Nonlinear Fokker–Planck equations as smooth Hilbertian gradient flows

Viorel Barbu* Michael Röckner†

In memory of Giuseppe Da Prato

Abstract

Under suitable assumptions on $\beta:\mathbb{R}\to\mathbb{R}$, $D:\mathbb{R}^d\to\mathbb{R}^d$ and $b:\mathbb{R}^d\to\mathbb{R}$, the nonlinear Fokker–Planck equation $u_t-\Delta\beta(u)+\mathrm{div}(Db(u)u)=0$, in $(0,\infty)\times\mathbb{R}^d$ where $D=-\nabla\Phi$, can be identified as a smooth gradient flow $\frac{d^+}{dt}\,u(t)+\nabla E_{u(t)}=0$, $\forall t>0$. Here, $E:\mathcal{P}^*\cap L^\infty(\mathbb{R}^d)\to\mathbb{R}$ is the energy function associated to the equation, where $\mathcal{P}^*$ is a certain convex subset of the space of probability densities. $\mathcal{P}^*$ is invariant under the flow and ∇E_u is the gradient of E, that is, the tangent vector field to $\mathcal{P}$ at u defined by $\langle\nabla E_u,z_u\rangle_u=\mathrm{diff}\,E_u\cdot z_u$ for all vector fields z_u on $\mathcal{P}^*$, where $\langle\cdot,\cdot\rangle_u$ is a scalar product on a suitable tangent space $\mathcal{T}_u(\mathcal{P}^*)\subset\mathcal{D}'(\mathbb{R}^d)$.

MSC: 60H15, 47H05, 47J05.
Keywords: Fokker–Planck equation, gradient flow, semigroup, stochastic equations, tensor metric.

1 Introduction

We are concerned here with the nonlinear Fokker–Planck equation (NFPE)

$$\begin{aligned}&u_t-\Delta\beta(u)+\mathrm{div}(Db(u)u)=0 \text{ in } (0,\infty)\times\mathbb{R}^d,\\&u(0,x)=u_0(x),\ x\in\mathbb{R}^d,\end{aligned}\tag{1.1}$$

*Al.I. Cuza University and Octav Mayer Institute of Mathematics of Romanian Academy, Iaşi, Romania. Email: vbarbu41@gmail.com
†Faculty of Mathematics, Bielefeld University, 33615 Bielefeld, Germany. E-mail: roeckner@math.uni-bielefeld.de

where $\beta : \mathbb{R} \to \mathbb{R}$, $D : \mathbb{R}^d \to \mathbb{R}^d$, $d \ge 1$, and $b : \mathbb{R} \to \mathbb{R}$ are assumed to satisfy the following hypotheses

(i) $\beta \in C^1(\mathbb{R})$, $\beta(0) = 0$, $0 < \gamma_1 \le \beta'(r) \le \gamma_2 < \infty$, $\forall r \in \mathbb{R}$.

(ii) $b \in C_b(\mathbb{R}) \cap C^1(\mathbb{R})$ and $b(r) \ge b_0 > 0$, $|b'(r)r + b(r)| \le \gamma_3 < \infty$, $\forall r \in \mathbb{R}$.

(iii) $D \in L^\infty(\mathbb{R}^d; \mathbb{R}^d) \cap W^{1,1}_{\rm loc}(\mathbb{R}^d; \mathbb{R}^d)$ and $\operatorname{div} D \in L^2(\mathbb{R}^d) + L^\infty(\mathbb{R}^d)$.

(iv) $D = -\nabla\Phi$, where $\Phi \in C(\mathbb{R}^d) \cap W^{2,1}_{\rm loc}(\mathbb{R}^d)$, $\Phi \ge 1$, $\lim\limits_{|x|\to\infty} \Phi(x) = +\infty$, $\Phi^{-m} \in L^1(\mathbb{R}^d)$ for some $m \ge 2$.

NFPE (1.1) is modeling the so called *anomalous diffusion* in statistical physics (see, e.g., [13]) and also describes the dynamics of Itô stochastic processes in terms of their probability densities. In fact, if u is a distributional solution to (1.1), such that $t \to u(t)dx$ is weakly continuous and $u(t) \in \mathcal{P}$, $\forall t \ge 0$, then there is a probabilistically weak solution X_t to the McKean–Vlasov stochastic differential equation

$$dX_t = D(X_t)b(u(t,X_t))dt + \left(\frac{2\beta(u(t,X_t))}{u(t,X_t)}\right)^{\frac{1}{2}} dW_t, \tag{1.2}$$

on a probability space $(\Omega, \mathcal{F}, \mathbb{P}, W_t)$ with normal filtration $(\mathcal{F}_t)_{t\ge0}$. More exactly, one has $\mathcal{L}_{X_t} \equiv u(t,x)$, where $\mathcal{L}_{X_t}$ is the density of the marginal law $\mathbb{P} \circ X_t^{-1}$ of X_t with respect to the Lebesgue measure (see [7], [10]).

The function $u : [0,\infty) \times \mathbb{R}^d \to \mathbb{R}$ is called a *mild solution* to (1.1) if it is L^1-continuous, that is $u \in C([0,\infty); L^1(\mathbb{R}^d))$, and

$$u(t) = \lim_{h\to0} u_h(t) \text{ in } L^1(\mathbb{R}^d),\ \forall t \ge 0 \tag{1.3}$$

where, for each $T > 0$, $u_h : (0,T) \to L^1(\mathbb{R}^d)$ is defined by

$$\begin{aligned} u_h(t) &= u_h^j,\ t \in [jh,(j+1)h),\ j = 0,1,...,\left[\tfrac{T}{h}\right], \\ u_h^{j+1} + hAu_h^{j+1} &= u_h^j,\ j = 0,1,...,\left[\tfrac{T}{h}\right];\ u_h^0 = u_0. \end{aligned} \tag{1.4}$$

Here, $A : L^1(\mathbb{R}^d) \to L^1(\mathbb{R}^d)$ is the operator

$$\begin{aligned} Ay &= -\Delta\beta(y) + \operatorname{div}(Db(y)y) \text{ in } \mathcal{D}'(\mathbb{R}^d);\ y \in D(A), \\ D(A) &= \{y \in L^1(\mathbb{R}^d);\ -\Delta\beta(y) + \operatorname{div}(Db(y)y) \in L^1(\mathbb{R}^d)\}. \end{aligned} \tag{1.5}$$

As shown in [9] (see also [7]–[8], [10]), under the above hypotheses (as a matter of fact, for less restrictive assumptions), the domain $D(A)$ is dense in $L^1(\mathbb{R}^d)$, that is, $\overline{D(A)} = L^1(\mathbb{R}^d)$, and the operator A is m-accretive in $L^1(\mathbb{R}^d)$, which means that (see, e.g., [4], [5])

$$
\begin{aligned}
&R(I+\lambda A) = L^1(\mathbb{R}^d),\ \forall \lambda > 0,\\
&\|(I+\lambda A)^{-1}y_1 - (I+\lambda A)^{-1}y_2\|_{L^1(\mathbb{R}^d)} \le \|y_1-y_2\|_{L^1(\mathbb{R}^d)},\\
&\qquad\qquad \forall \lambda > 0,\ y_1,y_2 \in L^1(\mathbb{R}^d).
\end{aligned}
$$

Then, by the Crandall & Liggett theorem (see [4], [5], p. 56) the Cauchy problem

$$\frac{du}{dt} + Au = 0,\ t \ge 0;\quad u(0) = u_0, \tag{1.6}$$

has, for each $u_0 \in L^1(\mathbb{R}^d)$ a unique solution $u = u(t,u_0)$ in the mild sense (1.3)–(1.4). Equivalently,

$$u(t,u_0) = \lim_{n\to\infty} \left(I + \frac{t}{n}A\right)^{-n} u_0 \text{ in } L^1(\mathbb{R}^d), \tag{1.7}$$

uniformly on the compact intervals of $[0,\infty)$.

Moreover, the map $t \to u(t,u_0)$, denoted $S(t)u_0$, is a *continuous semi-group* of contractions on $L^1(\mathbb{R}^d)$, that is,

$$
\begin{aligned}
&S(t+s) = S(t)S(s) \text{ for all } s,t \ge 0,\\
&\|S(t)u_1 - S(t)u_2\|_{L^1(\mathbb{R}^d)} \le \|u_1-u_2\|_{L^1(\mathbb{R}^d)},\ \forall t \ge 0,\ u_1,u_2 \in L^1(\mathbb{R}^d),\\
&\lim_{t\to 0} S(t)u_0 = u_0 \text{ in } L^1(\mathbb{R}^d).
\end{aligned}
$$

Note also (see [7]–[10]) that

$$S(t)(L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)) \subset L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d),\ \forall t \ge 0, \tag{1.8}$$

$$S(t)(L^1(\mathbb{R}^d) \cap L^1(\mathbb{R}^d;\Phi dx)) \subset L^1(\mathbb{R}^d) \cap L^1(\mathbb{R}^d;\Phi dx), \tag{1.9}$$

$$S(t)u_0 \in L^\infty((0,T)\times\mathbb{R}^d),\ \forall T > 0,\ \forall u_0 \in L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d), \tag{1.10}$$

and $S(t)\mathcal{P} \subset \mathcal{P}$, $\forall t \ge 0$, where

$$\mathcal{P} = \left\{ y \in L^1(\mathbb{R}^d),\ y(x) \ge 0,\ \text{a.e. } x \in \mathbb{R}^d;\ \int_{\mathbb{R}^d} y(x)dx = 1 \right\}. \tag{1.11}$$

We also note that, though $t \to u(t) = S(t)u_0$ is not differentiable, it is, however, a Schwartz-distributional solution to (1.1), that is,

$$\begin{aligned}\int_0^\infty \int_{\mathbb{R}^d} (u\varphi_t + \beta(u)\Delta_x\varphi + b(u)uD\cdot\nabla_x\varphi)dx\,dt \\ + \int_{\mathbb{R}^d} u_0(x)\varphi(0,x)dx = 0,\end{aligned} \tag{1.12}$$

for all $\varphi \in C_0^\infty([0,\infty)\times\mathbb{R}^d)$.

