The Weighted Energy Dissipation principle for gradient flows

Luca Tamanini

Université Paris Ouest Nanterre La Défense

May 5, 2015

Contents

These notes are based on the course held by Giuseppe Savaré in Bonn for the workshop *Gradient flows and entropy methods* (17-20 February 2015) during the Junior Hausdorff Trimester Program on Optimal Transportation.

1 Introduction

1.1 Premise

It is well known that gradient flows in linear or metric spaces can be constructed by studying the limit of the discrete solutions obtained by the so-called Minimizing Movement (MM) scheme. The lectures will present an introduction to another variational method, consisting in a family of minimum problems for suitable integral functionals defined in the space of (absolutely) continuous paths. In Hilbert spaces the method corresponds to an *elliptic regularization* of the differential equation, already envisaged by Lions and Magenes, and it has been studied by many authors in the past, as for instance Ilmanen, De Giorgi, Mielke and Stefanelli, with various applications. After a short and heuristic introduction for finite-dimensional flows, we will discuss the general metric setting, showing the relationships with optimal control problems and Hamilton-Jacobi equations and proving some convergence results, obtained by Savaré in collaboration with Rossi, Segatti and Stefanelli, under various assumptions on the energy functional. We will eventually compare the Weighted Energy Dissipation (WED) approach with the MM one.

1.2 What is known

Given an ambient space X and a function $\varphi : X \to]-\infty, +\infty]$, our aim is to study the gradient flow generated by φ in X. As a first example, let us assume that $X = \mathbb{R}^d$ and $\varphi \in C^1(\mathbb{R}^d)$; then in this case we are interested in solving the Cauchy problem

$$\begin{cases} u'(t) + D\varphi(u(t)) = 0 \\ u(0) = u_0 \end{cases} \tag{1.1}$$

for any given initial condition u_0 and in what follows we will use this first problem as a model to understand more difficult situations. As a second example, let X be a Hilbert space and let φ be convex or also λ-convex; then the differential of φ in (1.1) has to be replaced by its subdifferential and the differential equation has to be restated as a differential inclusion, namely

$$\begin{cases} u'(t) \in -\partial\varphi(u(t)) \\ u(0) = u_0 \end{cases} \tag{1.2}$$

The theory of gradient flows in the Hilbert case goes back to Brezis, Crandall, Komma and Pazy in the '70s. As a third example, let us consider a slightly more difficult situation. It can be obtained by assuming that φ is the sum of a convex function and a C^1 perturbation or it can be expressed as the difference of two convex functions; for the space, let X be Banach. In this framework we have to pay attention to a peculiar aspect that was hidden in the two previous cases: the duality map; indeed, in $\mathbb{R}^d$ and in Hilbert spaces we always identify X and its dual, so that in (1.1) and (1.2) X' does not appear. However, in Banach spaces the subdifferential is a dual object and for this reason (1.2) has to be rephrased as

$$\begin{cases} Ju'(t) \in -\partial\varphi(u(t)) \\ u(0) = u_0 \end{cases}$$

where $J : X \to X'$ is the dual map. In the even more general framework of metric spaces (and this is the setting we are really interested in), (X, d) is complete and $\varphi : X \to]-\infty, +\infty]$ is lower semicontinuous. Several contributions to the theory have come from De Giorgi, Marino, Sacconi and De Giovanni in the '90s, after the study of the W_2-gradient flow in $\mathscr{P}_2(X)$ by Jordan, Kinderlehrer and Otto. In this setting, the main difficulties come from the fact that we do not have a linear structure, so that *a priori* (1.2) does not make sense; we have to replace the differential inclusion by a scalar condition. To this aim we will look at our model case, i.e. $\mathbb{R}^d$, and here it is well known that the differential equation in (1.1) can be equivalently rewritten as

$$\frac{d}{dt}\varphi(u(t)) = -|u'(t)|^2 = -|D\varphi(u(t))|^2 \tag{1.3}$$

because by the chain rule

$$\frac{d}{dt}\varphi(u(t)) = \langle D\varphi u(t), u'(t)\rangle$$

and the right-hand side is equal to $-|u'(t)|^2$ or to $-|D\varphi(u(t))|^2$ if and only if the differential equation in (1.1) holds. In an apparently weaker form, (1.3) can be equivalently restated in turn as

$$\frac{d}{dt}\varphi(u(t)) + \frac{1}{2}\Big(|u'(t)|^2 + |D\varphi(u(t))|^2\Big) \leq 0 \tag{1.4}$$

since the converse inequality always holds: indeed by applying first Cauchy-Schwarz inequality and then Youngs'one we get

$$\frac{d}{dt}\varphi(u(t)) \geq -|D\varphi u(t)||u'(t)| \geq -\frac{1}{2}|u'(t)|^2 - \frac{1}{2}|D\varphi(u(t))|^2$$

However, we will much more interested in the integrated version of (1.4), namely

$$\varphi(u(t)) + \frac{1}{2}\int_0^t \Big(|u'(s)|^2 + |D\varphi(u(s))|^2\Big) ds \leq \varphi(u_0), \quad \forall t \geq 0 \tag{1.5}$$

because we do not need to differentiate φ along trajectories and we only need to give a meaning to $|D\varphi|$ instead of $D\varphi$. This inequality holds almost *verbatim* for Hilbert spaces (we have to replace $|\cdot|$ by the Hilbert norm) and Banach ones, but in the latter case we have to be careful and distinguish between norm and dual norm, i.e. (1.5) reads as

$$\varphi(u(t)) + \frac{1}{2}\int_0^t \Big(\|u'(s)\|_X^2 + \|D\varphi(u(s))\|_{X'}^2\Big) ds \leq \varphi(u_0), \quad \forall t \geq 0$$

Thus, our aim is to generalize (1.5) to metric spaces, but first we need a digression on absolutely continuous curves. As a first step, let us recall that, given an interval $I \subset [0, +\infty[$, a curve $u : I \to X$ is said to be *absolutely continuous* and we write $u \in AC(I, X)$ if there exists a non-negative function $v \in L^1(I)$ such that

$$d(u(s), u(t)) \leq \int_s^t v(r)dr, \quad \forall s, t \in I,\ s < t \tag{1.6}$$

This definition can be slightly modified, since we can ask v to be in $L^p(I)$ for a given $p \in [1, +\infty]$ and in this case the space of curves we get will be denoted by $AC^p(I, X)$, but in what follows we will only focus on the case $p = 1$. The minimal function v for which inequality (1.6) holds is called *metric velocity* of u and the reason is given by the characterization stated in the theorem below.

Theorem 1.1. *In the setting we have described, given a curve $u \in AC(I, X)$, the following are true:*

(i) for $\mathscr{L}^1$-a.e. $t \in I$, the limit

$$|\dot{u}|(t) := \lim_{h\to 0} \frac{d(u(t), u(t+h))}{|h|}$$

exists;

(ii) the map $t \mapsto |\dot{u}|(t)$ belongs to $L^1(I)$, satisfies (1.6) and it is the minimal (in the class of $L^1(I)$ functions) function for which (1.6) holds, namely for every v satisfying (1.6) we have $|\dot{u}|(t) \leq v(t)$ for $\mathscr{L}^1$-a.e. $t \in I$.

It is important to stress that in the case of Hilbert or reflexive Banach spaces, given a curve $u \in AC(I, X)$, we know that u is a.e. differentiable, its derivative can be written in standard way as limit of incremental ratios, namely

$$u'(t) = \lim_{h\to 0} \frac{u(t+h) - u(t)}{h}$$

and this notion of derivative is compatible with the metric one, because

$$|\dot{u}|(t) = \|u'(t)\|_X$$

Moreover, the following integral representation holds

$$u(t) = u(s) + \int_s^t u'(r)dr$$

where the integral has to be meant in the Bochner sense. This is all we need to know about absolutely continuous curves valued in a metric space, because now we can replace $|u'(s)|$ by $|\dot{u}|(s)$ in (1.5). As a next step, we have to give a meaning to $|D\varphi|$; thus, the second digression we provide is about upper gradients and slopes.

For sake of information, the notion of upper gradient was first introduced by Cheeger, even if in a different setting. In order to understand the philosophy behind the forthcoming definition, let us begin by the smooth case and observe that

$$\left|\frac{d}{dt}\varphi(u(t))\right| = |\langle D\varphi(u(t)), u'(t)\rangle| \leq \|D\varphi(u(t))\|_{X'}\|u'(t)\|_X$$

so that we can look at $\|D\varphi\|_{X'}$ as an upper bound on the growth rate of φ along u. What we are going to do in a while is to replace $\|D\varphi\|_{X'}$ by an arbitrary function G without any particular regularity assumption. More precisely, we say that a Borel map $G : X \to [0, +\infty]$ is an *upper gradient* for φ if for every $u \in AC([0,1], X)$ such that

$$\int_0^1 G(u(t))|\dot{u}|(t)dt < +\infty \tag{1.7}$$

it holds

$$|\varphi(u(1)) - \varphi(u(0))| \leq \int_0^1 G(u(t))|\dot{u}|(t)dt \tag{1.8}$$

Coupled with (1.7) and by virtue of Lebesgue differentiation theorem, this condition entails the pointwise bound

$$\left|\frac{d}{dt}\varphi(u(t))\right| \leq G(u(t))|\dot{u}|(t)$$

for $\mathscr{L}^1$-a.e. $t \in [0,1]$, so that we have got exactly what we desired. On the other hand, the *(descending) slope* of φ is defined by

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

In the literature, the usual notation for the descending slope is $|\partial^-\varphi|$, since one can also define the ascending slope $|\partial^+\varphi|$, but in the present discussion we are not interested in the ascending counterpart and therefore no confusion will arise. On the contrary, the notation $|\partial^-\varphi|$ will be used for the lower semicontinuous relaxation of $|\partial\varphi|$, namely

$$\begin{aligned} |\partial^-\varphi|(x) &:= \Gamma - \liminf |\partial\varphi|(x) \\ &= \inf\left\{\liminf_{n\to\infty} |\partial\varphi|(y_n) \,:\, y_n \to x\right\} \end{aligned} \tag{1.9}$$

The notion of slope provides a reasonable upper bound on any upper gradient of φ, because if $\varphi \circ u$ is differentiable at t, then

$$\left|\frac{d}{dt}(\varphi \circ u)(t)\right| \leq |\partial\varphi|(u(t))|\dot{u}|(t)$$

and the set of t's where $\varphi \circ u$ is differentiable is of full measure. For this reason, in the definition of upper gradient we can ask G to satisfy the further condition $G \leq |\partial\varphi|$. Nevertheless, it is not always true that the pointwise inequality above implies the integrated version (1.8), because $|\partial\varphi|$ need not be an upper gradient (indeed (1.7) may fail). Now we have all the required tools to understand the statement of the next definition.

