Abstract McKean–Vlasov and Hamilton–Jacobi–Bellman Equations, Their Fractional Versions and Related Forward–Backward Systems on Riemannian Manifolds

V. N. Kolokoltsov\textsuperscript{a,b} and M. S. Troeva\textsuperscript{c}

Received February 24, 2021; revised April 14, 2021; accepted July 30, 2021

Abstract—We introduce a class of abstract nonlinear fractional pseudo-differential equations in Banach spaces that includes both the McKean–Vlasov type equations describing nonlinear Markov processes and the Hamilton–Jacobi–Bellman–Isaacs equation of stochastic control and games. This allows for a unified analysis of these equations, which leads to an effective theory of coupled forward–backward systems (forward McKean–Vlasov evolution and backward Hamilton–Jacobi–Bellman–Isaacs evolution) that are central to the modern theory of mean-field games.

DOI: 10.1134/S0081543821050096

1. INTRODUCTION

We introduce a class of abstract nonlinear fractional pseudo-differential equations in Banach spaces that includes both the McKean–Vlasov type equations describing nonlinear Markov processes and the Hamilton–Jacobi–Bellman–Isaacs equation of stochastic control and games, thus allowing for a unified analysis of these equations. Looking at these equations as evolving in dual Banach triples allows us to recast directly the properties of one type to the properties of another type, which leads to an effective theory of coupled forward–backward systems (forward McKean–Vlasov evolution and backward Hamilton–Jacobi–Bellman–Isaacs evolution) that are central to the modern theory of mean-field games. We are working with the mild solutions to the fractional nonlinear equations that are based on the Zolotarev integral representation for the Mittag-Leffler functions. The abstract setting developed allows us to include in our analysis the related nonlinear fractional equations and forward–backward systems on manifolds, which yields results that are possibly new even for classical (not fractional) equations. We obtain well-posedness results for these equations.

The present work is a continuation of the studies by the authors in [29–31].

We shall analyze the following problems: nonlinear Cauchy problems of the form

\[ \dot{b}(t) = Ab(t) + H(t, b(t), D_b(t), \alpha), \quad b(a) = Y, \quad t \geq a, \quad (1.1) \]

where \( A \) and \( D_1, \ldots, D_n \) are unbounded linear operators in a Banach space \( B \), \( D = (D_1, \ldots, D_n) \), \( \alpha \) is a parameter from another Banach space \( B^\text{par} \) and \( H \) is a continuous mapping \( \mathbb{R} \times B \times B^\alpha \times B^\text{par} \to B \); their fractional counterparts

\[ D^\beta_{a+} b(t) = Ab(t) + H(t, b(t), D_b(t), \alpha), \quad b(a) = Y, \quad t \geq a, \quad (1.2) \]

\textsuperscript{a} Department of Statistics, University of Warwick, Coventry, CV4 7AL, UK.
\textsuperscript{b} National Research University Higher School of Economics, Pokrovskii bul. 11, Moscow, 109028 Russia.
\textsuperscript{c} North-Eastern Federal University, ul. Belinskogo 58, Yakutsk, 677000 Russia.

E-mail addresses: v.kolokoltsov@warwick.ac.uk (V. N. Kolokoltsov), troeva@mail.ru (M. S. Troeva).
where \( D_{a+s}^\beta \) is the Caputo–Dzherbashyan fractional derivative of order \( \beta \in (0,1) \),
\[
D_{a+s}^\beta b(t) = \frac{1}{\Gamma(-\beta)} \int_0^{t-a} \frac{b(t-z) - b(t)}{z^{1+\beta}} \, dz + \frac{b(t) - b(a)}{\Gamma(1-\beta)(t-a)^\beta};
\]  
(1.3)

anticipating versions of these equations (where \( H \) depends additionally on the future values of \( b(s) \)); and the forward–backward systems of coupled equations of this type, which represent the main class of systems studied in the modern theory of mean-field games.

**Remark 1.1.** The above explicit formula for the fractional derivative is a consequence of its more standard definition as \( D_{a+s}^\beta b(t) = I_a^{1-\beta}(d/dt)b(t) \) via the fractional integral \( I^\beta \). Representation (1.3) is convenient for constructing generalized fractional derivatives by methods of probability theory. The corresponding generalizations will be presented in subsequent works of the authors.

Our main examples concern the case when \( B \) is a space of functions on \( \mathbb{R}^d \) and \( D \) is the gradient (derivative) operator. Specifically, the fractional Hamilton–Jacobi–Bellman–Isaacs equation of controlled Markov processes (with an external parameter) is an equation of the form
\[
D_{a+s}^\beta f(t,x) = Af(t,x) + H\left(t,x,f(t,x),\frac{\partial f}{\partial x}(t,x),\alpha\right),
\]  
(1.4)

for which the most natural Banach space is \( B = C_\infty(\mathbb{R}^d) \), the Banach space of continuous functions \( f: \mathbb{R}^d \to \mathbb{R} \) tending to zero at infinity equipped with the sup-norm.

The Hamiltonian \( H \) arising from optimal control usually even does not depend explicitly on \( f \), just on its gradient, and it writes down as
\[
H(t,x,f,p,\alpha) = H(t,x,p,\alpha) = \sup_{u \in U}[J(t,x,u,\alpha) + g(t,x,u,\alpha)p],
\]  
(1.5)

with some functions \( J \) and \( g \), where \( U \) is a compact set of controls (or with inf sup instead of just sup in case of Isaacs equations). For such \( H \) the fractional equation (1.4) was derived in [32] as a Bellman equation for optimal control of scaled limits of continuous time random walks.

The fractional versions of the McKean–Vlasov type equations (describing nonlinear Markov processes in the sense of [22]) are quasi-linear equations of the type
\[
D_{a+s}^\beta f(t,x) = A^*f(t,x) + \sum_{j=1}^d h_j(t,x,\{f(t,\cdot)\},\alpha) \frac{\partial f}{\partial x_j}(t,x),
\]  
(1.6)

for which the most natural Banach space is \( L_1(\mathbb{R}^d) \) (or the space of Borel measures on \( \mathbb{R}^d \)). In these equations, \( A \) is the generator of a Feller process in \( \mathbb{R}^d \) and \( A^* \) is its dual operator. While \( H \) in (1.4) depends on the pointwise values of \( f \), the functions \( h_j \) in (1.6) usually depend on some integrals of \( f \).

The abstract framework of equations (1.2) allows one to treat these cases in a unified way, as well as to include in the theory in a more or less straightforward way important new cases, for instance, fractional Hamilton–Jacobi–Bellman–Isaacs or McKean–Vlasov type equations on manifolds.

Fully analogous results, of course, hold for the backward versions of the Cauchy problems above, namely, for the problems
\[
\dot{b}(t) = -Ab(t) + H(t,b(t),Db(t),\alpha), \quad b(T) = Y, \quad t \leq T,
\]  
(1.7)

and their fractional counterparts
\[
D^\beta_{T-s} b(t) = -Ab(t) + H(t,b(t),Db(t),\alpha), \quad b(T) = Y, \quad t \leq T,
\]  
(1.8)
where $D_{T-a}^\beta$ is the right Caputo–Dzherbashyan fractional derivative of order $\beta \in (0, 1)$:

$$D_{T-a}^\beta b(t) = \frac{1}{\Gamma(-\beta)} \int_0^{T-t} \frac{b(t + z) - b(t)}{z^{1+\beta}} \, dz + \frac{b(t) - b(T)}{\Gamma(1-\beta)(T-t)^\beta}. \quad (1.9)$$

The content of the paper is as follows. In Sections 2 and 3, we recall preliminaries from the theory of the Mittag-Leffler functions and fixed point principles. In Section 4 we present results on the well-posedness of equations (1.1) and (1.2) in the sense of mild solutions. In Section 5 we prove the local well-posedness for the anticipating versions of equations (1.1) and (1.2) for sufficiently small $T - a$. In Sections 6 and 7 we present results on the well-posedness of the abstract coupled forward–backward system and the fractional version of this forward–backward system.

In Section 8 we specify an abstract model for coupled forward–backward systems and those fractional analogs, looking at these equations as evolving in dual Banach triples. We obtain local well-posedness results for the fractional coupled forward–backward system consisting of the coupled McKean–Vlasov (forward) and Hamilton–Jacobi–Bellman (backward) equations for $t \in [a, T]$.

In Sections 9 and 10, the abstract setting developed in Section 8 allows us to prove well-posedness results for nonlinear fractional equations and fractional coupled forward–backward systems on manifolds.

We shall work everywhere with mild solutions, but the standard arguments allow one to get natural conditions ensuring that mild solutions are in fact classical (see [25, Theorem 6.1.3]).

Note that fractional equations have become a popular subject of research due to their wide applicability in various fields of natural sciences (see, e.g., [15, 36, 40, 41, 44]). We specially note the works [1] and [3], where the generalized Euler–Lagrange equations and linear fractional differential equations are considered. Numerical methods for fractional equations are presented in [6, 27]. The problem with two-sided fractional derivatives was analyzed in [16]. Generalized fractional equations were considered in [18, 19, 24].

As mentioned above, the theory of coupled forward–backward systems (forward McKean–Vlasov evolution and backward Hamilton–Jacobi–Bellman–Isaacs evolution) is central to the modern theory of mean-field games. Mean-field games represent an intensively developing part of game theory. Mean-field games were initiated by Lasry and Lions [35] and Huang, Malhame and Caines [17]. For recent surveys see, e.g., [4, 7–9, 13, 28, 39] and numerous references therein. Forward–backward systems on manifolds arise in the analysis of quantum mean-field games [26].

It is worth mentioning some works in which Hamilton–Jacobi–Bellman (HJB) equations on manifolds were investigated [5, 12, 38, 45]. These papers are mainly devoted to the viscosity solutions of the HJB equation on Riemannian manifolds, while the main focus of our work is on mild solutions.

Fractional forward–backward systems arise in the analysis of fractional mean-field games (see, e.g., [10, 43]).

2. PRELIMINARIES ON THE MITTAG-LEFFLER FUNCTIONS

By $E_\beta(x)$ we denote the standard Mittag-Leffler function of index $\beta$:

$$E_\beta(x) = \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(\beta k + 1)}.$$

For our purpose the most convenient formula for the Mittag-Leffler function is its integral representation (Zolotarev formula, or Zolotarev–Pollard formula)

$$E_\beta(s) = \frac{1}{\beta} \int_0^\infty e^{sx} x^{-1/\beta} G_\beta(1, x^{-1/\beta}) \, dx, \quad (2.1)$$
where
\[ G_\beta(t, x) = \frac{1}{\pi} \Re \int_0^\infty \exp\{ipx - tp^\beta e^{i\pi\beta/2}\} \, dp \]
is the heat kernel (solution with the Dirac initial condition) of the equation
\[ \frac{\partial G}{\partial t}(t, x) = -\frac{\partial^\beta}{\partial x^\beta} G(t, x), \]
or, in probabilistic terms, the transition probability density of the stable Lévy subordinator of index \( \beta \). The convenience of this formula is due to the fact that it allows one to define \( E_\beta(A) \) for an operator \( A \) whenever \( A \) generates a semigroup, so that \( e^{At} \) is well defined.