Moreover, as shown in [9] (see also [10]), $S(t)u_0$ is the unique distributional solution to NFPE (1.1) in the class of functions $u \in L^1((0,\infty)\times\mathbb{R}^d)\cap L^\infty((0,\infty)\times\mathbb{R}^d)$ such that $t \to u(t)dx$ is weakly continuous on $[0,\infty)$. In particular, this implies (see, e.g., [9] and [10], Chapter 5) that the McKean–Vlasov equation (1.2) has a unique strong solution X_t with the marginal law $u(t,\cdot)$.

The purpose of this work is to represent the solution $t \to S(t)u_0$ to (1.1) as a *subgradient flow* of the entropy functional (energy)

$$\begin{aligned}E(u) &= \int_{\mathbb{R}^d} (\eta(u(x)) + \Phi(x)u(x))dx,\ u \in \mathcal{P}\cap L^\infty(\mathbb{R}^d)\cap L^1(\mathbb{R}^d;\Phi dx),\\ \eta(r) &= \int_0^r \int_1^s \frac{\beta'(\tau)}{\tau b(\tau)}\,d\tau\,ds,\ r \ge 0,\end{aligned} \tag{1.13}$$

with the tangent space $\mathcal{T}_u(\mathcal{P}^*) \subset \mathcal{D}'(\mathbb{R}^d)$ defined in (3.1) below, for $u \in \mathcal{P}^*$. Here,

$$\mathcal{P}^* = \left\{\begin{array}{r} u \in \mathcal{P}\cap L^\infty\cap L^1(\mathbb{R}^d;\Phi dx);\sqrt{u}\in H^1(\mathbb{R}^d), \dfrac{\psi}{u}\in L^1(\mathbb{R}^d) \\ \text{for some } \psi \in \mathcal{X}\end{array}\right\}, \tag{1.14}$$

where we set $\frac{1}{0} := +\infty$ and

$$\mathcal{X} = \left\{\psi \in C^2(\mathbb{R}^d)\cap C_b(\mathbb{R}^d)\cap L^1(\mathbb{R}^d), \psi > 0, \frac{\nabla\psi}{\psi}\in L^\infty(\mathbb{R}^d), \frac{1}{\psi}\in L^1_{\rm loc}(\mathbb{R}^d)\right\}. \tag{1.15}$$

We also note that the function E is convex and lower semicontinuous on $L^2(\mathbb{R}^d)$ with the domain

$$D(E) = \{u \in \mathcal{P}\cap L^\infty(\mathbb{R}^d)\cap L^\infty\cap L^1(\mathbb{R}^d;\Phi dx)\}.$$

The class $\mathcal{X}$ is clearly nonempty and, in particular, it contains all functions ψ of the form $\psi(x) = (\alpha_1|x|^m + \alpha_2)^{-1}$, α_1, α_2 and $m > d$ and, therefore, since $\mathcal{X}$ is an algebra containing the constants, it is a rich class of functions. Hence, so is $\mathcal{P}^*$ since if $\psi \in \mathcal{X}$, $\psi > 0$, $u := \psi^2 \left(\int_{\mathbb{R}^d} \psi^2 dx\right)^{-1}$ is easily checked to be in $\mathcal{P}^*$. We also note that $\mathcal{P}^*$ is convex. The gradient flow representation means that, for $u(t) = S(t)u_0$, $u_0 \in \mathcal{P}^*$, we have

$$\frac{d}{dt}\, u(t) = -\nabla E_{u(t)},\ t > 0, \tag{1.16}$$

where $\nabla E_u \in \mathcal{T}_u(\mathcal{P}^*)$ is the gradient of E in the sense of the Riemannian type geometry of $\mathcal{P}$ to be defined later on. Such a result was recently established in [17] (see also [1], [2], [19]) on the manifold $\mathcal{P}$ endowed with the topology of weak convergence of probability measures and tangent bundle $L^2(\mathbb{R}^d; \mathbb{R}^d; \mu)_{\mu \in \mathcal{P}}$ and in the fundamental work [16] for the classical porous media equation. But we want to emphasize that we consider here the smaller space $\mathcal{P}^* \subset \mathcal{P}$ with the tangent bundle $(\mathcal{T}_u(\mathcal{P}^*))_{u \in \mathcal{P}^*}$ defined in (3.1) and scalar product (3.2) which is different from the one in [1], [2], [16], [17], [19]. Herein, we shall obtain a representation of the form (1.16) for NFPE (1.1). This result is based on the smoothing effect on initial data of the semigroup $S(t)$ in the space $H^{-1}(\mathbb{R}^d)$ which will be proved in Section 1. As a matter of fact, the space $H^{-1}(\mathbb{R}^d)$ is a viable alternative to $L^1(\mathbb{R}^d)$ for proving the well-posedness of NFPE (1.1). In fact, as seen below, the operator (1.5) has a quasi-m-accretive version in $H^{-1}(\mathbb{R}^d)$, which generates a C_0-semigroup of quasi-contractions which coincides with $S(t)$ on $L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$.

We recall that (see, e.g., [4], [5]), if H is a Hilbert space with the scalar product $(\cdot,\cdot)_H$ and norm $|\cdot|$, the operator $B : D(B) \subset H \to H$ is said to be m-accretive if

$$(Bu_1 - Bu_2, u_1 - u_2) \geq 0,\ \forall u_i \in D(B),\ i = 1, 2,$$

and $R(I + \lambda B) = H$, $\forall \lambda > 0$. It is said to be quasi m-accretive if $B + \omega f$ is m-accretive for some $\omega \geq 0$.

Notation. $L^p(\mathbb{R}^d)$, $1 \leq p \leq \infty$ (denoted L^p) is the space of Lebesgue measurable and p-integrable functions on $\mathbb{R}^d$, with the standard norm $|\cdot|_p$. $(\cdot,\cdot)_2$ denotes the inner product in L^2. By $L^p_{\rm loc}$ we denote the corresponding local space. Let $C^k(\mathbb{R})$ denote the space of continuously differentiable functions up to order k and $C_b(\mathbb{R})$ the space of continuous and bounded functions on $\mathbb{R}$.

For any open set $\mathcal{O} \subset \mathbb{R}^m$ let $W^{k,p}(\mathcal{O})$, $k \geq 1$, denote the standard Sobolev space on $\mathcal{O}$ and by $W^{k,p}_{\rm loc}(\mathcal{O})$ the corresponding local space. We set $W^{1,2}(\mathcal{O}) = H^1(\mathcal{O})$, $W^{2,2}(\mathcal{O}) = H^2(\mathcal{O})$, $H^1_0(\mathcal{O}) = \{u \in H^1(\mathcal{O}),\ u = 0 \text{ on } \partial\mathcal{O}\}$, where $\partial\mathcal{O}$ is the boundary of $\mathcal{O}$. By $H^{-1}(\mathcal{O})$ we denote the dual space of $H^1_0(\mathcal{O})$ (of $H^1(\mathbb{R}^m)$, respectively, if $\mathcal{O} = \mathbb{R}^m$). We shall also set $H^1 = H^1(\mathbb{R}^d)$ and $H^{-1} = H^{-1}(\mathbb{R}^d)$. $C^\infty_0(\mathcal{O})$ is the space of infinitely differentiable real-valued functions with compact support in $\mathcal{O}$ and $\mathcal{D}'(\mathcal{O})$ is the dual of $C^\infty_0(\mathcal{O})$, that is, the space of Schwartz distributions on $\mathcal{O}$. Lip$(\mathbb{R})$ is the space of real-valued Lipschitz functions on $\mathbb{R}$ with the norm denoted by $|\cdot|_{\rm Lip}$. The space H^{-1} is endowed with the scalar product

$$\langle y_1, y_2\rangle_{-1} = ((I-\Delta)^{-1}y_1, y_1)_2,\ \forall y_1, y_2 \in H^{-1},$$

and the Hilbert norm $|y|^2_{-1} = \langle y, y\rangle_{-1}$. By ${}_{H^{-1}}(\cdot,\cdot)_{H^1}$ we denote the duality pairing on $H^1 \times H^{-1}$. If Y is a Banach space, then $C([0,\infty);Y)$ is the space of continuous functions $y : [0,\infty) \to Y$ and $C_w([0,\infty);Y)$ is the space of weakly continuous Y-valued functions. Furthermore, let $C^\infty_0([0,\infty)\times\mathbb{R}^d)$ denote the space of all $\varphi \in C^\infty([0,\infty)\times\mathbb{R}^d)$ such that *support* $\varphi \subset K$, where K is compact in $[0,\infty)\times\mathbb{R}^d$. If $u : [0,\infty) \to H^{-1}$ is a given function, we shall denote its H^{-1}-strong derivative in t by $\frac{du}{dt}(t)$, and the right derivative by $\frac{d^+}{dt}u(t)$. We shall also use the following notations

$$\beta'(r) \equiv \frac{d}{dr}\beta(r),\ b'(r) = \frac{d}{dr}b(r),\ b^*(r) \equiv b(r)r,\ r \in \mathbb{R},$$
$$y_t = \frac{\partial}{\partial t}y,\ \nabla y = \left\{\frac{\partial y}{\partial x_i}\right\}_{i=1}^d,\ \Delta y = \sum_{i=1}^d \frac{\partial^2}{\partial^2 x_i}y,$$

$$\operatorname{div} y = \sum_{i=1}^d \frac{\partial y_i}{\partial x_i},\ y = \{y_i\}_{i=1}^d,$$

for $y = y(t,x)$, $(t,x) \in [0,\infty)\times\mathbb{R}^d$, where Δ and div are taken in the sense of the distribution space $\mathcal{D}'(\mathbb{R}^d)$.