Definition 1.2. *Let (X, d) be a complete metric space, $\varphi : X \to]-\infty, +\infty]$ be lower semicontinuous and $G : X \to [0, +\infty]$ be an upper gradient of φ. Given $u_0 \in D(\varphi)$, where $D(\varphi) := \{u \in X : \varphi(u) < +\infty\}$, a curve $u : [0, +\infty[\to X$ is a* gradient flow of φ (w.r.t. G) starting from u_0 *if $u \in AC^2_{loc}([0, +\infty[, X)$, $u(0) = u_0$ and*

$$\varphi(u(t)) + \frac{1}{2}\int_0^t \Big(|\dot{u}|^2(s) + G^2(u(s))\Big)ds \leq \varphi(u_0), \quad \forall t \geq 0 \tag{1.10}$$

With this notion we can close the quick overview on what is already known about gradient flows and we can now turn our attention to the problem of their existence.

2 The WED method

In the general theory of gradient flows, the result one is looking for is of the following kind, but it is worth stressing that, while in the smooth or Hilbertian case the notion of differential is available (so that there is a privileged choice for the upper gradient), in the metric case no canonical/optimal choice for upper gradients is known yet.

Theorem 2.1. *Let (X, d) be a complete metric space, $\varphi : X \to]-\infty, +\infty]$ be lower semicontinuous and $G : X \to [0, +\infty]$ be an upper gradient of φ. Let us assume that:*

(i) there exist $a, b \geq 0$ and $u_* \in X$ such that, for every $u \in X$,

$$\varphi(u) \geq -a - bd^2(u, u_*)$$

(ii) φ has compact sublevels, i.e. if $(u_n)_n \subset X$ is a sequence such that $\varphi(u_n) \leq C$ for every $n \in \mathbb{N}$, then there exists a convergent subsequence $(u_{n_k})_k$;

(iii) G is lower semicontinuous.

Then, for every $\overline{u} \in D(\varphi)$ there exists a curve $u \in AC^2_{loc}([0, +\infty[, X)$ which is a gradient flow for φ starting from $\overline{u}$.

In the case the relaxed slope $|\partial^-\varphi|$ defined at (1.9) is an upper gradient of φ, we can choose G to be $|\partial^-\varphi|$ (indeed, the relaxed slope is lower semicontinuous by definition) and the MM method tells us that a gradient flow exists, so that Theorem 2.1 holds. On the contrary, the WED method involves an object different from the relaxed slope, which will be defined much later and whose importance will be clarified at the end.

Remark 2.2. A quick comment on the assumptions: the first one is necessary, the second one can actually be relaxed and the third one is the most demanding hypothesis in practical situations. About the possibility to relax assumption (ii), we may find pleasant to work with a topology different from the metric one; in the case of lower semicontinuity w.r.t. a weaker topology, we have to make precise the assumptions on the weak topology. In more details, if we denote by σ the topology weaker than the one induced by the distance, then we have to assume that:

(a) d is σ-lower semicontinuous;

(b) φ has σ-compact sublevels, i.e. if $(u_n)_n \subset X$ is a bounded sequence and $\varphi(u_n) \leq C$ for every $n \in \mathbb{N}$, then there exists a σ-convergent subsequence $(u_{n_k})_k$;

(c) for every $x \in X$ and every σ-neighbourhood U of x, there exist a σ-open neighbourhood Y of x and $\delta > 0$ such that $B(y, \delta) \subset U$ for every $y \in Y$.

The third condition is a compatibility property between σ and d, because by the fact that σ is weaker than the distance topology, for any $x \in X$ and U as before we are always able to find a radius $\delta > 0$ such that $B(x, \delta) \subset U$, but here we are asking something more. For instance, if X is a Banach space and σ is the weak topology, then all the conditions above are satisfied and the same happens if (X, d) is the Wasserstein space $\mathscr{P}_2(\tilde{X})$ built over a metric space $(\tilde{X}, \tilde{d})$ and σ is the narrow topology. However, we will not work under these relaxed assumptions; on the contrary, from now on we will work in the framework of Theorem 2.1. $\Diamond$

In the literature there are many ways to build gradient flows, as for instance the MM scheme, but in this course our interest is mainly focused on the variational WED formulation (differences and analogies between the two approaches will be pointed out in the discussion). As a first step for the description of such method, let $\varepsilon > 0$ be given, let us define

$$\mu_\varepsilon := \frac{e^{-t/\varepsilon}}{\varepsilon} \mathscr{L}^1$$

which is a probability measure on $[0, +\infty[$ and, by means of μ_ε, let us also define an energy-type functional

$$\begin{aligned} I_\varepsilon(u) &:= \int_0^{+\infty} \left(\frac{\varepsilon}{2} |\dot{u}|^2(t) + \varphi(u(t)) \right) d\mu_\varepsilon(t) \\ &= \int_0^{+\infty} e^{-t/\varepsilon} \left(\frac{1}{2} |\dot{u}|^2(t) + \frac{1}{\varepsilon} \varphi(u(t)) \right) dt \end{aligned}$$

As we will see later, this functional is well-defined on $AC^2_{loc}([0, +\infty[, X)$. For the moment, let us briefly sketch the strategy of the proof of Theorem 2.1 that we will adopt in the future.

Strategy 2.3. Let us prove the following facts:

- for any given $\varepsilon > 0$, prove the existence of a minimizer of I_ε among all curves $u \in AC^2_{loc}([0,+\infty[, X)$ with $u(0) = \overline{u}$;
- find a decreasing sequence $(\varepsilon_k)_k$ such that $\varepsilon_k \downarrow 0$ and $u_{\varepsilon_k}(t) \to u(t)$ for every $t \geq 0$ as $k \to \infty$ for a suitable curve u;
- prove that u is a gradient flow for φ w.r.t. G.

Before we start with the proof, let us explain the interest in the WED approach and its history. If X is a Hilbert space and φ is quadratic, then I_ε had already been introduced by Lions and Magenes in the theory of elliptic regularization. The Hilbert case with φ convex has been studied later by Ilmanen in 1994, De Giorgi in 1996 and by Mielke, Ortiz and Stefanelli more recently. Let us see, in the case of smooth functions on $\mathbb{R}^d$, why the WED procedure provides a reasonable approximation technique and why we can talk about elliptic regularization.

Assume that a minimizer u_ε of I_ε exists; then the Euler equation associated to I_ε and satisfied by u_ε is the following

$$-\varepsilon u_\varepsilon''(t) + u_\varepsilon'(t) + D\varphi(u_\varepsilon(t)) = 0 \tag{2.1}$$

which is an elliptic problem and the second order term can be regarded as a perturbation of the classical gradient flow equation. This explains the expression *elliptic regularization* and also shows the approximation procedure lying behind the minimizers of I_ε: indeed, by the lower semicontinuity of $D\varphi$, by letting $\varepsilon \to 0$ we get that (up to pass to a suitable sequence $(\varepsilon_k)_k$) u_ε converges to a function u which is a gradient flow for φ, since the second order perturbation term vanishes. In the metric setting, the lower semicontinuity of the upper gradient is much more delicate: indeed, as we have already said, it is the most demanding assumption of Theorem 2.1. Now let us show why (2.1) is the Euler equation associated to I_ε; the technique we are going to perform is completely standard. Fix $v \in C_c^\infty([0,+\infty[, \mathbb{R}^d)$ and consider the perturbation $u_\varepsilon + \delta v$, where δ is a small parameter; then $I_\varepsilon(u_\varepsilon) \leq I_\varepsilon(u_\varepsilon + \delta v)$ for every δ, so that

$$\frac{\partial I_\varepsilon}{\partial \delta}(u_\varepsilon + \delta v)\bigg|_{\delta=0} = 0$$

and if we compute explicitly this derivative, we get

$$\begin{aligned}
\frac{\partial I_\varepsilon}{\partial \delta}(u_\varepsilon + \delta v)\bigg|_{\delta=0} &= \frac{\partial}{\partial \delta}\int_0^{+\infty}\left(\frac{\varepsilon}{2}|\dot{u}_\varepsilon + \delta \dot{v}|^2 + \varphi(u_\varepsilon + \delta v)\right)d\mu_\varepsilon(t)\bigg|_{\delta=0} \\
&= \int_0^{+\infty}\Big(\varepsilon\langle \dot{u}_\varepsilon(t), \dot{v}(t)\rangle + \langle D\varphi(u_\varepsilon(t)), v(t)\rangle\Big)d\mu_\varepsilon(t) \\
&= \int_0^{+\infty}\langle -\varepsilon u_\varepsilon''(t) + u_\varepsilon'(t) + D\varphi(u_\varepsilon(t)), v(t)\rangle d\mu_\varepsilon(t)
\end{aligned}$$

where the last identity is motivated by the following integration by parts formula, valid for any $w \in C^1([0,+\infty[,\mathbb{R}^d)$,

$$\int_0^{+\infty} \varepsilon\langle w(t), \dot{v}(t)\rangle d\mu_\varepsilon(t) = \int_0^{+\infty} \langle -\varepsilon\dot{w}(t) + w(t), v(t)\rangle d\mu_\varepsilon(t) + \varepsilon\langle w(0), v(0)\rangle$$

and by the fact that we can ask $v(0) = 0$. Pay attention to the fact that in the previous formula, we have to consider also the time derivative of μ_ε. Therefore, by the arbitrariness of v we precisely get (2.1).

2.1 Existence of minimizers

In this subsection we focus our attention on the proof of the existence of minimizers for I_ε (i.e. the first point of the strategy we outlined before), so that ε is kept fix. The procedure we will adopt is the standard direct method of Calculus of Variations.

