Introduction to Optimal Transportation and its applications, part III (prelim. version)

Luigi Ambrosio

Scuola Normale Superiore, Pisa
http://cvgmt.sns.it
l.ambrosio@sns.it

Outline

Gradient flows

(M, g) Riemannian (or Finsler) manifold, $F : M \to \mathbb{R}$.

$$\text{(GF)} \qquad x'(t) = -\nabla F\big(x(t)\big), \qquad x(0) = \bar{x}$$

Ingredients. Energy F, Metric (distance) g.

The metric g is implicitly used in (GF) to relate $d_x F \in T_x^* M$ (a covector) to $\nabla F \in T_x M$ (a vector):

$$dF_x(v) = g_x(\nabla F(x), v) \qquad \forall v \in T_x M.$$

SCUOLA NORMALE SUPERIORE

Hilbertian theory

Let H be Hilbert and $\lambda \geq 0$. We shall consider convex or more generally λ-convex functions $F : X \to \mathbb{R} \cup \{+\infty\}$:

$$F\big((1-t)x + ty\big) \leq (1-t)F(x) + tF(y) - \frac{\lambda}{2} t(1-t)|x-y|^2.$$

This corresponds to $\nabla^2 F \geq \lambda I$.
The subdifferential of F satisfies the *monotonicity inequality*

$$\langle \xi - \eta, x - y \rangle \geq \lambda |x-y|^2.$$

Also, the subdifferential inequality holds in a stronger form:

$$F(y) \geq F(x) + \langle v, y - x \rangle + \frac{\lambda}{2}|y-x|^2 \qquad \forall y \in H,\ v \in \partial F(x).$$

We denote, as usual, by $\nabla F(x)$ the element of minimal norm of $\partial F(x)$. It exists and is unique by convexity of $\partial F(x)$.

SCUOLA NORMALE SUPERIORE

Gradient flow as a differential inclusion

For convex functions a much more flexible formulation of (GF), replacing the equation with a subdifferential inclusion, is:

$$\text{(GFI)} \qquad \begin{cases} x'(t) \in -\partial F(x(t)) & \text{for a.e. } t > 0 \\ x \in AC^2_{\mathrm{loc}}\big((0,+\infty); H\big) \\ \lim_{t\downarrow 0} x(t) = \bar{x}. \end{cases}$$

We now summarize the main existence and uniqueness result in this context.

Theorem 1. (Brezis, Pazy) *Let* $F : H \to \mathbb{R} \cup \{+\infty\}$ *be convex and l.s.c.*

(i) *(Existence and uniqueness) for all* $\bar{x} \in \overline{D(F)}$ (GFI) *has a unique solution;*

(ii) *(Minimal selection and energy identity) for* $\mathscr{L}^1$*-a.e.* t, $x'(t) = -\nabla F(x(t))$, *so that also (GF) has a unique solution; in addition* $F(x(t)) \in AC_{\rm loc}(0, +\infty)$ *and*

$$F(x(s)) - F(x(t)) = \int_s^t |\nabla F(x(\tau))|^2 \, d\tau \qquad 0 < s \le t < \infty;$$

(iii) *(Regularizing effects)* $x'_+(t) = -\nabla F(x(t))$ *and* $(F \circ x)'_+(t) = -|\nabla F(x(t))|^2$ *for all* $t > 0$*. Finally*

$$F(x(t)) \le \inf_{v \in D(F)} F(v) + \frac{1}{2t}|v - \bar{x}|^2,$$

$$|\nabla F(x(t))|^2 \le \inf_{v \in D(\partial F)} |\nabla F(v)|^2 + \frac{1}{t}|v - \bar{x}|^2.$$

SCUOLA NORMALE SUPERIORE

(iv) (Asymptotic behaviour)

$$F(x(t)) - F(x_{\min}) \leq \big(F(\bar{x}) - F(x_{\min})\big) e^{-2\lambda t} \qquad t \geq 0.$$

In particular, if $\lambda > 0$, the (pointwise) *energy inequality*

$$F(x) \geq F(x_{\min}) + \frac{\lambda}{2}|x - x_{\min}|^2$$

gives

$$|x(t) - x_{\min}| \leq \sqrt{\frac{2(F(x) - F(x_{\min}))}{\lambda}} e^{-\lambda t}.$$

SCUOLA NORMALE SUPERIORE

Gradient flows in $\mathscr{P}_2(\mathbb{R}^n)$

Having in mind the differentiable structure of $\mathscr{P}_2(\mathbb{R}^n)$, we may define μ_t a *gradient flow* of $\mathcal{E}$ if

$$\text{(GF)} \quad \begin{cases} \frac{d}{dt}\mu_t + \nabla\cdot\big(v_t\mu_t\big) = 0 \\ -v_t \in \partial^W\mathcal{E}(\mu_t) \subset \mathrm{Tan}_{\mu_t}\big(\mathscr{P}_2(\mathbb{R}^n)\big) \text{ for } \mathscr{L}^1\text{-a.e. } t > 0. \end{cases}$$

