Rend. Lincei Mat. Appl. 36 (2025), 199–223
DOI 10.4171/RLM/1069

Partial Differential Equations. – *Nonlinear Fokker–Planck equations as smooth Hilbertian gradient flows*, by Viorel Barbu and Michael Röckner, accepted on 19 September 2025.

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 $u_t - \Delta\beta(u) + \operatorname{div}(Db(u)u) = 0$, 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 with 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 $\langle \nabla E_u, z_u \rangle_u = \operatorname{diff} E_u \cdot z_u$ for all vector fields z_u on $\mathcal{P}^*$, where $\langle \cdot,\cdot \rangle_u$ is a scalar product on a suitable tangent space $\mathcal{T}_u(\mathcal{P}^*) \subset \mathcal{D}'(\mathbb{R}^d)$.

Keywords. – Fokker–Planck equation, gradient flow, semigroup, stochastic equations, tensor metric.

Mathematics Subject Classification 2020. – 60H15 (primary); 35A15 (secondary).

1. Introduction

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

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

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

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

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

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

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

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 $u(t) \in \mathcal{P}$, $\forall t \geq 0$, then there is a probabilistically weak solution X_t to the McKean–Vlasov stochastic differential equation

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

on a probability space $(\Omega, \mathcal{F}, \mathbb{P}, W_t)$ with normal filtration $(\mathcal{F}_t)_{t\geq 0}$. 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

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

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

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

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

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

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])

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

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

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

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,

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

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

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

$$S(t+s)=S(t)S(s) \quad \text{for all } s,t\ge 0,$$

$$\big\|S(t)u_1-S(t)u_2\big\|_{L^1(\mathbb{R}^d)}\le \|u_1-u_2\|_{L^1(\mathbb{R}^d)}, \quad \forall t\ge 0,\ u_1,u_2\in L^1(\mathbb{R}^d),$$

$$\lim_{t\to 0}S(t)u_0=u_0 \quad \text{in } L^1(\mathbb{R}^d).$$

Note also (see [7–10]) that

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

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

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

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

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

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,

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

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

$$u\in L^1\big((0,\infty)\times\mathbb{R}^d\big)\cap L^\infty\big((0,\infty)\times\mathbb{R}^d\big)$$

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)

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

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,

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

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

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

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

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

The class $\mathcal{X}$ is clearly nonempty and, in particular, it contains all functions ψ of the form $\psi(x) = (\alpha_1|x|^m + \alpha_2)^{-1}, \alpha_1, \alpha_2$ 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 $\psi \in \mathcal{X}$, $\psi > 0$, $u := \psi^2(\int_{\mathbb{R}^d} \psi^2 dx)^{-1}$ is easily checked to be in $\mathcal{P}^*$. We also note that $\mathcal{P}^*$ is convex.

We recall (see, e.g., [20]) that if $F : \mathcal{X} \to]-\infty, +\infty]$ is a lower semicontinuous function on the metric space $(\mathcal{X}, \rho)$, a gradient flow associated with F is the function $u : (0, \infty) \to \mathcal{X}$ given by $u(t) = \lim_{h \to 0} u_h(t)$ uniformly on compact intervals where

$$
\begin{aligned}
&u_h(t) = u_h^j, \quad \forall t \in \big[(j-1)h, jh\big),\ j = 1, 2, \ldots],\\
&y_h^{j+1} \in \arg\min \left\{ F(y) + \frac{1}{2h}\rho^2(y, u_h^j);\ u \in \mathcal{X} \right\}.
\end{aligned}
\tag{1.16}
$$

Formally, the flow $t \to u(t)$ is a generalized solution (in the sense of a finite difference scheme approximation) to the evolution equation

$$
\frac{du(t)}{dt} = -\nabla F_{u(t)}, \quad t > 0,
\tag{1.17}
$$

where ∇F_u is the gradient of F at u in some generalized sense.

