A Γ-convergence result and an application to the derivation of the Monge-Ampère gravitational model

Luigi Ambrosio

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

Overview

[1] L.A., AYMERIC BARADAT, YANN BRENIER: *Monge-Ampère gravitation as a Γ-limit of good rate functions.* Preprint, 2020.

[2] YANN BRENIER: *A double large deviation principle for Monge-Ampère gravitation.* Bull. Inst. Math. Acad. Sin., **11** (2016), 23–41.

We derive the discrete version of the Vlasov-Monge-Ampère system starting from a stochastic model of a Brownian point cloud

$$\lim_{\varepsilon\to 0}\lim_{\eta\to 0} X^{\varepsilon,\eta}, \qquad dX_t^{\varepsilon,\eta} = v_\varepsilon(t, X^{\varepsilon,\eta})dt + \eta(t)dB_t,$$

where the inner limit is based on the Freidlin-Wentzell theorem and the outer limit relies on Γ-convergence.

Compared to the paper [2] the new contribution is on the Γ-convergence result, which makes the ϵ-limit more rigorous.

SCUOLA NORMALE SUPERIORE

Plan

1. Action functionals induced by convex functions
2. The Γ-convergence result
3. The Vlasov-Monge-Ampère gravitational model
4. Derivation of VMA via large deviations and Γ-convergence

SCUOLA NORMALE SUPERIORE

Action functionals induced by convex functions

Let H be a Hilbert space, $\lambda \in \mathbb{R}$, $f : H \to \mathbb{R} \cup \{+\infty\}$ λ-convex, l.s.c. (and proper). For $h_0,\ h_1 \in H$ we consider the action functional

$$\Lambda_f(h_0, h_1) : C([0,1]; H) \to \mathbb{R} \cup \{+\infty\}$$

defined by

$$\begin{cases} \int_0^1 |x'(t)|^2 + |\nabla f(x(t))|^2 \, dt & \text{if } x \in AC^2([0,1]; H),\ x(i) = h_i,\ i = 0,\ 1 \\ +\infty & \text{otherwise.} \end{cases}$$

The goal is to analyze the stability of $\Lambda_f(h_0, h_1)$ w.r.t. variational convergence of f and convergence of the endpoints h_i.

SCUOLA NORMALE SUPERIORE

Meaning of ∇f

In the case $\lambda \geq 0$, for $x \in D(f) = \{f < +\infty\}$, $\nabla f(x)$ is the element with minimal norm in the subdifferential $\partial f(x)$:

$$\partial f(x) := \{p \in H : \; f(y) \geq f(x) + \langle p, y - x \rangle \;\; \forall y \in H\}.$$

However, a "variational" characterization of $|\nabla f(x)|$ can be provided

$$|\nabla f(x)| = \sup_{y \neq x} \frac{[f(x) - f(y)]^+}{|x - y|}.$$

It yields that $x \mapsto |\nabla f(x)|$ is lower semicontinuous in H, a very useful property also in non-Hilbertian contexts.

SCUOLA NORMALE SUPERIORE

Lack of continuity of L

In general terms, the Lagrangian $L(x,p) = |p|^2 + |\nabla f(x)|^2$ is only l.s.c. w.r.t. x, even in finite dimensions. So, the regularity of minimizers of Λ (ensured, e.g., by a coercitivity assumption on f) is problematic.

One can prove that

$$f \text{ Lipschitz on bounded sets} \qquad \Longrightarrow \qquad |x'| \in L^\infty(0,1)$$

thanks to the Du Bois-Reymond equation

$$\frac{d}{dt}\big[x'(t)L_p(x(t),x'(t)) - L(x(t),x'(t))\big] = 0$$

that can be obtained just performing variations in the independent variable.

Can we derive a EL equation, formally $x''(t) = \nabla^2 f(x(t))\nabla f(x(t))$, or get higher regularity?

A particular case

In the case of the application to VMA, $H = \mathbb{R}^{Nd}$ and, for some $A = (a_1, \dots, a_N) \in H$, f is the semiconvex function

$$f(x) = -\frac{1}{2} \min_{\sigma \in \Sigma_N} |x - A^\sigma|^2,$$

with the notation $A^\sigma = (a_{\sigma(1)}, \dots, a_{\sigma(N)})$.

Notice that, out of singularities of the distance, one has $|\nabla f|^2 = |f|^2$, hence we may replace Λ_f by the simpler functional

$$\Lambda_f'(x) := \int_0^1 |x'(t)|^2 + |f(x(t))|^2 \, dt.$$

However, the "effective" functional will be the more difficult one with $|\nabla f(x(t))|^2$!

Resolvent map $J_\tau f$

For $-2\tau\lambda < 1$, $J_\tau f(x) = (Id + \tau\partial f)^{-1}(x)$ is the unique minimizer of the map

$$y \mapsto f(y) + \frac{1}{2\tau}|y - x|^2$$

and the minimal value $f_\tau(x)$ is the Moreau-Yosida approximation of f.

