An introduction to the Aubry-Mather theory

Andrey Biryuk and Diogo A. Gomes

Departamento de Matemática and Center for Mathematical Analysis, Geometry and Dynamical Systems, IST, Lisboa, Portugal

Abstract. This paper is a self-contained introduction to the Aubry-Mather theory and its connections with the theory of viscosity solutions of Hamilton-Jacobi equations. Our starting point is Mañe's variational approach using holonomic measures [Mn96]. We present the Legendre-Fenchel-Rockafellar theorem from convex analysis and discuss the basic theory of viscosity solutions of first order Hamilton-Jacobi equations. We apply these tools to study the Aubry-Mather problem following the ideas in [EG01]. Finally, in the last section, we present a new proof of the invariance under the Euler-Lagrange flow of the Mather measures using ideas from calculus of variations.

1. Motivation and the statement of the problem

Let $\mathbb{T}^d$ be the d-dimensional standard torus. Consider a Lagrangian $L(x, v)$, $L : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$, smooth in both variables, strictly convex in the velocity v, and coercive, that is,

$$\lim_{|v| \to \infty} \inf_x \frac{L(x, v)}{|v|} = +\infty.$$

The minimal action principle of classical mechanics asserts that the trajectories $\mathbf{x}(t)$ of mechanical systems are critical points or minimizers of the action

$$\int_0^T L(\mathbf{x}, \dot{\mathbf{x}}) ds. \tag{1}$$

These critical points are then solutions to the Euler-Lagrange equations

$$\frac{d}{dt} D_v L(\mathbf{x}, \dot{\mathbf{x}}) - D_x L(\mathbf{x}, \dot{\mathbf{x}}) = 0. \tag{2}$$

Mather's problem is a relaxed version of this variational principle, and consists in minimizing the action

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} L(x, v) d\mu(x, v) \tag{3}$$

among a suitable class of probability measures $\mu(x, v)$. Originally, in [Mat91], this minimization was performed over all measures invariant under the Euler-Lagrange equations (2). However, as realized by [Mn96], it is more convenient to consider a larger class of measures, the holonomic measures. It turns out that both problems are equivalent as any holonomic minimizing measure is automatically invariant under the Euler-Lagrange equations. In what follows, we will define this class of measures and provide the motivation for it.

Let $\mathbf{x}(t)$ be a trajectory on $\mathbb{T}^d$. Define a measure $\mu_{\mathbf{x}}^T$ on $\mathbb{T}^d \times \mathbb{R}^d$ by its action on test functions $\psi \in C_c(\mathbb{T}^d \times \mathbb{R}^d)$, $\psi(x, v)$, (continuous with compact support) as follows:

$$\langle \psi, \mu_{\mathbf{x}}^T \rangle = \frac{1}{T} \int_0^T \psi\big(\mathbf{x}(t), \dot{\mathbf{x}}(t)\big) dt.$$

If $\mathbf{x}(t)$ is globally Lipschitz, the family $\{\mu_{\mathbf{x}}^T\}_{T>0}$ has support contained in a fixed compact set, and therefore is weakly-$*$ compact. Consequently one can extract a limit measure $\mu_{\mathbf{x}}$ which encodes some of the asymptotic properties of the trajectory $\mathbf{x}$.

Let $\varphi \in C^1(\mathbb{T}^d)$. For $\psi(x, v) = v \cdot \nabla \varphi(x)$ we have

$$\langle \psi, \mu_{\mathbf{x}} \rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T \dot{\mathbf{x}} \cdot \nabla \varphi(\mathbf{x}) dt = \lim_{T \to \infty} \frac{\varphi\big(\mathbf{x}(T)\big) - \varphi\big(\mathbf{x}(0)\big)}{T} = 0.$$

