem 5.2:

Theorem 5.8. *Let $((\mathscr{X}, \mathbb{d}), \tau)$ be a proper path space and $(\mathbb{T}(t))_{t\geq 0}$ be an evolutionary semigroup with expectation operator $\mathbb{E}$. The following are equivalent:*

(i) $\mathbb{T}(t)C_b(\mathscr{X}, \mathscr{F}_{0-}) \subset \hat{C}_b(\mathscr{X}^-)$ *for all* $t \geq 0$*;*
(ii) $\mathbb{T}(t)\hat{C}_b(\mathscr{X}^-) \subset \hat{C}_b(\mathscr{X}^-)$ *for all* $t \geq 0$*;*
(iii) $\mathbb{E}C_b(\mathscr{X}) \subset \hat{C}_b(\mathscr{X}^-)$*;*
(iv) $\mathbb{E}C_b(\mathscr{X}, \mathscr{F}([0,\infty))) \subset \hat{C}_b(\mathscr{X}^-)$.

If these equivalent conditions are satisfied, $\mathbb{T}$ induces (in the sense of (4.4)) a C_b-semigroup on $C_b(\mathscr{X}^-)$.

Proof. Assume condition (i) and fix a sequence $(\mathbb{x}_n) \subset \mathscr{X}^-$ converging to $\mathbb{x} \in \mathscr{X}^-$. We denote by $\mathbb{p}_t$ the kernel of $\mathbb{T}(t)$ and define the measures μ_n on $\mathscr{X}^-$ by setting $\mu_n(A) := \mathbb{p}_t(\mathbb{x}_n, \tau^{-1}(A))$ and μ similarly, replacing $\mathbb{x}_n$ by $\mathbb{x}$. Condition (i) implies that

$$\int_{\mathscr{X}^-} F(\mathbb{y})\, d\mu_n(\mathbb{y}) = [\mathbb{T}(t)\hat{F}](\mathbb{x}_n) \to [\mathbb{T}(t)\hat{F}](\mathbb{x}) = \int_{\mathscr{X}^-} F(\mathbb{y})\, d\mu(\mathbb{y}) \tag{5.2}$$

for every $F \in C_{b,\mathrm{ext}}(\mathscr{X}^-)$. As this set is convergence determining by Proposition 5.6(c) , the convergence in (5.2) still holds true for all $F \in C_b(\mathscr{X}^-)$ and this means precisely that $\mathbb{T}(t)\hat{C}_b(\mathscr{X}^-) \subset \hat{C}_b(\mathscr{X}^-)$. This proves the implication (i) $\Rightarrow$ (ii), the converse implication (ii) $\Rightarrow$ (i) is trivial.

To prove (i) $\Rightarrow$ (iii), let $F \in C_b(\mathscr{X}, \mathscr{F}_{t-})$ for some $t > 0$. Then $\mathbb{E}F = \mathbb{T}(t)\Theta_{-t}F$ is continuous on $\mathscr{X}^-$ by (i). As $t > 0$ was arbitrary, it follows for all F in the set $\mathscr{D}$ from Corollary 5.7 that $\mathbb{E}F$ is continuous on $\mathscr{X}^-$. As this set is convergence determining by Corollary 5.7, (iii) follows by a similar argument as above.

The implication (iii) $\Rightarrow$ (i) follows immediately from the identity $\mathbb{T}(t)F = \mathbb{E}\Theta_t F$.

The remaining equivalenct (iii) $\Leftrightarrow$ (iv) follows as in the proof of Theorem 5.2, using the set $\mathscr{P}$ from Corollary 5.7.

Let us now assume that (i) – (iv) are satisfied. We fix a sequence $(\mathbb{x}_n) \subset \mathscr{X}^-$ converging to $\mathbb{x} \in \mathscr{X}^-$. Denoting the kernel of $\mathbb{E}$ by $\mathbb{k}$, it follows from (iii) that $\mathbb{k}(\mathbb{x}_n, \cdot) \rightharpoonup \mathbb{k}(\mathbb{x}, \cdot)$. Given a sequence $(t_n) \subset [0, \infty)$ converging to t and $F \in C_b(\mathscr{X}, \mathscr{F}_{0-})$, note that $\Theta_{-t_n}F \to \Theta_{-t}F$ with respect to β_0 as a consequence of Proposition 3.2. It follows that

$$[\mathbb{T}(t_n)F](\mathbb{x}_n) = \int_{\mathscr{X}} \Theta_{-t_n}F(\mathbb{y})\mathbb{k}(\mathbb{x}_n, d\mathbb{y}) \to \int_{\mathscr{X}} \Theta_{-t}F(\mathbb{y})\mathbb{k}(\mathbb{x}, d\mathbb{y}) = [\mathbb{T}(t_n)F](\mathbb{x}).$$

Using that $C_{b,\mathrm{ext}}(\mathscr{X}^-)$ is convergence determining for $\mathscr{X}^-$, the same convergence result holds true for $F \in \hat{C}_b(\mathscr{X}^-)$, proving the addendum. $\square$

Similarly to the case of continuous paths, we call an evolutionary semigroup $\mathbb{T}$ on a proper path space an *evolutionary C_b-semigroup* if it satisfies the equivalent conditions of Theorem 5.8.

A characterization of evolutionary semigroups through the generator in the spirit of Proposition 5.4 is also possible in the general case. However, as $D_-(\mathbb{D}_0)$ is not contained in the domain of the C_b-generator (as $\mathbb{D}F_a^b(f) \notin C_b(\mathscr{X})$), we have to use a suitable substitute. We put

$$D_-(\mathbb{D}) := D(\mathbb{D}) \cap C_b(\mathscr{X}, \mathscr{F}_{0-}).$$

Proposition 5.9. *Let $((\mathscr{X}, \mathbb{d}), \tau)$ be a proper path space and $\mathbb{T}$ be a C_b-semigroup on $C_b(\mathscr{X}^-)$ with C_b-generator $\mathbb{A}$. Then $\mathbb{T}$ is induced by an evolutionary semigroup if and only if $D_-(\mathbb{D}) \subset D(\mathbb{A})$ and $\mathbb{A}U = \mathbb{D}U$ for all $U \in D_-(\mathbb{D})$.*

Proof. The proof is identical to that of Proposition 5.4 with $D_-(\mathbb{D}_0)$ replaced by $D_-(\mathbb{D})$. We only need to know that $D_-(\mathbb{D})$ is β_0-dense in $C_b(\mathscr{X}^-)$. But this follows from Proposition 5.6(a), noting that the function F_r defined there belongs to $D(\mathbb{D})$ by Proposition A.4. $\square$

We end this section with some results in the case where $\mathscr{X} = \mathscr{X}_\mathrm{D}$ (or $\mathscr{X} = \mathscr{X}_\mathrm{C}$, where they hold trivially). We recall that $t^\star = t-$ in this case and we have $F_t^\star(f) = F_{t^\star}(f)$ whenever $f \in C_b(X)$ by Lemma 3.4.

Lemma 5.10. *Let $\mathscr{X} = \mathscr{X}_\mathrm{D}$ and $\mathbb{T}$ be an evolutionary C_b-semigroup with expectation operator $\mathbb{E}$. Then*

$$\mathbb{T}(t)F_{0^\star}(f) = \mathbb{E}F_t(f) = \mathbb{E}F_{t^\star}(f)$$

for all $f \in B_b(X)$ *and* $t \geq 0$*. In particular,* $\mathbb{E}F_0(f) = F_{0^\star}(f)$ *for all* $f \in B_b(X)$*.*

Proof. Fix $f \in C_b(X)$ and $t \geq 0$. From the relation of $\mathbb{T}$ and $\mathbb{E}$ it follows that for $s > 0$

$$\mathbb{T}(t+s)F_{0^\star}(f) = \mathbb{E}\Theta_{t+s}F_{0^\star}(f) = \mathbb{E}F_{(t+s)^\star}(f).$$

As $F_{0^\star}(f) \in C_b(\mathscr{X}^-)$, upon $s \downarrow 0$, the left-hand side converges pointwise to $\mathbb{T}(t)F_{0^\star}(f)$ by continuity of the semigroup. As for the right-hand side, note that $F_{(t+s)^\star}(f) \to F_t(f)$ as $s \downarrow 0$ by right-continuity of the paths. Using dominated convergence, $\mathbb{T}(t)F_{0^\star}(f) = \mathbb{E}F_t(f)$ follows. By a bp-closure argument, this equality extends to $f \in B_b(X)$. □

Corollary 5.11. *In the situation of Lemma 5.10, let* $\Bbbk$ *denote the kernel of the expectation operator* $\mathbb{E}$ *and write* $\mathbb{P}^{\mathbb{x}}$ *for the measure* $\Bbbk(\mathbb{x}, \cdot)$*. Given* $\mathbb{x} \in \mathscr{X}_{\mathrm{D}}$*, define*

$$C_{\mathbb{x}} = \{\mathbb{y} : \mathbb{y}(0) = \mathbb{y}(0^\star) = \mathbb{x}(0^\star)\} = \{\mathbb{y} : \mathbb{y} \textit{ is continuous at } 0 \textit{ and } \mathbb{y}(0) = \mathbb{x}(0^\star)\}.$$

Then $\mathbb{P}^{\mathbb{x}}(C_{\mathbb{x}}) = 1$ *for every* $\mathbb{x} \in \mathscr{X}$*.*

Proof. It is $A_- := \{\mathbb{y} : \mathbb{y}(0^\star) \neq \mathbb{x}(0^\star)\} \in \mathscr{F}_{0-}$. It follows from Proposition 4.1(iii) that $\Bbbk(\mathbb{x}, A_-) = \delta_{\mathbb{x}}(A_-) = 0$, proving that $\Bbbk(\mathbb{x}, \{\mathbb{y} : \mathbb{y}(0^\star) = \mathbb{x}(0^\star)\}) = 1$. Now let $n \in \mathbb{N}$ and define $f_n(x) := 1 \wedge nd(x, \mathbb{x}(0^\star))$. Then $f_n \in C_b(X)$ and Lemma 5.10 yields

$$\Bbbk(\mathbb{x}, \{\mathbb{y} : d(\mathbb{y}(0), \mathbb{x}(0^\star)) \geq n^{-1}\}) \leq [\mathbb{E}F_0(f_n)](\mathbb{x}) = [F_{0^\star}(f_n)](\mathbb{x}) = 0$$

for every $n \in \mathbb{N}$. This shows $\Bbbk(\mathbb{x}, \{\mathbb{y} : \mathbb{y}(0) \neq \mathbb{x}(0^\star)\}) = 0$ and yields the claim. □

6. Examples

6.1. **Deterministic evolutions.** In this subsection, we are interested in the situation where the expectation operator $\mathbb{E}$ is deterministic, i.e. for every $\mathbb{x} \in \mathscr{X}$ there is a unique $\varphi(\mathbb{x})$ such that $[\mathbb{E}F](\mathbb{x}) = F(\varphi(\mathbb{x}))$. We note that if $\mathbb{E}$ is of this form, then $\mathbb{E}$ is multiplicative, i.e. $\mathbb{E}(FG) = (\mathbb{E}F)(\mathbb{E}G)$. It is well-known, see [10, Theorem II.2.3], that any multiplicative $\mathbb{E}$ is of this form. Throughout, we will work on the path space $((\mathscr{X}, \mathbb{d}), \tau) = ((\mathscr{X}_{\mathrm{C}}, \mathbb{d}_{\mathrm{C}}), \tau_{\mathrm{C}})$ of continuous paths as these seem more appropriate for our purposes, see Remark 6.3 below.

Definition 6.1. An *evolution map* is a continuous mapping $\varphi : \mathscr{X} \to \mathscr{X}$ such that

(i) $\varphi(\mathbb{x}) = \varphi(\tau(\mathbb{x}))$ for all $\mathbb{x} \in \mathscr{X}$;
(ii) $\tau(\varphi(\mathbb{x})) = \tau(\mathbb{x})$;
(iii) $\varphi(\vartheta_t\varphi(\mathbb{x})) = \vartheta_t\varphi(\mathbb{x})$ for all $\mathbb{x} \in \mathscr{X}$ and $t \geq 0$.

Proposition 6.2. *Let* $\mathbb{T}$ *be an evolutionary* C_b*-semigroup with expectation operator* $\mathbb{E}$*. We denote the* C_b*-generator of* $\mathbb{T}$ *by* $\mathbb{A}$*. The following are equivalent:*

(i) $\mathbb{A}$ *is a derivation;*
(ii) $\mathbb{T}(t)$ *is multiplicative for every* $t \geq 0$*;*
(iii) $\mathbb{E}$ *is multiplicative;*
(iv) *There is an evolution map* φ *such that* $[\mathbb{E}F](\mathbb{x}) = F(\varphi(\mathbb{x}))$ *for all* $\mathbb{x} \in \mathscr{X}$*.*

Proof. (i) ⇔ (ii). This follows from [22, Theorem 3.4].

(ii) ⇔ (iii). If $\mathbb{T}$ is multiplicative and F, G are determined before $s > 0$ then

$$\begin{aligned}\mathbb{E}(FG) &= \mathbb{T}(s)\Theta_{-s}(FG) = \mathbb{T}(s)(\Theta_{-s}F)(\Theta_{-s}G)\\ &= [\mathbb{T}(s)\Theta_{-s}F][\mathbb{T}(s)\Theta_{-s}G] = [\mathbb{E}F][\mathbb{E}G].\end{aligned}$$

Thus, $\mathbb{E}$ is multiplicative for all functions from the set $\mathscr{D}$ of Corollary 5.7. As this set is β_0-dense in $C_b(\mathscr{X})$, the above equality extends to arbitrary $F, G \in C_b(\mathscr{X})$.

If we conversely assume that $\mathbb{E}$ is multiplicative, then for $F, G \in C_b(\mathscr{X}, \mathscr{F}_0)$, we have

$$\mathbb{T}(t)(FG) = \mathbb{E}\Theta_t(FG) = \mathbb{E}(\Theta_tF)(\Theta_tG) = [\mathbb{E}\Theta_tF][\mathbb{E}\Theta_tG] = [\mathbb{T}(t)F][\mathbb{T}(t)G].$$

(iii) ⇔ (iv). If $\mathbb{E}$ is multiplicative, it follows from [10, Theorem II.2.3] that there exists a continuous φ such that $[\mathbb{E}F](\varkappa) = F(\varphi(\varkappa))$. As $\mathbb{E}F = \mathbb{E}F\circ\tau$ by Proposition 4.1 and Lemma 2.5, we must have $\varphi = \varphi\circ\tau$ and thus (i) in Definition 6.1. Moroever, $\mathbb{E}F = F$ for all $F \in C_b(\mathscr{X}, \mathscr{F}_0)$. Thus, for $F \in C_b(\mathscr{X}, \mathscr{F}_0)$,

$$(\mathbb{E}F)(\varkappa) = F(\varphi(\varkappa)) = F(\tau(\varphi(\varkappa))) = F(\tau(\varkappa)) = F(\varkappa).$$

As F was arbitrary, $\tau\circ\varphi = \tau$ follows. Last, as $\mathbb{E}$ is homogeneous, Lemma 4.4 with $\Bbbk(\varkappa, \cdot) = \delta_{\varphi(\varkappa)}$ yields

$$\mathbb{1}_A(\varphi(\varkappa)) = \Bbbk(\varkappa, A) = \Bbbk(\vartheta_t\varphi(\varkappa), \vartheta_t(A)) = \mathbb{1}_{\vartheta_t A}(\varphi(\vartheta_t\varphi(\varkappa))) = \mathbb{1}_A(\vartheta_{-t}(\varphi(\vartheta_t\varphi(\varkappa)))).$$

This proves $\varphi(\varkappa) = \vartheta_{-t}(\varphi(\vartheta_t\varphi(\varkappa)))$ which is equivalent with $\varphi(\vartheta_t\varphi(\varkappa)) = \vartheta_t\varphi(\varkappa)$. Thus, φ is an evolution map. The converse is trivial. □

Remark 6.3. Let us consider for a moment the case of càdlàg paths, i.e. $\mathscr{X} = \mathscr{X}_{\mathrm{D}}$. In view of Corollary 5.11, for every evolution map φ that induces an evolutionary C_b-semigroup, $\varphi(\varkappa)$ is necessarily continuous at 0 for every $\varkappa \in \mathscr{X}$. But then condition (iii) of Definition 6.1 implies that $\varphi(\varkappa)$ is continuous at every $t \geq 0$.

With the help of Theorem 5.2(b), we can now determine a β_0-core for the C_b-generator of a semigroup $\mathbb{T}$ that satisfies the equivalent conditions of Proposition 6.2. In what follows, we write $\varphi(t, \varkappa)$ shorthand for $[\varphi(\varkappa)](t)$. We recall that $\mathbb{D}_0 F_a^b(f) = F_b(f) - F_a(f)$. Applying $\mathbb{E}$, we obtain for $0 \leq a < b$ and $f \in C_b(X)$,

$$[\mathbb{E}F_a^b(f)](\varkappa) = \int_a^b f(\varphi(s,\varkappa))\,ds \in D(\mathbb{A}) \quad \text{and}$$
$$[\mathbb{A}\mathbb{E}F_a^b(f)](\varkappa) = [\mathbb{E}(F_b(f) - F_a(f))](\varkappa) = f(\varphi(b,\varkappa)) - f(\varphi(a,\varkappa)).$$

Via approximation, we obtain more elements of $D(\mathbb{A})$. Compare the following Proposition with [22, Proposition 2.12] which concerns deterministic equations without delay.