2 The H^1-regularity of the semigroup $S(t)$

Consider the semigroup $S(t) : L^1 \to L^1$ defined earlier by the exponential formula (1.7). Define the operator $A^* : H^{-1} \to H^{-1}$,

$$A^*y = -\Delta\beta(y) + \operatorname{div}(Db^*(y)),\ \forall y \in D(A^*), \tag{2.1}$$

with the domain $D(A^*) = H^1$. More precisely, for each $y \in H^1$, $A^* y \in H^{-1}$ is defined by

$$ {}_{H^{-1}}(A^* y, \varphi)_{H^1} = \int_{\mathbb{R}^d} (\nabla \beta(y) - D b^*(y)) \cdot \nabla \varphi \, dx, \ \forall \varphi \in H^1. \tag{2.2} $$

As mentioned earlier, the semigroup $S(t)$ is not differentiable in L^1, but as shown below it is, however, H^{-1}-differentiable on the right on $(0, \infty)$.

Namely, we have

Theorem 2.1. *Assume that Hypotheses* (i)–(iv) *hold. Then, for each* $u_0 \in \mathcal{P} \cap L^\infty$, *the function* $u(t) = S(t)u_0$ *is in* $C([0, \infty); H^{-1}) \cap C_w([0, \infty); L^2)$, *it is* H^{-1}*-right differentiable on* $(0, \infty)$ *with* $\frac{d^+}{dt}\, u(t)$ *being* H^{-1}*-continuous from the right on* $(0, \infty)$, $S(t)u_0 \in H^1$, $t > 0$, *and*

$$ \frac{d^+}{dt}\, S(t)u_0 + A^* S(t)u_0 = 0, \ \forall t > 0. \tag{2.3} $$

Furthermore, $S(t)u_0 \in \mathcal{P} \cap L^\infty$, $\forall t \geq 0$, $\frac{d}{dt}\, S(t)u_0$ *exists on* $(0, \infty) \setminus N$, *where* N *is an at most countable subset of* $(0, \infty)$,

$$ \frac{d}{dt}\, S(t)u_0 + A^* S(t)u_0 = 0, \ \forall t \in (0, \infty) \setminus N, \tag{2.4} $$

and $t \to A^* S(t)u_0$ *is* H^{-1}*-continuous on* $(0, \infty) \setminus N$.

Moreover, $\sqrt{S(t)u_0} \in H^1(\mathbb{R}^d)$, *a.e.* $t > 0$, *that is,*

$$ \frac{\nabla S(t)u_0}{\sqrt{S(t)u_0}} \in L^2, \ \text{a.e. } t \in (0, \infty), \tag{2.5} $$

$$ E(S(t)u_0) < \infty, \ \text{a.e. } t \in (0, \infty). \tag{2.6} $$

for all $u_0 \in \mathcal{P}$ *such that* $u_0 \log u_0 \in L^1$. *If* $u_0 \in H^1$, *then* (2.3) *holds for all* $t \geq 0$, $t \to S(t)u_0$ *is locally* H^{-1}*-Lipschitz, on* $[0, \infty)$ *and* $u(t) \in H^1$, $\forall t \geq 0$.

Finally, if $u_0 \in \mathcal{P}^*$, *then*

$$ S(t)u_0 \in \mathcal{P}^*, \ \forall t \geq 0. \tag{2.7} $$

In particular, it follows by Theorem 2.1 that the semigroup $S(t)$ is generated by the opertor $-A^*$ in the space H^{-1}.

We shall prove Theorem 2.1 in several steps, the first one being the following lemma.

Lemma 2.2. *The operator A^* is quasi-m-accretive in H^{-1}, that is, $A^*+\omega I$ is m-accretive for some $\omega \ge 0$.*

Proof. We have

$$\begin{aligned}
&\langle A^*y_1 - A^*y_2, y_1-y_2\rangle_{-1} \qquad\qquad (2.8)\\
&\quad = (\beta(y_1)-\beta(y_2), y_1-y_2)_2 - (\beta(y_1)-\beta(y_2), (I-\Delta)^{-1}(y_1-y_2))_2\\
&\qquad\qquad + (D(b^*(y_1)-b^*(y_2)), \nabla(I-\Delta)^{-1}(y_1-y_2))_2\\
&\quad \ge \gamma_1|y_1-y_2|_2^2 - \gamma_2|y_1-y_2|_{-1}|y_1-y_2|_2 - |D|_\infty|b^*|_{\rm Lip}|y_1-y_2|_2|y_1-y_2|_{-1}\\
&\quad \ge -\omega|y_1-y_2|_{-1}^2,\ \forall y_1,y_2\in D(A^*),
\end{aligned}$$

for a suitable chosen $\omega\ge 0$ and so $A^*+\omega I$ is accretive in H^{-1}. (Here, we have used the inequality $|\nabla(I-\Delta)^{-1}(y_1-y_2)|_2 \le |y_1-y_2|_{-1}$.) Now, we shall prove that $R(I+\lambda A^*) = H^{-1}$ for $\lambda\in(0,\lambda_0)$, where λ_0 is suitably chosen. For this purpose, we fix $f\in H^{-1}$ and consider the equation

$$y-\lambda\Delta\beta(y)+\lambda{\rm div}(Db^*(y)) = f \text{ in } \mathcal{D}'(\mathbb{R}^d),\ y\in L^2. \qquad (2.9)$$

The latter can be written as

$$G_\lambda(y) = (I-\Delta)^{-1}f, \qquad (2.10)$$

where $G_\lambda : L^2\to L^2$ is the operator

$$G_\lambda(y) = \lambda\beta(y) + (I-\Delta)^{-1}y - \lambda(I-\Delta)^{-1}{\rm div}(Db^*(y)) - \lambda(I-\Delta)^{-1}\beta(y),\ \forall y\in L^2,$$

which by Hypotheses (i)–(iii) is continuous. Then, by (i)–(iii) we have

$$\begin{aligned}
&(G_\lambda(y_1)-G_\lambda(y_2), y_1-y_2)_2\\
&\qquad \ge \lambda\gamma_1|y_1-y_2|_2^2 + |y_1-y_2|_{-1}^2 - \lambda\gamma_2|y_1-y_2|_{-1}|y_1-y_2|_2\\
&\qquad\qquad - \lambda|D|_\infty|b^*|_{\rm Lip}|y_1-y_2|_2|y_1-y_2|_{-1}\\
&\qquad \ge \frac{1}{2}(\lambda\gamma_1 - \gamma_2^2\gamma\lambda^2 - \lambda^2|D|_\infty^2|b^*|_{\rm Lip}^2)|y_1-y_2|_2^2 + \frac{1}{2}|y_1-y_2|_{-1}^2\\
&\qquad \ge \alpha|y_1-y_2|_2^2,\ \forall y_1,y_2\in L^2,
\end{aligned}$$

for some $\alpha>0$ and $0<\lambda<\lambda_0$ with λ_0 sufficiently small. Hence, the operator G_λ is monotone and coercive in the space L^2. Since it is also continuous, we infer that it is surjective (see, e.g., [4], p. 37) and, therefore, $R(G_\lambda) = L^2$ for $0 < \lambda < \lambda_0$). Hence, (2.10) (equivalently (2.9)) has a solution $y \in L^2$ for $\lambda \in (0, \lambda_0)$ and $\beta(y) \in H^1$. Then, by (i) it follows that $y \in H^1$ and so $y \in D(A^*)$. Hence, A^* is quasi-m-accretive in H^{-1}. $\square$

Lemma 2.2 implies that there is a C_0-continuous nonlinear semigroup $S^*(t) : H^{-1} \to H^{-1}$, $t \geq 0$, which is generated by $-A^*$. This means (see, e.g., [4], p. 146 or [5], p. 56) that

$$S^*(t)u_0 = \lim_{n\to\infty} \left(I + \frac{t}{n}\, A^*\right)^{-n} u_0 \;\text{ in } H^{-1},\ \forall t \geq 0,\ \forall u_0 \in H^{-1}, \qquad (2.11)$$

uniformly on compact intervals. Moreover, for all $u_0 \in D(A^*) = H^1$ we have $S^*(t)u_0 \in D(A^*)$, $[0,\infty) \ni t \mapsto S^*(t)u_0 \in H^{-1}$ is locally Lipschitz and,

$$\begin{aligned} \frac{d^+}{dt}\, S^*(t)u_0 + A^*S^*(t)u_0 &= 0, \quad \forall t \geq 0, & (2.12)\\ \frac{d}{dt}\, S^*(t)u_0 + A^*S^*(t)u_0 &= 0, \quad \text{a.e. } t > 0, & (2.13)\end{aligned}$$

and the function $t \to \frac{d^+}{dt}\, S^*(t)u_0$ is continuous from the right in the H^{-1}-topology. Taking into account (2.2), we can rewrite (2.13) as

$$\begin{aligned} &\frac{d}{dt}\int_{\mathbb{R}^d} (S^*(t)u_0)(x)\varphi(x)dx + \int_{\mathbb{R}^d} (\nabla\beta(S^*(t)u_0)(x)) \\ &\quad -D(x)b^*((S^*(t)u_0)(x)) \cdot \nabla\varphi(x)dx = 0,\ \text{a.e. } t > 0,\ \forall \varphi \in H^1. \end{aligned} \qquad (2.14)$$