For instance, if

$$\mathcal{F}(\rho) = \int \rho\ln\rho + \rho V\,dx,$$

we know that $\partial^W\mathcal{F}(\rho) = \frac{\nabla\rho}{\rho} + \nabla V$, hence (GF) reduces to the Fokker-Planck equation

$$\frac{d}{dt}\rho_t = \nabla\cdot\big(\big(\frac{\nabla\rho_t}{\rho_t} + \nabla V\big)\rho_t\big) = \Delta\rho_t + \nabla\cdot(\nabla V\rho_t).$$

SCUOLA NORMALE SUPERIORE

Energy Dissipation Inequality

Both the system (GF) and the energy identity are encoded in a single inequality:

$$\frac{d}{dt}F(x(t)) \leq -\frac{1}{2}|\nabla F|^2(x(t)) - \frac{1}{2}|x'(t)|^2.$$

Indeed, along *any* curve $y(t)$, we have

$$\begin{aligned} &\frac{d}{dt}F(y(t)) = \langle \nabla F(y(t)), y'(t) \rangle \\ \geq \quad & -|\nabla F(y(t))||y'(t)| \quad (= \text{iff } -y'(t) \text{ is parallel to } \nabla F(y(t))) \\ \geq \quad & -\frac{1}{2}|\nabla F|^2(y(t)) - \frac{1}{2}|y'(t)|^2 \quad (= \text{iff } |\nabla F|(y(t)) = |y'(t)|.) \end{aligned}$$

SCUOLA NORMALE SUPERIORE

Energy Dissipation Inequality

It is technically more convenient to consider the inequality in an integral form, namely

$$\text{(EDI)}\quad \frac{1}{2}\int_0^t |x'(r)|^2\,dr + \frac{1}{2}\int_0^t |\nabla F(x(r))|^2\,dr \le F(\bar{x}) - F(x(t)).$$

We are going to show that:

- (EDI) makes sense also in metric spaces;
- (EDI) has a discrete in time counterpart.

Before doing that we turn to an even stronger formulation, the *Evolution Variational Inequality*. It relies on two ingredients: the energy inequality (derived from convexity) and the derivative of distance squared.

SCUOLA NORMALE SUPERIORE

Evolution Variational Inequality

In Hilbert spaces both ingredients are easy: if x solves (GFI) we have

$$\begin{aligned}\frac{d}{dt}\frac{1}{2}|x(t)-y|^2 &= \langle x'(t), x(t)-y\rangle = \langle -x'(t), y-x(t)\rangle \\ &\le F(y)-F(x(t))-\frac{\lambda}{2}|y-x(t)|^2.\end{aligned}$$

This gives:

Definition. (E,d) metric, $x:(0,+\infty)\to E$ locally absolutely continuous. We say that x is an (EVI) solution if

$$\text{(EVI)}\quad \frac{d}{dt}\frac{1}{2}d^2(x(t),y)\le F(y)-F(x(t))-\frac{\lambda}{2}d^2(x(t),y)\quad \text{for } \mathscr{L}^1\text{-a.e. } t>0$$

for all $y\in D(F)$.

Under suitable more restrictive assumptions on (d,F), we will see that (EVI) has a discrete version as well.

The (EVI) formulation is very strong and it leads easily to stability (as we will see, even with respect to convergence of the energies) and to contractivity.

Theorem 2. *Let x, y be solutions to (EVI) starting from $\bar{x}$ and $\bar{y}$ respectively. Then $d(x(t), y(t)) \leq d(\bar{x}, \bar{y})e^{-\lambda t}$ for all $t \geq 0$.*

Proof. Insert $y = y(t)$ in

$$\frac{d}{dt}d^2(x(t), y) \leq 2F(y) - 2F(x(t)) - \lambda d^2(x(t), y)$$

and $x = x(t)$ in

$$\frac{d}{dt}d^2(x, y(t)) \leq 2F(x) - 2F(y(t)) - \lambda d^2(x, y(t))$$

to obtain that $\frac{d}{dt}d^2(x(t), y(t)) \leq -2\lambda d^2(x(t), y(t))$. This argument can be made rigorous with Kruzkhov method of doubling of variables.

SCUOLA NORMALE SUPERIORE

Implicit Euler scheme

$\tau > 0$ time step, $\bar{x}$ initial condition.

We define (x_n) recursively setting $x_0 := \bar{x}$ and choosing x_{n+1}, at each step, among the minimizers of

$$y \mapsto F(y) + \frac{1}{2\tau} d^2(y, x_n).$$

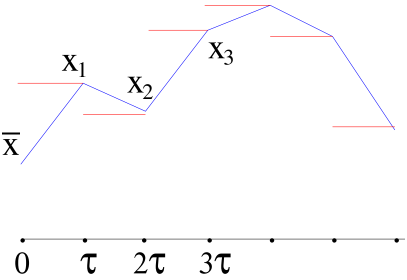

Figure: Discrete solution

$x_\tau(t)$ piecewise constant

$$x_\tau(t) := x_{n+1},\ t \in (n\tau, (n+1)\tau]$$

The Hilbert, convex case

The minimality of x_{n+1} gives the *discrete Euler equation*

$$(*) \qquad \frac{x_{n+1} - x_n}{\tau} \in -\partial F(x_{n+1}),$$

so that $x_{n+1} = (Id + \tau\partial F)^{-1}(x_n)$.
In terms of the piecewise affine interpolant $\tilde{x}_\tau(t)$, (*) reads

$$\tilde{x}'_\tau(t) \in -\partial F\big((Id + \tau\partial F)^{-1}(\tilde{x}_\tau(\tau[t/\tau]))\big).$$