Strategy 2.4. Let us check the following two properties:

(1) the functionsl I_ε is lower semicontinuous w.r.t. pointwise convergence in $AC^2_{loc}([0,+\infty[,X)$;

(2) if $(u_n)_n \subset AC^2_{loc}([0,+\infty[,X)$ is such that $(u_n(0))_n \subset X$ is bounded and $I_\varepsilon(u_n) \leq C < +\infty$ for every $n \in \mathbb{N}$, then there exist a curve $u \in AC^2_{loc}([0,+\infty[,X)$ and a subsequence $(u_{n_k})_k$ such that $u_{n_k} \to u$ pointwise;

If we are able to perform this strategy, then we will be able to produce the desired minimizer u, which belongs to $AC^2_{loc}([0,+\infty[,X)$. Indeed, as a first remark observe that if we take $\overline{u} \in D(\varphi)$, as it is in Theorem 2.1, and we consider the curve u defined by $u(t) \equiv \overline{u}$ for every $t \geq 0$, then $|\dot{u}| \equiv 0$ and $I_\varepsilon(u) = \varphi(\overline{u}) < +\infty$, so that $I_\varepsilon \not\equiv +\infty$ and this allows us to consider minimizing sequences in the set of curves belonging to $AC^2_{loc}([0,+\infty[,X)$ with initial condition $\overline{u}$. As a second step, once we have chosen such a minimizing sequence, by property (2) we can pass to a pointwise convergent subsequence and by property (1) the limit curve is a minimizer of I_ε in the set $AC^2_{loc}([0,+\infty[,X)$. As one can guess, property (2) follows by hypothesis (ii) in Theorem 2.1; thus if we choose to relax (ii) by considering a weak topology σ, then properties (1) and (2) have to be slightly modified and we have to replace pointwise convergence by the σ-pointwise one.

2.1.1 Lower semicontinuity

The first property we will prove is the lower semicontinuity of I_ε, which is a consequence of a good property of μ_ε: a global Poincaré inequality. Let $w \in AC^2_{loc}([0,+\infty[)$ (when we omit the target, we assume it is $\mathbb{R}$) and observe that in this case $AC^2_{loc}([0,+\infty[) = W^{1,2}_{loc}(0,+\infty)$, so that w has an additional regularity; let us assume also that w satisfies the following conditions

$$\begin{cases} \displaystyle\int_0^{+\infty} |w'(t)|^2 d\mu_\varepsilon(t) < +\infty \\ w(0) = 0 \end{cases} \tag{2.2}$$

Then

$$\frac{1}{4\varepsilon^2}\int_0^{+\infty} |w(t)|^2 d\mu_\varepsilon(t) \leq \int_0^{+\infty} |w'(t)|^2 d\mu_\varepsilon(t) \tag{2.3}$$

and this inequality holds also locally, i.e. on every interval $[0,T[$ for any $T \in [0,+\infty]$. Since w satisfies (2.2), this Poincaré inequality tells us that w is not only locally square-integrable (as by assumption), but also square-integrable *tout court*, i.e. $w \in L^2(0,+\infty)$, so that in a rough way we can say that

$$W^{1,2}_{loc}(0,+\infty) + (2.2) = W^{1,2}(0,+\infty)$$

In addition, its L^2-norm is controlled by the L^2-norm of its distributional derivative. Now let us prove (2.3) and let us start with the following remark: it is sufficient to consider the case $\varepsilon = 1$; indeed, if we have (2.3) for $\varepsilon = 1$, then the general case is deduced as below.

$$\begin{aligned} \varepsilon^2\int_0^{+\infty} |w'(t)|^2 d\mu_\varepsilon(t) &= \varepsilon^2\int_0^{+\infty} |w'(t)|^2 \frac{e^{-t/\varepsilon}}{\varepsilon} dt \\ &= \int_0^{+\infty} e^{-s}\varepsilon^2 |w'(\varepsilon s)|^2 ds \\ &= \int_0^{+\infty} e^{-s}|\tilde{w}'(s)|^2 ds \geq \frac{1}{4}\int_0^{+\infty} e^{-s}|\tilde{w}'(s)|^2 ds \\ &= \frac{1}{4}\int_0^{+\infty} |w(t)|^2 \frac{e^{-t/\varepsilon}}{\varepsilon} dt = \int_0^{+\infty} |w(t)|^2 d\mu_\varepsilon(t) \end{aligned}$$

For the second identity we simply used the change of variable $s = t/\varepsilon$ and in the third one we introduced the time-rescaled curve $\tilde{w}(s) := w(\varepsilon s)$. Thanks to this remark, the forthcoming computations will be simplified. Let us define $f(t) := e^{-t/2}w(t)$, so that

$$w(t) = e^{t/2}f(t), \quad w'(t) = e^{t/2}f'(t) + \frac{1}{2}e^{t/2}f(t)$$

As a consequence, we can observe that, for any $T > 0$,

$$\begin{aligned} \int_0^T |w'(t)|^2 e^{-t} dt &= \int_0^T \left| f'(t) + \frac{1}{2}f(t)\right|^2 dt \\ &= \int_0^T \left(|f'(t)|^2 + \frac{1}{4}|f(t)|^2 + f(t)f'(t)\right) dt \end{aligned}$$

$$
\begin{aligned}
&= \int_0^T \Big(|f'(t)|^2 + \frac{1}{4}|f(t)|^2\Big)dt + \frac{1}{2}f(T)^2 - \frac{1}{2}f(0)^2 \\
&\geq \frac{1}{4}\int_0^T |f(t)|^2 dt = \frac{1}{4}\int_0^T |w(t)|^2 e^{-t} dt
\end{aligned}
$$

where the only remarkable passage is the last one, where we used the fact that $f(0) = 0$ (since $w(0) = 0$ by assumption (2.2)). In this way we have just shown the localized version of the Poincaré inequality; by the arbitrariness of T, also (2.3) follows.

As an immediate consequence, if we choose $\lambda < 1/4\varepsilon^2$, then the functional

$$
w \mapsto \int_0^{+\infty} \Big(|w'(t)|^2 - \lambda|w(t)|^2\Big) d\mu_\varepsilon(t)
$$

is lower semicontinuous w.r.t. the weak topology of $W^{1,2}(0, +\infty)$, because it is continuous in the strong topology and the continuity follows from the fact that, by Poincaré inequality, the functional defines a *non-negative* quadratic form. Since we are in the Hilbert space $W^{1,2}(0, +\infty)$, this implies the convexity of the functional and thus the continuity.

Before the proof of the lower semicontinuity of I_ε, we need to point out two more remarks. The first one is quite technical and allows us to simplify some computations. More explicitly, we would like to restate hypothesis (i) of Theorem 2.1, by substituting u_* with the initial condition $\overline{u}$.

Remark 2.5. As a first step, by triangle and Young inequalities, we have

$$
\begin{aligned}
d^2(u, u_*) &\leq \Big(d(u, \overline{u}) + d(\overline{u}, u_*)\Big)^2 \\
&= d^2(u, \overline{u}) + d^2(\overline{u}, u_*) + 2d(u, \overline{u})d(\overline{u}, u_*) \\
&\leq 2d^2(u, \overline{u}) + 2d^2(\overline{u}, u_*)
\end{aligned}
$$

so that

$$
\varphi(u) \geq -a - bd^2(u, u_*) \geq -a - 2bd^2(u, \overline{u}) - 2bd^2(\overline{u}, u_*)
$$

If we set $A := a + 2bd^2(\overline{u}, u_*)$ and $B := 2b$, then

$$
\varphi(u) \geq -A - Bd^2(u, \overline{u}), \quad \forall u \in X \tag{2.4}
$$

Hence, up to change the constants a, b, the modification is licit. $\Diamond$

As a second remark, let us finally prove that I_ε is well-defined; the very important point is the way we can rewrite the functional properly and for that, given a curve $u \in AC^2_{loc}([0, +\infty[, X)$, we need to introduce the *length function*

$$
L(t) := \int_0^t |\dot{u}|(s) ds
$$

and observe that $L(t) \geq d(u(t), \overline{u})$, $L \in AC_{loc}(0, +\infty)$ and $L'(t) = |\dot{u}|(t)$. With this premise, we can write

$$I_\varepsilon(u) = \underbrace{\int_0^{+\infty} \frac{\varepsilon}{2}|\dot{u}|^2(t)d\mu_\varepsilon(t) - \lambda\int_0^{+\infty}\frac{\varepsilon}{2}L^2(t)d\mu_\varepsilon(t)}_{I_\varepsilon^{(1)}(u)} + \underbrace{\int_0^{+\infty}\left(\frac{\lambda\varepsilon}{2}L^2(t) + \varphi(u(t)) + A\right)d\mu_\varepsilon(t)}_{I_\varepsilon^{(2)}(u)} - A$$

and notice that, on the one hand, $I_\varepsilon^{(1)}(u) \geq 0$ whenever $\lambda < 1/4\varepsilon^2$ by Poincaré inequality applied to L, whereas on the other hand, if we choose $\lambda = 1/8\varepsilon^2$, then

$$\frac{\lambda\varepsilon}{2} = \frac{1}{16\varepsilon} > B$$

for ε small enough and, as a consequence,

$$I_\varepsilon^{(2)}(u) \geq \int_0^{+\infty}\left(Bd^2(u(t), \overline{u}) + \varphi(u(t)) + A\right)d\mu_\varepsilon(t) \geq 0 \tag{2.5}$$

where the second inequality is due to (2.4). Therefore, $I_\varepsilon(u) \geq -A$, whence the good definition. It is worth stressing that with the choice $\lambda = 1/8\varepsilon^2$ the lower estimate for $I_\varepsilon^{(1)}$ can actually be refined, because

$$I_\varepsilon^{(1)}(u) \geq \int_0^{+\infty}\frac{\varepsilon}{4}|\dot{u}|^2(t)d\mu_\varepsilon(t) \tag{2.6}$$

Now we can finally prove that I_ε is lower semicontinuous w.r.t. pointwise convergence. Let $(u_n)_n \subset AC^2_{loc}([0, +\infty[, X)$ be a sequence pointwise convergent to a certain curve u and, without loss of generality, let us assume that $I_\varepsilon(u_n) \leq C < +\infty$ for every $n \in \mathbb{N}$. By (2.6), this entails that

$$\int_0^{+\infty}|\dot{u}_n|^2(t)d\mu_\varepsilon(t) \leq C' < +\infty \tag{2.7}$$

for a suitable constant C', whence $(L_n')_n \subset L^2_{loc}(0, +\infty)$ is a bounded sequence (here L_n is the length function associated to u_n). This implies that, up to extract a suitable subsequence, $L_n' \rightharpoonup v$ in $L^2_{loc}(0, +\infty)$, where $v \in L^2_{loc}(0, +\infty)$ is a suitable limit function and $L_n \to \tilde{L}$ pointwise, where

$$\tilde{L}(t) := \int_0^t v(s)ds$$

because

$$L_n(t) = \int_0^t L_n'(s)ds \to \int_0^t v(s)ds = \tilde{L}(t)$$

We can also observe that, by the fact that $L_n' = |\dot{u}_n|$,

$$d(u_n(t), u_n(s)) \le \int_s^t L_n'(r)dr$$

and passing to the limit as $n \to \infty$ we obtain

$$d(u(t), u(s)) \le \int_s^t v(r)dr$$

This inequality shows that the metric derivative of u actually exists, $u \in AC^2_{loc}([0, +\infty[, X)$ and $|\dot{u}| \le v$ by Theorem 1.1. With this premise we can estimate the inferior limit of $I_\varepsilon(u_n)$ as follows.

