

Nonlinear Lévy and nonlinear Feller processes: an analytic introduction ^{*†}

Vassili N. Kolokoltsov[‡]

April 5, 2011

Abstract

The program of studying general nonlinear Markov processes was put forward in [6]. This program was developed by the author in monograph [10], where, in particular, nonlinear Lévy processes were introduced. The present paper is an invitation to the rapidly developing topic of this monograph. We provide a quick (and at the same time more abstract) introduction to the basic analytical aspects of the theory developed in Part II of [10].

1 Introduction

Nonlinear Lévy processes were introduced by the author in [10]. We provide a quick introduction to the basic analytical aspects of the theory developed in Part II of [10] giving more concise and more general formulations of some basic facts on well-posedness and sensitivity of nonlinear processes. For general background in Lévy and Markov processes we refer to books [1], [11], [12].

For sensitivity of the nonlinear jump-type processes, e.g. Boltzmann or Smoluchovski, we refer to papers [5] and [2].

Loosely speaking, a nonlinear Markov evolution is just a dynamical system generated by a measure-valued ordinary differential equation (ODE) with the specific feature of preserving positivity. This feature distinguishes it from a general Banach space valued ODE and yields a natural link with probability theory, both in interpreting results and in the tools of analysis. Technical complications for the sensitivity analysis, again compared with the standard theory of vector-valued ODE, lie in the specific unboundedness of generators that causes the derivatives of the solutions to nonlinear equations (with respect to parameters or initial conditions) to live in other spaces, than the evolution itself. From the probabilistic point of view, the first derivative with respect to initial data (specified by the linearized evolution around a path of nonlinear dynamics) describes the interacting particle approximation to this nonlinear dynamics (which, in turn, serves as the dynamic law of large numbers to this approximating Markov system of interacting particles), and

^{*}<http://arxiv.org/abs/1103.5591>

[†]Supported by the AFOSR grant FA9550-09-1-0664 'Nonlinear Markov control processes and games'

[‡]Department of Statistics, University of Warwick, Coventry CV4 7AL UK, Email: v.kolokoltsov@warwick.ac.uk

the second derivative describes the limit of fluctuations of the evolution of particle systems around its law of large numbers (probabilistically the dynamic central limit theorem). In this paper we concentrate only on the analytic aspects of the theory referring to [10] for probabilistic interpretation.

Recall first the definition of a propagator. For a set S , a family of mappings $U^{t,r}$, from S to itself, parametrized by the pairs of real numbers $r \leq t$ (resp. $t \leq r$) from a given finite or infinite interval is called a *forward propagator* (resp. a *backward propagator*), if $U^{t,t}$ is the identity operator in S for all t and the following *chain rule*, or *propagator equation*, holds for $r \leq s \leq t$ (resp. for $t \leq s \leq r$): $U^{t,s}U^{s,r} = U^{t,r}$. If the mappings $U^{t,r}$ forming a backward propagator depend only on the differences $r - t$, then the family $T^t = U^{0,t}$ forms a semigroup. That is why, propagators are sometimes referred to as two-parameter semigroups. By a propagator we mean a forward or a backward propagator (which should be clear from the context).

Let $\tilde{\mathcal{M}}(X)$ be a dense subset of the space $\mathcal{M}(X)$ of finite (positive Borel) measures on a polish (complete separable metric) space X (considered in its weak topology). By a nonlinear *sub-Markov* (resp. *Markov*) *propagator* in $\tilde{\mathcal{M}}(X)$ we shall mean any propagator $V^{t,r}$ of possibly nonlinear transformations of $\tilde{\mathcal{M}}(X)$ that do not increase (resp. preserve) the norm. If $V^{t,r}$ depends only on the difference $t - r$ and hence specifies a semigroup, this semigroup is called nonlinear or generalized *sub-Markov* or *Markov* respectively.

The usual, linear, Markov propagators or semigroups correspond to the case when all the transformations are linear contractions in the whole space $\mathcal{M}(X)$. In probability theory these propagators describe the evolution of averages of Markov processes, i.e. processes whose evolution after any given time t depends on the past $X_{\leq t}$ only via the present position X_t . Loosely speaking, to any nonlinear Markov propagator there corresponds a process whose behavior after any time t depends on the past $X_{\leq t}$ via the position X_t of the process and its distribution at t .

More precisely, consider the nonlinear kinetic equation

$$\frac{d}{dt}(g, \mu_t) = (A[\mu_t]g, \mu_t) \quad (1)$$

with a certain family of operators $A[\mu]$ in $C(X)$ depending on μ as a parameter and such that each $A[\mu]$ specifies a uniquely defined Markov process (say, via solution to the corresponding martingale problem, or by generating a Feller semigroup).

Suppose that the Cauchy problem for equation (1) is well posed and specifies the weakly continuous Markov semigroup T_t in $\mathcal{M}(X)$. Suppose also that for any weakly continuous curve $\mu_t \in \mathcal{P}(X)$ (the set of probability measures on X) the solutions to the Cauchy problem of the equation

$$\frac{d}{dt}(g, \nu_t) = (A[\mu_t]g, \nu_t) \quad (2)$$

define a weakly continuous propagator $V^{t,r}[\mu]$, $r \leq t$, of linear transformations in $\mathcal{M}(X)$ and hence a Markov process in X , with transition probabilities $p_{r,t}^{[\mu]}(x, dy)$. Then to any $\mu \in \mathcal{P}(X)$ there corresponds a (usual linear, but time non-homogeneous) Markov process X_t^ν in X (ν stands for an initial distribution) such that its distributions ν_t solve equation (2) with the initial condition ν . In particular, the distributions of X_t^μ (with the initial condition μ) are $\mu_t = T_t(\mu)$ for all times t and the transition probabilities $p_{r,t}^{[\mu]}(x, dy)$

specified by equation (2) satisfy the condition

$$\int_{X^2} f(y) p_{r,t}^{[\mu]}(x, dy) \mu_r(dx) = (f, V^{t,r} \mu_r) = (f, \mu_t). \quad (3)$$

We shall call the family of processes X_t^μ a *nonlinear Markov process*. When each $A[\mu]$ generates a Feller semigroup and T_t acts on the whole $\mathcal{M}(X)$ (and not only on its dense subspace), the corresponding process can be also called *nonlinear Feller*. Allowing for the evolution on subsets $\tilde{\mathcal{M}}(X)$ is however crucial, as it often occurs in applications, say for the Smoluchovski or Boltzmann equation with unbounded rates.

Thus a nonlinear Markov process is a semigroup of the transformations of distributions such that to each trajectory is attached a “tangent” Markov process with the same marginal distributions. The structure of these tangent processes is not intrinsic to the semigroup, but can be specified by choosing a stochastic representation for the generator, that is of the r.h.s. of (2).

In this paper we shall prove a general well-posedness result for nonlinear Markov semigroups that will cover, as particular cases,

(i) *nonlinear Lévy processes* specified by the families

$$\begin{aligned} A_\mu f(x) &= \frac{1}{2} (G(\mu) \nabla, \nabla) f(x) + (b(\mu), \nabla f)(x) \\ &+ \int [f(x+y) - f(x) - (y, \nabla f(x)) \mathbf{1}_{B_1}(y)] \nu(\mu, dy), \end{aligned} \quad (4)$$

where, for each probability measure μ on \mathbf{R}^d , $\nu(\mu, \cdot)$ is a Lévy measure (i.e. a Borel measure on \mathbf{R}^d without a mass point at the origin and such that the function $\min(1, |y|^2)$ is integrable with respect to it), $G(\mu)$ is a symmetric non-negative $d \times d$ -matrix, $b(\mu)$ a vector in \mathbf{R}^d and B_1 is the unit ball in \mathbf{R}^d with $\mathbf{1}_{B_1}$ being the corresponding indicator function;

(ii) *processes of order at most one* specified by the families

$$A_\mu f(x) = (b(x, \mu), \nabla f(x)) + \int_{\mathbf{R}^d} (f(x+y) - f(x)) \nu(x, \mu, dy), \quad (5)$$

where the Lévy measure ν is supposed to have a finite first moment;

(iii) *mixtures of possibly degenerate diffusions and stable-like processes* and processes generated by the operators of order at most one, explicitly defined below in Proposition 4.1.

It is worth noting that equations of type (2) that appear naturally as dynamic Law of Large Numbers for interacting particles, can be deduced, on the other hand, from the mere assumption of positivity preservation, see [10] and [14]. In case of diffusion (partial second order) operators $A[\mu]$, the corresponding evolution (1) was first analyzed by McKean and is often called the *McKean* or *McKean-Vlasov* diffusion. Its particular case that arises as the limit of grazing collisions in the Boltzmann collision model is sometimes referred to as the *Landau-Fokker-Planck* equation, see [4] for some recent results. The case of $A[\mu]$ being a Hamiltonian vector field is often called a *Vlasov-type* equation, as it contains the celebrated *Vlasov equation* from plasma physics. The case of $A[\mu]$ being pure integral operators comprises a large variety of models from statistical mechanics (say, Boltzmann and Smoluchovski equations) to evolutionary games (replicator dynamics), see [10] for a

comprehensive review and papers [3], [15], [16] for the introduction to nonlinear Markov evolutions from the physical point of view.