Let $\gamma(v)$ be a continuous function, $\gamma : \mathbb{R}^d \to \mathbb{R}$, such that $\inf \frac{\gamma(v)}{1+|v|} > 0$, and $\lim_{|v| \to \infty} \frac{\gamma(v)}{1+|v|} = \infty$. A measure μ in $\mathbb{T}^d \times \mathbb{R}^d$ is admissible if $\int_{\mathbb{T}^d \times \mathbb{R}^d} \gamma(v) d\mu < \infty$. An admissible measure μ on $\mathbb{T}^d \times \mathbb{R}^d$ is called holonomic if for all $\varphi \in C^1(\mathbb{T}^d)$ we have

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} v \cdot \nabla \varphi d\mu = 0. \tag{4}$$

Mather's problem consists in minimizing (3) under all probability measures that satisfy (4). As pointed out before, however, this problem was

introduced by Mañe in [Mn96] in his study of Mather's original problem [Mat91].

Most of the results in this paper are not original or are small adaptations of other proofs. The only relevant exception is the proof that any holomonic minimizing measure is invariant under the Euler Lagrange flow, which is original and relies on a simple calculus of variations argument: the construction of holonomy preserving variations. This invariance result was first established by Mañe [Mn96] (which, in fact, was not a full proof but was completed by J. Mather in [Mat]). Another proof was given by Fathi and Siconolfi [FS04] using quite different ideas.

In the original approach by Mather [Mat91], instead of action minimizing holonomic measures, the problem was to find minimal action measures invariant under the Euler-Lagrange flow. This approach has the obvious advantage that any such minimizing measure is automatically invariant under the Euler-Lagrange flow. However, properties such as the graph theorem are harder to prove and, furthermore, this approach does not generalize easily to more general problems, as is the example of the stochastic Mather problem studied in [Gom02].

This paper is organized as follows: firstly, in section 2, we establish the existence of minimizing holonomic measures; in the next section we present the proof of the classical Legendre-Fenchel-Rockafellar duality theorem which is an essential tool to understanding our problem; viscosity solutions of Hamilton Jacobi equations are the subject of section 4, where we give a self contained presentation of the all the results that we will need; next, in section 5, with the tools that were developed in the previous sections, we will compute the dual of Mather's problem and identify its value; then we, in section 6 we revisit the results from [EG01] which establish regularity for viscosity solutions of Hamilton-Jacobi equations; in section 7 we construct a special class of variations for measures that preserve the holonomy constraint and that will be needed in the last section to prove that the Mather measures are invariant under the Euler-Lagrange dynamics.

2. **Minimizing measures**

In this section we prove the existence of minimizing holonomic measures and establish that the minimizing measures are supported in a graph. The existence result is due to Mañe [Mn96]. The graph theorem in the original formulation is due to Mather [Mat91].

By coercivity, we may assume that the Lagrangian L is positive and separated away from zero by adding to it, if necessary, a suitable constant. To be precise, we assume that $L \geqslant 1$. Set $\gamma(v) = \inf_{x \in \mathbb{T}^d} L(x, v)$. Let $\mathcal{M}$ be the set of signed σ-finite Borel measures in $\mathbb{T}^d \times \mathbb{R}^d$ with finite γ-weighted

total variation, that is,

$$\mathcal{M} = \left\{ \mu \; : \; \|\mu\| = \int_{\mathbb{T}^d \times \mathbb{R}^d} \gamma(v) d|\mu| < \infty \right\},$$

where μ is a signed measure in $\mathbb{T}^d \times \mathbb{R}^d$

Recall, that the total variation measure $|\mu|$ is defined as

$$|\mu|(E) = \sup \sum_{i=1}^{\infty} |\mu(E_i)|,$$

where the supremum is taken over all (Borel) partitions $\{E_i\}$ of the Borel set E. By definition μ is σ-finite iff $|\mu|$ is.