$$\begin{aligned}
\liminf_{n\to\infty} I_\varepsilon(u_n) &\ge \liminf_{n\to\infty} I_\varepsilon^{(1)}(u_n) + \liminf_{n\to\infty} I_\varepsilon^{(2)}(u_n) - A \\
&\ge \int_0^{+\infty} \left(\frac{\varepsilon}{2}v^2(t) - \frac{\lambda\varepsilon}{2}\tilde{L}^2(t)\right) d\mu_\varepsilon(t) \\
&\quad + \int_0^{+\infty} \left(\frac{\lambda\varepsilon}{2}\tilde{L}^2(t) + \varphi(u(t)) + A\right) d\mu_\varepsilon(t) - A \\
&= \int_0^{+\infty} \left(\frac{\varepsilon}{2}v^2(t) + \varphi(u(t))\right) d\mu_\varepsilon(t) \\
&\ge \int_0^{+\infty} \left(\frac{\varepsilon}{2}|\dot{u}|^2(t) + \varphi(u(t))\right) d\mu_\varepsilon(t)
\end{aligned}$$

At the second inequality, in $I_\varepsilon^{(1)}$ we simply used the fact that $L_n' \to v$ in $L^2_{loc}(0, +\infty)$ for the first summand in the integral and $L_n \to \tilde{L}$ pointwise together with Lebesgue dominated convergence theorem for the second one; in $I_\varepsilon^{(2)}$ we used Fatou's lemma and this is licit because in (2.5) we have shown that the integrand is non-negative, hence bounded from below. The last inequality simply follows from the fact that $|\dot{u}| \le v$.

This concludes the proof of the lower semicontinuity of I_ε.

2.1.2 Compactness of the sublevels

It remains to prove property (2) of Strategy 2.4 and this will be performed via an Ascoli-Arzelà procedure, which relies on two facts: the equicontinuity of $(u_n)_n$ in the topology of $AC^2_{loc}([0, +\infty[, X)$ and the "pointwise" inclusion of the sequence in a compact set, i.e. $(u_n(t))_n \subset K$ with K compactly included in X for any time t. Our only assumptions are the following: $(u_n(0))_n$ is bounded and $I_\varepsilon(u_n) \le C < +\infty$ for any $n \in \mathbb{N}$. The first assumptions seems a bit relaxed in the present context, because the minimizing sequences for I_ε we are working with have the same initial condition, but in the last part of the course (when $\varepsilon \downarrow 0$) this degree of generality will be fundamental. For the second assumption, as already noticed above, it implies (2.7), which in turn entails

$$\int_0^T |\dot{u}_n|^2(t)dt \le C_T \tag{2.8}$$

for any $T \geq 0$ (pay attention to the fact that we are no longer integrating w.r.t. μ_ε; this is the reason why (2.8) can not be independent of T). Indeed

$$\int_0^T |\dot{u}_n|^2(t)dt \leq \varepsilon e^{T/\varepsilon} \int_0^T |\dot{u}_n|^2 d\mu_\varepsilon(t) \leq \varepsilon e^{T/\varepsilon} C' =: C_T$$

As a consequence, by (2.8)

$$\begin{aligned} d(u_n(t), u_n(s)) &\leq \int_s^t |\dot{u}_n|(r)dr \leq \sqrt{|t-s|}\left(\int_s^t |\dot{u}_n|^2(r)dr\right)^{1/2} \\ &\leq C_T^{1/2}\sqrt{|t-s|} \end{aligned} \tag{2.9}$$

for any $s, t \in [0, T]$ and this means that the sequence $(u_n)_n$ is uniformly Hölder continuous with parameter $\alpha = 1/2$ on any bounded interval $[0, T]$, whence the desired uniform continuity in $AC^2_{loc}([0, +\infty[, X)$. Another consequence of (2.8) and of the fact that $I_\varepsilon(u_n) \leq C$ for any $n \in \mathbb{N}$ is that, for any $T \geq 0$,

$$\int_0^T \varphi(u_n(t))dt \leq C'_T$$

The inequality is still true even if we consider only the positive part of φ, namely

$$\int_0^T \Big(\varphi(u_n(t))\Big)^+ dt \leq C''_T \tag{2.10}$$

because by (2.4) the negative part is uniformly bounded. At this point, in the smooth setting the Aubin-Lions-Simon lemma would be sufficient to get the conclusion, but in a purely metric setting we do not have it. For this reason we have to perform an Ascoli-Arzelà procedure and the most difficult part is to find a suitable countable dense set $D \subset [0, T]$ such that $(u_n(t))_n$ is convergent (up to extract a suitable subsequence) for every $t \in D$; as we are going to see soon, the structure of the set D is induced by (2.10). Fix a countable base J of open intervals of $]0, T[$, that is J is made up by intervals of the form $]a_k, b_k[\subset]0, T[$ with $k \in \mathbb{N}$. Secondly, observe that (2.10) holds also on the smaller interval $]a_k, b_k[$ and then, by Fatou's lemma, we get

$$\int_{a_k}^{b_k} \liminf_{n\to\infty} \Big(\varphi(u_n(t))\Big)^+ dt \leq C''_T$$

Since the Lebesgue measure of $]a_k, b_k[$ is strictly positive, there exists $t_k \in]a_k, b_k[$ such that

$$\liminf_{n\to\infty} \Big(\varphi(u_n(t_k))\Big)^+ < +\infty$$

Hence, there exists a suitable subsequence $(n_h)_h \subset \mathbb{N}$ such that $(\varphi(u_{n_h}(t_k)))_h$ is bounded; as a consequence, by hypothesis (ii) of Theorem 2.1 we can say that, up to extract a further subsequence, $(u_{n_h}(t_k))_h$ is a convergent sequence in $[0, +\infty[$, whose limit will be denoted by $w(t_k)$. Now the argument goes on in a standard way, because by induction and by a Cantor diagonal argument we are able to find a sequence of indices, still denoted by $(n_h)_h \subset \mathbb{N}$, and a countable family of points $D = \{t_k\}_{k\in\mathbb{N}}$ with $t_k \in]a_k, b_k[$ such that

$$\lim_{h\to+\infty} u_{n_h}(t_k) = w(t_k), \quad \forall k \in \mathbb{N} \tag{2.11}$$

Let us stress that the existence of the final subsequence $(u_{n_h})_h$ with the desired property is a consequence of the fact that J is countable, which grants the validity of Cantor diagonal argument. In addition, D is dense in $[0, T]$ because J has been chosen as a basis of open subsets of $]0, T[$. Now we need only to extend w from D to $[0, T]$ in a continuous way to get the conclusion; to this aim, observe that by the uniform Hölder continuity of $(u_{n_h})_h$ we have

$$d(w(t_k), w(t_j)) \le \liminf_{h\to\infty} d(u_{n_h}(t_k), u_{n_h}(t_j)) \le C_T^{1/2}\sqrt{|t_k - t_j|}$$

so that we can actually extend w from D to $[0, T]$ in Hölder continuous way and (2.11) becomes

$$\lim_{h\to\infty} u_{n_h}(t) = w(t), \quad \forall t \ge 0$$

The function w we have built satisfies the thesis only on $[0, T]$, but again by induction on $[0, nT]$ and by a Cantor diagonal argument we can extend w to $[0, +\infty[$ and get the conclusion. In the case we are working with a topology σ weaker than the metric one (see Remark 2.2), the compatibility assumption (c) is needed to adjust this Ascoli-Arzelà argument.

2.2 A priori estimates

Thus, for any $\varepsilon > 0$ we know that a minimizer u_ε of I_ε exists. As a next step, we have to prove that, as $\varepsilon \downarrow 0$, the minimizers u_ε are good approximated solutions of the gradient flow equation and for this purpose good *a priori* estimates are needed.

2.2.1 The inner variation estimate

The first one we are going to describe is the so-called *inner variation estimate*, which tells us that the map

$$t \mapsto \varphi(u_\varepsilon(t)) - \frac{\varepsilon}{2}|\dot{u}_\varepsilon|^2(t) \tag{2.12}$$

belongs to $W^{1,1}_{loc}(0, +\infty)$ and, moreover, it satisfies the following equation

$$\frac{d}{dt}\Big(\varphi(u_\varepsilon) - \frac{\varepsilon}{2}|\dot{u}_\varepsilon|^2\Big) + |\dot{u}_\varepsilon|^2 = 0 \tag{2.13}$$

where the time derivative is meant in the distributional sense. In the smooth case, i.e. with $X = \mathbb{R}^d$ and $\varphi \in C^1(\mathbb{R}^d)$, if we multiply the Euler equation (2.1) by $u'(t)$, then it is straightforward to get

$$-\frac{\varepsilon}{2}\frac{d}{dt}|u'_\varepsilon|^2 + |u'_\varepsilon|^2 + \frac{d}{dt}\varphi(u_\varepsilon) = 0$$

so that (2.13) is completely coherent with the smooth case. However, as this equation is a byproduct of the Euler's one, the information it encodes is less; indeed, as we are going to see in a while, the inner variation allows us to consider only time variations.

