Steepest descend algorithm in the Wasserstein metric for a sandpile model

Viorel Barbu

(Iași, Romania)

OCERTO 2016. Cortona, June 20-24, 2016

Table of Contents

The problem

$$
\begin{aligned}
&\rho_t(t,x) - \Delta(\beta(x,\rho(t,x))) \ni 0 \quad && \text{in } (0,T)\times\mathbb{R}^n,\\
&\rho(0,x) = \rho_0(x) && \text{in } \mathbb{R}^n,
\end{aligned} \tag{1}
$$

$$\beta(x,r) = rH(r-\rho_c(x)),\ \forall r\in\mathbb{R},\ \ x\in\mathbb{R}^n,$$

H is the Heaviside function

$$\rho_0,\rho_c \in L^1(\mathbb{R}^n)\cap L^\infty(\mathbb{R}^n),$$
$$0<\rho_c(x)\le\rho_0(x),\quad \text{a.e. } \ x\in\mathbb{R}^n.$$

For $n = 2$, equation (1) describes the dynamic of the sandpile model of self-organized criticality which we briefly present below. If $\rho = \rho(t, x_1, x_2)$ is the energy (or mass density) assigned to the site $x = (x_1, x_2)$ of a square lattice at time t, then, if $\rho(t, x_1, x_2)$ exceeds the critical value $\rho_c(x_1, x_2)$, the site becomes unstable and an avalanche develops according to

$$\rho(t+1, x_1, x_2) = \rho(t, x_1, x_2 \pm 1) + \frac{1}{4}\,\rho(t, x_1, x_2),$$
$$\rho(t+1, x_1 \pm 1, x_2) = \rho(t, x_1 \pm 1, x_2) + \frac{1}{4}\,\rho(t, x_1, x_2).$$

If $\rho(t, j)$ is the energy of the j-th cell at time t,

$$\begin{aligned} \rho(t+1, k) &= \rho(t, k) - \rho(t, k) H(\rho(t, k) - \rho_c(k)) \\ &\quad + \frac{1}{4} \sum_{j \neq k} \rho(t, j) H(\rho(t, j) - \rho_c(j)). \end{aligned} \tag{2}$$

The initial configuration ρ_0 of the system is over the critical state $\rho_c = \rho_c(x)$, which is a natural hypothesis if one takes into account that the evolution begins in the unstable zone $\{x;\ \rho(t,x) \geq \rho_c(x)\}$.
Weak solution to the Cauchy problem (1)

$$\rho \in L^1((0,T) \times \mathbb{R}^n)$$

$$\frac{\partial \rho}{\partial t} - \Delta \eta = 0 \quad \text{in } \mathcal{D}'((0,T) \times \mathbb{R}^n), \tag{3}$$

$$\eta \in L^1_{\mathrm{loc}}((0,T) \times \mathbb{R}^n) \to \mathbb{R},$$

$$\eta(t,x) \in \beta(x, \rho(t,x)), \quad \text{a.e. } (t,x) \in (0,T) \times \mathbb{R}^n,$$

$$\lim_{t \to 0} \int_{\mathbb{R}^n} |\rho(t,x) - \rho_0(x)| dx = 0.$$

Let $A : D(A) \subset L^1(\mathbb{R}^n) \to L^1(\mathbb{R}^n)$

$$\begin{aligned} Ay &= yH(y-\rho_c),\ \forall y \in D(A) \\ D(A) &= \{y \in L^1(\mathbb{R}^n);\ y \geq 0,\ \text{a.e. in } \mathbb{R}^n,\ \exists\, \eta \in L^1_{\text{loc}}(\mathbb{R}^n); \\ &\qquad \eta(x) \in y(x)H(y(x)-\rho_c(x)), \\ &\qquad \text{a.e. } x \in \mathbb{R}^n,\ \Delta(yH(y-\rho_c)) \in L^1(\mathbb{R}^n)\}. \end{aligned}$$