One spectacular application of this idea was developed in the work of Jordan, Kinderlehrer, and Otto [14] where the linear Fokker–Planck equation with $\beta(u) \equiv u$, $b(u) \equiv 1$ is represented in the variational form (1.17), where $\mathcal{X}$ is the space $\mathcal{P}$ endowed with the Wasserstein metric and F is the function (1.13). Later on, Otto [16] extended this construction to porous media equations. The general theory of gradient flows in the Wasserstein spaces was developed in the book of Ambrosio, Gigli, and Savaré [7]. More precisely, the gradient flow representation means that, for $u(t) = S(t)u_0$, $u_0 \in \mathcal{P}^*$, we have

$$\frac{d}{dt} u(t) = -\nabla E_{u(t)}, \quad t > 0, \tag{1.18}$$

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] on the manifold $\mathcal{P}$ endowed with the topology of weak convergence of probability measures and tangent bundle $L^2(\mathbb{R}^d; \mathbb{R}^d; \mu)_{\mu \in \mathcal{P}}$ (see also [1, 2, 15, 18, 19]) and in the fundamental work [16] for the classical porous media equation. Here, to represent the Fokker–Planck equation in the variational form (1.17), we shall proceed in a different way. Namely, we organize the space $\mathcal{P}^*$ as a Riemannian manifold by endowing its tangent space $\mathcal{T}_u(\mathcal{P}^*)$ at $u \in \mathcal{P}^*$ with a convenient Hilbertian structure with scalar product (metric tensor $\langle \cdot, \cdot \rangle_u$). Then one defines the gradient $\nabla E_u \in \mathcal{F}_u(\mathcal{P}^*)$ by the formula

$$\langle \nabla E_u, z \rangle_u = E'(u, z), \quad \forall z \in \mathcal{T}_u(\mathcal{P}^*), \quad u \in \mathcal{P}^*,$$

where $E'(u, z)$ is the directional derivative of E_u and

$$\nabla E_u = -\Delta \beta(u) + \operatorname{div}\big(Db^*(u)\big) \in H^{-1}.$$

This implies also that the flow $u(t) = S(t)u_0$ given by (1.7) is a.e. differentiable on $(0, \infty)$ in the norm of the Sobolev space $H^{-1} = H^{-1}(\mathbb{R}^d)$. This is the principal difference of our result compared with the Wasserstein based construction (1.16). In the later case, the gradient flow $S(t)$ is only continuous in L^1, but not differentiable in any convenient Banach space. 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 $L^1(\mathbb{R}^d) \cap L^\infty(\mathbb{R}^d)$.

One limitation of our approach might be that the gradient ∇E_u is defined on a subset $\mathcal{P}^*$ of the space $\mathcal{P}$ which implies that the flow $u(t) = S(t)u_0$ can be identified as a gradient flow only for solutions $u(t)$ with the property that $u(t) \in \mathcal{P}^*$, $\forall t \geq 0$. This happens, however, for equation (1.1) under hypotheses (i)–(iv).

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

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

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 by 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

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

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:

$$\beta'(r) \equiv \frac{d}{dr}\beta(r), \quad b'(r) = \frac{d}{dr}b(r), \quad b^*(r) \equiv b(r)r, \quad r \in \mathbb{R},$$

$$y_t = \frac{\partial}{\partial t}y, \quad \nabla y = \left\{\frac{\partial y}{\partial x_i}\right\}_{i=1}^d, \quad \Delta y = \sum_{i=1}^d \frac{\partial^2}{\partial^2 x_i}y,$$

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

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

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

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

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

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 the following theorem.