Theorem. *($\lambda \geq 0$) $J_\tau f$ is a contraction, f_τ is convex and $f_\tau \in C^{1,1}(H)$ with $\mathrm{Lip}(\nabla f_\tau) \leq \tau^{-1}$. Moreover*

$$p \in \partial f(x) \qquad \Longleftrightarrow \qquad p = \nabla f_\tau(x + \tau p).$$

In particular, choosing $p = \nabla f(x)$ gives

$$|\nabla f(x)| = |\nabla f_\tau|(x + \tau \nabla f(x)).$$

Variants of Λ

It is not hard to include non-autonomous variants of Λ or even replace the action by

$$\int_0^1 \left|x'(t) - \nabla f(x(t))\right|^2 dt.$$

This is due to the fact that a nonsmooth chain rule (for instance Thm. 1.2.5 [AGS]) gives that $t \mapsto f(x(t))$ is absolutely continuous in $[0,1]$ (in particular $h_i \in D(f)$)) whenever $|x'|$ and $|\nabla f(x)|$ belong to $L^2(0,1)$, with

$$\frac{d}{dt} f(x(t)) = \langle \nabla f(x(t)), x'(t) \rangle \qquad \text{a.e. in } (0,1).$$

Therefore, the product term is a null Lagrangian. Playing with the convexity parameter λ one can consider also

$$\int_0^1 \left|x'(t) - (\lambda x(t) - \nabla f(x(t)))\right|^2 dt.$$

Mosco convergence versus Γ-convergence

Definition. *We say that $f_n \to f$ Mosco-converge to f if:*

(i) $\liminf_n f_n(x_n) \geq f(x)$ *whenever $x_n \to x$ weakly in H;*

(ii) *for all $x \in H$ there exist $x_n \to x$ strongly, with $\limsup_n f_n(x_n) \leq f(x)$.*

• In finite dimensions, no difference w.r.t. the usual version of Γ-convergence. In infinite dimensions, it is more appropriate, as it ensures strong convergence of resolvents: $J_\tau f_n(x) \to J_\tau f(x)$.

• Under an equi-coercitivity assumption w.r.t. the strong topology of H, again the two versions of Γ-convergence become equivalent and, in addition, the infimum of $\Lambda_f(h_0, h_1)$ is always attained.

SCUOLA NORMALE SUPERIORE

Main Γ-convergence result

Theorem. *If* $f_n : H \to \mathbb{R} \cup \{+\infty\}$ *are* λ*-convex and l.s.c., with* $f_n \to f$ *w.r.t. Mosco convergence, and if*

$$h_{n,i} \to h_i \ \textit{strongly}, \qquad \sup_n |\nabla f_n(h_{n,i})| < \infty, \qquad i = 0,\ 1,$$

then $\Lambda_{f_n}(h_{n,0}, h_{n,1})$ Γ*-converge to* $\Lambda_f(h_0, h_1)$ *in the* $C([0,1]; H)$ *topology.*

Sketch of proof ($\lambda = 0$): $\Gamma - \liminf$

The $\Gamma - \liminf$ inequality follows immediately from the variational characterization of $|\nabla f(x)|$, which yields the *joint* lower semicontinuity

$$\liminf_{n\to\infty} |\nabla f_n(x_n)|^2 \geq |\nabla f(x)|^2 \qquad \text{whenever } x_n \to x \text{ strongly.}$$

This would not work if the weak convergence of the x_n were weak and this fact forces the use of the $C([0,1];H)$ topology.

The proof of the $\Gamma - \limsup$ inequality (construction of the recovery sequence) uses the strong convergence of resolvents, and for this reason Mosco convergence is needed.

Sketch of proof ($\lambda = 0$): $\Gamma - \limsup$

Fix $x(t)$ with $\Lambda_f(h_0, h_1)(x(\cdot)) < \infty$, $\tau > 0$ and set

$$x_n^\tau(t) = J_\tau f_n(x(t)), \qquad x^\tau(t) = J_\tau f(x(t)).$$

The monotonicity properties

$$|\nabla f(J_\tau f(h))| \leq \frac{|h - J_\tau f(h)|}{\tau} \leq |\nabla f(h)| \qquad h \in H$$

together with the contractivity of $J_\tau f$ yield

$$\begin{aligned}
\limsup_{n\to\infty} \int_0^1 |(x_n^\tau)'|^2 + |\nabla f_n(x_n^\tau)|^2 \, dt &\leq \limsup_{n\to\infty} \int_0^1 |x'|^2 + \frac{|x - x_n^\tau|^2}{\tau^2} \, dt \\
&= \int_0^1 |x'|^2 + \frac{|x - x^\tau|^2}{\tau^2} \, dt \\
&\leq \int_0^1 |x'|^2 + |\nabla f|^2(x) \, dt.
\end{aligned}$$

Adjustement of the endpoints

The recovery sequence $x_n(t) = x_n^{\tau(n)}(t)$ can be obtained by a diagonal argument, except for the fact that the endpoint conditions $x_n(0) = h_{n,0}$, $x_n(1) = h_{n,1}$ a priori are not satisfied.