A is accretive and

$$D(A) \subset \bigcap_{\lambda>0} R(I+\lambda A). \tag{4}$$

Consider the finite difference scheme for (1)

$$\begin{gathered} u_\varepsilon(t) = y_k^\varepsilon \text{ for } t \in [k\varepsilon,(k+1)\varepsilon),\ k = 0,1,\dots,N_\varepsilon = \left[\tfrac{T}{\varepsilon}\right] \\ y_{k+1}^\varepsilon + \varepsilon A y_{k+1}^\varepsilon = y_k^\varepsilon,\ k = 0,1,\dots, \end{gathered}$$

By the Crandall & Liggett exponential formula, $\{u_\varepsilon\}$ is convergent to a weak (or mild) solution to (1). Our aim is the convergence to ρ of the steepest descent type algorithm generated by a variational problem associated with the Wasserstein metric.

Consider the set of probability densities

$$\mathcal{P} = \Big\{\rho : \mathbb{R}^n \to [0,\infty),\ \text{Lebesgue measurable}, \int_{\mathbb{R}^n} \rho(x)dx = 1,\ \int_{\mathbb{R}^n} |x|^2 \rho(x)dx < \infty\Big\}.$$

On $\mathcal{P}$, we define the second order Wasserstein distance d

$$d^2(\rho^1,\rho^2) = \inf_{\mu\in\mathcal{M}(\rho^1,\rho^2)} \left\{\frac{1}{2}\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2 \mu(dx,dy)\right\},\ \forall \rho^1,\rho^2 \in \mathcal{P},$$

where $\mathcal{M}(\rho^1,\rho^2)$ is the set of all probability measures on $\mathbb{R}^n \times \mathbb{R}^n$ coupling the probability densities ρ^1 and ρ^2

$$\begin{aligned}
\int_{\mathbb{R}^n\times\mathbb{R}^n} \zeta(x)\mu(dx,dy) &= \int_{\mathbb{R}^n} \zeta(x)\rho^1(x)dx,\\
\int_{\mathbb{R}^n\times\mathbb{R}^n} \zeta(y)\mu(dx,dy) &= \int_{\mathbb{R}^n} \zeta(y)\rho^2(y)dy,\ \forall \zeta \in C_0^0(\mathbb{R}^n).
\end{aligned}$$

We associate with (1) the steepest descent algorithm

$$\rho_k^h(t) = \rho_k^h \ \text{ for } t \in [kh, (k+1)h), \ k = 0, 1, ..., N = \left[\frac{T}{h}\right], \qquad (5)$$

$$\begin{aligned} \rho_k^h &= \arg \min_{\rho \in \mathcal{P}} \left\{ \frac{1}{h} d^2(\rho, \rho_{k-1}^h) + \mathbb{E}(\rho) \right\}, \ k = 1, ..., N, \qquad (6) \\ \mathbb{E}(\rho) &= \int_{\mathbb{R}^n} g(x, \rho(x)) dx, \ \forall \rho \in \mathcal{P}. \end{aligned}$$

Here $g : \mathbb{R}^n \times \mathbb{R} \to \overline{\mathbb{R}} = (-\infty, +\infty]$, is the convex function

$$g(x, r) = \begin{cases} r \ln \left(\frac{r}{\rho_c(x)} \right) & \text{for} \quad r \geq \rho_c(x), \\ 0 & \text{for} \quad 0 < r < \rho_c(x). \end{cases}$$

In the following, we shall simply write ρ^h instead of ρ_k^h.

The main result is that, for $h \to 0$, the sequence $\{\rho^h\} \subset L^1((0,T) \times \mathbb{R}^n)$ is strongly convergent to the weak solution ρ to the Cauchy problem (1). For the linear Fokker–Planck equation

$$\rho_t - \Delta\rho + \mathrm{div}(\nabla\psi\rho) = 0 \text{ in } (0,T) \times \mathbb{R}^n,$$

a similar result was established by R. Jordan, D. Kinderlehrer and F. Otto, while, for the porous media equation, that is for $\beta(r) = ar^m$, $m \geq 2$, this problem was studied by F. Otto.