This is the *explicit* time discretization scheme for the ODE $y' = -(\partial F)_\tau(y)$, where

$$(\partial F)_\tau := \frac{Id - (Id + \tau\partial F)^{-1}}{\tau} = \partial F \circ (Id + \tau\partial F)^{-1}.$$

This ODE is used in the classical existence proof of (GF) by approximation, since $(\partial F)_\tau$ is Lipschitz.

SCUOLA NORMALE SUPERIORE

In order to show that (x_τ) converge as $\tau \downarrow 0$ to a continuous solution (GF) we have to read (EDI) or (EVI) *inside* the Euler scheme.

SCUOLA NORMALE SUPERIORE

How (EDI) can be read in the Euler scheme

Definition (Slope) Let $F : E \to \mathbb{R} \cup \{+\infty\}$ and $x \in D(F)$. We set

$$|\partial F|(x) := \limsup_{y \to x} \frac{[F(x) - F(y)]^+}{d(x,y)}.$$

Equivalently, $|\partial F|(x)$ is the smallest $C \geq 0$ satisfying

$$F(y) \geq F(x) - Cd(x,y) + o\big(d(x,y)\big).$$

With this characterization, a simple application of the Hahn-Banach theorem gives

$$|\partial F|(x) = |\nabla F(x)| \ \big(= \min\{|\xi| : \ \xi \in \partial F(x)\}\big)$$

when $E = H$ and F is convex and l.s.c.

How (EDI) can be read in the Euler scheme

Using the concepts of metric derivative $|x'(t)|$ and slope $|\partial F|(x(t))$ we can formulate (EDI) for absolutely continuous maps $x : [0, +\infty) \to E$:

$$\frac{1}{2}\int_0^t |x'|^2(r)\,dr + \frac{1}{2}\int_0^t |\partial F|^2(x(r))\,dr \le F(\bar{x}) - F(x(t)) \quad \forall t \ge 0.$$

Lemma 3. (First discrete Euler equation) *Let $y \in E$ be a minimizer of $v \mapsto F(v) + d^2(v, x)/2\tau$. Then*

$$(*) \qquad |\nabla F|(y) \le \frac{d(x, y)}{\tau}.$$

Proof.

$$F(y) - F(\tilde{y}) \le \frac{1}{2\tau}\left\{d^2(\tilde{y}, x) - d^2(y, x)\right\} \le \frac{d(y, \tilde{y})}{2\tau}\big(d(x, y) + d(x, \tilde{y})\big).$$

Dividing both sides by $d(y, \tilde{y})$ and taking the limit as $\tilde{y} \to y$ yields (*). □

SCUOLA NORMALE SUPERIORE

How (EDI) can be read in the Euler scheme

Now we can interpolate between x and y *in a variational way*, as follows: For $\sigma \in (0,\tau)$ we choose y_σ among the minimizers of

$$\tilde{y} \mapsto F(\tilde{y}) + \frac{1}{2\sigma} d^2(\tilde{y}, x).$$

Lemma 4. *$g(\sigma) := F(y_\sigma) + d^2(y_\sigma, x)/2\sigma \in \mathrm{Lip}_{\mathrm{loc}}((0,\tau])$ and*

$$g'(\sigma) = -\frac{1}{2\sigma^2} d^2(y_\sigma, x) \qquad \text{for } \mathscr{L}^1\text{-a.e. } \sigma \in (0,\tau].$$

Proof. $g(\sigma+h) - g(\sigma)$ can be bounded from above by

$$\Big(F(y_\sigma) + \frac{d^2(y_\sigma,x)}{2(\sigma+h)}\Big) - \Big(F(y_\sigma) + \frac{d^2(y_\sigma,x)}{2\sigma}\Big) = -h\frac{d^2(y_\sigma,x)}{2\sigma^2} + o(h).$$

At any differentiability point σ we get $g'(\sigma) = -d^2(y_\sigma, x)/2\sigma^2$. □

How (EDI) can be read in the Euler scheme

Since $g(0+) = F(x)$ and $g(\tau) = F(x) + d^2(x,y)/2\tau$, by integration between 0 and τ Lemma 4 gives

$$\begin{aligned} F(x) - F(y) &= \frac{d^2(x,y)}{2\tau} + \int_0^\tau \frac{d^2(y_\sigma, x)}{2\sigma^2}\, d\sigma \\ &\geq \frac{\tau}{2}\frac{d^2(x,y)}{\tau^2} + \frac{1}{2}\int_0^\tau |\partial F|^2(y_\sigma)\, d\sigma \\ &\sim \frac{\tau}{2}|x'|^2 + \frac{\tau}{2}|\partial F|^2. \end{aligned}$$

Adding all these inequalities with $x = x_n$, $y = x_{n+1}$ one obtains that the discrete solution (variationally interpolated in time) fulfils a discrete version of (EDI). Suitable lower semicontinuity assumptions on the slope ensure that limit curves fulfil (EDI).