The following basic notations will be used:

$C_\infty(\mathbf{R}^d) \subset C(\mathbf{R}^d)$ consists of f such that $\lim_{x \rightarrow \infty} f(x) = 0$,

$C^k(\mathbf{R}^d)$ (resp. $C_\infty^k(\mathbf{R}^d)$) is the Banach space of k times continuously differentiable functions with bounded derivatives on \mathbf{R}^d (resp. its closed subspace of functions f with $f^{(l)} \in C_\infty(\mathbf{R}^d)$, $l \leq k$) with

$$\|f\|_{C^k(\mathbf{R}^d)} = \sum_{l=0}^k \|f^{(l)}\|_{C(\mathbf{R}^d)},$$

$\mathcal{P}(\mathbf{R}^d)$ the set of probability measures on \mathbf{R}^d .

$\|A\|_{D \rightarrow B}$ denotes the norm of an operator A in the Banach space $\mathcal{L}(D, B)$ of bounded linear operators between Banach spaces D and B , and $\|\xi\|_B$ denotes the norm of ξ as an element of the Banach space B .

2 Dual propagators

A backward propagator $\{U^{t,r}\}$ of uniformly (for t, r from a compact set) bounded linear operators on a Banach space B is called *strongly continuous* if the family $U^{t,r}$ depends strongly continuously on t and r .

For a strongly continuous backward propagator $\{U^{t,r}\}$ of bounded linear operators on a Banach space B with a common invariant domain $D \subset B$, which is itself a Banach space with the norm $\|\cdot\|_D \geq \|\cdot\|_B$, let $\{A_t\}$, $t \geq 0$, be a family of bounded linear operators $D \rightarrow B$ depending strongly measurably on t (i.e. the function $t \mapsto A_t f \in B$ is measurable for each $f \in D$). Let us say that the family $\{A_t\}$ *generates* $\{U^{t,r}\}$ on the invariant domain D if the equations

$$\frac{d}{ds} U^{t,s} f = U^{t,s} A_s f, \quad \frac{d}{ds} U^{s,r} f = -A_s U^{s,r} f, \quad t \leq s \leq r, \quad (6)$$

hold a.s. in s for any $f \in D$, that is there exists a negligible subset S of \mathbf{R} such that for all $t < r$ and all $f \in D$ equations (6) hold for all s outside S , where the derivatives exist in the Banach topology of B . In particular, if the operators A_t depend strongly continuously on t (as bounded operators $D \rightarrow B$), this implies that equations (6) hold for all s and $f \in D$, where for $s = t$ (resp. $s = r$) it is assumed to be only a right (resp. left) derivative.

For a Banach space B or a linear operator A one usually denotes by B^* or A^* its *Banach dual* (space or operator respectively). Alternatively the notations B' and A' are in use.

Theorem 2.1. (Basic duality)

Let $\{U^{t,r}\}$ be a strongly continuous backward propagator of bounded linear operators in a Banach space B with a common invariant domain D , which is itself a Banach space with the norm $\|\cdot\|_D \geq \|\cdot\|_B$, and let the family $\{A_t\}$ of bounded linear operators $D \rightarrow B$ generate $\{U^{t,r}\}$ on D . Then

(i) the family of dual operators $V^{s,t} = (U^{t,s})^*$ forms a weakly- \star continuous in s, t propagator of bounded linear operators in B^* (contractions if all $U^{t,r}$ are contractions) such that

$$\frac{d}{dt}V^{s,t}\xi = -V^{s,t}A_t^*\xi, \quad \frac{d}{ds}V^{s,t}\xi = A_s^*V^{s,t}\xi, \quad t \leq s, \quad (7)$$

hold weakly- \star in D^* , i.e., say, the second equation means

$$\frac{d}{ds}(f, V^{s,t}\xi) = (A_s f, V^{s,t}\xi), \quad t \leq s, \quad f \in D; \quad (8)$$

(ii) $V^{s,t}\xi$ is the unique solution to the Cauchy problem of equation (8) in B^* , i.e. if $\xi_t = \xi$ for a given $\xi \in B^*$ and $\xi_s, s \in [t, r]$, is a weakly- \star continuous family in B^* satisfying

$$\frac{d}{ds}(f, \xi_s) = (A_s f, \xi_s), \quad t \leq s \leq r, \quad f \in D, \quad (9)$$

then $\xi_s = V^{s,t}\xi$ for $t \leq s \leq r$.

(iii) $U^{s,r}f$ is the unique solution to the inverse Cauchy problem of the second equation in (6).

Proof. Statement (i) is a direct consequence of duality.

(ii) Let $g(s) = (U^{s,r}f, \xi_s)$ for a given $f \in D$. Writing

$$\begin{aligned} & (U^{s+\delta,r}f, \xi_{s+\delta}) - (U^{s,r}f, \xi_s) \\ &= (U^{s+\delta,r}f - U^{s,r}f, \xi_s) + (U^{s,r}f, \xi_{s+\delta} - \xi_s) \\ &+ (U^{s+\delta,r}f - U^{s,r}f, \xi_{s+\delta} - \xi_s) \end{aligned}$$

and using (6), (8) and the invariance of D , allows one to conclude that

$$\frac{d}{ds}g(s) = -(A_s U^{s,r}f, \xi_s) + (U^{s,r}f, A_s^* \xi_s) = 0,$$

because a.s. in s

$$\left(\frac{U^{s+\delta,r}f - U^{s,r}f}{\delta}, \xi_{s+\delta} - \xi_s \right) \rightarrow 0,$$

as $\delta \rightarrow 0$ (since the family $\delta^{-1}(U^{s+\delta,r}f - U^{s,r}f)$ is relatively compact, being convergent, and ξ_s is weakly continuous). Hence $g(r) = (f, \xi_r) = g(t) = (U^{t,r}f, \xi_t)$, showing that ξ_r is uniquely defined.

(iii) is proved similar to (ii). □

Remark 1. In addition to the statement of Theorem 2.1 let us note (as one sees directly from duality), that (i) $V^{s,t}\xi$ depend weakly- \star continuous on s, t uniformly for bounded ξ and (ii) $V^{s,t}$ is a weakly- \star continuous operator, that is $\xi_n \rightarrow \xi$ weakly- \star implies $V^{s,t}\xi_n \rightarrow V^{s,t}\xi$ weakly- \star .

Remark 2. Working with discontinuous A_t is crucial for the development of the related theory of SDE with nonlinear noise, see [8] and [9]. In this paper we shall use only continuous families of generators $\{A_t\}$.

We deduce now some corollaries of Theorem 2.1: on the extension of the operators $V^{s,t}$ to D^* , and on their stability with respect to a perturbation of the family A_t .

Theorem 2.2. *Under the assumptions of Theorem 2.1 suppose additionally that*

(i) $\{U^{t,s}\}$ is a strongly continuous backward propagator of uniformly bounded operators in D ;

(ii) there exists another subspace $\tilde{D} \subset D$, dense in D , which is itself a Banach space with the norm $\|\cdot\|_{\tilde{D}} \geq \|\cdot\|_D$ such that the mapping $t \mapsto A_t$ is a continuous mapping $t \rightarrow \mathcal{L}(\tilde{D}, D)$;

(iii) B^* is dense in D^* (which holds automatically in case of reflexive D).

Then the operators $V^{s,t} : B^* \rightarrow B^*$ extend to the operators $V^{s,t} : D^* \rightarrow D^*$ forming a weakly- \star continuous propagator in D^* that solves equation (8) weakly- \star in \tilde{D}^* , that is, for any $\xi \in D^*$, equation (8) holds for all $f \in \tilde{D}$.