Proposition 6.4. *Let $\mathbb{T}$ be an evolutionary semigroup satisfying the equivalent conditions of Proposition 6.2. Denote its C_b-generator by $\mathbb{A}$ and its evolution map by φ. For $f \in C_b(X)$ and $G \in C_b(\mathscr{X}, \mathscr{F}_0)$ the following are equivalent:*

(i) *$F_0(f) \in D(\mathbb{A})$ and $\mathbb{A}F_0(f) = G$;*
(ii) *For every $\varkappa \in \mathscr{X}$ we have*

$$\lim_{t\to 0} \frac{f(\varphi(t,\varkappa)) - f(\varkappa(0))}{t} = G(\varkappa).$$

Proof. (i) ⇒ (ii). Condition (i) is equivalent with

$$f(\varphi(t,\varkappa)) - f(\varkappa(0)) = [\mathbb{T}(t)F_0(f) - F_0(f)](\varkappa) = \int_0^t [\mathbb{T}(s)G](\varkappa)\,ds,$$

for all $\varkappa \in \mathscr{X}$ and $h > 0$. Dividing by t and taking the limit as $t \to 0$, using the continuity of the integrand on the right, (ii) follows.

(ii) ⇒ (i). We have $t^{-1}F_0^t(f) \in D(\mathbb{D})$ and $t^{-1}F_0^t(f) \to F_0(f)$ pointwise as $t \to 0$. Moreover, $\mathbb{D}t^{-1}F_0^t(f) = t^{-1}(F_t(f) - F_0(f))$. Applying the evolution operator $\mathbb{E}$, it follows from Proposition 5.4 that

$$U_t := \mathbb{E}t^{-1}F_0^t(f) \in D(\mathbb{A}) \quad \text{and} \quad U_t(\varkappa) \to [F_0(f)](\varkappa)$$

as $t \to 0$ for all $\varkappa \in \mathscr{X}$. Moreover,

$$[\mathbb{A}U_t](\varkappa) = [\mathbb{E}t^{-1}(F_t(f) - F_0(f))](\varkappa) = \frac{f(\varphi(t,\varkappa)) - f(\varkappa(0))}{t} \to G(\varkappa)$$

by (ii). The closedness of $\mathbb{A}$ yields (i). □

A class of examples that falls into the situation described in this section concerns *delay differential equations*. Let us discuss this example in more detail. We limit ourselves to a finite time horizon here. A treatment of an infinite time horizon is also possible but in that case it is more appropriate to consider a slightly different path space, see Example 2.3. Thus, fix $h \in (0, \infty)$ and $d \in \mathbb{N}$ and define

$$\mathscr{C}_h := C([-h, 0]; \mathbb{R}^d),$$

which is endowed with the supremum norm.

Given a continuous function $y : [-h, \infty) \to \mathbb{R}$ and $t \in [0, \infty)$, we define $y_t \in \mathscr{C}_h$ by setting $y_t(s) := y(t + s)$. Then y_t is the *past at time* t. Given a continuous map $b : \mathscr{C}_h \to \mathbb{R}^d$ and $\xi \in \mathscr{C}_h$, we consider the following delay differential equation:

$$\begin{cases} y'(t) = b(y_t) & \text{for } t \geq 0, \\ y_0 = \xi. \end{cases} \tag{6.1}$$

A *solution* of (6.1) is a continuously differentiable function $y : [-h, \infty) \to \mathbb{R}^d$, such that (6.1) is satisfied. It follows from [4, Theorem II.4.3.1], that if b is Lipschitz continuous, then for every $\xi \in \mathscr{C}_h$, the delay differential equation (6.1) has a unique solution y^ξ. The proof of [4, Theorem II.4.3.1], which is based on Banach's fixed point theorem, immediately yields continuity of the map $\xi \mapsto y^\xi$.

Setting $X := \mathbb{R}^d$ and $\mathscr{X} := C(\mathbb{R}; X)$, we now construct an evolutionary C_b-semigroup $\mathbb{T}$ associated to (6.1). To that end, define the map $\varphi : \mathscr{X} \to \mathscr{X}$ by setting

$$[\varphi(\varkappa)](t) := \begin{cases} \varkappa(t), & \text{if } t \leq 0, \\ y^{\varkappa_0}(t), & \text{if } t > 0. \end{cases} \tag{6.2}$$

As $\xi \mapsto y^\xi$ is continuous, it follows that φ is continuous. That φ satisfies conditions (i) and (ii) in Definition 6.1 is obvious and condition (iii) follows from uniqueness of solutions to (6.1). Thus, φ is an evolution map and it follows from Proposition 6.2 that $[\mathbb{E}F](\varkappa) := F(\varphi(\varkappa))$ is the expectation operator of a multiplicative evolutionary semigroup $\mathbb{T}$. We also note that $t \mapsto \varphi(t, \varkappa)$ is differentiable on $[0, \infty)$.

Proposition 6.5. *Let $b : \mathscr{C}_h \to \mathbb{R}^d$ be Lipschitz continuous, define φ by (6.2) and set $[\mathbb{E}F](\varkappa) = F(\varphi(\varkappa))$. We also define the semigroup $\mathbb{T}$ by setting $\mathbb{T}(t) = \mathbb{E}\Theta_t$ and denote its C_b-generator by $\mathbb{A}$.*

(a) *For every $t > 0$, it is $\mathbb{T}(t)B_b(\mathscr{X}, \mathscr{F}([-h, 0])) \subset B_b(\mathscr{X}, \mathscr{F}([-h, 0]))$.*

(b) *For every $f \in C_c^1(\mathbb{R}^d)$, it is $F_0(f) \in D(\mathbb{A})$ and*

$$[\mathbb{A}F_0(f)](\varkappa) = (\nabla f)(\varkappa(0)) \cdot b(\varkappa_0). \tag{6.3}$$

(c) *$\mathbb{T}$ is uniquely determined by* (b) *in the following sense. If $\mathbb{S}$ is a multiplicative evolutionary semigroup with generator $\mathbb{B}$, and for every $f \in C_c^1(\mathbb{R}^d)$, it holds $F_0(f) \in D(\mathbb{B})$ with $\mathbb{B}F_0(f) = \mathbb{A}F_0(f)$ as given by (6.3), then $\mathbb{S} = \mathbb{T}$.*

Proof. (a). Clearly, $\varphi(\varkappa)|_{[0,\infty)}$ is $\mathscr{F}([-h, 0])$-measurable. Thus, (a) follows from Lemma 4.8.

(b). As $[0, \infty) \ni t \mapsto \varphi(t, \varkappa)$ solves (6.1), the chainrule shows that $t \mapsto f(\varphi(t, \varkappa))$ is differentiable with derivative $(\nabla f)(\varphi(t, \varkappa)) \cdot b(\varphi(\varkappa)_t)$. In particular, the derivative at 0 is given by $(\nabla f)(\varkappa(0)) \cdot b(\varkappa_0)$. Note that the latter is a bounded function, as b is Lipschitz continuous and f has compact support. Thus, (b) follows from Proposition 6.4.

(c). Let $\mathbb{S}$ and $\mathbb{B}$ as in the statement. By Proposition 6.2, the expectation operator of $\mathbb{S}$ is induced by an evolution map ψ. We only have to prove that $\psi = \varphi$. By assumption, $F_0(f) \in D(\mathbb{B})$ with $[\mathbb{B}F_0(f)](\varkappa) = (\nabla f)(\varkappa(0)) \cdot b(\varkappa_0)$ for all $f \in C_c^1(\mathbb{R}^d)$. Applying this for functions f satisfying $f(x) = x_j$ on $[-n, n]^d$ for some $n \in \mathbb{N}$ and $j = 1, \ldots, d$, the characterization of the C_b-generator in (iv) of Theorem

A.12 implies that for every $\mathbb{x} \in \mathscr{X}$, the map $t \mapsto \psi(t, \mathbb{x})$ is differentiable from the right in 0 with derivative $b(\psi(\mathbb{x})_0)$. Now, fix $t > 0$ and set $\mathbb{y} = \vartheta_t \psi(\mathbb{x})$. Note that $\mathbb{y}(0) = \psi(\mathbb{x}, t) = \vartheta_t \psi(\mathbb{x}, 0)$. By Condition (iii) in Definition 6.1, it is

$$\begin{aligned}\frac{\psi(t+\varepsilon, \mathbb{x}) - \psi(t, \mathbb{x})}{\varepsilon} &= \frac{\vartheta_{t+\varepsilon}\psi(0, \mathbb{x}) - \vartheta_t \psi(0, \mathbb{x})}{\varepsilon} \\ &= \frac{\vartheta_\varepsilon(0, \psi(\vartheta_t\psi(\mathbb{x})) - \psi(0, \vartheta_t\psi(\mathbb{x}))}{\varepsilon} = \frac{\psi(\varepsilon, \mathbb{y}) - \psi(0, \mathbb{y})}{\varepsilon},\end{aligned}$$

for all $\varepsilon > 0$. Letting $\varepsilon \to 0$, it follows that $t \mapsto \psi(t, \mathbb{x})$ is differentiable from the right at every point $t \geq 0$ with derivative $b(\psi(\mathbb{x})_t)$. This shows that $t \mapsto \psi(t, \mathbb{x})$ is a $W^{1,\infty}_{\mathrm{loc}}$-solution of (6.1) with $\xi = \mathbb{x}_0$. Thus [4, Theorem II.4.3.1] (which works also for this weaker type of solutions) yields $\psi(\mathbb{x}) = \varphi(\mathbb{x})$. □

Remark 6.6. It is worth to point out that $D(\mathbb{A}) \setminus D_-(\mathbb{D})$ also contains elements that are not measurable with respect to $\mathscr{F}(\{0\})$. For example, arguing similar to the proof of (ii) ⇒ (i) in Proposition 6.4, one can show that for every $f \in C^1_c(\mathbb{R}^d)$ and $t > 0$, the function $F_{\varphi,t}(f) : \mathbb{x} \mapsto f(\varphi(t, \mathbb{x}))$ belongs to $D(\mathbb{A})$ and

$$[\mathbb{A}F_{\varphi,t}(f)](\mathbb{x}) = (\nabla f)(\varphi(t, \mathbb{x})) \cdot b(\varphi(\mathbb{x})_t).$$

Then $F_{\varphi,t}$ is $\mathscr{F}([-h, 0])$-measurable, but not, in general, $\mathscr{F}(\{0\})$-measurable.

Let us compare the evolutionary semigroup $\mathbb{T}$ constructed above with other semigroup approaches to the delay differential equation (6.1). We first consider the situation where the map b in (6.1) is linear.

Following [3] (see also [19, Section VI.6]), there is a semigroup on the space $\mathscr{C}_h$ whose generator is given by

$$Au = u' \quad \text{for} \quad u \in D(A) = \{v \in C^1([-h, 0]; \mathbb{R}^d) : v'(0) = b(v)\}. \tag{6.4}$$

Note that for $D(A)$ to be a linear space, it is essential that b is linear. We can identify $C_b(\mathscr{C}_h)$ with $C_b(\mathscr{X}, \mathscr{F}([-h, 0]))$ as follows. We define the extension map $\rho : \mathscr{C}_h \to \mathscr{X} = C(\mathbb{R}; \mathbb{R}^d)$ by

$$[\rho(\mathbb{x})](t) = \begin{cases} \mathbb{x}(-r), & \text{if } t \leq r, \\ \mathbb{x}(t), & \text{if } t \in (-r, 0), \\ \mathbb{x}(0), & \text{if } t \geq 0. \end{cases}$$

The map $F \mapsto \tilde{F} := F \circ \rho$ is a homeomorphism between $\mathscr{C}_h$ and $C_b(\mathscr{X}, \mathscr{F}([-h, 0]))$. Note that the latter is invariant under $\mathbb{T}$ by Proposition 6.5. Using [19, Corollary VI.6.3], we see that the relationship between $\mathbb{T}$ and T is given by

$$(\mathbb{T}(t)\tilde{F})(\rho(\mathbb{x})) = F(T(t)\mathbb{x}) \qquad \text{for all } \mathbb{x} \in \mathscr{C}_h.$$

In [4, Theorem II.4.2], a semigroup is constructed on the product space

$$M^p := \mathbb{R}^d \times L^p((-h, 0); \mathbb{R}^d)$$

for $1 \leq p < \infty$. Thus, an element u of M^p has two components $u_1 \in \mathbb{R}^d$ and $u_2 \in L^p((-h, 0); \mathbb{R}^d)$. Note that u_2 need not have a trace in 0, but one should think of u_1 as a replacement for that trace. The generator of this semigroup is given by

$$\begin{cases} A(u_1, u_2) = (b(u_2), u_2') \\ \qquad D(A) = \{v \in M^p : u_2 \in W^{1,p}((-h, 0); \mathbb{R}^d) \text{ and } u_1 = u_2(0)\}. \end{cases} \tag{6.5}$$

Again, linearity of b is necessary for the linearity of A.

In the case where b is nonlinear, one can describe solutions of (6.1) in terms of the semigroup for $b = 0$ by means of the variation of constants formula. This was done, for example, in [24], see in particular Section 3 therein for the definition of the generator. In [24], the representation of the solution via the variation of constants formula was used to establish regularity of viscosity solutions for the associated

HJB equation. In our approach, which may be regarded as a Koopman approach to such equations, even for nonlinear b, we obtain a semigroup directly describing the solutions, without relying on the variation of constants formula.

We note that in both semigroup approaches, the *time derivative* plays an important role. In (6.4), the generator A itself is a realization of the derivative, whereas in (6.5), it is the second component of the generator. It is well-known that the generator of (various versions of) the shift semigroup on $C([-h,0];\mathbb{R}^d)$ or $L^p((-h,0);\mathbb{R}^d)$ is a suitable realization of this time derivative. Thus, in our setting, the appropriate generalization of the time derivative is the operator $\mathbb{D}$, which enters the generator $\mathbb{A}$ of $\mathbb{T}$ through the requirement $D_-(\mathbb{D}_0)\subset D(\mathbb{A})$ and $\mathbb{A}F=\mathbb{D}F$ for $F\in D_-(\mathbb{D}_0)$, see Proposition 5.4. On the other hand, the delay differential equation (6.1) only appears 'at the point 0'. In (6.4) via the boundary condition in the generator, in (6.5) as the first component of A and in $\mathbb{A}$ via Proposition 6.5(b).

6.2. **Markov Processes.** In this section, we show how the concept of a classical Markov process (with continuous or càdlàg paths) fits into our general framework. Throughout this subsection, we consider $\mathscr{X}\in\{\mathscr{X}_{\mathrm{C}},\mathscr{X}_{\mathrm{D}}\}$. We recall that $0^\star=0$ if $\mathscr{X}=\mathscr{X}_{\mathrm{C}}$ and $0^\star=0-$ if $\mathscr{X}=\mathscr{X}_{\mathrm{D}}$. We note that

$$\sigma(\pi_{0^\star})=\bigcap_{\varepsilon>0}\mathscr{F}([-\varepsilon,0)).$$

With slight abuse of notation, we put $\mathscr{F}(\{0^\star\})=\sigma(\pi_{0^\star})$. Note that a function $F\in B_b(\mathscr{X})$ is $\mathscr{F}(\{0^\star\})$-measurable if and only if it is of the form $F_{0^\star}(f)$ for some $f\in B_b(X)$.

Theorem 6.7. *Let $\mathbb{T}$ be an evolutionary C_b-semigroup with expectation operator $\mathbb{E}$. Then the following are equivalent:*

(i) *$\mathbb{T}(t)F_{0^\star}(f)$ is $\mathscr{F}(\{0^\star\})$-measurable for all $t\geq 0$ and $f\in B_b(X)$;*
(ii) *$\mathbb{T}(t)F_{0^\star}(f)$ is $\mathscr{F}(\{0^\star\})$-measurable for all $t\geq 0$ and $f\in C_b(X)$;*
(iii) *$\mathbb{E}F$ is $\mathscr{F}(\{0^\star\})$-measurable for all $F\in B_b(\mathscr{X},\mathscr{F}([0,\infty)))$;*
(iv) *$\mathbb{E}F$ is $\mathscr{F}(\{0^\star\})$-measurable for all $F\in C_b(\mathscr{X},\mathscr{F}([0,\infty)))$.*

If these equivalent conditions are satisfied, the following hold true:

(a) *$\mathbb{T}$ induces a C_b-semigroup T on $C_b(X)$ in the sense that $\mathbb{T}(t)F_{0^\star}(f)=F_{0^\star}(T(t)f)$ for all $t>0$ and $f\in C_b(X)$.*
(b) *Denote the kernel associated to $\mathbb{E}$ by $\Bbbk$ and write $\mathbb{P}^{\mathbb{x}}$ for the probability measure $\Bbbk(\mathbb{x},\cdot)$. For every $\mathbb{x}\in\mathscr{X}$, under the measure $\mathbb{P}^{\mathbb{x}}$, the canonical process $(Z_t)_{t\geq0}$ given by $Z_t(\mathbb{x}):=\mathbb{x}(t)$ is a Markov process with transition semigroup T starting at $\mathbb{x}(0^\star)$.*

Proof. The implication (i) ⇒ (ii) is trivial, while the implication (ii) ⇒ (i) follows from a bp-closedness argument. Taking Lemma 5.10 into account, the equivalence (ii) ⇒ (iii) follows similar to the proof of Lemma 4.8. As $\mathbb{T}$ is assumed to be an evolutionary C_b-semigroup either Theorem 5.2 or Theorem 5.8 yields $\mathbb{E}F\in\hat{C}_b(\mathscr{X}^-)$ for all $F\in C_b(\mathscr{X})$. Taking this into account the implication (iii) ⇒ (iv) is trivial and the converse implication follows, once again, by a bp-closedness argument.

Now assume that the equivalent conditions are satisfied.