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*

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

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 $(0,\infty)$,

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

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,

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

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

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

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

In particular, it follows by Theorem 2.1 that the semigroup $S(t)$ is generated by the operator $-A^*$ 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 $\omega \geq 0$.*

PROOF. We have

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

for a suitable chosen $\omega \geq 0$ and so $A^* + \omega I$ is accretive in H^{-1}. (Here, we have used the inequality $|\nabla(I-\Delta)^{-1}(y_1 - y_2)|_2 \leq |y_1 - y_2|_{-1}$.) 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

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

The latter can be written as

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

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

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

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

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

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}. ∎

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

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

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

$$\frac{d^+}{dt} S^*(t)u_0 + A^*S^*(t)u_0 = 0, \quad \forall t \geq 0, \tag{2.12}$$

$$\frac{d}{dt} S^*(t)u_0 + A^*S^*(t)u_0 = 0, \quad \text{a.e. } t > 0, \tag{2.13}$$

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

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

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

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

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

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

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

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

where

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

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

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

By (i), this yields

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

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

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

Hence,

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

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, $S^*(t+s) = S^*(t)S^*(s)$, $t, s \ge 0$, we infer that $S^*(t)$ has a smoothing effect on initial data, that is,

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

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

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

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

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

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

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

This yields

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

and, therefore, by (2.8),

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

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

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

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 [12, proof of Lemma 3.1]), 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

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

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

PROOF OF THEOREM 2.1 (CONTINUED). We note first that

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

Indeed, by (2.17), it follows that if $u_0 \in L^1 \cap L^2$, then

$$|v_h^{j+1}|_1 \le |v_k^j|_1, \quad \forall j = 0, 1, \dots,$$

and, therefore,

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

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

$$\mathcal{X}_\delta(r) = \begin{cases} 1 & \text{for } r \ge \delta, \\ \frac{r}{\delta} & \text{for } r \in (-\delta, \delta), \\ -1 & \text{for } r \le -\delta. \end{cases}$$

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

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

which yields

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

Hence,

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

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 [8, Theorem 4.1] (see also [10, p. 161]), we have, for all $u_0 \in \mathcal{P}$ with $u_0\log u_0 \in L^1$,

$$E\big(S(t)u_0\big) + \int_0^t \Psi\big(S(\tau)u_0\big)d\tau \le E(u_0) < \infty, \quad \forall t \ge 0, \tag{2.23}$$

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 $S(t)(\mathcal{P}) \subset \mathcal{P}$, $\forall t \ge 0$, we have also that $u(t) \ge 0$, $\forall t \ge 0$, and $u(t) \in L^\infty$, $\forall t \ge 0$. So, it remains to prove that $\frac{\psi}{u(t)} \in L^1$, $\forall t \ge 0$. To this end, we consider the cut-off function

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

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

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

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

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

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

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

and so, by (2.14), we get

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

By (2.25), we have

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

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

On the other hand, we have by hypotheses (i)–(iii) that

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

and

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

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

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

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

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

This yields

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

while

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

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

Letting $n \to \infty$, we get

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

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 that

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

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:

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

(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)

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

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

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

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 $\operatorname{div}(b^*(u)\nabla(y_1 - y_2)) \equiv 0$. 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

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

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

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

and this yields

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

If we take $\varphi(x) = \eta(\frac{|x|^2}{n})$, 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

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

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 $\{y_n\} \subset W^{1,1}_{\mathrm{loc}}$ be such that

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

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

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

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

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

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

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

as claimed.

As a consequence, we have that

$$\big\{z = -\operatorname{div}\big(b^*(u)\nabla y\big);\ y \in C_0^\infty(\mathbb{R}^d)\big\} \text{ is dense in } \mathcal{T}_u(\mathcal{P}^*) \text{ for all } u \in \mathcal{P}^*. \tag{3.7}$$

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 $\mathcal{T}_u(\mathcal{P}^*)$-gradient $\nabla E_u : \mathcal{T}_u(\mathcal{P}^*) \to \mathcal{F}_u(\mathcal{P}^*) \subset H^{-1}$ of the energy function $E : L^2 \to]-\infty, +\infty]$ defined by (1.13), that is,

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

We note that E is convex, nonidentically $+\infty$, and we have the following result.

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)

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

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 $\lim_{n\to\infty} u_n = u$ 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)

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

and

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

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

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

Hence,

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

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 $\alpha \in (0, 1)$,

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

Hence, we find that

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

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

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

So, let us prove (4.6).