Proof. The fact that $V^{s,t}$ extend to linear operators in D^* follows without any additional assumption from the invariance of D under $U^{t,s}$. Assumption (i) implies that this extension is bounded and weakly- \star continuous in D^* . In order to prove that (8) holds for $f \in \tilde{D}$ and $\xi \in D^*$, observe that

$$(f, V^{r,t}\xi) = (f, \xi) + \int_t^r (A_s f, V^{s,t}\xi) ds \quad (10)$$

for $\xi \in B^*$, $f \in D$. Now, for a $\xi \in D^*$ and $f \in \tilde{D}$, let us pick up a sequence $\xi_n \in B^*$ such that $\xi_n \rightarrow \xi$ in the norm topology of D^* as $n \rightarrow \infty$ (which is possible by assumption (iii)). As $A_s f \in D$ (by assumption(ii)), we can pass to the limit in (10) with ξ_n instead of ξ (using dominated convergence) yielding (10) for $\xi \in D^*$ and $f \in \tilde{D}$. Finally, as $(A_s f, V^{s,t}\xi)$ is a continuous function of s (by assumption (ii) and the weak- \star continuity of $V^{s,t}$ in D^*), equation (10) implies (8) for $\xi \in D^*$ and $f \in \tilde{D}$. \square

Theorem 2.3. *Under the assumptions of Theorem 2.2 assume additionally that the backward propagator $\{U^{t,s}\}$ in D is generated by $\{A_t\}$ on the invariant domain \tilde{D} (in particular \tilde{D} is invariant and equations (6) hold in the norm topology of D for any $f \in \tilde{D}$). Then $V^{s,t}\xi$ represents the unique weakly- \star continuous in D^* solution of equation (8) in \tilde{D}^* . Moreover, for the propagator $\{U^{t,s}\}$ in D to be generated by $\{A_t\}$ on \tilde{D} it is sufficient to assume that $\{U^{t,s}\}$ is a strongly continuous family of bounded operators in \tilde{D} .*

Proof. The first statement is a direct consequence of Theorem 2.1 applied to the pair of spaces \tilde{D}, D . The last statement is proved as in the previous theorem. Namely, we first rewrite equation (6) in the integral form, i.e. as

$$U^{t,r} f = f + \int_t^r A_s U^{s,r} f ds, \quad U^{t,r} f = f + \int_t^r U^{t,s} A_s f ds. \quad (11)$$

These equations would imply (6) with the derivative defined in the norm topology of D , for $f \in \tilde{D}$, if we can prove that the functions $A_s U^{s,r} f$ and $U^{t,s} A_s f$ are continuous functions $s \mapsto D$. To see that this is true, say for the first function, we can write

$$A_{s+\delta} U^{s+\delta,r} f - A_s U^{s,r} f = A_{s+\delta} (U^{s+\delta,r} f - U^{s,r} f) + (A_{s+\delta} - A_s) U^{s,r} f.$$

The first term tends to zero in the norm topology of D , as $\delta \rightarrow 0$, by the strong continuity of $U^{s,r}$ in \tilde{D} , and the second term tends to zero by the continuity of the family A_s (assumption (ii) of Theorem 2.2). \square

We conclude this section with a simple result on the convergence of propagators.

Theorem 2.4. *Suppose we are given a sequence of backward propagators $\{U_n^{t,r}\}$, $n = 1, 2, \dots$, generated by the families $\{A_t^n\}$ and a backward propagator $\{U^{t,r}\}$ generated by the family $\{A_t\}$. Let all these propagators satisfy the same conditions as $U^{t,r}$ and A_t from Theorem 2.1 with the same D, B . Suppose also that all $U^{t,r}$ are uniformly bounded as operators in D .*

Assume finally that, for any t and any $f \in D$, $A_t^n f$ converge to $A_t f$, as $n \rightarrow \infty$, in the norm topology of B . Then $U_n^{t,r}$ converges to $U^{t,r}$ strongly in B . Moreover,

$$\|(V_n^{r,t} - V^{r,t})\xi\|_{D^*} \leq c \|A_s^n - A_s\|_{D \rightarrow B} \|\xi\|_{B^*}. \quad (12)$$

Proof. By the density argument (taking into account that $U_n^{t,r} g$ are uniformly bounded in B), in order to prove the strong convergence of $U_n^{t,r}$ to $U^{t,r}$, it is sufficient to prove that $U_n^{t,r} g$ converges to $U^{t,r} g$ for any $g \in D$. But if $g \in D$,

$$(U_n^{t,r} - U^{t,r})g = U_n^{t,s} U^{s,r} g \Big|_{s=t}^r = \int_t^r U_n^{t,s} (A_s^n - A_s) U^{s,r} g ds, \quad (13)$$

which converges to zero in the norm topology of B by the dominated convergence. Estimate (12) also follows from (13). \square

3 Perturbation theory for weak propagators

The main point of the perturbation theory is to build a propagator generated by the family of operators $\{A_t + F_t\}$, when a propagator $U^{t,r}$ generated by $\{A_t\}$ is given and $\{F_t\}$ are bounded. However, if $\{F_t\}$ are only bounded, then instead of the solutions to the equation

$$\frac{d}{ds} f = A_s f + F_s f, \quad t \leq s \leq r, \quad (14)$$

with a given terminal f_r , as desired, one can only construct the solutions to the so called *mild form* of this equation:

$$f_t = U^{t,r} f + \int_t^r U^{t,s} F_s f_s ds, \quad (15)$$

which is only formally equivalent to (14) (i.e. when a solution to the mild equation is regular enough which may not be the case).

Let us recall the simplest perturbation theory result for propagators, which clarifies this issue (a proof can be found e.g. in [11], Theorem 1.9.3, and simpler, but similar fact for semigroups is discussed in almost any text book on functional analysis).

Theorem 3.1. *(i) Let $U^{t,r}$ be a strongly continuous backward propagator of bounded linear operators in a Banach space B , and $\{F_t\}$ be a family of bounded operators in B depending strongly continuous on t . Set*

$$\Phi^{t,r} = U^{t,r} + \int_t^r U^{t,s} F_s U^{s,r} ds + \sum_{m=1}^{\infty} \int_{t \leq s_1 \leq \dots \leq s_m \leq r} U^{t,s_1} F_{s_1} U^{s_1,s_2} \dots F_{s_m} U^{s_m,r} ds_1 \dots ds_m. \quad (16)$$

It is claimed that this series converges in B and the family $\{\Phi^{t,r}\}$ also forms a strongly continuous propagator of bounded operators in B such that $f_t = \Phi^{t,s} f$ is the unique solution to equation (15).

(ii) Suppose additionally that a family of linear operators $\{A_t\}$ generates $\{U^{t,r}\}$ on the common invariant domain D , which is dense in B and is itself a Banach space under a norm $\|\cdot\|_D \geq \|\cdot\|_B$. Suppose that $U^{t,r}$ and $\{F_t\}$ are also uniformly bounded operators in D . Then D is invariant under $\{\Phi^{t,r}\}$ and the family $\{A_t + F_t\}$ generates $\{\Phi^{t,r}\}$ on D . Moreover, series (16) also converges in the operator norms of D and operators $\Phi^{t,r} f$ are bounded as operators in the Banach space D .

We presented this theorem, because for the sensitivity analysis of nonlinear equations we shall need non-homogeneous extensions of equations (9) of the form

$$\frac{d}{ds}(f, \xi_s) = (A_s f, \xi_s) + (F_s f, \xi_s), \quad t \leq s \leq r, \quad (17)$$

where F_s is a family of operators bounded in D , but, what is crucial and necessitates technical complications, not bounded in B .

Under the assumption of Theorem 2.3 and assuming $\{F_t\}$ is a bounded strongly continuous family of operators in D , it follows directly from Theorem 3.1 (ii) applied to the pair of Banach spaces (D, \tilde{D}) that the perturbation theory propagator (16) solves equation (14) in D and is generated on \tilde{D} by the family $\{A_t + F_t\}$. Hence, by Theorem 2.1, the dual propagator $\{\Psi^{r,t} = (\Phi^{t,r})'\}$ is weakly- \star continuous in D^* and yields a unique solution to (17) in \tilde{D}^* (i.e. so that, for $\xi_s = \Psi^{s,t} \xi_t$, equation (17) holds for all $f \in \tilde{D}$).

The next result proves the same fact, except for uniqueness, under weaker assumptions of Theorem 2.2.

Theorem 3.2. *Under the assumptions of Theorem 2.2 assume $\{F_t\}$ is a bounded strongly continuous family of operators in D . Let $\{\Phi^{t,r}\}$ be given by (16), which by Theorem 3.1 (i) (applied to the pair of Banach spaces (D, \tilde{D})) is a strongly continuous propagator in D , and let $\{\Psi^{r,t} = (\Phi^{t,r})'\}$, which is clearly a weakly- \star continuous backward propagator in D^* . Then the curve $\xi_s = \Psi^{s,t} \xi_t$ solves equation (17) in \tilde{D}^* with a given terminal condition ξ_t , that is (17) holds for all $f \in \tilde{D}$.*

Proof. From duality and (16) it follows that

$$\Psi^{r,t} = V^{r,t} + \sum_{m=1}^{\infty} \int_{t \leq s_1 \leq \dots \leq s_m \leq r} V^{r,s_m} F'_{s_m} \dots V^{s_2,s_1} F'_{s_1} V^{s_1,t} ds_1 \dots ds_m, \quad (18)$$

where F'_s are of course dual operators to F_s , and where the integral is understood in weak- \star sense and the series converges in the norm-topology of D^* (we need to take into account Remark 2 to see that the weak integral is well defined). To prove (17) for $f \in \tilde{D}$ we should now differentiate term by term the corresponding series $(f, \Psi^{r,t} \xi)$ with respect to r using Theorem 2.2. This term-by-term differentiation is then justified by the fact that the series of derivatives

$$(A_r f, V^{r,t} \xi_t) + \left[(F_r f, V^{r,t} \xi_t) + \int_t^r (A_r f, V^{r,s} F'_s V^{s,t}) ds \right] + \dots$$

converges uniformly in r . □

4 T -products

Here we shall recall the notion of T -products showing how they can be used to construct propagators generated by families of operators each of which generates a sufficiently regular semigroup.