**Remark 2.1.** For the Mittag-Leffler function, there exists a representation close to (2.1) which is given in terms of the Wright function (see, e.g., the books [15, 44]).

From (2.1) it follows that
\[ E_\beta'(s) = \frac{1}{\beta} \int_0^\infty e^{sx} x^{-1/\beta} G_\beta(1, x^{-1/\beta}) \, dx, \tag{2.2} \]
so that the integral on right-hand side is finite.

We also need the well-known formula for the Mellin transform of \( G_\beta \):
\[ \int_0^\infty x^{-\omega} G_\beta(1, x^{-1/\beta}) \, dx = \frac{\Gamma(1 - \omega + 1/\beta)}{\Gamma(\beta - \beta \omega + 1)}, \tag{2.3} \]
which is valid for \( \omega < 1 + 1/\beta \) (for a proof see, e.g., [25, Proposition 8.1.1]).

3. **PRELIMINARIES ON THE FIXED-POINT PRINCIPLE FOR INTEGRAL CURVES**

For a Banach space \( B \) and \( \tau < t \), we denote by \( C([\tau, t], B) \) the Banach space of continuous functions \( f : [\tau, t] \to B \) with the norm
\[ \|f\|_{C([\tau, t], B)} = \sup_{s \in [\tau, t]} \|f(s)\|_B \]
and by \( C_Y([\tau, t], B) \) its closed subset consisting of functions \( f \) such that \( f(\tau) = Y \), which is a complete metric space under the induced topology.

For a closed convex subset \( M \) of \( B \), \( C_Y([\tau, t], M) \) denotes a convex subset of \( C_Y([\tau, t], B) \) of functions with values in \( M \).

The following result is Theorem 2.1.3 from [25]. It is a version of the fixed-point principle specifically tailored to be used for nonlinear diffusions and fractional equations.

**Theorem 3.1.** Suppose that for any \( Y \in M \) and \( \alpha \in B^{\text{par}} \), with \( B^{\text{par}} \) another Banach space, a mapping \( \Phi_{Y, \alpha} : C([\tau, T], M) \to C_Y([\tau, T], M) \) is given with some \( T > \tau \) such that
\[
\begin{align*}
\| \Phi_{Y, \alpha}(\mu_1^1)(t) - \Phi_{Y, \alpha}(\mu_2^1)(t) \| & \leq L(Y) \int_{\tau}^{t} (t - s)^{-\omega} \| \mu_1^1 - \mu_2^1 \|_{C([s, s], B)} \, ds, \\
\| \Phi_{Y, \alpha}(\mu_1^2)(t) - \Phi_{Y, \alpha}(\mu_2^2)(t) \| & \leq \kappa \| Y_1 - Y_2 \| + \kappa_1 \| \alpha_1 - \alpha_2 \|,
\end{align*}
\]
for all \( t \in [\tau, T], \mu, \mu_1^1, \mu_2^1 \in C([\tau, T], M), \), \( Y_1, Y_2 \in M, \alpha, \alpha_1, \alpha_2 \in B^{\text{par}}, \) some constants \( \kappa, \kappa_1 \geq 0, \omega \in [0, 1], \) and a continuous function \( L \) on \( M \).
Then for any $Y \in M$ and $\alpha \in B_{\text{par}}$ the mapping $\Phi_{Y,\alpha}$ has a unique fixed point $\mu_{\tau}(Y, \alpha)$ in $C_\tau([\tau, T], M)$. Moreover, if $\omega > 0$, then for all $t \in [\tau, T]$
\[ \|\mu_{\tau}(Y, \alpha) - Y\| \leq E_{1-\omega}(L(Y)\Gamma(1-\omega)(t-\tau)^{1-\omega})\|\Phi_{Y,\alpha}(Y)(t) - Y\| \] (3.2)
and the fixed points $\mu_{\tau}(Y_1, \alpha_1)$ and $\mu_{\tau}(Y_2, \alpha_2)$ with different initial data $Y_1$, $Y_2$ and parameters $\alpha_1$, $\alpha_2$ enjoy the estimate (for $j = 1, 2$
\[ \|\mu_{\tau}(Y_1, \alpha_1) - \mu_{\tau}(Y_2, \alpha_2)\| \leq (\omega_1\|Y_1 - Y_2\| + \omega_1\|\alpha_1 - \alpha_2\|)E_{1-\omega}(L(Y_j)\Gamma(1-\omega)(t-\tau)^{1-\omega}). \] (3.3)

If $\omega = 0$, these estimates simplify to
\[ \|\mu_{\tau}(Y, \alpha) - Y\| \leq e^{(t-\tau)L(Y)}\|\Phi_{Y,\alpha}(Y)(t) - Y\|, \] (3.4)
\[ \|\mu_{\tau}(Y_1, \alpha_1) - \mu_{\tau}(Y_2, \alpha_2)\| \leq (\omega_1\|Y_1 - Y_2\| + \omega_1\|\alpha_1 - \alpha_2\|)\exp\{(t-\tau)\min\{L(Y_1), L(Y_2)\}\}. \] (3.5)

Here $\Phi_{Y,\alpha}(Y)$ means the application of the mapping $\Phi_{Y,\alpha}$ to the function that is identically equal to $Y$.

4. ABSTRACT FRACTIONAL MCKEAN–VLASOV AND HAMILTON–JACOBI–BELLMAN EQUATIONS

For two Banach spaces $B$ and $C$ we denote by $L(B, C)$ the Banach space of bounded linear operators $B \rightarrow C$ with the usual operator norm denoted by $\|\cdot\|_{B \rightarrow C}$.

A sequence of embedded Banach spaces $B_2 \subset B_1 \subset B$ with norms $\|\cdot\|_2$, $\|\cdot\|_1$ and $\|\cdot\|$, respectively, will be referred to as a Banach triple (of embedded spaces) or a Banach tower of order 3 if the norms are ordered, $\|\cdot\|_2 \geq \|\cdot\|_1 \geq \|\cdot\|$, and $B_2$ is dense in $B_1$ in the topology of $B_1$, while $B_1$ is dense in $B$ in the topology of $B$. The following setting will play the key role in this paper.

**Conditions A.** (i) Let $B_2 \subset B_1 \subset B$ be a Banach triple with norms $\|\cdot\|_2$, $\|\cdot\|_1$ and $\|\cdot\|$, respectively, and let
\[ D_i \in \mathcal{L}(B_1, B) \cap \mathcal{L}(B_2, B_1), \quad i = 1, \ldots, n. \]
Without loss of generality we assume that the norms of all $D_j$ are bounded by 1 in both $\mathcal{L}(B_1, B)$ and $\mathcal{L}(B_2, B_1)$ (which is usually the case in applications below).

(ii) Let $A \in \mathcal{L}(B_2, B)$ and $A$ generate a strongly continuous semigroup $e^{At}$ in both $B$ and $B_1$, so that
\[ \|e^{At}\|_{B \rightarrow B} \leq Me^{mt}, \quad \|e^{At}\|_{B_1 \rightarrow B_1} \leq M_1e^{m_1t} \] (4.1)
with some nonnegative constants $M, m$ and $M_1, m_1$, and let $B_2$ be an invariant core of the operator $A$ for this semigroup in $B$.

(iii) Let $B^\text{par}$ be another Banach space (of parameters) with norm $\|\cdot\|_\text{par}$ and $H: \mathbb{R} \times B \times B^\alpha \times B^\text{par} \rightarrow B$ be a continuous mapping that is Lipschitz continuous in the sense that
\[ \|H(t, b_0, b_1, \ldots, b_n, \alpha) - H(t, \tilde{b}_0, \tilde{b}_1, \ldots, \tilde{b}_n, \alpha)\| \leq L_H^\text{par}\sum_{j=0}^n \|b_j - \tilde{b}_j\| \left(1 + L_H^\text{par}\sum_{j=1}^n \|b_j\|\right), \] (4.2)
\[ \|H(t, b_0, b_1, \ldots, b_n, \alpha) - H(t, b_0, b_1, \ldots, b_n, \tilde{\alpha})\| \leq L_H^\text{par}\|\alpha - \tilde{\alpha}\| \left(1 + \sum_{j=1}^n \|b_j\|\right) \] (4.3)
and is of linear growth
\[ \|H(t, b_0, b_1, \ldots, b_n, \alpha)\| \leq L_H^\text{par}\left(1 + \sum_{j=1}^n \|b_j\|\right) \] (4.4)
with some constants $L_H$, $L_H'$ and $L_H^\text{par}$. 
Remark 4.1. For the classical HJB equations, estimate (4.2) holds with \( L_H' = 0 \), in which case (4.4) follows from (4.2) and (4.3) (with \( L_H \) that may depend on \( \alpha \)). However, for McKean–Vlasov type equations, the linear growth of the Lipschitz constant in (4.2) is not avoidable.

An important assumption in our analysis is the following smoothing property of the semigroup \( e^{At} \): for \( t > 0 \) it takes \( B \) to \( B_1 \) and
\[
\|e^{At}\|_{B \rightarrow B_1} \leq \kappa t^{-\omega}, \quad t \in (0, 1),
\]
with some constants \( \kappa > 0 \) and \( \omega \in (0, 1) \). Sometimes we use a similar condition for the pair \( B_1, B_2 \):
\[
\|e^{At}\|_{B_1 \rightarrow B_2} \leq \kappa_1 t^{-\omega}, \quad t \in (0, 1).
\]

Remark 4.2. For pseudo-differential operators \( A \) (including the generators of Feller semigroups) the deeper smoothing property (4.6) can be derived from (4.5) and the differentiability of the symbol \( A(x, p) \) of the operator \( A \) with respect to the first argument (see [25, Theorem 5.15.1]).

The so-called mild version of the Cauchy problem (1.1) is the integral equation
\[
b(t) = e^{A(t-a)}Y + \int_0^t e^{A(t-s)}H(s, b(s), Db(s), \alpha) \, ds, \quad t \geq a.
\]

It is well known (see, e.g., [25]) and easy to see that if \( b(t) \) is a classical solution to equation (1.1), then it also solves equation (4.7) (i.e., \( b(\cdot) \) is a continuous mapping from \( [a, \infty) \) to \( B_1 \) satisfying equation (4.7)), so that the uniqueness for (4.7) implies the uniqueness for (1.1).

The following theorem on the well-posedness of equation (4.7) is valid.

**Theorem 4.1.** Let Conditions A and the smoothing property (4.5) hold. Then equation (4.7) is well posed in \( B_1 \); that is, for any \( Y \in B_1 \) and \( \alpha \in B^{\text{par}} \) it has a unique global solution \( b(t) = b(t; Y, \alpha) \in B_1 \) that depends Lipschitz continuously on the initial data \( Y \) and the parameter \( \alpha \). In particular,
\[
\sup_{t \in [a, T]} \|b(t; Y, \alpha) - b(t; \bar{Y}, \bar{\alpha})\|_1 \leq K (\|\alpha - \bar{\alpha}\|_{\text{par}}(1 + \|Y\|_1) + \|Y - \bar{Y}\|_1)
\]

with a constant \( K \) depending on \( T - a \) and all constants appearing in the assumptions of the theorem.