We also note that the semigroup $S^*(t)$ is quasi-contractive on H^{-1}, that is,

$$|S^*(t)u_0 - S^*(t)\bar{u}_0|_{-1} \leq \exp(\omega t)|u_0 - \bar{u}_0|_{-1},\ \forall t \geq 0,\ \forall u_0, \bar{u}_0 \in H^{-1},$$

for some $\omega \geq 0$. Moreover, we have for all $u_0 \in L^2$ and $T > 0$,

$$|S^*(t)u_0|_2^2 + \int_0^t |\nabla(S^*(s)u_0)|_2^2 ds \leq C_T|u_0|_2^2,\ \forall t \in [0,T]. \qquad (2.15)$$

Here is the argument. By (2.11) we have, for all $T > 0$,

$$S^*(t)u_0 = \lim_{h\to 0} v_h(t) \;\text{ in } H^{-1},\ \forall t \in (0,T), \qquad (2.16)$$

where

$$\begin{aligned} v_h(t) &= v_h^j,\ \forall t \in [jh, (j+1)h),\ j = 0, 1, ..., N_h = \left[\tfrac{T}{h}\right], \\ v_h^{j+1} + hA^*v_h^{j+1} &= v_h^j,\ j = 0, 1, ..., N_h; \quad v_h^0 = u_0. \end{aligned} \qquad (2.17)$$

Since $v_h^0 = u_0 \in L^2$, we get by (2.17)

$$(\beta(v_h^{j+1}), v_h^{j+1} - v_h^j)_2 + h|\nabla\beta(v_h^{j+1})|_2^2 = h(\nabla\beta(v_h^{j+1}), Db^*(v_h^{j+1}))_2$$
$$\leq \frac{h}{2}\,|\nabla\beta(v_h^{j+1})|_2^2 + \frac{h}{2}\,(|D|^\infty|b^*|_\infty|v_h^{j+1}|_2)^2.$$

By (i), this yields

$$\int_{\mathbb{R}^d} j(v_h^{j+1})dx + \frac{1}{2}\,\gamma_1^2 h \sum_{k=1}^{j+1} |\nabla(v_h^k)|_2^2 \leq \int_{\mathbb{R}^d} j(u_0)dx + Ch\sum_{k=1}^{j+1} |v_h^k|_2^2,$$

where $j(r) = \int_0^r \beta(s)ds$. Since $\frac{1}{2}\,\gamma_1 r^2 \leq j(r) \leq \frac{1}{2}\,\gamma_2 r^2$, $\forall r \in \mathbb{R}$, we have

$$|v_h(t)|_2^2 + \int_0^t |\nabla v_h(s)|_2^2 ds \leq C\left(\int_0^t |v_h(s)|_2^2 ds + |u_0|_2^2\right),\ t \in (0,T).$$

Hence

$$|v_h(t)|_2^2 + \int_0^t |\nabla v_h(s)|_2^2 ds \leq C|u_0|_2^2,\ \forall t \in (0,T),\ h > 0.$$

Therefore, by (2.16) and by the weak-lower semicontinuity of the $L^2(0,T;H^1)$-norm, (2.15) follows. Hence, $S^*(t)u_0 \in H^1$, a.e. $t > 0$, and so, by the semigroup property, $S^*(t+s) = S^*(t)S^*(s)$, $t, s \geq 0$, we infer that $S^*(t)$ has a smoothing effect on initial data, that is,

$$S^*(t)u_0 \in H^1 = D(A^*),\ \forall t > 0, u_0 \in L^2. \tag{2.18}$$

Then, by (2.12) it follows that $t \mapsto S^*(t)u_0$ is H^{-1}-continuous on $(0,T)$ for all $u_0 \in L^2$, hence $t \mapsto |S^*(t)u_0|_2$ is lower semicontinuous on $(0,T)$. Furthermore, (2.18) implies

$$\frac{d^+}{dt}\,S^*(t)u_0 + A^*S^*(t)u_0 = 0,\ \forall u_0 \in L^2,\ \forall t > 0, \tag{2.19}$$

and that the function $t \to \frac{d^+}{dt}\,S^*(t)u_0 = -A^*S^*(t)u_0$ is H^{-1}-right continuous on $(0,\infty)$. Since $S^*(\cdot)u_0 \in L^\infty(0,T;L^2) \cap C([0,T];H^{-1})$, it follows that

$$\sup_{t\in[0,T]} |S^*(t)u_0|_2 \leq \operatorname*{ess\,sup}_{t>0} |S^*(t)u_0|_2 \vee |S^*(T)u_0|_2 + |u_0|_2 < \infty$$

and hence we obtain by the uniqueness of limits that the function $t \to S^*(t)u_0$ is L^2-weakly continuous, that is, $S^*(\cdot)u_0 \in C_w([0,T];L^2)$, $\forall T > 0$. We set $u_h(t) = u(t+h) - u(t)$, $u(t) \equiv S^*(t)u_0$, $\forall t \in [0,T]$, $h > 0$, $u_0 \in L^2$. By (2.19) we have

$$\frac{d^+}{dt}\, u_h(t) + A^*u(t+h) - A^*u(t) = 0,\ \forall t \in (0,T].$$

This yields

$$\frac{1}{2}\,\frac{d^+}{dt}\,|u_h(t)|^2_{-1} + \langle A^*u(t+h) - A^*u(t), u_h(t)\rangle_{-1} = 0$$

and, therefore, by (2.8)

$$\frac{1}{2}\,\frac{d^+}{dt}\,|u_h(t)|^2_{-1} \le \omega |u_h(t)|^2_{-1},\ \forall t \in (0,T].$$

Hence, for all $h > 0$, we have

$$|u_h(t)|_{-1}\exp(-\omega t) \le |u_h(s)|_{-1}\exp(-\omega s),\ 0 < s < t < T,$$

and, therefore, the function $t \to \exp(-\omega t)|A^*S^*(t)u_0|_{-1}$ is monotonically decreasing on $(0,\infty)$ and so it is everywhere continuous on $(0,\infty)$, except for a countable set $N \subset (0,\infty)$.

Since the continuity points of $t \to \exp(-\omega t)A^*S^*(t)u_0$ coincide with that of $t \to \exp(-\omega t)|A^*S^*(t)u_0|_{-1}$ (see the proof of Lemma 3.1 in [12]), we infer that the function $t \to \exp(-\omega t)A^*S^*(t)u_0$ has at most countably many discontinuities. Hence, for each $u_0 \in L^2$, the function $t \to S^*(t)u_0$ is H^{-1} differentiable on $(0,\infty) \setminus N$ and

$$\frac{d}{dt}\, S^*(t)u_0 + A^*S^*(t)u_0 = 0,\ \forall t \in (0,\infty) \setminus N, \tag{2.20}$$

where N is a countable subset of $(0,\infty)$.

Proof of Theorem 2.1 (continued). We note first that

$$S(t)u_0 = S^*(t)u_0,\ \forall t \ge 0,\ u_0 \in L^1 \cap L^2. \tag{2.21}$$

Indeed, by (2.17) it follows that if $u_0 \in L^1 \cap L^2$, then $|v_h^{j+1}|_1 \le |v_k^j|_1$, $\forall j = 0, 1, ...$, and, therefore,

$$|v_h^{j+1}|_1 \le |v_j^0|_1 = |v_0|_1,\ \forall j. \tag{2.22}$$

This follows by multiplying equation (2.17) with $\mathcal{X}_\delta(v_h^{j+1})$ and integrating over $\mathbb{R}^d$, where $\mathcal{X}_\delta$ is defined by

$$\mathcal{X}_\delta(r) = \begin{cases} 1 & \text{for} \quad r \geq \delta, \\ \dfrac{r}{\delta} & \text{for} \quad r \in (-\delta, \delta), \\ -1 & \text{for} \quad r \leq -\delta. \end{cases}$$

Taking into account that $v_h^{j+1} \in H^1$, we have by (2.1) that

$$\begin{aligned}&(A^* v_h^{j+1}, \mathcal{X}_\delta(v_h^{j+1}))_2 \\ &= \int_{\mathbb{R}^d} \beta'(v_h^{j+1})|\nabla v_h^{j+1}|^2 \mathcal{X}'_\delta(v_h^{j+1})dx + \int_{[x;|v_h^{j+1}(x)|\leq\delta]} (v_h^{j+1})(D \cdot \nabla v_h^{j+1})dx,\end{aligned}$$

which yields

$$\limsup_{\delta\to 0} (A^* v_h^{j+1}, \mathcal{X}_\delta(v_h^{j+1}))_2 \geq 0, \ \forall j = 0, 1, \dots .$$

Hence,

$$\limsup_{\delta\to\infty} \int_{\mathbb{R}^d} v_h^{j+1} \mathcal{X}_\delta(v_h^{j+1})dx \leq |v_h^j|_1, \ \forall j = 0, 1, ...,$$

and so (2.22) follows.

Comparing (2.17) with (1.4), we infer that $u_h \equiv v_h$, $\forall h$, and so, by (1.3) and (2.16), we get (2.21), as claimed. In particular, we have that, if $u_0 \in \mathcal{P} \cap L^\infty$, then by (1.8) it also follows that $S^*(t)u_0 \in \mathcal{P} \cap L^\infty$, $\forall t > 0$.

Roughly speaking, this means that the semigroup $S(t)$ is smooth on $L^1 \cap L^2$ in H^{-1}-norm. Then, by (2.3)–(2.4), (2.21) and the corresponding properties of $S(t)$ follow by (2.12), (2.19)–(2.20). As regards (2.5)–(2.6), we note first that by Theorem 4.1 in [8] (see also [10], p. 161), we have, for all $u_0 \in \mathcal{P}$ with $u_0 \log u_0 \in L^1$,

$$E(S(t)u_0) + \int_0^t \Psi(S(\tau)u_0)d\tau \leq E(u_0) < \infty, \ \forall t \geq 0, \tag{2.23}$$

where E is the energy functional (1.13) and

$$\Psi(u) \equiv \int_{\mathbb{R}^d} \left| \frac{\beta'(u)\nabla u}{\sqrt{b^*(u)}} - D\sqrt{b^*(u)} \right|^2 dx. \tag{2.24}$$

Hence, $\Psi(S^*(t)u_0) < \infty$, a.e. $t > 0$, which by (2.21) and Hypotheses (i)–(iii) implies (2.5) (see [8, Lemma 5.1]), as claimed. Moreover, by (2.23), also (2.6) holds.