We shall work with three Banach spaces B_0, B_1, B_2 with the norms denoted by $\|\cdot\|_i$, $i = 0, 1, 2$, such that $B_0 \subset B_1 \subset B_2$, B_0 is dense in B_1 , B_1 is dense in B_2 and $\|\cdot\|_0 \geq \|\cdot\|_1 \geq \|\cdot\|_2$.

Let $L_t : B_1 \mapsto B_2$, $t \geq 0$, be a family of uniformly (in t) bounded operators such that the closure in B_2 of each L_t is the generator of a strongly continuous semigroup of bounded operators in B_2 . For a partition $\Delta = \{0 = t_0 < t_1 < \dots < t_N = t\}$ of an interval $[0, t]$ let us define a family of operators $U_\Delta(\tau, s)$, $0 \leq s \leq \tau \leq t$, by the rules

$$\begin{aligned} U_\Delta(\tau, s) &= \exp\{(\tau - s)L_{t_j}\}, \quad t_j \leq s \leq \tau \leq t_{j+1}, \\ U_\Delta(\tau, r) &= U_\Delta(\tau, s)U_\Delta(s, r), \quad 0 \leq r \leq s \leq \tau \leq t. \end{aligned}$$

Let $\Delta t_j = t_{j+1} - t_j$ and $\delta(\Delta) = \max_j \Delta t_j$. If the limit

$$U(s, r)f = \lim_{\delta(\Delta) \rightarrow 0} U_\Delta(s, r)f \quad (19)$$

exists for some f and all $0 \leq r \leq s \leq t$ (in the norm of B_2), it is called the T -product (or *chronological exponent*) of L_t and is denoted by $T \exp\{\int_r^s L_\tau d\tau\}f$. Intuitively, one expects the T -product to give a solution to the Cauchy problem

$$\frac{d}{dt}\phi = L_t\phi, \quad \phi_0 = f, \quad (20)$$

in B_2 with the initial conditions f from B_1 .

Theorem 4.1. *Let a family $L_t f$, $t \geq 0$, of linear operators in B_2 be given such that*

(i) *each L_t generates a strongly continuous semigroup e^{sL_t} , $s \geq 0$, in B_2 with invariant core B_1 ,*

(ii) *L_t are uniformly bounded operators $B_0 \rightarrow B_1$ and $B_1 \rightarrow B_2$,*

(iii) *B_0 is also invariant under all e^{sL_t} and these operators are uniformly bounded as operators in B_0, B_1, B_2 , with the norms not exceeding e^{Ks} with a constant K (the same for all B_j and L_t),*

(iv) *$L_t f$, as a function $t \mapsto B_2$, depends continuously on t locally uniformly in f (i.e. for f from bounded subsets of B_1).*

Then

(i) *the T -product $T \exp\{\int_0^s L_\tau d\tau\}f$ exists for all $f \in B_2$, and the convergence in (19) is uniform in f on any bounded subset of B_1 ;*

(ii) *if $f \in B_0$, then the approximations $U_\Delta(s, r)$ converge also in B_1 ;*

(iii) *this T -product defines a strongly continuous (in t, s) family of uniformly bounded operators in both B_1 and B_2 ,*

(iv) *this T -product $T \exp\{\int_0^s L_\tau d\tau\}f$ is a solution of problem (20) for any $f \in B_1$.*

Proof. (i) Since B_1 is dense in B_2 and all $U_\Delta(s, r)$ are uniformly bounded in B_2 (by (iii)), the existence of the T -product for all $f \in B_2$ follows from its existence for $f \in B_1$. In the latter case it follows from the formula

$$U_\Delta(s, r) - U_{\Delta'}(s, r) = U_{\Delta'}(s, \tau)U_\Delta(\tau, r)|_{\tau=r}^{\tau=s} = \int_r^s \frac{d}{d\tau} U_{\Delta'}(s, \tau)U_\Delta(\tau, r) d\tau$$

$$= \int_r^s U_{\Delta'}(s, \tau)(L_{[\tau]_{\Delta}} - L_{[\tau]_{\Delta'}})U_{\Delta}(\tau, r) d\tau$$

(where we denoted $[s]_{\Delta} = t_j$ for $t_j \leq s < t_{j+1}$), because L_t are uniformly continuous (condition (iv)) and $U_{\Delta}(s, r)$ are uniformly bounded in B_2 and B_1 (by condition (iii)).

(ii) If $f \in B_0$, then the equations

$$U_{\Delta}(s, r) = \int_r^s L_{[\tau]_{\Delta}} U_{\Delta}(\tau, r) d\tau,$$

imply that the family $U_{\Delta}(s, t)$ is uniformly Lipschitz continuous in B_1 as a function of t , because L_s are uniformly bounded operators $B_0 \rightarrow B_1$ and $U_{\Delta}(s, r)$ are uniformly bounded in B_0 . Hence one can choose a subsequence, $U_{\Delta_n}(s, r)$, converging in $C([0, T], B_1)$. But the limit is unique (it is the limit in B_2), implying the convergence of the whole family $U_{\Delta}(s, t)$, as $\delta(\Delta) \rightarrow 0$.

(iii) It follows from (iii) that the limiting propagator is bounded. Strong continuity in B_1 is deduced first for $f \in B_0$ and then for all $f \in B_1$ by the density argument.

(iv) If $f \in B_0$, we can pass to the limit in the above approximate equations to obtain the equation

$$U(s, r)f = \int_r^s L_{\tau} U(\tau, r)f d\tau.$$

Since B_0 is dense in B_1 , we then deduce the same equation for an arbitrary $f \in B_1$. This implies that $U(s, r)f$ satisfies equation (20) by condition (iv) and the basic theorem of calculus. \square

To conclude the section we present a rather general example of a non-homogeneous generator of a strongly continuous Markov propagator specifying a time nonhomogeneous Feller process. This will be a time-nonhomogeneous possibly degenerate diffusion combined with a mixture of possibly degenerate stable-like processes and processes generated by the operators of order at most one, that is a process generated by an operator of the form

$$\begin{aligned} L_t f(x) &= \frac{1}{2} \text{tr}(\sigma_t(x) \sigma_t^T(x) \nabla^2 f(x)) + (b_t(x), \nabla f(x)) + \int (f(x+y) - f(x)) \nu_t(x, dy) \\ &+ \int_P (dp) \int_0^K d|y| \int_{S^{d-1}} a_{p,t}(x, s) \frac{f(x+y) - f(x) - (y, \nabla f(x))}{|y|^{\alpha_{p,t}(x,s)+1}} \omega_{p,t}(ds). \end{aligned} \quad (21)$$

Here $s = y/|y|$, $K > 0$ and (P, dp) is a Borel space with a finite measure dp and $\omega_{p,t}$ are certain finite Borel measures on S^{d-1} .

Proposition 4.1. *Let the functions σ, b, a, α and the finite measure $|y|\nu(x, dy)$ be of smoothness class C^5 with respect to all variables (the measure is smooth in the weak sense), and a_p, α_p take values in compact subintervals of $(0, \infty)$ and $(0, 2)$ respectively. Then the family of operators L_t of form (21) generates a backward propagator $U_{t,s}$ on the invariant domain $C_{\infty}^2(\mathbf{R}^d)$, and hence a unique Markov process.*

Proof. For a detailed proof (that uses several ingredients including Theorem 4.1 as a final step) we refer to the book [11]. \square

5 Nonlinear propagators

The following result from [10] represents the basic tool allowing one to build nonlinear propagators from infinitesimal linear ones.

Recall that $V^{s,t}$ denotes the dual of $U^{t,s}$ given by Theorem 2.1. Let M be a bounded subset of B^* that is closed in the norm topologies of both B^* and D^* . For a $\mu \in M$ let $C_\mu([0, r], M)$ be the metric space of the continuous in the norm D^* curves $\xi_s \in M$, $s \in [0, r]$, $\xi_0 = \mu$, with the distance

$$\rho(\xi, \eta) = \sup_{s \in [0, r]} \|\xi_s - \eta_s\|_{D^*}.$$

Theorem 5.1. (i) Let D be a dense subspace of a Banach space B that is itself a Banach space such that $\|f\|_D \geq \|f\|_B$, and let $\xi \mapsto A[\xi]$ be a mapping from B^* to bounded linear operators $A[\xi] : D \rightarrow B$ such that

$$\|A[\xi] - A[\eta]\|_{D \rightarrow B} \leq c\|\xi - \eta\|_{D^*}, \quad \xi, \eta \in B^*. \quad (22)$$

(ii) For any $\mu \in M$ and $\xi \in C_\mu([0, r], M)$, let the operator curve $A[\xi_t] : D \rightarrow B$ generate a strongly continuous backward propagator of uniformly bounded linear operators $U^{t,s}[\xi]$ in B , $0 \leq t \leq s \leq r$, on the common invariant domain D (in particular, (6) holds), such that

$$\|U^{t,s}[\xi]\|_{D \rightarrow D} \leq c, \quad t \leq s \leq r, \quad (23)$$

for some constant $c > 0$ and with their dual propagators $V^{s,t}[\xi]$ preserving the set M .