**Proof.** Solutions to (4.7) are fixed points of the mapping

\[
[\Phi_{Y,\alpha}(b(\cdot))](t) = e^{A(t-a)}Y + \int_a^t e^{A(t-s)}H(s, b(s), Db(s), \alpha) \, ds
\]

acting in \( C_Y([a, T], B_1) \) for any \( T > a \). The fact that it takes this space to itself follows directly from the assumptions of the theorem.

Assume first that \( L_H' = 0 \) in (4.2). Then it follows that
\[
\|([\Phi_{Y_1,\alpha}(b(\cdot))(t) - [\Phi_{Y_2,\alpha}(b(\cdot))](t)]_1 \leq M_1 e^{m_1(t-a)}\|Y_1 - Y_2\|_1
\]

and
\[
\|([\Phi_{Y,\alpha}(b^1(\cdot))](t) - [\Phi_{Y,\alpha}(b^2(\cdot))](t)]_1 \leq \kappa L_H[1 + (T - a)^\omega]M_1 e^{m_1(T-a)} \int_a^t (t-s)^{-\omega} (\|b^1(s) - b^2(s)\| + \sum_{j=1}^n \|D_j b^1(s) - D_j b^2(s)\|) \, ds
\]

\[
\leq \kappa L_H(n + 1)[1 + (T - a)^\omega]M_1 e^{m_1(T-a)} \int_a^t (t-s)^{-\omega} \|b^1(s) - b^2(s)\|_1 \, ds.
\]
Note that the coefficient \((T - a)^\omega\) arises when considering the case \(t - a > 1\), since the smoothing \((4.5)\) is assumed only for \(t \leq 1\). Due to these inequalities, the well-posedness for any \(\alpha\) follows from Theorem 3.1. Moreover, from \((4.3)\) we find that
\[
\|\Phi_{Y,\alpha}(b(\cdot))(t) - \Phi_{Y,\tilde{\alpha}}(b(\cdot))(t)\|_1 \leq \tilde{K}\|\alpha - \tilde{\alpha}\|_{\text{par}} (1 + \|b(\cdot)\|_{C([r,T],B_1)})
\]
therefore, if \(b(\cdot)\) is a fixed point of \(\Phi_{Y,\alpha}\), its growth is bounded by \((3.4)\), and thus
\[
\|\Phi_{Y,\alpha}(b(\cdot))(t) - \Phi_{Y,\tilde{\alpha}}(b(\cdot))(t)\|_1 \leq K\|\alpha - \tilde{\alpha}\|_{\text{par}} (1 + \|Y\|_1),
\]
with some constants \(\tilde{K}\) and \(K\) depending on the constants appearing in the assumptions of the theorem. Hence Theorem 3.1 (with \(\alpha\) from \((3.1)\) depending on \(\|Y\|_1\)) again applies to give \((4.8)\).

If \(L'_{H} \neq 0\) in \((4.2)\), an additional preliminary step is needed in the proof. Namely, one has to show that all iterations are uniformly bounded in \(B_1\) and then apply the previous argument. The boundedness of all iterations follows from the assumption of linear growth \((4.4)\). To prove this, let us assume, just to shorten the formulas, that \(T - a < 1\). Then we derive from \((4.4)\) and \((4.5)\) that
\[
\|\Phi_{Y,\alpha}(b(\cdot))(t)\|_1 \leq r \left( 1 + \int_a^t (t-s)^{-\omega}\|b(s)\|_1 ds \right) \leq \lambda (1 + (I_a^{1-\omega}b(t))),
\]
with constants \(r\) and \(\lambda\) depending on \(\|Y\|_1\), \(\omega\), \(M_1\) and \(m_1\), where \(I_a^{1-\omega}\) is a fractional integral of order \(1 - \omega\). Iterating and using the obvious formula
\[
E_\beta(\lambda(t-a)^\beta) = \left( \sum_{j=0}^{\infty} \lambda^j I_a^{\beta j} \right) 1(t),
\]
we find the following estimate for the iterations of \(\Phi\):
\[
\|\Phi_{Y,\alpha}(Y)(t)\|_1 \leq \lambda E_{1-\omega}(\lambda(t-a)^{1-\omega}) + \lambda^n I_a^{n\beta} 1 \|Y\|_1,
\]
which is uniformly bounded in \(n\). \(\square\)

**Remark 4.3.** Of course, Theorem 3.1 also supplies explicit estimates for the constant \(K\) and the growth of solutions in time.

To treat equation \((1.2)\), we recall that its mild form is the integral equation
\[
b(t) = E_\beta(A(t-a)^\beta)Y + \beta \int_a^t (t-s)^{\beta-1} E_\beta' (A(t-s)^\beta)H(s,b(s),Db(s),\alpha) ds, \tag{4.12}
\]
where \(E_\beta(A)\) is defined by \((2.1)\). Thus, more explicitly this equation writes down as
\[
b(t) = \frac{1}{\beta} \int_0^\infty e^{A(t-a)^\beta} x Y x^{-1-1/\beta} G_\beta(1, x^{-1/\beta}) dx
\]
\[
+ \int_a^t (t-s)^{\beta-1} \int_0^\infty e^{A(t-s)^\beta} x^{-1/\beta} G_\beta(1, x^{-1/\beta}) dx H(s,b(s),Db(s),\alpha) ds. \tag{4.13}
\]
Again one proves (see [25, Theorem 8.2.1]) that any solution of problem \((1.2)\) solves \((4.12)\).

**Theorem 4.2.** Let Conditions A and the smoothing property \((4.5)\) hold. Then equation \((4.12)\) is well posed in \(B_1\); that is, for any \(Y \in B_1\) and \(\alpha \in B_{\text{par}}\) it has a unique global solution \(b(t) = b(t;Y,\alpha) \in B_1\), which depends Lipschitz continuously on the initial data \(Y\) and parameter \(\alpha\) so that \((4.8)\) holds.
Proof. For simplicity, let us discuss only the case with $L'_H = 0$ in (4.2). Modifications needed for the general case are the same as in the previous theorem. Solutions to (4.12) are fixed points of the mapping

$$[\Phi_{Y,\alpha}(b(\cdot))(t) = E_{\beta}(A(t - a)\beta)Y + \beta \int_a^t (t - s)^{\beta - 1} E_{\beta}'(A(t - s)\beta)H(s, b(s), Db(s), \alpha) \, ds$$

acting in $C_Y([a, T], B_1)$ for any $T > a$. We have

$$||[\Phi_{Y_1,\alpha}(b(\cdot))] - [\Phi_{Y_2,\alpha}(b(\cdot))]||_1 \leq M_1 E_{\beta}(m_1(t - a)\beta)||Y_1 - Y_2||_1,$$  \hspace{1cm} (4.14)

which follows from (4.1) and (2.1). Next,

$$[\Phi_{Y,\alpha}(b^1(\cdot))] - [\Phi_{Y,\alpha}(b^2(\cdot))] = \int_a^t (t - s)^{\beta - 1} \int_0^\infty e^{A(t-s)\beta x} x^{-1/\beta} G_{\beta}(1, x^{-1/\beta}) \, dx \times \left[ H(s, b^1(s), Db^1(s), \alpha) - H(s, b^2(s), Db^2(s), \alpha) \right] \, ds.$$

Decomposing the double integral into two parts over the set $(t - s)^{\beta} x \leq 1$ and its complement, we use estimate (4.5) in the first integral and the second estimate from (4.1) in the second integral, which yields

$$||[\Phi_{Y_1,\alpha}(b^1(\cdot))] - [\Phi_{Y_2,\alpha}(b^2(\cdot))]||_1 \leq \chi L_H(n + 1) \int_a^t (t - s)^{\beta - 1} (t - s)^{-\omega \beta} \int_0^{\infty} x^{-\omega - 1/\beta} G_{\beta}(1, x^{-1/\beta}) \, dx ||b^1(s) - b^2(s)||_1 \, ds$$

$$+ \chi L_H M_1(n + 1) \int_a^t (t - s)^{\beta - 1} e^{m_1(t-s)\beta x} x^{-1/\beta} G_{\beta}(1, x^{-1/\beta}) \, dx ||b^1(s) - b^2(s)||_1 \, ds.$$

Let us stress for clarity that to estimate the second integral, we used the representation

$$e^{A(t-s)\beta x} = e^{A[(t-s)\beta x-1]} e^A$$

and applied estimate (4.5) to the second factor and estimate (4.1) to the first one.

Using (2.3) to estimate the first term and (2.2) to estimate the second term yields

$$||[\Phi_{Y,\alpha}(b^1(\cdot))] - [\Phi_{Y,\alpha}(b^2(\cdot))]||_1 \leq L_H(n + 1)(\chi + M_1)C(m_1, \beta, T - a) \int_a^t (t - s)^{-[1 - \beta(1 - \omega)]} ||b^1(s) - b^2(s)||_1 \, ds,$$  \hspace{1cm} (4.15)

with a constant $C$ depending on $m_1, \beta$ and $T - a$. Due to these inequalities and a similar inequality expressing the Lipschitz dependence of $\Phi$ on $\alpha$, the claim follows again from Theorem 3.1. \hspace{1cm} \square

Remark 4.4. Notice that $E_1(x) = e^x$ and thus $E'_1(x) = e^x$. Hence the fractional mild equation (4.12) turns into the classical mild equation (4.7) as $\beta \to 1$. So the discussion of the case of classical differential equations can be considered to be included in the fractional theory. In later sections, we shall sometimes talk about just fractional equations having in mind that the classical case is recovered automatically by setting $\beta = 1$.

