Annales de la Faculté des Sciences de Toulouse
MATHÉMATIQUES

VICTOR ISSA

Uniqueness of semi-concave weak Solutions for Hamilton–Jacobi Equations

Tome XXXIV, nº 4 (2025), p. 877–900.

https://doi.org/10.5802/afst.1825

Publication membre du centre
Mersenne pour l'édition scientifique ouverte
http://www.centre-mersenne.org/
e-ISSN : 2258-7519

Annales de la faculté des sciences de Toulouse Volume XXXIV, n° 4, 2025
pp. 877-900

Uniqueness of semi-concave weak Solutions for Hamilton–Jacobi Equations (*)

Victor Issa (1)

Abstract. — It is well known that when the nonlinearity is convex, the Hamilton–Jacobi PDE admits a unique semi-convex weak solution, which is the viscosity solution. In this paper, motivated by problems arising from spin glasses, we show that if the Hamilton–Jacobi PDE with strictly convex nonlinearity and regular enough initial condition admits a semi-concave weak solution, then this solution is the viscosity solution.

Résumé. — Il est bien connu que, lorsque la non-linéarité est convexe, l'équation d'Hamilton–Jacobi admet une unique solution faible semi-convexe et que cette solution est la solution de viscosité. Dans cet article, motivé par des problèmes liés aux verres de spin, nous montrons que si l'équation d'Hamilton–Jacobi avec non-linéarité strictement convexe et condition initiale régulière admet une solution faible semi-concave, alors celle-ci est la solution de viscosité.

1. Introduction

It is well known that the Hamilton–Jacobi equation with convex nonlinearity admits a unique semi-convex weak solution [2, 8], and that this unique solution is equal to the viscosity solution. More recently, it was shown that the Hamilton–Jacobi equation with arbitrary nonlinearity and smooth initial condition admits at most one convex weak solution [6] and that it is equal to the viscosity solution. Let $\mathcal{C}^{1,1}$ denote the space of differentiable functions $\mathbb{R}^d \to \mathbb{R}$ with Lipschitz gradient. In the present document, we prove an analog of those results for semi-concave weak solutions.

Theorem 1.1. — *Let $T^* \in [0, +\infty)$, $H : \mathbb{R}^d \to \mathbb{R}$ be a $\mathcal{C}^2$ strictly convex function and let $f : [0, T^*] \times \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz function. Assume that*

(*) Reçu le 14 mars 2024, accepté le 13 août 2024.
Keywords: Hamilton–Jacobi equations, viscosity solutions, weak solutions.
2020 *Mathematics Subject Classification:* 35F21, 35D30, 35D40.
(1) École Normale Supérieure de Lyon, Lyon, France — victor.issa@ens-lyon.fr
Article proposé par Radu Ignat.

$u_0 = f(0,\cdot)$ is $\mathcal{C}^{1,1}$, the function f satisfies $\partial_t f - H(\nabla_x f) = 0$ almost everywhere and there exists a positive constant $c > 0$ such that for all $t \in [0,T^*]$ the function $x \mapsto c|x|^2 - f(t,x)$ is convex. Then f coincides on $[0,T^*]\times\mathbb{R}^d$ with the unique viscosity solution of

$$\begin{cases} \partial_t u - H(\nabla_x u) = 0 \\ u(0,\cdot) = u_0. \end{cases} \tag{1.1}$$

If a function is both semi-concave and semi-convex on $\mathbb{R}^d$, then it is $\mathcal{C}^{1,1}$ [4, Corollary 3.3.8]. In Theorem 1.1, since f is equal to the viscosity solution, it is both semi-concave and semi-convex. As a consequence, f is $\mathcal{C}^{1,1}$ and so f is a classical solution of (1.1).

Let $\mathcal{P}_2(\mathbb{R}_+)$ be the set of Borel probability measures on $\mathbb{R}_+$ with finite second moment. In [12], it was pointed out that the limit free energy of the SK model is related to the viscosity solution of

$$\begin{cases} \partial_t p - \int (\partial_\mu p)^2 \mathrm{d}\mu = 0 & \text{on } \mathbb{R}_+ \times \mathcal{P}_2(\mathbb{R}_+), \\ p(0,\cdot) = \psi & \text{on } \mathcal{P}_2(\mathbb{R}_+), \end{cases} \tag{1.2}$$

where ψ is the cascade transform of the uniform measure on $\{-1,1\}$ [13, Definition 3.2]. More precisely, let p be the viscosity solution of (1.2), the Parisi formula [14, 15] at inverse temperature β for the SK model may be written [12, Theorem 1.1]

$$\lim_{N\to\infty} F_N(\beta) = -p(\beta^2/2, \delta_0) + \frac{\beta^2}{2} + \log 2. \tag{1.3}$$

According to this observation, it may be possible to prove the Parisi formula relying mainly on PDE tools. Devising such a proof was attempted in [13], but the method only yielded the following tight bound,

$$\lim_{N\to\infty} F_N(\beta) \leqslant -p(\beta^2/2, \delta_0) + \frac{\beta^2}{2} + \log 2. \tag{1.4}$$

One of the limitations that prevented a complete proof of the Parisi formula in [13] was the lack of a selection principle applicable in the context of spin glasses. We hope that the selection principle Theorem 1.1 may generalize to the Hamilton–Jacobi equations that arise when studying the SK model, as well as more general spin glass models with convex covariance structures. The PDE point of view proposed in [12] can also yield results in the context of non-convex spin glasses. For example, an analog of (1.4) has been proven this way for the bipartite model [11]. For this type of models, the free energy has yet to be rigorously identified. In Section 6, we give an example of a non-convex nonlinearity for which Theorem 1.1 fails. Some new ideas are needed, to further develop this method in the context of non-convex spin glasses.

Notation

For every $x \in \mathbb{R}^d$, we denote by x^i the i-th coordinate of x. Given $R > 0$, we denote by B_R the ball of center 0 and radius R in $\mathbb{R}^d$, with respect to the sup norm. Let $g : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}$ be a differentiable function and $t \in \mathbb{R}_+$, we denote by g^t the partial function $g(t, \cdot)$. The gradient of g with respect to the space variable x is denoted by $\nabla_x g$. In particular, $\nabla_x g : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$. The notation ∇ is reserved for the gradient with respect to all the variables. For example, here we have $\nabla g = (\partial_t g, \nabla_x g)$. We denote by $\mathcal{L}^d$ the Lebesgue measure on $\mathbb{R}^d$. If f is a Lipschitz function, we denote by $\mathrm{Lip}(f)$ the optimal Lipschitz constant of f. The push forward of a measure ν by a function f is written $f_*\nu$. When a measure ν_1 is absolutely continuous with respect to another measure, ν_2 we write $\nu_1 \ll \nu_2$. For $\varepsilon > 0$ we denote by η_ε a molifier, for example $\eta_\varepsilon(x) = \frac{1}{\varepsilon^d}\eta(\frac{x}{\varepsilon})$ where

$$\eta(x) = \begin{cases} C \exp\left(\frac{1}{|x|^2-1}\right) & \text{if } |x| < 1, \\ 0 & \text{otherwise,} \end{cases}$$

and C is the constant such that $\int \eta(x)\,\mathrm{d}x = 1$.

2. Proof Sketch

In this section, we will outline the strategy we will use to prove Theorem 1.1. The following seemingly weaker result actually implies Theorem 1.1.

THEOREM 2.1. — *Let $T^* \in [0, +\infty)$, $H : \mathbb{R}^d \to \mathbb{R}$ be a $\mathcal{C}^2$ strictly convex function and let $f : [0, T^*] \times \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz function. Assume that $u_0 = f(0, \cdot)$ is $\mathcal{C}^{1,1}$, the function f satisfies $\partial_t f - H(\nabla_x f) = 0$ almost everywhere and there exists a positive constant $c > 0$ such that for every $t \in [0, T^*]$ the function $x \mapsto c|x|^2 - f(t, x)$ is convex. Then, there exists $T \in (0, T^*]$ such that $f = u$ on $[0, T) \times \mathbb{R}^d$ where u is the viscosity solution of* (1.1).

Note that the hypotheses on H and u_0 in Theorems 1.1 and 2.1 guarantee the existence and uniqueness of a viscosity solution [2, Theorem 7.1]. Furthermore, the viscosity solution is Lipschitz and semi-convex [2, Theorems 8.2 & 8.3]. We will use those facts during the proofs of Theorems 1.1 and 2.1. Let us start by proving that Theorem 1.1 is a consequence of Theorem 2.1.

Proof of Theorem 1.1 using Theorem 2.1. — Suppose Theorem 2.1 is true, let u be the viscosity solution of (1.1) and define

$$T_0 = \sup\{T \in [0, T^*),\ f = u \text{ on } [0, T) \times \mathbb{R}^d\}.$$

Suppose $T_0 < T^*$, by definition of T_0 we have $f = u$ on $[0, T_0) \times \mathbb{R}^d$ and even on $[0, T_0] \times \mathbb{R}^d$ since both u and f are Lipschitz. Let $u_1 = f(T_0, \cdot\,) = u(T_0, \cdot\,)$, since $f(T_0, \cdot\,)$ is semi-concave and $u(T_0, \cdot\,)$ is semi-convex, the function u_1 is $\mathcal{C}^{1,1}$. Hence, $(t, x) \mapsto f(T_0 + t, x)$ satisfies the hypothesis of Theorem 2.1, so there exists $T > 0$ such that $f(T_0 + t, x) = u(T_0 + t, x)$ on $[0, T) \times \mathbb{R}^d$. Thus, $f = u$ on $[0, T_0 + T) \times \mathbb{R}^d$ which contradicts the definition of T_0. In conclusion, under Theorem 2.1 we must have $T_0 = T^*$, which means that $f = u$. $\square$

Let us now sketch the proof of Theorem 2.1. We will proceed in two steps, and use arguments involving characteristic curves. If g is a smooth solution of (1.1), then differentiating (1.1) with respect to the i-th space coordinate yields

$$\partial_t \partial_{x_i} g - \nabla H(\nabla_x g) \cdot \nabla_x(\partial_{x_i} g) = 0. \tag{2.1}$$

Hence, $\partial_{x_i} g$ solves a transport equation, with a vector field that depends on g. In particular, we may learn new information on g by studying the solutions of the associated characteristic ordinary differential equation $\dot{\phi} = -\nabla H(\nabla_x g(t, \phi))$.