Assume now that $u_0 \in \mathcal{P}^*$, hence $\frac{\psi}{u_0} \in L^1$ for some $\psi \in \mathcal{X}$, where $\mathcal{X}$ is defined by (1.15). We note that since $S(t)(\mathcal{P}) \subset \mathcal{P}$, $\forall t \geq 0$, we have also that $u(t) \geq 0$, $\forall t \geq 0$, and $u(t) \in L^\infty$, $\forall t \geq 0$. So, it remains to prove that $\frac{\psi}{u(t)} \in L^1$, $\forall t \geq 0$. To this end, we consider the cut-off function

$$\varphi_n(x) = \eta\left(\frac{|x|^2}{n}\right)\psi(x),\ \forall x \in \mathbb{R}^d,\ n \in \mathbb{N},$$

where $\eta \in C^2([0,\infty))$ is such that $0 \leq \eta \leq 1$ and

$$\eta(r) = 1,\ \forall r \in [0,1];\quad \eta(r) = 0,\ \forall r \geq 2. \tag{2.25}$$

Since $u : [0,\infty) \to H^{-1}$ is locally Lipschitz, $[0,\infty) \ni t \to {}_{H^{-1}}(u(t),\varphi)_{H^1}$ is locally Lipschitz for all $\varphi \in H^1$, and so almost everywhere differentiable. We also note the chain differentiation rule

$$\frac{d}{dt}\int_{\mathbb{R}^d} g(u(t,x))\varphi_n(x)dx = {}_{H^{-1}}\left(\frac{du}{dt}(t), \gamma(u(t))\varphi_n\right)_{H^1},\ \text{a.e. } t \in (0,T),$$

for all $T > 0$ and all $u \in L^2(0,T;H^1)$, with $\frac{du}{dt} \in L^2(0,T;H^{-1})$, where $\gamma \in C^1(\mathbb{R})$, $g(r) \equiv \int_0^r \gamma(s)ds$.

In the special case, where $\frac{du}{dt} \in L^2(0,T;L^2)$, this formula follows by [4, Lemma 4.4, p. 158]. If $\frac{du}{dt} \in L^2(0,T;H^{-1})$, this follows by approximating u by $u_\varepsilon = (I - \varepsilon\Delta)^{-1}u$ and letting $\varepsilon \to 0$. We also note that by (2.15) we have that $u = S^*(t)u_0 \in L^2(0,T;H^1)$.

Let $\varepsilon > 0$ be arbitrary, but fixed. Then, since $(u(\cdot)+\varepsilon)^{-1} \in L^2(0,T;H^1)$, we have

$$-\frac{d}{dt}\int_{\mathbb{R}^d} \frac{\varphi_n(x)}{u(t,x)+\varepsilon}\,dx = {}_{H^{-1}}\left(\frac{du}{dt}(t), \frac{\varphi_n}{(u(t)+\varepsilon)^2}\right)_{H^1},\ \text{a.e. } t > 0,$$

and so, by (2.14) we get

$$\begin{aligned}
&\frac{d}{dt}\int_{\mathbb{R}^d} \frac{\varphi_n(x)}{u(t,x)+\varepsilon}\,dx + 2\int_{\mathbb{R}^d} \frac{\beta'(u(t,x))\varphi_n(x)|\nabla u(t,x)|^2}{(u(t,x)+\varepsilon)^3}\,dx \\
&\quad = \int_{\mathbb{R}^d} \frac{\beta'(u(t,x))(\nabla\varphi_n(x)\cdot\nabla u(t,x))}{(u(t,x)+\varepsilon)^2}\,dx \\
&\qquad - \int_{\mathbb{R}^d} \frac{(D(x)\cdot\nabla\varphi_n(x))b(u(t,x))u(t,x)}{(u(t,x)+\varepsilon)^2}\,dx \\
&\qquad + 2\int_{\mathbb{R}^d} \frac{b(u(t,x))u(t,x))(D(x)\cdot\nabla u(t,x))\varphi_n(x)}{(u(t,x)+\varepsilon)^3}\,dx,\ \text{a.e. } t > 0.
\end{aligned} \tag{2.26}$$

By (2.25), we have

$$|\nabla \varphi_n(x)| \le \frac{4\psi(x)}{\sqrt{n}} \, |\eta'|_\infty + \varphi_n(x) g(x), \; x \in \mathbb{R}^d,$$

where $g(x) = \frac{|\nabla \psi(x)|}{\psi(x)}$.

On the other hand, we have by Hypotheses (i)—(iii) that

$$2 \int_{\mathbb{R}^d} \frac{\beta'(u(t,x))\varphi_n(x)|\nabla u(t,x)|^2}{(u(t,x)+\varepsilon)^3} \, dx \ge 2\gamma_1 \int_{\mathbb{R}^d} \frac{\varphi_n(x)|\nabla u(t,x)|^2}{(u(t,x)+\varepsilon)^3}, \qquad (2.27)$$

and

$$\begin{aligned}
&\int_{\mathbb{R}^d} \frac{\beta'(u(t,x))\nabla \varphi_n(x) \cdot \nabla u(t,x)}{(u(t,x)+\varepsilon)^2} \, dx \\
&\quad \le \gamma_2 \int_{\mathbb{R}^d} \frac{|\nabla u(t,x)|}{(u(t,x)+\varepsilon)^2} \left(\frac{4\psi(x)}{\sqrt{n}} \, |\eta'|_\infty + \varphi_n(x) g(x) \right) dx \\
&\quad \le C_1\gamma_2 \int_{\mathbb{R}^d} \frac{|\nabla u(t,x)|\varphi_n(x)}{(u(t,x)+\varepsilon)^2} \, dx + \frac{C_2\gamma_2}{\sqrt{n}} \int_{\mathbb{R}^d} \frac{|\nabla u(t,x)|\psi(x)}{(u(t,x)+\varepsilon)^2} \, dx \qquad (2.28)\\
&\quad \le \frac{\gamma_1}{2} \int_{\mathbb{R}^d} \frac{\varphi_n(x)|\nabla u(t,x)|^2}{(u(t,x)+\varepsilon)^3} \, dx + \frac{C_2\gamma_2}{\sqrt{n}} \int_{\mathbb{R}^d} \frac{|\nabla u(t,x)|\psi(x)}{(u(t,x)+\varepsilon)^2} \, dx \\
&\quad + C_3 \int_{\mathbb{R}^d} \frac{\varphi_n(x)}{u(t,x)+\varepsilon} \, dx.
\end{aligned}$$

$$\begin{aligned}
&\int_{\mathbb{R}^d} \frac{D(x)\cdot\nabla \varphi_n(x) b(u(t,x))u(t,x)}{(u(t,x)+\varepsilon)^2} \, dx \le C_4 \int_{\mathbb{R}^d} \frac{|\nabla \varphi_n(x)|}{u(t,x)+\varepsilon} \, dx \\
&\quad \le C_5 \int_{\mathbb{R}^d} \left(\frac{\varphi_n(x)}{u(t,x)+\varepsilon} + \frac{1}{\sqrt{n}(u(t,x)+\varepsilon)} \right) dx.
\end{aligned} \qquad (2.29)$$

$$\begin{aligned}
&2 \int_{\mathbb{R}^d} \frac{b(u(t,x))u(t,x)(D(x)\cdot\nabla u(t,x))\varphi_n(x)}{(u(t,x)+\varepsilon)^3} \, dx \\
&\quad \le C_6\gamma_3 \int_{\mathbb{R}^d} \frac{|\nabla u(t,x)|\varphi_n(x)}{(u(t,x)+\varepsilon)^2} \, dx \qquad (2.30)\\
&\quad \le \frac{\gamma_1}{2} \int_{\mathbb{R}^d} \frac{|\nabla u(t,x)|^2\varphi_n(x)}{(u(t,x)+\varepsilon)^3} \, dx + C_7 \int_{\mathbb{R}^d} \frac{\varphi_n(x)}{u(t,x)+\varepsilon} \, dx.
\end{aligned}$$

Then, by (2.27)–(2.30) and by Hypotheses (i)–(iii) it follows that

$$\begin{aligned}\frac{d}{dt}\int_{\mathbb{R}^d}\frac{\varphi_n(x)}{u(t,x)+\varepsilon}\,dx &+\gamma_1\int_{\mathbb{R}^d}\frac{\varphi_n(x)|\nabla u(t,x)|^2}{(u(t,x)+\varepsilon)^3}\,dx\\ &\le C_8\int_{\mathbb{R}^d}\frac{\varphi_n(x)}{u(t,x)+\varepsilon}\,dx+\frac{C_8}{\sqrt{n}}\int_{\mathbb{R}^d}\frac{1}{u(t,x)+\varepsilon}\,dx\\ &\quad+\tfrac{C_8}{\sqrt{n}}\int_{\mathbb{R}^d}\frac{|\nabla u(t,x)|\psi(x)}{(u(t,x)+\varepsilon)^2}\,dx,\ \text{a.e. } t>0.\end{aligned}$$

This yields

$$\begin{aligned}\int_{\mathbb{R}^d}\frac{\varphi_n(x)}{u(t,x)+\varepsilon}\,dx &+\gamma_1\int_0^t\int_{\mathbb{R}^d}\frac{\varphi_n(x)|\nabla u(s,x)|^2}{(u(s,x)+\varepsilon)^3}\,dxds\\ &\le\int_{\mathbb{R}^d}\frac{\varphi_n(x)}{u_0(x)+\varepsilon}\,dx+C_9\int_0^t ds\int_{\mathbb{R}^d}\frac{\varphi_n(x)}{u(s,x)+\varepsilon}\,dx\\ &\quad+\frac{C_9}{\sqrt{n}}\int_0^t ds\int_{\mathbb{R}^d}\frac{1}{u(s,x)+\varepsilon}\,dx\\ &\quad+\frac{C_9}{\sqrt{n}}\int_0^t ds\int_{\mathbb{R}^d}\frac{|\nabla u(s,x)|\psi(x)}{(u(s,x)+\varepsilon)^2}\,dxds,\ \forall t\ge 0,\end{aligned}$$

while

$$\begin{aligned}\int_0^T\int_{\mathbb{R}^d}\frac{|\nabla u(s,x)|\psi(x)}{(u(s,x)+\varepsilon)^2}\,dxds &\le\frac{1}{\varepsilon^2}\left(\int_0^T ds\int_{\mathbb{R}^d}|\nabla u(s,x)|^2dx\right)^{\frac12}\left(T\int_{\mathbb{R}^d}\psi^2(x)dx\right)^{\frac12}\\ &\le\frac{C_{10}}{\varepsilon^2},\end{aligned}$$

because by (2.15) we know that $\nabla u\in L^2(0,T;L^2)$.