Remark 4.5. All the results above have their straightforward counterparts for the backward problems under exactly the same conditions. The only difference is that, instead of the sets
Let us comment on how the results above are applied to the HJB equation (1.4) and McKean–Vlasov equation (1.6). For these cases, $B$ is a space of functions on $\mathbb{R}^d$, $A$ is the generator of a Feller process in $\mathbb{R}^d$ and $D$ is the gradient (derivative) operator. Specifically, for the HJB equation the natural Banach triple is $C^2_0(\mathbb{R}^d) \subset C^1_0(\mathbb{R}^d) \subset C_\infty(\mathbb{R}^d)$, and for the McKean–Vlasov equations a possible triple is $W_2(\mathbb{R}^d) \subset W_1(\mathbb{R}^d) \subset L_1(\mathbb{R}^d)$. Here $C^\infty(\mathbb{R}^d)$ denotes the space of continuous functions on $\mathbb{R}^d$ tending to zero at infinity, $C^2_0(\mathbb{R}^d)$ are their subsets of functions having derivatives of order up to $j$ in $C^\infty(\mathbb{R}^d)$, and $W_j(\mathbb{R}^d)$ denote the Sobolev spaces of functions with partial derivatives (understood in the sense of generalized functions) of order up to $j$ in $L_1(\mathbb{R}^d)$, equipped with the integral (Sobolev) norms

$$\|f\|_{L^1(\mathbb{R}^d)} = \int |f(x)| \, dx, \quad \|f\|_{W^1(\mathbb{R}^d)} = \|f\|_{L^1(\mathbb{R}^d)} + \sum_{j=1}^d \int \left| \frac{\partial f}{\partial x_j} \right| \, dx,$$

$$\|f\|_{W^j(\mathbb{R}^d)} = \|f\|_{W^1(\mathbb{R}^d)} + \sum_{i<j} \int \left| \frac{\partial^2 f}{\partial x_j \partial x_i} \right| \, dx.$$ 

The proof of the smoothing properties of $A$ is usually based on the properties of the Green functions of the operator $A$ (transition probabilities of the processes generated by $A$). For instance, (4.5) and (4.6) are known to hold with $\omega = 1/2$ for non-degenerate diffusion operators $A$ with sufficiently smooth coefficients. When $A = -|\Delta|^{\alpha}$ with $\alpha \in (1, 2)$, or more generally $A = -a(x)|\Delta|^{\alpha}$ with a smooth positive function $a$ bounded from above and below, or even more generally when $A$ is a pseudo-differential operator that generates a non-degenerate stable-like process with a symmetric spectral measure, that is,

$$A = \int_{S^{d-1}} |(\nabla, s)|^\alpha \mu(s) \, ds,$$

with a positive smooth function $\mu$ on the $(d - 1)$-dimensional sphere $S^{d-1}$, bounded from above and below, the semigroup generated by $A$ satisfies (4.5), (4.6) and (4.1) with $\omega = 1/\alpha$, as shown in [21] (see also the books [23, 25]). Similar properties also hold for various mixtures of diffusions and stable processes perturbed by pure jump processes. As shown in Theorem 8.1 below, from the properties of $A$ one can derive analogous properties for $A^*$ by duality arguments. Thus, for these $A$ the required assumptions on $A$ and $A^*$ from (1.4) and (1.6) hold. We also refer to [25, Sect. 5.15], where it is shown how one can derive (4.6) from the simpler estimate (4.5) and additional smoothness. The theory of fractional HJB equations was initially developed in [33], with more detail given in [25].

**Remark 4.6.** From an abstract point of view the spaces $C^\infty(\mathbb{R}^d)$ and $L^1(\mathbb{R}^d)$ are the basic examples of abstract $AL$ and $AM$ spaces (see [42]).

5. **ANTICIPATING EQUATIONS**

Anticipating versions of equations (1.1) and (1.2) can be stated as the problems

$$\dot{b}(t) = Ab(t) + H(t, b(t), Db(t), u(t, b(\cdot))), \quad b(a) = Y, \quad t \in [a, T],$$

(5.1)

$$D^d_{a^+, s} b(t) = Ab(t) + H(t, b(t), Db(t), u(t, b(\cdot))), \quad b(a) = Y, \quad t \in [a, T],$$

(5.2)

where $A, D_1, \ldots, D_n$ and $H$ are the same as in (1.1) and (1.2) and $u$ is a continuous mapping $\mathbb{R} \times C([a, T], B_1) \rightarrow B_{\text{par}}$ that is Lipschitz continuous in the second argument, so that

$$\|u(t, b(\cdot)) - u(t, \tilde{b}(\cdot))\|_{\text{par}} \leq L_u \|b(\cdot) - \tilde{b}(\cdot)\|_{C([a, T], B_1)}.$$ 

(5.3)
Usually, in applications \( u(t, \cdot) \) actually depends only on the future \( b(s): s \in [t, T] \), but this additional condition does not simplify the analysis.

For these problems, uniqueness usually does not hold globally (that is, for large \( T \)); only the existence can be proved under rather general assumptions (see, e.g., [34] for equations of type (5.1)). We shall prove here the local well-posedness, that is, for sufficiently small \( T - a \). Of course, we again work with mild solutions, relying on the fact that any solution of (5.1) is a fixed point of the mapping

\[
[\Phi_Y(b(\cdot))](t) = e^{A(t-a)}Y + \int_0^t e^{A(t-s)}H(s, b(s), Db(s), u(s, b(\cdot))) \, ds, \tag{5.4}
\]

and any solution of (5.2) is a fixed point of the mapping

\[
[\Phi_Y(b(\cdot))](t) = E_\beta(A(t-a)\beta)Y + \beta \int_0^t (t-s)^{\beta-1}E'_\beta(A(t-s)\beta)H(s, b(s), Db(s), u(s, b(\cdot))) \, ds. \tag{5.5}
\]

Because of the anticipating dependence of \( u \) on \( b \), Theorem 3.1 is not applicable. We shall use just the standard Banach contraction principle.

**Theorem 5.1.** Let Conditions A, the smoothing property (4.5) and the Lipschitz estimate (5.3) hold. Then, for any \( r > 0 \), there exists \( T_0 \) such that equations (5.1) and (5.2) have unique mild solutions \( b(t) = b(t; Y) \) (that is, they are fixed points of (5.4) and (5.5), respectively) for all \( T < T_0 \) and \( Y \) such that \( ||Y||_1 \leq r \), and these solutions depend Lipschitz continuously on \( Y \) inside the ball \( ||Y||_1 \leq r \).

**Proof.** Due to formula (2.1) expressing Mittag-Leffler functions in terms of exponents, the proof is essentially the same for the two equations. So let us discuss only equation (5.1). There are two steps. Firstly, due to the linear growth condition (4.4), one derives for \( ||Y||_1 \leq r \) the estimate

\[
||[\Phi_Y(b(\cdot))](t)||_1 \leq M_1 e^{m_1(T-a)}r + \epsilon \|b(\cdot)\|_{C([a,T],B_1)},
\]

where \( \epsilon \) depends on \( T_0 \) and \( \epsilon \to 0 \) for \( T_0 \to a \). So one can choose \( T_0 \) in such a way that \( \epsilon < 1 \). In this case, if \( \sup_t \|b(t)\|_1 \leq R \), then

\[
||[\Phi_Y(b(\cdot))](t)||_1 \leq M_1 e^{m_1(T-a)}r + \epsilon R \leq R
\]

whenever

\[
R > \frac{M_1 e^{m_1(T-a)}r}{1 - \epsilon}.
\]

Thus, for any such \( R \), \( \Phi_Y \) takes \( C_Y([a,T],B^R_1) \) to itself, where \( B^R_1 \) is the ball in \( B_1 \) of radius \( R \) centred at 0. Next, from the estimate

\[
||[\Phi_Y(b^1(\cdot))](t) - [\Phi_Y(b^2(\cdot))](t)||_1 \leq ||b^1(\cdot) - b^2(\cdot)||_{C([a,T],B_1)}(T-a)e^{m_1(T-a)}(k_1 + k_2R),
\]

where \( k_1 \) and \( k_2 \) do not depend on \( R \) and \( Y \), it follows that \( \Phi_Y \) is a contraction in the complete metric spaces \( C_Y([a,T],B^R_1) \) for sufficiently small \( T - a \).

The Lipschitz property in \( Y \) follows, as in Theorem 3.1, from the general fact that the closeness of mappings implies the closeness of their fixed points (see, e.g., [25, Proposition 9.1.3]).
6. FORWARD–BACKWARD SYSTEMS

The well-posedness result of this section is an abstract version of the results obtained in [34] (slightly extended in [25, Ch. 6]). This abstract presentation not only simplifies the exposition, but it is specifically designed for a more or less straightforward extension to the fractional case, as given in the next section, and which is our main concern here.

The paper [34] also contains a global existence result (without uniqueness) that can be recast in the present abstract setting, but we do not go into detail here.

To introduce our coupled forward–backward system, we need to introduce Conditions A both for the forward and backward parts of the system plus the coupling mechanism (interpreted as control in applications) including an appropriate setting of the parameter space to support this mechanism. That is, the forward–backward version of Conditions A reads as follows.

**Conditions AFB.** (i) Let \( B_2 \subset B_1 \subset B \) and \( B_{2b} \subset B_{1b} \subset B_b \) be two Banach triples (index ‘b’ coming from “backward”), with the norms in the second triple denoted by \( \| \cdot \|_{2b}, \| \cdot \|_{1b} \) and \( \| \cdot \|_b \). Let \( D = (D_1, \ldots, D_n) \) and \( D^b = (D_{1b}, \ldots, D_{nb}) \) be operators satisfying Conditions A(i) for the triples \( B_2 \subset B_1 \subset B \) and \( B_{2b} \subset B_{1b} \subset B_b \), respectively.

(ii) Let \( A \) and \( A^b \) be operators with properties described in Conditions A(ii) for the triples \( B_2 \subset B_1 \subset B \) and \( B_{2b} \subset B_{1b} \subset B_b \), respectively.

(iii) Let \( B^{\text{con}} \) be another Banach space (of control functions) with norm denoted by \( \| \cdot \|_{\text{con}} \). Let \( H: \mathbb{R} \times B \times B^n \times B^{\text{con}} \to B \) be a continuous mapping that is Lipschitz continuous in the sense that

\[
\| H(t, b_0, b_1, \ldots, b_n, u) - H(t, \tilde{b}_0, \tilde{b}_1, \ldots, \tilde{b}_n, u) \| \leq L_H \sum_{j=0}^n \| b_j - \tilde{b}_j \| \left( 1 + L_H' \sum_{j=1}^n \| b_j \| \right), \tag{6.1}
\]

\[
\| H(t, b_0, b_1, \ldots, b_n, u) - H(t, b_0, b_1, \ldots, b_n, \tilde{u}) \| \leq L_H^{\text{par}} \| u - \tilde{u} \|_{\text{con}} \left( 1 + L_H' \sum_{j=1}^n \| b_j \| \right). \tag{6.2}
\]

Let \( H^b: \mathbb{R} \times B_b \times (B_b)^n \times C([a, T], B) \to B_b \) be a continuous mapping that is Lipschitz continuous in the sense that

\[
\| H^b(t, f_0, f_1, \ldots, f_n, b(\cdot)) - H^b(t, \tilde{f}_0, \tilde{f}_1, \ldots, \tilde{f}_n, b(\cdot)) \|_b \leq L_H^b \sum_{j=0}^n \| f_j - \tilde{f}_j \|_b \left( 1 + (L_H^b)' \sum_{j=1}^n \| b_j \|_b \right), \tag{6.3}
\]

\[
\| H^b(t, f_0, f_1, \ldots, f_n, b(\cdot)) - H^b(t, f_0, f_1, \ldots, f_n, \tilde{b}(\cdot)) \|_b \leq L_H^b \| b(\cdot) - \tilde{b}(\cdot) \|_{C([a, T], B)} \left( 1 + (L_H^b)' \sum_{j=1}^n \| f_j \|_b \right). \tag{6.4}
\]

Finally, let both \( H \) and \( H^b \) be of linear growth (in the sense of condition (4.4)).