It should be emphasized that $\rho \to -\mathbb{E}(\rho)$ is the entropy functional associated with the sandpile process while equation (1) itself can be viewed as a nonlinear Fokker–Planck equation describing the particle transport in irregular media (the so called anomalous diffusion). In the limit case of linear diffusion, the energy $\mathbb{E}$ reduces to the Gibbs–Boltzmann entropy. The convergence of the steepest descent algorithms reveals that the dynamic of the self–organized criticality process involves the maximization of the system entropy at each time step.

The main result

The functional $\mathbb{E} : L^1(\mathbb{R}^n) \to [0, +\infty]$ is convex and lower semi-continuous on $\mathcal{P}$. Moreover, since

$$\lim_{|r|\to\infty} \frac{g(x,r)}{|r|} = +\infty, \quad \forall x \in \mathbb{R}^n,$$

it is easily seen by the Dunford–Pettis weak compactness criterium, that every level set $\{\rho \in \mathcal{P};\ \mathbb{E}(\rho) \leq \lambda\}$ is weakly compact in $L^1(\mathbb{R}^n)$. Since, for each $\rho_{k-1} \in \mathcal{P}$, the functional $\rho \to \frac{1}{h}\ d^2(\rho_{k-1}, \rho) + \mathbb{E}(\rho)$ is lower semicontinuous and strictly convex, it follows that there is a unique minimizer $\rho_k \in \mathcal{P}$ in (6). In particular, this means that the steepest descent algorithm is well defined.

Theorem 1

Theorem

Let $\rho_0 \in \mathcal{P} \cap L^\infty(\mathbb{R}^n)$ *and* $\{\rho^h\}$ *be the sequence of step functions defined by* (5). *Then, there is*

$$\lim_{h\to 0} \rho^h = \rho, \quad \textit{strongly in } L^1((0,T)\times\mathbb{R}^n),$$

where $\rho \in C([0,T]; L^1(\mathbb{R}^n)) \cap L^\infty((0,T)\times\mathbb{R}^n)$ *is the unique weak solution to the Cauchy problem* (1).

If

$$J(h)(\rho) = \arg \min_{u \in \mathcal{P}} \left\{ \frac{1}{h}\, d^2(\rho, u) + \mathbb{E}(u) \right\},\ \ \rho \in \mathcal{P},\ \ h > 0,$$

we may rewrite (5) as

$$\rho^h(t) = J^k(h)\rho_0 \text{ for } t \in [kh, (k+1)h),\ \ k = 0, 1, ..., N = \left[\tfrac{T}{h}\right],$$

and so, by Theorem 1,

$$\rho(t) = \lim_{k \to \infty} J^k\left(\frac{t}{k}\right)\rho_0, \quad \text{strongly in } L^1((0, T) \times \mathbb{R}^n).$$

For comparison, the convergence of finite difference scheme is equivalent to

$$\rho(t) = \lim_{k \to \infty} \left(I + \frac{t}{k}\, A\right)^{-k} \rho_0 \quad \text{strongly in } L^\infty(0, T; L^1(\mathbb{R}^n)).$$

Proof of Theorem 1

Let

$$\partial g(x,r) = \{\theta \in \mathbb{R};\ \theta(r-\bar{r}) \geq g(x,r) - g(x,\bar{r}),\ \forall \bar{r} \in \mathbb{R}^+\}.$$

$$\partial g(x,r) = \begin{cases} \ln\left(\frac{r}{\rho_c(x)}\right) + \rho_c(x) & \text{if} \quad r > \rho_c(x), \\ [0, \rho_c(x)] & \text{if} \quad r = \rho_c(x), \\ 0 & \text{if} \quad 0 \leq r < \rho_c(x). \end{cases}$$

$$\beta(x,r) = r\partial g(x,r) - g(r),\ \forall r \in \mathbb{R}^+,\ x \in \mathbb{R}^n.$$

Lemma 3.1

Lemma

$|\rho_k|_{L^\infty} \leq |\rho_0|_{L^\infty} + |\rho_c|_{L^\infty}, \ \forall k = 1, 2, \ldots$

Lemma 3.2

Lemma

$$\rho_k(x) \geq \rho_c(x), \ \ a.e. \ x \in \mathbb{R}^n, \ \forall k = 1, 2, ...$$