Letting $n\to\infty$, we get

$$\int_{\mathbb{R}^d}\frac{\psi(x)dx}{u(t,x)+\varepsilon}\le\int_{\mathbb{R}^d}\frac{\psi(x)dx}{u_0(x)+\varepsilon}+C_T\int_0^t ds\int_{\mathbb{R}^d}\frac{\psi(x)dx}{u(s,x)+\varepsilon},\ \forall t\in(0,T),$$

where $C_T>0$ is independent of ε, and so, for $\varepsilon\to 0$ it follows by Gronwall's lemma (which is applicable since $\psi\in L^1$) and by Fatou's lemma,

$$\int_{\mathbb{R}^d}\frac{\psi(x)dx}{u(t,x)}\le\exp(C_Tt)\int_{\mathbb{R}^d}\frac{\psi(x)dx}{u_0(x)}<\infty,\ \forall t\in(0,T),$$

as claimed. □

3 A new tangent space to $\mathcal{P}$

To represent NFPE (1.1) as a gradient flow as in [16], [17], we shall interpret the space $\mathcal{P}^*$ as a Riemannian manifold endowed with an appropriate tangent bundle with scalar product which is, however, different from the one in [16], [17]. To this purpose, we define the tangent space $\mathcal{T}_u(\mathcal{P}^*)$ at $u \in \mathcal{P}^* \subset \mathcal{P}$ as follows,

$$\mathcal{T}_u(\mathcal{P}^*) = \{z = -\mathrm{div}(b^*(u)\nabla y); y \in W^{1,1}_{\mathrm{loc}}(\mathbb{R}^d), \sqrt{u}\,\nabla y \in L^2\}\,(\subset H^{-1}). \quad (3.1)$$

(Here, $\mathcal{P}^*$ is defined by (1.14).)

The differential structure of the manifold $\mathcal{P}^*$ is defined by providing for $u \in \mathcal{P}^*$ the linear space $\mathcal{T}_u(\mathcal{P}^*)$ with the scalar product (metric tensor)

$$\begin{aligned} \langle z_1, z_2 \rangle_u &= \int_{\mathbb{R}^d} b^*(u)\nabla y_1 \cdot \nabla y_2\, dx, \\ z_i &= \mathrm{div}(b^*(u)\nabla y_i),\ i = 1, 2, \end{aligned} \quad (3.2)$$

and with the corresponding Hilbertian norm $\|\cdot\|_u$,

$$\|z\|_u^2 = \int_{\mathbb{R}^d} b^*(u)|\nabla y|^2 dx,\ z = -\mathrm{div}(b^*(u)\nabla y). \quad (3.3)$$

As a matter of fact, $\mathcal{T}_u(\mathcal{P}^*)$ is viewed here as a factor space by identifying in (3.2) two functions $y_1, y_2 \in W^{1,1}_{\mathrm{loc}}$ if $\mathrm{div}(b^*(u)\nabla(y_1 - y_2)) \equiv 0$. Note also that, since $b^*(u) \geq b_0 u$ and $u > 0$, a.e. on $\mathbb{R}^d$, $\|z\|_u = 0$ implies that $z \equiv 0$. Moreover, we have

$$\|z_1\|_u = \|z_2\|_u \text{ for } z_i = \mathrm{div}(b^*(u)\nabla y_i),\ i = 1, 2, \quad (3.4)$$

if $\mathrm{div}(b^*(u)\nabla(y_1 - y_2)) \equiv 0$ in H^{-1}. Indeed, for each $\varphi \in C_0^\infty(\mathbb{R}^d)$, we have in this case that

$$\int_{\mathbb{R}^d} b^*(u)\nabla(y_1 - y_2) \cdot \nabla(\varphi y_i) dx = 0,\ i = 1, 2,$$

and this yields

$$\int_{\mathbb{R}^d} b^*(u)\nabla(y_1 - y_2) \cdot (\varphi \nabla y_i + y_i \nabla \varphi) dx = 0,\ i = 1, 2. \quad (3.5)$$

If we take $\varphi(x) = \eta\left(\frac{|x|^2}{n}\right)$, where $\eta \in C^2([0, \infty))$; $\eta(r) = 1$ for $0 \leq r \leq 1$, $\eta(r) = 0$ for $r \geq 2$, and let $n \to \infty$ in (3.5), we get via the Lebesgue dominated convergence theorem that

$$\int_{\mathbb{R}^d} b^*(u)\nabla(y_1 - y_2)\cdot\nabla y_i\,dx = 0,\ i = 1,2,$$

which, as easily seen, implies (3.4), as claimed. Hence, the norm $\|z\|_u$ is independent of representation (3.2) for z. We should also note that $\mathcal{T}_u(\mathcal{P}^*)$ so defined is a Hilbert space, in particular, it is complete in the norm $\|\cdot\|_u$. Here is the argument.

Let $u \in \mathcal{P}^*$ and let $\{y_n\} \subset W^{1,1}_{\rm loc}$ be such that

$$\|z_n - z_m\|^2_u = \int_{\mathbb{R}^d} b^*(u)|\nabla(y_n - y_m)|^2 dx \to 0 \text{ as } n,m \to \infty.$$

This implies that the sequence $\{\sqrt{b^*(u)}\,\nabla y_n\}$ is convergent in L^2 as $n \to \infty$ and by Hypothesis (ii) so is $\{\sqrt{u}\,\nabla y_n\}$. Let

$$f = \lim_{n\to\infty} \sqrt{b^*(u)}\,\nabla y_n \text{ in } L^2. \tag{3.6}$$

As $\frac{\psi}{u} \in L^1$ for some $\psi \in \mathcal{X}$, we infer that $\{\nabla y_n\}$ is convergent in $L^1_{\rm loc}$ and so, by the Sobolev embedding theorem (see, e.g., [11], p. 278), the sequence $\{y_n\}$ is convergent in $L^{\frac{d}{d-1}}_{\rm loc}$ and, therefore, in $L^1_{\rm loc}$ too. Hence, as $n \to \infty$, we have

$$\begin{aligned} y_n &\longrightarrow y && \text{in } L^1_{\rm loc} \cap L^{\frac{d}{d-1}}_{\rm loc}, \\ \nabla y_n &\longrightarrow \nabla y && \text{in } (L^1_{\rm loc})^d. \end{aligned}$$

and hence, along a subsequence, a.e. on $\mathbb{R}^d$. So, by (3.6) we infer that $f = \sqrt{b^*(u)}\,\nabla y$, where $y \in W^{1,1}_{\rm loc}$. Hence, as $n \to \infty$, we have

$$\|z_n - z\|_u \to 0 \text{ for } z = -{\rm div}(b^*(u)\nabla y),\ y \in W^{1,1}_{\rm loc},$$

as claimed.

As a consequence, we have that

$$\{z = -{\rm div}(b^*(u)\nabla y); y \in C^\infty_0(\mathbb{R}^d)\} \text{ is dense in } \mathcal{T}_u(\mathcal{P}^*) \text{ for all } u \in \mathcal{P}^*. \tag{3.7}$$

To conclude, we have shown that, *for each* $u \in \mathcal{P}^*$, $\mathcal{T}_u(\mathcal{P}^*)$ *is a Hilbert space with the scalar product* (3.2) and, as mentioned earlier, this is just the *tangent space* to $\mathcal{P}^*$ at u.

4 The Fokker–Planck gradient flow on $\mathcal{P}^*$

We are going to define here the gradient of the energy function $E : L^2 \to]-\infty, +\infty]$ defined by (1.13). Namely,

$$E(u) = \begin{cases} \displaystyle\int_{\mathbb{R}^d} (\eta(u) + \Phi u) dx & \text{if } u \in \mathcal{P} \cap L^\infty(\mathbb{R}^d) \cap L^1(\mathbb{R}^d; \Phi dx) \\ +\infty & \text{otherwise.} \end{cases}$$

We note that E is convex, nonidentically $+\infty$ and we have:

Lemma 4.1. *E is lower-semicontinuous on L^2.*

Proof. We first note that if $u \in \mathcal{P} \cap L^\infty \cap L^1(\mathbb{R}^d; \Phi dx)$, then by the proof of (4.6) in [8] for all $\alpha \in [m/(m+1), 1)$, we have by Hypothesis (iv)

$$\int_{\mathbb{R}^d} \eta(u) dx \geq -C_\alpha \left(\int_{\mathbb{R}^d} \Phi u \, dx + 1 \right)^\alpha,$$

hence, since $r^\alpha \leq \frac{1}{2C_\alpha} r + C'_\alpha,\ r \geq 0$,

$$E(u) \geq \frac{1}{2} \int_{\mathbb{R}^d} \Phi \, u \, dx - C''_\alpha \tag{4.1}$$

for some $C_\alpha, C'_\alpha, C''_\alpha \in (0, \infty)$ independent of u.