In order to get (2.13), let $\{R_\tau\}_{\tau\geq 0}$ be a smooth family of diffeomorphisms of $]0,+\infty[$ such that $R_0 = \mathrm{Id}$, e.g.

$$R_\tau(t) := t + \tau\xi(t)$$

with $\xi \in C_c^\infty(0,+\infty)$, and define $\tilde{u}_\tau(s) := u_\varepsilon(R_\tau^{-1}(s))$. If we look now at $I_\varepsilon(\tilde{u}_\tau)$, by a change of variable $s = R_\tau(t)$ we can observe that

$$I_\varepsilon(\tilde{u}_\tau) = \int_0^{+\infty} \frac{e^{-R_\tau(t)/\varepsilon}}{\varepsilon}\Big(\frac{\varepsilon}{2}\frac{|\dot{u}_\varepsilon|^2(t)}{R'_\tau(t)} + R'_\tau(t)\varphi(u_\varepsilon(t))\Big)dt \tag{2.14}$$

and this shows, as just anticipated, that perturbations happen only in time. In addition, as $\tilde{u}_\tau = u_\varepsilon$ for $\tau = 0$ and u_ε is a minimizer of I_ε, then we have

$$\frac{d}{d\tau}I_\varepsilon(\tilde{u}_\varepsilon)\Big|_{\tau=0} = 0$$

and if we compute explicitly this derivative, then (2.13) will follow immediately. The only things we need are identity (2.14), integration by parts formula and the following relations.

$$R'_\tau(t) = 1 + \tau\xi'(t), \quad \partial_\tau R'_\tau(t) = \xi'(t)$$

By means of these tools, the following chain of identities is easy to check

$$\begin{aligned}
\frac{d}{d\tau}I_\varepsilon(\tilde{u}_\varepsilon)\Big|_{\tau=0} &= \int_0^{+\infty}\frac{e^{-t/\varepsilon}}{\varepsilon}\Big[\Big(-\frac{\varepsilon}{2}|\dot{u}_\varepsilon|^2(t)\xi'(t) + \xi'(t)\varphi(u_\varepsilon(t))\Big) \\
&\quad + \Big(-\frac{e^{-t/\varepsilon}\xi(t)}{\varepsilon^2}\Big(\frac{\varepsilon}{2}|\dot{u}_\varepsilon|^2(t) + \varphi(u_\varepsilon(t))\Big)\Big)\Big]dt \\
&= \int_0^{+\infty}\Big(\varphi(u_\varepsilon(t)) - \frac{\varepsilon}{2}|\dot{u}_\varepsilon|^2(t)\Big)\Big(\xi'(t) - \frac{\xi(t)}{\varepsilon}\Big)d\mu_\varepsilon(t) \\
&\quad - \int_0^{+\infty}\xi(t)|\dot{u}_\varepsilon|^2(t)d\mu_\varepsilon(t) \\
&= \int_0^{+\infty}\Big[-\frac{d}{dt}\Big(\varphi(u_\varepsilon(t)) - \frac{\varepsilon}{2}|\dot{u}_\varepsilon|^2(t)\Big) - |\dot{u}_\varepsilon|^2(t)\Big]\xi(t)d\mu_\varepsilon(t)
\end{aligned}$$

and, by the arbitrariness of ξ, (2.13) is proved.

2.2.2 The dynamic programming principle

After the inner variation one, more refined estimates are needed, because we are not able to pass to the limit as $\varepsilon \downarrow 0$ yet. The main idea is to interpret the

WED problem as a (very particular) optimal control problem and to this aim we have to think to the initial condition $\overline{u} \in D(\varphi)$ as not being fixed; in this way, we can introduce the following function

$$V_\varepsilon(x) := \min\{I_\varepsilon(u) \,:\, u \in AC^2_{loc}([0,+\infty[, X),\, u(0) = x\}$$

and for people at ease with optimal control theory its meaning is evident: it is the *value function* of an optimal control problem, more precisely an *infinite horizon problem*; indeed, using the velocity of the curve u as parameter control, it is immediate to see that in the smooth case V_ε can be rewritten as

$$V_\varepsilon(x) = \min_{v \in L^2_{loc}(0,+\infty)} \left\{ \int_0^{+\infty} e^{-t/\varepsilon} \mathscr{L}(u(t;x,v),v)dt \right\}$$

where $u(t;x,v)$ is a solution to the following Cauchy problem

$$\begin{cases} u'(t) = v(t) \\ u(0) = x \end{cases}$$

and the Lagrangian $\mathscr{L}$ is given by

$$\mathscr{L}(x,v) := \frac{|v|^2}{2} + \frac{1}{\varepsilon}\varphi(x) \tag{2.15}$$

Still working in the smooth case, let us explain why an approach tailored by optimal control theory will give us useful tools to prove Theorem 2.1.

- By the dynamic programming principle, we know that

  $$V_\varepsilon(x) = \min_{v \in L^2_{loc}(0,+\infty)} \left\{ \int_0^{T} e^{-t/\varepsilon} \mathscr{L}(u(t;x,v),v)dt + V_\varepsilon(u(T;x,v))e^{-T/\varepsilon} \right\}$$

  where the term $V_\varepsilon(u(T;x,v))e^{-T/\varepsilon}$ is usually called *discounted cost.*

- The value function V_ε satisfies the static Hamilton-Jacobi equation

  $$\frac{1}{\varepsilon}V_\varepsilon(x) + \mathscr{H}(x, DV_\varepsilon(x)) = 0 \tag{2.16}$$

  where $\mathscr{H}$ is the Hamiltonian, i.e. the conjugate of $\mathscr{L}$ w.r.t. the second variable; in more explicit terms,

  $$\mathscr{H}(x,p) := \sup_v\{\langle p, v\rangle - \mathscr{L}(x,v)\} = \mathscr{L}^*(x,-p) \tag{2.17}$$

  Since we are working with the Lagrangian $\mathscr{L}$ given by (2.15), the Hamiltonian can be explicitly computed and we obtain

  $$\mathscr{H}(x,p) = \frac{|p|^2}{2} - \frac{1}{\varepsilon}\varphi(x)$$

  so that the static Hamilton-Jacobi equation (2.16) can be rewritten as

  $$V_\varepsilon(x) + \frac{\varepsilon}{2}|DV_\varepsilon(x)|^2 = \varphi(x)$$

- Combining the dynamic programming principle and the static Hamilton-Jacobi equation, we get that the optimal v_* providing

$$\mathscr{H}(x, -DV_\varepsilon(x)) = \langle -DV_\varepsilon(x), v_* \rangle - \mathscr{L}(x, v_*)$$

(that is to say, attaining the supremum in (2.17) for $p = -DV_\varepsilon(x)$) is exactly the initial velocity of the optimal curve u_ε, i.e. $u'_\varepsilon(0)$. As on the other hand, by Pontryagin's optimality conditions, $v_* = \partial_p \mathscr{H}(x, -DV_\varepsilon(x))$, we deduce that the minimizer u_ε satisfies the following equation

$$u'_\varepsilon(t) = \partial_p \mathscr{H}(u_\varepsilon(t), -DV_\varepsilon(u_\varepsilon(t)))$$

Since we are working with the Lagrangian $\mathscr{L}$ given by (2.15), this equation turns out to be

$$u'_\varepsilon(t) = -DV_\varepsilon(u_\varepsilon(t)) \tag{2.18}$$

and this means that u_ε is a gradient flow of V_ε.

What is curious in the procedure we have just pointed out is the fact that at the end we have produced a gradient flow of V_ε instead of φ. Thus, we have to prove that V_ε converges to φ as $\varepsilon \downarrow 0$ in order to pass to the limit in (2.18) and we will directly do that in the metric case. Moreover, thanks to the static Hamilton-Jacobi equation (2.16) and the gradient flow equation (2.18), we can interpret the inner variation estimate in terms of the value function V_ε, because by these identities we have

$$\varphi(u_\varepsilon) - \frac{\varepsilon}{2}|u'_\varepsilon|^2 = \varphi(u_\varepsilon) - \frac{\varepsilon}{2}|DV_\varepsilon(u_\varepsilon)|^2 = \varphi(u_\varepsilon) - (\varphi - V_\varepsilon)(u_\varepsilon) = V_\varepsilon(u_\varepsilon)$$

whence (2.13) becomes

$$\frac{d}{dt}V_\varepsilon(u_\varepsilon) + |u'_\varepsilon|^2 = 0$$

This identity and the previous one hold also in the metric case, but the proof is much more sophisticated and will be treated in the final part of these notes. For the moment, remark that, again by (2.16) and (2.18), the equality above can be further rewritten as follows

$$\begin{aligned}\frac{d}{dt}V_\varepsilon(u_\varepsilon) &= -|u'_\varepsilon|^2 = -\frac{1}{2}|u'_\varepsilon|^2 - \frac{1}{2}|DV_\varepsilon(u_\varepsilon)|^2 \\ &= -\frac{1}{2}|u'_\varepsilon|^2 - \frac{\varphi - V_\varepsilon}{\varepsilon}(u_\varepsilon)\end{aligned}$$

and in a metric setting it reads as

$$\frac{d}{dt}V_\varepsilon(u_\varepsilon) = -\frac{1}{2}|\dot{u}_\varepsilon|^2 - \frac{\varphi - V_\varepsilon}{\varepsilon}(u_\varepsilon) \tag{2.19}$$

Remark 2.6. The function $t \mapsto V_\varepsilon(u_\varepsilon(t))$ is non-increasing, because of the very definition of V_ε. Take indeed the constant curve $w(t) = x$, which is an admissible competitor in the definition of V_ε, and observe that $\dot{w} = 0$, so that $V_\varepsilon(x) \leq \varphi(x)$; as a consequence, the right-hand side in (2.19) is non-positive. $\Diamond$

If we integrate (2.19) on $[0, T]$, then we get

$$V_\varepsilon(u_\varepsilon(T)) + \frac{1}{2}\int_0^T \left(|\dot{u}_\varepsilon|^2(t) + 2\frac{\varphi(u_\varepsilon(t) - V_\varepsilon(u_\varepsilon(t))}{\varepsilon}\right)dt = V_\varepsilon(\overline{u}) \qquad (2.20)$$

and the advantage of the computation is due to the fact that this condition resembles a lot De Giorgi's metric formulation of gradient flows (1.10), so that in our metric setting it will be sufficient to study it. If we are able to prove that $V_\varepsilon(\overline{u}) \to \varphi(\overline{u})$, $V_\varepsilon(u_\varepsilon(T)) \to \varphi(u_\varepsilon(T))$ and

$$\frac{1}{2}\int_0^T \left(|\dot{u}_\varepsilon|^2(t) + 2\frac{\varphi(u_\varepsilon(t) - V_\varepsilon(u_\varepsilon(t))}{\varepsilon}\right)dt \to \frac{1}{2}\int_0^T \left(|\dot{u}|^2(t) + G^2(u(t))\right)dt$$

where G is an upper gradient of φ, all the limits being understood in the sense of Γ-convergence, then at the limit we obtain (1.10), because we can not expect equality (2.20) to be preserved under Γ-convergence, and thus the proof that u is a gradient flow of φ in the sense of Definition 1.2. To estimate $V_\varepsilon(\overline{u})$ and $V_\varepsilon(u_\varepsilon(T))$ we only need lower semicontinuity, which is not particularly hard to prove, and the same is true for $|\dot{u}_\varepsilon|^2$ in the integral; on the contrary, it will be much more harder to show the lower semicontinuity of $(\varphi - V_\varepsilon)/\varepsilon$.

