Pacific Institute for the Mathematical Sciences

A least action principle for steepest descent in a non-convex landscape

Nassif Ghoussoub
Department of Mathematics
University of British Columbia
Vancouver, BC, Canada

Robert J. McCann
Department of Mathematics
University of Toronto
Toronto, ON, Canada

Preprint number: PIMS-04-2
Received on January 19, 2004
Revised on March 19, 2004

http://www.pims.math.ca/publications/preprints/

A least action principle for steepest descent in a non-convex landscape

Nassif Ghoussoub and Robert J. McCann

1. Introduction

Physical dynamics interpolate naturally between the dissipative and conservative extremes, in which friction either dominates or can be neglected. *Gradient flows* and *Hamiltonian systems* represent the archetypal examples of these two extremes. The orbits of a Hamiltonian system correspond to the critical paths of an *action* functional, but variational characterizations for the trajectories of a gradient flow are less familiar. For steepest descent into a *convex valley* such a characterization was formulated by Brezis-Ekeland [**4**], but their principle was not amenable to deducing existence of solutions. Later, Auchmuty [**2**] used min-max methods to establish the existence of solutions variationally, but under certain growth conditions on the convex potential. Recently, Ghoussoub and Tzou [**8**] deduced the existence of semigroup flows in full generality (for convex lower semicontinuous potentials) using a modified Brezis-Ekeland principle which is invariant under Bolza duality [**12**]. For convex gradient flows, their method now provides a direct alternative to proving existence by maximal monotone accretive operator theoretic semigroup methods or by time-step approximation; recent references for the latter techniques include [**3**] [**7**] [**10**] [**1**]. In a forthcoming paper [**9**], the first-named author develops further the scope of the duality method and proposes a general framework for a variational formulation of many equations which do not normally fit into standard Euler-Lagrange theory. This approach is based on a concept of *anti-self dual Lagrangians* which seems to be inherent in many important differential equations.

In the present article we streamline and broaden the scope of the Ghoussoub-Tzou approach to gradient flows, showing how it can be adapted to characterize the path of steepest descent of a *non-convex* potential as the global minimum of a *convex* action. This is achieved by *dynamically rescaling space*, to convert the energy landscape on which our descent takes place from a static but non-convex profile to a contracting convex one. This approach is inspired by the time-dependent change of coordinates used to find similarity variables in nonlinear partial differential equations, especially as employed by Otto to quantify long-time behaviour of porous medium flows [**11**]. Although that descent takes place on the 'Riemannian' manifold of probability measures metrized by Wasserstein distance, our present considerations will be restricted to the more traditional setting of a Hilbert space H with norm $|u| = \langle u, u\rangle^{1/2}$.

Let $[0, T]$ be a fixed real interval ($0 < T < +\infty$). Consider the Hilbert space $L^2_H := L^2([0,T]; H)$ of Bochner integrable functions from $[0,T]$ into H with norm denoted by $\|\cdot\|_{L^2_H}$,

2000 *Mathematics Subject Classification.* Primary 37L05; Secondary 34G20, 35K55, 47J35, 49N15.

Key words and phrases. variational principle, gradient flow, steepest descent, dissipative, evolution, nonlinear, semigroup, self-dual.

The first author's research was partially supported by a grant from the Natural Sciences and Engineering Research Council of Canada. He gratefully acknowledges the hospitality and support of the Centre de Recherches Mathematiques in Montreal where this work was completed.

The second author's research was supported in part by United States National Science Foundation Grant DMS 0074037 and Natural Sciences and Engineering Research Council of Canada Grant 217006-03 RGPIN. This work was initiated at the Banff International Research Station.

and the Sobolev space

$$A_H^2 := \{u : [0,T] \to H;\ \dot u \in L_H^2\} =: AC^2([0,T];H)$$

consisting of all absolutely continuous vector-valued arcs $u : [0,T] \to H$, equipped with the norm

$$\|u\|_{A_H^2} = \left(|u(0)|^2 + \int_0^T |\dot u|^2 dt\right)^{\frac{1}{2}}.$$

Our main result is formulated under a *semiconvexity* assumption, which means that the energy landscape differs from a convex valley by a smooth function. In particular, this assumption is satisfied whenever the landscape is sufficiently smooth.