(a). By assumption, given $t>0$ for every $f\in C_b(X)$, there exists a unique element $T(t)f\in C_b(X)$ such that $\mathbb{T}(t)F_{0^\star}(f)=F_{0^\star}(T(t)f)$. As $\mathbb{T}(t)F_{0^\star}(f)=\mathbb{E}\Theta_tF_{0^\star}(f)$, the fact that $\mathbb{E}$ is a kernel operator implies that $T(t)$ is a kernel operator and from the semigroup property of $\mathbb{T}$ one easily deduces the semigroup property of T. Next, let $\iota: X\to\mathscr{X}$ be defined by $[\iota(x)](t)\equiv x$. Then ι is continuous and $T(t)f(x)=[\mathbb{T}(t)F_{0^\star}(f)](\iota(x))$. Thus, the continuity of the map $(t,x)\mapsto T(t)f(x)$ follows from that of $(t,\mathbb{x})\mapsto[\mathbb{T}(t)F_{0^\star}(f)](\mathbb{x})$. This proves (a).

(b). Applying Proposition 4.7 to $F = F_{0^\star}(f)$ for some $f \in C_b(X)$, it follows from Lemma 5.10 that for every $s, t \geq 0$,

$$\big[\mathbb{E}^{\mathbb{x}}[f(Z_{t+s})|\mathscr{F}_t]\big](\mathbb{y}) = \big[\mathbb{E}^{\mathbb{x}}[\Theta_{t+s}F_0(f)|\mathscr{F}_t]\big](\mathbb{y}) = [\mathbb{T}(s)F_{0^\star}(f)](\vartheta_t\mathbb{y}) = (T(s)f)(Z_t(\mathbb{y}))$$

for $\mathbb{P}^{\mathbb{x}}$-almost all $\mathbb{y} \in \mathscr{X}$. ☐

It is a natural question, if every Markovian C_b-semigroup T on $C_b(X)$ can be lifted to an evolutionary semigroup $\mathbb{T}$. For this, it is certainly necessary (by Theorem 6.7(b)) that the associated Markovian process can be realized with continuous, resp. càdlàg paths. We will prove next that this condition is also sufficient. To that end, we put $\mathscr{X}^+ := C([0,\infty); X)$ if $\mathscr{X} = \mathscr{X}_{\mathrm{C}}$ and $\mathscr{X}^+ := D([0,\infty); X)$ if $\mathscr{X} = \mathscr{X}_{\mathrm{D}}$. Let $T = (T(t))_{t\geq 0}$ be a Markovian C_b-semigroup on X and denote by p_t the kernel associated to $T(t)$.

Definition 6.8. Let T be a C_b-semigroup on $C_b(X)$. We say that the associated Markov process *can be realized with paths in* $\mathscr{X}^+$ if

(i) for every $x \in X$ we find a measure $\mathbf{P}^x$ on $\mathscr{X}^+$ such that, under this measure, the canonical process is a Markov process with transition semigroup T starting at x;
(ii) the map $x \mapsto \mathbf{P}^x$ is weakly continuous.

Remark 6.9. Assume that X is locally compact and that T is a *Feller semigroup*, i.e. it leaves the space $C_0(X)$ of continuous functions vanishing at infinity invariant and is a strongly continuous semigroup on that space. Then the associated Markov process can be realized with paths in $\mathscr{X}^+ = D([0,\infty); X)$. Indeed, the existence of the measure $\mathbf{P}^x$ follows from [20, Theorem 4.2.7] and the continuity of the map $x \mapsto \mathbf{P}^x$ follows from [20, Theorem 4.2.5].

Theorem 6.10. *Let $T = (T(t))_{t\geq 0}$ be a Markovian C_b-semigroup such that the associated Markov process can be realized with paths in $\mathscr{X}^+$. Then there exists an evolutionary C_b-semigroup $\mathbb{T} = (\mathbb{T}(t))_{t\geq 0}$ such that $\mathbb{T}(t)F_0(f) = F_0(T(t)f)$ for all $f \in C_b(X)$.*

Proof. We denote by $(\mathbf{P}^x)_{x\in X}$ the family of probability measures on $\mathscr{X}^+$ such that under these measures the canonical process on $\mathscr{X}^+$ is a Markov process with transition semigroup T starting at x. Motivated by Remark 4.3, we want to construct a kernel $\mathbb{k}$ with $\mathbb{k}_+(\mathbb{x}, \cdot) = \mathbf{P}^{\mathbb{x}(0^\star)}$. To that end, we let $A_- \in \mathscr{F}((-\infty, 0))$ and $A_+ \in \mathscr{F}([0,\infty))$. Note that we can (and shall) identify these with elements of $\mathfrak{B}(\mathscr{X}^-)$ and $\mathfrak{B}(\mathscr{X}^+)$ respectively. Given $\mathbb{x} \in \mathscr{X}$, we set

$$\mathbb{k}(\mathbb{x}, A_- \cap A_+) := \delta_{\mathbb{x}}(A_-)\mathbf{P}^{\mathbb{x}(0^\star)}(A_+).$$

Note that, in particular, every cylinder set can be written as an intersection $A_- \cap A_+$ with A_-, A_+ as above. Arguing as in the proof of Fubini's theorem (see [6, Theorem 3.3.1]), $\mathbb{k}(\mathbb{x}, \cdot)$ can be extended to a measure on all of $\mathfrak{B}(\mathscr{X})$. Also, it is straightforward to show that $\mathbb{x} \mapsto \mathbb{k}(\mathbb{x}, A)$ is measurable for every $A \in \mathfrak{B}(\mathscr{X})$, so $\mathbb{k}$ is a kernel.

We prove that $\mathbb{k}$ satisfies Proposition 4.1(iii). To that end, let $A_-, B_- \in \mathscr{F}_{0-}$ and $B_+ \in \mathscr{F}([0,\infty))$. Setting $A = B_- \cap B_+$, it follows that

$$\begin{aligned}\mathbb{k}(\mathbb{x}, A_- \cap A) &= \mathbb{k}(\mathbb{x}, (A_- \cap B_-) \cap B_+) = \delta_{\mathbb{x}}(A_- \cap B_-)\mathbf{P}^{\mathbb{x}(0^\star)}(B_+)\\ &= \delta_{\mathbb{x}}(A_-)\delta_{\mathbb{x}}(B_-)\mathbf{P}^{\mathbb{x}(0^\star)}(B_+)\\ &= \delta_{\tau(\mathbb{x})}(A_-)\delta_{\tau(\mathbb{x})}(B_-)\mathbf{P}^{(\tau(\mathbb{x}))(0^\star)}(B_+)\\ &= \delta_{\tau(\mathbb{x})}(A_-)\mathbb{k}(\tau(\mathbb{x}), A).\end{aligned}$$

By a monotone class argument, this equality generalizes to arbitrary $A \in \mathfrak{B}(\mathscr{X})$. Thus, by Proposition 4.1, $\mathbb{k}$ is the kernel of an expectation operator. To conclude that $\mathbb{E}$ is the expectation operator of an evolutionary semigroup, it remains to verify that $\mathbb{E}$ is homogeneous. This is easily verified using Lemma 4.4 (note that it suffices to consider a cylinder set A in that Lemma) and the Chapman–Kolmogorov equations for the transition kernels p_t of $T(t)$. In the case of càdlàg paths, one additionally uses the observation that under $\mathbf{P}^x$ the measure of a cylinder set does not change, if a condition at a time $t > 0$ is instead imposed at time $t^\star = t-$. That the evolutionary semigroup associated to the expectation operator $\mathbb{E}$ is a C_b-semigroup follows from Theorem 5.2 in case of continuous paths and from Theorem 5.8 in case of càdlàg paths. Indeed for $F \in C_b(\mathscr{X}, \mathscr{F}([0,\infty))) \simeq C_b(\mathscr{X}^+)$, we have

$$[\mathbb{E}F](\mathbb{x}) = \int_{\mathscr{X}^+} F(\mathbb{y})\mathbf{P}^{\mathbb{x}(0-)}(d\mathbb{y})$$

and the latter depends continuously on $\mathbb{x}(0-)$. □

Example 6.11. An example that fits into the setting described in Theorem 6.10 is given by stochastic differential equations. Consider the equation

$$\begin{cases} dY(t) = b(Y(t))dt + \sigma(Y(t))dB(t), & \text{for } t \geq 0, \\ Y(0) = y. \end{cases} \tag{6.6}$$

Here, $(B(t))_{t\geq 0}$ is an m-dimensional Brownian motion, defined on a fixed probability space $(\Omega, \Sigma, \mathbf{P})$ and the functions $b\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are globally Lipschitz continuous.

Under this assumption it is well-known (see, e.g., [42, Theorem 5.2.1]) that for every $y \in \mathbb{R}^d$, there exists a unique strong solution $(Y^y(t))_{t\geq 0}$ of (6.6) which is adapted to the filtration $(\mathscr{G}_t)_{t\in[0,T]}$ generated by $\{B(s) : s \leq t\}$. The process $(Y^y(t))_{t\geq 0}$ is called a (time-homogeneous) *Itô diffusion process.*

It follows from [42, Theorem 7.1.2], that $(Y^y(t))_{t\geq 0}$ is Markovian. Defining for each $y \in \mathbb{R}^d$ the measure $\mathbf{P}^y$ on $\mathscr{X}^+ := C([0,\infty); \mathbb{R}^d)$ by $\mathbf{P}^y(A_+) := \mathbf{P}(Y^y \in A_+)$, [42, Lemma 8.1.4] implies weak continuity of the map $y \mapsto \mathbf{P}^y$. Thus, by Theorem 6.10, we can construct an evolutionary C_b-semigroup from the measures $(\mathbf{P}^y)_{y\in\mathbb{R}^d}$. To describe the evolution operator $\mathbb{E}$, we use the representation from Remark 4.3 with $\mathbb{k}_+(\mathbb{x}, \cdot) = \mathbf{P}^{\mathbb{x}(0)}$. By Equation (4.2), we have

$$(\mathbb{E}F)(\mathbb{x}) = \int_{\mathscr{X}^+} F(\mathbb{x}_- \otimes_0 \mathbb{y}_+)\mathbf{P}^{\mathbb{x}(0)}(d\mathbb{y}_+) = \mathbf{E}F\big(\mathbb{x}_- \otimes_0 Y^{\mathbb{x}(0)}\big). \tag{6.7}$$

Note that the integrals are well-defined as the process $(Y^y(t))_{t\geq 0}$ has $\mathbf{P}$-almost surely continuous paths.

For the rest of this subsection, let $\mathbb{T}$ be an evolutionary C_b-semigroup satisfying the equivalent conditions of Theorem 6.7 and denote by $T = (T(t))_{t\geq 0}$ the induced semigroup on $C_b(X)$. It is a rather natural question, how the C_b-generator $\mathbb{A}$ of $\mathbb{T}$ and the C_b-generator A of T are related. To study this question, we introduce the following notation. Denote the transition kernels of $(T(t))_{t\geq 0}$ by $(p_t)_{t\geq 0}$. Given $f_1, \dots f_n \in C_b(X)$, we put

$$F(s_1, \dots, s_n, x) = \int_{X^n} p_{s_1}(x, dy_1) \prod_{j=2}^n p_{s_j}(y_j, dy_{j-1}) \prod_{k=1}^n f_k(y_k).$$

Moreover, for $n \in \mathbb{N}$ and $0 \leq a < b < \infty$, we put

$$D_n(a,b) := \{(s_1, \dots, s_n) \in [0,\infty)^n : a \leq s_1 \leq s_2 \leq \dots \leq s_n \leq b\}.$$

Lemma 6.12. *With the above notation, the function*

$$x \mapsto \int_{D_n(a,b)} F(s_1, \dots, s_n; x)$$

belongs to the domain of the generator A of T and

$$A\int_{D_n(a,b)} F(s_1,\dots,s_n) = \int_{D_{n-1}(a,b)} F(s_1,\dots,s_{n-1},b) - \int_{D_{n-1}(a,b)} F(a,s_2,\dots,s_n).$$

Proof. Making use of the semigroup law and substitution, we see that

$$T(h)\int_{D_n(a,b)} F(s_1,\dots,s_n) = \int_{D_n(a,b)} F(s_1+h,s_2,\dots,s_n) = \int_{D_n(a+h,b+h)} F(s_1,\dots,s_n).$$

We now define the sets $\overline{D}_n$ and $\underline{D}_n$ by

$$\overline{D}_n(a,b,h) := \{(s_1,\dots,s_n) : a\le s_1\le\cdots\le s_{n-1}\le b\le s_n\le b+h\}$$
$$\underline{D}_n(a,b,h) := \{(s_1,\dots,s_n) : a\le s_1\le a+h, s_1\le s_2\cdots\le s_n\le b\}.$$

Using these, we see that

$$\begin{aligned}(T(h)-I)\int_{D_n(a,b)} &= \int_{D_n(a+h,b+h)} - \int_{D_n(a,b)}\\ &= \int_{D_n(a+h,b+h)} - \int_{D_n(a+h,b)} - \int_{\underline{D}_n(a,b,h)}\\ &= \int_{\overline{D}_n(a,b,h)} - \int_{\underline{D}_n(a,b,h)} + o(h).\end{aligned}$$

Using dominated convergence, it is easy to see that

$$\frac{1}{h}\int_{\overline{D}_n(a,b,h)} F(s_1,\dots,s_n) \to \int_{D_{n-1}(a,b)} F(s_1,\dots,s_{n-1},b)$$

and

$$\frac{1}{h}\int_{\underline{D}_n(a,b,h)} F(s_1,\dots,s_n) \to \int_{D_{n-1}(a,b)} F(a,s_2,\dots,s_n).$$

Combining this with the above, the claim follows. □

Theorem 6.13. *Let $\mathbb{T}$ be an evolutionary C_b-semigroup satisfying the equivalent conditions of Theorem 6.7. We denote by $\mathbb{A}$ the C_b-generator of $\mathbb{T}$ and by A the C_b-generator of the induced semigroup T on $C_b(X)$.*

(a) *For every $u\in D(A)$, we have $F_{0^\star}(u)\in D(\mathbb{A})$ and $\mathbb{A}F_{0^\star}(u) = F_{0^\star}(Au)$.*
(b) *Functions of the form*

$$U = \sum_{j=1}^n U_j F_{0^\star}(u_j) \tag{6.8}$$

where $U_1,\dots,U_n\in D_-(\mathbb{D})$ and $u_1,\dots,u_n\in D(A)$ form
  (i) *a β_0-core for $\mathbb{A}$ in the case $\mathscr{X}=\mathscr{X}_{\mathrm{C}}$;*
  (ii) *a bp-core for $\mathbb{A}_{\mathsf{full}}$ in the case $\mathscr{X}=\mathscr{X}_{\mathrm{D}}$.*

Proof. (a). Let $u\in D(A)$ with $Au=f$. By Proposition A.4, this is equivalent with $T(t)u-u=\varphi_t := \int_0^t T(s)f\,ds$. Consequently,

$$\begin{aligned}\mathbb{T}(t)F_{0^\star}(u) - F_{0^\star}(u) &= F_{0^\star}(T(t)u) - F_{0^\star}(u) = F_{0^\star}(\varphi_t)\\ &= \mathbb{E}\int_0^t F_{s^\star}(f)\,ds = \int_0^t \mathbb{E}F_{s^\star}(f)\,ds = \int_0^t \mathbb{T}(s)F_{0^\star}(f)\,ds.\end{aligned}$$

By Proposition A.4 for $\mathbb{T}$, it follows that $F_{0^\star}(u)\in D(\mathbb{A})$ and $\mathbb{A}F_{0^\star}(u) = F_{0^\star}(Au)$.