The first step of the proof of Theorem 2.1 will be to show that there exists $T \in (0, T^*]$ such that $f \geqslant u$ on $[0, T) \times \mathbb{R}^d$. This is done by studying the behavior of f and u along the solutions of

$$\begin{cases} \dot{\phi} = -\nabla H(\nabla_x u(t, \phi)), \\ \phi(0) = x, \end{cases} \tag{2.2}$$

and showing that if ϕ is a solution of (2.2) then $t \mapsto (f - u)(t, \phi(t))$ is non-decreasing. The viscosity solution is $\mathcal{C}^{1,1}$ on $[0, T_0)$ for some $T_0 > 0$, so the existence of solutions to (2.2) in short time is not an issue for this step of the proof.

For the second step of the argument, we will use the same idea. However, this time we will show that $f - u$ is non-increasing along the solutions of

$$\begin{cases} \dot{\phi} = -\nabla H(\nabla_x f(t, \phi)), \\ \phi(0) = x. \end{cases} \tag{2.3}$$

Here the existence of solutions to (2.3) is not clear since the right-hand side is not fully Lipschitz. This kind of problem has already been investigated in [1, 3, 10]. We will give a fully self-contained existence argument, relying on ideas borrowed from [1, Theorem 6.2]. The precise construction is given in Section 4 and uses some results from the theory of Young measures (see Theorem 4.2). To construct a solution of (2.3), we will build solutions of problems that approximate (2.3) and have smooth right-hand sides. Then, using the semi-concavity of f, We will check that this sequence of solutions converges to a solution of (2.3) in an appropriate sense.

Once the existence of solutions is obtained, we can show that $f \leqslant u$ on the set E of points of $[0,T) \times \mathbb{R}^d$ reached by the solutions of (2.3) built previously. Finally, since we already know that $f \geqslant u$ on $[0,T) \times \mathbb{R}^d$, we will deduce that $f = u$ on E, and using uniqueness of solutions in short time for (2.2) we will deduce that $E = [0,T) \times \mathbb{R}^d$ which will end the proof of Theorem 2.1. This final part is treated in Section 5. Note, if c is a semi-concavity constant of f and $H(p) = |p|^2/2$, then the right-hand side of (2.3) satisfies,

$$(-\nabla_x f(t,y) - (-\nabla_x f(t,x))) \cdot (y-x) \geqslant -c|y-x|^2. \tag{2.4}$$

This means that the distance between two solutions of (2.3) started at x and y is lower bounded by $|x-y| \exp(\frac{-ct}{2})$. So, the solutions of (2.3) tend to repel each other.

3. Lower bound

There exists a short time $T_0 \in (0,T^*]$ such that u is $\mathcal{C}^{1,1}$ on $[0,T_0) \times \mathbb{R}^d$ [9, Proposition 12.1]. In this section, we will use this smoothness result to prove the existence of $T \in (0,T^*]$ such that $f \geqslant u$ on $[0,T) \times \mathbb{R}^d$. Define $a = -\nabla H(\nabla_x u)$ so that the space derivatives of u satisfy the following almost everywhere on $[0,T_0) \times \mathbb{R}^d$,

$$\partial_t(\partial_{x_i} u) + a \cdot \nabla \partial_{x_i} u = 0. \tag{3.1}$$

We are interested in the characteristic equation of this PDE, namely

$$\dot{\phi}(t) = a(t, \phi(t)). \tag{3.2}$$

Let us now define $T = \min(T_0, 1/\operatorname{Lip}(\nabla H(\nabla u_0)))$, for every $t < T$ and $x \in \mathbb{R}^d$, let $W(t,x) = x - t\nabla H(\nabla u_0(x))$. For every $t < T$, the map W^t is bi-Lipschitz and equal to $id_{\mathbb{R}^d}$ up to a Lipschitz perturbation with Lipschitz constant smaller than 1. From this, we deduce the following lemma, which will enable us to prove that $f - u$ is a decreasing function of time along W.

Lemma 3.1. — *On $[0,T)$, the push forward of $\mathcal{L}^{d+1}$ by the function $(t,x) \mapsto (t, W(t,x))$ is absolutely continuous with respect to $\mathcal{L}^{d+1}$ and for every $t < T$, the function W^t is surjective.*

Proof. — Define $W_0(t,x) = (t, W(t,x))$, let $B \subset [0,T) \times \mathbb{R}^d$ be a set of measure 0. We have $(W_0)^{-1}(B) = \{(t,x), (t, W(t,x)) \in B\}$, define $B^t =$

$\{x, (t,x) \in B\} \subset \mathbb{R}^d$. By Fubini's theorem, we have

$$\begin{aligned}\mathcal{L}^{d+1}((W_0)^{-1}(B)) &= \int_0^T \int_{\mathbb{R}^d} 1_B(t, W(t,x))\,\mathrm{d}x\mathrm{d}t \\ &= \int_0^T \mathcal{L}^d((W^t)^{-1}(B^t))\,\mathrm{d}t.\end{aligned}$$

On the other hand, we have $0 = \mathcal{L}^{d+1}(B) = \int_0^T \mathcal{L}^d(B^t)\,\mathrm{d}t$ so $\mathcal{L}^d(B^t) = 0$ t-almost everywhere. In particular, if we show that the pushforward of $\mathcal{L}^d$ by W^t is absolutely continuous with respect to $\mathcal{L}^d$ and that W^t is surjective, then the lemma is proven.

Since $t < T$, the function $x \mapsto W^t(x) - x$ is a contraction. Hence, it suffices to show the following. For every Lipschitz function $g : \mathbb{R}^d \to \mathbb{R}^d$ with Lipschitz constant $L < 1$ with respect to the sup norm, the function $h : x \mapsto x + g(x)$ is surjective and the pushforward of $\mathcal{L}^d$ by h is absolutely continuous with respect to $\mathcal{L}^d$.

Let $y \in \mathbb{R}^d$, the function $x \mapsto y - g(x)$ has a unique fixed point x^*, and by construction $h(x^*) = y$, so h is surjective. Now, we wish to show that $h_*\mathcal{L}^d \ll \mathcal{L}^d$. Let $|\cdot|_\infty$ denote the ℓ^∞ norm on $\mathbb{R}^d$, denote by $B(x,r) = x + (-r,r)^d$ the ball of center x and radius r in $\mathbb{R}^d$ with respect to $|\cdot|_\infty$. First, note that for every $x, y \in \mathbb{R}^d$ we have $|h(x) - h(y)|_\infty \geqslant |x-y|_\infty - |g(x) - g(y)|_\infty \geqslant (1-L)|x-y|_\infty$ with $1 - L > 0$. Since $|h(x) - h(y)|_\infty \geqslant (1-L)|x-y|_\infty$ we have $h^{-1}(B(h(x), r)) \subset B(x, \frac{1}{1-L})$. In particular, for every ball B using the surjectivity of h we have $\mathcal{L}^d(h^{-1}(B)) \leqslant \frac{1}{(1-L)^d}\mathcal{L}^d(B)$. Now let $A \subset \mathbb{R}^d$ be a set of measure 0. By definition of the outer Lebesgue measure, for every $\delta > 0$ there exists a sequence of balls $(B_n)_n$ such that, $A \subset \bigcup B_n$ and $\sum \mathcal{L}^d(B_n) < \delta$. So we have

$$\begin{aligned}\mathcal{L}^d(h^{-1}(A)) &\leqslant \sum_n \mathcal{L}^d(h^{-1}(B_n)) \\ &\leqslant \sum_n \frac{1}{(1-L)^d}\mathcal{L}^d(B_n) \\ &< \frac{\delta}{(1-L)^d}.\end{aligned}$$

This is true for all $\delta > 0$, so $\mathcal{L}^d(h^{-1}(A)) = 0$. Thus, $h_*\mathcal{L}^d \ll \mathcal{L}^d$. □

Lemma 3.2. — *For $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$, $W(\,\cdot\,, x)$ is the unique integral solution on $[0,T)$ of (3.2) with initial condition $\phi(0) = x$.*

Proof. — Since a is Lipschitz on $[0,T) \times \mathbb{R}^d$, the Picard–Lindelöf theorem applies and (3.2) admits a unique maximal solution. Let

$$D = \{(t,x) \in [0,T) \times \mathbb{R}^d, \nabla u \text{ is not differentiable at } (t,x)\}, \tag{3.3}$$

by Rademacher's theorem the set D is $\mathcal{L}^{d+1}$ negligible. According to Lemma 3.1, this means that the set

$$\{(t,x) \in [0,T) \times \mathbb{R}^d, \nabla u \text{ is not differentiable at } (t, W^t(x))\}$$

is also $\mathcal{L}^{d+1}$ negligible. In particular, for $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$, the function $t \mapsto \nabla u(t, W(t,x))$ is differentiable $\mathcal{L}^1$-almost everywhere on $[0,T)$. Let $t \in [0,T)$ such that ∇u is differentiable at $(t, W(t,x))$, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}&\nabla_x u(t, W(t,x))\\ &= \nabla_x u_t(t, W(t,x)) + a(t, W(t,x)) \cdot \nabla_x(\nabla_x u(t, W(t,x))) = 0. \quad (3.4)\end{aligned}$$