Then the weak nonlinear Cauchy problem

$$\frac{d}{dt}(f, \mu_t) = (A[\mu_t]f, \mu_t), \quad \mu_0 = \mu, \quad f \in D, \quad (24)$$

is well posed in M . More precisely, for any $\mu \in M$ it has a unique solution $T_t(\mu) \in M$, and the transformations T_t of M form a semigroup for $t \in [0, r]$ depending Lipschitz continuously on time t and the initial data in the norm of D^* , i.e.

$$\|T_t(\mu) - T_t(\eta)\|_{D^*} \leq c(r, M)\|\mu - \eta\|_{D^*}, \quad \|T_t(\mu) - \mu\|_{D^*} \leq c(r, M)t \quad (25)$$

with a constant $c(r, M)$.

Proof. Since

$$(f, (V^{t,0}[\xi^1] - V^{t,0}[\xi^2])\mu) = (U^{0,t}[\xi^1]f - U^{0,t}[\xi^2]f, \mu)$$

and

$$\begin{aligned} U^{0,t}[\xi^1] - U^{0,t}[\xi^2] &= U^{0,s}[\xi^1]U^{s,t}[\xi^2] \Big|_{s=0}^t \\ &= \int_0^t U^{0,s}[\xi^1](A[\xi_s^1] - A[\xi_s^2])U^{s,t}[\xi^2] ds, \end{aligned}$$

and taking into account (22) and (23) one deduces that

$$\begin{aligned} \|(V^{t,0}[\xi^1] - V^{t,0}[\xi^2])\mu\|_{D^*} &\leq \|U^{0,t}[\xi^1] - U^{0,t}[\xi^2]\|_{D \rightarrow B} \|\mu\|_{B^*} \\ &\leq tc(r, M) \sup_{s \in [0, r]} \|\xi_s^1 - \xi_s^2\|_{D^*} \end{aligned}$$

(of course we used the assumed boundedness of M), implying that for $t \leq t_0$ with a small enough t_0 the mapping $\xi_t \mapsto V^{t,0}[\xi_t]$ is a contraction in $C_\mu([0, t], M)$. Hence by the contraction principle there exists a unique fixed point for this mapping. To obtain the unique global solution one just has to iterate the construction on the next interval $[t_0, 2t_0]$, then on $[2t_0, 3t_0]$, etc. The semigroup property of T_t follows directly from uniqueness.

Finally, if $T_t(\mu) = \mu_t$ and $T_t(\eta) = \eta_t$, then

$$\mu_t - \eta_t = V^{t,0}[\mu_t]\mu - V^{t,0}[\eta_t]\eta = (V^{t,0}[\mu_t] - V^{t,0}[\eta_t])\mu + V^{t,0}[\eta_t](\mu - \eta).$$

Estimating the first term as above yields

$$\sup_{s \leq t} \|\mu_s - \eta_s\|_{D^*} \leq c(r, M)(t \sup_{s \leq t} \|\mu_s - \eta_s\|_{D^*} + \|\mu - \eta\|_{D^*}),$$

which implies the first estimate in (25) first for small times, which is then extended to all finite times by the iteration. The second estimate in (25) follows from (8). \square

Remark 3. For our purposes, the basic examples are given by $B = C_\infty(\mathbf{R}^d)$, $M = \mathcal{P}(\mathbf{R}^d)$, and $D = C_\infty^2(\mathbf{R}^d)$ or $D = C_\infty^1(\mathbf{R}^d)$. In order to see that $\mathcal{P}(\mathbf{R}^d)$ is closed in the norm topology of D^* for $D = C_\infty^k(\mathbf{R}^d)$ with any natural k , observe that the distance d on $\mathcal{P}(\mathbf{R}^d)$ induced by its embedding in $(C_\infty^k(\mathbf{R}^d))'$ is defined by

$$d(\mu, \eta) = \sup\{|(f, \mu - \eta)| : f \in C_\infty^2(\mathbf{R}^d), \|f\|_{C_\infty^2(\mathbf{R}^d)} \leq 1\}.$$

and hence

$$d(\mu, \eta) = \sup\{|(f, \mu - \eta)| : f \in C^2(\mathbf{R}^d), \|f\|_{C^2(\mathbf{R}^d)} \leq 1\}.$$

Consequently, convergence $\mu_n \rightarrow \mu$, $\mu_n \in \mathcal{P}(\mathbf{R}^d)$, with respect to this metric implies the convergence $(f, \mu_n) \rightarrow (f, \mu)$ for all $f \in C^k(\mathbf{R}^d)$ and hence for all $f \in C_\infty(\mathbf{R}^d)$ and for f being constants. This implies tightness of the family μ_n and that the limit $\mu \in \mathcal{P}(\mathbf{R}^d)$.

Theorem 4.1 supplies a useful criterion for condition (ii) of the previous theorem, thus yielding the following corollary.

Theorem 5.2. Under the assumption (i) of Theorem 5.1 assume instead of (ii) the following:

(ii') There exists another Banach space \tilde{D} , which is a dense subspace of D , so that all $A[\mu]$, $\mu \in M$, are uniformly bounded operators $\tilde{D} \rightarrow D$ and $D \rightarrow B$.

(iii') For any $\mu \in M$ the operator $A[\mu] : D \rightarrow B$ generates a strongly continuous semigroup $e^{tA[\mu]}$ in B with invariant core D , such that \tilde{D} is also invariant under all $e^{sA[\mu]}$, and these operators are uniformly bounded as operators in \tilde{D}, D, B , with the norms not exceeding e^{Ks} with a constant K ,

(iv') the set M is invariant under all dual semigroups $(e^{tA[\mu]})'$.

Then condition (ii) and hence the conclusion of Theorem 5.1 hold. Moreover, the operators $U^{t,s}[\mu]$ form a strongly continuous propagator of bounded operators in D .

Proof. For $\xi_t \in C_\mu([0, r], M)$, the operator curve $L_s = A[\xi_s] : D \rightarrow B$ clearly satisfies conditions (i)-(iii) of Theorem 4.1. To check its last condition (iv) we have to show that $A[\xi_t]f$ as a function $t \mapsto B$ is continuous uniformly for f from a bounded domain of D . And this follows from (22), as it implies

$$\|(A[\xi_t] - A[\xi_s])f\|_B \leq c\|\xi_t - \xi_s\|_{D^*}\|f\|_D.$$

Hence Theorem 4.1 is applicable to the curve $L_s = A[\xi_s] : D \rightarrow B$, implying condition (ii) of Theorem 5.1. \square

As a preliminary step in studying sensitivity, let us prove a simple stability result for the above nonlinear semigroups T_t with respect to the small perturbations of the generator.

Theorem 5.3. *Under the assumptions of Theorem 5.1 suppose $\xi \mapsto \tilde{A}[\xi]$ is another mapping from B^* to bounded operators $D \rightarrow B$ satisfying the same condition as A with the corresponding $\tilde{U}^{t,s}, \tilde{V}^{s,t}$ satisfying the same conditions as $U^{t,s}, V^{s,t}$. Suppose*

$$\|\tilde{A}[\xi] - A[\xi]\|_{D \rightarrow B} \leq \kappa, \quad \xi \in M \quad (26)$$

with a constant κ . Then

$$\|\tilde{T}_t(\eta) - T_t(\mu)\|_{D^*} \leq c(r, M)(\kappa + \|\mu - \eta\|_{D^*}). \quad (27)$$

Proof. As in the proof of Theorem 5.1, denoting $T_t(\mu) = \mu_t$ and $\tilde{T}_t(\eta) = \tilde{\eta}_t$ one can write

$$\mu_t - \tilde{\eta}_t = (V^{t,0}[\mu] - \tilde{V}^{t,0}[\tilde{\eta}])\mu + \tilde{V}^{t,0}[\tilde{\eta}](\mu - \eta)$$

and then

$$\sup_{s \leq t} \|\mu_s - \tilde{\eta}_s\|_{D^*} \leq c(r, M) \left(t \sup_{s \leq t} \|\mu_s - \tilde{\eta}_s\|_{D^*} + \kappa \right) + \|\mu - \eta\|_{D^*},$$

which implies (27) first for small times, and then for all finite times by iterations. \square

6 Linearized evolution around a path of a nonlinear semigroup

Both for numerical simulations and for the application to interacting particles, it is crucial to analyze the dependence of the solutions to nonlinear kinetic equations on some parameters and on the initial data. Ideally we would like to have smooth dependence.