THEOREM 1.1 (Least action descent in a non-convex landscape). *Let $W : H \longrightarrow \mathbf{R} \cup \{+\infty\}$ be* semiconvex*, meaning for some $k \geq 0$ the function $\overline{W}(u) := W(u) + k|u|^2/2$ is strictly convex, lower semicontinuous on H, and not identically infinity. Then $V(t,v) := e^{-2kt}\overline{W}(e^{kt}v)$ is convex at each instant in time so let*

$$V^*(t,u) := \sup_{v \in H} \langle u, v\rangle - V(t,v)$$

be its Legendre-Fenchel transform. For any $u_0 \in \operatorname{dom} \partial W$, consider the functional

$$\Phi[u] = \frac{1}{2}(|u(0)|^2 + |u(T)|^2) - 2\langle u(0), u_0\rangle + |u_0|^2 + \int_0^T [V(t,u(t)) + V^*(t, -\dot u(t))]\, dt \tag{1}$$

on the path space A_H^2. Then there exists a unique v in A_H^2 such that

$$\Phi[v] = \inf_{u \in A_H^2} \Phi[u] = 0. \tag{2}$$

Moreover, the path $w(t) := e^{kt}v(t)$ is the unique solution in A_H^2 to

$$\begin{cases} -\dot w(t) & \in & \partial W(w(t)) \quad \text{a.e.} \quad \text{on} \quad [0,T] \\ w(0) & = & u_0. \end{cases} \tag{3}$$

Here the set $\partial W(u)$ is related to the subdifferential of $\overline{W}$ by $\partial\overline{W}(u) = \partial W(u) + ku$, and $\operatorname{dom} \partial W := \{u \in H \mid W(u) < +\infty \text{ and } \partial W(u) \neq \emptyset\}$. An alternate representation (9) of the action $\Phi[\,\cdot\,]$ shows why it is minimized by the evolution (3), but conceals the convexity manifest in (1).

This theorem can of course be deduced from well-known existence results concerning semi-group flows. Our purpose here is to give a simple variational proof which implies these existence results. At the same time, we highlight the convex-analytic properties of our parabolic rescaling of space: it commutes with both Legendre transformation and Yosida regularization.

2. Duality and gradient flow in an evolving landscape

The path space A_H^2 is also a Hilbert space that can be identified with the product space $H \times L_H^2$, while its dual $(A_H^2)^*$ can be identified with $H \times L_H^2$. The duality is given by the formula:

$$\langle u, (a,p)\rangle_{A_H^2, H\times L_H^2} = \langle u(0), a\rangle_H + \int_0^T \langle \dot u(t), p(t)\rangle dt.$$

THEOREM 2.1 (Least action descent into an evolving valley). *Let $V : [0,T] \times H \to \mathbf{R} \cup \{+\infty\}$ be a measurable function with respect to the σ-field in $[0,T] \times H$ generated by the products of Lebesgue sets in $[0,T]$ and Borel sets in H. Assume V satisfies the following conditions:*

- *For every $t \in [0,T]$, the function $V(t,\cdot)$ is convex and lower semicontinuous on H.*
- $\int_0^T V^*(t,0)dt < \infty.$
- *There is $\alpha \in L^\infty[0,T]$ such that:*

$$V(t,x) \leq \alpha(t)(1 + |x|^2) \quad \textit{for } (t,x) \in [0,T] \times H. \tag{4}$$

For any $u_0 \in H$, the functional

$$\Phi[u] = \frac{1}{2}(|u(0)|^2 + |u(T)|^2) - 2\langle u(0), u_0\rangle + |u_0|^2 + \int_0^T [V(t, u(t)) + V^*(t, -\dot{u}(t))]\, dt \tag{5}$$

admits a unique minimizer v in A^2_H and

$$\Phi[v] = \inf_{u \in A^2_H} \Phi[u] = 0. \tag{6}$$

Among paths in A^2_H, v is the unique solution to

$$\begin{cases} -\dot{v}(t) & \in & \partial V(t, v(t)) \quad \text{a.e. on} \quad [0, T] \\ v(0) & = & u_0. \end{cases} \tag{7}$$

Proof. First, we notice that $\Phi[\,\cdot\,]$ is convex and that:

$$\Phi[u] \geq 0 \quad \text{for all } u \text{ in } A^2_H. \tag{8}$$

Indeed, it is clear that

$$\Phi[u] = |u(0) - u_0|^2 + \int_0^T [V(t, u(t)) + V^*(t, -\dot{u}(t)) + \langle u(t), \dot{u}(t)\rangle]\, dt. \tag{9}$$

By the Fenchel-Young inequality, we have

$$V(t, u(t)) + V^*(t, -\dot{u}(t)) \geq \langle u(t), -\dot{u}(t)\rangle = -\frac{1}{2}\frac{d}{dt}|u(t)|^2 \quad \text{a.e. on } [0, T] \tag{10}$$

with equality holding if and only if $-\dot{u}(t) \in \partial V(t, u(t))$.

This also yields that $\Phi[u] \geq |u(0) - u_0|^2 \geq 0$, which means that (6) would automatically imply (7).