Lemma 3.3

Lemma

Let $\{\rho_k\}$ *be the sequence defined by* (6). *Then*

$$\frac{1}{h}\int_{\mathbb{R}^n}(x-y)\cdot\xi(x)\mu_k(dx,dy)=\int_{\mathbb{R}^n}\eta_k(x)\mathrm{div}\,\xi(x)dx,\ \forall\xi\in C_0^\infty(\mathbb{R}^n,\mathbb{R}^n),$$

$\eta_k\in L^1(\mathbb{R}^n)$, $\eta_k(x)\in\beta(x,\rho_k(x))$, *a.e.* $x\in\mathbb{R}^n$, *and* μ_k *is the optimal probability measure coupling* ρ_{k-1} *and* ρ_k.

Proof. Consider the diffeomorphism $\phi(t,x) : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$ generated by the equation

$$\begin{aligned} \frac{d}{dt}\,\phi(t,x) &= \xi(\phi(t,x)),\ t \geq 0,\ x \in \mathbb{R}^n, \\ \phi(0,x) &= x \end{aligned}$$

and define $\rho_t \in \mathcal{P}$ by the Monge-Ampère equation

$$\det(D\phi(t,x))\rho_t(\phi(t,x)) = \rho_k(x),\ t \in (0,1),\ x \in \mathbb{R}^n.$$

We have

$$\begin{aligned} \frac{d}{dt}\int_{\mathbb{R}^n} g(x,\rho_t(x))dx = \int_{\mathbb{R}^n} \zeta_k(t,y)\frac{d}{dt}(\det D\phi(t,y))^{-1}\det D\phi(t,y)) \\ +g(y,\rho_k(\phi^{-1}(t,y))(\det D\phi(t,y))^{-1})\frac{d}{dt}(\det D\phi(t,y)))dy, \end{aligned}$$

$$\zeta_k(t,y) \in \partial g(y,\rho_k(\phi^{-1}(t,y)\det(D\phi(t,y))^{-1}).$$

We obtain that

$$\frac{d}{dt}\int_{\mathbb{R}^n} g(x,\rho_t(x))dx\Big|_{t=0} = \int_{\mathbb{R}^n} \eta_k(y)\mathrm{div}\,\xi(y)dy,$$

where $\eta_k(y) \in \beta(y,\rho_k(y))$, a.e. $y \in \mathbb{R}^n$.

Now, if $\mu_t \in \mathcal{M}(\rho_{k-1}, \rho_t)$ is the optimal measure coupling ρ_{k-1} and ρ_t,

$$\begin{aligned} d^2(\rho_{k-1}, \rho_t) &= \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} |x - y|^2 \mu_t(dx, dy) \\ &= \frac{1}{2} \int_{\mathbb{R}^n \times \mathbb{R}^n} |\phi(t, x) - y|^2 \mu_k(dx, dy) \end{aligned}$$

because

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} \psi(x, y) \mu_t(dx, dy) = \int_{\mathbb{R}^n \times \mathbb{R}^n} \psi(x, \phi(t, y)) \mu_k(dx, dy), \ \forall \psi \in C_0^0(\mathbb{R}^n \times \mathbb{R}^n).$$

This yields

$$\frac{1}{2}(d^2(\rho_{k-1}, \rho_k) - d^2(\rho_{k-1}, \rho_k)) \leq \frac{1}{2t} \int_{\mathbb{R}^n \times \mathbb{R}^n} (|\phi(t, x) - y|^2 - |x - y|^2) \mu_k(dx, dy),$$

and, letting $t \to 0$,

$$\frac{1}{h} \limsup_{t \to 0} \left(\frac{d^2(\rho_{k-1}, \rho_t) - d^2(\rho_{k-1}, \rho_k)}{t} \right) \leq \int_{\mathbb{R}^n \times \mathbb{R}^n} \frac{x - y}{h} \cdot \xi(x) \mu_k(dx, dy), \quad \forall h > 0.$$