(iv) Let \( u: \mathbb{R} \times B_b \times (B_b)^n \to B^{\text{con}} \) be a continuous function that is Lipschitz continuous in the sense that

\[
\| u(t, f_0, f_1, \ldots, f_n) - u(t, \tilde{f}_0, \tilde{f}_1, \ldots, \tilde{f}_n) \|_{\text{con}} \leq L_u \sum_{j=0}^n \| f_j - \tilde{f}_j \|_b \tag{6.5}
\]

with a constant \( L_u \).
The forward–backward system we are analyzing here is of the form

\[
\begin{align*}
\dot{b}(t) &= Ab(t) + H(t, b(t), Db(t), u(t, f(t), D^bf(t))), \quad b(a) = Y, \quad t \in [a, T], \\
\dot{f}(t) &= -A^b f(t) + H^b(t, f(t), D^bf(t), b_{\geq t}), \quad f(T) = Z, \quad t \in [a, T].
\end{align*}
\]  

(6.6)

The notation \(b_{\geq t}\) indicates the assumption that \(H^b(t, \cdot, \cdot)\) depends only on the future values \(\{b(s), s \in [t, T]\}\) of curves \(b(\cdot) \in C([a, T], B)\).

Our approach to the analysis of system (6.6) is based on its reduction to a single anticipating equation of type (5.1). Namely, by Theorem 4.1 (and Remark 4.5), given a curve \(b(\cdot) \in C_Y([a, T], B)\) and \(Z \in B_{1b}\), there exists a unique mild solution \(f(t) = f(t; Z, b(\cdot)) \in C([a, T], B_{1b})\) of the second equation of system (6.6). Substituting this solution into the first equation of system (6.6), we get the equation

\[
\dot{b}(t) = Ab(t) + H(t, b(t), Db(t), u(t, f(t; Z, b(\cdot)), D^bf(t; Z, b(\cdot))))), \quad t \in [a, T].
\]  

(6.7)

This equation is of type (5.1).

**Theorem 6.1.** Let Conditions AFB and the smoothing property (4.5) for \(A\) and \(A^b\) hold. Then, for any \(Y \in B_1\) and \(Z \in B_{1b}\) there exists \(T_0\) such that the forward–backward system (6.6) has a unique mild solution \((b(t), f(t))\) for all \(T < T_0\), which depends locally Lipschitz continuously on \(Y\) and \(Z\), that is, when \(Y\) and \(Z\) are taken from bounded sets.

**Proof.** The discussion above shows that the statement is reduced to the well-posedness of equation (6.7). To apply Theorem 5.1, we just have to check the Lipschitz estimate (5.3), which in the present case has the form

\[
\|u(t, f(t; Z, b(\cdot)), D^bf(t; Z, b(\cdot))) - u(t, f(t; Z, \tilde{b}(\cdot)), D^bf(t; Z, \tilde{b}(\cdot)))\|_{C([a, T], B)} \leq \|b(\cdot) - \tilde{b}(\cdot)\|_{C([a, T], B_1)}. \tag{6.8}
\]

But this holds, because the function \(u\) is a Lipschitz continuous function of its arguments (by (6.5)) and \(f\) and \(D^bf\) are Lipschitz functions of \(b(\cdot)\) by Theorem 4.1. Let us stress that this application was the main reason for us to get the continuous dependence of the solutions of the HJB and McKean–Vlasov equations in a “deeper” space \(B_1\) (rather than just in \(B\)).

The Lipschitz property of solutions with respect to \(Y\) and \(Z\) is obtained similarly to Theorems 3.1 and 5.1. \(\Box\)

### 7. FRACTIONAL FORWARD–BACKWARD SYSTEMS

The fractional analog of system (6.6) is, of course, the system

\[
\begin{align*}
D^\alpha_{a+} b(t) &= Ab(t) + H(t, b(t), Db(t), u(t, f(t), D^bf(t))), \quad b(a) = Y, \quad t \in [a, T], \\
D^\beta_{T-} f(t) &= -A^b f(t) + H^b(t, f(t), D^bf(t), b_{\geq t}), \quad f(T) = Z, \quad t \in [a, T].
\end{align*}
\]  

(7.1)

Under the same Conditions AFB, we can conclude from Theorem 4.2 that for a given curve \(b(\cdot) \in C_Y([a, T], B)\) and \(Z \in B_{1b}\) there exists a unique mild solution \(f(t) = f(t; Z, b(\cdot)) \in C([a, T], B_{1b})\) of the second equation of system (7.1). Substituting this solution into the first equation of system (7.1), we get the equation

\[
D^\alpha_{a+} b(t) = Ab(t) + H(t, b(t), Db(t), u(t, f(t; Z, b(\cdot)), D^bf(t; Z, b(\cdot))))), \quad t \in [a, T],
\]  

(7.2)

which is of type (5.2).

The following result is proved by exactly the same argument as Theorem 6.1.
Theorem 7.1. Let Conditions AFB and the smoothing property (4.5) for \( A \) and \( A^b \) hold. Then, for any \( Y \in B_1 \) and \( Z \in B_{1b} \) there exists \( T_0 \) such that the forward-backward system (7.1) has a unique mild solution \((b(t), f(t))\) for all \( T < T_0 \), which depends locally Lipschitz continuously on \( Y \) and \( Z \).

8. DUAL BANACH TRIPLES AND RELATED NONLINEAR EQUATIONS

In applications the two Banach triples participating in Conditions AFB are usually linked by duality, and the operators \( A \) and \( A^b \) are dual, which allows one to derive the required properties of \( A^b \) from the corresponding properties of \( A \).

It is convenient to introduce an appropriate abstract setting. Following [25], we shall call a pair of Banach spaces \((B, C)\) a dual pair if each of these spaces is a closed subset of the dual of the other space that separates the points of the latter. Of course, if \( B' \) is the Banach dual of \( B \), then both \((B, B')\) and \((B', B)\) are dual pairs. But our symmetric notion also includes pairs like \((C_{\infty}(\mathbb{R}^d), L_1(\mathbb{R}^d))\), which are crucial for our present discussion. The duality is a bilinear form \((B, C) \rightarrow \mathbb{R}\), which will always be denoted by \((b, c)\), \( b \in B, c \in C \). Recall that the requirement that \( B \) is a subset of the dual of \( C \) means the following formula for the norm:

\[
\|b\|_B = \sup_{\|c\| \leq 1} |(b, c)|.
\]

Let us say that a Banach triple \( B_2 \subset B_1 \subset B \) is generated by a vector-valued operator \( D = (D_1, \ldots, D_n) \) if

\[
\|b\|_1 = \|b\| + \sum_{j=1}^n \|D_j b\|, \quad \|b\|_2 = \|b\|_1 + \sum_{i \leq j} \|D_i D_j b\|.
\] (8.1)

Let two Banach triples \( B_2 \subset B_1 \subset B \) and \( B_{2b} \subset B_{1b} \subset B_b \) (with the norms of the second denoted by \( \|\cdot\|_{2b}, \|\cdot\|_{1b} \) and \( \|\cdot\|_b \)) be generated by operators \( D = (D_1, \ldots, D_n) \) and \( D^b = (D^b_1, \ldots, D^b_n) \), respectively. Let us say that the triples are dual if \( (B, B_b) \) forms a dual pair and the operators \( D_j \) and \( -D^b_j \) are dual in the sense that

\[
(D_j b, f) = -(b, D^b_j f), \quad b \in B_1, \quad f \in B_{1b}, \quad j = 1, \ldots, n.
\]

With some abuse of notation, we shall write just \( D \) for the operator \( D^b \) as well, so that the previous equation becomes \((D_j b, f) = -(b, D_j f)\), and we shall say that the dual triples are generated by \( D \). Of course, in all applications, both \( D \) and \( D^b \) are the gradients, so that this unified notation is fully natural.

When discussing duality, it is useful to work with two versions of the smoothing property (4.5). Namely, let a Banach triple \( B_2 \subset B_1 \subset B \) be generated by \( D = (D_1, \ldots, D_n) \) and \( A \) generate a strongly continuous semigroup \( e^{At} \) in \( B \). Let us say that the semigroup \( e^{At} \) is smoothing from the right in the first order if (4.5) holds. Since the triple is generated by \( D \), this is equivalent to the estimate

\[
\|D_j e^{At} b\| \leq \varkappa t^{-\omega} \|b\|, \quad j = 1, \ldots, n, \quad t \in (0, 1],
\] (8.2)

for all \( b \in B \) (possibly with a constant \( \varkappa \) other than in (4.5), but with the same \( \omega \)). Let us say that the semigroup \( e^{At} \) is smoothing from the left in the first order if for any \( t > 0 \) the operators \( e^{At} D_j \) defined on \( B_1 \) can be extended to bounded operators in \( B \) such that

\[
\|e^{At} D_j b\| \leq \varkappa t^{-\omega} \|b\|, \quad j = 1, \ldots, n, \quad t \in (0, 1],
\] (8.3)

for all \( b \in B \).
For a Banach triple generated by $D$, let us say that the semigroup $e^{At}$ is smoothing from the right in the second order if
\[ \| D_j e^{At} b \|_1 \leq \zeta t^{-\omega} \| b \|_1, \quad j = 1, \ldots, n, \quad t \in (0, 1], \]  
holds, and it is smoothing from the left in the second order if
\[ \| e^{At} D_j b \|_1 \leq \zeta t^{-\omega} \| b \|_1, \quad j = 1, \ldots, n, \quad t \in (0, 1], \]  
holds.

The next result supplies some links between various smoothing properties.

**Theorem 8.1.** Let two Banach triples $B_2 \subset B_1 \subset B$ and $B_{2b} \subset B_{1b} \subset B_b$ generated by operators $D = (D_1, \ldots, D_n)$ be dual. Let operators
\[ A \in \mathcal{L}(B_2, B) \quad \text{and} \quad A^b \in \mathcal{L}(B_{2b}, B_b) \]
be dual, $A^b = A^*$, in the sense that
\[ (Ab, f) = (b, A^* f), \quad b \in B_2, \quad f \in B_{2b}, \]
and let $A$ and $A^b = A^*$ generate strongly continuous semigroups $e^{At}$ in $B$ and $e^{A^* t}$ in $B_b$, with the cores of the operators $A$ and $A^*$ contained in $B_1$ and $B_{1b}$, respectively.

(i) If $e^{At}$ is smoothing in the first order from the right (respectively, from the left), then $e^{A^* t}$ is smoothing in the first order from the left (respectively, from the right), with the same parameter $\omega$, and vice versa.

(ii) If the commutators $[D_j, e^{At}]$ extend to uniformly bounded (in $t \in [0, T]$ for any $T$) operators in $B$, then

(a) $e^{At}$ is a bounded semigroup in $B_1$ (that is, the second estimate of (4.1) follows from the first one), and

(b) $e^{At}$ is smoothing in the first order from the right if and only if it is smoothing in the first order from the left.

(iii) The commutators $[D_j, e^{At}]$ extend to uniformly bounded operators in $B$ if and only if $[D_j, e^{A^* t}]$ extend to uniformly bounded operators in $B_b$.