The rest of the section deals with the reverse inequality and the existence of a minimum. We follow the method of Ghoussoub-Tzou [**8**] by showing that $\Phi[\,\cdot\,]$ "behaves lower semicontinuously" with respect to certain perturbations. For that, we associate the following functional Ψ defined on $(A^2_H)^* = H \times L^2_H$ as:

$$\Psi[a, y] = \inf_{u \in A^2_H} \Big\{ \tfrac{1}{2}(|u(0) + a|^2 + |u(T)|^2) - 2\langle u(0) + a, u_0\rangle + |u_0|^2 + \int_0^T V(t, u(t) + y(t)) + V^*(t, -\dot{u}(t))dt \Big\} \tag{11}$$

so that

$$\Psi[0, 0] = \inf_{u \in A^2_H} \Phi[u]. \tag{12}$$

The following lemma establishes a key duality-symmetry between the two functionals.

Lemma 2.2 (Self-duality). *Defining $\Psi[\cdot, \cdot]$ by (11), the hypotheses of Theorem 2.1 imply*

$$\Psi^*[v] = \Phi[-v] \quad \text{for all } v \in A^2_H,$$

where Ψ^* is the Legendre transform of Ψ in the duality $(H \times L^2_H, A^2_H)$.

Proof of Lemma 2.2. For $v \in A^2_H$, write:

$$\Psi^*[v] = \sup_{a \in H} \sup_{y \in L^2_H} \sup_{u \in A^2_H} \Big\{ \langle a, v(0)\rangle - \tfrac{1}{2}(|u(0) + a|^2 + |u(T)|^2) + 2\langle u(0) + a, u_0\rangle - |u_0|^2 + \int_0^T [\langle y(t), \dot{v}(t)\rangle - V(t, u(t) + y(t)) - V^*(t, -\dot{u}(t)]dt \Big\}.$$

Making a substitution

$$u(0) + a = \tilde{a} \in H \quad \text{and } u + y = \tilde{y} \in L^2_H,$$

we obtain

$$\Psi^*[v] \geq \sup_{\tilde{a} \in H} \sup_{\tilde{y} \in L^2_H} \sup_{u \in A^2_H} \Big\{ \langle \tilde{a} - u(0), v(0)\rangle - \tfrac{1}{2}(|\tilde{a}|^2 + |u(T)|^2) + 2\langle \tilde{a}, u_0\rangle - |u_0|^2 + \int_0^T [\langle \tilde{y}(t) - u(t), \dot{v}(t)\rangle - V(t, \tilde{y}(t)) - V^*(t, -\dot{u}(t))]dt \Big\}.$$

Since $\dot{u} \in L^2_H$ and $u \in L^2_H$, we have:

$$\int_0^T \langle u, \dot{v}\rangle\, dt = \langle v(T), u(T)\rangle - \langle v(0), u(0)\rangle - \int_0^T \langle \dot{u}, v\rangle\, dt,$$

which implies

$$\Psi^*[v] \geq \sup_{\tilde{a}\in H} \sup_{\tilde{y}\in L^2_H} \sup_{u\in A^2_H} \Big\{ \quad \langle \tilde{a}, v(0)\rangle - \tfrac{1}{2}(|\tilde{a}|^2 + |u(T)|^2) + 2\langle \tilde{a}, u_0\rangle - |u_0|^2 - \langle u(T), v(T)\rangle \\ + \int_0^T [\langle \tilde{y}, \dot{v}\rangle + \langle \dot{u}, v\rangle - V(t, \tilde{y}(t)) - V^*(t, -\dot{u}(t))]\, dt \Big\}.$$

It is now convenient to identify A^2_H with $H \times L^2_H$ via the correspondence:

$$\begin{aligned} (c, x) \in H \times L^2_H \quad &\mapsto \quad u(t) = c + \int_t^T x(s)\, ds \in A^2_H \\ u \in A^2_H \quad &\mapsto \quad \big(u(T), -\dot{u}(t)\big) \in H \times L^2_H. \end{aligned}$$

We finally obtain

$$\begin{aligned} \Psi^*[v] \quad &\geq \quad -|u_0|^2 + \sup_{\tilde{a}\in H} \sup_{c\in H} \Big\{ \langle \tilde{a}, v(0) + 2u_0\rangle + \langle -c, v(T)\rangle - \frac{1}{2}(|\tilde{a}|^2 + |c|^2) \Big\} \\ &\qquad + \sup_{\tilde{y}\in L^2_H} \sup_{x\in L^2_H} \Big\{ \int_0^T (\langle \tilde{y}, \dot{v}\rangle + \langle x, -v\rangle - V(t, \tilde{y}(t)) - V^*(t, x(t)) dt \Big\} \\ &= \quad -|u_0|^2 + \frac{1}{2}(|v(0) + 2u_0|^2 + |v(T)|^2) + \int_0^T [V^*(t, \dot{v}(t)) + V(t, -v(t))]\, dt \\ &= \quad \frac{1}{2}(|v(0)|^2 + |v(T)|^2) + 2\langle v(0), u_0\rangle + |u_0|^2 + \int_0^T [V(t, -v(t)) + V^*(t, \dot{v}(t))]\, dt \\ &= \quad \Phi[-v]. \end{aligned}$$