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

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

Hence, choosing ε small enough and α close enough to 1, so that $(\frac{1}{\alpha}-\varepsilon)/(1-\alpha) \geq m$, hypothesis (iv) implies (4.6). ∎

By Lemma 4.1, we have for E that its directional derivative

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

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

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

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

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

We recall that if E is Gâteaux differentiable at u, then ∂E_u reduces to the L^2-gradient $\nabla E_u \in L^2$ of E at u defined by

$$E'(u,z) = (\nabla E_u, z)_2, \quad \forall z \in L^2. \tag{4.9}$$

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 $z = z_2 = -\operatorname{div}(b^*(u)\nabla y_2)$, where $y_2 \in C_0^\infty(\mathbb{R}^d)$, then

$$z = -b^*(u)\Delta y_2 - \nabla y_2 \cdot \big(b'(u)u + b(u)\big)\nabla u$$

and so, by (i) and (1.14), it follows that $z \in L^2$. Then, by (4.7), we have

$$\begin{aligned} E'(u,z) &= \operatorname{diff} E_u \cdot z = \lim_{\lambda \downarrow 0} \frac{1}{\lambda}\big(E(u+\lambda z_2) - E(u)\big) \\ &= \int_{\mathbb{R}^d} \left(\frac{\nabla\beta\big(u(x)\big)}{b^*\big(u(x)\big)} - D(x)\right) b^*\big(u(x)\big) \cdot \nabla y_2(x)\,dx \\ &= \int_{\mathbb{R}^d} b^*\big(u(x)\big)\nabla y_2(x) \cdot \nabla\left(\int_0^{u(x)} \frac{\beta'(s)}{b^*(s)}\,ds + \Phi(x)\right) dx. \end{aligned} \tag{4.10}$$

We claim that

$$x \mapsto \int_0^{u(x)} \frac{\beta'(s)}{b^*(s)}\,ds + \Phi(x) \quad \text{in } W^{1,1}_{\text{loc}}. \tag{4.11}$$

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

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

Hence,

$$\frac{\gamma_1}{|b|_\infty}\log u \le \int_0^u \frac{\beta'(s)}{b^*(s)}\,ds \le \frac{\gamma_2}{b_0}\log u. \tag{4.12}$$

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

$$\begin{aligned} \int_{K_n} (\log u)^- dx &\le \left(\int_{K_n} (\log u)^2 u\,dx\right)^{\frac12} \big(\inf_K \psi\big)^{-\frac12} \left(\int_K \frac{\psi}{u}\,dx\right)^{\frac12} \\ &\le \sup_K \big((\log u)^- u\big)\left(\int_{K_n} (\log u)^- dx\right)^{\frac12} \big(\inf_k \psi\big)^{-\frac12} \left(\int_K \frac{\psi}{u}\,dx\right)^{\frac12}. \end{aligned}$$

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

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

Letting $\varepsilon \to 0$ yields $|\nabla \log u| \in L^1_{\text{loc}}$, and (4.11) is proved by hypothesis (iv). Hence, by (4.10),

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

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

$$z_2 \mapsto E'(u, z_2) = \operatorname{diff} E_u z_2$$

extends by continuity to all $z = z_2 \in \mathcal{T}_u(\mathcal{P}^*)$ and by (3.2) it follows that for $u^* \in \mathcal{P}^*$ there is $\nabla E_u \in \mathcal{T}_u(\mathcal{P}^*) \subset H^{-1}$ such that

$$\operatorname{diff}(E_u z_2) = (\nabla E_u, z_2)_u$$

and

$$\begin{aligned} \nabla E_u &= -\operatorname{div}\left(b^*(u)\nabla\Big(\int_0^u \frac{\beta'(s)}{b^*(s)}\, ds + \Phi\Big)\right) \\ &= -\Delta\beta(u) + \operatorname{div}\big(Db^*(u)\big) \in H^{-1}. \end{aligned} \tag{4.13}$$

This is by definition the $\mathcal{T}_u(\mathcal{P}^*)$-gradient of E at u. Though, for simplicity, we have denoted it again by ∇E_u, it should be mentioned that it does not coincide with the L^2-gradient defined by (4.9).