(b). Here, we make use of Theorem 5.2(b) (in the case $\mathscr{X} = \mathscr{X}_{\mathrm{C}}$) and Proposition 4.10(b) (in the case $\mathscr{X} = \mathscr{X}_{\mathrm{D}}$). In both cases, it suffices to show that every function in $D(\mathbb{A}_0)$ is of the form in (6.8). To that end, it suffices to compute $\mathbb{E}V$ and $\mathbb{E}\mathbb{D}V$ for $V \in D_+(\mathbb{D}_0)$, i.e. when V is a product of functions of the form $F_a^b(f)$ (see Proposition 3.2). Note that any such V can be written as the sum of integrals of the form

$$I(\varkappa) = \int_{[a,b]^n} \prod_{k=1}^n f_k(\varkappa(s_k)),$$

where the parameters $n \in \mathbb{N}$ and $0 \le a < b$ may vary from summand to summand. By linearity, it suffices to consider a single integral I. By Proposition 3.2,

$$\mathbb{D}I(\varkappa) = \sum_{k=1}^n (f_k(\varkappa(b)) - f_k(\varkappa(a))) \int_{[a,b]^{n-1}} \prod_{j \neq k} f_j(\varkappa(s_j)) = \sum_{k=1}^n W_k(\varkappa).$$

To compute $\mathbb{E}I$ and $\mathbb{E}\mathbb{D}I$, the order of the variables $s_1, \ldots, s_n$ has to be taken into account. To that end, we make use of the symmetric group S_n and decompose

$$I(\varkappa) = \sum_{\sigma \in S_n} \int_{D_n(a,b)} \prod_{k=1}^n f_k(\varkappa(s_{\sigma(k)})).$$

It follows that

$$\mathbb{E}I(\varkappa) = \sum_{\sigma \in S_n} \int_{D_n(a,b)} F(s_{\sigma(1)}, \ldots, s_{\sigma(n)}; \varkappa(0^\star)).$$

It is a consequence of Lemma 6.12 that, as a function of $\varkappa(0^\star)$, this is an element of $D(A)$ and

$$\begin{aligned}
(A\mathbb{E}I)(\varkappa(0^\star)) &= \sum_{\sigma \in S_n} \int_{D_{n-1}(a,b)} F(s_{\sigma(1)}, \ldots, s_{\sigma(n-1)}, b; \varkappa(0^\star)) \\
&\qquad - \int_{D_{n-1}(a,b)} F(a, s_{\sigma(2)}, \ldots, s_{\sigma(n)}; \varkappa(0^\star)) \\
&= \sum_{k=1}^n \sum_{\substack{\sigma \in S_n \\ \sigma(n)=k}} \int_{D_{n-1}(a,b)} F(s_{\sigma(1)}, \ldots, s_{\sigma(n-1)}, b; \varkappa(0^\star)) \\
&\qquad - \sum_{\substack{\sigma \in S_n \\ \sigma(1)=k}} \int_{D_{n-1}(a,b)} F(a, s_{\sigma(2)}, \ldots, s_{\sigma(n)}; \varkappa(0^\star)) \\
&= \sum_{k=1}^n \mathbb{E}W_k = \mathbb{E}\mathbb{D}I.
\end{aligned}$$

Here we have used in the last line that as σ runs trough all of S_n with either the first or the last entry fixed, we run through all of S_{n-1} in the remaining entries. Thus, decomposing the integral in W_k as the integral in I above, the equality follows. □

Example 6.14. In the situation of Example 6.11, the generator of the semigroup T on $C_b(\mathbb{R}^d)$ related to the Itô diffusion Y^y is given by

$$(Au)(y) = b(y) \cdot \nabla u(y) + \frac{1}{2} \operatorname{tr}\Big((\sigma\sigma^\top)(y) D^2 u(y)\Big) \quad \text{for } y \in \mathbb{R}^d,$$

for sufficiently smooth u, e.g. for $u \in C_c^2(\mathbb{R}^d)$ (see [42, Theorem 7.3.3]). Theorems 6.7 and 6.13 relate the 'classical' semigroup T and its generator A to the evolutionary semigroup $\mathbb{T}$ and its C_b-generator $\mathbb{A}$ as follows:

$$\begin{aligned}
\mathbb{T}(t)F_0(f) &= F_0(T(t)f) \quad && \text{for } t > 0 \text{ and } f \in C_b(X), \\
\mathbb{A}F_0(u) &= F_0(Au) && \text{for } u \in D(A).
\end{aligned}$$

Remark 6.15. Although the underlying dynamics are Markovian, the extension of the transition semigroup to the path space allows to consider path-dependent functions and their evolutions. For example, in finance, this enables the dynamic pricing of financial derivatives, whose payoffs depend not only on the final value of the underlying asset but also on its entire historical trajectory.

Indeed, if the risk-neutral dynamics $(Z_t)_{t\geq 0}$ are Markovian, then the fair price of a derivative $f(Z_\tau)$ with maturity $\tau \geq 0$ at time $0 \leq t \leq \tau$ is given by

$$u(t, Z_t(\mathbb{y})) = \big[\mathbb{E}^{\mathbb{x}}[f(Z_\tau(\mathbb{y})) \mid \mathscr{F}_t]\big](Z_t(\mathbb{y}))$$

for $\mathbb{P}^{\mathbb{x}}$-almost all $\mathbb{y} \in \mathscr{X}$. Under appropriate assumptions, the value function u satisfies the final value problem

$$\begin{cases} \partial_t u(t,x) = -Au(t,x), \\ \quad u(\tau, x) = f(x), \end{cases}$$

for all $0 \leq t \leq \tau$ and $x \in X$. In particular, the value function is given by $u(t,\cdot) = T(\tau - t)f$, where T denotes the associated transition semigroup with generator A.

Using the extension $\mathbb{T}$ on $\mathscr{X}^+$ from Theorem 6.10, the value function

$$\Theta_t U(t) = \mathbb{E}^{\mathbb{x}}[F \mid \mathscr{F}_t]$$

corresponding to the fair price of a path-dependent derivative $F \in B_b(\mathscr{X}, \mathscr{F}_\tau)$ is given by $U(t) = \mathbb{T}(\tau - t)F$. For further details, we refer to our forthcoming article [14], where we provide a detailed analysis of the connection between such value functions, martingales, final value problems, and path-dependent PDEs.

Path-dependent functions also play an important role in control theory, where path-dependent stochastic optimization problems are usually formulated by means of backward stochastic differential equations [44]. If the controlled function is path-dependent, then the value function satisfies a semi-linear path-dependent parabolic equation. For a detailed survey on path-dependent PDEs we refer to [46], in particular to [45] for Sobolev solutions and [11, 18] for viscosity solutions.

6.3. **Stochastic delay equations.** In this section, we show that stochastic delay equations give rise to evolutionary C_b-semigroups on the path space $((\mathscr{X}, \mathbb{d}), \tau) = ((\mathscr{X}_{\mathrm{C}}, \mathbb{d}_{\mathrm{C}}), \tau_{\mathrm{C}})$ of continuous functions. We fix $h \in (0,\infty)$ and use the space $\mathscr{C}_h := C([-h,0];\mathbb{R}^d)$ introduced in Section 6.1. Also, for a stochastic process $(Y(t))_{t\geq -h}$, we denote the history at time t by $Y_t \in \mathscr{C}_h$.

Given an m-dimensional Brownian motion $(B(t))_{t\geq 0}$ on the probability space $(\Omega, \Sigma, \mathbf{P})$, consider the stochastic delay equation

$$\begin{cases} dY(t) = b(Y_t)dt + \sigma(Y_t)dB(t), \quad \text{for } t \geq 0, \\ \quad Y_0 = \xi, \end{cases} \tag{6.9}$$

where $\xi \in \mathscr{C}_h$ is given and $b\colon \mathscr{C}_h \to \mathbb{R}^d$ and $\sigma\colon \mathscr{C}_h \to \mathbb{R}^{d\times m}$ are Lipschitz continuous. Again, it is well-known that for every $\xi \in \mathscr{C}_h$, the stochastic delay equation (6.9) has a unique strong solution $(Y^\xi(t))_{t\geq 0}$ which is adapted to the filtration $(\mathscr{G}_t)_{t\geq 0}$ generated by $(B(t))_{t\geq 0}$ and has almost surely continuous paths (see, e.g., [41, Chapter 5, Theorem 2.2]).

In the notation of Remark 4.3, we define

$$\mathbf{P}^\xi(A_+) := \mathbf{P}(Y^\xi \in A_+)$$

for $\xi \in \mathscr{C}_h$ and $A_+ \in \mathfrak{B}(\mathscr{X}^+)$ as well as (analogously to Equation (6.7))

$$(\mathbb{E}F)(\mathbb{x}) := \mathbf{E}^{\mathbb{x}_0}\left[F(\mathbb{x}_- \otimes_0 \cdot)\right] = \int_{\mathscr{X}^+} F(\mathbb{x}_- \otimes_0 \mathbb{y}_+)\mathbf{P}^{\mathbb{x}_0}(d\mathbb{y}_+) = \mathbf{E}F(\mathbb{x}_- \otimes_0 Y^{\mathbb{x}_0}).$$

Obviously, $\mathbb{E}$ is a Markovian kernel operator satisfying condition (ii) of Proposition 4.1 and therefore an expectation operator.

To prove that $\mathbb{E}$ is homogeneous, we impose an additional non-degeneracy condition: For every $\xi \in \mathscr{C}_h$, the matrix $\sigma(\xi)$ is invertible and $\sup_{\xi\in\mathscr{C}_h} \|\sigma(\xi)^{-1}\| < \infty$. Under this assumption, it is proved in [8] that the stochastic delay equation (6.9) induces a Markov process on the state space $\mathscr{C}_h$. We can use this to establish the homogeneity of $\mathbb{E}$. To that end, we introduce the following notation: For $\xi \in \mathscr{C}_h$ and $t \geq 0$, let $Y^{t,\xi}$ be the unique strong solution of

$$\begin{cases} Y^{t,\xi}(s) = \xi(0) + \int_t^s b(Y_r^{t,\xi})dr + \int_t^s \sigma(Y_r^{t,\xi})dB(r), \quad \text{for } s \geq t, \\ \quad Y_t^{t,\xi} = \xi. \end{cases} \tag{6.10}$$

We also set

$$\big(\mathbb{x} \otimes_t Y^{t,\mathbb{x}_t}\big)(\omega, s) := \begin{cases} \mathbb{x}(s), & \text{if } s \leq t, \\ Y^{t,\mathbb{x}_t}(\omega, s), & \text{if } s \geq t. \end{cases}$$

With this notation, we have $(\mathbb{E}F)(\mathbb{x}) = \mathbf{E}F(\mathbb{x} \otimes_0 Y^{0,\mathbb{x}_0})$. Also, for $t \geq 0$, we have

$$(\mathbb{E}_t F)(\mathbb{x}) = \mathbf{E}F\big[\vartheta_{-t}(\vartheta_t\mathbb{x} \otimes_0 Y^{0,(\vartheta_t\mathbb{x})_0})\big] = \mathbf{E}F(\mathbb{x} \otimes_t Y^{t,\mathbb{x}_t}).$$

The second equality holds due to the uniquenss of the solution of (6.10) and the fact that $B(t+\cdot) - B(t)$ and $B(\cdot)$ have the same distribution. Using this, we obtain

$$(\mathbb{E}\mathbb{E}_t F)(\mathbb{x}) = \mathbf{E}\Big[\mathbf{E}F(\mathbb{y} \otimes_t Y^{t,\mathbb{y}_t})\Big|_{\mathbb{y}=\mathbb{x}\otimes_0 Y^{0,\mathbb{x}_0}}\Big].$$

It was shown in [8, Proof of Proposition 4.1] that

$$\mathbf{E}F(\mathbb{y} \otimes_t Y^{t,\mathbb{y}_t})\Big|_{\mathbb{y}=\mathbb{x}\otimes_0 Y^{0,\mathbb{x}_0}} = \mathbf{E}\Big[F(\mathbb{x} \otimes_0 Y^{0,\mathbb{x}_t})\Big|\mathscr{G}_t\Big] \quad \mathbf{P}\text{-a.s.}$$

From this, we obtain

$$(\mathbb{E}\mathbb{E}_t F)(\mathbb{x}) = \mathbf{E}\Big[\mathbf{E}\big[F(\mathbb{x} \otimes_0 Y^{0,\mathbb{x}_0})|\mathscr{G}_t\big]\Big] = \mathbf{E}F(\mathbb{x} \otimes_0 Y^{0,\mathbb{x}_0}) = (\mathbb{E}F)(\mathbb{x})$$

by the tower property. Hence, $\mathbb{E}$ is homogeneous and therefore generates an evolutionary semigroup on $C_b(\mathscr{X})$.

Similar to the case of of deterministic delay equations (see the discussion at the end of Subsection 6.1), a stochastic variation of constants formula can be used to describe the solution of the above SDDE by means of the shift semigroup. In the product space $M^2 = \mathbb{R}^d \times L^2((-h, 0); \mathbb{R}^d)$, where the generator of the shift semigroup is given by (6.5) with $b \equiv 0$, this approach was used, e.g., in [30], [13], [49]. Working in M^2 allows us to use results from [12] on abstract SDEs in Hilbert spaces. However, working in a Hilbert space is not always natural from the point of view of applications, as the coefficients b and σ are ill-adapted to this setting. In this situations, it is preferable to work on spaces of continuous functions (or even càdlàg functions) which, however, poses other problems. We refer the reader to [27], [23], [25] for recent developments in this directions.

We point out that, similar to the deterministic equation, the semigroup used in the stochastic variation of constants formula corresponds to an equation with linear coefficients. Once again, the evolutionary semigroup in our appropach serves as a Koopman semigroup and yields a direct description of solutions even for nonlinear coefficients.

6.4. **Stochastic flows driven by Lévy processes.** In this subsection, we consider a generalization of the deterministic equations considered in Section 6.1, where the evolution of our system is influenced by a random parameter which lives in an external probability space. We will work on the càdlàg path space $((\mathscr{X}, \mathbb{d}), \tau) = ((\mathscr{X}_{\mathrm{D}}, \mathbb{d}_{J_1}), \tau_D)$ endowed with the J_1-topology with values in a complete separable metric space X.

Definition 6.16. Let $\Omega := D([0,\infty);\mathbb{R}^d)$ be the space of all càdlàg functions $\omega\colon [0,\infty) \to \mathbb{R}^d$ endowed with the σ-algebra $\mathscr{G} := \sigma(\omega(s)\colon s \in [0,\infty))$ and the filtration $\mathscr{G}_t := \sigma(\omega(s)\colon s \in [0,t])$, $t \in [0,\infty)$. A *random evolution map* is a measurable function $\phi : \Omega \times \mathscr{X} \to \mathscr{X}$ such that for fixed $\omega \in \Omega$, the map $\mathbb{x} \mapsto \phi(\omega,\mathbb{x})$ is continuous on $\mathscr{X}^-$ and for every $\omega \in \Omega$ and $\mathbb{x} \in \mathscr{X}$, the following conditions are satisfied:

(i) $\phi(\omega,\mathbb{x}) = \phi(\omega,\tau(\mathbb{x}))$;
(ii) $\tau(\phi(\omega,\mathbb{x})) = \tau(\mathbb{x})$;
(iii) $\phi(\omega,\mathbb{x}) = \phi(\omega + c,\mathbb{x})$ for all $c \in \mathbb{R}^d$;
(iv) $\omega \mapsto \tau_t\phi(\omega,\mathbb{x})$ is $\mathscr{G}_t$-measurable for all $t \in [0,\infty)$;
(v) $\phi(\vartheta_t\omega, \vartheta_t\phi(\omega,\mathbb{x})) = \vartheta_t\phi(\omega,\mathbb{x})$ for all $t \in [0,\infty)$.

Proposition 6.17. *Let ϕ be a random evolution map and $\mathbf{P}$ be a probability measure on $(\Omega,\mathscr{G})$ such that the canonical process $Z_t(\omega) := \omega(t)$ is a Lévy process starting at 0. Then, the operator $\mathbb{E}\colon B_b(\mathscr{X}) \to B_b(\mathscr{X})$, given by $[\mathbb{E}F](\mathbb{x}) := \mathbf{E}\left[F(\phi(Z,\mathbb{x}))\right]$, is a homogeneous expectation operator. The induced evolutionary semigroup is a C_b-semigroup.*

Proof. First, we show that $\mathbb{E}F$ is $\mathscr{F}_{0-}$-measurable for all $F \in B_b(\mathscr{X})$. Indeed, condition (i) implies $[\mathbb{E}F](\mathbb{x}) = \mathbf{E}\left[F(\phi(Z,\mathbb{x}))\right] = \mathbf{E}\left[F(\phi(Z,\tau(\mathbb{x})))\right] = [\mathbb{E}F](\tau(\mathbb{x}))$ for all $\mathbb{x} \in \mathscr{X}$, and the claim follows directly from Lemma 2.5(c).

Second, we note that $\mathbb{E}F = F$ whenever $F \in B_b(\mathscr{X},\mathscr{F}_{0-})$. Indeed, since $F = F \circ \tau$, it follows from condition (ii) that for every $\mathbb{x} \in \mathscr{X}$,

$$[\mathbb{E}F](\mathbb{x}) = \mathbf{E}\left[F\left(\tau(\phi(Z,\mathbb{x}))\right)\right] = \mathbf{E}\left[F(\tau(\mathbb{x}))\right] = F(\tau(\mathbb{x})) = F(\mathbb{x}).$$

Third, we verify that $\mathbb{E}$ is homogeneous. Let $F \in B_b(\mathscr{X})$. By definition of $\mathbb{E}_t$, we have

$$[\mathbb{E}_tF](\mathbb{x}) = [\Theta_t\mathbb{E}\Theta_{-t}F](\mathbb{x}) = \mathbf{E}\left[F(\vartheta_{-t}\phi(Z,\vartheta_t\mathbb{x}))\right]$$

and therefore

$$[\mathbb{E}\mathbb{E}_tF](\mathbb{x}) = \mathbf{E}\Big[\tilde{\mathbf{E}}\big[F\big(\vartheta_{-t}\phi\big(\tilde{Z},\vartheta_t\phi(Z,\mathbb{x})\big)\big)\big]\Big],$$

where $\tilde{Z}$ is an independent copy of Z. Moreover, using first the tower property of conditional expectation, then condition (v) and then condition (i), it follows that

$$\begin{aligned}
[\mathbb{E}F](\mathbb{x}) &= \mathbf{E}\left[F(\phi(Z,\mathbb{x}))\right] \\
&= \mathbf{E}\big[\mathbf{E}[F(\phi(Z,\mathbb{x})) \mid \mathscr{G}_t]\big] \\
&= \mathbf{E}\Big[\mathbf{E}\big[F\big(\vartheta_{-t}\phi(Z_{t+\cdot},\vartheta_t\phi(Z,\mathbb{x}))\big) \mid \mathscr{G}_t\big]\Big] \\
&= \mathbf{E}\Big[\mathbf{E}\big[F\big(\vartheta_{-t}\phi(Z_{t+\cdot} - Z_t + Z_t,\vartheta_t\tau_t\phi(Z,\mathbb{x}))\big) \mid \mathscr{G}_t\big]\Big] \\
&= \mathbf{E}\Big[\tilde{\mathbf{E}}\big[F\big(\vartheta_{-t}\phi\big(\tilde{Z} + Z_t,\vartheta_t\tau_t\phi(Z,\mathbb{x})\big)\big)\big]\Big] \\
&= \mathbf{E}\Big[\tilde{\mathbf{E}}\big[F\big(\vartheta_{-t}\phi\big(\tilde{Z},\vartheta_t\phi(Z,\mathbb{x})\big)\big)\big]\Big] = [\mathbb{E}\mathbb{E}_tF](\mathbb{x}),
\end{aligned}$$

where $\tilde{Z}$ is an independent copy of Z. Here, in the fifth equality, we used that the mapping $\omega \mapsto \vartheta_t\tau_t\phi(Z(\omega),\mathbb{x})$ is $\mathscr{G}_t$-measurable due to condition (iv) and that $Z_{t+\cdot} - Z_t$ is independent of $\mathscr{G}_t$. The sixth equality is valid due to condition (iii).