(iv) If $e^{A^* t}$ is smoothing from the left in the first order, then $e^{At}$ is smoothing from the left in the second order.

(v) If the commutators $[D_j, e^{At}]$ extend to uniformly bounded (in $t \in [0, T]$ for any $T$) operators in $B_1$, then

(a) $e^{At}$ is a bounded semigroup in $B_2$, and

(b) $e^{At}$ is smoothing in the second order from the right if and only if it is smoothing in the second order from the left.

**Proof.** (i) Let $e^{At}$ be smoothing in the first order from the right. If $f \in B_{1b}$ belongs to the core of the generator $A^*$ for the semigroup $e^{A^* t}$, then
\[ \| D_j^b e^{A^* t} f \|_b = \sup_{\| b \|_1 \leq 1} |(b, D_j^b e^{A^* t} f)| = \sup_{\| b \|_1 \leq 1, b \in B_1} |(b, D_j^b e^{A^* t} f)| = \sup_{\| b \|_1 \leq 1, b \in B_1} |(D_j b, e^{A^* t} f)| = \sup_{\| b \|_1 \leq 1, b \in B_1} |(e^{At} D_j b, f)| \leq \zeta e^{-\omega t} \| f \|_b. \]

Since the core of the generator $A^*$ is dense, the operator $D_j e^{A^* t}$ extends to a bounded operator $B_b \rightarrow B_b$. The other statements are proved in a similar way.
(ii) Let us prove (a). If \( b \in B_1 \), then for \( t \in [0, T] \),
\[
\|D_t e^{At} b\| \leq \|e^{At} D_j b\| + \|[D_j, e^{At}] b\| \leq M e^{mt}\|b\|_1 + C(T)\|b\|
\]
and thus
\[
\|e^{At} b\|_1 \leq n M e^{mt}\|b\|_1 + (1 + C(T))\|b\| \leq (n M e^{mt} + 1 + C(T))\|b\|_1,
\]
so that the operators \( e^{At} \) are bounded in \( B_1 \) uniformly for \( t \in [0, T] \). Statement (b) is straightforward because \( D_j e^{At} = e^{At} D_j + [D_j, e^{At}] \).

(iii) This follows from the duality relation
\[
([D_j, e^{At}] b, f) = (b, [D_j, e^{At}] f).
\]

If the commutator \([D_j, e^{At}]\) extends to a uniformly bounded operator in \( B \), then the left-hand side of this equality defines a bounded linear functional of \( b \in B \) for any \( f \in B_b \). In order to identify the right-hand side, one can check it for \( b \in B_1 \) and \( f \in B_{1b} \).

(iv) To estimate \( \|e^{At} D_j b\|_1 \), we need to estimate \( \|D_t e^{At} D_j b\| \). We have
\[
\|D_t e^{At} D_j b\| = \sup_{f \in B_{1b}, \|f\|_1 \leq 1} |(D_t e^{At} D_j b, f)| = \sup_{f \in B_{1b}, \|f\|_1 \leq 1} |(D_j b, e^{At} D_t f)|
\]
\[
\leq \|b\|_1 \sup_{\|f\|_1 \leq 1} \|e^{At} D_t f\|_1 \leq \omega \|b\|_1.
\]

(v) Let us prove (a). In order to estimate \( \|e^{At} b\|_2 \), we need an estimate for \( \|D_j e^{At} b\| \). We have
\[
\|D_j e^{At} b\| = \|D_j e^{At} D_j b\| + \|D_j e^{At} [D_j, b]\| \leq \|e^{At} D_j b\| + \|[D_j, e^{At}] D_j b\| + \|[D_j, e^{At}] [D_j, b]\|
\]
\[
\leq \|e^{At}\|_{B \to B} \|b\|_2 + \|D_j [e^{At}]_{B_1 \to B_1} \|b\|_2 + \|[D_j, e^{At}]_{B_1 \to B_1} \|b\|_1
\]
\[
\leq (\|e^{At}\|_{B \to B} + 2 \|[D_j, e^{At}]_{B_1 \to B_1} \|b\|_2.)
\]

Let us prove (b). Assume that \( e^{At} \) is smoothing in the second order from the left. To show that it is smoothing in the second order from the right, we need to estimate \( \|D_j e^{At} b\|_1 \) and thus \( \|D_j e^{At} b\| \). We have
\[
\|D_j e^{At} b\| \leq \|D_j e^{At} D_j b\| + \|D_j [D_j, e^{At}] b\|,
\]
and both terms are estimated by \( \|b\|_1 \), as required. \( \square \)

For dual Banach triples and dual generators, we can reformulate Theorems 6.1 and 7.1 almost without any additional assumptions on \( A^* \). The following result is a direct consequence of Theorems 6.1, 7.1 and 8.1.

**Theorem 8.2.** Let two Banach triples \( B_2 \subset B \subset B_2^* \subset B_{2b} \subset B_{1b} \subset B_b \) generated by operators \( D = (D_1, \ldots, D_n) \) be dual. Let operators
\[
A \in \mathcal{L}(B_2, B) \quad \text{and} \quad A^b \in \mathcal{L}(B_{2b}, B_b)
\]
be dual, \( A^b = A^* \), in the sense that
\[
(A b, f) = (b, A^* f), \quad b \in B_2, \quad f \in B_{2b},
\]
and let \( A \) and \( A^b \) generate strongly continuous semigroups \( e^{At} \) in \( B \) and \( e^{At^*} \) in \( B_b \), with the cores of the operators \( A \) and \( A^* \) contained in \( B_1 \) and \( B_{1b} \), respectively. Moreover, \( A \) is smoothing in the first order from the right or from the left and the operators \( [e^{At}, D_j] \) extend to uniformly bounded operators \( B \to B \). Let assumptions (iii) and (iv) about \( H, H^b \) and \( u \) of Conditions AFB hold.
Then, for any \( Y \in B_1 \) and \( Z \in B_{1b} \) there exists \( T_0 \) such that the forward–backward systems (6.6) and (7.1) with \( A^b = A^* \) have unique mild solutions for all \( T < T_0 \).

Let us specify the abstract setting to a more concrete forward–backward system consisting of the coupled McKean–Vlasov (forward) and HJB (backward) equations for \( t \in [a, T] \):

\[
\begin{align*}
D_{a+}^t b(t, x) &= A^* b(t, x) + \sum_{j=1}^d h_j(t, b(t, \cdot), u(t, x, \frac{\partial f}{\partial x_j}(t, x))) \frac{\partial b}{\partial x_j}(t, x), \quad b(a, \cdot) = Y, \\
D_{T-}^t f(t, x) &= -Af(t, x) + H(t, x, \frac{\partial f}{\partial x}(t, x), b_{\geq t}), \quad f(T, \cdot) = Z.
\end{align*}
\]  

(8.6)

Here the dual Banach triples are \( W_2(\mathbb{R}^d) \subset W_1(\mathbb{R}^d) \subset L_1(\mathbb{R}^d) \) for the first (forward) equation and \( C_{\infty}^2(\mathbb{R}^d) \subset C_{\infty}^1(\mathbb{R}^d) \subset C_{\infty}(\mathbb{R}^d) \) for the second (backward) equation, both generated by the derivative operator \( D = \partial / \partial x \). The corresponding norms were defined at the end of Section 4. Instead of the Banach space \( B^{con} \) from Conditions AFB, we take the subset of \( U \)-valued functions in the Banach space of continuous functions on \( \mathbb{R}^d \), where \( U \) is a compact subset of the Euclidean space. At the end of Section 4 we gave examples of generators \( A \) that satisfy the requirements of Theorem 8.2, which provide local well-posedness conditions for system (8.6).

9. FRACTIONAL MCKEAN–VLASOV AND HAMILTON–JACOBI–BELLMAN EQUATIONS ON MANIFOLDS

Let

\[
\Delta_{LB} \phi = \text{div}(\nabla \phi) = \frac{1}{\sqrt{\det g}} \sum_{j,k} \frac{\partial}{\partial x_j} \left( \sqrt{\det g} g^{jk} \frac{\partial \phi}{\partial x_k} \right)
\]

(9.1)

denote the Laplace–Beltrami operator on a compact Riemannian manifold \((M, g)\) of dimension \( N \), with the Riemannian metric given by the matrix \( g = (g_{jk}(x)) \) and its inverse matrix \( G = (g^{jk}(x)) \). Let \( K(t, x, y) \) be the corresponding heat kernel, that is, \( K(t, x, y) = \Delta_{LB} K \) as a function of \( (t > 0, x \in M) \) and has the Dirac initial condition \( K(0, x, y) = \delta_y(x) \). It is well known (see, e.g., [2, 11, 14]) that the Cauchy problem for this heat equation is well posed in \( M \) and the resolving operators

\[
S_t f(x) = e^{t \Delta_{LB}} f(x) = \int_M K(t, x, y) f(y) \, dv(y),
\]

(9.2)

where \( dv(y) \) is the Riemannian volume on \( M \), form a strongly continuous semigroup of contractions (the Markovian semigroup of the Brownian motion in \( M \)) in the space \( C(M) \) of bounded continuous functions on \( M \), equipped with the sup-norm.

Let \( C^1(M) \) denote the space of continuously differentiable functions on \( M \) equipped with the norm

\[
\| f \|_{C^1(M)} = \| f \|_{C(M)} + \sup_x \| \nabla f(x) \| x,
\]

(9.3)

where in local coordinates

\[
\| \nabla f(x) \|^2 = (\nabla f(x), G(x) \nabla f(x)) = \sum_{j,k} g^{jk}(x) \frac{\partial f}{\partial x_j} \frac{\partial f}{\partial x_k},
\]

(9.4)

and \( C^2(M) \) denote the space of twice continuously differentiable functions on \( M \) equipped with the norm

\[
\| f \|_{C^2(M)} = \| f \|_{C^1(M)} + \sup_x \| \nabla^2 f(x) \| x,
\]

(9.5)
where in local coordinates

$$\|\nabla^2 f(x)\|_x^2 = \sum_{j,k} g^{jk}(x) \frac{\partial^2 f}{\partial x_j \partial x_k} \frac{\partial^2 f}{\partial x_l \partial x_m}.$$  \hspace{1cm} (9.6)

The gradient $\nabla f(x)$ is an element of $T^*M_x$, the cotangent space to $M$ at $x$, and $\nabla^2 f$ is a tensor of type $(0, 2)$. Formulas (9.4) and (9.6) represent the standard lifting of the Riemannian metric to tensors.