Since $t \mapsto \nabla_x u(t, W(t,x))$ is a Lipschitz function, it is equal to the integral of its derivative, so it is constant on $[0,T)$. It follows that, for every $t \in [0,T)$,

$$W(t,x) = x + \int_0^t \nabla H(\nabla u_0(x))\,\mathrm{d}\tau = x + \int_0^t \nabla H(\nabla_x u(\tau, W(\tau,x)))\,\mathrm{d}\tau. \quad (3.5)$$

Thus, $W(\,\cdot\,, x)$ is the unique integral solution of (3.2) on $[0,T)$. $\square$

Let us now study the behavior of f and u along the solutions of (3.2). By Rademacher's theorem, f and u are differentiable almost everywhere. According to Lemma 3.1, the push forward of $\mathcal{L}^{d+1}$ by $(t,x) \mapsto (t, W(t,x))$ is absolutely continuous with respect to $\mathcal{L}^{d+1}$. Hence, for $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$ the function $t \mapsto (f-u)(t, W(t,x))$ is differentiable $\mathcal{L}^1$-almost everywhere. Using the convexity of H the following holds $\mathcal{L}^{d+1}$-almost everywhere,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}&(f-u)(t, W(t,x))\\ &= (\partial_t(f-u) - \nabla H(\nabla_x u) \cdot (\nabla_x f - \nabla_x u))(t, W(t,x))\\ &= \big(H(\nabla_x f) - H(\nabla_x u) + \nabla H(\nabla_x u) \cdot (\nabla_x u - \nabla_x f)\big)(t, W(t,x))\\ &\geqslant 0.\end{aligned}$$

Since f, u and W are Lipschitz functions, the mapping $t \mapsto (f-u)(t, W(t,x))$ is equal to the integral of its derivative. So, we have $(f-u)(t, W^t(x)) \geqslant 0$ for every $t < T$ and $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$. By continuity of f, u and W this last bound is actually true for every $t < T$ and for every $x \in \mathbb{R}^d$. According to Lemma 3.1, for every $t < T$ the function W^t is surjective. In conclusion, we have proven the following lemma.

Lemma 3.3. — *Let f and H be functions satisfying the hypothesis of Theorem 2.1, and let u be the viscosity solution of* (1.1). *Let $T_0 > 0$ be such that $u \in \mathcal{C}^{1,1}([0,T_0) \times \mathbb{R}^d)$, and define $T = \min(T_0, 1/\operatorname{Lip}(\nabla H(\nabla u_0)))$. Then, for every $t < T$ and $x \in \mathbb{R}^d$, $f(t,x) \geqslant u(t,x)$.*

4. Construction of the characteristics

As explained in Section 2, if we define $b = -\nabla H(\nabla_x f)$ and if f is twice differentiable, we have for all i,

$$\partial_t(\partial_{x_i} f) + b \cdot \nabla_x(\partial_{x_i} f) = 0. \tag{4.1}$$

This means that, if they exist, the derivatives of f each satisfy a transport equation with a vector field b that depends on f. The aim of this section is to construct appropriate solutions of the associated characteristic equation,

$$\dot{\phi}(t) = b(t, \phi(t)). \tag{4.2}$$

Note that this equation can be studied even if f is not twice differentiable, as there are only derivatives of order at most one that appear in the definition of the vector field b and those always exist, at least almost everywhere, since f is Lipschitz. We define $f_\varepsilon(t,x) = (f^t * \eta_\varepsilon)(x)$ and $b_\varepsilon = -\nabla H(\nabla_x f_\epsilon)$, so that b_ε is a smooth approximation of b. Furthermore, if ϕ_ε is a maximal solution of

$$\begin{cases} \dot{\phi}_\varepsilon(t) = b_\varepsilon(t, \phi(t)), \\ \phi_\varepsilon(0) = x. \end{cases} \tag{4.3}$$

Then, $|\phi_\varepsilon(t)| \leqslant x + t\|b_\varepsilon\|_{L^\infty([0,T^*]\times\mathbb{R}^d)}$, so the maximal solution of (4.3) does not blow up in finite time, hence it is defined on $[0, T^*)$. Define $X_\varepsilon(t,x)$ as the value at $t \in [0, T^*)$ of the solution of (4.3) started at $x \in \mathbb{R}^d$. We will show that the sequence $(X_\varepsilon)_\varepsilon$ converges, in some sense, to a probability measure on the set of solution of (4.2) as $\varepsilon \to 0$. To do so, we will use the theory of Young measures. We start by showing that the X_ε^t's do not compress too much the size of measurable sets.

To proceed, we will use the fact that the divergence of the vector field b_ε is bounded from below, this is a consequence of the fact that f is semi-concave and the following computation. Start by noting that,

$$\nabla_x b_\varepsilon^t = -\nabla^2 H(\nabla_x f_\varepsilon^t)\nabla_x^2 f_\varepsilon^t,$$

and according to the hypothesis of Theorem 2.1, there exists a positive constant c_0 such that $x \mapsto c_0|x|^2 - f(t,x)$ is convex for all $t \geqslant 0$. So, for all $\varepsilon > 0$, $\nabla_x^2 f_\varepsilon \leqslant c_0$. The function H is convex, so $\nabla^2 H(\nabla_x f_\varepsilon^t(x))$ is a positive semi-definite matrix for all x. Let $M(x)$ denote the square root of this matrix, we have

$$\begin{aligned} \operatorname{div} b_\varepsilon^t(x) &= \operatorname{Tr}(\nabla_x b_\varepsilon^t(x)) \\ &= \operatorname{Tr}(-\nabla^2 H(\nabla_x f_\varepsilon^t(x))\nabla_x^2 f_\varepsilon^t(x)) \\ &= \operatorname{Tr}(M(x)(c_0 - \nabla_x^2 f_\varepsilon^t(x))M(x)^*) - c_0 \operatorname{Tr}(\nabla^2 H(\nabla_x f_\varepsilon^t(x))) \\ &\geqslant 0 - c_0 \Delta H(\nabla_x f_\varepsilon^t(x)). \end{aligned}$$

Thus setting $c = c_0 \|\Delta H\|_{L^\infty(B(0,\mathrm{Lip}(f)))}$, we have

$$\operatorname{div} b_\varepsilon^t(x) \geqslant -c. \tag{4.4}$$

In fact the results of this section hold as long as (4.4) holds, and asking for f to be semi-concave is merely a way of imposing this condition.

Recall that given $R \in (0,\infty]$, B_R denotes the ball of center 0 and radius R in $\mathbb{R}^d$ with respect to the sup norm and with the understanding that $B_\infty = \mathbb{R}^d$.

Lemma 4.1. — *There exists a constant $c > 0$ depending on H and f such that, for every $R \in (0,\infty]$, $\varepsilon > 0$, $t \in [0, T^*)$ and non-negative continuous function $\beta : \mathbb{R}^d \to \mathbb{R}$,*

$$\int_{B_R} \beta(X_\varepsilon^t(x))\,\mathrm{d}x \leqslant e^{ct} \int_{B_{R+t|b|_\infty}} \beta(x)\,\mathrm{d}x. \tag{4.5}$$

Furthermore, for every measurable set $A \subset \mathbb{R}^d$, we have $\mathcal{L}^d((X_\varepsilon^t)^{-1}(A)) \leqslant e^{ct}\mathcal{L}^d(A)$ and $(X_\varepsilon)_*\mathcal{L}^d \ll \mathcal{L}^d$.

Proof. — A change of variable yields

$$\int_{B_R} \beta(X_\varepsilon^t(x))\,\mathrm{d}x = \int_{X_\varepsilon^t(B_R)} \beta(x) \det(\nabla_x X_\varepsilon^t((X_\varepsilon^t)^{-1}(x)))^{-1}\,\mathrm{d}x. \tag{4.6}$$

We have $X_\varepsilon^t(B_R) \subset B_{R+t|b|_\infty}$, so it is enough to show that $\det(\nabla_x X_\varepsilon^t(x))^{-1} \leqslant e^{ct}$. It follows from the definition of X_ε that

$$\frac{\mathrm{d}}{\mathrm{d}t} \nabla_x X_\varepsilon^t(x) = \nabla_x b_\varepsilon^t(X_\varepsilon^t(x)) \nabla_x X_\varepsilon^t(x). \tag{4.7}$$

If $A(t)$ is a matrix that is a differentiable function of a parameter t, then $\det(A(t))$ is a differentiable function of t and

$$\frac{\mathrm{d}}{\mathrm{d}t} \det(A(t)) = \mathrm{Tr}(\mathrm{adj}(A(t)) \frac{\mathrm{d}}{\mathrm{d}t} A(t)), \tag{4.8}$$

where $\mathrm{adj}(A(t))$ is the adjugate of $A(t)$ (transpose of the cofactor matrix), in particular $A(t)\,\mathrm{adj}(A(t)) = \det(A(t)) I_d$. Applying this to the matrix $A(t) = \nabla_x X_\varepsilon^t(x)$ yields