SCUOLA NORMALE SUPERIORE

Convexity and negative curvature in metric spaces

Definition. (Convexity in length metric spaces) $F : E \to \mathbb{R} \cup \{+\infty\}$ is said to be λ-convex if $t \mapsto F(\gamma(t))$ is $\lambda d^2\big(\gamma(0), \gamma(1)\big)$-convex along all constant speed geodesics $\gamma : [0,1] \to E$. Equivalently, for all $t \in [0,1]$:

$$F(\gamma(t)) \le (1-t)F(\gamma(0)) + tF(\gamma(1)) - \frac{\lambda}{2}t(1-t)d^2\big(\gamma(0), \gamma(1)\big).$$

Definition (Non Positively Curved metric spaces) A length space (E, d) is said to be NPC if $\frac{1}{2}d^2(\cdot, z)$ is 1-convex for all $z \in E$.

The Hilbertian identity

(Hid)

$$\frac{1}{2}\big|(1-t)x + ty - z\big|^2 = \frac{1-t}{2}|x-z|^2 + \frac{t}{2}|y-z|^2 - \frac{t(1-t)}{2}|x-y|^2$$

shows that Hilbert spaces are NPC.

Compatibility of energy and distance

The theory of gradient flows works well for convex functionals in NPC spaces (Jost, Majer). But, the space we are interested in, namely $\mathscr{P}_2(H)$, is PC!

Theorem 5. ($\mathscr{P}_2(H)$ is PC) *For all constant speed geodesics* $\mu_t : [0,1] \to \mathscr{P}_2(H)$ *and all* $\sigma \in \mathscr{P}_2(H)$ *we have*

$$\frac{1}{2} W_2^2(\mu_t, \sigma) \geq \frac{1-t}{2} W_2^2(\mu_0, \sigma) + \frac{t}{2} W_2^2(\mu_1, \sigma) - \frac{t(1-t)}{2} W_2^2(\mu_0, \mu_1).$$

Proof. We use the Hilbertian identity (Hid) and the following fact: for all $t \in [0,1]$, $\pi \in \Gamma_0(\mu_0, \mu_1)$, $\pi' \in \Gamma_0(\mu_t, \sigma)$ there exists $\eta \in \mathscr{P}(H^3)$ with

$$(\pi_1, \pi_2)_\sharp \eta = \pi, \qquad \left((1-t)\pi^1 + t\pi^2, \pi^3\right)_\sharp \eta = \pi'$$

(notice that η has μ_0, μ_1, σ as marginals).

Using η we can estimate

$$
\begin{aligned}
W_2^2(\mu_t,\sigma) = \int |x_1 - x_2|^2 \, d\pi' = \int \big|(1-t)x_1 + tx_2 - x_3\big|^2 \, d\eta \\
= \quad (1-t)\int |x_1 - x_3|^2 \, d\eta + t\int |x_2 - x_3|^2 \, d\eta - t(1-t)\int |x_1 - x_2|^2 \\
= \quad (1-t)W_2^2(\mu_0,\sigma) + tW_2^2(\mu_1,\sigma) - t(1-t)W_2^2(\mu_0,\mu_1),
\end{aligned}
$$

because $(\pi_1,\pi_3)_\sharp\eta \in \Gamma(\mu_0,\sigma)$, $(\pi_2,\pi_3)_\sharp\eta \in \Gamma(\mu_1,\sigma)$ and $(\pi_1,\pi_2)_\sharp\eta \in \Gamma_0(\mu_0,\mu_1)$.
It remains to show the existence of η. To this aim, remember that $\Gamma_0(\mu_t,\mu_0)$ contains only an element, induced by a transport T_0, and $\Gamma_0(\mu_t,\mu_1)$ contains only an element, induced by a transport T_1. Furthermore, by cyclical monotonicity, the map

$$(x,y) \mapsto z_t(x,y) = (1-t)x + ty$$

is injective on $\operatorname{supp}\pi$ and its inverse is precisely (T_0, T_1) on $z_t(\operatorname{supp}\pi) \supset \operatorname{supp}\mu_t$ and $\pi = (T_0, T_1)_\sharp\mu_t$.

SCUOLA NORMALE SUPERIORE

Setting $\Psi(z, x_3) := \big(T_0(z), T_1(z), y\big)$ and $\eta := \Psi_\sharp \pi'$, we get

$$\big((1-t)\pi_1 + t\pi_2, \pi_3\big)_\sharp \eta = \big((1-t)T_0 + tT_1, y)_\sharp \pi' = \pi'$$

because $(1-t)T_0(z) + tT_1(z) = z$ and

$$(\pi_1, \pi_2)_\sharp \eta = (T_0, T_1)_\sharp \pi' = (T_0, T_1)_\sharp \mu_t = \pi. \ \square$$

Definition. (Compatibility of energy and distance) We say that F and d are *compatible* if $\forall x,\ y,\ z \in E$ there exists a continuous curve $\gamma : [0,1] \to E$ with $\gamma(0) = x$, $\gamma(1) = y$ and

(i) $F(\gamma(t)) \leq (1-t)F(x) + tF(y) - \frac{\lambda}{2} t(1-t) d^2(x,y)$;

(ii) $\frac{1}{2} d^2(\gamma(t), z) \leq \frac{1-t}{2} d^2(x,z) + \frac{t}{2} d^2(y,z) - \frac{\lambda}{2} t(1-t) d^2(x,y)$ for all $t \in [0,1]$.