The dual Banach triples used for the analysis of equations on manifolds are the natural analogs of the triples used for equations in $\mathbb{R}^d$. Namely, these are the triple $C^2(M) \subset C^1(M) \subset C(M)$ (with the norms introduced above) and the triple of Sobolev function spaces $W_2(M) \subset W_1(M) \subset L_1(M)$, with the norms

$$\|f\|_{L_1(M)} = \int_M |f(x)| \, dv(x), \quad \|f\|_{W_1(M)} = \|f\|_{L_1(M)} + \int_M \|\nabla f(x)\|_x \, dv(x),$$

$$\|f\|_{W_2(M)} = \|f\|_{W_1(M)} + \int_M \|\nabla^2 f(x)\|_x \, dv(x).$$

It is known that the semigroup $S_t$ in $C(M)$ has $C^2(M)$ as its invariant core. This semigroup $S_t$ is also strongly continuous as a semigroup in $L_1(M)$ (actually in all $L_p(M)$, $p \geq 1$; see [2] and references therein).

**Remark 9.1.** The identification of $\nabla f$ with the collection of $N$ functions $\partial f/\partial x_j$, which is required to comply with the setting of Sections 4 and 8, can be justified only in local coordinates, but not globally. In order to have an invariant theory, we cannot assume that the operator $D = (D_1, \ldots, D_N)$ acts from $B = C(M)$ to $B^N$. In the invariant exposition the gradient $D = \nabla$ maps $C(M)$ to the space $\mathcal{F}T^*M$ of covector fields on $M$, and $D^2 = \nabla^2$ maps $C(M)$ to the symmetric tensor fields of type $(0, 2)$. However, the topologies of $C^1(M)$ and $C^2(M)$ (as well as $W_1(M)$ and $W_2(M)$) remain to be generated by $D = \nabla$ and $D^2$, respectively, according to (9.3) and (9.5), in analogy with (8.1). Thus the whole theory is recovered in this slightly generalized setting.

The key smoothing and smoothness preservation properties of this semigroup needed for our theory are collected in the next result.

**Theorem 9.1.** (i) The operators $S_t$ are smoothing from the right and from the left:

$$\|\nabla S_t f\|_{C(M)} \leq Ct^{-1/2}\|f\|_{C(M)}, \quad \|S_t \nabla f\|_{C(M)} \leq Ct^{-1/2}\|f\|_{C(M)},$$  \hspace{1cm} (9.7)

with a constant $C$, uniformly for any compact time interval.

(ii) The operators $S_t$ are smoothness preserving:

$$\|S_t f\|_{C^1(M)} \leq C\|f\|_{C^1(M)},$$  \hspace{1cm} (9.8)

with a constant $C$, uniformly for any compact time interval.

(iii) The commutators $[\nabla, S_t]$ extend to bounded operators in $C(M)$, uniformly for any compact time interval.

**Proof.** The first estimate in (9.7) is a consequence of the well-known estimate for the derivatives of the heat kernel on a compact Riemannian manifold (see [11, Theorem 6]):

$$\|\nabla K(t, x, y)\|_M \leq C(\delta, N)t^{-N/2}t^{1/2} \exp \left\{-\frac{d^2(x, y)}{(4 + \delta)t} \right\},$$  \hspace{1cm} (9.9)
with any \( \delta > 0 \) and a constant \( C(\delta, N) \), where the derivative \( \nabla \) is taken with respect to \( x \), and where \( d \) is the Riemannian distance in \( M \). Indeed, differentiating (9.2) and using (9.9) yields the first estimate (9.7).

As shown in Theorem 8.1, the second estimate in (9.7) follows from the first estimate and (iii). Also (ii) follows from (iii). Thus it remains to show (iii).

First of all note that, due to the possibility of iteration, it is sufficient to prove (iii) for \( t \leq t_0 \) with any small \( t_0 \). Next, by covering \( M \) with charts, it is sufficient to prove (iii) for \( f \) with support in any small ball of \( M \). Thus it is required to show that the integral operator with the integral kernel

\[
\frac{\partial K}{\partial x}(t, x, y) - \frac{\partial K}{\partial y}(t, x, y)
\]

is bounded in \( C(M_0) \), where \( M_0 \) is a closed ball in \( M \) covered by a coordinate chart.

To this end one can use a well-known asymptotic expansion of the heat kernel \( K(t, x, y) \) (see, e.g., [14, 20] and the most modern presentation in [37]). Namely,

\[
K(t, x, y) \sim K_{as}(t, x, y) \sum_{k=0}^{\infty} t^j \Phi_j(x, y),
\]

where \( \Phi_0 > 0 \), all \( \Phi_j \) are smooth functions and

\[
K_{as}(t, x, y) = (4\pi t)^{-N/2} \exp\left\{-\frac{d^2(x, y)}{4t}\right\},
\]

with \( d(x, y) \) being the Riemannian distance on \( M \) arising from the Riemannian metric \( g \), and this expansion can be differentiated yielding the expansion for the derivatives in \( t, x, y \) of any order. From this fact it follows that to prove the boundedness of the operator with the kernel (9.8) it is sufficient to prove the boundedness of the operator with the kernel

\[
k(t, x, y) = \frac{\partial K_{as}}{\partial x}(t, x, y) - \frac{\partial K_{as}}{\partial y}(t, x, y),
\]

because the differentiation of the expansion does not create any singularity.

To prove this boundedness, we can use the well-known small distance asymptotics representation of the distance (see, e.g., [20]),

\[
d^2(x, y) = (x - y, g(y)(x - y)) + O(|x - y|^3).
\]

It follows that (the cancellation of the main term occurs in subtraction)

\[
k(t, x, y) = \frac{1}{4t} K_{as}(t, x, y) O(|x - y|^2).
\]

By changing the integration variable \( y \) to \( z = (x - y)/\sqrt{t} \) in the formula

\[
\int_M k(t, x, y) f(y) \, dy, \quad f \in C(M_0),
\]

we can conclude that this operator is bounded. □

The following result is the direct consequence of Theorems 9.1 and 8.1 and the well-known fact that the operator \( \Delta_{LB} \) is self-dual with respect to the coupling given by the integration with respect to the volume measure \( dv \) on \( M \).
Theorem 9.2. The smoothing properties of the operators $S_t$ also hold in the integral norms; that is,
\begin{align}
\|\nabla S_t f\|_{L^1(M)} &\leq Ct^{-1/2}\|f\|_{L^1(M)}, \\
\|S_t \nabla f\|_{L^1(M)} &\leq Ct^{-1/2}\|f\|_{L^1(M)},
\end{align}
(9.15)
and the commutators $[\nabla, S_t]$ extend to bounded operators in $L^1(M)$, uniformly for any compact time interval.

For the stochastic control of diffusions on $(M, g)$ with the second-order part being fixed as $\Delta_{LB}$, in the case where control is carried out via the drift only (drift control of the Brownian motion on $M$), the corresponding HJB equation is the equation
\[
\frac{\partial f}{\partial t}(t, x) = \Delta_{LB} f(t, x) + H(x, \nabla f(t, x)),
\]
(9.17)
where the Hamiltonian $H(x, p)$, $x \in M$, $p \in T^* M_x$, is a function on the cotangent bundle $T^* M$ of the form
\[
H(x, p) = \sup_{u \in U} [(g(x, u), p) + J(x, u)];
\]
(9.18)
here $U$ is a compact set of possible controls and $J$ and $g$ are a continuous function and a vector field depending on $u$ as a parameter. In the case of zero-sum stochastic two-player games with the so-called Isaacs condition, the Hamiltonian function takes the form
\[
H(x, p) = \sup_{u \in U} \inf_{v \in V} [(g(x, u, v), p) + J(x, u, v)] = \inf_{v \in V} \sup_{u \in U} [(g(x, u, v), p) + J(x, u, v)].
\]
(9.19)
The possibility to exchange sup and inf here is called the Isaacs condition. It is fulfilled, in particular, when the control of two players can be separated in the sense that the Hamiltonian becomes
\[
H(x, p) = \sup_{u \in U} [(g_1(x, u), p) + J_1(x, u)] + \inf_{v \in V} [(g_2(x, v), p) + J_2(x, v)] + J_0(x).
\]
(9.20)

The fractional version of equation (9.17) is the fractional HJB equation
\[
D_a^\beta f(t, x) = \Delta_{LB} f(t, x) + H(x, \nabla f(t, x))
\]
(9.21)
(the fractional derivative acts on the variable $t$, while $\nabla$ and $\Delta_{LB}$ act on the variable $x \in M$), which describes the cost functions for the scaled continuous-time random walks with the spatial motion governed by $\Delta_{LB}$ (see [32]).

The McKean–Vlasov equation on $M$ describing a nonlinear diffusion on $M$ (with the operator $\Delta_{LB}$ as the main part and a nonlinear drift $h$) is the equation of the form
\[
\frac{\partial f}{\partial t}(t, x) = \Delta_{LB} f(t, x) + (h(x, f(t, \cdot)), \nabla f(t, x)),
\]
(9.22)
where $h$ maps pairs $(x, f(t, \cdot))$ to the elements of the tangent space $T_x M$ of $M$ at $x$ (so that the coupling $(h(x, f(t, \cdot)), \nabla f(t, x))$ yields a well-defined function on $M$). Thus $h$ can be looked at as a vector field on $M$ depending on $f$ as a parameter. The dependence of $h$ on $f$ via certain integrals, that is,
\[
h(x, f(t, \cdot)) = h \left( x, \int_M r(y)f(t, y) \, dv(y) \right),
\]
(9.23)
where $r$ is some bounded function on $M$, is typical in applications.
The fractional version of equation (9.22) is, of course, the equation
\[ D_{a+}^\beta f(t, x) = \Delta_{LB} f(t, x) + (h(x, f(t, \cdot)), \nabla f(t, x)). \] (9.24)

The following results are direct consequences of Theorems 9.1, 9.2, 4.1 and 4.2 (and Remark 9.1).

**Theorem 9.3.** (i) Let \( H(x, p) \) be a continuous function on the cotangent bundle \( T^*M \) of the compact Riemannian manifold \( (M, g) \) such that
\[ |H(x, p_1) - H(x, p_2)| \leq L_H \|p_1 - p_2\|_x \] (9.25)
for all \( x \) with a constant \( L_H \). Then for any \( Y \in C^1(M) \) there exists a unique mild solution of the Cauchy problem for equation (9.17) and a unique solution of the Cauchy problem for equation (9.21) with the initial condition \( f(0, \cdot) = Y \). These solutions depend Lipschitz continuously on the initial data in the norm of \( C^1(M) \).

(ii) Let \( h(x, f(t, \cdot)) \) be a uniformly bounded continuous vector field on \( M \) depending Lipschitz continuously on \( f \) in the norm of \( L_1(M) \):
\[ \|h(x, f_1(t, \cdot)) - h(x, f_2(t, \cdot))\|_x \leq L_h \|f_1(t, \cdot) - f_2(t, \cdot)\|_{L_1(M)} \] (9.26)
for all \( x \in M \) with a constant \( L_h \). Then for any \( Y \in W_1(M) \) there exists a unique mild solution of the Cauchy problem for equation (9.22) and a unique solution of the Cauchy problem for equation (9.24) with the initial condition \( f(0, \cdot) = Y \). These solutions depend Lipschitz continuously on the initial data in the norm of \( W_1(M) \).