Let now $u, u_n \in L^2$, $n \in \mathbb{N}$, such that $\lim\limits_{n\to\infty} u_n = u$ in L^2. We may assume that

$$\liminf_{n\to\infty} E(u_n) = \lim_{n\to\infty} E(u_n) < \infty$$

and that $E(u_n) < \infty$ for all $n \in \mathbb{N}$. Then, by (4.1)

$$\sup_{n\in\mathbb{N}} \int_{\mathbb{R}^d} \Phi \, u_n \, dx < \infty. \tag{4.2}$$

Now, suppose that

$$E(u) > \lim_{n\to\infty} E(u_n). \tag{4.3}$$

Then

$$\begin{aligned} E(u) &> \liminf_{n\to\infty} \int_{\mathbb{R}^d} \eta(u_n) dx + \liminf_{n\to\infty} \int_{\mathbb{R}^d} \Phi \, u_n \, dx \\ &\geq \liminf_{n\to\infty} \int_{\mathbb{R}^d} \eta(u_n) dx + \int_{\mathbb{R}^d} \Phi \, u \, dx, \end{aligned}$$

where we applied Fatou's lemma to the second summand in the last inequality. If we can also apply it to the first, then we get a contradiction to (4.3) and the lemma is proved. To justify the application of Fatou's lemma to the first summand, it is enough to prove that there exist $f_n \in L^1$, $n \in \mathbb{N}$, $f_n \geq 0$, such that (along a subsequence)

$$f_n \to f \text{ in } L^1, \tag{4.4}$$

and

$$\eta(u_n) \geq -f_n, n \in \mathbb{N}. \tag{4.5}$$

To find such f_n, $n \in \mathbb{N}$, we use (4.2). Recall from (4.4) in [8] that, for some $c \in (0, \infty)$,

$$\eta(r) \geq -cr \log^-(r) - cr,\ r \geq 0.$$

Hence,

$$\eta(u_n) \geq -cu_n \log^-(u_n) - cu_n,\ n \in \mathbb{N}.$$

Since $u_n \to u$ in L^2 and thus in L^1_{loc}, it follows by(4.2) and our assumptions on Φ that (again along a subsequence) $u_n \to u$ in L^1. Furthermore, for all $\alpha \in (0,1)$,

$$\begin{aligned} -f \log^-(r) &= 1_{[0,1]}(r) r \log r = 1_{[0,1]}(r) r\, \tfrac{1}{1-\alpha}\, r^\alpha \underbrace{r^{1-\alpha} \log r^{1-\alpha}}_{\geq -e^{-1}} \\ &\geq -\frac{1}{(1-\alpha)e}\, r^\alpha,\ r \geq 0. \end{aligned}$$

Hence, we find that

$$\eta(u_n) \geq -\frac{c}{(1-\alpha)e}\, u_n^\alpha - cu_n,\ n \in \mathbb{N}.$$

But, since $u_n \to u$ in L^2 and thus $u_n^\alpha \to u^\alpha$ in L^1_{loc}, by Hypothesis (iv) it remains to show that, for some $\varepsilon, \alpha \in (0,1)$,

$$\sup_{n \in \mathbb{N}} \int_{\mathbb{R}^d} u_n^\alpha\, \Phi^\varepsilon\, dx < \infty, \tag{4.6}$$

to conclude that (along a subsequence) $u_n^\alpha \to u^\alpha$ in L^1, and then (again selecting a subsequence of $\{u_n\}$ if necessary) (4.4) and (4.5) hold with

$$f_n := \frac{c}{(1-\alpha)e}\, u_n^\alpha + cu_n,\ n \in \mathbb{N}.$$

So, let us prove (4.6).

Applying Hölder's inequality with $p := \frac{1}{\alpha}$, we find that

$$\int_{\mathbb{R}^d} u_n^\alpha \, \Phi^\varepsilon \, dx \le \left(\int_{\mathbb{R}^d} u_n \Phi \, dx \right)^\alpha \left(\int_{\mathbb{R}^d} \Phi^{-\left(\frac{1}{\alpha}-\varepsilon\right)/(1-\alpha)} \, dx \right)^{1-\alpha}.$$

Hence, choosing ε small enough and α close enough to 1, so that $\left(\frac{1}{\alpha} - \varepsilon\right)/(1-\alpha) \ge m$, Hypothesis (iv) implies (4.6). □

By Lemma 4.1 we have for E that its directional derivative

$$E'(u,z) = \lim_{\lambda \to 0} \frac{1}{\lambda}(E(u+\lambda z) - E(u))$$

exists for all $u \in \mathcal{P}^*$ and $z \in L^2$ (it is unambiguously either a real number or $+\infty$) (see, e.g., [6], p. 86).

In the following, we shall take $u \in \mathcal{P}^* \subset D(E) = \{u \in L^2; E(u) < \infty\}$ and $z \in \mathcal{T}_u(\mathcal{P})$ and obtain that

$$\begin{aligned} E'(u,z) &= \lim_{\lambda \downarrow 0} \frac{1}{\lambda}(E(u+\lambda z) - E(u)) \\ &= \int_{\mathbb{R}^d} z(x) \left(\int_1^{u(x)} \frac{\beta'(\tau)}{b^*(\tau)} \, d\tau + \Phi(x) \right) dx. \end{aligned} \tag{4.7}$$

Moreover, the subdifferential $\partial E_u : L^2 \to L^2$ of E at u is expressed as (see [6, Proposition 2.39])

$$\partial E_u = \{y \in L^2;\ (z,y)_2 \le E'(u,z);\ \forall z \in L^2\}. \tag{4.8}$$

We recall that if E is Gâteaux differentiable at u, then ∂E_u reduces to the gradient ∇E_u of E at u and

$$E'(u,z) = (\nabla E_u, z)_2,\ \forall z \in L^2.$$

Any element $y \in \partial E_u$ is called a *subgradient* of E at u. In the following, we shall denote, for simplicity, again by ∇E_u any subgradient of E at u and we shall keep the notation diff $E_u \cdot z$ for $E'(u,z)$.

If $z \in \mathcal{T}_u(\mathcal{P}^*)$ is of the form $z = z_2 = -\mathrm{div}(b^*(u)\nabla y_2)$, where $y_2 \in C_0^\infty(\mathbb{R}^d)$, then $z = -b^*(u)\Delta y_2 - \nabla y_2 \cdot (b'(u)u + b(u))\nabla u$ and so, by (i) and (1.14), it follows that $z \in L^2$ and hence

$$
\begin{aligned}
E'(u,z) &= \operatorname{diff} E_u \cdot z = \lim_{\lambda \downarrow 0} \frac{1}{\lambda}(E(u+\lambda z_2) - E(u)) \\
&= \int_{\mathbb{R}^d} \left(\frac{\nabla \beta(u(x))}{b^*(u(x))} - D(x) \right) b^*(u(x)) \cdot \nabla y_2(x)\, dx \qquad (4.9)\\
&= \int_{\mathbb{R}^d} b^*(u(x)) \nabla y_2(x) \cdot \nabla \left(\int_0^{u(x)} \frac{\beta'(s)}{b^*(s)}\, ds + \Phi(x) \right) dx.
\end{aligned}
$$

We claim that

$$
x \mapsto \int_0^{u(x)} \frac{\beta'(s)}{b^*(s)}\, ds + \Phi(x) \ \text{ is in } W^{1,1}_{\mathrm{loc}}. \tag{4.10}
$$

To prove this, we first note that by Hypotheses (i) and (ii)

$$
\frac{\gamma_1}{|b|_\infty} \frac{1}{s} \le \frac{\beta'(s)}{b^*(s)} \le \frac{\gamma_2}{b_0} \frac{1}{s}, \ \ s > 0.
$$

Hence,

$$
\frac{\gamma_1}{|b|_\infty} \log u \le \int_0^u \frac{\beta'(s)}{b^*(s)}\, ds \le \frac{\gamma_2}{b_0} \log u. \tag{4.11}
$$

Now, let $\psi \in \mathcal{X}$ such that $\frac{\psi}{u} \in L^1$. then, for every compact $K \subset \mathbb{R}^d$ and $K_n := \left\{ \frac{1}{n} \le u \le 1 \right\}$, $n \in \mathbb{N}$,

$$
\begin{aligned}
\int_{K_n} (\log u)^- dx &\le \left(\int_{K_n} (\log u)^2 u\, dx \right)^{\frac{1}{2}} (\inf\nolimits_K \psi)^{-\frac{1}{2}} \left(\textstyle\int_K \frac{\psi}{u}\, dx \right)^{\frac{1}{2}} \\
&\le \sup_K ((\log u)^- u) \left(\int_{K_n} (\log u)^- dx \right)^{\frac{1}{2}} \left(\inf_k \psi \right)^{-\frac{1}{2}} \left(\int_K \frac{\psi}{u}\, dx \right)^{\frac{1}{2}}.
\end{aligned}
$$

Dividing by $\left(\int_{K_n} (\log u)^- dx \right)^{\frac{1}{2}}$ and letting $n \to \infty$ yields $\log u \in L^1_{\mathrm{loc}}$, since trivially $(\log u)^+ \in L^1_{\mathrm{loc}}$, since $u \in L^\infty$. Furthermore, for $\varepsilon > 0$,

$$
\int_K |\nabla \log(u+\varepsilon)| dx \int_K \frac{|\nabla u|}{u+\varepsilon}\, dx \le \left(\int_K \frac{|\nabla u|^2}{u}\, dx \right)^{\frac{1}{2}} \left(\inf_K \psi \right)^{-\frac{1}{2}} \left(\int_K \frac{\psi}{u}\, dx \right)^{\frac{1}{2}}.
$$