Here we have used that $V(t,\cdot)^{**} = V(t,\cdot)$ and that for any $L : [0, T] \times H \times H \to \mathbf{R} \cup \{+\infty\}$ convex and lower semicontinuous, we have:

$$\int_0^T L^*(t, s(t), v(t)) dt = \sup_{x,y\in L^2_H} \int_0^T [\langle y(t), s(t)\rangle + \langle x(t), v(t)\rangle - L(t, y(t), x(t))] dt$$

where L^* is the Legendre dual of L in both state variables. □

End of proof of Theorem 2.1. First we prove that the convex functional Ψ is subdifferentiable at $(0, 0)$. For that, it is sufficient to show that Ψ is bounded on neighborhoods of zero in $H \times L^2_H$. Note that

$$\begin{aligned} \Psi[a, y] \quad &\leq \quad 2|u_0| \cdot |a| + \frac{|a|^2_H}{2} + |u_0|^2 + \int_0^T [V(t, y(t)) + V^*(t, 0)] dt \\ &\leq \quad 2|u_0| \cdot |a| + \frac{|a|^2_H}{2} + |u_0|^2 + \int_0^T \big[\alpha(t)(|y(t)|^2 + 1) + V^*(t, 0)\big]\, dt \end{aligned}$$

which means that Ψ is bounded in a neighborhood of $(0, 0)$ in the space $H \times L^2_H$, hence it is subdifferentiable at $(0, 0)$.

Now take $-v \in \partial\Psi[0, 0] \in A^2_H$. Then again by Young's inequality,

$$\Psi[0, 0] + \Psi^*[-v] = 0 = \Phi[v] + \inf_{u\in A^2_H} \Phi[u]$$

It follows that:

$$-\inf_{A^2_H} \Phi = \Phi[v] \geq \inf_{A^2_H} \Phi$$

which means that $\inf_{A^2_H} \Phi \leq 0$. In view of (8), we must have $\inf_{A^2_H} \Phi = 0 = \Phi[v]$. Thus the minimum is zero and is attained at v. □

3. Yosida regularization and gradient flow of a semiconvex potential

Consider again a measurable function $v : [0,T] \times H \to \mathbf{R} \cup \{+\infty\}$ such that $V(t,\cdot)$ is convex and lower semicontinuous on H for every $t \in [0,T]$, but without the bound (4) of Theorem 2.1, which is quite restrictive and not satisfied by most potentials of interest. One way to remedy this is to regularize V by using inf-convolution. That is, for each $\lambda > 0$ we define

$$V_\lambda(t,x) = \inf\{V(t,y) + \frac{1}{2\lambda}|x-y|_H^2;\ y \in H\},$$

so that

$$V_\lambda(t,x) \le V(t,0) + \frac{1}{2\lambda}|x|^2$$

while its conjugate is given by

$$V_\lambda^*(t,y) = V^*(t,y) + \frac{\lambda}{2}|y|^2. \tag{13}$$

The V_λ now satisfy the hypothesis of Theorem 2.1 (as long as $\int_0^T V(t,0) + V^*(t,0)dt < \infty$) and therefore the corresponding evolution equations

$$\begin{cases} \dot v(t) + \partial V_\lambda(t, v(t)) &= 0 \quad \text{a.e. on} \quad [0,T] \\ v(0) &= u_0 \end{cases} \tag{14}$$

have unique solutions $v_\lambda(t)$ in A_H^2 that minimize

$$\Phi_\lambda[v] := |v(0)-u_0|^2 + \int_0^T [V_\lambda(t,v(t)) + V_\lambda{}^*(t,-\dot v(t)) + \langle v(t), \dot v(t)\rangle]\,dt. \tag{15}$$

Now we need to argue that $(v_\lambda)_\lambda$ converges as $\lambda \to 0$ to a solution of the original problem. This analysis is reminiscent of the approach via the resolvent theory of Hille-Yosida, but is much easier here since the variational argument does not require uniform convergence of $(v_\lambda)_\lambda$ or their time-derivatives.

One still needs an upper bound on the L^2-norm of $(\dot u_\lambda(t))_\lambda$ however. This is straightforward when V is a time-independent convex potential (as shown for example in [**8**]), but not always possible for general time-dependent potentials. However, we shall be able to provide such an estimate in the special case when the time-dependent potential is of the form $V(t,x) = e^{-2kt}\overline{W}(e^{kt}x)$ with $\overline{W}$ being an appropriate convex function.

Let us summarize the properties of infimal convolution recapitulated from Evans [**7**, §9.6.1].