10. FRACTIONAL FORWARD–BACKWARD SYSTEMS ON MANIFOLDS

Finally, we study the analog of system (8.6) on manifolds with the generator \( A = \Delta_{LB} \):
\[
\begin{cases}
D_{a+}^\beta b(t, x) = \Delta_{LB} b(t, x) + h(t, x, b(t, \cdot), u(t, x, \nabla f(t, x))) \nabla b(t, x), & b(a, \cdot) = Y, \\
D_{T-a}^\beta f(t, x) = -\Delta_{LB} f(t, x) + H(t, x, \nabla f(t, x), b_{\geq t}), & f(T, \cdot) = Z.
\end{cases}
\] (10.1)

The following result is a consequence of Theorems 8.2, 9.1 and 9.2.

**Theorem 10.1.** Let \( U \) be a compact set in a Euclidean space and \( u(t, x, p) \in U \) be a continuous function of the triple \( t \in \mathbb{R} \), \( x \in M \), \( p \in T^*M_x \) which is Lipschitz continuous in the last argument:
\[ |u(t, x, p) - u(t, x, \tilde{p})| \leq L_u \|p - \tilde{p}\|_x. \] (10.2)

Let \( h: \mathbb{R} \times M \times L_1(M) \times U \to TM \) be a continuous mapping such that \( h(t, x, b(t, \cdot), u) \in TM_x \) for all \( x, f \) and \( u \) which is Lipschitz continuous in the sense that
\[ \|h(t, x, b(t, \cdot), u) - h(t, x, \tilde{b}(t, \cdot), \tilde{u})\|_x \leq L_h \|b(t, \cdot) - \tilde{b}(t, \cdot)\|_{L_1(M)} + L_{hu}\|u - \tilde{u}\|. \] (10.3)

Let \( H(t, x, p, b_{\geq t}) \) be a continuous function with the arguments \( t \in \mathbb{R} \), \( x \in M \), \( p \in T^*M_x \) and \( b(\cdot) \in C([a, T], L_1(M)) \) which is Lipschitz continuous in the sense that
\[ |H(t, x, p, b_{\geq t}) - H(t, x, \tilde{p}, \tilde{b}_{\geq t})| \leq L_H \|p - \tilde{p}\|_x, \] (10.4)
\[ |H(t, x, p, b_{\geq t}) - H(t, x, p, \tilde{b}_{\geq t})| \leq L_H \|b_{\geq t} - \tilde{b}_{\geq t}\|_{C([a, T], L_1(M))}(1 + \|p\|_x). \] (10.5)

Then for all \( Y \) and \( Z \) there exists \( T_0 \) such that system (10.1) has a unique solution for all \( T \leq T_0 \).

**ACKNOWLEDGMENTS**

The authors are grateful to the referee for going through this article very carefully and providing them with a number of useful comments.
FUNDING

The work of V. N. Kolokoltsov (Sections 1–7) was supported by the Russian Science Foundation under grant no. 20-11-20119. The work of M. S. Troeva (Sections 8–10) was supported by the Ministry of Science and Higher Education of the Russian Federation (grant no. FSRG-2020-0006).

REFERENCES

1. O. P. Agrawal, “Generalized variational problems and Euler–Lagrange equations,” Comput. Math. Appl. 59 (5), 1852–1864 (2010).

2. D. Applebaum and R. S. Brockway, “$L_2$ properties of Lévy generators on compact Riemannian manifolds,” J. Theor. Probab. 34 (2), 1029–1042 (2021); arXiv:1907.11123v2 [math.PR].

3. T. Atanackovic, D. Dolicanin, S. Pilipovic, and B. Stankovic, “Cauchy problems for some classes of linear fractional differential equations,” Fract. Calc. Appl. Anal. 17 (4), 1039–1059 (2014).

4. Yu. Averboukh, “Deterministic limit of mean field games associated with nonlinear Markov processes,” Appl. Math. Optim. 81 (3), 711–738 (2020).

5. D. Azagra, J. Ferrera, and F. López-Mesas, “Nonsmooth analysis and Hamilton–Jacobi equations on Riemannian manifolds,” J. Funct. Anal. 220 (2), 304–361 (2005).

6. D. Baleanu, K. Diethelm, E. Scalas, and J. J. Trujillo, Fractional Calculus: Models and Numerical Methods, 2nd ed. (World Scientific, Hackensack, NJ, 2017), Ser. Complex. Nonlinearity Chaos 5.

7. A. Bensoussan, J. Frehse, and Ph. Yam, Mean Field Games and Mean Field Type Control Theory (Springer, New York, 2013), Springer Briefs Math.

8. R. Carmona and F. Delarue, Probabilistic Theory of Mean Field Games with Applications. I: Mean Field FBSDEs, Control, and Games (Springer, Cham, 2018), Probab. Theory Stoch. Model. 83.

9. R. Carmona and F. Delarue, Probabilistic Theory of Mean Field Games with Applications. II: Mean Field Games with Common Noise and Master Equations (Springer, Cham, 2018), Probab. Theory Stoch. Model. 84.

10. A. Cesaroni, M. Cirant, S. Dipierro, M. Novaga, and E. Valdinoci, “On stationary fractional mean field games,” J. Math. Pures Appl. 122, 1–22 (2019).

11. E. B. Davies, “Pointwise bounds on the space and time derivatives of heat kernels,” J. Oper. Theory 21 (2), 367–378 (1989).

12. T. E. Duncan and B. Pasik-Duncan, “Solvable stochastic differential games in rank one compact symmetric spaces,” Int. J. Control 91 (11), 2445–2450 (2018).

13. D. A. Gomes, E. A. Pimentel, and V. Voskanyan, Regularity Theory for Mean-Field Game Systems (Springer, Cham, 2016), Springer Briefs Math.

14. A. Grigor’yan, Heat Kernel and Analysis on Manifolds (Am. Math. Soc., Providence, RI, 2009), AMS/IP Stud. Adv. Math. 47.

15. Handbook of Fractional Calculus with Applications, Vol. 2: Fractional Differential Equations, Ed. by A. Kochubei and Yu. Luchko (De Gruyter, Berlin, 2019).

16. M. E. Hernández-Hernández and V. N. Kolokoltsov, “On the solution of two-sided fractional ordinary differential equations of Caputo type,” Fract. Calc. Appl. Anal. 19 (6), 1393–1413 (2016).

17. M. Huang, R. Malhamé, and P. E. Caines, “Large population stochastic dynamic games: Closed-loop McKean–Vlasov systems and the Nash certainty equivalence principle,” Commun. Inf. Syst. 6 (3), 221–252 (2006).

18. V. Kiryakova, Generalized Fractional Calculus and Applications (Longman Scientific & Technical, Harlow, 1994), Pitman Res. Notes Math. Ser. 301.

19. V. N. Kolokoltsov, Nonlinear Markov Processes and Kinetic Equations (Cambridge Univ. Press, Cambridge, 2010), Cambridge Tracts Math. 182.

20. V. N. Kolokoltsov, Markov Processes, Semigroups and Generators (De Gruyter, Berlin, 2011), De Gruyter Stud. Math. 38.
25. V. Kolokoltsov, Differential Equations on Measures and Functional Spaces (Birkhäuser, Cham, 2019), Birkhäuser Adv. Texts. Basler Lehrbücher.
26. V. N. Kolokoltsov, “Quantum mean-field games with the observations of counting type,” Games 12 (1), 7 (2021).
27. V. Kolokoltsov, F. Lin, and A. Mijatović, “Monte Carlo estimation of the solution of fractional partial differential equations,” Fract. Calc. Appl. Anal. 24 (1), 278–306 (2021).
28. V. N. Kolokoltsov and O. A. Malafeyev, Many Agent Games in Socio-Economic Systems: Corruption, Inspection, Coalition Building, Network Growth, Security (Springer, Cham, 2019), Springer Ser. Oper. Res. Financ. Eng.
29. V. N. Kolokoltsov and M. Troeva, “Regularity and sensitivity for McKean–Vlasov SPDEs,” AIP Conf. Proc. 1907, 030046 (2017).
30. V. N. Kolokoltsov and M. S. Troeva, “Regularity and sensitivity for McKean–Vlasov type SPDEs generated by stable-like processes,” Probab. Anal. Issues Anal. 7 (2), 69–81 (2018).
31. V. N. Kolokoltsov and M. Troeva, “On mean field games with common noise and McKean–Vlasov SPDEs,” Stoch. Anal. Appl. 37 (4), 522–549 (2019).
32. V. N. Kolokoltsov and M. A. Veretennikova, “A fractional Hamilton Jacobi Bellman equation for scaled limits of controlled continuous time random walks,” Commun. Appl. Ind. Math. 6 (1), e-484 (2014).
33. V. N. Kolokoltsov and M. A. Veretennikova, “Well-posedness and regularity of the Cauchy problem for nonlinear fractional in time and space equations,” Fract. Differ. Calc. 4 (1), 1–30 (2014); arXiv:1402.6735v1 [math.AP].
34. V. Kolokoltsov and W. Yang, “Existence of solutions to path-dependent kinetic equations and related forward–backward systems,” Open J. Optim. 2 (2), 39–44 (2013).
35. J.-M. Lasry and P.-L. Lions, “Jeux à champ moyen. I: Le cas stationnaire,” C. R., Math., Acad. Sci. Paris 343 (9), 619–625 (2006).
36. N. N. Leonenko, M. M. Meerschaert, and A. Sikorskii, “Correlation structure of fractional Pearson diffusions,” Comput. Math. Appl. 66 (5), 737–745 (2013).
37. M. Ludewig, “Strong short-time asymptotics and convolution approximation of the heat kernel,” Ann. Global Anal. Geom. 55 (2), 371–394 (2019).
38. C. Mantegazza and A. C. Mennucci, “Hamilton–Jacobi equations and distance functions on Riemannian manifolds,” Appl. Math. Optim. 47 (1), 1–25 (2002).
39. D. T. Nguyen, S. L. Nguyen, and N. H. Du, “On mean field systems with multi-classes,” Discrete Contin. Dyn. Syst. 40 (2), 683–707 (2020).
40. I. Podlubny, Fractional Differential Equations: An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and Some of Their Applications (Academic Press, San Diego, CA, 1999), Math. Sci. Eng. 198.
41. A. V. Pskhu, “Initial-value problem for a linear ordinary differential equation of noninteger order,” Sb. Math. 202 (4), 571–582 (2011) [transl. from Mat. Sb. 202 (4), 111–122 (2011)].
42. H. H. Schaefer, Banach Lattices and Positive Operators (Springer, Berlin, 1974), Grundl. Math. Wiss. 215.
43. Q. Tang and F. Camilli, “Variational time-fractional mean field games,” Dyn. Games Appl. 10 (2), 573–588 (2020).
44. Y. Zhou, Fractional Evolution Equations and Inclusions: Analysis and Control (Elsevier, Amsterdam, 2016).
45. X. Zhu, “The optimal control related to Riemannian manifolds and the viscosity solutions to Hamilton–Jacobi–Bellman equations,” Syst. Control Lett. 69, 7–15 (2014).

This article was submitted by the authors simultaneously in Russian and English.