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 $u(t) \equiv S(t)u_0$,

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

$$\frac{d^+}{dt} S(t)u_0 = \Delta\beta\big(S(t)u_0\big) - \operatorname{div}\big(Db^*\big(S(t)u_0\big)\big), \quad \forall t > 0, \tag{4.14}$$

$$\frac{d}{dt} S(t)u_0 = \Delta\beta\big(S(t)u_0\big) - \operatorname{div}\big(Db^*\big(S(t)u_0\big)\big), \quad \forall t \in (0,\infty) \setminus N, \tag{4.15}$$

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 (4.13), we may rewrite (4.14)–(4.15) as the gradient flow on $\mathcal{P}^*$ endowed with the metric tensor (3.2). Namely, we have the following theorem.

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*

$$\frac{d}{dt}u(t) = -\nabla E_{u(t)}, \quad a.e.\ t > 0, \tag{4.16}$$

$$\frac{d^+}{dt}u(t) = -\nabla E_{u(t)}, \quad \forall t > 0, \tag{4.17}$$

$$\frac{d}{dt}u(t) = -\nabla E_{u(t)}, \quad \forall t \in (0,\infty) \setminus N, \tag{4.18}$$

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

By (3.2), we may rewrite (4.17) as

$$\frac{d^+}{dt}E\big(S(t)u_0\big) = -\left\|\frac{d^+}{dt}S(t)u_0\right\|^2_{u(t)}, \quad \forall t > 0. \tag{4.19}$$

Equivalently,

$$\frac{d^+}{dt}E\big(S(t)u_0\big) + A\big(S(t)u_0\big) = 0, \quad \forall t > 0, \tag{4.20}$$

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

$$\frac{d}{dt}E\big(S(t)u_0\big) = -\left\|\frac{d}{dt}S(t)u_0\right\|^2_{u(t)} = \Psi\big(S(t)u_0\big), \quad \forall t \in (0,\infty) \setminus N. \tag{4.21}$$

As a matter of fact, the energy dissipation formula (4.21) 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 $-\Delta\beta(u_\infty) + \operatorname{div}(Db(u_\infty)u_\infty) = 0$.

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

$$A^*u = B_u \operatorname{diff} E_u, \quad \forall u \in D(A^*) = H^1, \tag{4.22}$$

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

$$\begin{aligned} &B_u(y) = -\operatorname{div}(b^*(u)\nabla y), \quad \forall y \in D(B_u), \\ &D(B_u) = \{y \in l^2,\ \sqrt{u}\,\nabla y \in L^2\}. \end{aligned} \tag{4.23}$$

This means that ∇E_u can be equivalently written as

$$\nabla E_u = B_u(\operatorname{diff} E_u). \tag{4.24}$$

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

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

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.

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

REFERENCES

[1] S. ALBEVERIO – Y. G. KONDRATIEV – M. RÖCKNER, Differential geometry of Poisson spaces. *C. R. Acad. Sci. Paris Sér. I Math.* **323** (1996), no. 10, 1129–1134. Zbl 0866.58011 MR 1423438

[2] S. ALBEVERIO – Y. G. KONDRATIEV – M. RÖCKNER, Analysis and geometry on configuration spaces. *J. Funct. Anal.* **154** (1998), no. 2, 444–500. Zbl 0914.58028 MR 1612725

[3] L. AMBROSIO – N. GIGLI – G. SAVARÉ, *Gradient flows in metric spaces and in the space of probability measures*. Lectures Math. ETH, Birkhäuser, Basel, 2005. Zbl 1090.35002 MR 2129498

[4] V. BARBU, *Nonlinear differential equations of monotone types in Banach spaces*. Springer Monogr. Math., Springer, New York, 2010. Zbl 1197.35002 MR 2582280