Lemma 3.1. *Let $\overline{W} : H \longrightarrow \mathbf{R} \cup \{+\infty\}$ be convex lower semicontinuous, $\overline{W}(u_0) < \infty$, and $\lambda > 0$. Then*

$$\overline{W}_\lambda(w) := \inf\{\overline{W}(u) + \frac{1}{2\lambda}|w-u|_H^2;\ u \in H\}, \tag{16}$$

is convex on H and bounded by

$$\overline{W}_\lambda(w) \le \overline{W}(u_0) + \frac{1}{2\lambda}|w-u_0|^2. \tag{17}$$

There exist Lipschitz maps $\nabla\overline{W}_\lambda : H \longrightarrow H$ and $J_\lambda : H \longrightarrow \operatorname{dom}\partial\overline{W}$ such that for each $w \in H$:
(i) (differentiability): $\partial\overline{W}_\lambda(w) = \{\nabla\overline{W}_\lambda(w)\}$;
(ii) (nonlinear resolvent): the infimum (16) is uniquely attained at $u = J_\lambda(w)$;
(iii) (first order condition):

$$\nabla\overline{W}_\lambda(w) = \frac{w - J_\lambda(w)}{\lambda} \in \partial\overline{W}(J_\lambda(w)); \tag{18}$$

(iv) (uniform gradient bound):

$$|\nabla\overline{W}_\lambda(w)| \le \inf_{u \in \partial\overline{W}(w)} |u|; \tag{19}$$

(v) (Lipschitz contractions): both J_λ and $\nabla(\lambda\overline{W}_\lambda/2)$ are contractions on H.

We shall call $\overline{W}_\lambda$ the Yosida λ-regularization of $\overline{W}$. Now we note the following useful property which says that the Yosida regularization actually commutes with our rescaling of space.

LEMMA 3.2 (Rescaling commutes with dualization and Yosida regularization). *Let $\overline{W}(u)$ be a lower semicontinuous convex function on H and consider the time dependent convex potential $V(t,v) := e^{-2kt}\overline{W}(e^{kt}v)$ and its Legendre-Fenchel transform $V^*(t,u)$ for each time t. Then*

$$V^*(t,u) = e^{-2kt}(\overline{W})^*(e^{kt}u) \tag{20}$$

where $(\overline{W})^*$ is the Legendre-Fenchel dual of $\overline{W}$. Moreover, if $V_\lambda(t,v)$ denotes the Yosida λ-regularization of $V(t,v)$ for each time t, and if $J_\lambda(t,v)$ is the corresponding attainment map, then

$$V_\lambda(t,v) := e^{-2kt}\overline{W}_\lambda(e^{kt}v) \quad \text{and} \quad J_\lambda(t,v) = e^{-kt}\tilde{J}_\lambda(e^{kt}v) \tag{21}$$

where $\overline{W}_\lambda$ is the Yosida λ-regularization of $\overline{W}$ and $\tilde{J}_\lambda$ its corresponding attainment map.

PROOF. Note

$$\begin{aligned} V^*(t,v) &= \sup\{\langle v,x\rangle - e^{-2kt}\overline{W}(e^{kt}x); x \in H\} \\ &= e^{-2kt}\sup\{\langle e^{kt}v, e^{kt}x\rangle - \overline{W}(e^{kt}x); x \in H\} \\ &= e^{-2kt}\sup\{\langle e^{kt}v, w\rangle - \overline{W}(w); w \in H\} \\ &= e^{-2kt}(\overline{W})^*(e^{kt}v). \end{aligned}$$

which proves (20). On the other hand,

$$\begin{aligned} (V_\lambda)^*(t,x) &= V^*(t,x) + \frac{\lambda}{2}|x|^2 \\ &= e^{-2kt}(\overline{W})^*(e^{kt}x) + \frac{\lambda}{2}|x|^2 \\ &= e^{-2kt}\left((\overline{W})^*(e^{kt}x) + \frac{\lambda}{2}|e^{kt}x|^2\right) \\ &= e^{-2kt}(\overline{W}_\lambda)^*(e^{kt}x). \end{aligned}$$

Use now (20) with $(\overline{W}_\lambda)^*$ instead of $\overline{W}$ to conclude that

$$V_\lambda(t,v) := e^{-2kt}\overline{W}_\lambda(e^{kt}v) \tag{22}$$

which means Yosida regularization commutes with our rescaling of space. □

Next we establish the required a priori estimate.