Of course, F and d are always compatible if E is NPC and F is convex (the interpolating curves being geodesics from x to y, independently of z).

SCUOLA NORMALE SUPERIORE

Lemma 6. (Second discrete Euler equation) *If F and d are compatible, and y minimizes $z \mapsto F(z) + d^2(z,x)/2\tau$, then*

$$\frac{1}{2\tau}\big(d^2(y,z) - d^2(x,z)\big) \le F(z) - F(y) \qquad \forall z \in E.$$

Proof. Let $\gamma(t)$ from y to z given by the compatibility condition (so that $t \mapsto d^2(\gamma(t),z)/2$ is 1-convex);

$$\begin{aligned}
&F(y) + \frac{1}{2\tau}d^2(y,x) \le F(\gamma(t)) + \frac{1}{2\tau}d^2(\gamma(t),x)\\
&(1-t)F(y) + tF(z) + \frac{(1-t)}{2\tau}d^2(y,x) + \frac{t}{2\tau}d^2(z,x)\\
&-\frac{t(1-t)}{2\tau}d^2(y,z).
\end{aligned}$$

SCUOLA NORMALE SUPERIORE

Rearrangement of terms and division by $t > 0$ gives

$$F(y) - F(z) \leq \frac{1}{2\tau} d^2(x,z) - \frac{1-t}{2\tau} d^2(y,z)$$

and we can let $t \downarrow 0$ to conclude. $\square$

Now, using the discrete EVI property just proved

$$(\mathrm{EVI}_\tau) \quad \frac{1}{2\tau}\big(d^2(x_{n+1},z) - d^2(x_n,z)\big) \leq F(z) - F(x_{n+1}) \quad \forall z \in E$$

we can recover precisely the Hilbertian theory, stability and even error estimates, under the following assumptions:

• (E,d) metric complete;
• $F \geq 0$, compatible with d;
• The discrete semigroup $S_\tau \bar{x}(t)$ (equal to x_{n+1} in $(n\tau,(n+1)\tau]$) exists for τ sufficiently small.

Theorem 7. (Existence) *For all* $\bar{x} \in \overline{D(F)}$ *there exists a (unique) solution* $S\bar{x}$ *of (EVI) starting from* $\bar{x}$*. If* $\bar{x} \in D(F)$ *we have the apriori error estimate*

$$\sup_{t \geq 0} d\big(S_\tau \bar{x}(t), S\bar{x}(t)\big) \leq 8\sqrt{\tau}\sqrt{F(\bar{x})}.$$

Theorem 8. (Regularizing effects and pointwise formulation) $x(t) = S\bar{x}(t)$ *satisfies:*

(i) *the right metric derivative* $|x'_+|(t)$ *exists for all* $t > 0$ *and coincides with* $|\partial F|\big(x(t)\big)$*;*

(ii) $t \mapsto F(x(t))$ *is locally absolutely continuous in* $(0, +\infty)$*, right differentiable and*

$$\frac{d}{dt+} F(x(t)) = -|\partial F|^2\big(x(t)\big) \quad \forall t > 0.$$

SCUOLA NORMALE SUPERIORE

(iii)

$$F(x(t)) \leq \inf_{v\in D(F)} F(v) + \frac{1}{2t}|v-\bar{x}|^2,$$

$$|\partial F|^2(x(t)) \leq \inf_{v\in D(|\partial F|)} |\nabla F|^2(v) + \frac{1}{t}|v-\bar{x}|^2.$$

Let us prove convergence of S_τ and the apriori error estimate when $\bar{x}\in D(F)$. The strategy is to compare S_τ and $S_{\tau/2}$ using $(\mathrm{EVI})_\tau$.

Lemma 9. *For $t = n\tau$, $n\geq 1$, we have*

$$d^2\big(S_\tau\bar{x}(t), S_{\tau/2}\bar{y}(t)\big) - d^2(\bar{x},\bar{y}) \leq 2\tau F(\bar{x}).$$

Proof. (*Step 1.*) First we prove the inequality for $t=\tau$, in a stronger form:

$$d^2\big(S_\tau\bar{x}(\tau), S_{\tau/2}\bar{y}(\tau)\big) - d^2(\bar{x},\bar{y}) \leq 2\tau\big[F(\bar{x}) - F(S_\tau\bar{x}(\tau))\big].$$