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} \det(\nabla_x X_\varepsilon^t(x)) &= \mathrm{Tr}(\mathrm{adj}(\nabla_x X_\varepsilon^t(x)) \nabla_x b_\varepsilon^t(X_\varepsilon^t(x)) \nabla_x X_\varepsilon^t(x)) \\
&= \mathrm{Tr}(\nabla_x X_\varepsilon^t(x)\,\mathrm{adj}(\nabla_x X_\varepsilon^t(x)) \nabla_x b_\varepsilon^t(X_\varepsilon^t(x))) \\
&= \det(\nabla_x X_\varepsilon^t(x)) \div b_\varepsilon^t(X_\varepsilon^t(x)).
\end{aligned}$$

Hence, $\det(\nabla_x X_\varepsilon^t(x)) = \exp\Big(\int_0^t \operatorname{div} b_\varepsilon^s(X_\varepsilon^s(x))\,\mathrm{d}s\Big)$. Furthermore, according to (4.4) the divergence of b_ε^t is bounded from below by a constant. In conclusion,

$$\det(\nabla_x X_\varepsilon^t(x))^{-1} = \exp\left(-\int_0^t \operatorname{div} b_\varepsilon^s(X_\varepsilon^s(x))\,\mathrm{d}s\right) \leqslant e^{ct}.$$

Thus, the desired bound is proven. The second inequality, $\mathcal{L}^d((X_\varepsilon^t)^{-1}(A)) \leqslant e^{ct}\mathcal{L}^d(A)$ is obtained by taking continuous positive approximations of 1_A in (4.5) and using dominated convergence. In particular, if $\mathcal{L}^d(A) = 0$ then $\mathcal{L}^d((X_\varepsilon^t)^{-1}(A)) = 0$, so $(X_\varepsilon)_*\mathcal{L}^d \ll \mathcal{L}^d$. □

We will now tackle the problem of showing that the sequence (X_ε) converges up to extraction. Let E be a topological space, we denote by $\mathcal{P}(E)$ the set of Borel probability measures on E. We recall the following basic result on the theory of Young measures [1, Theorem 6.5].

Theorem 4.2. — *Let K be a compact metric space and $A \subset \mathbb{R}^d$ be a measurable set. Let $(\eta_h)_h$ be a sequence of measurable maps $A \to \mathcal{P}(K)$. There exists an extraction $(\varepsilon_k)_k$ and a measurable map $\eta : A \to \mathcal{P}(K)$ such that the following holds. For every bounded function $\phi = \phi(x,u) : A \times K \to \mathbb{R}$ that is continuous with respect to u and measurable with respect to x, we have*

$$\lim_{k\to\infty} \int_A \int_K \phi(x,u)\,\mathrm{d}\eta_{\varepsilon_k,x}(u)\,\mathrm{d}x = \int_A \int_K \phi(x,u)\,\mathrm{d}\eta_x(u)\,\mathrm{d}x. \tag{4.9}$$

Fix $T \in (0,T^*)$, and define $\Gamma = \mathcal{C}^0([0,T],\mathbb{R}^d)$. By the Arzelà–Ascoli theorem, the closure $\mathcal{K}$ of $\{X_\varepsilon(\,\cdot\,,x), x \in B(0,R), \varepsilon > 0\}$ in Γ is a compact set for the topology of uniform convergence. For every $\varepsilon > 0$, consider

$$\eta_\varepsilon : \begin{cases} \mathbb{R}^d \longrightarrow \mathcal{P}(\mathcal{K}) \\ x \longmapsto \delta_{X_\varepsilon(\,\cdot\,,x)}. \end{cases} \tag{4.10}$$

By Theorem 4.2, up to extraction, the sequence $(\eta_\varepsilon)_\varepsilon$ converges to some measurable map $\eta : \mathbb{R}^d \to \mathcal{P}(\mathcal{K})$. More precisely, there exists a sequence $(\varepsilon_k)_k$ such that for every bounded function $\phi = \phi(x,\gamma) : \mathbb{R}^d \times \Gamma \to \mathbb{R}$ that is continuous in γ and measurable in x, we have

$$\lim_{k\to\infty} \int_{\mathbb{R}^d} \phi(x, X_{\varepsilon_k}(\,\cdot\,,x))\,\mathrm{d}x = \int_{\mathbb{R}^d} \int_\Gamma \phi(x,\gamma)\,\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x. \tag{4.11}$$

The remainder of this section is devoted to showing that for almost all $x \in \mathbb{R}^d$, the probability measure η_x is well-behaved and concentrated on the set of integral solutions of (4.2) with initial condition x.

Lemma 4.3. — *Let $c > 0$ be the constant of Lemma 4.1. For every $t \in [0, T]$, $R \in (0, \infty]$ and non-negative continuous bounded map $\beta : \mathbb{R}^d \to \mathbb{R}$, we have*

$$\int_{B_R} \int_{\Gamma} \beta(\gamma(t))\,\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x \leqslant e^{ct} \int_{B_{M_R}} \beta(x)\,\mathrm{d}x, \tag{4.12}$$

where $M_R = R + T|b|_\infty$. Furthermore, if $A \subset \mathbb{R}^d$ is a measurable set, then

$$\int_{\mathbb{R}^d} \int_{\Gamma} 1_A(\gamma(t))\,\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x \leqslant e^{ct} \mathcal{L}^d(A). \tag{4.13}$$

Proof. — Consider, $\phi(x, \gamma) = \beta(\gamma(t))$, then using Lemma 4.1 we have,

$$\int_{B_R} \phi(x, X_\varepsilon(\,\cdot\,, x))\,\mathrm{d}x = \int_{B_R} \beta(X_\varepsilon(t, x))\,\mathrm{d}x \leqslant e^{ct} \int_{B_{R+t|b|_\infty}} \beta(x)\,\mathrm{d}x. \tag{4.14}$$

Letting $\varepsilon \to 0$ along the subsequence $(\varepsilon_k)_k$ we obtain (4.12). Let $A \subset \mathbb{R}^d$ be a measurable set, still according to Lemma 4.1, we have

$$\int_{\mathbb{R}^d} \int_{\Gamma} 1_A(\gamma(t))\,\mathrm{d}\eta_{\varepsilon,x}\,\mathrm{d}x = \int_{\mathbb{R}^d} 1_A(X_\varepsilon^t(x))\,\mathrm{d}x \leqslant e^{ct} \mathcal{L}^d(A). \tag{4.15}$$

Again, letting $\varepsilon \to 0$ along the subsequence $(\varepsilon_k)_k$ we obtain (4.13). □

For $x \in \mathbb{R}^d$, let Γ_x^b denote the set of integral solutions of (4.2) with initial condition x, that is

$$\Gamma_x^b = \{\gamma \in \Gamma, \forall\, t \in [0, T], \gamma(t) = x + \int_0^t b(\tau, \gamma(\tau))\,\mathrm{d}\tau\}. \tag{4.16}$$

Lemma 4.4. — *For $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$, we have $\eta_x(\Gamma - \Gamma_x^b) = 0$.*

Proof. — Fix $t \in [0, T]$ and let $c : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ be a bounded smooth vector field. Define

$$\phi^c(x, \gamma) = \left|\gamma(t) - x - \int_0^t c(\tau, \gamma(\tau))\,\mathrm{d}\tau\right|. \tag{4.17}$$

Let $\varepsilon > 0$ and $R > 0$, using Fubini's theorem and Lemma 4.1, we have

$$\begin{aligned}
\int_{B_R} \phi^c(x, X_\varepsilon(\,\cdot\,, x))\,\mathrm{d}x &\leqslant \int_{B_R} \left|X_\varepsilon(t, x) - x - \int_0^t c(\tau, X_\varepsilon(\tau, x))\,\mathrm{d}\tau\right| \mathrm{d}x \\
&\leqslant \int_{B_R} \int_0^t |b_\varepsilon - c|(\tau, X_\varepsilon(\tau, x))\,\mathrm{d}\tau\mathrm{d}x \\
&\leqslant \int_0^t \int_{B_R} |b_\varepsilon - c|(\tau, X_\varepsilon(\tau, x))\,\mathrm{d}x\mathrm{d}\tau \\
&\leqslant \int_0^t e^{c\tau} \int_{B_{M_R}} |b_\varepsilon - c|(\tau, x)\,\mathrm{d}x\mathrm{d}\tau \\
&\leqslant e^{ct} \int_0^t |b_\varepsilon^\tau - c^\tau|_{L^1(B_{M_R})}\mathrm{d}\tau.
\end{aligned}$$

By dominated convergence, letting $\varepsilon \to 0$ along the subsequence $(\varepsilon_k)_k$ and using the definition of the map η, we obtain

$$\int_{B_R} \int_\Gamma \phi^c(x,\gamma)\,\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x \leqslant e^{ct} \int_0^t |b^\tau - c^\tau|_{L^1(B_{M_R})} \mathrm{d}\tau. \tag{4.18}$$

Taking $c = b_\varepsilon$ leads to

$$\lim_{\varepsilon \to 0} \int_{B_R} \int_\Gamma \phi^{b_\varepsilon}(x,\gamma)\,\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x = 0. \tag{4.19}$$

By Fatou's lemma,

$$\int_{B_R} \liminf_{\varepsilon \to 0} \int_\Gamma \phi^{b_\varepsilon}(x,\gamma)\,\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x = 0. \tag{4.20}$$

So, for $\mathcal{L}^d$-almost every $x \in B_R$,

$$\liminf_{\varepsilon \to 0} \int_\Gamma \phi^{b_\varepsilon}(x,\gamma)\,\mathrm{d}\eta_x(\gamma) = 0. \tag{4.21}$$

Letting $R \to \infty$ along a countable set, we obtain that (4.21) is true for $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$.