PROPOSITION 3.3 (Uniform speed limit). *Let $\overline{W}(u)$ be a lower semicontinuous convex function on H with $u_0 \in \operatorname{dom}\partial W$. Let $\overline{W}_\lambda$ be its Yosida regularization for each $\lambda > 0$ and define the time dependent potential $V_\lambda(v,t) := e^{-2kt}\overline{W}_\lambda(e^{kt}v)$ and its corresponding energy functional $\Phi_\lambda[\,\cdot\,]$ as in (15). If $v_\lambda \in A^2_H$ satisfies $\Phi_\lambda[v_\lambda] = 0$, then for a.e. $t \in [0,T]$, we have*

$$|\dot{v_\lambda}(t)| \leq C_0 := 2(k|u_0| + \inf_{u \in \partial\overline{W}(u_0)} |u|). \tag{23}$$

PROOF. Suppose $v_\lambda \in A^2_H$ satisfies $\Phi_\lambda[v_\lambda] = 0$. As in Theorem (2.1) we conclude that:

$$\begin{cases} -\dot{v_\lambda}(t) &= \nabla V_\lambda(v_\lambda(t),t) \qquad \text{for a.e. } t \in [0,T] \\ v_\lambda(0) &= u_0. \end{cases} \tag{24}$$

To obtain the desired estimate, it is convenient to exploit our dynamical rescaling of space to find an autonomous gradient flow equivalent to (24). Indeed, note that if $\tilde{W}(w) := \overline{W}_\lambda(w) - k|w|^2/2$, then $v_\lambda(t) = e^{-kt}w_\lambda(t)$ satisfies (24) if and only if

$$\begin{cases} -\dot{w_\lambda}(t) &= \nabla\tilde{W}(w_\lambda(t)) \qquad \text{for a.e. } t \in [0,T] \\ w_\lambda(0) &= u_0. \end{cases} \tag{25}$$

Next we establish that if $w_\lambda \in A^2_H$ satisfies (25) then

$$|\dot{w_\lambda}(t)| \leq e^{kt}|\nabla\tilde{W}(u_0)| \qquad \text{a.e. on } [0,T]. \tag{26}$$

Indeed, consider $f(t) = |\nabla \tilde{W}(w_\lambda(t))|^2/2$ which is absolutely continuous on $[0,T]$. Using $D^2\tilde{W} \geq -kI$ (assuming $\tilde{W}$ is C^2), we compute

$$\begin{aligned} f'(t) = &\quad \langle \nabla \tilde{W}(w_\lambda(t)), D^2\tilde{W}(w_\lambda(t))\dot{w_\lambda}(t)\rangle \\ = &- \langle \nabla \tilde{W}(w_\lambda(t)), D^2\tilde{W}(w_\lambda(t))\nabla \tilde{W}(w_\lambda(t))\rangle \\ \leq &\quad 2kf(t). \end{aligned}$$

Gronwall's inequality yields $f(t) \leq e^{2kt}f(0)$ whence $|\nabla \tilde{W}(w_\lambda(t))| \leq e^{kt}|\nabla \tilde{W}(w_\lambda(0))|$. This establishes (26) in light of (25).

Finally, we can address our main claim (23). Let $v_\lambda \in A^2_H$ satisfy (24) so that $w_\lambda(t) := e^{kt}v_\lambda(t)$ satisfies (26). Integrating that estimate yields

$$\begin{aligned} |w_\lambda(t) - w_\lambda(0)| &\leq \int_0^t |\dot{w_\lambda}(\tau)| d\tau && (27) \\ &\leq \frac{e^{kt}-1}{k}|\nabla \tilde{W}(u_0)|. && (28) \end{aligned}$$

Applying (26) and (28) to $\dot{v_\lambda}(t) = e^{-kt}(\dot{w_\lambda}(t) - kw_\lambda(t))$ gives:

$$\begin{aligned} |\dot{v_\lambda}(t)| &\leq e^{-kt}(e^{kt}|\nabla \tilde{W}(u_0)| + k|w_\lambda(0)| + (e^{kt}-1)|\nabla \tilde{W}(u_0)|) \\ &= (2 - e^{-kt})|\nabla \tilde{W}(u_0)| + ke^{-kt}|u_0| \\ &\leq 2|\nabla \overline{W}_\lambda(u_0)| + 2k|u_0| \\ &=: C_\lambda(u_0), \end{aligned}$$

where $\nabla \tilde{W}(u_0) = \nabla \overline{W}_\lambda(u_0) - ku_0$ has been used. Finally, the constants $C_\lambda(u_0) \leq C_0$ are bounded independently of λ by Lemma 3.1(iv), completing the proposition. □