To verify that the induced evolutionary semigroup is a C_b-semigroup, we use Theorem 5.8. If $(\mathbb{x}_n)_{n\in\mathbb{N}}$ is a sequence in $\mathscr{X}^-$ converging to $\mathbb{x} \in \mathscr{X}^-$, then by assumption, we have $\phi(\omega,\mathbb{x}_n) \to \phi(\omega,\mathbb{x})$. Given $F \in C_b(\mathscr{X})$, it follows from the dominated convergence theorem that

$$[\mathbb{E}F](\mathbb{x}_n) = \mathbf{E}[F(\phi(Z,\mathbb{x}_n))] \to \mathbf{E}[F(\phi(Z,\mathbb{x}))] = [\mathbb{E}F](\mathbb{x}). \qquad \square$$

As an application, we consider a stochastic delay equation, driven by additive Lévy noise. To that end, let

$$\mathscr{D}_h := D((-h,0);\mathbb{R}^d).$$

For fixed $\omega \in \Omega$ and $\xi \in \mathscr{D}_h$, we consider the (deterministic) delay equation

$$\begin{cases} dy(t) = b(y_t)dt + d\omega(t), & \text{for } t \geq 0, \\ \quad y_0 = \xi. \end{cases} \tag{6.11}$$

We interpret this as an integral equation and require the solution to be continuous at 0. Thus, (6.11) is equivalent to $y(t) = \xi(t)$ for $t \leq 0$ and

$$y(t) = \xi(0-) + \int_0^t b(y_s)\,ds + \omega(t) - \omega(0) \tag{6.12}$$

for all $t > 0$. If b is Lipschitz continuous, then, arguing similar as in [4, Theorem II.4.3.1], we see that (6.11) has a unique solution $y^\xi(\omega) \in D((-h,\infty);\mathbb{R}^d)$. Given $\omega \in \Omega$ and $\mathbb{x} \in \mathscr{X}$, we define

$$[\phi(\omega,\mathbb{x})](t) := \begin{cases} \mathbb{x}(t), & \text{if } t < 0, \\ y^{\mathbb{x}_0}, & \text{if } t \geq 0. \end{cases}$$

We claim that ϕ is a random evolution map in the sense of Definition 6.16. Here conditions (i) – (iv) can be directly verified. To prove (v), we fix $t > 0$ and set $\mathbb{y} = \vartheta_t\phi(\omega,\mathbb{x})$. Then we have $\mathbb{y}(s) = [\phi(\omega,\mathbb{x})](t+s)$. It follows that

$$\begin{aligned} \mathbb{y}(s) &= \mathbb{x}(0-) + \int_0^{t+s} b(\phi(\omega,\mathbb{x})_r)\,dr + \omega(t+s) - \omega(0) \\ &= \mathbb{x}(0-) + \int_0^t b(\phi(\omega,\mathbb{x})_r)\,dr + \int_t^{t+s} b(\phi(\omega,\mathbb{x})_r)\,dr + \omega(t+s) - \omega(0) \\ &= [\phi(\omega,\mathbb{x})](t-) + \int_0^s b(\phi(\omega,\mathbb{x})_{t+r})\,dr + \omega(t+s) - \omega(t) \\ &= \mathbb{y}(0-) + \int_0^s b(\mathbb{y}_r)\,dr + (\vartheta_t\omega)(s) - (\vartheta_t\omega)(0). \end{aligned} \tag{6.13}$$

By uniqueness, $\phi(\vartheta_t\omega,\mathbb{y}) = \mathbb{y}$, which is exactly (v).

Example 6.18. Let $X = \mathbb{R}^d$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ be a bounded, continuous function. For fixed $\omega \in \Omega$ and $\mathbb{x} \in \mathscr{X}$, consider the delay equation

$$dy(t) = b(y(t-1))dt + d\omega(t), \quad y_0 = \mathbb{x}_0. \tag{6.14}$$

Of course, we interpret this equation as an integral equation and require the solution to be continuous at 0 (see Remark 6.3). Thus, (6.14) is equivalent to $y(t) = \mathbb{x}(t)$ for $t \leq 0$ and

$$y(t) = \mathbb{x}(0-) + \int_0^t b(y(s-1))\,ds + \omega(t) - \omega(0)$$

for all $t > 0$.

This equation can be solved iteratively, yielding the following random evolution map:

$$\phi(\omega,\mathbb{x},t) = \begin{cases} \mathbb{x}(t), & \text{if } t < 0, \\ \mathbb{x}(0-) + \int_0^t b(\mathbb{x}(s-1))\,ds + \omega(t) - \omega(0), & \text{if } t \in [0,1), \\ \phi(\omega,\mathbb{x},1-) + \int_1^t b(\phi(\omega,\mathbb{x},s-1))\,ds + \omega(t) - \omega(0), & \text{if } t \in [1,2), \\ \dots \end{cases}$$

This is indeed a random evolution map. Conditions (i) – (iv) can be directly verified and for (v) one can argue as in Equation (6.13).

Now, let $\mathbf{P}$ be a measure on $(\Omega, \mathscr{G})$ such that the coordinate process $Z_t(\omega) := \omega(t)$ is a Lévy process starting at 0. We denote by $(S(t))_{t\geq 0}$ the transition semigroup of this process on $B_b(\mathbb{R}^d)$. It is well known, that S is a C_b-semigroup (it is actually *Feller*, in the sense that it restricts to a strongly continuous semigroup on $C_0(\mathbb{R}^d)$) and we denote its C_b-generator by A. Moreover, we denote the expectation operator associated to ϕ and $\mathbf{P}$ via Proposition 6.17 by $\mathbb{E}$ and the induced evolutionary semigroup by $\mathbb{T}$.

We also define the map $\varphi : \mathscr{X} \to \mathscr{X}$ by setting $\varphi(\varkappa, t) := \varkappa(t)$ for $t < 0$ and $\varphi(\varkappa, t) := \varkappa(0-) + \int_0^t b(\varkappa(s-1))\, ds$ for $t \in [0,1)$. Afterwards, we define φ recursively, setting $\varphi(\varkappa, t) := \varphi(\varkappa, n-) + \int_n^t b(\varphi(\varkappa, s-1))\, ds$ for $t \in [n, n+1)$.

It follows for $t \in (0,1)$ and $\varkappa \in \mathscr{X}$ that $\phi(Z, \varkappa, t) = \varphi(\varkappa, t) + Z_t$ $\mathbf{P}$-almost surely, so that for any $f \in B_b(\mathbb{R}^d)$,

$$\mathbb{T}(t)F_{0^\star}(f)(\varkappa) = \mathbf{E}[f(\phi(Z,\varkappa,t))] = \mathbf{E}f(\varphi(\varkappa,t) + Z_t) = (S(t)f)(\varphi(\varkappa,t)).$$

Note that $t \mapsto \varphi(\varkappa, t)$ is differentiable with derivative $b(\varphi(\varkappa, t-1))$ at any point where $\varkappa$ is continuous. If $f \in C_c^\infty(\mathbb{R}^d)$, then $t \mapsto S(t)f$ is differentiable with respect to $\|\cdot\|_\infty$ and the derivative is $S(t)Af$. With the chain rule, it follows that

$$\begin{aligned}\partial_t \mathbb{T}(t)F_{0^\star}(f)(\varkappa) &= S(t)Af(\varphi(\varkappa,t)) + S(t)\big[\nabla f(\varphi(\varkappa,t)) \cdot b(\varphi(\varkappa,t-1))\big] \\ &= \mathbb{T}(t)\Big[F_{0^\star}(Af) + \sum_{j=1}^d F_{0^\star}(\partial_j f) F_{-1}(b_j)\Big](\varkappa)\end{aligned}$$

for all but finitely many $t \in (0,1)$. Proceeding recursively, one can prove the same result for almost all $t > 0$. Now Proposition A.4 shows that

$$\big[F_{0^\star}(f), F_{0^\star}(Af) + \sum_{j=1}^d F_{0^\star}(\partial_j f)F_{-1}(b_j)\big] \in \mathbb{A}_{\mathsf{full}},$$

where $\mathbb{A}_{\mathsf{full}}$ denotes the full generator of $\mathbb{T}$.

Remark 6.19. The stochastic delay equation from Example 6.18 is a special case of the equations considered in [48], namely we here only consider additive Lévy noise whereas the authors of [48] also allow a multiplicative noise term depending on the past. In [48], the authors construct the transition semigroup of the solution on the state space $\mathscr{X}_{[-h,0]} = D([-h,0])$ of càdlàg paths on $[-h,0]$. As they discuss on page 1416, their semigroup is not a C_b-semigroup. In fact, it neither leaves the space $C_b(\mathscr{X}_{[-h,0]})$ invariant nor does it satisfy the continuity condition from Definition A.9. Inspecting their argument, one sees that the problem is that the function π_{-h}, which is continuous on $\mathscr{X}_{[-h,0]}$ because of the special role of the point $-h$, is mapped by the shift semigroup to the non-continous function π_{-h+t}. In our setting, this problem does not occur as π_{-h} is not continuous on $\mathscr{X}^-$. Indeed, Proposition 6.17 shows that we obtain a C_b-semigroup and using Lemma 4.8 one can easily see that it leaves $C_b(\mathscr{X}, \mathscr{F}([-h,0)))$ invariant.

Appendix A. Transition semigroups

In many cases, the concept of a strongly continuous semigroup, see for example [19], is too strong to study transition semigroups of Markovian processes. Over the years, several suggestions have been made for alternative semigroup theories that can be used instead but, so far, no single theory has prevailed. We mention weakly integrable semigroups [33], weakly continuous semigroups [9], π-semigroups [47], bi-continuous semigroups [36, 21] and sequentially equicontinuous semigroups [34]. Let us also mention [28, 29] where transition semigroups of certain stochastic processes were studied; here the so-called *strict* or *mixed* topology plays an important role.

In this article, we use a very flexible semigroup theory, that can be seen as a special case of the theory of 'semigroups on norming dual pairs' introduced in

[37, 38]. On the other hand, we have also opportunity to make use of the 'full generator' of a semigroup that goes back to the book of Ethier and Kurtz [20, Section 1.5]. As there is no single reference available where all needed results can be found, we introduce in this appendix the relevant notions and state the results that are used in the main part; there are also some new results.

Throughout this appendix, X is a Polish space and $\mathfrak{B}(X)$ denotes its Borel σ-algebra. The spaces of bounded real valued Borel measurable functions and bounded continuous functions on X are denoted by $B_b(X)$ and $C_b(X)$ respectively. Both are Banach spaces with respect to the supremum norm $\|\cdot\|_\infty$. The space of bounded signed measures is denoted by $\mathscr{M}(X)$. We note that $\mathscr{M}(X)$ can be isometrically identified with a closed subspace of the norm dual $B_b(X)^*$ and also a closed subspace of $C_b(X)^*$ by means of the identification

$$\langle f, \mu\rangle := \int_X f(x)d\mu(x).$$

A.1. **Kernel operators and bp-convergence.** A *kernel* on X is a map $k : X \times \mathfrak{B}(X) \to \mathbb{R}$ such that

(i) the map $x \mapsto k(x, A)$ is measurable for every $A \in \mathfrak{B}(X)$;
(ii) $k(x, \cdot)$ is a signed measure for every $x \in X$;
(iii) $\sup_x |k|(x, X) < \infty$ where $|k|(x, \cdot)$ refers to the total variation of $k(x, \cdot)$.

A kernel k is called *positive* (*Markovian*) if every measure $k(x, \cdot)$ is a positive measure (a probability measure).

Given a kernel k on X, we can define a bounded linear operator K on $B_b(X)$ by setting

$$[Kf](x) := \int_X f(y)k(x, dy) \quad \text{for } f \in B_b(X) \text{ and } x \in X. \tag{A.1}$$

We call an operator of this form a *kernel operator* and k the kernel *associated* to K and, conversely, K the operator associated to k.

It may happen, that in (A.1) the function Kf is continuous, whenever $f \in C_b(X)$. In fact, this is the case if and only if the map $x \mapsto k(x, \cdot)$ is continuous with respect to the weak topology $\sigma(C_b(X), \mathscr{M}(X))$. In this case, K defines a bounded linear operator on $C_b(X)$. We call such an operator *kernel operator on* $C_b(X)$. We note that any kernel operator on $C_b(X)$ can be extended to a kernel operator on $B_b(X)$ (by merely using (A.1) for general $f \in B_b(X)$ instead of $f \in C_b(X)$). We will not distinguish between kernel operators on $C_b(X)$ and their (unique) extension to $B_b(X)$.

We now treat the cases $C_b(X)$ and $B_b(X)$ simultaneously and let V denote either of these spaces. It turns out that a bounded linear operator on V is a kernel operator if and only if it is continuous with respect to the weak topology $\sigma := \sigma(V, \mathscr{M}(X))$, see e.g. [38, Proposition 3.5]. We write $\mathscr{L}(V, \sigma)$ for the space of σ-continuous linear operators on V. For sequences, σ-convergence is nothing else than *bp-convergence* (bp is short for *b*ounded and *p*ointwise). Indeed, by dominated convergence, bp-convergence implies σ-convergence and the converse follows by testing against Dirac measures and using the uniform boundedness principle.

Lemma A.1. *Let $V \in \{C_b(X), B_b(X)\}$ and $K \in \mathscr{L}(V)$. The following are equivalent:*

(i) *K is a kernel operator.*
(ii) *K is σ-continuous, i.e. $K \in \mathscr{L}(V, \sigma)$.*
(iii) *K is* bp-continuous, *i.e. if $(f_n)_{n\in\mathbb{N}} \subset V$ is a sequences that bp-converges to $f \in V$, then Kf_n bp-converges to Kf.*
(iv) *the norm-adjoint K^* leaves the space $\mathscr{M}(X)$ invariant.*

In this case, the σ-adjoint of K is given by $K' := K^*|_{\mathscr{M}(X)}$.

Proof. The equivalence of (ii) and (iv) (as well as the addendum) follow from [38, Proposition 3.1] applied to the norming dual pair $(V, \mathscr{M}(X))$. As already mentioned, the equivalence of (ii) and (i) is [38, Proposition 3.5]. For the equivalence of (i) and (iii), see [7, Lemma A.1] in the case $V = C_b(X)$ and [2, Lemma 5.1] in the case $V = B_b(X)$. □

Remark A.2. In the case $V = B_b(X)$, the statements of Lemma A.1 remain valid also for general measurable spaces.

Let us also briefly recall the notion of *bp-closure*, see [20, Section 3.4]. A subset $M \subset B_b(X)$ is called *bp-closed* if whenever $(f_n)_{n\in\mathbb{N}} \subset M$ bp-converges to $f \in B_b(X)$, it follows that $f \in M$. The *bp-closure* of a set $S \subset B_b(X)$ is the smallest bp-closed set that contains S. If the bp-closue of S is all of $B_b(X)$, we say that S is *bp-dense* in $B_b(X)$. By [20, Proposition 3.4.2], $C_b(X)$ is bp-dense in $B_b(X)$.

These notions help overcome a weakness of working with the weak topology σ. As $C_b(X)$ is σ-dense in $B_b(X)$, for every $f \in B_b(X)$ there is a net $(f_\alpha) \subset C_b(X)$ converging to f with respect to σ. However, there need not be a sequence $(f_n)_{n\in\mathbb{N}} \subset C_b(X)$ that bp-converges (i.e. σ-converges) to f. Using bp-density and bp-closures allows us to work only with sequences nevertheless.

A.2. **The full generator of a transition semigroup.**

Definition A.3. A *transition semigroup* on X is a family of Markovian kernel operators $T = (T(t))_{t\geq 0} \subset \mathscr{L}(B_b(X), \sigma)$ such that

(i) $T(t+s) = T(t)T(s)$ for all $t, s \geq 0$ and $T(0) = I$;
(ii) For every $f \in X$ the map $(t, x) \mapsto [T(t)f](x)$ is jointly measurable.

If the same is true for a family $T = (T(t))_{t\in\mathbb{R}} \subset \mathscr{L}(B_b(X), \sigma)$ with index set $\mathbb{R}$, we call T a *transition group*.

If T is a transition semigroup, then for every $f \in B_b(X)$ and $\mu \in \mathscr{M}(X)$ the map $t \mapsto \langle T(t)f, \mu\rangle$ is measurable and, given $\lambda > 0$, there exists an operator $R(\lambda) \in \mathscr{L}(B_b(X), \sigma)$ such that

$$\langle R(\lambda)f, \mu\rangle = \int_0^\infty e^{-\lambda t}\langle T(t)f, \mu\rangle dt \quad \text{for } f \in B_b(X) \text{ and } \mu \in \mathscr{M}(X). \tag{A.2}$$

This shows that, in the terms of [38, Definition 5.1], a transition semigroup is an integrable semigroup on the norming dual pair $(B_b(X), \mathscr{M}(X))$. By [38, Proposition 5.2], the family $(R(\lambda))_{\lambda>0}$ is a *pseudo-resolvent*, i.e. it satisfies the resolvent identity

$$R(\lambda) - R(\mu) = (\mu - \lambda)R(\lambda)R(\mu) \quad \text{for } \lambda, \mu > 0.$$

For more information on pseudo-resolvents we refer to [19, Section III.4.a].