Theorem 1. *(Riesz representation) Let $\mu \in \mathcal{M}$, define $F_\mu : C_0^\gamma \to \mathbb{R}$ by $F_\mu(f) = \int f d\mu$. Then $F_\mu \in \left(C_0^\gamma\right)'$ and the map $\mu \mapsto F_\mu$ is an isometric isomorphism of $\mathcal{M}$ onto $\left(C_0^\gamma\right)'$.*

Proof. For $\gamma \equiv 1$ this is the standard Riesz representation theorem, see [Con90, Chpt 3, Thm 5.7]. The general case follows by an isometry $C_0 \to C_0^\gamma$, given by the linear operator of multiplication by function $f \mapsto f\gamma$. $\square$

Applying the Riesz representation theorem, as stated in the previous theorem, we see that the set $\mathcal{M}$ is the dual of the set $C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)$ of continuous functions ϕ that satisfy

$$\|\phi\|_\gamma = \sup_{\mathbb{T}^d \times \mathbb{R}^d} \left| \frac{\phi}{\gamma} \right| < \infty, \qquad \lim_{|v| \to \infty} \frac{\phi(x,v)}{\gamma(v)} \to 0.$$

We should recall the following fact that any sequence of measures such that $\int_{\mathbb{T}^d \times \mathbb{R}^d} \gamma d|\mu_n|$ is bounded admits a subsequence (still denoted by μ_n) which converges weakly, that is, there exists a measure μ, such that for all $\phi \in C_0^\gamma$ we have

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \phi d\mu_n \to \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi d\mu.$$

In particular, since $1 \in C_0^\gamma$, if each of the measures μ_n is a probability measure (i.e. nonegative, with total mass equal to 1) then μ is also a probability measure.

Theorem 2. *There exists a holonomic probability measure μ on $\mathbb{T}^d \times \mathbb{R}^d$ which minimizes*

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} L(x,v) d\mu(x,v).$$

Proof. Consider a minimizing sequence of holonomic probability measures μ_n. Clearly, the sequence $\int_{\mathbb{T}^d \times \mathbb{R}^d} L(x,v) d\mu_n$ is bounded, and so is $\int_{\mathbb{T}^d \times \mathbb{R}^d} \gamma(v) d\mu_n$. Therefore, through a subsequence, μ_n converges in $(C_0^\gamma)'$ to a measure μ. The measure μ is necessary a probability measure as each of μ_n is. Furthermore, for any $\varphi \in C^1(\mathbb{T}^d)$, $vD_x\varphi(x) \in C_0^\gamma$. Therefore

$$0 = \int_{\mathbb{T}^d \times \mathbb{R}^d} vD_x\varphi(x) d\mu_n \to \int_{\mathbb{T}^d \times \mathbb{R}^d} vD_x\varphi(x) d\mu,$$

and so μ is holonomic. It remains to prove that it is minimizer. Let $L_k(x,v)$ be a sequence in C_0^γ increasing pointwise to L. Then

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} L d\mu_n \geq \int_{\mathbb{T}^d \times \mathbb{R}^d} L_k d\mu_n.$$

Therefore

$$\inf \geq \int_{\mathbb{T}^d \times \mathbb{R}^d} L_k d\mu.$$

Applying Fatou's lemma to the right hand side of the previous inequality yields the result. $\qquad\square$

Remark. Another proof will be given using the Fenchel-Legendre-Rockafellar Theorem, Theorem 4.

Theorem 3. *Any minimizing measure is supported in the closure of a graph. That is, there exists a measurable function $\boldsymbol{v} : \mathbb{T}^d \to \mathbb{R}^d$ such that*

$$\operatorname{supp}\mu \in \operatorname{cl}\{x, v \in \mathbb{T}^d \times \mathbb{R}^d : v = \boldsymbol{v}(x)\}.$$

Proof. Proof is given by contradiction. Let $\mu(x,v)$ be a minimizing measure. Then, by the desintegration of measures result (slicing measures) in [Eva90, p.14, Thm 10] there exists a probability measure θ in $\mathbb{T}^d$ and for θ-almost every x there exists a probability measure $\eta(dv; x)$ in $\mathbb{R}^d$ such that $d\mu = \theta(dx)\eta(dv; x)$. If μ is supported on a graph then there exists a function $\boldsymbol{v}(x)$ such that $\eta(v; x) = \delta_{\boldsymbol{v}(x)}(v)$. Otherwise, define the function $\boldsymbol{v}(x)$ in the following way

$$\boldsymbol{v}(x) = \int v \, \eta(dv; x).$$

Furthermore, define the measure $\tilde{\eta}(dv; x) = \delta_{\boldsymbol{v}(x)}(v)$. Then, by the strict convexity of the Lagrangian we have

$$\int_{\mathbb{R}^d} L(x,v)\tilde{\eta}(dv; x) < \int_{\mathbb{R}^d} L(x,v)\eta(dv; x),$$

for all points x in which $\tilde{\eta}(dv; x) \neq \eta(dv; x)$. Therefore, by integrating against θ, we conclude that