PROOF OF THEOREM 1.1. Start with $W : H \longrightarrow \mathbf{R} \cup \{+\infty\}$ semiconvex, meaning that for some $k \geq 0$ the function $\overline{W}(u) := W(u) + k|u|^2/2$ is strictly convex, lower semicontinuous on H, and not identically infinity. Taking k larger if necessary ensures $\overline{W}(u)$ grows quadratically and hence attains its minimum. Set $V(t,v) := e^{-2kt}\overline{W}(e^{kt}v)$ and let $V^*(t,u)$ denote its Legendre-Fenchel transform for each time t. For any $u_0 \in \mathrm{dom}\,\partial W$, consider the functional

$$\Phi[u] = \frac{1}{2}(|u(0)|^2 + |u(T)|^2) - 2\langle u(0), u_0\rangle + |u_0|^2 + \int_0^T [V(t,u(t)) + V^*(t, -\dot{u}(t))]\,dt$$

on the space of curves A^2_H. We need to show that there exists v in A^2_H such that:

$$\Phi[v] = \inf_{u \in A^2_H} \Phi[u] = 0.$$

Uniqueness of v then follows from strict convexity of $\Phi[\,\cdot\,]$, and it is easy to see that the path $w(t) := e^{kt}v(t)$ satisfies

$$\begin{cases} \dot{w}(t) + \partial W(w(t)) = 0 \quad \text{a.e. on } [0,T] \\ w(0) = u_0 \end{cases}$$

from the equality conditions in Young's inequality (10).

Let $\overline{W}_\lambda$ be the Yosida regularization of $\overline{W}$ for each $\lambda > 0$ and its associated map $\tilde{J}_\lambda$ from Lemma 3.1. We know from Lemma 3.2 that the λ-regularization of V satisfies $V_\lambda(v,t) := e^{-2kt}\overline{W}_\lambda(e^{kt}v)$ and that its corresponding attainment map $J_\lambda(t,v) = e^{-kt}\tilde{J}_\lambda(e^{kt}v)$ where $\tilde{J}_\lambda$ is the attainment map for $\overline{W}_\lambda$. If Φ_λ denotes the energy functional (15), then Theorem 2.1 yields $v_\lambda \in A^2_H$ such that $\Phi_\lambda[v_\lambda] = 0$ for each $\lambda > 0$. That is

$$\begin{cases} -\dot{v_\lambda}(t) = \nabla V_\lambda(t, v_\lambda(t)) \qquad \text{for a.e. } t \in [0,T] \\ v_\lambda(0) = u_0. \end{cases} \tag{29}$$

By Proposition 3.3, we have for a.e. $t \in [0,T]$, the estimate

$$|\dot{v_\lambda}(t)| \leq C_0 := 2(k|u_0| + \inf_{u \in \partial \overline{W}(u_0)} |u|). \tag{30}$$

It follows that a subsequence $(v_{\lambda_j})_j$ is converging weakly in A^2_H to a path v. The projection of this path onto any vector in H lies in the real Sobolev space $A^2_{\mathbf{R}} \subset L^2[0,T]$ of Hölder-1/2 functions, hence converges pointwise. For each $t \in [0,T]$, it follows that $v_{\lambda_j}(t) \to v(t)$ weakly in H as $\lambda_j \to 0$. From Lemma 3.1, we have for any $\lambda > 0$ and any $t \geq 0$,

$$|v_\lambda(t) - J_\lambda(t, v_\lambda(t))| = \lambda |\nabla V_\lambda(t, v_\lambda)| = \lambda |\dot{v_\lambda}(t)| \leq \lambda C_0,$$

thus $J_{\lambda_j}(t, v_{\lambda_j}(t)) \to v(t)$ weakly in H for every $t \in [0,T]$.

Now $V(t,\cdot)$ and $V^*(t,\cdot)$ are weakly lower semi-continuous on H, and $V(t,\cdot)$ is bounded below by $\inf_H \overline{W}(u) > -\infty$. Using Fatou's lemma one easily deduces:

$$\begin{aligned} \int_0^T V(t, \quad v(t))dt &\leq \varliminf_j \int_0^T V(t, J_{\lambda_j}(t, v_{\lambda_j}(t)))dt, \\ \int_0^T V^*(t, -\dot{v}(t))dt &\leq \varliminf_j \int_0^T V^*(t, -\dot{v}_{\lambda_j}(t))dt, \end{aligned} \tag{31}$$

where $(u_0, p_0) \in \partial \overline{W}$, (20), convexity of $V^*(t,\cdot)$, and the bound

$$\begin{aligned} V^*(t, -\dot{v}(t)) &\geq e^{-2kt}[(\overline{W})^*(p_0) + \langle u_0, -e^{kt}\dot{v}(t) - p_0\rangle] \\ &\geq -|(\overline{W})^*(p_0) - \langle u_0, p_0 \rangle| - C_0|u_0| \end{aligned}$$

have been used to establish strong and hence weak lower semicontinuity (31). Moreover,

$$|v(0)|^2 \leq \varliminf_j |v_{\lambda_j}(0)|^2 \quad \text{and} \quad |v(T)|^2 \leq \varliminf_j |v_{\lambda_j}(T)|^2,$$

$$\int_0^T \frac{|v_{\lambda_j}(t) - J_{\lambda_j}(t, v_{\lambda_j}(t))|^2}{\lambda_j} \to 0,$$

and

$$\int_0^T \lambda_j |\dot{v}_{\lambda_j}(t)|^2 dt \leq C_0^2 T \lambda_j \to 0.$$