We get

$$\int_{\mathbb{R}^n\times\mathbb{R}^n} \frac{x-y}{h}\cdot\xi(x)\mu_k(dx,dy) \geq \int_{\mathbb{R}^n} \eta_k(x)\mathrm{div}\,\xi(x)dx,$$

and changing ξ in $-\xi$ one obtains the desired equation.
Consider now $\{\rho^h\} \subset L^1((0,T)\times\mathbb{R}^n)$ defined by

$$\eta^h(t,x) = \eta_k(x) \text{ for } t\in[kh,(k+1)h),\ k=0,1,...,N=\left[\frac{T}{h}\right],$$

where η_k are as in Lemma 3.3. We have

$$\eta^h(t,x)\in\beta(x,\rho^h(t,x)),\ \text{a.e. } (t,x)\in(0,T)\times\mathbb{R}^n.$$

Lemma 3.4

Lemma

There is C independent of k such that

$$\int_{\mathbb{R}^n} |x|^2 \rho_k(x) dx \leq C. \tag{7}$$

Lemma

The sequence $\{\rho_h\}_{h>0}$ is strongly compact in $L^1((0,T)\times\mathbb{R}^n)$.

Proof. We note that, for all $\xi \in (L^2(\mathbb{R}^n))^2$, we have

$$\left|\int_{\mathbb{R}^n} \nabla\eta_k(x)\cdot\xi(x)dx\right| \le \frac{1}{h}\left(\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2\mu_k(dx,dt)\right)^{\frac{1}{2}}\left(\int_{\mathbb{R}^n}|\xi|^2dx\right)^{\frac{1}{2}}$$
$$\le \frac{\sqrt{2}}{h}\, d(\rho_{k-1},\rho_k)\left(\int_{\mathbb{R}^n}|\xi|^2dx\right)^{\frac{1}{2}}.$$

This yields

$$\int_{\mathbb{R}^n} |\nabla\eta_k(x)|^2dx \le \frac{2}{h^2}\, d^2(\rho_{k-1},\rho_k),\ \ k=1,2,\dots,N,$$
$$\int_0^T\int_{\mathbb{R}^n} |\nabla\eta^h(t,x)|^2dx\,dt \le 2\mathbb{E}(\rho_0),\ \ \forall h.$$

$$\int_0^T \|\eta^h\|^2_{H^1(\mathbb{R}^n)} dt + \|\eta^h\|_{L^\infty((0,T)\times\mathbb{R}^n)} + \|\eta^h\|_{L^1(0,T)\times\mathbb{R}^n)} \leq C.$$

$$\lim_{|\bar{x}|\to 0} \int_0^T \int_{B_R} |\rho^h(t, x+\bar{x}) - \rho^h(t,x)| dx\, dt = 0, \ \forall R > 0,$$

uniformly in h.

$$\lim_{R\to\infty} \int_0^T \int_{B_R^c} |\rho^h(t,x)|^2 dx\, dt = 0.$$

Our aim is to derive the compactness of $\{\rho_h\}$ in $L^1((0,T)\times\mathbb{R}^n)$ by Kolmogorov's compactness theorem in L^p. To this end, we are going to prove that, for each $R > 0$,

$$\lim_{|\bar{x}|,|x|\to 0} \int_0^T \int_{B_R} |\rho^h(t+s, x+\bar{x}) - \rho^h(t,x)| dx\, dt = 0$$

uniformly in h.

For all $m, n \in \mathbb{N}$,

$$d^2(\rho_k, \rho_{k+m}) = \frac{1}{2}\int_{\mathbb{R}^n} \rho_{k+m}(x)|x - \nabla\psi(x)|^2 dx,$$
$$\int_{\mathbb{R}^n} \rho_k(x)\zeta(x)dx = \int_{\mathbb{R}^n} \rho_{k+m}(x)\zeta(\nabla\psi^*(x))dx,\ \forall \zeta \in C_0^0(\mathbb{R}^n).$$

We take $\zeta = \varphi(\eta_{k+m} - \eta_k)$, where $\varphi \in C_0^\infty(\mathbb{R}^n)$ and $\varphi \geq 0$ on $\mathbb{R}^n$.