Letting $\varepsilon \to 0$ yields $|\nabla \log u| \in L^1_{\mathrm{loc}}$, and (4.10) is proved by Hypothesis (iv). Hence,

$$
E'(u, z_2) = \int_{\mathbb{R}^d} b^*(u(x)) \nabla y_2(x) \cdot \nabla y_1(x) dx = - \langle z_1, z_2 \rangle_u,
$$

where $z_1 = -\mathrm{div}(b^*(u)\nabla y_1)$, $y_1 = \int_0^u \frac{\beta'(s)}{b^*(s)}\,ds + \Phi$. Therefore,

$$z_2 \mapsto E'(u, z_2) = \mathrm{diff}\, E_u z_2 = \langle \nabla E_u, z_2 \rangle$$

extends to all $z \in \mathcal{T}_u(\mathcal{P}^*)$ by continuity and by (3.2) it follows that for $u \in \mathcal{P}^*$ any subgradient $\nabla_u E$ of E is given by

$$\begin{aligned} \nabla E_u &= -\mathrm{div}\left(b^*(u)\nabla \left(\int_0^u \frac{\beta'(s)}{b^*(s)}\,ds + \Phi \right)\right) \\ &= -\Delta\beta(u) + \mathrm{div}(Db^*(u)) \in H^{-1}. \end{aligned} \tag{4.12}$$

In particular, this means that ∂E_u is single valued and $\partial E_u = \nabla E_u$.

On the other hand, by Theorem 2.1 we know that, for $u_0 \in \mathcal{P}^*$ with $u_0 \log u_0 \in L^1$, we have for the flow $u(t) \equiv S(t)u_0$,

$$S(t)u_0 \in H^1 \cap \mathcal{P},\ \forall t > 0,\ \frac{\nabla(S(t)u_0)}{\sqrt{S(t)u_0}} \in L^2, \text{ a.e. } t > 0,$$

$$\frac{d^+}{dt}\, S(t)u_0 = \Delta\beta(S(t)u_0) - \mathrm{div}(Db^*(S(t)u_0)),\ \forall t > 0, \tag{4.13}$$

$$\frac{d}{dt}\, S(t)u_0 = \Delta\beta(S(t)u_0) - \mathrm{div}(Db^*(S(t)u_0)),\ \forall t \in (0,\infty) \setminus N, \tag{4.14}$$

where N is at most countable set of $(0,\infty)$. Moreover, if $u_0 \in \mathcal{P}^*$, then, as seen in Theorem 2.1, it follows that $S(t)u_0 \in \mathcal{P}^*$, $\forall t > 0$, and $\nabla E_{u(t)}$ is well defined, a.e. $t > 0$. Taking into account (4.12), we may rewrite (4.13)-(4.14) as the gradient flow on $\mathcal{P}^*$ endowed with the metric tensor (3.2). Namely, we have

Theorem 4.2. *Under Hypotheses* (i)–(iv)*, for each $u_0 \in \mathcal{P}^*$, the function $u(t) = S(t)u_0 \in \mathcal{P}^*$, $\forall t > 0$, and it is the solution to the gradient flow*

$$\frac{d}{dt}\, u(t) = -\nabla E_{u(t)},\ \textit{a.e. } t > 0, \tag{4.15}$$

$$\frac{d^+}{dt}\, u(t) = -\nabla E_{u(t)},\ \forall t > 0, \tag{4.16}$$

$$\frac{d}{dt}\, u(t) = -\nabla E_{u(t)},\ \forall t \in (0,\infty) \setminus N, \tag{4.17}$$

where N is at most countable set of $0,\infty)$.

By (3.2) we may rewrite (4.16) as

$$\frac{d^+}{dt}\,E(S(t)u_0) = -\left\|\frac{d^+}{dt}\,S(t)u_0\right\|^2_{u(t)},\ \forall t>0. \tag{4.18}$$

Equivalently,

$$\frac{d^+}{dt}\,E(S(t)u_0) + A(S(t)u_0) = 0,\ \forall t>0, \tag{4.19}$$

where A^* is the generator (2.1) of the Fokker–Planck semigroup $S^*(t)$ (equivalently, $S(t)$) in H^{-1}. Similarly, by (3.3) and (2.23)–(2.24) we can write

$$\frac{d}{dt}\,E(S(t)u_0) = -\left\|\frac{d}{dt}\,S(t)u_0\right\|^2_{u(t)} = \Psi(S(t)u_0),\ \forall t\in(0,\infty)\setminus N. \tag{4.20}$$

As a matter of fact, the energy dissipation formula (4.20) was used in [8] (see also [5], Chapter 4) to prove that $S(t)u_0 \to u_\infty$ strongly in L^1 as $t\to\infty$, where u_∞ is the unique solution to equilibrium equation $-\Delta\beta(u_\infty)+\mathrm{div}(Db(u_\infty)u_\infty)=0$.

Remark 4.3. Taking into account (4.7), we see also that the operator A^* defined by (2.1) can be expressed as

$$A^*u = B_u\,\mathrm{diff}\,E_u,\ \forall u\in D(A^*)=H^1, \tag{4.21}$$

where $B_u : H^1 \to H^{-1}$ is the linear symmetric operator defined by

$$\begin{aligned} B_u(y) &= -\mathrm{div}(b^*(u)\nabla y),\ \forall y\in D(B_u),\\ D(B_u) &= \{y\in l^2,\ \sqrt{u}\ \nabla y\in L^2\}. \end{aligned} \tag{4.22}$$

This means that ∇E_u can be equivalently written as

$$\nabla E_u = B_u(\mathrm{diff}\ E_u). \tag{4.23}$$

In the special case $b(r)\equiv 1$,

$$E_u \equiv \int_1^u \frac{\beta'(\tau)}{\tau}\,d\tau + \Phi$$

and so $u(t)=S(t)u_0$ is the Wasserstein gradient flow of the functional E defined by the time-discretized scheme

$$\begin{aligned} u_h(t) &= u_h^j,\ t\in[jh,(j+1)h),\ j=0,1,...,\\ u_h^j &= \min_u\left\{\frac{1}{2h}\ d_2(u,u_h^{j-1}) + E(u)\right\} \end{aligned}$$

where d_2 is the Wasserstein distance of order two (see [3], [14], [16]). However, in the general case considered here, this is not the case.

Acknowledgement. This work was funded by the DFG (German Research Foundation) Project ID 317210226-SFB 1283.

References

[1] Albeverio, S., Kondratiev, Y., Röckner, M., Differential geometry of Poisson spaces, *C.R.Acad. Sci. Paris*, Ser. I. Math. 323 (1996), 1129-1134.

[2] Albeverio, S., Kondratiev, Y., Röckner, M., Analysis and geometry on configuration spaces, *J. Funct. Anal.*, 154 (1998), 444-500.

[3] Ambrosio, L., Gigli, N., Savaré, G., *Gradient Flows: in Metric Spaces and in the Space of Probability Measures*, Lectures in Mathematics, Birkhäuser, Zurich, 2005.

[4] Barbu, V., *Nonlinear Differential Equations of Monotone Type in Banach Spaces*, Springer, Berlin. Heidelberg. New York, 2010.

[5] Barbu, V., *Semigroup Approach to Nonlinear Diffusion Equations*, World Scientific, New Jersey, London, Singapore, Shanghai, Hong Kong, 2021.

[6] Barbu, V.,Precupanu, T., *Convexity and Optimization in Banach Spaces*, Springer, Berlin. Heidelberg. New York, 2012.

[7] Barbu, V., Röckner, M., From Fokker–Planck equations to solutions of distribution dependent SDE, *Annals of Probability*, 48 (2020), 1902-1920.

[8] Barbu, V., Röckner, M., The evolution to equilibrium of solutions to nonlinear Fokker-Planck equations, *Indiana Univ. Math. J.*, 72 (1) (2023), 89-131.

[9] Barbu, V., Röckner, M., Uniqueness for nonlinear Fokker-Planck equations and for McKean-Vlasov SDEs: The degenerate case, J. Funct. Anal, 285 (3) (2023).

[10] Barbu, V., Röckner, M., *Nonlinear Fokker–Planck Flows and their Probabilistic Counterparts*, Lecture Notes in Mathematics, vol. 2353, Springer, 2024.

[11] Brezis, H., *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York, Dordrecht. Heidelberg. London, 2010.

[12] Crandall, M.G., Pazy, A., Semigroups of nonlinear contractions and dissipative sets, *J. Funct. Analysis*, 3 (1969), 376-418.

[13] Franck, T.D., *Nonlinear Fokker–Planck Equations*, Springer, Berlin. Heidelberg. New York, 2005.

[14] Jordan, R., Kinderlehrer, D., Otto, F., The variational formulation of the Fokker–Planck equation, *SIAM J. Math. Anal.*, 29 (1) (1998), 1-17.

[15] Kondratyev, S., Vorotnikov, D., Nonlinear Fokker–Planck equations with reaction as gradient flows of the free energy, *J. Funct. Anal.*, 278 (2020), 108310.

[16] Otto, F., The geometry of dissipative evolution equations: the porous medium equation, *Comm. Partial Differential Equations*, 26 (1-2) (2001), 101-174.

[17] Rehmeier, M., Röckner, Nonlinear Fokker–Planck–Kolmogorov equations as gradient flows on the space of probability measures, arXiv:2306.09530 [math.AP]

[18] Ren, P., Röckner, Wang, F.Y., Linearization of nonlinear Fokker–Planck equations and applications, *J. Diff. Equations*, 322 (2022), 1-37.

[19] Röckner, M., *Dirichlet Forms on Infinite Dimensional Manifold-like State Spaces. A survey of Recent Results and Some Prospective for the Future*, *Probability Towards 2000*, 287-306, Springer. New York, 1998.