However, calling $h^*_{n,i} := x_n(i) = J_{\tau(n)} f_n(h_i)$ the “wrong” terminal values, we at least have $h^*_{n,i} \to h_i$. In addition, the monotonicity properties of the resolvent grant

$$\limsup_{n\to\infty} |\nabla f_n(h^*_{n,i})|^2 < \infty$$

provided $\tau(n) \to 0$ sufficiently slowly.

This, combined with the assumption $\limsup_n |\nabla f_n(h_{n,i})|^2 < \infty$, grant the possibility to interpolate, in small intervals, $[-\delta_n, 0]$, $[1, 1+\delta_n]$ between $h_{n,i}$ and $h^*_{n,i}$ with a small cost.

Finally, a rescaling of $[-\delta_n, 1+\delta_n]$ to $[0, 1]$ gives the result.

SCUOLA NORMALE SUPERIORE

A nonlinear interpolation lemma

Lemma. *Let* $f : H \to \mathbb{R} \cup \{+\infty\}$ *be convex, l.s.c. Then for all* $\tau > 0$ *and all* $q_0, q_1 \in D(\partial f)$ *the infimum of* $\Lambda_f(q_0, q_1)$ *can be estimated from above by*

$$2|\nabla f(q_0)|^2 + \left(8 + \frac{4}{\tau^2}\right)|q_1 - q_0|^2 + \left(4 + 8\tau^2\right)|\nabla f(q_0) - \nabla f(q_1)|^2.$$

Sketch of proof. Having in mind that $p \in \partial f(x)$ iff $p = \nabla f_\tau(x + \tau p)$, we first interpolate linearly between $q_i + \tau\nabla f(q_i)$, $i = 0, 1$

$$\gamma(t) := (1 - t)\big(q_0 + \tau\nabla f(q_0)\big) + t\big(q_1 + \tau\nabla f(q_1)\big)$$

and then we go back to the “original variables” to get an admissible curve $x(t)$ from q_0 to q_1:

$$x(t) := \gamma(t) - \tau\nabla f_\tau(\gamma(t)).$$

The Lipschitz bound on ∇f_τ and the equivalence $\nabla f(x) = \nabla f_\tau(\gamma)$ give the result.

The VMA system

The Vlasov-Monge-Ampère system, introduced by Y.Brenier in 2011

$$\partial_t f(t,x,\xi) + \mathrm{div}_x\big(\xi f(t,x,\xi)\big) - \mathrm{div}_\xi\big(\nabla\phi(t,x) f(t,x,\xi)\big) = 0$$

$$\det\big(I + \nabla^2\phi(t,x)\big) = \varrho(t,x), \qquad \varrho(t,x) = \int f(t,x,\xi)\,d\xi$$

can be viewed as a nonlinear variant of the classical Vlasov-Poisson system, as

$$\det\big(I + \nabla^2\phi(t,x)\big) \sim 1 + \Delta\phi(t,x).$$

Formally, it can also be viewed (A.-Gangbo) as an Hamiltonian ODE in the Wasserstein space of probabilities in phase space, with velocity

$$v_t(x,\xi) := \big(\xi, -\nabla\phi(t,x)\big)$$

and $\frac{1}{2}|x|^2 + \phi(t,\cdot)$ is the K-potential of the optimal transport problem from $\varrho(t,\cdot)$ to the uniform measure.

Discrete version of the VMA system

If we replace (say on the torus) the uniform measure by a family of N points $A = (a_1, \ldots, a_N)$ with N large, and the continuous densities by discrete ones $(x_1, \ldots, x_N)$, formally VMA corresponds to

$$(*) \qquad x_i''(t) = x_i(t) - a_{\sigma_{\rm opt}}(i) \qquad i = 1, \ldots, N$$

because the dynamic is ruled by the discrete optimal transport problem

$$\min_{\sigma \in \Sigma_N} \sum_{i=1}^{N} |x_i - a_{\sigma(i)}|^2.$$

Strongly inspired by Brenier's paper (Bull. Inst. Mat. Sin. 2016), we want to derive a more rigorous version of (*) starting from a Brownian point cloud, by applying LDP and Γ-convergence.

Problems arise from the lack of well-posedness of the ODE, that can be attacked within the DiPerna-Lions theory, but only in the a.e. sense.