More precisely, suppose we are given a family of operators $A^\alpha[\mu]$, depending on a real parameter α and satisfying the assumptions of Theorem 5.1 for each α . For $\mu_t^\alpha = \mu_t^\alpha(\mu_0^\alpha)$, a solution to corresponding (1) with the initial condition μ_0^α , we are interested in the derivative

$$\xi_t(\alpha) = \frac{\partial \mu_t^\alpha}{\partial \alpha}. \quad (28)$$

In this section we shall start with the analysis of the linearized evolution around a path of a nonlinear semigroup. Namely, differentiating (1) (at least formally for the moment) with respect to α yields the equation

$$\frac{d}{dt}(g, \xi_t(\alpha)) = (A^\alpha[\mu_t^\alpha]g, \xi_t(\alpha)) + (D_{\xi_t(\alpha)}A^\alpha[\mu_t^\alpha]g, \mu_t^\alpha) + \left(\frac{\partial A^\alpha[\mu_t^\alpha]}{\partial \alpha} g, \mu_t^\alpha \right), \quad (29)$$

with the initial condition

$$\xi_0 = \xi_0(\alpha) = \frac{\partial \mu_0^\alpha}{\partial \alpha}, \quad (30)$$

where

$$D_\eta A^\alpha[\mu] = \lim_{s \rightarrow 0^+} \frac{1}{s} (A^\alpha[\mu + s\eta] - A^\alpha[\mu]) \quad (31)$$

denotes the Gateaux derivatives of $A[\mu]$ as a mapping $D^* \rightarrow \mathcal{L}(D, B)$, assuming that the definition of $A^\alpha[\mu]$ can be extended to a neighborhood of M in D^* .

This section is devoted to the preliminary analysis of the solutions to equation (29). In the next section we shall explore their connections with the derivatives from the r.h.s. of (28).

Let $\tilde{D} \subset D \subset B$ be, as above, three Banach spaces such that $\|\cdot\|_{\tilde{D}} \geq \|\cdot\|_D \geq \|\cdot\|_B$, D is dense in B in the topology of B and \tilde{D} is dense in D in the topology of B ; and let M and $C_\mu([0, r], M)$ be defined as in Section 5.

Theorem 6.1. (i) Let, for each α , $\xi \mapsto A^\alpha[\xi]$ be a mapping from B^* to linear operators $A^\alpha[\xi]$ that are uniformly bounded as operators $D \rightarrow B$ and $\tilde{D} \rightarrow D$ and such that

$$\|A[\xi] - A[\eta]\|_{D \rightarrow B} \leq c\|\xi - \eta\|_{D^*}, \quad \xi, \eta \in B^* \quad (32)$$

for a constant $c > 0$.

(ii) For any α , $\mu \in M$ and $\xi \in C_\mu([0, r], M)$, let the operator curve $A^\alpha[\xi_t]$ generate a strongly continuous backward propagator of uniformly bounded linear operators $U^{t,s;\alpha}[\xi]$, $0 \leq t \leq s \leq r$, in B on the common invariant domain D , and with the dual propagator $V^{s,t;\alpha}[\xi]$ preserving the set M .

(iii) Let the propagators $\{U^{t,s;\alpha}[\xi]\}$, $t \leq s$, are strongly continuous and bounded propagators in both B and D .

(iv) Let the derivatives $\partial A^\alpha[\mu_t^\alpha]/\partial \alpha$ exist in the norm topologies of $\mathcal{L}(D, B)$ and $\mathcal{L}(\tilde{D}, D)$, and represent also bounded operators in $\mathcal{L}(D, B)$ and $\mathcal{L}(\tilde{D}, D)$.

(v) Let $A^\alpha[\mu]$ can be extended to a mapping $D^* \rightarrow \mathcal{L}(D, B)$ such that the limit in (45) exists in the norm topology of $\mathcal{L}(D, B)$ for any $\mu \in B^*$, $\xi \in D^*$. Moreover, the Gateaux derivatives $\xi \mapsto D_\xi A^\alpha[\mu]$ is continuous in μ (taken in the norm topology of B^*) and defines a bounded linear operator $D^* \rightarrow \mathcal{L}(D, B)$, that is

$$\|D_\xi A^\alpha[\mu]\|_{D \rightarrow B} \leq c\|\mu\|_{B^*}\|\xi\|_{D^*} \quad (33)$$

with a constant c .

(vi) Finally, suppose there exists a representation

$$(D_\xi A^\alpha[\mu]g, \mu) = (F^\alpha[\mu]g, \xi) \quad (34)$$

with $F^\alpha[\mu]$ being a continuous mapping $D^* \rightarrow \mathcal{L}(D, D)$.

Then, for each α , $\mu \in M$, there exists a weakly- \star continuous in D^* family of propagator $\Pi^{s,t}[\alpha, \mu]$ (constructed below) solving equation (29) in \tilde{D}^* , that is, for any $\xi_0 \in D^*$, $\xi_t^\alpha = \Pi^{s,t}[\alpha, \mu]\xi_0$ satisfies (29) for any $f \in \tilde{D}$.

Remark 4. Condition (vi) causes no trouble. In fact it follows from duality and additional weak continuity assumption on D_ξ . We shall not formulate this assumption by two reasons. (i) In case of reflexive B it is satisfied automatically. (ii) Though in case we are most interested in, that is for B^* being the space of Borel measures, B is not reflexive, in applications to Markov semigroup representation (34) again arises automatically, due to the special structure of $A[\mu]$ (of the Lévy-Khintchin type).

Remark 5. Construction of propagators from condition (ii) can naturally be carried out via Theorem 5.2, that is via T -products.

Proof. Theorem 5.1 implies that, for any α , the weak nonlinear Cauchy problem

$$\frac{d}{dt}(f, \mu_t^\alpha) = (A^\alpha[\mu_t^\alpha]f, \mu_t^\alpha), \quad \mu_0 = \mu, \quad f \in D, \quad (35)$$

is well posed in M , and its resolving semigroup T_t^α satisfies (25) uniformly in α .

Next, the equation

$$\frac{d}{dt}(g, \xi_t(\alpha)) = (A^\alpha[\mu_t^\alpha]g, \xi_t(\alpha)) + (D_{\xi_t(\alpha)}A^\alpha[\mu_t^\alpha]g, \mu_t^\alpha) \quad (36)$$

has form (17) with F_s specified by (34), i.e.

$$(F_s g, \xi) = (F^\alpha[\mu_s^\alpha]g, \xi) = (D_\xi A^\alpha[\mu_s^\alpha]g, \mu_s^\alpha).$$

From (33) it follows that

$$\|F_s\|_{D \rightarrow D} = \sup_{\|g\|_D \leq 1} \sup_{\|\xi\|_{D^*} \leq 1} (D_\xi A^\alpha[\mu_s^\alpha]g, \mu_s^\alpha) \leq c \|\xi\|_{D^*} \|\mu\|_{B^*}^2, \quad (37)$$

which is uniformly bounded for $\mu_s^\alpha \in M$. Consequently, Theorem 3.2 yields a construction of the strongly continuous family $\{\Phi^{t,r}\}$ in D such that its dual propagator $\{\Psi^{r,t} = (\Phi^{t,r})'\}$ solves the Cauchy problem for equation (36).

By the Duhamel principle, the solution to equation (29) for $r \geq t$ with the initial condition ξ_t can be written as

$$(g, \Pi^{r,t}[\alpha, \mu]\xi_t) = (\Phi^{t,r}[\alpha, \mu]g, \xi_t) + \int_t^r \left(\frac{\partial A^\alpha[\mu_s^\alpha]}{\partial \alpha} \Phi^{s,r}[\alpha, \mu]g, \mu_s^\alpha \right) ds. \quad (38)$$

□

Theorem 6.2. *Under the assumptions of Theorem 6.1, assume additionally that the backward propagators $\{U^{t,s;\alpha}[\xi]\}$, $t \leq s$, represent strongly continuous bounded propagators also in \tilde{D} (and hence, by the last statement of Theorem 2.3, the family $A^\alpha[\xi_t] : D \rightarrow B$ also generates $\{U^{t,s;\alpha}\}$, as a propagator in D , on \tilde{D}). Then, for each $\alpha, \mu \in M, \xi_0 \in D^*$, the curve $\Pi^{s,t}[\alpha, \mu]\xi_0$ represents the unique weakly- \star continuous in D^* solution to equation (29) in \tilde{D}^* .*

Proof. This is a straightforward extension of Theorem 6.1, obtained by taking into account the simple arguments given before Theorem 3.2. □

We shall not further pay attention to somewhat complicated details arising under the conditions of Theorem 6.1, but will use more natural conditions of Theorem 6.2.

We complete this section by an additional stability result for $\Pi^{s,t}$.