SCUOLA NORMALE SUPERIORE

By $(\mathrm{EVI})_{\tau/2}$ we get

$$d^2(S_{\tau/2}\bar{y}(\tau/2), z) - d^2(\bar{y}, z) \le \tau\big[F(z) - F(S_{\tau/2}\bar{y}(\tau/2))\big],$$

$$d^2(S_{\tau/2}\bar{y}(\tau), z) - d^2(S_{\tau/2}\bar{y}(\tau/2), z) \le \tau\big[F(z) - F(S_{\tau/2}\bar{y}(\tau))\big],$$

whose sum (taking into account that $F(S_{\tau/2}\bar{y}(\tau))$ is smaller than $F(S_{\tau/2}\bar{y}(\tau/2))$) gives

$$(1) \qquad d^2(S_{\tau/2}\bar{y}(\tau), z) - d^2(\bar{y}, z) \le 2\tau\big[F(z) - F(S_{\tau/2}\bar{y}(\tau))\big].$$

By $(\mathrm{EVI})_{\tau}$ we also get

$$(2) \qquad d^2(S_\tau\bar{x}(\tau), z) - d^2(\bar{x}, z) \le 2\tau\big[F(z) - F(S_\tau\bar{x}(\tau))\big].$$

SCUOLA NORMALE SUPERIORE

Now, set $z = \bar{x}$ in (1) and $z = S_{\tau/2}\bar{y}(\tau)$ in (2) and add, to get

$$(3) \quad d^2\big(S_\tau\bar{x}(\tau), S_{\tau/2}\bar{y}(\tau)\big) - d^2(\bar{x},\bar{y}) \le 2\tau\big[F(\bar{x}) - F(S_\tau\bar{x}(\tau))\big].$$

(*Step 2.*) By (3) we get

$$\begin{aligned} &d^2\big(S_\tau\bar{x}(2\tau), S_{\tau/2}\bar{y}(2\tau)\big) - d^2(S_\tau\bar{x}(\tau), S_{\tau/2}\bar{y}(\tau)) \\ &\le 2\tau\big[F(S_\tau\bar{x}(\tau)) - F(S_\tau\bar{x}(2\tau))\big] \end{aligned}$$

and we can add this to (3), obtaining

$$d^2\big(S_\tau\bar{x}(2\tau), S_{\tau/2}\bar{y}(2\tau)\big) - d^2(\bar{x},\bar{y}) \le 2\tau\big[F(\bar{x}) - F(S_\tau\bar{x}(2\tau))\big].$$

Iterating this procedure the inequality is achieved at all times $t = n\tau$ with n integer. $\Box$

SCUOLA NORMALE SUPERIORE

Convergence of S_τ

Lemma 9 with $\bar{x} = \bar{y}$ gives

$$d\big(S_{\tau/2^i}\bar{x}(t), S_{\tau/2^{i+1}}\bar{x}(t)\big) \leq 2^{(1-i)/2}\sqrt{\tau}\sqrt{F(\bar{x})}.$$

As a consequence, $S_{\tau/2^i} \to S^\tau$ as $i \to \infty$, with

$$d\big(S_\tau\bar{x}(t), S^\tau(t)\big) \leq \sum_{i=0}^{\infty} 2^{(1-i)/2}\sqrt{\tau}\sqrt{F(\bar{x})}.$$

We conclude showing that S^τ solves (EVI), and therefore does not depend on the initial τ.

SCUOLA NORMALE SUPERIORE

Convergence of S_τ

To this aim, it suffices to read $(\mathrm{EVI})_\tau$ as follows

$$\frac{d}{dt}\frac{1}{2}d^2\big(S_\tau\bar{x}(t),z\big) \le \tau\sum_{n=0}^{\infty}\big(F(z)-F(S_\tau\bar{x}((n+1)\tau))\big)\delta_{n\tau}$$

in the distribution sense.
Since $\sum_n \tau\delta_{n\tau}$ weakly converge to $\mathscr{L}^1$ and (here we replace $(n+1)\tau$ by the interpolation $\tau+\tau[t/\tau]$)

$$\limsup_{i\to\infty} -F\big(S_{\tau/2^i}\bar{x}(\tau+\tau[t/\tau])\big) \le -F(S^\tau\bar{x}(t))$$

we get

$$\frac{d}{dt}\frac{1}{2}d^2\big(S^\tau\bar{x}(t),z\big) \le \big(F(z)-F(S^\tau\bar{x}(t))\big)\mathscr{L}^1. \square$$

SCUOLA NORMALE SUPERIORE

Proof of regularizing effects

Fix $v \in E$. Integrate in $(0,t)$ (EVI) and use monotonicity of $t \mapsto F(x(t))$ to get

$$\frac{1}{2}\left(d^2(x(t),v) - d^2(\bar{x},v)\right) \le \int_0^t F(v) - F(x(s))\, ds \le t(F(v) - F(x(t)).$$

In particular

$$(*) \qquad F(x(t)) \le F(v) + \frac{1}{2t} d^2(v,\bar{x}).$$

In order to prove the regularization of $|\partial F|$ we use the slope estimate

$$\frac{F(u) - F(v)}{d(u,v)} \le |\partial F|(u)$$

(a consequence of the convexity of F), the monotonicity of $|\partial F|(x(s))$ and

$$\lim_{t \downarrow 0} t\, F(x(t)) = 0.$$

The latter is easily implied by (*) and $\bar{x} \in \overline{D(F)}$.