First step: adding noise to the a_i

By adding noise to the a_i, $a_i \mapsto a_i + \sqrt{\varepsilon} B_t^i$, and viewing the evolution problem modulo permutations (or, equivalently, in the space of empirical measures), the probability density $\varrho_\epsilon(t,x)$ for the point cloud is given in $\mathbb{R}^n$ by

$$\frac{1}{N! \sqrt{2\pi\epsilon t}^{dN}} \sum_{\sigma \in \Sigma_N} \exp\left(-\sum_{i=1}^N \frac{|x_i - a_{\sigma(i)}|^2}{2\varepsilon t} \right).$$

In PDE terms, it is a continuity equation $\partial_t \varrho_\varepsilon + \mathrm{div}(v_\varepsilon \varrho_\varepsilon) = 0$, where the driving vector field v_ε is representable by

$$v_\varepsilon(t,x) := \frac{x - \nabla f_\varepsilon(t,x)}{2t},$$

$$f_\varepsilon(t,x) = \varepsilon t \log\left[\frac{1}{N!} \sum_{\sigma \in \Sigma_N} \exp\left(\frac{\langle x, A^\sigma \rangle}{\varepsilon t} \right) \right].$$

Laplace and large deviation principles

By Laplace's principle, $f_\varepsilon(t,\cdot)$ converge to the convex function

$$f(x) := \max_{\sigma\in\Sigma_N} \langle x, A^\sigma \rangle,$$

so that (at least at differentiability points x of f)

$$\lim_{\varepsilon\to 0} v_\varepsilon(t,x) = \frac{x - \nabla f(x)}{2t} = \frac{1}{2t}\nabla D^2(x), \quad D^2(x) := \frac{1}{2}\min_{\sigma\in\Sigma_N} |x - A^\sigma|^2.$$

Definition. *A family (η_n) of probability measures in X satisfies the LDP with speed (a_n) and rate functional I if*

$$\liminf_{n\to\infty} a_n \log \eta_n(A) \geq -\inf_{x\in A} I(x)$$

for $A \subset X$ open, and

$$\limsup_{n\to\infty} a_n \log \eta_n(C) \leq -\inf_{x\in C} I(x)$$

for $C \subset X$ closed.

Second step: adding noise to the ODE driven by v_ε

Now, by adding noise not only to the lattice, but also the points

$$dX_t^{\varepsilon,\eta} = v_\varepsilon(t, X^{\varepsilon,\eta})dt + \frac{\sqrt{\eta}}{\sqrt{t}} d\tilde{B}_t$$

the Freidlin-Wentzell theorem ensures that, with ϵ fixed, the conditioned laws of $X^{\varepsilon,\eta}$ in $[t_0, t_1] \subset (0, \infty)$ satisfy a LDP principle in $X = C([t_0, t_1]; \mathbb{R}^N)$ with speed η and rate functional

$$I_\varepsilon(z) := \int_{t_0}^{t_1} t|z'(t) - v_\varepsilon(t, z(t))|^2 \, dt$$

set to $+\infty$ if $z \notin H^1([t_0, t_1]; \mathbb{R}^{Nd})$.

Remembering the limit of v_ε, we may take the limit also as $\varepsilon \to 0$.

Convergence to an effective functional

A time-dependent version of the general Γ-convergence result, gives:

Theorem. *The functionals $I_\varepsilon(z)$, even with prescribed boundary conditions,* Γ*-converge in* $X = C([t_0, t_1]; \mathbb{R}^{Nd})$ *to*

$$I(z) = \int_{t_0}^{t_1} t \left| z'(t) - \frac{z(t) - \nabla f(z(t))}{2t} \right|^2 dt.$$

If $z(t)$ does touch singularities of f, minimizers of I satisfy, after an exponential rescaling, exactly the discrete VMA system.

The advantage of this “variational” derivation is that it makes sense regardless of this assumption, so that we may consider minimizers of I as the “true” solutions to the discrete VMA which, as stated, is ill posed.

A 1-dimensional regularity result

Set $d = 1$, let, as usual,

$$f(x) := \max_{\sigma \in \Sigma_N} \langle A^\sigma, x \rangle \qquad x \in \mathbb{R}^N$$

and let $\pi(x)$ the equivalence relation in $\{1, \ldots, N\}$ induced by $i \sim j$ iff $x_i = x_j$.

Theorem. *For any minimizer $z(t)$ of the functional*

$$\int_0^T |z'(t)|^2 + |z(t) - \nabla f(z(t))|^2 \, dt$$

with endpoint conditions $z(0) \in \{P\}^\sigma$, $z(T) \in \{Q\}^\sigma$, there exist

$$0 < t_1 < \cdots < t_k < T$$

such that z is smooth and $\pi(z(t))$ is constant in the intervals (t_i, t_{i+1}).

The d-dimensional case is open.

Thank you for the attention!

Slides available upon request