Let $N \subset (0,T)\times\mathbb{R}^d$ be the complement of $\overline{N} = \{(t,x), \lim_{k\to\infty} b_{\varepsilon_k}(t,x) = b(t,x)\}$. For every $t \in (0,T)$, define $N_t = \{x \in \mathbb{R}^d, (t,x) \in N\}$. The set N is $\mathcal{L}^{d+1}$-negligible, so for $\mathcal{L}^1$-almost every $t \in (0,T)$, $\mathcal{L}^d(N_t) = 0$. For every $t \in (0,T)$ such that, $\mathcal{L}^d(N_t) = 0$, according to Lemma 4.3 we have $\int_{\mathbb{R}^d} \int_\Gamma 1_{N_t}(\gamma(t))\,\mathrm{d}\eta_x(\gamma) dx = 0$. Therefore,

$$\int_0^T \int_{\mathbb{R}^d} \int_\Gamma 1_{N_t}(\gamma(t))\,\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x\mathrm{d}t = 0. \tag{4.22}$$

By Fubini's theorem,

$$\int_{\mathbb{R}^d} \int_\Gamma \int_0^T 1_{N_t}(\gamma(t))\,\mathrm{d}t\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x = 0. \tag{4.23}$$

So for $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$,

$$\eta_x\left(\left\{\gamma \in \Gamma, \int_0^T 1_{N_t}(\gamma(t))\,\mathrm{d}t = 0\right\}\right) = 1. \tag{4.24}$$

Now, let $x \in \mathbb{R}^d$ be such that (4.21) and (4.24) are satisfied, and let $\gamma \in \Gamma$ be a point in the support of η_x such that $\int_0^T 1_{N_t}(\gamma(t))\,\mathrm{d}t = 0$. Then, for $\mathcal{L}^1$-almost every $t \in (0,T)$, we have $\gamma(t) \notin N_t$, thus for $\mathcal{L}^1$-almost every $t \in (0,T)$,

$$\lim_{k\to\infty} b_{\varepsilon_k}(t,\gamma(t)) = b(t,\gamma(t)). \tag{4.25}$$

By dominated convergence, it follows that

$$\lim_{k\to\infty}\int_0^T b_{\varepsilon_k}(t,\gamma(t))\,\mathrm{d}t=\int_0^T b(t,\gamma(t))\,\mathrm{d}t. \tag{4.26}$$

And thus,

$$\lim_{k\to\infty}\left|\gamma(t)-x-\int_0^t b_{\varepsilon_k}(\tau,\gamma(\tau))\,\mathrm{d}\tau\right|=\left|\gamma(t)-x-\int_0^t b(\tau,\gamma(\tau))\,\mathrm{d}\tau\right|. \tag{4.27}$$

Let $(\varepsilon_{k(l)})_l$ be a subsequence of (ε_k) along which the lim inf in (4.21) is reached. Combining (4.21) with (4.27), and taking the limit as $l\to\infty$, we obtain

$$\int_\Gamma\left|\gamma(t)-x-\int_0^t b(\tau,\gamma(\tau))\,\mathrm{d}\tau\right|\mathrm{d}\eta_x(\gamma)=0. \tag{4.28}$$

In conclusion, for $\mathcal{L}^d$-almost every $x\in\mathbb{R}^d$,

$$\eta_x\left(\left\{\gamma\in\Gamma,\gamma(t)=x+\int_0^t b(\tau,\gamma(\tau))\,\mathrm{d}\tau\right\}\right)=1. \tag{4.29}$$

Thus, for $\mathcal{L}^d$-almost every $x\in\mathbb{R}^d$,

$$\eta_x\left(\left\{\gamma\in\Gamma,\forall\, t\in[0,T]\cap\mathbb{Q},\gamma(t)=x+\int_0^t b(\tau,\gamma(\tau))\,\mathrm{d}\tau\right\}\right)=1. \tag{4.30}$$

Since, the functions in Γ are continuous with respect to t and $[0,T]\cap\mathbb{Q}$ is dense in $[0,T]$, we have,

$$\Gamma_x^b=\left\{\gamma\in\Gamma,\forall\, t\in[0,T]\cap\mathbb{Q},\gamma(t)=x+\int_0^t b(\tau,\gamma(\tau))\,\mathrm{d}\tau\right\}. \tag{4.31}$$

Finally, we obtain $\eta_x(\Gamma-\Gamma_x^b)=0$ for $\mathcal{L}^d$-almost every $x\in\mathbb{R}^d$. □

5. Proof of the main result

The goal of this section is to use the Young measure η built in Section 4 to complete the proof of Theorem 2.1. Recall that in Section 3 we have already shown the existence of $T\in(0,T^*)$ such that $f\geqslant u$ on $[0,T)\times\mathbb{R}^d$. Also recall that u is $\mathcal{C}^{1,1}$ on $[0,T)\times\mathbb{R}^d$ and W^t is a surjective function for $t<T$. Here we will try to prove the converse inequality $f\leqslant u$ using an argument similar to the one used to prove Lemma 3.3 in which the vector field W will be replaced by η. This approach will only lead to a bound valid on the image of $(t,x)\mapsto(t,\gamma(t))$ for $\mathcal{L}^d$-almost every $x\in\mathbb{R}^d$ and η_x-almost every $\gamma\in\Gamma_x^b$. But, using this bound, we will be able to show that $\eta_x=\delta_{W(\,\cdot\,,x)}$, except on

a negligible set. Then, using the surjectivity of W^t for $t < T$, we can finish the proof of Theorem 2.1.

Lemma 5.1. — *For $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$, and η_x-almost every $\gamma \in \Gamma_x^b$, we have*

$$\forall\, t \in [0,T],\ f(t,\gamma(t)) = u(t,\gamma(t)). \tag{5.1}$$

Proof. — Fix $t \in (0,T)$, let

$$\begin{aligned}
\mathcal{D}_t &= \left\{x \in \mathbb{R}^d, f \text{ is not differentiable at } (t,x)\right\} \subset \mathbb{R}^d\\
A_t &= \left\{(x,\gamma) \in \mathbb{R}^d \times \Gamma, f \text{ is not differentiable at } (t,\gamma(t))\right\} \subset \mathbb{R}^d \times \Gamma.
\end{aligned}$$

By Rademacher's theorem, for $\mathcal{L}^1$-almost every $t \in (0,T)$, the set $\mathcal{D}_t$ is $\mathcal{L}^d$-negligible. Hence, according to Lemma 4.3, for $\mathcal{L}^1$-almost every $t \in (0,T)$, the following holds,

$$\int_{\mathbb{R}^d}\int_\Gamma 1_{A_t}(x,\gamma)\,\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x = \int_{\mathbb{R}^d}\int_\Gamma 1_{\mathcal{D}_t}(\gamma(t))\,\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x = 0. \tag{5.2}$$

Integrating (5.2) with respect to $t \in (0,T)$ and invoking Fubini's theorem, we obtain

$$\int_{\mathbb{R}^d}\int_\Gamma\int_0^T 1_{A_t}(x,\gamma)\,\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x = \int_0^t\int_{\mathbb{R}^d}\int_\Gamma 1_{\mathcal{D}_t}(\gamma(t))\,\mathrm{d}t\mathrm{d}\eta_x(\gamma)\,\mathrm{d}x = 0. \tag{5.3}$$

So, for $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$ and η_x-almost every $\gamma \in \Gamma_x^b$, $\int_0^T 1_{A_t}(x,\gamma)\,\mathrm{d}t = 0$. In particular, for $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$ and η_x-almost every $\gamma \in \Gamma_x^b$, the function $t \mapsto (f-u)(t,\gamma(t))$ is differentiable $\mathcal{L}^1$-almost everywhere. Fix $x \in \mathbb{R}^d$ and $\gamma \in \Gamma_x^b$ such that for $\mathcal{L}^1$-almost every $t \in (0,T)$, f is differentiable at $(t,\gamma(t))$. We have, for $\mathcal{L}^1$-almost every $t \in (0,T)$,

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}(f-u)(t,\gamma(t))\\
&\qquad= \big(\partial_t f - \partial_t u + \gamma'(t)\cdot(\nabla f - \nabla u)(t,\gamma(t))\big)\\
&\qquad= H(\nabla f) - H(\nabla u) - \nabla H(\nabla f)\cdot(H(\nabla f) - H(\nabla u))(t,\gamma(t))\\
&\qquad= -(T_{\nabla f}H(\nabla u) - H(\nabla u))(t,\gamma(t))\\
&\qquad\leqslant 0,
\end{aligned}$$

where $T_pH(q) = H(p) + \nabla H(p)\cdot(q-p)$ is the tangent of H at p evaluated at q. For this choice of x and γ, the function $t \mapsto (f-u)(t,\gamma(t))$ vanishes at zero and is non-increasing. In addition, according to Lemma 3.3 the function $t \mapsto (f-u)(t,\gamma(t))$ is non-negative. Therefore, for every $t \in (0,T)$, $f(t,\gamma(t)) = u(t,\gamma(t))$. $\square$

LEMMA 5.2. — *For $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$, η_x-almost every $\gamma \in \Gamma_x^b$, and $\mathcal{L}^1$-almost every $t \in (0,T)$, the function f is differentiable at $(t,\gamma(t))$ and we have*

$$\nabla f(t,\gamma(t)) = \nabla u(t,\gamma(t)). \tag{5.4}$$

Proof. — Let $x \in \mathbb{R}^d$ and $\gamma \in \Gamma_x^b$ such that the previous lemma holds. The function $t \mapsto (f-u)(t,\gamma(t))$ is constant equal to zero, therefore it is differentiable, and its derivative is equal to zero. In addition, inspecting the proof of the previous lemma, we see that for $\mathcal{L}^1$-almost all $t \in (0,T)$, f and u are differentiable at $(t,\gamma(t))$. For every $t \in (0,T)$ such that f and u are differentiable at $(t,\gamma(t))$, we have

$$0 = \frac{\mathrm{d}}{\mathrm{d}t}(f-u)(t,\gamma(t)) = -(T_{\nabla f}H(\nabla u) - H(\nabla u))(t,\gamma(t)). \tag{5.5}$$