$$\begin{aligned}
&\int_{\mathbb{R}^n} (\rho_{k+m}-\rho_k)(\eta_{k+m}-\eta_k)\varphi\, dx \\
&\quad = \int_{\mathbb{R}^n} \rho_{k+m}(x)(\zeta(x)-\zeta(\nabla\psi^*(x)))dx \\
&\quad \leq \int_0^1 \left(\int_{\mathbb{R}^n} \rho_{k+m}(x)(\nabla\zeta((1-\lambda)x - \lambda\nabla\psi(x))dx)\right)^{\frac{1}{2}} d\lambda \\
&\qquad\qquad \left(\int_{\mathbb{R}^n} \rho_{k+m}(x)|x - \nabla\psi^*(x)|^2 dx\right)^{\frac{1}{2}} \\
&\quad \leq \frac{C}{h}\left(d(\rho_{k+m-1}, \rho_{k+m}) + d(\rho_{k-1}, \rho_k)\right)\right) d(\rho_k, \rho_{k+m}).
\end{aligned}$$

We obtain

$$h \sum_{k \in \mathbb{N}} \int_{\mathbb{R}^n} \varphi |\rho_{k+m} - \rho_k|^2 dx \leq 2N \sum_{k \in \mathbb{N}} d^2(\rho_{k-1}, \rho_k) \leq CNh,$$

$$\int_0^{T-s} \int_{\mathbb{R}^n} \varphi(x)(\rho^h(t+s,x) - \rho^h(t,x))^2 dx\, dt \leq Cs,$$

for all $\varphi \in C_0^\infty(\mathbb{R}^n)$ and $h > 0$. Hence

$$\int_0^{T-s} \int_{\mathbb{B}_R} |\rho^h(t+s,x) - \rho^h(t,x)| dx\, dt \leq C_R |s|^{\frac{1}{2}},$$

for all $R > 0$ and $s \in (0,1)$.

Proof of Theorem 1 (continued).

Since the sequence $\{\rho^h\}_h$ is compact in $L^1((0,T)\times\mathbb{R}^n)$, and bounded in $L^\infty((0,T)\times\mathbb{R}^n)\cap L^2(0,T;H^1(\mathbb{R}^n))$, on a subsequence $\{h_n\}\to 0$, we have

$$\begin{aligned} \rho^{h_n} &\to \rho \quad \text{strongly in } L^1((0,T)\times\mathbb{R}^n) \\ &\qquad\ \ \text{and weak-star in } L^\infty((0,T)\times\mathbb{R}^n), \\ \eta^{h_n} &\to \eta \quad \text{weak-star in } L^\infty((0,T)\times\mathbb{R}^n), \\ &\qquad\ \ \text{and weakly in } L^2(0,T;H^1(\mathbb{R}^n)). \end{aligned}$$

By Lemma 3.4,

$$\int_{\mathbb{R}^n} |x|^2\rho(t,x)dt < \infty, \quad \text{a.e. } t\in(0,T),$$

while $\rho\geq\rho_c$, $\rho(t)\in\mathcal{P}$, a.e. $t\in(0,T)$.

Consider the function $j : \mathbb{R}^n \times \mathbb{R}^+ \to \mathbb{R}$

$$j(x,r) = \begin{cases} \frac{1}{2}\, r^2 & \text{for} \quad r \geq \rho_c(x), \\ \frac{1}{2}\, \rho_c^2(x) & \text{for} \quad r < \rho_c(x), \end{cases} \tag{8}$$

which is just the potential of β, that is, $\beta(x,r) = \partial_r j(x,r)$, $\forall r \in \mathbb{R}$.
We have

$$\eta^{h_n}(t,x)(\rho^{h_n}(t,x) - r(t,x)) \geq j(x, \rho^{h_n}(t,x)) - j(x, r(t,x)),$$
$$\text{a.e. } (t,x) \in (0,T) \times \mathbb{R}^n, \ \forall r \in L^1((0,T) \times \mathbb{R}^n).$$