Since for every $t \in [0,T]$ and any $\lambda > 0$ and any $v, x \in H$, we have:

$$V(t, J_\lambda(t, v(t))) = V_\lambda(t, v(t)) + \frac{|v(t) - J_\lambda(t, v(t))|^2}{2\lambda}$$

and

$$V_\lambda^*(t, x) = V^*(t, x) + \frac{\lambda}{2}|x|^2,$$

it follows that

$$\begin{aligned} \Phi[v] &= \frac{1}{2}(|v(0)|^2 + |v(T)|^2) - 2\langle v(0), u_0\rangle + |u_0|^2 + \int_0^T [V(t, v(t)) + V^*(t, -\dot{v}(t))]\, dt \\ &\leq |u_0|^2 + \varliminf_j \frac{1}{2}(|v_{\lambda_j}(0)|^2 + |v_{\lambda_j}(T)|^2) - 2\langle v_{\lambda_j}(0), u_0\rangle \\ &\quad + \varliminf_j \int_0^T (\frac{|v_{\lambda_j}(t) - J_{\lambda_j}(t, v_{\lambda_j}(t))|^2}{2\lambda_j} + \frac{\lambda_j}{2}|\dot{v}_{\lambda_j}(t)|^2)dt \\ &\quad + \varliminf_j \int_0^T V(t, J_{\lambda_j}(t, v_{\lambda_j}(t)))dt + \varliminf_j \int_0^T V^*(t, -\dot{v}_{\lambda_j}(t))dt \\ &\leq |u_0|^2 + \varliminf_j \frac{1}{2}(|v_{\lambda_j}(0)|^2 + |v_{\lambda_j}(T)|^2) - 2\langle v_{\lambda_j}(0), u_0\rangle \\ &\quad + \varliminf_j \int_0^T V_{\lambda_j}(t, v_{\lambda_j}(t)) + (V_{\lambda_j})^*(t, -\dot{v}_{\lambda_j}(t))dt \\ &\leq \varliminf_j \Big(\frac{1}{2}(|v_{\lambda_j}(0)|^2 + |v_{\lambda_j}(T)|^2) - 2\langle v_{\lambda_j}(0), u_0\rangle + |u_0|^2 \\ &\qquad + \int_0^T V_{\lambda_j}(t, v_{\lambda_j}(t)) + V_{\lambda_j}^*(t, -\dot{v}_{\lambda_j}(t))dt\Big) \\ &= 0 \end{aligned}$$

From (9) we have the opposite inequality $\Phi[\,\cdot\,] \geq 0$ so the theorem is proved. □

References

[1] L.A. Ambrosio, N. Gigli, G. Savare. *Gradient flows in metric spaces and in the Wasserstein space of probability measures.* To appear: Birkhäuser, 2004.

[2] G. Auchmuty. *Saddle points and existence-uniqueness for evolution equations*, Differential Integral Equations, **6** (1993), 1161–1171.

[3] H. Brezis. *Operateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North Holland, Amsterdam-London, 1973.

[4] H. Brezis, I. Ekeland, *Un principe variationnel associé à certaines equations paraboliques. Le cas independant du temps*, C.R. Acad. Sci. Paris Sér. A **282** (1976), 971–974.

[5] H. Brezis, I. Ekeland, *Un principe variationnel associé à certaines equations paraboliques. Le cas dependant du temps*, C.R. Acad. Sci. Paris Sér. A **282** (1976), 1197–1198.

[6] C. Castaing, M. Valadier, *Convex Analysis and Measurable Multifunctions*, Springer-Verlag, New York, 1977.

[7] L.C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, vol. 19, Amer. Math. Soc., Providence, 1998.

[8] N. Ghoussoub, L. Tzou. *A variational principle for gradient flows*, (preprint 2003), To appear in Math. Annalen.

[9] N. Ghoussoub *Anti-self dual Lagrangians and variational principles for non Euler-Lagrange PDEs and evolution equations*, (2004). In preparation.

[10] R. Jordan, D. Kinderlehrer, F. Otto. *The variational formulation of the Fokker-Planck equation*, SIAM J. Math. Anal. **29** (1998), 1–17.

[11] F. Otto. *The geometry of dissipative evolution equations: the porous medium equation* Comm. Partial Differential Equations **26** (2001), 101–174.

[12] R. T. Rockafellar. *Existence and duality theorems for convex problems of Bolza*. Trans. Amer. Math. Soc. **159** (1971), 1–40.

Department of Mathematics, University of British Columbia, Vancouver BC Canada V6T 1Z2
E-mail address: nassif@math.ubc.ca

Department of Mathematics, University of Toronto, Toronto Ontario Canada M5S 3G3
E-mail address: mccann@math.toronto.edu