By strict convexity of H, we obtain $\nabla f(t,\gamma(t)) = \nabla u(t,\gamma(t))$. ☐

Proof of Theorem 2.1. — Let $x \in \mathbb{R}^d$ and $\gamma \in \Gamma_x^b$ be such that Lemma 5.2 holds. We have for every $t \in (0,T)$,

$$\begin{aligned}\gamma(t) &= x - \int_0^t \nabla H(\nabla f(\tau,\gamma(\tau)))\,\mathrm{d}\tau \\ &= x - \int_0^t \nabla H(\nabla u(\tau,\gamma(\tau)))\,\mathrm{d}\tau.\end{aligned}$$

Therefore γ is an integral solution of (3.2) with initial condition x, so by Lemma 3.2 we have $\gamma = W(\,\cdot\,,x)$. So, for $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$, η_x is supported on $\{W(\,\cdot\,,x)\}$, this means that $\eta_x = \delta_{W(\,\cdot\,,x)}$. Applying Lemma 5.1 we obtain, for $\mathcal{L}^d$-almost every $x \in \mathbb{R}^d$ and every $t \in (0,T)$ that,

$$f(t,W(t,x)) = u(t,W(t,x)). \tag{5.6}$$

By continuity of W, f and u, the identity in the above display is actually true for every $x \in \mathbb{R}^d$. Finally, according to Lemma 3.1, for every $t \in [0,T)$ the map $W(t,\cdot\,)$ is surjective, thus $f = u$ on $[0,T) \times \mathbb{R}^d$. ☐

6. A counter-example for non-convex nonlinearities

The aim of this section is to show that when the nonlinearity is not assumed to be convex, Theorem 1.1 no longer holds. To do so, we will consider the conservation law naturally associated with (1.1). In order to keep our notation consistent with the literature on conservation laws, we choose to write this equation with the + sign convention rather than the possibly more natural choice of writing it with the − sign convention. We start by stating a corollary of Theorem 1.1. We say that a function $\phi : \mathbb{R} \to \mathbb{R}$ is one-sided Lipschitz when there exists $c \geqslant 0$ such that, for all $x \geqslant y$:

$$\phi(x) - \phi(y) \geqslant -c(x-y). \tag{6.1}$$

One-sided Lipschitz functions are exactly the functions whose anti-derivatives are semi-concave.

Corollary 6.1. — *Let $T^* \in [0,+\infty)$, $H : \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^2$ strictly convex function and let $v : [0,T^*] \times \mathbb{R} \to \mathbb{R}$ be a bounded function. Assume that the function $v_0 = v(0,\cdot\,)$ is Lipschitz, that v is differentiable almost everywhere, satisfies $\partial_t v - \partial_x(H(v)) = 0$ almost everywhere, and that there exists a positive constant $c > 0$ such that for all $t \in [0,T^*]$ the function $v(t,\cdot\,)$ is one-sided Lipschitz with constant c. Then v coincides on $[0,T^*] \times \mathbb{R}$ with the unique entropy solution of*

$$\begin{cases} \partial_t v + \partial_x(H(v)) = 0, \\ v(0,\cdot\,) = v_0. \end{cases} \tag{6.2}$$

Proof. — There is a formal correspondence between the solutions of (6.2) and the solutions of the following Hamilton–Jacobi initial value problem,

$$\begin{cases} \partial_t u - \overline{H}(\partial_x u) = 0, \\ u(0,\cdot\,) = u_0, \end{cases} \tag{6.3}$$

where $\overline{H}(p) = H(-p)$. The correspondence is given by $v \mapsto -\int_{-\infty}^{\cdot} v(\,\cdot\,,y)\,\mathrm{d}y$ and $u \mapsto -\partial_x u$. This correspondence sends the viscosity solution of (6.3) to the entropy solution of (6.2) and vice versa [9, Theorem 16.1], [7, Section 2]. Hence, if v satisfies the hypothesis of the corollary, then the function $u : (t,x) \mapsto -\int_{-\infty}^{x} v(t,y)\,\mathrm{d}y$ satisfies the hypotheses of Theorem 1.1 with respect to the equation (6.3), where $u_0(x) = -\int_{-\infty}^{x} v(0,y)\,\mathrm{d}y$. In particular u is the viscosity solution of (6.3) and so v is the entropy solution of (6.2). $\square$

To construct a counter-example to Theorem 1.1 in the absence of convexity, we will construct a Lipschitz initial condition $v_0 : \mathbb{R} \to \mathbb{R}$, a $\mathcal{C}^2$ non-convex function $H : \mathbb{R} \to \mathbb{R}$ and $v : [0,T^*] \times \mathbb{R} \to \mathbb{R}$ a bounded and one-sided Lipschitz almost everywhere solution of (6.2) such that v is not the entropy solution of (6.2).

We define a non-convex even function H by

$$H(p) = \begin{cases} \frac{5}{4}p^3 + \frac{19}{8}p^2 + \frac{15}{16}p + \frac{5}{32} & \text{if } p \leqslant -\frac{1}{2} \\ \frac{1}{2}p^2 & \text{if } -\frac{1}{2} \leqslant p \leqslant \frac{1}{2} \\ -\frac{5}{4}p^3 + \frac{19}{8}p^2 - \frac{15}{16}p + \frac{5}{32} & \text{if } \frac{1}{2} \leqslant p, \end{cases} \tag{6.4}$$

and a Lipschitz initial condition

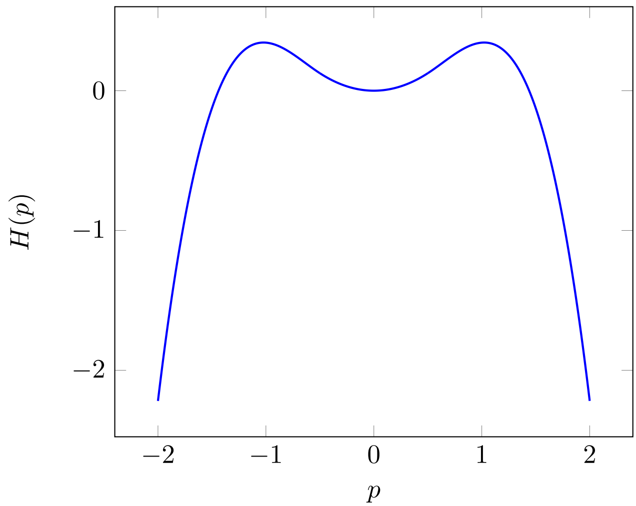

Figure 6.1. Graph of the non-convex function H

$$v_0(x) = \begin{cases} -\frac{3}{2} & \text{if } x \leqslant -\frac{3}{2} \\ x & \text{if } -\frac{3}{2} \leqslant x \leqslant \frac{3}{2} \\ \frac{3}{2} & \text{if } \frac{3}{2} \leqslant x \leqslant L \\ -x+L+\frac{3}{2} & \text{if } L \leqslant x \leqslant L+1 \\ \frac{1}{2} & \text{if } L+1 \leqslant x. \end{cases} \tag{6.5}$$

The exact value of the parameter L is to be fixed later, for now we imagine that L is large, for example $L > 30 \times H'(-\frac{3}{2})$, so that the effect of the asymmetry of v_0 is felt close to the origin by the solution only after a long enough time.

Proposition 6.2. — *There exists $L \geqslant \frac{3}{2}$ and a bounded function $v : [0,T^*] \times \mathbb{R} \to \mathbb{R}$ such that the following holds.*

(1) *$v(0,\cdot)$ is equal to the function v_0 defined by* (6.5)*.*
(2) *v is differentiable almost everywhere and satisfies $\partial_t v - \partial_x(H(v)) = 0$ almost everywhere, where H is the function defined by* (6.4)*.*
(3) *There exists $c > 0$ such that for all $t \in [0,T^*]$ the function $v(t,\cdot)$ is one-sided Lipschitz with constant c.*
(4) *v is* not *the entropy solution of* (6.2)*.*

If g is a piecewise continuous function whose discontinuities lie along a curve $\Gamma = \{(z(t),t)\}$, then the jump of g along Γ is denoted $[g]$, it is

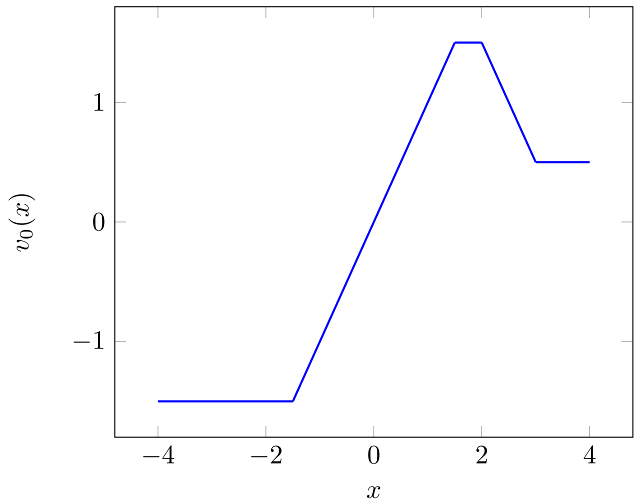

Figure 6.2. Graph of the Lipschitz initial condition v_0 for $L = 2$

defined by

$$[g] = \lim_{\omega_+} g - \lim_{\omega_-} g, \tag{6.6}$$

where $\omega_- = \{x < z(t)\}$ and $\omega_+ = \{x > z(t)\}$ for simplicity we write $g_\pm$ instead of $\lim_{\omega_\pm} g$.