In general, the operators $(R(\lambda))_{\operatorname{Re}\lambda>0}$ are not injective and hence are not the resolvent of a single-valued operator. However, there is a multi-valued operator L_{full} (called the *full generator*) such that $(\lambda - L_{\mathsf{full}})^{-1} = R(\lambda)$ for $\lambda > 0$, see [31, Appendix A] (also for more information about multi-valued operators). The *domain* of L_{full} is given by $D(L_{\mathsf{full}}) := \{f \in B_b(X) : (f, g) \in L_{\mathsf{full}} \text{ for some } g \in B_b(X)\}$.

Proposition A.4. *Let T be a transition semigroup with full generator L_{full} and $f, g \in B_b(X)$. The following are equivalent:*

(i) $(f, g) \in L_{\mathsf{full}}$.
(ii) *For all $t > 0$ and $x \in X$ it is*

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

(iii) *For all $x \in X$ the map $t \mapsto T(t)f(x)$ belongs to $W^{1,\infty}_{\mathrm{loc}}([0,\infty))$ and $\partial_t T(t)f(x) = T(t)g(x)$ for almost every t.*

Proof. The equivalence of (i) and (ii) is proved in [38, Proposition 5.7]. Note that this shows that our terminology is consistent with that used in the book of Ethier–Kurtz [20, Section 1.5] who use (ii) for the definition. In view of the fundamental theorem of calculus for Sobolev functions, (iii) is merely a reformulation of (ii). □

Lemma A.5. *Let T be a transition semigroup with full generator L_{full}. The following hold true:*

(a) *L_{full} is bp-closed, i.e. if $((f_n, g_n))_{n\in\mathbb{N}} \subset L_{\mathsf{full}}$ bp-converges to $(f,g) \in B_b(X) \times B_b(X)$, then $(f,g) \in L_{\mathsf{full}}$;*
(b) *If $D(L_{\mathsf{full}})$ is bp-dense in $B_b(X)$, then T is uniquely determined by L_{full}.*

Proof. (a). Let $((f_n, g_n))_{n\in\mathbb{N}} \subset L_{\mathsf{full}}$. By Proposition A.4, this is equivalent to

$$T(t)f_n(x) - f_n(x) = \int_0^t (T(s)g_n)(x)\, ds \tag{A.3}$$

for all $t > 0$ and $x \in X$. If (f_n, g_n) bp-converges to (f, g) then the bp-continuity of the operators $T(t)$ imply that the left hand side of (A.3) converges to $T(t)f(x)-f(x)$ and the right-hand side to $\int_0^t T(s)g(x)\, ds$. Using Proposition A.4 again, (a) follows.

(b). It follows from Proposition A.4 that $\|T(t)f - T(s)f\|_\infty \le |t-s|\|g\|_\infty$ for $(f,g) \in L_{\mathsf{full}}$ and $t, s > 0$. This implies that, for $f \in D(L_{\mathsf{full}})$ the orbit $t \mapsto T(t)f$ is $\|\cdot\|_\infty$-continuous and thus in particular pointwise continuous. Now let T_1 and T_2 be transition semigroups with the same full generator L_{full} and fix $f \in D(L_{\mathsf{full}})$. It follows that

$$\int_0^\infty e^{-\lambda t}(T_1(t)f)(x)\, dt = \int_0^\infty e^{-\lambda t}(T_2(t)f)(x)\, dt = (\lambda - L_{\mathsf{full}})^{-1}f(x)$$

for all $\lambda > 0$ and $x \in X$. By the uniqueness theorem for the Laplace transform [1, Theorem 1.7.3], $(T_1(t)f)(x) = (T_2(t)f)(x)$ for almost every $t > 0$. By continuity, $T_1(t)f(x) = T_2(t)f(x)$ for all $t > 0$. Thus, for every $t > 0$, the operators $T_1(t)$ and $T_2(t)$ coincide on the set $D(L_{\mathsf{full}})$. Now fix $t > 0$. As the operators $T_1(t), T_2(t)$ are bp-continuous, the set of f for which $T_1(t)f = T_2(t)f$ is bp-closed. Since $D(L_{\mathsf{full}})$ is bp-dense in $B_b(X)$, $T_1(t) = T_2(t)$ follows. □

Following [35], we say that a subset $C \subset L_{\mathsf{full}}$ is a *bp-core*, if L_{full} is the bp-closure of C. We next give a criterion for a bp-core.

Lemma A.6. *Let $(T(t))_{t\ge 0}$ be a transition semigroup with full generator L_{full}. Moreover, let $C \subset L_{\mathsf{full}}$ be a subspace with the following properties:*

(i) *If $(f,g) \in C$ then $(T(t)f, T(t)g) \in C$ for all $t > 0$;*
(ii) *The set $S := \{f \in B_b(X) : (f,g) \in C \text{ for some } g \in B_b(X)\}$ is bp-dense in $B_b(X)$;*
(iii) *For every $x \in X$ and $g \in B_b(X)$ such that $(f,g) \in C$ for some $f \in B_b(X)$, the map $t \mapsto T(t)g(x)$ is continuous at almost every point in $(0,\infty)$.*

Then C is a bp-core for L_{full}.

Proof. We denote the bp-closure of C by M. Since C is a vector space, so is M. Moreover, L_{full} is bp-closed by Lemma A.5(a), so $M \subset L_{\mathsf{full}}$. We thus only need to prove that $L_{\mathsf{full}} \subset M$.

Fix $(f,g) \in C$ and $t_0 > 0$. We claim that for every $\lambda > 0$

$$\Big(\int_0^{t_0} e^{-\lambda t}T(t)f\, dt, \int_0^{t_0} e^{-\lambda t}T(t)g\, dt\Big) \in M.$$

To see this, let $x \in X$ and define the functions $\varphi(t) := \mathbb{1}_{(0,t_0)}(t)e^{-\lambda t}[T(t)f](x)$ and $\varphi_n(t) := \sum_{k=1}^n \mathbb{1}_{[\frac{k-1}{n}t_0, \frac{k}{n}t_0)}(t)e^{-\lambda\frac{k-1}{n}}\big[T(\frac{k-1}{n}t_0)g\big](x)$ and the functions ψ and ψ_n similarly, replacing f with g. It is a consequence of condition (i) and the fact that M is a vector space that $(\int_0^{t_0} \varphi_n(t)\, dt, \int_0^{t_0} \psi_n(t)\, dt) \in M$. Since $f \in D(L_{\mathsf{full}})$, the orbit $t \mapsto T(t)f$ is $\|\cdot\|$-continuous (see the proof of Lemma A.5(b)) and hence pointwise continuous. It follows that φ is Riemann-integrable and $\int_0^{t_0} \varphi_n(t)\,dt$, which is nothing else than Riemannian sums for the integral of φ converge to $\int_0^{t_0} \varphi(t)\,dt$. It is a consequence of condition (iii), that also the function ψ is Riemann integrable and thus the integrals of ψ_n converge to that of ψ, proving the claim.

Next note that since $(f, g) \in L_{\mathsf{full}}$, we have

$$\int_0^{t_0} e^{-\lambda t} T(t) g\,dt = \lambda \int_0^{t_0} e^{-\lambda t} T(t) f\,dt - f + e^{-\lambda t_0} T(t_0) f,$$

see, e.g., [38, Proof of Proposition 5.7]. For $t_0 \to \infty$, the last term tends (in norm) to 0 and it follows that

$$(R(\lambda, L_{\mathsf{full}})f, \lambda R(\lambda, L_{\mathsf{full}})f - f) \in M$$

for every $f \in S$. Since S is bp-dense in $B_b(X)$ by (ii), the same is true for arbitrary $f \in B_b(X)$. Now let $(f_0, g_0) \in L_{\mathsf{full}}$. Note that $R(\lambda, L_{\mathsf{full}})(\lambda f_0 - g_0) = f_0$. Thus, applying the above to $f = \lambda f_0 - g_0$, it follows that $(f_0, g_0) \in M$, finishing the proof. ☐

Sometimes, it is more convenient to work with single-valued operators. To that end, we introduce the following terminology:

Definition A.7. Let $\mathscr{L}$ be a (potentially) multi-valued operator on $B_b(X)$. A *slice of* $\mathscr{L}$ is a single-valued operator $L_0 : D(L_0) \to B_b(X)$ such that $(f, L_0 f) \in \mathscr{L}$ for every $f \in D(L)$.

From Lemma A.6, we immediately obtain:

Corollary A.8. *Let $(T(t))_{t \geq 0}$ be a transition semigroup with full generator L_{full} and L_0 be a slice of L_{full} such that:*

(i) *For every $t > 0$ and $f \in D(L)$ it is $T(t)f \in D(L_0)$ and $L_0 T(t) f = T(t) L_0 f$;*
(ii) *$D(L_0)$ is bp-dense in $B_b(X)$;*
(iii) *for every $f \in D(L_0)$and $x \in X$ the map $t \mapsto T(t) L_0 f(x)$ is continuous at almost every point in $(0, \infty)$.*

Then $\{(f, L_0 f) : f \in D(L_0)\}$ is a bp-core for L_{full}. By slight abuse of notation, we will say that $D(L_0)$ is a bp-core *for L_{full}.*

A.3. **C_b-semigroups and the C_b-generator.**

Definition A.9. A transition semigroup T is called *C_b-semigroup* if

(i) for every $t \geq 0$ the operator $T(t)$ leaves the space $C_b(X)$ invariant;
(ii) for every $f \in C_b(X)$ the map $(t, x) \mapsto [T(t)f](x)$ is continuous on $[0, \infty) \times X$.

A transition group T is called *C_b-group* if the above hold true for every $t \in \mathbb{R}$.

Note that the continuity assumption in Definition A.9 implies via a bp-closedness argument the measurability assumption in Definition A.3.

We can give an equivalent description of C_b-semigroups using an additional locally convex topology on $C_b(X)$, the so-called *strict topology* β_0. This topology is defined as follows: We denote by $\mathscr{F}_0(X)$ the space of all functions $\varphi : X \to \mathbb{R}$ that *vanish at infinity*, i.e. given $\varepsilon > 0$ we find a compact subset K of X such that $|\varphi(x)| \leq \varepsilon$ for all $x \in X \setminus K$. The strict topology β_0 is generated by the seminorms $\{p_\varphi : \varphi \in \mathscr{F}_0(X)\}$, where $p_\varphi(f) = \|\varphi f\|_\infty$. This topology is consistent with the duality $(C_b(X), \mathscr{M}(X))$, i.e. the dual space $(C_b(X), \beta_0)'$ is $\mathscr{M}(X)$, see [32, Theorem 7.6.3]. In fact, it is the Mackey topology of the dual pair $(C_b(X), \mathscr{M}(X))$, i.e. the finest locally convex topology on $C_b(X)$ which yields $\mathscr{M}(X)$ as a dual space. This is a consequence of Theorems 4.5 and 5.8 of [50]. Consequently, a kernel operator on $C_b(X)$ (i.e. a σ-continuous operator by Lemma A.1) is automatically β_0-continuous.

By [32, Theorem 2.10.4] β_0 coincides on $\|\cdot\|_\infty$-bounded sets with the topology of uniform convergence on compact subsets of X.

We recall the following characterization from [37, Theorem 4.4].

Theorem A.10. *Let T be a transition semigroup such that every operator $T(t)$ leaves $C_b(X)$ invariant. The following are equivalent:*

(i) *T is a C_b-semigroup, i.e. the map $(t,x) \mapsto [T(t)f](x)$ is continuous on $[0,\infty)\times X$ for every $f \in C_b(X)$.*
(ii) *T is β_0-continuous on $C_b(X)$, i.e. for every $f \in C_b(X)$, the orbit $t \mapsto T(t)f$ is β_0-continuous on $[0,\infty)$.*
(iii) *T is β_0-continuous on $C_b(X)$ and locally β_0-equicontinuous.*

We note that many transition semigroups are in fact C_b-semigroups, see, e.g., the references [29, 2, 35, 7] for concrete examples.

Using the dominated convergence theorem, it is easy to see that if T is a transition semigroup such that every $T(t)$ leaves the space $C_b(X)$ invariant, then also its Laplace transform $R(\lambda)$ leaves the space $C_b(X)$ invariant. If additionally the continuity assumption in Definition A.9(ii) is satisfied, $R(\lambda)|_{C_b(X)}$ is injective and thus the resolvent of a unique (single-valued) operator.

Definition A.11. Let T be a C_b-semigroup and $(R(\lambda))_{\operatorname{Re}\lambda>0}$ be its Laplace transform. The *C_b-generator* of T is the unique single-valued operator L such that $(\lambda - L)^{-1} = R(\lambda)|_{C_b(X)}$ for all $\operatorname{Re}\lambda > 0$.

As the resolvent of L is merely the restriction of that of L_{full} to $C_b(X)$, the C_b-generator L is the *part* of L_{full} in $C_b(X)$, i.e. $f \in D(L)$ and $Lf = g$ if and only if $(f,g) \in L_{\mathsf{full}}$ and $f, g \in C_b(X)$. We give some alternative characterizations of the C_b-generator.

Theorem A.12. *Let T be a C_b-semigroup with C_b-generator L. For $f, g \in C_b(X)$ the following are equivalent:*

(i) *$f \in D(L)$ and $Af = g$;*
(ii) *$t^{-1}(T(t)f - f) \to g$ with respect to σ as $t \to 0$;*
(iii) *$t^{-1}(T(t)f - f) \to g$ with respect to β_0 as $t \to 0$;*
(iv) *$\sup_{t\in(0,1)} \|t^{-1}(T(t)f - f)\|_\infty < \infty$ and $t^{-1}(T(t)f(x) - f(x)) \to g(x)$ for all $x \in X$.*

Proof. See [7, Theorem A.5]. □

The full generator can be reconstructed from the C_b-generator. In fact, we have:

Corollary A.13. *Let $(T(t))_{t\geq 0}$ be a C_b-semigroup with C_b-generator L and full generator L_{full}. Then $D(L)$ is a bp-core for L_{full}.*

Proof. If $f \in D(L)$ and $x \in X$, then $Lf \in C_b(X)$ and the map $t \mapsto [T(t)Lf](x)$ is continuous on all of $(0,\infty)$ proving (iii) in Corollary A.8. Condition (i) can easily be obtained from [38, Proposition 5.7]. It remains to prove condition (ii). To that end, note that for every $f \in C_b(X)$, it holds $\lambda R(\lambda, L)f \in D(L)$ and $\lambda R(\lambda, L)f \to f$ pointwise as $\lambda \to \infty$, see [37, Theorem 2.10]. This proves that the bp-closure of $D(L)$ contains $C_b(X)$ and hence $B_b(X)$. Now Corollary A.8 yields the claim. □

Note that the concept of bp-core is only appropriate on the space $B_b(X)$. We also introduce a suitable concept of a core for the C_b-generator. If T is a C_b-semigroup with C_b-generator L and $D \subset D(L)$, we say that D is a *β_0-core* for L if for every $f \in D(L)$ there is a net $(f_\alpha) \subset D$ such that $f_\alpha \to f$ and $Lf_\alpha \to Lf$ with respect to β_0.

Lemma A.14. *Let T be a C_b-semigroup with C_b-generator L and $D \subset D(L)$ be such that*

(i) *D is dense in $C_b(X)$ with respect to β_0 and*
(ii) *$T(t)D \subset D$ for all $t > 0$.*

Then D is a β_0-core for L.

Proof. This follows along the lines of [39, Proposition 2.3] which is concerned with the full generator. □

Appendix B. Examples of path spaces

B.1. **Continuous paths.** Let (X, d) be a separable, complete metric space and put $\mathscr{X}_\mathrm{C} := C(\mathbb{R}; X)$, the set of all continuous functions from $\mathbb{R}$ to X, and

$$\mathbb{d}_\mathrm{C}(\mathbb{x}, \mathbb{y}) := \sum_{n=1}^{\infty} 2^{-n}\big[1 \wedge \sup_{t\in[-n,n]} d(\mathbb{x}(t), \mathbb{y}(t))\big].$$

We also set

$$[\tau_\mathrm{C}(\mathbb{x})](t) := \begin{cases} \mathbb{x}(t), & \text{if } t < 0, \\ \mathbb{x}(0), & \text{if } t \geq 0. \end{cases}$$

Proposition B.1. *The pair $((\mathscr{X}_\mathrm{C}, \mathbb{d}_\mathrm{C}), \tau_\mathrm{C})$ is a path space. Moreover, in this case the map τ_C and every evaluation map π_t, for $t \in \mathbb{R}$, is continuous.*

Proof. The evaluation maps π_t are clearly continuous from $\mathscr{X}_\mathrm{C}$ to X and it is well known that they generate the Borel σ-algebra, proving (P1). We point out that, by the continuity of the paths, we have $\mathscr{F}_t = \mathscr{F}_{t-}$ in this case. In what follows, we work with $\mathscr{F}_t$ instead of $\mathscr{F}_{t-}$ whenever $\mathscr{X} = \mathscr{X}_\mathrm{C}$.