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} L(x, v)\tilde{\eta}(dv; x)\theta(dx) < \int_{\mathbb{T}^d \times \mathbb{R}^d} L(x, v)d\mu.$$

Finally, it remains to show that the measure $d\tilde{\mu} = \tilde{\eta}(dv; x)\theta(dx)$ is holonomic. But this follows easily from the fact that

$$\int_{\mathbb{R}^d} vD_x\varphi(x)\tilde{\eta}(dv; x) \stackrel{\text{def } \tilde{\eta}}{=} \boldsymbol{v}(x)D_x\varphi(x) \stackrel{\text{def } \boldsymbol{v}}{=} \int_{\mathbb{R}^d} vD_x\varphi(x)\eta(dv; x).$$

$\square$

Later we will prove that the function $v(x)$ is uniformly continuous μ-almost everywhere (in fact Lipschitz) and so the closure of the graph is in fact a graph.

3. Legendre-Fenchel-Rockafellar duality theorem

In this section, for the convenience of the reader, we give a proof of the Legendre-Fenchel-Rockafellar duality Theorem which is based on the one presented in [Vil03].

Let E be a locally convex topological vector space with dual E'. The duality pairing between E and E' is denoted by $(\cdot, \cdot)$. Let $h : E \to (-\infty, +\infty]$ be a convex function. The Legendre-Fenchel transform $h^* : E' \to [-\infty, +\infty]$ of h is defined by

$$h^*(y) = \sup_{x \in E}\big((x, y) - h(x)\big),$$

for $y \in E'$. In a similar way, if $g : E \to [-\infty, +\infty)$ is concave we define

$$g^*(y) = \inf_{x \in E}\big((x, y) - g(x)\big).$$

Theorem 4 (Fenchel-Legendre-Rockafellar)**.** *Let E be a locally convex topological vector space over $\mathbb{R}$ with dual E'. Let $h : E \to (-\infty, +\infty]$ be a convex function and $g : E \to [-\infty, +\infty)$ a concave function. Then, if there exists a point x_0 where both g and h are finite and at least one of them is continuous,*

$$\min_{y \in E'}[h^*(y) - g^*(y)] = \sup_{x \in E}[g(x) - h(x)]. \tag{5}$$

Remark. It is part of the theorem that the infimum in the left-hand side above is a minimum.

Proof. First we show the "$\geqslant$" inequality in (5). Recall that

$$\inf_{y \in E'} [h^*(y) - g^*(y)] = \inf_{y \in E'} \sup_{x_1, x_2 \in E} [g(x_1) - h(x_2) - (y, x_1 - x_2)].$$

By choosing $x_1 = x_2 = x$ we conclude that

$$\inf_{y \in E'} [h^*(y) - g^*(y)] \geq \sup_{x \in E} [g(x) - h(x)].$$

The opposite inequality is more involved and requires the use of Hahn-Banach's theorem. Let

$$\lambda = \sup_{x \in E} [g(x) - h(x)].$$

If $\lambda = +\infty$ there is nothing to prove, thus we may assume $\lambda < +\infty$. We just need to show that there exists $y \in E'$ such that for all x_1 and x_2 we have

$$g(x_1) - h(x_2) - (y, x_1 - x_2) \leq \lambda, \tag{6}$$

since then, by taking the supremum over x_1 and x_2 yields

$$h^*(y) - g^*(y) \leq \lambda.$$

From $\lambda \geqslant g(x) - h(x)$ it follows $g(x) \leq \lambda + h(x)$. Hence the following convex subsets of $E \times \mathbb{R}$:

$$C_1 = \left\{ (x_1, t_1) \in E \times \mathbb{R} : t_1 < g(x_1) \right\}$$

and

$$C_2 = \left\{ (x_2, t_2) \in E \times \mathbb{R} : \lambda + h(x_2) < t_2 \right\}.$$

are disjoint. Let x_0 as in the statement of the theorem. We will assume that g is continuous at x_0 (for the case in which h is the continuous function the argument is similar). Since $(x_0, g(x_0) - 1) \in C_1$ and g is continuous at x_0, C_1 has non empty interior. Therefore, see [KF75, Chpt 4, sect 14.5], the sets C_1 and C_2 can be separated by a nonzero linear function, i.e., there exists a nonzero vector $z = (w, \alpha) \in E' \times \mathbb{R}$ such that

$$\inf_{c_1 \in C_1} (z, c_1) \leq \sup_{c_2 \in C_2} (z, c_2),$$

that is, for any x_1 such that $g(x_1) > -\infty$ and for any x_2 s.t. $h(x_2) < +\infty$ we have

$$(w, x_1) + \alpha t_1 \leq (w, x_2) + \alpha t_2,$$

whenever $t_1 < g(x_1)$ and $\lambda + h(x_2) < t_2$.

Note that α can not be zero. Otherwise by using $x_2 = x_0$ and taking x_1 in a neighborhood of x_0 where g is finite we deduce that w is also zero. Therefore $\alpha > 0$, otherwise, by taking $t_1 \to -\infty$ we would obtain a contradiction. Dividing w by α and letting $y = -\frac{w}{\alpha}$, we would obtain

$$-(y, x_1) + g(x_1) \leq -(y, x_2) + h(x_2) + \lambda.$$

This is equivalent to (6) and thus we completed the proof.   □

Remark. The condition of continuity at x_0 can be relaxed to the condition of "Gâteaux continuity" or directional continuity, that is the function $t \mapsto f(x_0 + tx)$ is continuous at $t = 0$ for any $x \in E$. Here f stands for either h or g.

4. Viscosity Solutions

In this section we present some basic results on viscosity solutions and its connections with optimal control problems. Most of these results are standard and can be found, for instance in [Eva98], [BCD97] or [FS93].

Let $\alpha > 0$. The infinite horizon optimal control problem consists in minimizing

$$u_\alpha(x) = \inf_{\mathbf{x}:\mathbf{x}(0)=x} \int_{-\infty}^{0} e^{\alpha s} L(\mathbf{x}, \dot{\mathbf{x}}) ds, \tag{7}$$

among all globally Lipschitz trajectories $\mathbf{x}$ with initial condition $\mathbf{x}(0) = x$. Let $T \in \mathbb{R}$. A similar problem, the initial value problem, consists in minimizing

$$V(x, t) = \inf_{\mathbf{x}:\mathbf{x}(t)=x} \int_{-T}^{t} L(\mathbf{x}, \dot{\mathbf{x}}) ds + \psi(\mathbf{x}(-T)), \tag{8}$$

for $t \geq -T$, among all globally Lipschitz trajectories $\mathbf{x}$ with initial condition $\mathbf{x}(t) = x$. Both problems will be useful in the sequel.

Proposition 5. *For any $\alpha > 0$ the function u_α satisfies the dynamic programming principle, that is, for any $T > 0$,*

$$u_\alpha(x) = \inf_{\mathbf{x}:\mathbf{x}(0)=x} \left(\int_{-T}^{0} e^{\alpha s} L(\mathbf{x}, \dot{\mathbf{x}}) ds + e^{-\alpha T} u_\alpha\big(\mathbf{x}(-T)\big) \right). \tag{9}$$

Similarly, V also satisfies

$$V(x, t) = \inf_{\mathbf{x}:\mathbf{x}(t)=x} \left(\int_{-\tilde{t}}^{t} L(\mathbf{x}, \dot{\mathbf{x}}) dt + V\big(\mathbf{x}(-\tilde{t}), -\tilde{t}\big) \right), \tag{10}$$

for all $-T \leq -\tilde{t} \leq t$.