Proof of Proposition 6.2. — The characteristic curves $(t, X^t(x))$ of (6.2) satisfy $X^t(x) = x + tH'(v_0(x))$. Setting $a = H'(-\frac{3}{2}) = -H'(\frac{3}{2}) > 0$, explicitly we have,

$$X^t(x) = \begin{cases} x + ta & \text{if } x \leqslant -\frac{3}{2} \\ x + tH'(x) & \text{if } -\frac{3}{2} \leqslant x \leqslant -\frac{1}{2} \\ x + tx & \text{if } -\frac{1}{2} \leqslant x \leqslant \frac{1}{2} \\ x + tH'(x) & \text{if } \frac{1}{2} \leqslant x \leqslant \frac{3}{2} \\ x - ta & \text{if } \frac{3}{2} \leqslant x \leqslant L \\ x + tH'(\frac{3}{2} + L - x) & \text{if } L \leqslant x \leqslant L + 1 \\ x + tH'(\frac{1}{2}) & \text{if } L + 1 \leqslant x \end{cases} \tag{6.7}$$

We have $a = H'(-\frac{3}{2}) > 0$, so the characteristic curves started at $x \leqslant -\frac{3}{2}$ are increasing functions of time and their graphical representation in the (x, t) plane leans to the right. When x ranges from $-\frac{3}{2}$ to $-\frac{1}{2}$ the quantity $H'(x)$ goes from positive to negative. The characteristic curves started at those x's collide at a unique point denoted A. When x ranges from $-\frac{1}{2}$ to $\frac{1}{2}$, $H'(x)$ goes from negative to positive and the distance between the characteristic started from those points increases. For the characteristic curves started from $x \in [\frac{1}{2}, \frac{3}{2}]$, a similar observation than for those started from $x \in [-\frac{3}{2}, -\frac{1}{2}]$ can be made, in particular they collide at a unique point denoted B. For $x \in [\frac{3}{2}, L]$, the characteristic curves behave like in the region $x \in (-\infty, -\frac{3}{2}]$, but mirrored: they lean to the left. When x ranges from L to $L+1$, the quantity $\frac{3}{2} + L - x$ ranges from $\frac{3}{2}$ from to $\frac{1}{2}$, therefore $H'(\frac{3}{2} + L - x)$ goes from negative to positive, and we observe a ratification fan like in the $[-\frac{1}{2}, \frac{1}{2}]$ range. Finally, for $x \geqslant L+1$, since $H'(\frac{1}{2}) > 0$ the characteristic curves lean to the right.

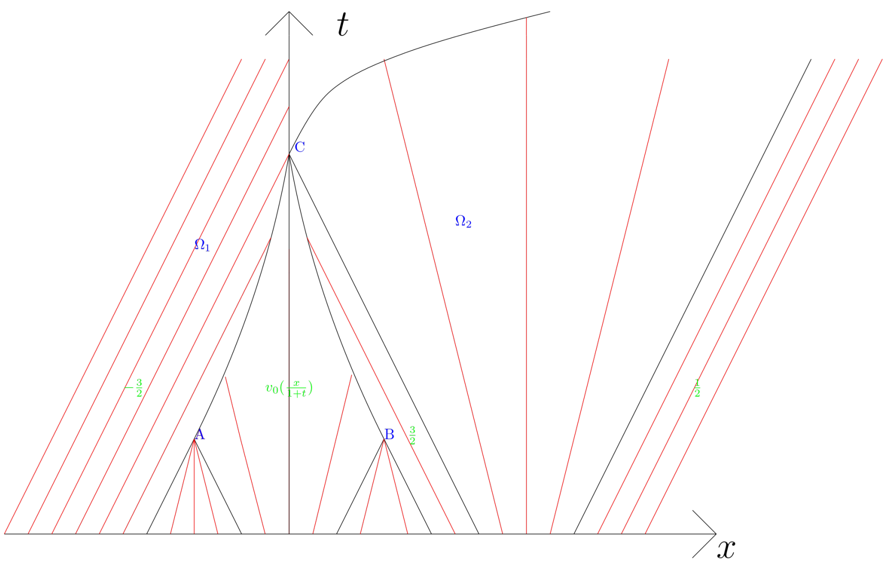

Figure 6.3. Graph of $(x,t) \mapsto (x + tH(v_0(x)), t)$ when $L = at_1$

This behavior is summarized in Figure 6.3 where the characteristic curves are drawn along with three other curves started at the points A, B and C. Those curves define several regions in the (t,x)-plane, two of which are labeled Ω_1 and Ω_2 in Figure 6.3.

We can build a solution v of (6.2) using the characteristic lines and Figure 6.3. Let $(t,y) \in \mathbb{R}_+ \times \mathbb{R}$, if (t,y) does not belong to Ω_1 we set $v(t,y) = v_0(x)$ where $y = X^t(x)$ and $s \mapsto X^s(x)$ is the unique characteristic curve that reaches (t,y) on Figure 6.3, otherwise we set $v(t,y) = -\frac{3}{2}$. As we will see below (see (6.16)), the entropy solution is not constant equal to $-\frac{3}{2}$ in Ω_1.

We start by checking that the solution v we have built satisfies the Rankine–Hugoniot condition and that the curves along which the discontinuities evolve look like the curves on Figure 6.3. The function v exhibits discontinuities, they appear at time $t = t_0$ at $x = -x_0$ and $x = x_0$ where $x_0 = \frac{3}{2} - \frac{a}{a+\frac{1}{2}} \simeq 0.68$ and $t_0 = \frac{1}{a+\frac{1}{2}} \simeq 0.36$ (point A and B on Figure 6.3). Those discontinuities continue to exist at times $t > t_0$ along some curves $\Gamma_A = \{(t, z_A(t))\}$ and $\Gamma_B = \{(t, z_B(t))\}$. Furthermore, z_A and z_B satisfy the Rankine–Hugoniot condition, that is

$$[H(v)]_A = \frac{\mathrm{d}z_A}{\mathrm{d}t}[v]_A, \tag{6.8}$$

$$[H(v)]_B = \frac{\mathrm{d}z_B}{\mathrm{d}t}[v]_B, \tag{6.9}$$

where $[\,\cdot\,]_A$ and $[\,\cdot\,]_B$ respectively denote the jump of $\cdot$ along Γ_A and Γ_B. Using the expression of v in terms of v_0 in each region,

$$\frac{\mathrm{d}z_A}{\mathrm{d}t} = \frac{H((\frac{z_A}{1+t})) - H(-\frac{3}{2})}{\frac{z_A}{1+t} - (-\frac{3}{2})} \geqslant 0, \tag{6.10}$$

$$\frac{\mathrm{d}z_B}{\mathrm{d}t} = \frac{H(\frac{3}{2}) - H(\frac{z_B}{1+t})}{\frac{3}{2} - \frac{z_B}{1+t}} \leqslant 0. \tag{6.11}$$

In particular, z_A is a non-decreasing function of time, meaning Γ_A is curved to the right. Furthermore, let $g(p)$ denote the slope between $(-\frac{3}{2}, H(-\frac{3}{2}))$ and $(p, H(p))$, we have $z_A'(t) = g(z_A(t))$, in addition there exists $m > 0$ such that for all $p \in [-x_0, 0]$ we have $g(p) \geqslant m$. In particular, $z_A'(t) \geqslant m > 0$ so z_A goes to 0 at a speed greater than some $m > 0$. The curve Γ_B is the symmetrical of the curve Γ_A with respect to the vertical axis, so similar observations can be made for the function, z_B in particular z_B also goes to 0 at a speed greater than $m > 0$. In particular, this shows that as time elapses, the two discontinuities move toward each other in the (t, x)-space and the speed at which they move toward each other is bounded from below by a positive constant. Thus, the two discontinuities meet at some point C at a time, t_1 and by symmetry the point C must lie on the vertical axis. Choosing $L = at_1$, the first characteristic started from $x \in [L, L+1]$ to hit the vertical axis is $X^t(L) = L + tH'(\frac{3}{2})$ which hits the vertical axis at $t_2 = \frac{L}{-H'(\frac{3}{2})} = t_1$. Hence, the influence of the asymmetry of the initial condition v_0 is felt by the discontinuities only after the moment they meet at the point C. Before time t_1 the discontinuities appearing at A and B behave as if the initial condition was

$$\tilde{v_0}(x) = \begin{cases} -\frac{3}{2} & \text{if } x \leqslant -\frac{3}{2}, \\ x & \text{if } -\frac{3}{2} \leqslant x \leqslant \frac{3}{2}, \\ \frac{3}{2} & \text{if } \frac{3}{2} \leqslant x. \end{cases} \tag{6.12}$$

In particular, by symmetry, the point C must lie on the vertical axis. After time t_1 only one discontinuity remains and it evolves along the curve $\Gamma_C = \{(z_C(t), t)\}$. The Rankine–Hugoniot condition imposes that after time t_1, the point C evolves along the curve $\Gamma_C = \{(z_C(t), t)\}$ where z_C satisfies

$$[H(v)]_C = \frac{\mathrm{d}z_C}{\mathrm{d}t}[v]_C. \tag{6.13}$$

After time t_1, along Γ_C we have $v_+ = \frac{3}{2} + L - x$ for some $x \in [L, L+1]$ so $v_+ \in [\frac{1}{2}, \frac{3}{2}]$ hence $z_C'(t)$ is non-negative and Γ_C leans to the right. As the discontinuity deviates to the right, v_+ becomes closer to $\frac{1}{2}$ and $z_C'(t)$ becomes bigger. Thus, the speed at which the discontinuity moves away from the vertical axis is an increasing function of time. Hence, at some time t_3 we have $z_C(t_3) = z$ where z is the maximum of H on $[\frac{1}{2}, \frac{3}{2}]$.