Note that τ_C is continuous and thus certainly measurable. Moreover, as $\mathscr{X}_\mathrm{C}^-$ is closed in $\mathscr{X}_\mathrm{C}$, it is complete with respect to the original metric $\mathbb{d}_\mathrm{C}$. To finish the proof of (P2), it remains to prove the measurability requirement. To that end, let $t_1 < t_2 < \ldots < t_k \leq 0 < t_{k+1} < \ldots < t_n$ and $A_j \in \mathfrak{B}(X)$ for $j = 1, \ldots, n$. Consider the cylinder set

$$A = \{\mathbb{x} : \mathbb{x}(t_j) \in A_j \text{ for } j = 1, \ldots, n\} \in \mathfrak{B}(\mathscr{X}_\mathrm{C}).$$

Then

$$(\tau_\mathrm{C})^{-1}(A) = \Big\{\mathbb{x} : \mathbb{x}(t_j) \in A_j \text{ for } j = 1, \ldots, k \text{ and } \mathbb{x}(0) \in \bigcap_{j=k+1}^{n} A_j\Big\} \in \mathscr{F}_0.$$

As the cylinder sets generate $\mathfrak{B}(\mathscr{X}_\mathrm{C})$, it follows that τ_C is $\mathscr{F}_0$-measurable. Note that in the above situation, the cylinder set A belongs to $\mathscr{F}_0$ if and only if $k = n$. In that case $\tau_\mathrm{C}^{-1}(A) = A$. Once again, cylinder sets of this form generate $\mathscr{F}_0$ so that $(\tau_\mathrm{C})^{-1}(A) = A$ for all $A \in \mathscr{F}_0$, proving (P2).

To prove (P3), let $t_n \to t_\infty$ and $\mathbb{x}_n \to \mathbb{x}_\infty$. Given $\varepsilon > 0$, pick N_0 so large that $\sum_{n\geq N_0} 2^{-n} \leq \varepsilon$ and then N_1 so large that $t + t_n \in [-N_1, N_1]$ for all $t \in [-N_0, N_0]$ and $n \in \mathbb{N} \cup \{\infty\}$.

As the set $\{\mathbb{x}_n : n \in \mathbb{N} \cup \{\infty\}\}$ is compact, it follows from the Arzelà–Ascoli Theorem that we find $\delta > 0$ such that $d(\mathbb{x}_n(t), \mathbb{x}_n(s)) \leq \varepsilon$ for all $n \in \mathbb{N} \cup \{\infty\}$ and $t, s \in [-N_1, N_1]$ with $|t - s| \leq \delta$. Next pick n_0 so large, that $|t_n - t_\infty| \leq \delta$ and $\sup_{t\in[-N_1,N_1]} d(\mathbb{x}_n(t), \mathbb{x}_\infty(t)) \leq \varepsilon$ for all $n \geq n_0$. Then for $n \geq n_0$

$$\begin{aligned} d((\vartheta_{t_n}\mathbb{x}_n)(t), &(\vartheta_{t_\infty}\mathbb{x}_\infty)(t)) = d(\mathbb{x}_n(t + t_n), \mathbb{x}_\infty(t + t_\infty)) \\ &\leq d(\mathbb{x}_n(t + t_n), \mathbb{x}_n(t + t_\infty)) + d(\mathbb{x}_n(t + t_\infty), \mathbb{x}_\infty(t + t_\infty)) \leq 2\varepsilon \end{aligned}$$

for all $t \in [-N_0, N_0]$. Altogether, it follows that $\mathbb{d}_\mathrm{C}(\vartheta_{t_n}\mathbb{x}_n, \vartheta_{t_\infty}\mathbb{x}_\infty) \leq 3\varepsilon$ for $n \geq n_0$. This proves that $(t, \mathbb{x}) \mapsto \vartheta_t\mathbb{x}$ is continuous and thus (P3). □

B.2. **Càdlàg paths in the J_1-topology.** Let (X, d) be a complete, separable metric space. Without loss of generality, we assume that $d(x, y) \le 1$ for all $x, y \in X$. We consider the set $\mathscr{X}_{\mathrm{D}} = D(\mathbb{R}; X)$ of all *càdlàg functions* $\mathbb{x} : \mathbb{R} \to X$, i.e. for every $t \in \mathbb{R}$, we have

$$\mathbb{x}(t) = \lim_{s \downarrow t} \mathbb{x}(s) \quad \text{and} \quad \mathbb{x}(t-) := \lim_{s \uparrow t} \mathbb{x}(s) \text{ exists.}$$

As is well known (see e.g. [20, Lemma 3.5.1]), any $\mathbb{x} \in \mathscr{X}_{\mathrm{D}}$ has at most countably many discontinuities. There are several ways to extend Skorohod's J_1-topology, originally defined for càdlàg functions on $[0, 1]$, to unbounded time intervals. Billingsley [5] and Lindvall [40] use, similar to the case of continuous functions, a series over the distance of restrictions of the functions to compact time intervals. We note that in the J_1-topology, convergence of a càdlàg function on a compact time interval $[a, b]$ implies convergence of the values at the endpoints a and b. In the approach of Billingsley and Lindvall, the restrictions to the compact time intervals have to be slightly modified so that convergence with respect to the metric on the unbounded time interval does not entail convergence of the values in these (technically chosen) endpoints.

On the other hand, Whitt [51] and Ethier and Kurtz [20] use an integral instead of a series and no such alteration is needed. We will here follow this second approach. We should also point out that [51] is one of the few references where also the time interval $\mathbb{R}$ (instead of $[0, \infty)$) is considered.

Given $-\infty < a < b < \infty$, we denote by Λ_a^b the collection of all strictly increasing, bijective and Lipschitz continuous functions λ from $[a, b]$ to $[a, b]$ so that

$$\|\lambda\|_L := \sup_{a \le s < t \le b} \left| \log \frac{\lambda(t) - \lambda(s)}{t - s} \right| < \infty.$$

On $D([a, b], X)$, the space of X-valued càdlàg functions on the interval $[a, b]$, we define the metric

$$\mathbb{d}_a^b(\mathbb{x}, \mathbb{y}) = \inf_{\lambda \in \Lambda_a^b} \big[\|\lambda\|_L \vee \sup_{t \in [a,b]} d(\mathbb{x}(t), \mathbb{y}(\lambda(t))) \big\}. \tag{B.1}$$

This is a complete metric that induces Skorohod's J_1-topology on $D([a, b], X)$, see [5, Sect. 12]. We note that $\mathbb{d}_a^b(\mathbb{x}, \mathbb{y}) \le 1$, as $d(x, y) \le 1$ for all $x, y \in X$ and the choice $\lambda(t) = t$ yields $\|\lambda\|_L = 0$.

Given $\mathbb{x}, \mathbb{y} \in D(\mathbb{R}; X)$, we denote their restrictions to $[a, b]$ by $\mathbb{x}|_a^b$ and $\mathbb{y}|_a^b$ respectively and define

$$\mathbb{d}_{J_1}(\mathbb{x}, \mathbb{y}) := \int_{-\infty}^0 \int_0^\infty e^{s-t} \mathbb{d}_s^t(\mathbb{x}|_s^t, \mathbb{y}|_s^t)\, dt\, ds.$$

This integral is well defined by [51, Lemma 2.4]. By [51, Theorem 2.6] $(\mathscr{X}_{\mathrm{D}}, \mathbb{d}_{\mathrm{D}})$ is a complete separable metric space (see [20, Theorem 3.5.6]) which induces Skorohod's J_1-topology on $D(\mathbb{R}; X)$ in the sense that $\mathbb{x}_n \to \mathbb{x}$ in $(\mathscr{X}_{\mathrm{D}}, \mathbb{d}_{\mathrm{D}})$ if and only if $\mathbb{x}_n|_a^b \to \mathbb{x}|_a^b$ in $D([a, b], X)$ whenever a and b are continuity points of $\mathbb{x}$, see [51, Theorem 2.5].

To simplify notation, we will not distinguish between a function $\mathbb{x} \in \mathscr{X}_{\mathrm{D}}$ and its restriction to some compact interval $[a, b]$ in what follows. It will be clear from the context what the domain of definition is.

On $\mathscr{X}_{\mathrm{D}}$, we define the stopping map τ_{D} by

$$[\tau_{\mathrm{D}}(\mathbb{x})](t) = \begin{cases} \mathbb{x}(t), & \text{if } t < 0, \\ \mathbb{x}(0-), & \text{if } t \ge 0. \end{cases}$$

To prove that also the càdlàg functions form a path space, we make use of the following '*modulus of continuity*':

$$\omega'(\mathbb{x}, \delta, T) := \inf_{(t_0,\dots,t_n)} \max_{j=1,\dots,n} \sup_{s,t\in[t_{j-1},t_j)} d(\mathbb{x}(s), \mathbb{x}(t)),$$

where the infimum is taken over all partitions $-T = t_0 < t_1 < \ldots < t_n = T$ satisfying $\min_j(t_j - t_{j-1}) \geq \delta$. Similar to the Arzelà–Ascoli theorem, this modulus of continuity can be used to characterize compact subsets of $\mathscr{X}_{\mathrm{D}}$ (see [20, Theorem 3.6.3]).

In order to prove that $((\mathscr{X}_{\mathrm{D}}, \mathbb{d}_{J_1}), \tau_{\mathrm{D}})$ is a path space, we first establish a Lemma.

Lemma B.2. *Let $\mathbb{x} \in D(\mathbb{R}; X)$ be a piecewise constant function with finitely many jumps, i.e.*

$$\mathbb{x}(t) = x_0 \mathbb{1}_{(-\infty,r_0)}(t) + \sum_{j=1}^{n} x_j \mathbb{1}_{[r_{j-1},r_j)}(t) + x_{n+1}\mathbb{1}_{[r_n,\infty)}(t),$$

where $-\infty < r_0 < r_1 < \ldots < r_n < \infty$ satisfies $\min_j(r_j - r_{j-1}) \geq \delta > 0$ and $x_0, \ldots, x_{n+1} \in X$. For $|t| < \delta$ we put

$$f_\delta(t) := \max\Big\{ \log\frac{\delta}{\delta - |t|}, \log\Big(1 + \frac{|t|}{\delta}\Big)\Big\}.$$

Then for $|t| < \delta/2$ we have $\mathbb{d}_{J_1}(\mathbb{x}, \vartheta_t\mathbb{x}) \leq f_\delta(t)$. In particular, $\vartheta_t\mathbb{x} \to \mathbb{x}$ as $t \to 0$.

Proof. Fix real numbers $a < b$ and put

$$\{t_0, \ldots, t_k\} := (\{r_0, \ldots, r_n\} \cap [a, b]) \cup \{a, b\},$$

so that $a = t_0 < t_1 < \cdots < t_k = b$. Now, let $|t| < \delta/2$. Then, for every $j = 2, \ldots, k-2$ we have $t_j - t \in (t_{j-1}, t_{j+1})$ but if $t > 0$ it is possible that $t_1 - t < t_0$ and in the case where $t < 0$ it is possible that $t_{k-1} - t > t_k$. Define $s_0 < \ldots < s_k$ by setting $s_0 = t_0 = a$ and $s_k = t_k = b$ and

$$s_j = \begin{cases} t_1 - t, & \text{if } j = 1 \text{ and } t_1 - t > t_0,\\ t_1, & \text{if } j = 1 \text{ and } t_1 - t \leq t_0,\\ t_j - t, & \text{if } j = 2, \ldots, k-2,\\ t_{k-1} - t, & \text{if } j = k-1 \text{ and } t_{k-1} - t < t_k,\\ t_{k-1}, & \text{if } j = 1 \text{ and } t_{t-1} - t \geq t_k. \end{cases}$$

Now define $\lambda_t \in \Lambda_a^b$ in such a way, that $\lambda_t(s_j) = t_j$ and λ_t is piecewise linear, i.e.

$$\lambda_t(s) = t_j + \frac{t_j - t_{j-1}}{s_j - s_{j-1}}(s - s_{j-1}) = t_j + d_j(t)(s - s_{j-1})$$

for $s \in [s_{j-1}, s_j]$ We note that $\lambda_t'(s) = d_j(t)$ for $s \in (s_{j-1}, s_j)$ and $d_j(t) \equiv 1$ for $j = 2, \ldots, k-2$. Also in cases $j = 1$ and $j = k-1$ it is possible that $d_j(t) = 1$ and this is the case when $s_j = t_j$. However, a brief computation shows that in any case $|\log d_j(t)| \leq f_\delta(t)$ for $j = 1$ and $k-1$ so that $\|\lambda\|_L \leq f_\delta(t)$.

Noting that $\vartheta_t \mathbb{1}_{[t_{j-1},t_j)} = \mathbb{1}_{[t_{j-1}-t,t_j-t)}$ one sees that $(\vartheta_t\mathbb{x})(\lambda_t(s)) = \mathbb{x}(s)$ for all $s \in [a, b]$ so that $\mathbb{d}_a^b(\vartheta_t\mathbb{x}, \mathbb{x}) \leq \|\lambda_t\|_L \leq f_\delta(t)$. As a, b were arbitrary, the claim follows from the definition of $\mathbb{d}_{J_1}$. ☐

We are now ready to prove:

Proposition B.3. *The pair $((\mathscr{X}_{\mathrm{D}}, \mathbb{d}_{J_1}), \tau_{\mathrm{D}})$ is a path space.*

Proof. If $\mathbb{x}_n \to \mathbb{x}$, then $\pi_t(\mathbb{x}_n) \to \pi_t(\mathbb{x})$ whenever t is a continuity point of $\mathbb{x}$ ([20, Proposition 3.5.2]). The fact that the Borel σ-algebra is generated by the evaluation maps π_t follows from [20, Proposition 3.7.1]. The measurability requirements in (P2) can be proved in the same way as in Proposition B.1. To see that $\mathscr{X}_{\rm D}^-$ is Polish, define for $\mathbb{x}, \mathbb{y} \in \mathscr{X}_{\rm D}^-$

$$\mathbb{d}_{J_1}^-(\mathbb{x},\mathbb{y}) := \int_0^\infty e^{-t}\mathbb{d}_{-t}^0(\mathbb{x},\mathbb{y})\,dt.$$

As every $\mathbb{x} \in \mathscr{X}_{\rm D}^-$ is continuous at 0, it follows from the results of [51, Sect. 2] already used above that $\mathbb{d}_{J_1}^-$ defines a complete metric on $\mathscr{X}_{\rm D}^-$ such that $\mathbb{x}_n \to \mathbb{x}$ if and only if the restrictions to $[a,0]$ converge in $D([a,0],X)$ for every $a<0$ that is a continuity point of $\mathbb{x}$. Noting that every element of $\mathscr{X}_{\rm D}^-$ is constant on $[0,b]$ for every $b>0$ it is easy to see that $\mathbb{x}_n \to \mathbb{x}$ in $D([a,0],X)$ if and only if $\mathbb{x}_n \to \mathbb{x}$ in $D([a,b],X)$ for all $b>0$; to see that convergence in $D([a,b],X)$ implies convergence in $D([a,0];X)$ use the continuity of $\mathbb{x}$ in 0 and [51, Lemma 2.2]. This proves that a sequence $(\mathbb{x}_n) \subset \mathscr{X}_{\rm D}^-$ converges to some $\mathbb{x} \in \mathscr{X}_{\rm D}^-$ with respect to $\mathbb{d}_{J_1}^-$ if and only if it converges with respect $\mathbb{d}_{J_1}$. Consequently, $\mathbb{d}_{J_1}^-$ induces the same topology on $\mathscr{X}_{\rm D}^-$ as $\mathbb{d}_{J_1}$ the proof of (P2) is finished.

To prove (P3), we show that for any compact subset K of $\mathscr{X}_{\rm D}$ it is $\vartheta_t\mathbb{x} \to \vartheta_s\mathbb{x}$ as $t \to s$ uniformly for $\mathbb{x} \in K$. We may assume without loss of generality that $s=0$. Given $\varepsilon>0$, pick $T>0$ such that $\int_{T-1}^\infty e^{-t}\,dt \le \varepsilon$. As K is compact, there is $\delta>0$ such that $\omega'(\mathbb{x},\delta,T) \le \varepsilon$ for all $\mathbb{x} \in K$, see [20, Theorem 3.6.3]. For fixed $\mathbb{x} \in K$ we find a partition $\pi = \pi(\mathbb{x})$ of $[-T,T]$, say $\pi = (-T = t_0 < t_1 < \ldots < t_n = T)$, with $\min_j(t_j - t_{j-1}) \ge \delta$ and $\max_j \sup_{t,s\in[t_{j-1},t_j)} d(\mathbb{x}(t),\mathbb{x}(s)) \le 2\varepsilon$. Define $\mathbb{x}_\pi$ by setting

$$\mathbb{x}_\pi(s) = \mathbb{x}(t_0-)\mathbb{1}_{(-\infty,t_0)}(s) + \sum_{j=1}^n \mathbb{x}(t_{j-1})\mathbb{1}_{[t_{j-1},t_j)}(s) + \mathbb{x}(t_n)\mathbb{1}_{[t_n,\infty)}.$$

Then $d(\mathbb{x}_\pi(s),\mathbb{x}(s)) \le 2\varepsilon$ for every $s \in [-(T-1),T-1]$ and, using $\lambda(s)=s$, it follows that $\mathbb{d}_a^b(\mathbb{x}_\pi(s),\mathbb{x}(s)) \le 2\varepsilon$ for all $-(T-1) \le a < 0 < b \le T-1$ and thus altogether $\mathbb{d}_{J_1}(\mathbb{x}_\pi,\mathbb{x}) \le 2\varepsilon$.