Theorem 6.3. *Under the assumptions of Theorem 6.2, suppose that*

(i) *in addition to (32) and (33), one has the same properties for the pair (\tilde{D}, D) , i.e.*

$$\|A[\xi] - A[\eta]\|_{\tilde{D} \rightarrow D} \leq c \|\xi - \eta\|_{D^*}, \quad \xi, \eta \in B^*, \quad (39)$$

$$\|D_\xi A^\alpha[\mu]\|_{\tilde{D} \rightarrow D} \leq c \|\mu\|_{B^*} \|\xi\|_{D^*}, \quad (40)$$

(ii) derivatives of $A^\alpha[\mu]$ are Lipschitz in the norm-topology of D^\star , more precisely:

$$\left\| \frac{\partial A^\alpha[\mu]}{\partial \alpha} - \frac{\partial A^\alpha[\eta]}{\partial \alpha} \right\|_{\tilde{D} \rightarrow D} \leq c \|\mu - \eta\|_{D^\star}, \quad (41)$$

$$\|D_\xi(A^\alpha[\mu] - A^\alpha[\nu])\|_{D \rightarrow B} \leq c \|\mu - \eta\|_{D^\star} \|\xi\|_{D^\star}. \quad (42)$$

Suppose now that $\mu_0^\alpha(n) \rightarrow \mu_0^\alpha$ in the norm-topology of D^\star , as $n \rightarrow \infty$ for each α . Then $\Pi^{s,t}[\alpha, \mu_0^\alpha(n)]\xi_0 \rightarrow \Pi^{s,t}[\alpha, \mu_0^\alpha]\xi_0$ weakly- \star in D^\star and in the norm topology of \tilde{D}^\star .

Proof. We shall use the notation for propagators introduced above adding dependence on n for all objects constructed from $\mu_0^\alpha(n)$.

By (25) we conclude that $T_t^\alpha \mu_0^\alpha(n) \rightarrow T_t^\alpha \mu_0^\alpha$, as $n \rightarrow \infty$, in the norm-topology of D^\star uniformly in t, α . Hence, by (39) and Theorem 2.4 (applied to the pair of spaces (\tilde{D}, D)),

$$U^{t,s;\alpha}[T_t^\alpha \mu_0^\alpha(n)] \rightarrow U^{t,s;\alpha}[T_t^\alpha \mu_0^\alpha]$$

in the norm-topology of $\mathcal{L}(D, D)$. Similarly, by (40) and (41),

$$\begin{aligned} & |(D_\xi A^\alpha[\mu]g, \mu) - (D_\xi A^\alpha[\eta]g, \eta)| \\ & \leq |(D_\xi(A^\alpha[\mu] - A^\alpha[\eta])g, \mu)| + |(D_\xi A^\alpha[\eta]g, \mu - \eta)| \\ & \leq c \|\mu - \eta\|_{D^\star} \|g\|_{\tilde{D}} \|\xi\|_{D^\star} (\|\mu\|_{B^\star} + \|\eta\|_{B^\star}), \end{aligned}$$

so that

$$\|F_s[\mu]g - F_s[\eta]g\|_D \leq c \|\mu - \eta\|_{D^\star} \|g\|_{\tilde{D}} (\|\mu\|_{B^\star} + \|\eta\|_{B^\star}).$$

and thus by Theorem 2.4,

$$\Phi^{t,s}[\alpha, T_t^\alpha \mu_0^\alpha(n)] \rightarrow \Phi^{t,s}[\alpha, T_t^\alpha \mu_0^\alpha], \quad n \rightarrow \infty,$$

in the norm-topology of $\mathcal{L}(D, D)$. Consequently, again by Theorem 2.4,

$$\Psi^{s,t}[\alpha, T_t^\alpha \mu_0^\alpha(n)]\xi \rightarrow \Psi^{s,t}[\alpha, T_t^\alpha \mu_0^\alpha]\xi$$

weakly- \star in D^\star and in the norm-topology of \tilde{D}^\star , for any $\xi \in D^\star$.

Finally, from (38) it follows that

$$\begin{aligned} & (g, \Pi^{r,0}[\alpha, \mu](n)\xi - \Pi^{r,0}[\alpha, \mu]\xi) \\ & = ((\Phi^{0,r}(n) - \Phi^{0,r})g, \xi) + \int_0^r \left(\frac{\partial A^\alpha[\mu_s^\alpha(n)]}{\partial \alpha} (\Phi^{s,r}(n) - \Phi^{s,r})g, \mu_s^\alpha \right) \\ & + \int_0^r \left(\frac{\partial A^\alpha[\mu_s^\alpha(n)]}{\partial \alpha} \Phi^{s,r}(n)g, \mu_s^\alpha(n) - \mu_s^\alpha \right) + \int_0^r \left(\left(\frac{\partial A^\alpha[\mu_s^\alpha(n)]}{\partial \alpha} - \frac{\partial A^\alpha[\mu_s^\alpha]}{\partial \alpha} \right) \Phi^{s,r}g, \mu_s^\alpha \right), \end{aligned}$$

which allows one to conclude that

$$\|\Pi^{r,0}[\alpha, \mu](n)\xi - \Pi^{r,0}[\alpha, \mu]\xi\|_{\tilde{D}^\star} \rightarrow 0,$$

as $n \rightarrow \infty$, as required. \square

7 Sensitivity analysis for nonlinear propagators

Our final question is whether the solution ξ_t constructed in Theorem 6.1 does in fact yield the derivative (28). The difference with the standard case, discussed in textbooks on ODE in Banach spaces, lies in the fact that the solution to the linearized equation (29) exists in a different space than the nonlinear curve μ_t itself.

Theorem 7.1. *Under the assumptions of Theorem 6.3, let $\xi_0 = \xi \in B^*$ and is defined by (30), where the derivative exists in the norm-topology of \tilde{D}^* and weakly- \star in D^* . Then the unique solution $\xi_t[\alpha] = \Pi^{t,0}[\alpha, \mu_0^{\alpha}] \xi$ of equation (29) constructed in the Theorem 6.2 satisfies (28), where the derivative exists in the norm-topology of \tilde{D}^* and weakly- \star in D^* .*

Proof. The main idea is to approximate A_s^{α} by bounded operators, use the standard sensitivity theory for vector valued ODE and then obtain the required result by passing to the limit. To carry out this program, let us pick up a family of operators $A_s^{\alpha}(n)$, $n = 1, 2, \dots$, bounded in B and D , that satisfy all the same conditions as A_s^{α} and such that $\|(A_s^{\alpha}(n) - A_s^{\alpha})g\|_B \rightarrow 0$ for all $g \in D$ and uniformly for all α and g from bounded subsets of \tilde{D} . As such approximation, one can use either standard Yosida approximation (which is convenient in abstract setting) or, in case of the generators of Feller Markov processes, generators of approximating pure-jump Markov processes. As in the proof of Theorem 6.3, we shall use the notation for propagators introduced in the previous section adding dependence on n for all objects constructed from $A_s^{\alpha}(n)$.

Since $(A_s^{\alpha}(n))'$ are bounded linear operators in B^* and D^* , the equation for μ_t and ξ_t are both well posed in the strong sense in both B^* and D^* . Hence the standard result on the differentiation with respect to initial data is applicable (see e.g. [13] or Appendix D in [10]) leading to the conclusion that $\xi_t[\alpha](n)$ represent the derivatives of $\mu_t^{\alpha}(n)$ in both B^* and D^* .

Consequently

$$\mu_t^{\alpha}(n) - \mu_t^{\alpha_0}(n) = \int_{\alpha_0}^{\alpha} \xi_t[\beta](n) d\beta \quad (43)$$

holds as an equation in D^* (and in B^* whenever $\xi \in B^*$).

Using Theorem 5.3 we deduce the convergence of $\mu_t^{\alpha}(n)$ to μ_t^{α} in the norm-topology of D^* . Consequently, using Theorem 6.3 we can deduce the convergence of $\xi_t^{\alpha}(n)$ to ξ_t^{α} in the norm-topology of \tilde{D}^* . Hence, we can pass to the limit $n \rightarrow \infty$ in equation (43) in the norm topology of \tilde{D}^* yielding the equation

$$\mu_t^{\alpha} - \mu_t^{\alpha_0} = \int_{\alpha_0}^{\alpha} \xi_t[\beta] d\beta, \quad (44)$$

where all objects are well defined in $(C_{\infty}^1(\mathbf{R}^d))^*$.

This equation together with continuous dependence of ξ_t on α (which is proved in literally the same way as continuous dependence on μ in Theorem 6.3) implies (28) in the sense required. \square

Applying Theorem 7.1 for the case of A_s not depending on any additional parameter, we obtain directly the smooth dependence of the nonlinear evolution μ_t on the initial data. Namely, for $\mu_t = \mu_t(\mu_0)$, a solution to (1) with the initial condition μ_0 , we can define the Gateaux derivatives

$$\xi_t(\mu_0, \xi) = D_{\xi} \mu_t(\mu_0) = \lim_{s \rightarrow 0^+} \frac{1}{s} (\mu_t(\mu_0 + s\xi) - \mu_t(\mu_0)) \quad (45)$$

Differentiating (1) with respect to initial data yields

$$\frac{d}{dt}(g, \xi_t(\mu_0, \xi)) = (A[\mu_t]g, \xi_t(\mu_0, \xi)) + (D_{\xi_t(\mu_0, \xi)}A[\mu_t]g, \mu_t), \quad (46)$$

which represents a simple particular case of equation (29). Hence, Theorem 7.1 implies that, under the assumptions of this theorem (that do not involve the dependence on α), the derivative (45) does exist and is given by the unique solution to equation (46) with the initial condition $\xi_0 = \xi$. However, this existence and well-posedness hold weakly- \star in \tilde{D}^* , not in B^* , as the nonlinear evolution itself.