Proof. We will establish (9), the proof of (10) is similar. Clearly, for any trajectory $\mathbf{x}(\cdot)$,

$$\int_{-\infty}^0 e^{\alpha s} L(\mathbf{x},\dot{\mathbf{x}})ds = \int_{-T}^0 e^{\alpha s} L(\mathbf{x},\dot{\mathbf{x}})dt \; + \; \int_{-\infty}^{-T} e^{\alpha s} L(\mathbf{x},\dot{\mathbf{x}})ds$$

$$= \int_{-T}^0 e^{\alpha s} L(\mathbf{x},\dot{\mathbf{x}})ds \; +$$

$$e^{-\alpha T}\int_{-\infty}^0 e^{\alpha s} L\big(\mathbf{x}(s-T),\dot{\mathbf{x}}(s-T)\big)ds$$

$$\geq \int_{-T}^0 e^{\alpha s} L(\mathbf{x},\dot{\mathbf{x}})ds \; + \; e^{-\alpha T}u_\alpha\big(\mathbf{x}(-T)\big).$$

Thus by taking the infimum we obtain that

$$u_\alpha(x) \geq \inf_{\mathbf{x}:\mathbf{x}(0)=x}\left(\int_{-T}^0 e^{\alpha s} L(\mathbf{x},\dot{\mathbf{x}})ds + e^{-\alpha T}u_\alpha\big(\mathbf{x}(-T)\big)\right).$$

To obtain the other inequality, fix x, fix $\epsilon > 0$ and consider a trajectory $\mathbf{x}$ such that $\mathbf{x}(0) = x$ and

$$\inf_{\mathbf{x}:\mathbf{x}(0)=x}\left(\int_{-T}^0 e^{\alpha s} L(\mathbf{x},\dot{\mathbf{x}})ds + e^{-\alpha T}u_\alpha\big(\mathbf{x}(-T)\big)\right) \geq$$

$$\int_{-T}^0 e^{\alpha s} L(\mathbf{x},\dot{\mathbf{x}})ds + e^{-\alpha T}u_\alpha\big(\mathbf{x}(-T)\big) - \epsilon.$$

Let $\mathbf{y}$ be a trajectory such that $\mathbf{y}(0) = \mathbf{x}(-T)$ and

$$u_\alpha\big(\mathbf{x}(-T)\big) \geq \int_{-\infty}^0 e^{\alpha s} L(\mathbf{y},\dot{\mathbf{y}})ds - \epsilon.$$

Denote by $\mathbf{z}$ the concatenation of these two trajectories:

$$\mathbf{z}(t) = \begin{cases} \mathbf{x}(t) & \text{if } t \geq -T \\ \mathbf{y}(t+T) & \text{if } t \leq -T. \end{cases}$$

Then

$$u_\alpha(x) \leq \int_{-\infty}^0 e^{\alpha s} L(\mathbf{z},\dot{\mathbf{z}})ds \leq \int_{-T}^0 e^{\alpha s} L(\mathbf{z},\dot{\mathbf{z}})ds + e^{-\alpha T}u_\alpha\big(\mathbf{z}(-T)\big) + \epsilon$$

$$\leq \inf_{\mathbf{x}:\mathbf{x}(0)=x}\left(\int_{-T}^0 e^{\alpha s} L(\mathbf{x},\dot{\mathbf{x}})ds + e^{-\alpha T}u_\alpha\big(\mathbf{x}(-T)\big)\right) + 2\epsilon.$$

Thus, by sending $\epsilon \to 0$ we obtain the opposite inequality. $\square$

Proposition 6. *Both infima in proposition 5 are in fact minima.*

Proof. This follows from the compactness of the torus and a standard calculus of variations argument. □

We recall that the Hamiltonian H is the Legendre transform of L given by

$$H(p, x) = \sup_v \big(p \cdot v - L(x, v)\big).$$

If L is strictly convex and coercive in v then so is $H(p, x)$ as a function of p. Furthermore, for $p = D_v L(x, v)$ we have

$$v = D_p H(p, x), \tag{11}$$

and

$$D_x L(x, v) = -D_x H(p, x). \tag{12}$$

The proof of these identities can be found in [Eva98].