Remark 2.7. Observe that there is a strong analogy between the value function V_ε and the Moreau-Yosida approximation coming into play in the MM scheme. Let us recall that the Moreau-Yosida approximation is given by

$$Y_t(x) := \inf_{y \in X}\left\{\varphi(y) + \frac{d^2(x,y)}{2t}\right\} \qquad (2.21)$$

and solves the following (dynamic) Hamilton-Jacobi equation

$$\begin{cases} \partial_t Y_t + \frac{1}{2}|DY_t|^2 = 0 \\ Y_0 = \varphi \end{cases}$$

On the other hand, as we have already seen, the value function V_ε also solves an Hamilton-Jacobi equation, even if static. $\Diamond$

Thus, once we have (2.20) it is clear what we have to do, but how can we deduce (2.20) in a purely metric setting? Indeed the tools we used, i.e. (2.16) and (2.18) are no longer available. As a first step, let us point out two important properties of the value function V_ε:

(i) V_ε is lower semicontinuous;

(ii) V_ε is non-increasing monotone in ε, that is if $\varepsilon_1 < \varepsilon_2$, then $V_{\varepsilon_1}(x) \geq V_{\varepsilon_2}(x)$ for every $x \in X$.

The first property is a straightforward consequence of the lower semicontinuity of I_ε, because if we take a sequence $(x_n)_n \subset X$ such that $x_n \to x$ and for any

$n \in \mathbb{N}$ we choose an optimal curve $u_n \in AC^2_{loc}([0,+\infty[,X)$ with initial condition x_n, which means that

$$V_\varepsilon(x_n) = I_\varepsilon(u_n)$$

then, by what we have already proved, we can find a subsequence u_{n_k} pointwise converging to a suitable curve u such that u is a minimizer of I_ε and $u(0) = x$. This means that

$$V_\varepsilon(x) = I_\varepsilon(u) \leq \liminf_{k\to\infty} I_\varepsilon(u_{n_k}) = \liminf_{k\to\infty} V_\varepsilon(x_{n_k})$$

For the second property, just observe that via a change of variable V_ε can be equivalently rewritten as

$$V_\varepsilon(x) = \min_{\substack{w\in AC^2_{loc}([0,+\infty[,X)\\ w(0)=x}} \int_0^{+\infty} e^{-t}\left(\frac{1}{2\varepsilon}|\dot{w}|^2(t) + \varphi(w(t))\right)dt$$

and the optimal curve w attaining the minimum is $w(t) = u_\varepsilon(\varepsilon t)$, that is any ε-rescaled minimizer of I_ε.

After this premise, let us deduce the dynamic programming principle in the metric case. Let $x \in X$ be the initial condition for our minimization problem and let u be a curve in $AC^2_{loc}([0,+\infty[,X)$ starting from x. Then for any $T \geq 0$ we can observe that

$$\begin{aligned}
I_\varepsilon(u) &= \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\left(\frac{\varepsilon}{2}|\dot{u}|^2(t) + \varphi(u(t))\right)dt \\
&\quad + \int_T^{+\infty} \frac{e^{-t/\varepsilon}}{\varepsilon}\left(\frac{\varepsilon}{2}|\dot{u}|^2(t) + \varphi(u(t))\right)dt \\
&= \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\left(\frac{\varepsilon}{2}|\dot{u}|^2(t) + \varphi(u(t))\right)dt \\
&\quad + \int_0^{+\infty} \frac{e^{-(t+T)/\varepsilon}}{\varepsilon}\left(\frac{\varepsilon}{2}|\dot{u}|^2(t+T) + \varphi(u(t+T))\right)dt \\
&= \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\left(\frac{\varepsilon}{2}|\dot{u}|^2(t) + \varphi(u(t))\right)dt \\
&\quad + e^{-T/\varepsilon}\int_0^{+\infty} \frac{e^{-t/\varepsilon}}{\varepsilon}\left(\frac{\varepsilon}{2}|\dot{u}|^2(t+T) + \varphi(u(t+T))\right)dt \\
&= \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\left(\frac{\varepsilon}{2}|\dot{u}|^2(t) + \varphi(u(t))\right)dt + e^{-T/\varepsilon} I_\varepsilon(\tilde{u})
\end{aligned}$$

where $\tilde{u}(t) = u(t+T)$. Hence, if we first minimize I_ε at the left-hand side among all curves starting from x and then we minimize I_ε at the right-hand side among all curves starting from $u(T)$, we immediately deduce that

$$V_\varepsilon(x) \leq \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon}\left(\frac{\varepsilon}{2}|\dot{u}|^2(t) + \varphi(u(t))\right)dt + e^{-T/\varepsilon} V_\varepsilon(u(T)) \tag{2.22}$$

On the other hand, if u_ε is a minimizer of I_ε relative to the initial condition x, then the shift $\tilde{u}_\varepsilon(t) := u_\varepsilon(t+T)$ is optimal for I_ε with initial condition $u_\varepsilon(T)$. In fact, if this were not the case, it would be sufficient to glue u_ε on $[0,T]$ with any optimal curve starting from $u_\varepsilon(T)$ to get a strictly better curve for I_ε with initial condition x and this is impossible by the optimality of u_ε. Hence, for minimizers of I_ε starting at x equality actually holds in (2.22), so that the dynamic programming principle is true also in the metric case and reads as

$$V_\varepsilon(x) = \min_{\substack{u \in AC^2_{loc}([0,+\infty[,X) \\ u(0)=x}} \left\{ \int_0^T \frac{e^{-t/\varepsilon}}{\varepsilon} \left(\frac{\varepsilon}{2}|\dot{u}|^2(t) + \varphi(u(t)) \right) dt + e^{-T/\varepsilon} V_\varepsilon(u(T)) \right\}$$

Hence, the strategy described by Bardi and Capuzzo Dolcetta in *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations* for the finite-dimensional case perfectly fits to the metric setting too and can be summarized as follows: the minimum $V_\varepsilon(x)$ is achieved if and only if

(1) let the system evolve for a finite time T along an arbitrary path;

(2) pay the corresponding cost, that is the integral part at the right-hand side above;

(3) pay in the best possible way what remains to pay, that is $e^{-T/\varepsilon} V_\varepsilon(u(T))$;

(4) minimize over all possible paths.

As a next step, let us investigate the regularity of V_ε along absolutely continuous curves where φ is integrable, that is suppose that, for given $0 \le a < b$, $w \in AC^2_{loc}(]a,b[,X)$ and $\varphi \circ w \in L^1_{loc}(a,b)$. Then the map $t \mapsto V_\varepsilon(w(t))$ belongs to $AC_{loc}(]a,b[)$ and

$$\left| \frac{d}{dt} V_\varepsilon(w(t)) \right| \le |\dot{w}|(t) \sqrt{\frac{\varphi(w(t)) - V_\varepsilon(w(t))}{\varepsilon}} \tag{2.23}$$

so that $\sqrt{(\varphi - V_\varepsilon)/\varepsilon}$ can be regarded as an upper gradient for V_ε. We will omit the details of the proof, but let us say that this fact relies on the dynamic programming principle, because if we apply it with $x = w(s)$ and $T = t - s$, where $w \in AC^2_{loc}([0,+\infty[,X)$ is a given curve, then by a change of variable we obtain

$$\begin{aligned} V_\varepsilon(w(s)) - V_\varepsilon(w(t)) \quad \le \quad & \int_s^t \left(\frac{\varepsilon}{2}|\dot{w}|^2(r) + \varphi(w(r)) \right) \frac{e^{-(r-s)/\varepsilon}}{\varepsilon} dr \\ & + \left(e^{-(t-s)/\varepsilon} - 1 \right) V_\varepsilon(w(t)) \end{aligned}$$

In a similar fashion, by reversing the time interval T we are able to estimate the difference $V_\varepsilon(w(t)) - V_\varepsilon(w(s))$ too, so that we infer that the map $t \mapsto V_\varepsilon(w(t))$ belongs to $AC_{loc}(]a,b[)$. As a particular case, let us take $w = u_\varepsilon$, minimizer of

I_ε with initial datum x, and observe that the inequality above turns out to be an identity, so that if we multiply both sides by $e^{-s/\varepsilon}$ we get

$$e^{-s/\varepsilon}V_\varepsilon(u_\varepsilon(s)) = \int_s^t \left(\frac{\varepsilon}{2}|\dot u_\varepsilon|^2(r) + \varphi(u_\varepsilon(r))\right)\frac{e^{-r/\varepsilon}}{\varepsilon}dr + e^{-t/\varepsilon}V_\varepsilon(u_\varepsilon(t))$$

which can be rewritten as

$$\frac{V_\varepsilon(u_\varepsilon(s))e^{-s/\varepsilon} - V_\varepsilon(u_\varepsilon(t))e^{-t/\varepsilon}}{t-s} = \frac{1}{t-s}\int_s^t \left(\frac{\varepsilon}{2}|\dot u_\varepsilon|^2(r) + \varphi(u_\varepsilon(r))\right)\frac{e^{-r/\varepsilon}}{\varepsilon}dr$$

As $V_\varepsilon \circ u_\varepsilon$ is absolutely continuous, we can choose s to be a Lebesgue point for $|\dot u_\varepsilon|^2$ and $\varphi(u_\varepsilon)$; this allows us to pass to the limit as $t \to s$ in the identity above, whence

$$-\frac{d}{dt}\Big(V_\varepsilon(u_\varepsilon(t))e^{-t/\varepsilon}\Big)\Big|_{t=s} = \left(\frac{\varepsilon}{2}|\dot u_\varepsilon|^2(s) + \varphi(u_\varepsilon(s))\right)\frac{e^{-s/\varepsilon}}{\varepsilon}$$

By developing the derivative at the left-hand side we finally get

$$-\frac{d}{dt}V_\varepsilon(u_\varepsilon(t))\Big|_{t=s} = \frac{1}{2}|\dot u_\varepsilon|^2(s) + \frac{\varphi(u_\varepsilon(s)) - V_\varepsilon(u_\varepsilon(s))}{\varepsilon}$$

and so we have proved (2.19). Hence, even without the static Hamilton-Jacobi equation (2.16) and the gradient flow equation (2.18) we have been able to show that minimizers of I_ε are gradient flow of V_ε w.r.t. the upper gradient G_ε defined by

$$\frac{1}{2}G_\varepsilon^2(u) := \frac{\varphi(u) - V_\varepsilon(u)}{\varepsilon}$$

This upper gradient is not the exact pointwise slope of V_ε, but nevertheless it provides good controls when we pass to the limit as $\varepsilon \downarrow 0$; for this reason, G_ε can be thought as an approximated slope.