SCUOLA NORMALE SUPERIORE

Proof of regularizing effects

$$
\begin{aligned}
\frac{t^2}{2}|\partial F|^2(x(t)) &\leq \int_0^t s|\partial F|^2(x(s))\,ds = -\int_0^t s(F\circ x)'(s)\,ds \\
&= \int_0^t F(x(s))\,ds - tF(x(t)) \\
&\overset{(EVI)}{\leq} tF(v) + \frac{1}{2}\big(d^2(\bar{x},v) - d^2(v,x(t))\big) - tF(x(t)) \\
&\leq t|\partial F|(x(t))d(v,x(t)) + \frac{1}{2}\big(d^2(\bar{x},v) - d^2(x(t),v)\big) \\
&\leq \frac{t^2}{2}|\partial F|^2(v) + \frac{1}{2}d^2(\bar{x},v)\ \Box.
\end{aligned}
$$

SCUOLA NORMALE SUPERIORE

Stability of (EVI) solutions

Theorem 9. (Stability) (E, d) *metric,* F_h, $F : E \to \overline{\mathbb{R}}$. *Assume that*

(i) (F_h, d) *and* (F, d) *are compatible;*

(ii) $\Gamma(d)\text{-}\limsup_{h\to\infty} F_h \leq F$*;*

(iii) $\Gamma(\sigma)\text{-}\liminf_{h\to\infty} F_h \geq F$ *for some weak topology* σ *for which* d *is sequentially l.s.c;*

(iv) F_h *are equi-coercive in bounded subsets of* E *for the topology* σ*.*

(v) $x_h \to x$ *for the topology* σ *and* $d(x_h, \bar{x}) \to d(x, \bar{x})$ *implies* $d(x_n, x) \to 0$.

Then, if $\bar{x}_h \in D(F_h)$ *satisfy* $d(\bar{x}_h, \bar{x}) \to 0$ *and* $F_h(\bar{x}_h) \to F(\bar{x}) \in \mathbb{R}$, *the corresponding (EVI) solutions converge locally uniformly in* $[0, +\infty)$*.*

In the applications σ is the topology of weak convergence: it is weaker than the metric topology of $\mathscr{P}(H)$, by the moment convergence condition.

SCUOLA NORMALE SUPERIORE

Stability of (EVI) solutions

Sketch of proof. By the universal error estimate

$$d\big(S\bar{x}(t), S_\tau\bar{x}(t)\big) \leq 8\sqrt{\tau}\sqrt{F(\bar{x})}$$

we need only to show pointwise convergence of the discrete semigroups. By induction, we are led to the following statement: if $d(x_h, x) \to 0$, and y_h minimizes $z \mapsto F_h(z) + d^2(z, x_h)/2\tau$, then $d(y_h, y) \to 0$, where y minimizes

$$z \mapsto F(z) + \frac{1}{2\tau} d^2(z, x).$$

The assumptions we made on F_h imply, whenever $x_h \to x$,

$$\Gamma(\sigma) - \lim_{h\to\infty} F_h(\cdot) + \frac{1}{2\tau} d^2(\cdot, x_h) = F(\cdot) + \frac{1}{2\tau} d^2(\cdot, x)$$

and therefore convergence of minimizers to minimizers.

Applications of the theory in $\mathscr{P}_2(H)$

It turns out that all the concept introduced so far (subdifferential inclusion, energy dissipation, evolution variational inequality) turn out to be equivalent:

Theorem 10. *If $\mathcal{E} : \mathscr{P}_2(H) \to [0, +\infty]$ is convex along geodesics, then the (GF), (EDI) and (EVI) formulations are equivalent.*
If, in addition, $\mathcal{E}$ is convex along all interpolating curves, then $\mathcal{E}$ and W_2 are compatible, hence we have error estimates for the Euler scheme.

We now illustrate a convex and two nonconvex examples of gradient flows in $\mathscr{P}_2(H)$.

SCUOLA NORMALE SUPERIORE

Examples of gradient flows: the Fokker-Planck equations

We have seen that the Relative Entropy functional $\mathcal{H}(\cdot|\gamma)$ is convex along geodesics and along interpolating curves if γ is log-concave. Therefore the general theory of gradient flows in $\mathscr{P}_2(H)$ is applicable:

Theorem. (A-Savaré-Zambotti) *Let $\gamma \in \mathscr{P}(H)$ be log-concave. Then the gradient flow $S^\gamma\mu(t)$ of $\mathcal{H}(\cdot|\gamma)$ defines a continuous contraction semigroup in $\mathscr{P}_2(H)$. Furthermore, $\mathcal{S}^\gamma$ is stable with respect to weak convergence of γ and finite-dimensional approximations.*

This provides the natural extension of the finite-dimensional FP equations. Indeed, if $H = \mathbb{R}^n$, $\gamma = e^{-V}\mathscr{L}^n$, and $S^\gamma\mu(t) = \rho_t\mathscr{L}^n = u_t\gamma$, then

$$\frac{d}{dt}\rho_t = \nabla\cdot(\nabla\rho_t + \rho_t\nabla V), \qquad \frac{d}{dt}u_t = \Delta_\gamma u_t.$$