[5] V. BARBU, *Semigroup approach to nonlinear diffusion equations*. World Scientific Publishing, Hackensack, NJ, 2022. Zbl 1528.35001 MR 4359030

[6] V. BARBU – T. PRECUPANU, *Convexity and optimization in Banach spaces*. 4th edn., Springer Monogr. Math., Springer, Dordrecht, 2012. Zbl 1244.49001 MR 3025420

[7] V. BARBU – M. RÖCKNER, From nonlinear Fokker-Planck equations to solutions of distribution dependent SDE. *Ann. Probab.* **48** (2020), no. 4, 1902–1920. Zbl 1469.60216 MR 4124528

[8] V. BARBU – M. RÖCKNER, The evolution to equilibrium of solutions to nonlinear Fokker-Planck equation. *Indiana Univ. Math. J.* **72** (2023), no. 1, 89–131. Zbl 1514.35436 MR 4557070

[9] V. Barbu – M. Röckner, Uniqueness for nonlinear Fokker-Planck equations and for McKean-Vlasov SDEs: The degenerate case. *J. Funct. Anal.* **285** (2023), no. 4, article no. 109980. Zbl 1514.60071 MR 4583740

[10] V. Barbu – M. Röckner, *Nonlinear Fokker-Planck flows and their probabilistic counterparts*. Lecture Notes in Math. 2353, Springer, Cham, 2024. Zbl 1548.35261 MR 4789626

[11] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*. Universitext, Springer, New York, 2011. Zbl 1218.46002 MR 2759829

[12] M. G. Crandall – A. Pazy, Semi-groups of nonlinear contractions and dissipative sets. *J. Functional Analysis* **3** (1969), 376–418. Zbl 0182.18903 MR 0243383

[13] T. D. Frank, *Nonlinear Fokker-Planck equations. Fundamentals and applications*. Springer Ser. Synergetics, Springer, Berlin, 2005. Zbl 1071.82001 MR 2118870

[14] R. Jordan – D. Kinderlehrer – F. Otto, The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.* **29** (1998), no. 1, 1–17. Zbl 0915.35120 MR 1617171

[15] S. Kondratyev – D. Vorotnikov, Nonlinear Fokker-Planck equations with reaction as gradient flows of the free energy. *J. Funct. Anal.* **278** (2020), no. 2, article no. 108310. Zbl 1430.49049 MR 4030267

[16] F. Otto, The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations* **26** (2001), no. 1–2, 101–174. Zbl 0984.35089 MR 1842429

[17] M. Rehmeier – M. Röckner, Nonlinear Fokker–Planck–Kolmogorov equations as gradient flows on the space of probability measures. [v1] 2023, [v4] 2024, arXiv:2306.09530v4.

[18] P. Ren – M. Röckner – F.-Y. Wang, Linearization of nonlinear Fokker-Planck equations and applications. *J. Differential Equations* **322** (2022), 1–37. Zbl 1486.60104 MR 4398417

[19] M. Röckner, Dirichlet forms on infinite-dimensional "manifold-like" state spaces: a survey of recent results and some prospects for the future. In *Probability towards 2000 (New York, 1995)*, pp. 287–306, Lect. Notes Stat. 128, Springer, New York, 1998. Zbl 0978.31007 MR 1632608

[20] F. Santambrogio, {Euclidean, metric, and Wasserstein} gradient flows: an overview. *Bull. Math. Sci.* **7** (2017), no. 1, 87–154. Zbl 1369.34084 MR 3625852

Received 25 February 2025,
and in revised form 18 September 2025

Viorel Barbu
Alexandru Ioan Cuza University
700506 Iaşi
Octav Mayer Institute of Mathematics, Romanian Academy
700505 Iaşi, Romania
vbarbu41@gmail.com

Michael Röckner
Faculty of Mathematics, Bielefeld University
33615 Bielefeld, Germany
Academy of Mathematics and Systems Science, Chinese Academy of Sciences
100190 Beijing, P. R. China
roeckner@math.uni-bielefeld.de