2.3 Construction of a gradient flow

From now on we have to look both at (2.20), or equivalently at

$$V_\varepsilon(u_\varepsilon(T)) + \frac{1}{2}\int_0^T \Big(|\dot u_\varepsilon|^2(t) + G_\varepsilon^2(u_\varepsilon(t))\Big)dt = V_\varepsilon(\overline{u}) \tag{2.24}$$

and at our target, namely the gradient flow inequality (1.10). In order to pass to the limit as $\varepsilon \downarrow 0$, we shall prove the following three facts:

(a) prove that

$$\int_0^T |\dot u_\varepsilon|^2(t)dt \le C_T, \quad \int_0^T \varphi(u_\varepsilon(t))dt \le C'_T$$

for suitable constants C_T, C'_T independent of ε; thanks to these uniform bounds, it will be possible to find a sequence $(\varepsilon_n)_n$ converging to 0 such that $u_{\varepsilon_n} \to u$ pointwise for a suitable limit function u and $|\dot{u}_{\varepsilon_n}| \rightharpoonup v$ in $L^2_{loc}(0, +\infty)$ with $v \geq |\dot{u}|$, as it has already been proved for the existence of minimizers of I_ε (indeed, observe that in that proof, and more precisely in the compactness of the sublevels part, we did not use the fact that ε was fixed);

(b) since $u_{\varepsilon_n} \to u$ pointwise by the previous point, we have to understand what is the possible limit of $V_{\varepsilon_n}(u_{\varepsilon_n})$. Hence, show that $\Gamma - \lim_{\varepsilon \to 0} V_\varepsilon = \varphi$; more precisely, we will prove that, on the one hand, if $x_n \to x$ and $\varepsilon_n \to 0$, then

$$\liminf_{n\to\infty} V_{\varepsilon_n}(x_n) \geq \varphi(x) \tag{2.25}$$

and on the other hand

$$\lim_{\varepsilon \downarrow 0} V_\varepsilon(x) = \varphi(x) \tag{2.26}$$

which means that the constant sequence $x_n = x$ works as recovering sequence. Actually (2.26) follows for free from (2.25) and from the fact that $V_\varepsilon \leq \varphi$, which has already been pointed out in Remark 2.6;

(c) study the behaviour of

$$\Gamma - \liminf_{\varepsilon\downarrow 0} G_\varepsilon(x) := \inf \left\{ \liminf_{n\to\infty} G_{\varepsilon_n}(x_n) \, : \, x_n \to x \right\}$$

This Γ-liminf will be also denoted by G^-, since by definition it can be seen as relaxation of the upper gradient G. In more details, we will prove that

$$|\partial^- \varphi| \leq G^- \leq G := \limsup_{\varepsilon\downarrow 0} G_\varepsilon \leq |\partial\varphi| \tag{2.27}$$

where $|\partial^-\varphi|$ has already been defined at (1.9) as the relaxation of $|\partial\varphi|$. This shows that G^- is not an object as bad as it seems, because it can be controlled in terms of slopes and if $|\partial\varphi|$ is lower semicontinuous, then all the inequalities above turn into equalities.

Hence, by property (a) we can handle the convergence of the first summand in the integral in (2.24) and by property (b) we are able to control $V_\varepsilon(u_\varepsilon(T))$ with $\varphi(u(T))$, even if equality turns into an inequality. Finally, by definition of G^- the desired result will be achieved and (2.24) will become

$$\varphi(u(T)) + \frac{1}{2}\int_0^T \Big(|\dot{u}|^2(t) + (G^-)^2(u(t)) \Big) dt \leq \varphi(\overline{u})$$

Thus, if G^- is an upper gradient for φ, we have been able to produce a gradient flow for φ w.r.t. G^- and this concludes the proof of Theorem 2.1. Pay attention to the fact that an analogous inequality is obtained by the MM technique, but in that case the Γ-liminf of G_ε is replaced by the Γ-liminf of the descending slope.

Let us now demonstrate the properties claimed above, in the same order we have presented them.

2.3.1 Uniform bounds

For (a), let us observe first that if we forget G_ε^2 in (2.24) we immediately get

$$\int_0^T \frac{1}{2}|\dot u_\varepsilon|^2(t)dt \le V_\varepsilon(\overline{u}) - V_\varepsilon(u_\varepsilon(T))$$

and the right-hand side can be easily estimated, because $V_\varepsilon(\overline{u}) \le \varphi(\overline{u})$. Hence, in the case φ is non-negative, the first desired estimate follows, because V_ε is non-negative too and therefore

$$\int_0^T |\dot u_\varepsilon|^2(t)dt \le 2\varphi(\overline{u}) =: C$$

which is an even stronger bound, because C does not depend on time. In the general case, it is sufficient to use the fact that

$$V_\varepsilon(x) \ge -a - bd^2(x, \overline{u})$$

so that

$$\int_0^T \frac{1}{2}|\dot u_\varepsilon|^2(t)dt \le \varphi(\overline{u}) + a + bd^2(u_\varepsilon(T), \overline{u})$$

and it is now sufficient to apply Gronwall's lemma to get a uniform bound independent of ε, namely

$$\int_0^T \frac{1}{2}|\dot u_\varepsilon|^2(t)dt \le C_T \tag{2.28}$$

In order to get a uniform estimate also for the integral of $\varphi(u_\varepsilon)$, consider the inner variation equation (2.13) and observe that it can be rewritten in terms of the value function V_ε, because by comparing the inner variation estimate (2.13) and the distributional derivative of the value function (2.19) we got in the metric case, we can guess that

$$V_\varepsilon(u_\varepsilon) = \varphi(u_\varepsilon) - \frac{\varepsilon}{2}|\dot u_\varepsilon|^2 \tag{2.29}$$

In order to prove this claim, notice that the right-hand side is a Sobolev function, hence it belongs to $AC^2_{loc}([0,+\infty[)$. Let then W_ε be its (unique) absolutely continuous representative and observe that

$$W_\varepsilon' = -|\dot u_\varepsilon|^2 \tag{2.30}$$

Thanks to this fact, for any $0 \le a \le b < +\infty$ we have

$$\begin{aligned} \int_a^b \Big(\frac{\varepsilon}{2}|\dot u_\varepsilon|^2 + \varphi(u_\varepsilon)\Big)d\mu_\varepsilon &= \int_a^b \Big[\varepsilon|\dot u_\varepsilon|^2 + \Big(\varphi(u_\varepsilon) - \frac{\varepsilon}{2}|\dot u_\varepsilon|^2\Big)\Big]d\mu_\varepsilon(t) \\ &= \int_a^b \Big(-\varepsilon W_\varepsilon' + W_\varepsilon\Big)d\mu_\varepsilon \end{aligned}$$

$$= \quad W_\varepsilon(a)e^{-a/\varepsilon} - W_\varepsilon(b)e^{-b/\varepsilon}$$

where the last identity follows by an integration by parts. As a next step, $W_\varepsilon(b)e^{-b/\varepsilon} \to 0$ as $b \to +\infty$, because by (2.30) $t \mapsto W_\varepsilon(t)$ is non-increasing and in addition

$$\int_0^{+\infty} W_\varepsilon(t)e^{-t/\varepsilon}dt < +\infty$$

Thus, if we choose $a = 0$ and we let $b \to +\infty$, we entail that

$$\int_0^{+\infty} \left(\frac{\varepsilon}{2}|\dot{u}_\varepsilon|^2 + \varphi(u_\varepsilon)\right)d\mu_\varepsilon = W_\varepsilon(0)$$

but the left-hand side coincides with $V_\varepsilon(u_\varepsilon(0))$ by definition, so that $V_\varepsilon(u_\varepsilon(0)) = W_\varepsilon(0)$ and by the dynamic programming principle we can say that $V_\varepsilon \circ u_\varepsilon = W_\varepsilon$ *tout court*, whence (2.29). As a byproduct, we get that

$$\frac{d}{dt}V_\varepsilon(u_\varepsilon) = -|\dot{u}_\varepsilon|^2 \tag{2.31}$$

even in the metric case, as already anticipated. The advantage of the identification is due to the fact that $\varphi \circ u_\varepsilon$ can be expressed as

$$\varphi(u_\varepsilon) = V_\varepsilon(u_\varepsilon) + \frac{\varepsilon}{2}|\dot{u}_\varepsilon|^2$$

On the one hand, $V_\varepsilon(u_\varepsilon)$ is upper bounded since non-increasing (this follows from (2.31), but it has already been noticed in Remark 2.6) and $u_\varepsilon(0) = \overline{u}$ for every $\varepsilon > 0$. On the other hand, $\varepsilon|\dot{u}_\varepsilon|^2/2$ is locally uniformly integrable in ε by (2.28). As a consequence, we immediately deduce that

$$\int_0^T \varphi(u_\varepsilon(t))dt \leq C'_T$$

and so also the second desired inequality in (a) is established.

2.3.2 Γ-limit of V_ε

Let $(x_n)_n \subset X$ be a sequence converging to x, $(\varepsilon_n)_n$ be such that $\varepsilon \downarrow 0$ and, for any $n \in \mathbb{N}$, let u_n be a minimizer of I_{ε_n} with initial condition x_n. Up to extraction, let us assume that

$$V_{\varepsilon_n}(x_n) \to \liminf_{n\to\infty} V_{\varepsilon_n}(x_n) \tag{2.32}$$

and, without loss of generality, let us also suppose the liminf above to be finite (otherwise, (2.25) is trivially true). By the boundedness of $(x_n)_n$ and by (a), up to extract a further subsequence we can assume that $(u_n)_n$ is pointwise convergent to u and its metric derivative is weakly convergent in $L^2_{loc}(0, +\infty)$. However, a precisation is required, because (a) can not be directly applied; indeed, in the previous subsection we assumed that all the minimizers had the same starting point $\overline{u}$, while now the initial point is not fixed, although remains bounded. This problem can be easily avoided, since, as already said, up to pass to a suitable subsequence we can assume $(V_{\varepsilon_n}(x_n))_{n\in\mathbb{N}}$ to be bounded and this is sufficient, because $V_\varepsilon(\overline{u}) \le \varphi(\overline{u})$ can be replaced by $V_{\varepsilon_n}(x_n) \le \tilde{C}$ for some constant $\tilde{C}$.