Basically the same calculation shows $\mathbb{d}_{J_1}(\vartheta_t\mathbb{x}_\pi,\vartheta_t\mathbb{x}) \le 2\varepsilon$ whenever $|t| \le 1$. Taking Lemma B.2 into account, it follows that

$$\mathbb{d}_{J_1}(\vartheta_t\mathbb{x},\mathbb{x}) \le \mathbb{d}_{J_1}(\vartheta_t\mathbb{x},\vartheta_t\mathbb{x}_\pi) + \mathbb{d}_{J_1}(\vartheta_t\mathbb{x}_\pi,\mathbb{x}_\pi) + \mathbb{d}_{J_1}(\mathbb{x}_\pi,\mathbb{x}) \le 6\varepsilon + f_\delta(t)$$

for all $|t| \le \delta/2$. Picking $|t|$ so small that $f_\delta(t) \le \varepsilon$ (note that this only depends on δ and may thus be done indepently of the particular $\mathbb{x} \in K$), $\mathbb{d}_{J_1}(\vartheta_t\mathbb{x},\mathbb{x}) \le 7\varepsilon$. As $\varepsilon>0$ was arbitrary, (P3) is proved. □

B.3. **Càdlàg paths as proper path space.** We now verify that $\mathscr{X}_{\rm D}$, the space of all càdlàg paths endowed with the J_1-topology is indeed a proper path space so that the results of Section 5.2 apply. We start with a Lemma.

Lemma B.4. *Let $K \subset \mathscr{X}_{\rm D}^-$ be compact. Then, for any $\varepsilon>0$ there exists $\delta>0$ such that $d(\mathbb{x}(t),\mathbb{x}(0)) \le \varepsilon$ for all $t \in [-\delta,0]$ and $\mathbb{x} \in K$.*

Proof. If the conclusion is wrong, we find $\varepsilon_0>0$, a sequence $(\mathbb{x}_n) \subset K$ and a sequence $t_n \uparrow 0$ with $d(\mathbb{x}_n(t_n),\mathbb{x}(0)) \ge \varepsilon_0$ for all $n \in \mathbb{N}$. As K is compact, passing to a subsequence, we may and shall assume that $\mathbb{x}_n$ converges to some $\mathbb{x} \in K$. In particular, $\mathbb{x}$ is continuous at 0 so that $\mathbb{x}_n(0) \to \mathbb{x}(0)$ by [20, Proposition 3.5.2]. But then [20, Proposition 3.6.5] implies $\mathbb{x}_n(t_n) \to \mathbb{x}(0)$ which yields $\mathbb{x}_n(t_n) \to \mathbb{x}(0)$. This is a contradiction. □

Proposition B.5. *The space $((\mathscr{X}_{\rm D},\mathbb{d}_{J_1}),\tau_{\rm D})$ is a proper path space.*

Proof. Let $\mathbb{x}_n \to \mathbb{x}$. By (P3) $\vartheta_t\mathbb{x}_n \to \vartheta_t\mathbb{x}$ for all $t \in \mathbb{R}$. Note that $\vartheta_t\mathbb{x}$ is continuous at 0 for almost all $t \in \mathbb{R}$. Using [20, Proposition 3.6.5], it is easy to see that $\tau(\vartheta_t\mathbb{x}_n) \to \tau(\vartheta_{\mathbb{x}})$ whenever $\vartheta_t\mathbb{x}$ is continuous at 0. This proves (P4).

The proof of (P5) is similar to that of (P3). Given $\varepsilon>0$ and $K \subset \mathscr{X}_{\rm D}^-$, Lemma B.4 allows us to pick $\delta>0$ so that $d(\mathbb{x}(t),\mathbb{x}(0)) \le \varepsilon$ for all $t \in [-\delta,0]$ and $\mathbb{x} \in K$.

Next, choose $T > 0$ so large that $\int_{T-1}^{\infty} e^{-t}\,dt \le \varepsilon$. Picking a smaller δ if necessary, we may and shall assume that $\omega'(\mathbb{x}, 2\delta, T) \le \varepsilon$ for all $\mathbb{x} \in K$.

Now, fix $\mathbb{x} \in K$ and pick a partition $\pi = (t_0 < \ldots < t_n)$ of $[-T, T]$ such that $\min_j(t_j - t_{j-1}) \ge 2\delta$ and $\max_j \sup_{t,s\in[t_{j-1},t_j)} d(\mathbb{x}(t), \mathbb{x}(s)) \le 2\varepsilon$. As $\mathbb{x}(t) = \mathbb{x}(0) = \mathbb{x}(0-)$ for all $\mathbb{x} \in \mathscr{X}_{\mathrm{D}}^{-}$ we may and shall assume that $t_{n-1} \le 0$ so that the only partition point larger than 0 is $t_n = T$. Next, π is modified to $\tilde{\pi}$ as follows. If $|t_{n-1}| < \delta$, we replace t_{n-1} by $\tilde{t}_{n-1} := -\delta$ whereas all other partition points are unchanged. This results in a partition $\tilde{\pi}$ satisfying $\min_j(\tilde{t}_j - \tilde{t}_{j-1}) \ge \delta$ and $\max_j \sup_{t,s\in[\tilde{t}_{j-1},\tilde{t}_j)} d(\mathbb{x}(t), \mathbb{x}(s)) \le 2\varepsilon$. Repeating the arguments from above, $\mathbb{d}_{\mathrm{D}}(\mathbb{x}_{\tilde{\pi}}, \mathbb{x}) \le 3\varepsilon$.

Now let $0 \le t \le \delta/2$. Note that $(\vartheta_{-t}\mathbb{x}_{\tilde{\pi}})(s) = \mathbb{x}_{\tilde{\pi}}(s - t) = \mathbb{x}(\tilde{t}_{n-1})$ for all $s \ge \tilde{t}_{n-1} + t$ and thus, in particular, for all $s \ge 0$. It follows that $\tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x}_{\tilde{\pi}}) = \vartheta_{-t}\mathbb{x}_{\tilde{\pi}}$ and Lemma B.2 yields $\mathbb{d}_{\mathrm{D}}(\mathbb{x}_{\tilde{\pi}}, \tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x}_{\tilde{\pi}})) \le f_\delta(t)$.

Finally, we estimate $\mathbb{d}_{\mathrm{D}}(\tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x}_{\tilde{\pi}}), \tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x}))$. If $-T < s < 0$ it is $\tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x})(s) = \mathbb{x}(s - t)$ and thus $d(\tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x}_{\tilde{\pi}})(s), \tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x})(s)) = d(\mathbb{x}_{\tilde{\pi}}(s - t), \mathbb{x}(s - t)) \le 2\varepsilon$ by the above. On the other hand, for $s \ge 0$ it is $d(\tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x}_{\tilde{\pi}})(s), \tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x})(s)) = d(\mathbb{x}(\tilde{t}_{n-1}), \mathbb{x}(-t-)) \le \varepsilon$ by the initial choice of δ. Thus $\mathbb{d}_{\mathrm{D}}(\tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x}_{\tilde{\pi}}), \tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x})) \le 3\varepsilon$. Altogether,

$$\mathbb{d}_{\mathrm{D}}(\mathbb{x}, \tau_{\mathrm{D}}(\vartheta_{-t}\mathbb{x})) \le 6\varepsilon + f_\delta(t)$$

whenever $\mathbb{x} \in K$ and $0 \le t \le \delta/2$. This implies (P5).

Finally, we prove (P6). As $\mathbb{d}_a^b$ and thus $\mathbb{d}_{J_1}$ is bounded by 1 we find $t^\star > 0$ such that

$$\mathbb{d}_{t^\star}(\mathbb{x}, \mathbb{y}) := \int_{-\infty}^{0} \int_{0}^{t^\star} e^{t-s} \mathbb{d}_s^t(\mathbb{x}, \mathbb{y})\,dt\,ds \ge \frac{1}{2}\mathbb{d}_{J_1}(\mathbb{x}, \mathbb{y})$$

for all $\mathbb{x}, \mathbb{y} \in \mathscr{X}$. For $t_0 > t^\star$, Lemma 2.5(e) yields $\mathbb{d}_{t^*}(\tau_{t_0}(\mathbb{x}), \tau_{t_0}(\mathbb{y})) = \mathbb{d}_{t^\star}(\mathbb{x}, \mathbb{y})$ and thus

$$\mathbb{d}(\tau_{t_0}(\mathbb{x}), \tau_{t_0}(\mathbb{y})) \ge \mathbb{d}_{t^\star}(\mathbb{x}, \mathbb{y}) \ge \frac{1}{2}\mathbb{d}_{J_1}(\mathbb{x}, \mathbb{y}).$$

This finishes the proof. □

References

[1] W. Arendt, C. J. K. Batty, M. Hieber, and F. Neubrander, *Vector-valued Laplace transforms and Cauchy problems*, vol. 96 of Monographs in Mathematics, Birkhäuser/Springer Basel AG, Basel, second ed., 2011.

[2] W. Arendt, S. Kunkel, and M. Kunze, *Diffusion with nonlocal boundary conditions*, J. Funct. Anal., 270 (2016), pp. 2483–2507.

[3] A. Bátkai and S. Piazzera, *Semigroups for delay equations*, vol. 10 of Research Notes in Mathematics, A K Peters, Ltd., Wellesley, MA, 2005.

[4] A. Bensoussan, G. Da Prato, M. C. Delfour, and S. K. Mitter, *Representation and control of infinite dimensional systems*, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, second ed., 2007.

[5] P. Billingsley, *Convergence of probability measures*, Wiley Series in Probability and Statistics: Probability and Statistics, John Wiley & Sons, Inc., New York, second ed., 1999. A Wiley-Interscience Publication.

[6] V. I. Bogachev, *Measure theory. Vol. I, II*, Springer-Verlag, Berlin, 2007.

[7] K. Bogdan and M. Kunze, *The fractional Laplacian with reflections*, Potential Anal., 61 (2024), pp. 317–345.

[8] O. Butkovsky and M. Scheutzow, *Invariant measures for stochastic functional differential equations*, Electron. J. Probab., 22 (2017), pp. Paper No. 98, 23.

[9] S. Cerrai, *A Hille-Yosida theorem for weakly continuous semigroups*, Semigroup Forum, 49 (1994), pp. 349 – 367.

[10] J. B. Cooper, *Saks spaces and applications to functional analysis*, vol. 139 of North-Holland Mathematics Studies, North-Holland Publishing Co., Amsterdam, second ed., 1987. Notas de Matemática [Mathematical Notes], 116.

[11] A. Cosso and F. Russo, *Crandall-Lions viscosity solutions for path-dependent PDEs: the case of heat equation*, Bernoulli, 28 (2022), pp. 481–503.

[12] G. Da Prato and J. Zabczyk, *Stochastic equations in infinite dimensions*, vol. 152 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, second ed., 2014.

[13] F. de Feo, S. Federico, and A. Święch, *Optimal control of stochastic delay differential equations and applications to path-dependent financial and economic models*, SIAM J. Control Optim., 62 (2024), pp. 1490–1520.

[14] R. Denk, M. Kunze, and M. Kupper, *Martingales and path-dependent PDEs via evolutionary semigroups.* Preprint, arXiv:2507.01845, 2025.

[15] O. Diekmann, S. A. van Gils, S. M. Verduyn Lunel, and H.-O. Walther, *Delay equations*, vol. 110 of Applied Mathematical Sciences, Springer-Verlag, New York, 1995. Functional, complex, and nonlinear analysis.

[16] S. Drapeau, A. Jamneshan, M. Karliczek, and M. Kupper, *The algebra of conditional sets and the concepts of conditional topology and compactness*, J. Math. Anal. Appl., 437 (2016), pp. 561–589.

[17] T. Eisner, B. Farkas, M. Haase, and R. Nagel, *Operator theoretic aspects of ergodic theory*, vol. 272 of Graduate Texts in Mathematics, Springer, Cham, 2015.

[18] I. Ekren, N. Touzi, and J. Zhang, *Viscosity solutions of fully nonlinear parabolic path dependent PDEs: Part II*, Ann. Probab., 44 (2016), pp. 2507–2553.

[19] K.-J. Engel and R. Nagel, *One-parameter semigroups for linear evolution equations*, vol. 194 of Graduate Texts in Mathematics, Springer-Verlag, New York, 2000. With contributions by S. Brendle, M. Campiti, T. Hahn, G. Metafune, G. Nickel, D. Pallara, C. Perazzoli, A. Rhandi, S. Romanelli and R. Schnaubelt.

[20] S. N. Ethier and T. G. Kurtz, *Markov processes*, Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics, John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.

[21] B. Farkas, *Perturbations of bi-continuous semigroups on $C_b(H)$ with applications to the Ornstein-Uhlenbeck semigroup*, Semigroup Forum, 68 (2004), pp. 329 – 353.

[22] B. Farkas and H. Kreidler, *Towards a Koopman theory for dynamical systems on completely regular spaces*, Philos. Trans. Roy. Soc. A, 378 (2020), pp. 20190617, 15.

[23] S. Federico, *A stochastic control problem with delay arising in a pension fund model*, Finance Stoch., 15 (2011), pp. 421–459.

[24] S. Federico, B. Goldys, and F. Gozzi, *HJB equations for the optimal control of differential equations with delays and state constraints, I: regularity of viscosity solutions*, SIAM J. Control Optim., 48 (2010), pp. 4910–4937.

[25] F. Flandoli and G. Zanco, *An infinite-dimensional approach to path-dependent Kolmogorov equations*, Ann. Probab., 44 (2016), pp. 2643–2693.

[26] D. H. Fremlin, D. J. H. Garling, and R. G. Haydon, *Bounded measures on topological spaces*, Proc. London Math. Soc. (3), 25 (1972), pp. 115–136.

[27] M. Fuhrman, F. Masiero, and G. Tessitore, *Stochastic equations with delay: optimal control via BSDEs and regular solutions of Hamilton-Jacobi-Bellman equations*, SIAM J. Control Optim., 48 (2010), pp. 4624–4651.

[28] B. Goldys and M. Kocan, *Diffusion semigroups in spaces of continuous functions with mixed topology*, J. Differential Equations, 173 (2001), pp. 17–39.

[29] B. Goldys, M. Nendel, and M. Röckner, *Operator semigroups in the mixed topology and the infinitesimal description of Markov processes*, J. Differential Equations, 412 (2024), pp. 23–86.

[30] F. Gozzi and C. Marinelli, *Stochastic optimal control of delay equations arising in advertising models*, in Stochastic partial differential equations and applications—VII, vol. 245 of Lect. Notes Pure Appl. Math., Chapman & Hall/CRC, Boca Raton, FL, 2006, pp. 133–148.

[31] M. Haase, *The functional calculus for sectorial operators*, no. v. 169 in Operator theory, advances and applications, Birkhäuser Verlag, 2006.

[32] H. Jarchow, *Locally convex spaces*, B. G. Teubner, Stuttgart, 1981. Mathematische Leitfäden. [Mathematical Textbooks].

[33] B. Jefferies, *Weakly integrable semigroups on locally convex spaces*, J. Funct. Anal., 66 (1986), pp. 347 – 364.

[34] K. Kruse, J. Meichsner, and C. Seifert, *Subordination for sequentially equicontinuous equibounded C_0-semigroups*, J. Evol. Equ., 21 (2021), pp. 2665–2690.

[35] F. Kühn and M. Kunze, *Feller generators with measurable lower order terms*, Positivity, 26 (2022), pp. Paper No. 85, 41.

[36] F. Kühnemund, *A Hille-Yosida theorem for bi-continuous semigroups*, Semigroups Forum, 67 (2003), pp. 205 – 225.

[37] M. Kunze, *Continuity and equicontinuity of semigroups on norming dual pairs*, Semigroup Forum, 79 (2009), pp. 540–560.

[38] ———, *A Pettis-type integral and applications to transition semigroups*, Czechoslovak Math. J., 61(136) (2011), pp. 437–459.
[39] ———, *Perturbation of strong Feller semigroups and well-posedness of semilinear stochastic equations on Banach spaces*, Stochastics, 85 (2013), pp. 960–986.
[40] T. Lindvall, *Weak convergence of probability measures and random functions in the function space* $D(0,\infty)$, J. Appl. Probability, 10 (1973), pp. 109–121.
[41] X. Mao, *Stochastic differential equations and their applications*, Horwood Publishing Series in Mathematics & Applications, Horwood Publishing Limited, Chichester, 1997.
[42] B. Øeksendal, *Stochastic differential equations*, Universitext, Springer-Verlag, Berlin, sixth ed., 2003. An introduction with applications.
[43] S. E. A. Mohammed, *Stochastic functional differential equations*, vol. 99 of Research Notes in Mathematics, Pitman (Advanced Publishing Program), Boston, MA, 1984.
[44] E. Pardoux and S. G. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett., 14 (1990), pp. 55–61.
[45] S. Peng and Y. Song, *G-expectation weighted Sobolev spaces, backward SDE and path dependent PDE*, J. Math. Soc. Japan, 67 (2015), pp. 1725–1757.
[46] S. Peng, Y. Song, and F. Wang, *Survey on path-dependent PDEs*, Chinese Ann. Math. Ser. B, 44 (2023), pp. 837–856.
[47] E. Priola, *On a class of Markov type semigroups in spaces of uniformly continuous and bounded functions*, Studia Math., 136 (1999), pp. 271 – 295.
[48] M. Reiss, M. Riedle, and O. van Gaans, *Delay differential equations driven by Lévy processes: stationarity and Feller properties*, Stochastic Process. Appl., 116 (2006), pp. 1409–1432.
[49] M. Rosestolato and A. Święch, *Partial regularity of viscosity solutions for a class of Kolmogorov equations arising from mathematical finance*, J. Differential Equations, 262 (2017), pp. 1897–1930.
[50] F. Sentilles, *Bounded continuous functions on a completely regular space*, Trans. Amer. Math. Soc., 168 (1972), pp. 311 – 336.
[51] W. Whitt, *Some useful functions for functional limit theorems*, Math. Oper. Res., 5 (1980), pp. 67–85.

Email address: robert.denk@uni-konstanz.de

Email address: markus.kunze@uni-konstanz.de

Email address: kupper@uni-konstanz.de

Universität Konstanz, Fachbereich Mathematik und Statistik, 78357 Konstanz, Germany