SCUOLA NORMALE SUPERIORE

Examples of gradient flows: the Fokker-Planck equations

It is useful to compare with the traditional “linear” viewpoint: first, one proves the existence of the semigroup P_t in $L^2(\gamma)$ generated by the Dirichlet form

$$\mathcal{E}(u,v) := \int_H \langle \nabla u, \nabla v \rangle \, d\gamma,$$

namely

$$\frac{d}{dt}\langle P_t u, v\rangle = -\mathcal{E}_\gamma(P_t u, v) \qquad \forall v \in D(\mathcal{E}_\gamma)$$

(this requires the proof that $\mathcal{E}_\gamma$ is *closable*).

Then, if P_t fulfils additional regularizing properties, one can define a dual semigroup P_t^* in the space of measures (which gives, in particular, the transition probabilities):

$$\langle P_t^* \mu, \phi\rangle := \langle \mu, P_t \phi\rangle.$$

Examples of gradient flows: the Fokker-Planck equations

In our approach, a direct construction of the dual semigroup P_t^*, which coincides with S_t^γ, is provided. Furthermore, even if we want to follow the conventional viewpoint, to prove closability of $\mathcal{E}_\gamma$ we need an input from optimal transport theory: for $u \geq 0$ with $\int u^2 \, d\gamma = 1$, we use the representation

$$4\mathcal{E}_\gamma(u,u) = \int \left| \frac{\nabla u^2}{u^2} \right|^2 u^2 \, d\gamma = \|\nabla^W \mathcal{H}(\cdot|\gamma)(u^2)\|^2_{L^2(u^2\gamma)}$$

and the fact that, for convex functionals Φ, $\mu \mapsto \|\nabla^W \Phi(\mu)\|^2_{L^2(\mu)}$ is lower semicontinuous.

Extensions to Gaussian Wiener spaces have been given recently by Fang-Shao-Sturm and Maas.

SCUOLA NORMALE SUPERIORE

Examples of gradient flows: a fourth order PDE

Quantum drift diffusion equation (Derrida-Lebowitz-Spohn)

$$\text{(DLS)} \qquad \frac{d}{dt}u + \nabla \cdot \left(u\nabla(\frac{\Delta\sqrt{u}}{\sqrt{u}})\right) = 0 \quad \text{in } \Omega \times (0,\infty)$$

Bleher-Lebowitz-Spohn, Jüngel-Pinnau, Carillo-Toscani, Dolbeault-Gentil-Jüngel,...
(short time existence, 1-dimensional solutions, asymptotics)

With suitable variational boundary conditions, (DLS) can be interpreted as the Wasserstein gradient flow of the Fisher information functional

$$\mathcal{F}(u) := 4\int_\Omega |\nabla\sqrt{u}|^2\,dx = \int |\nabla \ln u|^2 u\,dx.$$

Indeed, it turns out that, formally,

$$\nabla^W \mathcal{F}(u) = -\nabla(\frac{\Delta\sqrt{u}}{\sqrt{u}}).$$

SCUOLA NORMALE SUPERIORE

Notice that

$$\mathcal{F}(u) = \int \Big|\frac{\nabla \rho}{\rho}\Big|^2 \rho \, dx$$

so it can also be intepreted as $\|\nabla^W \int \rho \ln \rho \, dx\|^2_{L^2(\rho)}$, i.e. the energy dissipation rate in the heat equation. Similar remarks apply if one changes the reference measure from $\mathscr{L}^n$ to a Gaussian γ.

Even though $\mathcal{F}(u)$ *is not* convex along geodesics (and so only the (EDI) formulation of the gradient flow is applicable), Gianazza-Savaré-Toscani used this interpretation and the Euler scheme to derive global existence results for (DLS), as well as energy dissipation identities.

Examples of gradient flows: the CRS model

Chapman-Rubinstein-Schatzman model:

$$\frac{d}{dt}\mu_t - \nabla \cdot \left(\nabla h_{\mu_t}\mu_t\right) = 0 \qquad \begin{cases} -\Delta h_\mu + h_\mu = \mu & \text{in } \Omega \\ h_\mu = 1 & \text{on } \partial\Omega \end{cases}$$

in $\Omega \times (0, +\infty)$ London equation

It can be viewed (A-Serfaty) as the gradient flow of

$$\mathcal{E}(\mu) := \frac{\lambda}{2}\mu(\Omega) + \frac{1}{2}\int_\Omega |\nabla h_\mu|^2 + (h_\mu - 1)^2.$$

This interpretation yields entropies, existence, (partial) uniqueness results.

$\mathcal{E}$ arises (Sandier-Serfaty) as the mean field limit of

$$G_\varepsilon(u, A) := \frac{1}{2}\int_\Omega |(\nabla - iA)u|^2 + |\mathrm{curl}\, A - h_{ex}^\varepsilon|^2 + \frac{(1-|u|^2)^2}{2\varepsilon^2}.$$

SCUOLA NORMALE SUPERIORE