Remark 2.8. For the limit curve u to exist, the inequalities established in (a) are required and can not be avoided. Indeed, one could observe that, by writing

$$V_{\varepsilon_n}(x_n) = \int_0^{+\infty} \left(\frac{\varepsilon_n}{2}|\dot{u}_n|^2 + \varphi(u_n)\right) d\mu_\varepsilon \tag{2.33}$$

and by (2.32), the right-hand side in (2.33) is uniformly bounded w.r.t. $n \in \mathbb{N}$, so that it would be tempting to repeat the same argument we performed in the subsection *Compactness of the sublevels.* Unfortunately, in the present situation (2.8) is no longer true, because

$$C_T := \varepsilon e^{T/\varepsilon} C' \to +\infty$$

as $\varepsilon \downarrow 0$. Thus, a shortcut is not possible. $\Diamond$

For sake of simplicity let us finally assume that $\varphi \ge 0$ (without this hypothesis some technicalities are required, but we will not deal with them[1]). By a simple rescaling we have

$$\begin{aligned} V_{\varepsilon_n}(x_n) &= \int_0^{+\infty} \left(\frac{\varepsilon_n}{2}|\dot{u}_n|^2 + \varphi(u_n)\right) d\mu_{\varepsilon_n} \\ &= \int_0^{+\infty} \left(\frac{1}{2\varepsilon_n}|\dot{u}_n|^2(\varepsilon_n s) + \varphi(u_n(\varepsilon_n s))\right) e^{-s} ds \\ &\ge \int_0^{+\infty} \varphi(u_n(\varepsilon_n s)) e^{-s} ds \end{aligned}$$

In order to handle this inequality as $n \to \infty$, observe that by triangle inequality and by the fact that the sequence $(u_n)_n$ is uniformly Hölder continuous with exponent $\alpha = 1/2$ (the argument we used at (2.9) is still true in the present situation), it holds

$$d(u_n(\varepsilon_n s), x) \le d(u_n(\varepsilon_n s), x_n) + d(x_n, x) \le C_t^{1/2}\sqrt{t} + d(x_n, x)$$

so that $u_n(\varepsilon_n s) \to x$ as $n \to \infty$ for any $s \ge 0$. Therefore, by Fatou's lemma,

$$\liminf_{n\to\infty} V_{\varepsilon_n}(x_n) \ge \int_0^{+\infty} \varphi(x) e^{-s} ds = \varphi(x)$$

and this proves (2.25). The fact that (2.26) also holds has already been commented, so that it remains only (c).

[1] For sake of information, we have to add and subtract the square of the length function L_n associated to u_n, as we have already done in order to prove the good definition of I_ε.

2.3.3 Behaviour of G^-

In order to study the behaviour of G_ε and prove that $|\partial^-\varphi| \le G^-$, $G \le |\partial\varphi|$ pointwise, it is useful to rewrite the integral functional I_ε; to this aim let us introduce the *energy functional* E relative to u as follows

$$E(t) := \int_0^t |\dot{u}|^2(s)ds$$

and observe that $E'(t) = |\dot{u}|^2(t)$ for $\mathscr{L}^1$-a.e. t, so that by a simple integration by parts and by the fact that $E(t)e^{-t/\varepsilon} \to 0$ as $t \to +\infty$ we get

$$\int_0^{+\infty} \frac{\varepsilon}{2}|\dot{u}|^2\, d\mu_\varepsilon = \int_0^{+\infty} \frac{1}{2}E\, d\mu_\varepsilon$$

Hence, the functional I_ε can be rewritten as

$$I_\varepsilon(u) = \int_0^{+\infty} \left(\frac{1}{2}E(t) + \varphi(u(t))\right) d\mu_\varepsilon(t)$$

for any curve $u \in AC^2_{loc}([0,+\infty[, X)$. After this premise, let us show first that $G \le |\partial\varphi|$ pointwise and note that, by definition, in order to get an upper bound on G as $\varepsilon \downarrow 0$ we need a lower bound on V_ε. Thus, let u be optimal (for this reason, from now on we switch to the usual notation and u will be denoted by u_ε) with $u_\varepsilon(0) = x$, whence

$$V_\varepsilon(x) = \int_0^{+\infty} \left(\frac{1}{2}E(t) + \varphi(u_\varepsilon(t))\right) d\mu_\varepsilon(t) \tag{2.34}$$

By Jensen's inequality

$$E(t) = t \fint_0^t |\dot{u}_\varepsilon|^2(s)ds \ge t\left(\fint_0^t |\dot{u}_\varepsilon|(s)ds\right)^2 \ge \frac{d^2(x, u_\varepsilon(t))}{t}$$

and by replacing this inequality into (2.34) it follows that

$$V_\varepsilon(x) \ge \int_0^{+\infty} \left(\frac{d^2(x, u_\varepsilon(t))}{2t} + \varphi(u_\varepsilon(t))\right) d\mu_\varepsilon(t) \ge \int_0^{+\infty} Y_t(x) d\mu_\varepsilon(t)$$

where Y_t is the Moreau-Yosida approximation we defined at (2.21). As a consequence

$$\frac{\varphi - V_\varepsilon}{\varepsilon}(x) \le \int_0^{+\infty} \frac{\varphi(x) - Y_t(x)}{\varepsilon} d\mu_\varepsilon(t) = \int_0^{+\infty} \frac{\varphi(x) - Y_{\varepsilon s}(x)}{\varepsilon s} s e^{-s} ds$$

and thus, taking into account the characterization of the descending slope in terms of the Yosida approximations

$$\limsup_{\eta \downarrow 0} \frac{\varphi(x) - Y_\eta(x)}{\eta} = \frac{1}{2}|\partial\varphi|^2(x)$$

we infer that

$$\limsup_{\varepsilon\downarrow 0} \frac{\varphi - V_\varepsilon}{\varepsilon}(x) \leq \int_0^{+\infty} \frac{1}{2}|\partial\varphi|^2(x) s e^{-s} ds = \frac{1}{2}|\partial\varphi|^2(x)$$

namely $G \leq |\partial\varphi|$ pointwise. In the opposite direction, we have to prove $|\partial^-\varphi| \leq G^-$. To this aim, let $x \in X$ be an arbitrary point and let $w \in AC^2_{loc}(]a,b[, X)$ be a curve such that $w(t_0) = x$ for a suitable $t_0 \in]a,b[$ and $\varphi \circ w \in L^1_{loc}(a,b)$. Then

$$\begin{aligned} \frac{1}{2}G^2(x) &= \limsup_{\varepsilon\downarrow 0} \frac{\varphi(x) - V_\varepsilon(x)}{\varepsilon} \\ &\geq \liminf_{h\downarrow 0} \frac{\varphi(w(t_0)) - \varphi(w(t_0+h))}{h} - \limsup_{h\downarrow 0} \frac{1}{2h}\int_{t_0}^{t_0+h} |\dot{w}|^2(t)dt \end{aligned}$$

and if we assume t_0 to be a differentiation point for w and $\varphi \circ w$, this inequality can be restated as

$$\frac{1}{2}G^2(x) \geq (\varphi \circ w)'(t_0) + \frac{1}{2}|\dot{w}|^2(t_0)$$

The crucial fact is that the bound holds for any absolutely continuous curve; hence, if we speed up w by a factor $\delta > 0$, then we get

$$\frac{1}{2}G^2(x) \geq (\varphi \circ w)'(t_0)\delta + \frac{1}{2}|\dot{w}|^2(t_0)\delta^2$$

and now if we optimize this inequality w.r.t δ assuming $(\varphi\circ w)'(t_0)$ to be positive, i.e. we choose

$$\delta = \frac{(\varphi \circ w)'(t_0)}{|\dot{w}|^2(t_0)}$$

(pay attention to the fact that by the previous assumption $\delta > 0$ is satisfied) then we obtain

$$\frac{1}{2}G^2(x) \geq \frac{1}{2}\frac{[(\varphi \circ w)'(t_0)]^2}{|\dot{w}|^2(t_0)}$$

whence

$$|(\varphi \circ w)'(t_0)| \leq G(x)|\dot{w}|(t_0)$$

Thus, G is a good candidate to be a slope of φ, at least along curves (because of this restriction we can not actually talk about *slope*), because it is bounded from above by the global descending slope and bounds from above any derivative of φ along curves. By means of this intermediate step, it is possible to prove that $|\partial^-\varphi| \leq G^-$ pointwise, but unfortunately, because of lack of time, we can not give the details.

Let us only say that if the relaxation $|\partial^-\varphi|$ defined as $\Gamma - \liminf |\partial\varphi|$ is an upper gradient of φ, then $G^- \equiv |\partial\varphi|$, which means that G^- is an upper gradient of φ too, and the limit function u satisfies the gradient flow condition w.r.t. G^- with equality, i.e.

$$\varphi(u(t)) + \frac{1}{2}\int_0^t \Big(|\dot{u}|^2(s) + (G^-)^2(u(s))\Big)ds = \varphi(\overline{u})$$

or equivalently

$$\frac{d}{dt}\varphi(u(t)) = -|\dot{u}|^2(t) = -(G^-)^2(u(t)) = -|\partial^-\varphi|^2(u(t))$$

In particular, this is the case if φ is λ-geodesically convex in X. Under this strong assumption, $|\partial^-\varphi| = |\partial\varphi|$ and as a consequence all the inequalities in (2.27) turn into equalities. The same conclusion is true if $|\partial\varphi|$ is lower semicontinuous. In the general case, more refined estimates on the behaviour of G^- are required and this is a hard task.

3 Why the WED approach

As a motivating remark, the WED method is probably more complicated than the MM one, but one of the advantages is that in many situations, as for instance the infinite-dimensional case, it is easier to control the derivative of φ along curves (if φ is λ-geodesically convex, then the behaviour of φ is tested along geodesics) rather than the difference of φ at two different points, which are not necessarily connected by a path. In other terms, it is easier to work with G rather than with $|\partial\varphi|$, because the slope is defined globally, whereas for $G(x)$ the set of competitors is smaller, since we need only to consider those points connected to x by an absolutely continuous curve. And *a posteriori* this is natural, since our aim is to find a gradient flow starting from a given initial datum; hence it is useless to check the incremental ratio of φ along all the descending directions: it is more convenient to focus our attention to the descending directions given by absolutely continuous curves.