We highlight that before time t_1, v is in fact equal to the entropy solution, it is at some time after time t_1 that v will stop being the entropy solution. Note that the jumps along Γ_A, Γ_B and Γ_C of v are all non-negative, so the space derivative of v is bounded below (but not above) at all times by a constant that does not depend on time. In other words, there exists a constant c such that $v(t, \cdot\,)$ is one-sided Lipschitz with constant c. Furthermore, for every, $(y, t) \in \mathbb{R} \times \mathbb{R}_+$ there exists $x \in \mathbb{R}$ such that $v(t, y) = v_0(x)$ so v is a bounded function and $\|v\|_{L^\infty} \leqslant \|v_0\|_{L^\infty}$.

Let us now show that v is not an entropy solution. By contradiction, suppose that v is an entropy solution of (6.2), the curve Γ_C must satisfy the following additional condition [16, Proposition 2.3.7],

$$[F(v)] \leqslant \frac{\mathrm{d}z_C}{\mathrm{d}t}[E(v)] \tag{6.14}$$

for every entropy-entropy flux pair (E, F), that is pairs of function (E, F) such that E is convex and $F' = H'E'$. It is well known that it is necessary and sufficient to check (6.14) only for pairs (E, F) where $E = |x - k|$, $k \in \mathbb{R}$ for, v to be an entropy solution [16, Proposition 2.3.7]. In particular, v is an entropy solution if and only if (6.14) holds for all pairs (E, F) with $E = (x - k)_+$ and $E = (x - k)_-$, $k \in \mathbb{R}$. Hence, if v is an entropy solution, we have for all $k \in (v_-, v_+)$,

$$\frac{H(v_+) - H(k)}{v_+ - k} \leqslant \frac{H(v_+) - H(v_-)}{v_+ - v_-} \leqslant \frac{H(k) - H(v_-)}{k - v_-}. \tag{6.15}$$

The condition (6.15) is not satisfied by the solution v that we built. Indeed, recall that z denotes the maximum of H on $[\frac{1}{2}, \frac{3}{2}]$, the point $(-\frac{3}{2}, H(-\frac{3}{2}))$ is strictly above the line defined by the points, $(0, H(0))$ and $(z, H(z))$ this means that,

$$\frac{H(z) - H(0)}{z - 0} > \frac{H(z) - H(-\frac{3}{2})}{z - (-\frac{3}{2}))}. \tag{6.16}$$

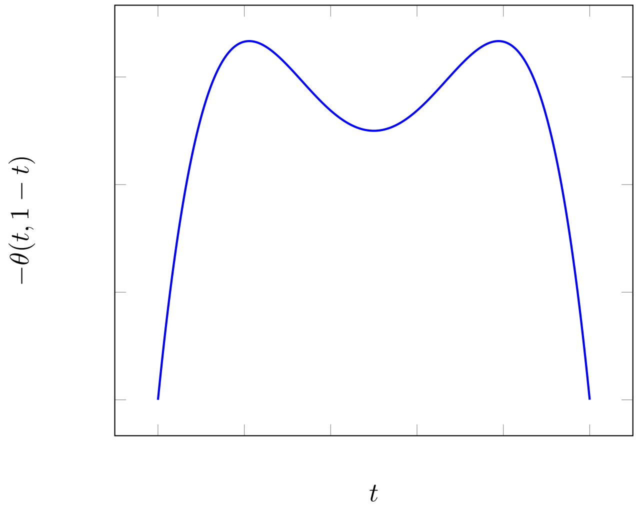

Figure 6.4. Graphical representation of the function θ along the line $t \mapsto (t, 1-t)$ for $t \in [0, 1]$.

But (6.16) contradicts the inequality on the left-hand side of (6.15) for $k = 0$ and time $t = t_3$. Hence, the solution v that we have built is not an entropy solution. $\square$

Finally, we go back to the statistical physics motivations discussed in the Introduction. To highlight the challenge posed by the counter-example of Proposition 6.2, we define a two species spin model with a double-well covariance function resembling Figure 6.1. Let $\sigma = (\sigma_1, \sigma_2) \in \{-1, 1\}^N \times \{-1, 1\}^N$, set

$$E_N(\sigma) = \frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} J^1_{ij} \sigma_{1i} \sigma_{1j} + \frac{1}{\sqrt{N}} \sum_{i,j=1}^{N} J^2_{ij} \sigma_{2i} \sigma_{2j} \\ + \frac{\sqrt{6}}{N^{\frac{3}{2}}} \sum_{i,j,k,l=1}^{N} J_{ijkl} \sigma_{1i} \sigma_{1j} \sigma_{2k} \sigma_{2l},$$

where the J^1_{ij}'s, J^2_{ij}'s and the J_{ijkl}'s are independent and identically distributed standard Gaussian random variables. For every $\sigma, \tau \in \{-1, 1\}^N \times \{-1, 1\}^N$, we have

$$\mathbb{E}[E_N(\sigma) E_N(\tau)] = N \theta\Big(\frac{\sigma_1 \cdot \tau_1}{N}, \frac{\sigma_2 \cdot \tau_2}{N}\Big), \tag{6.17}$$

where $\theta(x, y) = x^2 + y^2 + 6x^2y^2$. The covariance function of the SK model is the square function, which is the function appearing in (1.2). The same holds true for more general models, in the context of (6.17), the behavior of the nonlinearity of the associated equation will be governed by the behavior of the function θ. If we plot a slice of the function θ along the line $t \mapsto (1-t, t)$ we see that the function θ has a double-well structure. This is the same pathological behavior than the one exhibited by (6.4).

This means, that the behaviors we have used to build a counter-example to Theorem 1.1 in Proposition 6.2 cannot be easily ruled out for weak solutions arising in the context of spin glasses. Nonetheless, it was recently proven [5, Theorem 1.1], that even for non-convex models the limit free energy, if it exists, stays in the wavefront. This highlights that even thought Theorem 1.1 is not true for non-convex H, the weak solution arising in the study of non-convex spin glass models is in a sense well-behaved. The precise PDE mechanism that leads to this property has yet to be uncovered.

Acknowledgement

I warmly thank Jean-Christophe Mourrat, Pierre Cardaliaguet and Stefano Bianchini for the help they provided during the conception and the writing of this paper. I also thank the anonymous reviewers for giving several suggestions.

Bibliography

[1] L. Ambrosio, "Transport Equation and Cauchy Problem for BV Vector fields", *Invent. Math.* **158** (2004), no. 2, p. 227-260.

[2] G. Barles, *An Introduction to the Theory of Viscosity Solutions for First-Order Hamilton–Jacobi Equations and Applications*, Lecture Notes in Mathematics, vol. 2074, Springer, 2013, 49-109 pages.

[3] F. Bouchut, F. James & S. Mancini, "Uniqueness and weak stability for multidimensional transport equations with one-sided Lipschitz coefficient", *Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5)* **4** (2005), no. 1, p. 1-25.

[4] P. Cannarsa & C. Sinestrari, *Semiconcave functions, Hamilton–Jacobi equations, and Optimal Control*, Progress in Nonlinear Differential Equations and their Applications, vol. 58, Birkhäuser, 2004, xii+304 pages.

[5] H.-B. Chen & J.-C. Mourrat, "On the free energy of vector spin glasses with non-convex interactions", 2023, `https://arxiv.org/abs/2311.08980`.

[6] H.-B. Chen, J.-C. Mourrat & J. Xia, "Statistical inference of finite-rank tensors", *Ann. Henri Lebesgue* **5** (2022), p. 1161-1189.

[7] L. Corrias, M. Falcone & R. Natalini, "Numerical schemes for conservation laws via Hamilton–Jacobi equations", *Math. Comput.* **64** (1995), no. 210, p. 555-580.

[8] L. C. Evans, *Partial Differential Equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, 2010, xxi+749 pages.

[9] P.-L. Lions, *Generalized solutions of Hamilton–Jacobi equations*, Research Notes in Mathematics, vol. 69, Pitman, 1982, 317 pages.

[10] P.-L. Lions & B. Seeger, “Transport equations and flows with one-sided Lipschitz velocity fields”, 2023, `https://arxiv.org/abs/2306.13288`.

[11] J.-C. Mourrat, “Nonconvex interactions in mean-field spin glasses”, *Probab. Math. Phys.* **2** (2021), no. 2, p. 61-119.

[12] ———, “The Parisi formula is a Hamilton–Jacobi equation in Wasserstein space”, *Can. J. Math.* **74** (2022), no. 3, p. 607-629.

[13] ———, “Free energy upper bound for mean-field vector spin glasses”, *Ann. Inst. Henri Poincaré, Probab. Stat.* **59** (2023), no. 3, p. 1143-1182.

[14] D. Panchenko, *The Sherrington–Kirkpatrick Model*, Springer Monographs in Mathematics, Springer, 2013, xii+156 pages.

[15] G. Parisi, “Infinite Number of Order Parameters for Spin-Glasses”, *Phys. Rev. Lett.* **43** (1979), no. 23, p. 1754-1756.

[16] D. Serre, *Systems of Conservation Laws 1: Hyperbolicity, Entropies, Shock Waves*, Cambridge University Press, 2003, xxii+263 pages.