Integrating on $(0, T) \times \mathbb{R}^n$ and letting $h_n \to 0$,

$$\int_0^T \int_{\mathbb{R}^n} \eta(t,x)(\rho(t,x) - r(t,x))dx\, dt \geq \int_0^T \int_{\mathbb{R}^n} j(x, \rho(t,x)) - j(x, r(t,x))dx\, dt.$$

$$\eta(t,x)(\rho(t,x) - r) \geq j(\rho(t,x)) - j(x,r), \ \forall r \in \mathbb{R}^n, \quad \text{a.e. } (t,x) \in (0,T) \times \mathbb{R}^n.$$

$$\eta(t,x) \in \beta(x, \rho(t,x)), \text{ a.e. } (t,x) \in (0,T) \times \mathbb{R}^n. \tag{9}$$

On the other hand, for all $\zeta \in C_0^\infty(\mathbb{R}^n)$,

$$\left| \int_{\mathbb{R}^n} \rho_k(x)\zeta(x) - \int_{\mathbb{R}^n} \rho_{k-1}(y)\zeta(y)dy - \int_{\mathbb{R}^n \times \mathbb{R}^n} (x-y) \cdot \nabla\zeta(y)\mu_k(dx\, dt) \right|$$
$$= \left| \int_{\mathbb{R}^n} \zeta(x) - \zeta(y) - (x-y) \cdot \nabla\zeta(y)\mu_k(dx, dy) \right|$$
$$\leq C \int_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^2 \mu_k(dx, dt) \leq Cd^2(\rho_k, \rho_{k-1}), \ \forall k = 1, \ldots,$$

where μ_k is the optimal measure coupling ρ_{k-1} and ρ_k.

For $\xi = \nabla\zeta$, $\zeta \in C_0^\infty(\mathbb{R}^n)$,

$$\left|\int_{\mathbb{R}^n} \rho_k(x)\zeta(x)dx - \int_{\mathbb{R}^n} \rho_{k-1}(y)\zeta(y)dy - \int_{\mathbb{R}^n} \eta_k(x)\Delta\zeta(x)dx\right| \\ \leq Cd^2(\rho_k, \rho_{k-1}),$$

and, for any $\varphi \in C_0^\infty((0, T) \times \mathbb{R}^n)$, we have

$$\left|\int_0^T \int_{\mathbb{R}^n} \rho^h(t,x)\,\frac{\varphi(t+h,x)-\varphi(t,x)}{h}\,dt\,dx + \int_h^T \int_{\mathbb{R}^n} \eta^h(t,x)\Delta\varphi(t,x)dt\,dx\right| \\ \leq C\sum_{k=1}^{\infty} d^2(\rho_k, \rho_{k-1} \leq Ch, \ \forall h.$$

Letting $h = h_n \to 0$, we get

$$\int_0^T \int_{\mathbb{R}^n} (\rho(t,x)\varphi_t(t,x) + \eta(t,x)\Delta\varphi(t,x))dt\,dx = 0,$$
$$\forall \varphi \in C_0^\infty((0,T)\times\mathbb{R}^n).$$

We also note that

$$\lim_{t\to 0}\int_{\mathbb{R}^n}(\rho(t,x)-\rho_0(x))\zeta(x)dx = 0,\ \forall\zeta\in L^2(\mathbb{R}^n).$$

It follows that $\eta \in L^2(0,T;H^1(\mathbb{R}^n))$

$$\frac{d}{dt}\rho(t) - \Delta\eta(t) = 0, \quad \text{a.e. in } (0,T), \tag{10}$$

where $\frac{d}{dt}\rho$ is the strong derivative of $\rho : [0,T] \to H^{-1}(\mathbb{R}^n)$.