8 Back to nonlinear Markov semigroups

We developed the theory in the most abstract form, for general nonlinear evolutions in Banach spaces, not even using positivity. This unified exposition allows one to obtain various concrete evolutions as a direct consequence of one general result. The main application we have in mind concerns the families $A[\mu]$ of the Lévy-Kchintchin type form (with variable coefficients):

$$\begin{aligned} A[\mu]u(x) &= \frac{1}{2}(G_\mu(x)\nabla, \nabla)u(x) + (b_\mu(x), \nabla u(x)) \\ &+ \int [u(x+y) - u(x) - (y, \nabla u(x))\mathbf{1}_{B_1}(y)]\nu_\mu(x, dy), \end{aligned} \quad (47)$$

where $\nu_\mu(x, \cdot)$ is a Lévy measure for all $x \in \mathbf{R}^d, \mu \in \mathcal{P}(\mathbf{R}^d)$. The basic examples were given in the introduction.

Applied to *nonlinear Lévy process* specified by the families (4), our general results yield the following.

Theorem 8.1. *Suppose the coefficients of a family (4) depend on μ Lipschitz continuously in the norm of the Banach space $(C_\infty^2(\mathbf{R}^d))'$ dual to $C_\infty^2(\mathbf{R}^d)$, i.e.*

$$\begin{aligned} \|G(\mu) - G(\eta)\| + \|b(\mu) - b(\eta)\| + \int \min(1, |y|^2)|\nu(\mu, dy) - \nu(\eta, dy)| \\ \leq \kappa \|\mu - \eta\|_{(C_\infty^2(\mathbf{R}^d))'} = \kappa \sup_{\|f\|_{C_\infty^2(\mathbf{R}^d)} \leq 1} |(f, \mu - \eta)| \end{aligned} \quad (48)$$

with constant κ . Then there exists a unique nonlinear Lévy semigroup generated by A_μ , and hence a unique nonlinear Lévy process.

Proof. The well-posedness of all intermediate propagators is obvious in case of Lévy processes, because they are constructed via Fourier transform, literally like Lévy semigroup (details are given in [10]). Of course here $M = \mathcal{P}(\mathbf{R}^d)$, $D = C_\infty^2(\mathbf{R}^d)$, $\tilde{D} = C_\infty^4(\mathbf{R}^d)$. \square

Remark 6. *Condition (48) is not at all weird. It is satisfied, for instance, when the coefficients G, b, ν depend on μ via certain integrals (possibly multiple) with smooth enough densities, i.e. in a way that is usually met in applications.*

Applied to processes of order at most one specified by the families (5), our general results yield the following.

Theorem 8.2. *Assume that for any $\mu \in \mathcal{P}(\mathbf{R}^d)$, $b(\cdot, \mu) \in C^1(\mathbf{R}^d)$ and $\nabla \nu(x, \mu, dy)$ (gradient with respect to x) exists in the weak sense as a signed measure and depends weakly continuous on x . Let the following conditions hold.*

(i) *boundedness:*

$$\sup_{x, \mu} \int \min(1, |y|) \nu(x, \mu, dy) < \infty, \quad \sup_{x, \mu} \int \min(1, |y|) |\nabla \nu(x, \mu, dy)| < \infty, \quad (49)$$

(ii) *tightness: for any $\epsilon > 0$ there exists a $K > 0$ such that*

$$\sup_{x, \mu} \int_{\mathbf{R}^d \setminus B_K} \nu(x, \mu, dy) < \epsilon, \quad \sup_{x, \mu} \int_{\mathbf{R}^d \setminus B_K} |\nabla \nu(x, \mu, dy)| < \epsilon, \quad (50)$$

$$\sup_{x, \mu} \int_{B_{1/K}} |y| \nu(x, \mu, dy) < \epsilon, \quad (51)$$

(iii) *Lipschitz continuity:*

$$\sup_x \int \min(1, |y|) |\nu(x, \mu_1, dy) - \nu(x, \mu_2, dy)| \leq c \|\mu_1 - \mu_2\|_{(C_\infty^1(\mathbf{R}^d))^*}, \quad (52)$$

$$\sup_x |b(x, \mu_1) - b(x, \mu_2)| \leq c \|\mu_1 - \mu_2\|_{(C_\infty^1(\mathbf{R}^d))^*} \quad (53)$$

uniformly for bounded μ_1, μ_2 .

Then the weak nonlinear Cauchy problem (1) with A_μ given by (5) is well posed, i.e. for any $\mu \in \mathcal{M}(\mathbf{R}^d)$ it has a unique solution $T_t(\mu) \in \mathcal{M}(\mathbf{R}^d)$ (so that (5) holds for all $g \in C_\infty^1(\mathbf{R}^d)$) preserving the norm, and the transformations T_t of $\mathcal{P}(\mathbf{R}^d)$, $t \geq 0$, form a semigroup depending Lipschitz continuously on time t and the initial data in the norm of $(C_\infty^1(\mathbf{R}^d))^$.*

Proof. Here we use $M = \mathcal{P}(\mathbf{R}^d)$, $D = C_\infty^1(\mathbf{R}^d)$, $\tilde{D} = C_\infty^2(\mathbf{R}^d)$. The corresponding auxiliary propagators required in Theorem 2.1 are constructed in [10] (Chapter 4) and [11] (Chapter 5). \square

In both cases above, straightforward additional smoothness assumptions on the coefficients of the generator yield smoothness with respect to parameters and/or initial data via Theorem 7.1.

Similarly one gets the well-posedness for mixtures of nonlinear diffusions and stable-like processes given by (21) with coefficients depending on distribution μ . Our theory also applies to nonlinear stable-like processes on manifolds, see [10] (Section 11.4), and to nonlinear dynamic quantum semigroups, see [10] (Section 11.3).

Let us stress again, referring to [10], [7] and [6], that the first and second derivatives of nonlinear Markov semigroups with respect to initial data (for simplicity, we dealt only with the first derivative here) describe the dynamic law of large numbers for interacting particle systems and the corresponding central limit theorem for fluctuations, respectively.

References

- [1] D. Applebaum. Lévy Processes and Stochastic Calculus. Cambridge studies in advanced mathematics, v. 93. Cambridge Univ. Press, 2004.
- [2] I. Bailleul. Sensitivity for Smoluchovski equation. Preprint 2009. <http://www.statslab.cam.ac.uk/~ismael/files/Sensitivity.pdf>
- [3] T.D. Frank. Nonlinear Markov processes. Phys. Lett. A **372:25** (2008), 4553-4555.
- [4] H. Guérin, S. Méléard, E. Nualart. Estimates for the density of a nonlinear Landau process. Journal of Functional Analysis **238** (2006), 649-677.
- [5] V. N. Kolokoltsov. On the regularity of solutions to the spatially homogeneous Boltzmann equation with polynomially growing collision kernel. Advanced Studies in Contemp. Math. **12** (2006), 9-38.
- [6] V. N. Kolokoltsov. Nonlinear Markov Semigroups and Interacting Lévy Type Processes. Journ. Stat. Physics **126:3** (2007), 585-642.
- [7] V. N. Kolokoltsov. The central limit theorem for the Smoluchovski coagulation model. arXiv:0708.0329v1[math.PR] 2007. Prob. Theory Relat. Fields 146: 1 (2010), 87-153.
- [8] V. N. Kolokoltsov. The Lévy-Khintchine type operators with variable Lipschitz continuous coefficients generate linear or nonlinear Markov processes and semigroups. Probability Theory Related Fields. Online First DOI 10.1007/s00440-010-0293-8.
- [9] V. N. Kolokoltsov. Stochastic integrals and SDE driven by nonlinear Lévy noise. In D. Crisan (Ed.) "Stochastic Analysis in 2010", Springer 2011, p. 227-242.
- [10] V. N. Kolokoltsov. Nonlinear Markov processes and kinetic equations. Cambridge University Press, 2010.
- [11] V. N. Kolokoltsov. Markov Processes, Semigroups and Generators. DeGruyter, 2011.
- [12] A. E. Kyprianou. Introductory Lectures on Fluctuations of Lévy Processes with Applications. Universitext. Springer 2006.
- [13] R.H. Martin. Nonlinear operators and differential equations in Banach spaces. New York, 1976.
- [14] D.W. Stroock. Markov Processes from K. Ito's Perspective. Annals of Mathematics Studies. Princeton University Press 2003.
- [15] M. Zak. Dynamics of intelligent systems. Int. J. Theor. Phys. **39:8** (2000), 2107-2140.
- [16] M. Zak. Quantum Evolution as a Nonlinear Markov Process. Foundations of Physics Letters **15:3** (2002), 229-243.