Consider the functional $f : H^{-1}(\mathbb{R}^n) \to [0, \infty]$ defined by

$$f(u) = \begin{cases} \int_{\mathbb{R}^n} j(x, u(x))dx & \text{if } u \in L^1(\mathbb{R}^n) \cap H^{-1}(\mathbb{R}^n), \\ +\infty & \text{otherwise,} \end{cases}$$

If endow $H^{-1}(\mathbb{R}^n)$ with the norm

$$\|u\|_{H^{-1}(\mathbb{R}^n)} = \left(\int_{\mathbb{R}^n} u(I - \Delta)^{-1} u \, dx \right)^{\frac{1}{2}},$$

then the subdifferential $\partial f : H^{-1}(\mathbb{R}^n) \to H^{-1}(\mathbb{R}^n)$

$$\partial f(u) = \{(I - \Delta)w; \ w \in H^{-1}(\mathbb{R}^n), \ w(x) \in \beta(x, u(x)), \ \text{a.e. } x \in \mathbb{R}^n\},$$

and so

$$\frac{d}{dt}\rho + (I - \Delta)\eta = \eta, \ \text{a.e. } t \in (0, T),$$

$$\eta(t) \in \partial f(\rho(t)), \ \text{a.e. } t \in (0, T),$$

$$\frac{d}{dt} f(\rho(t)) + \int_{\mathbb{R}^n} |\eta(t, x)|^2 dx = \int_{\mathbb{R}^n} \eta(t, x)(I - \Delta)^{-1}\eta(t, x)dx, \ \text{a.e. } t \in (0, T).$$

Taking into account that $\eta(I-\Delta)^{-1}\eta \in L^1(0,T;L^2(\mathbb{R}^n))$, it follows that $t \to f(\rho(t))$ is absolutely continuous on $[0,T]$ and, recalling that

$$\rho(t,x) \geq \rho_c(x), \text{ a.e. } (t,x) \in (0,T) \times \mathbb{R}^n,$$

we infer that the function $t \to \int_{\mathbb{R}^n} \rho^2(t,x)dx$ is continuous on $[0,T]$. Since $\lim_{t\to 0} \rho(t) = \rho_0$ weakly in $L^2(\mathbb{R}^n)$, it follows that

$$\lim_{t\to 0} \rho(t) = \rho_0 \text{ strongly in } L^2(\mathbb{R}^2),$$

and so in $L^1(\mathbb{R}^n)$, too. Hence, $\rho \in L^1(0,T;L^1(\mathbb{R}^n))$ is $L^1(\mathbb{R}^n)$-continuous and so it is a weak solution to the Cauchy problem (1). The Cauchy problem (1) has at most one weak solution $\rho \in L^\infty((0,T)\times\mathbb{R}^n)$. This implies that

$$\rho = \lim_{h\to 0} \rho^h \text{ in } L^1((0,T)\times\mathbb{R}^n).$$

References

- P. Bak, C. Tang, K. Wiesenfeld, Self-organized criticality, *Phys. Review.*, 38 (1988), 364-374.
- P. Bantay, I.M. Janosi, Self-organization and anomalous diffusion, *Physica*, A, 185 (1992), 11-18.
- Ph. Benilan, H. Brezis, M.G. Crandall, A semilinear elliptic equation in $L^1(\mathbb{R}^n)$, *Ann. Scuola Normale Superiore Pisa*, Serie IV, II (1975), 523-555.
- H. Brezis, M.G. Crandall, Uniqueness of solutions of the initial-value problem for $u_t - \Delta\varphi(u) = 0$, *J. Math. Pures Appl.*, 58 (1979), 153-163.
- R. Cafiero, V. Loreto, L. Pietronero, A. Vespignani, S. Zapperi, Local rigidity and self-organized criticality, *Europhys. Lett.*, 29 (12) (1995), 111-116.

- E.A. Carlen, W. Gangbo, *Annals of Mathematics*, 157 (2003), 807-846.
- L.C. Evans, Partial differential equations and Monge–Kantorovich mass transfer, *Current Development in Mathematics*, International Press, boston, MA, 1999.
- R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker-Planck equations, *SIAM J. Math. Anal.*, 29 (1998), 1-17.
- F. Otto, Dynamics of labyrinthine pattern formation in magnetic fluids: a mean-field theory, *Arch. Rational Mech. Anal.*, 141 (1998), 63-103.
- F. Otto, The geometry of dissipative evolution equations: the porous medium equations, *Comm. Partial Different. Equations*, 26 (2001), 101-174.