Proposition 7. *Let $\varphi(x)$ be a smooth function with bounded first and second derivatives. Suppose $\alpha \geq 0$. Denote by $\Phi(x, t)$ the unique classical solution to the time dependent Hamilton-Jacobi equation*

$$\Phi_t + \alpha \Phi + H(D_x \Phi, x) = 0 \tag{13}$$

on the time interval $[-T, 0]$, with initial data $\Phi(x, -T) = \varphi(x)$. This solution exists for small time T, by the method of characteristics. Then, for all $t \geq -T$,

$$e^{\alpha t} \Phi(x, t) = \inf_{\mathbf{x}:\mathbf{x}(t)=x} \left(\int_{-T}^t e^{\alpha s} L(\mathbf{x}(s), \dot{\mathbf{x}}(s)) ds + e^{-\alpha T} \varphi\big(\mathbf{x}(-T)\big) \right).$$

Proof. Observe that for any Lipschitz trajectory $\mathbf{x}$ we have

$$-e^{\alpha t} \Phi(\mathbf{x}(t), t) + e^{-\alpha T} \varphi(\mathbf{x}(-T)) = -\int_{-T}^t \frac{d}{ds} \left(e^{\alpha s} \Phi(\mathbf{x}(s), s) \right) ds$$

$$= -\int_{-T}^t \left(\alpha e^{\alpha s} \Phi(\mathbf{x}(s), s) + e^{\alpha s} D_x \Phi(\mathbf{x}(s), s) \cdot \dot{\mathbf{x}} + e^{\alpha s} \Phi_s(\mathbf{x}(s), s) \right) ds.$$

Adding $\int_{-T}^t e^{\alpha s} L(\mathbf{x}(s), \dot{\mathbf{x}}(s)) ds + e^{\alpha t} \Phi(\mathbf{x}(t), t)$ to the above equality and taking the infimum over all trajectories $\mathbf{x}$, we obtain

$$\inf_{\mathbf{x}:\mathbf{x}(t)=x} \left(\int_{-T}^t e^{\alpha s} L(\mathbf{x}(s), \dot{\mathbf{x}}(s)) ds + e^{-\alpha T} \varphi\big(\mathbf{x}(-T)\big) \right) = e^{\alpha t} \Phi(\mathbf{x}(t), t) +$$

$$\inf_{\mathbf{x}:\mathbf{x}(t)=x} \left(\int_{-T}^{t} e^{\alpha s}\big(-\alpha\Phi(\mathbf{x}(s),s) - \Phi_s(\mathbf{x}(s),s)\big) + e^{\alpha s}\big(L(\mathbf{x}(s),\dot{\mathbf{x}}(s)) - \right.$$

$$\left. D_x\Phi(\mathbf{x}(s),s)\cdot\dot{\mathbf{x}}\big)ds \right).$$

Now recall that $-H(p,x) \leq L(x,v) - p\cdot v$, therefore

$$\inf_{\mathbf{x}:\mathbf{x}(t)=x} \left(\int_{-T}^{t} e^{\alpha s}L(\mathbf{x}(s),\dot{\mathbf{x}}(s))ds + e^{-\alpha T}\varphi\big(\mathbf{x}(-T)\big) \right) \geq e^{\alpha t}\Phi(\mathbf{x}(t),t) +$$

$$\inf_{\mathbf{x}:\mathbf{x}(t)=x} \left(\int_{-T}^{t} e^{\alpha s}\big(-\alpha\Phi(\mathbf{x}(s),s) - \Phi_s(\mathbf{x}(s),s) - H(D_x\Phi(\mathbf{x}(s),s),\mathbf{x}(s))\big)ds \right)$$

$$= e^{\alpha t}\Phi(\mathbf{x}(t),t).$$

Now consider the trajectory $\mathbf{x}$ given by solving the following differential equation

$$\dot{\mathbf{x}}(s) = D_pH(D_x\Phi(\mathbf{x}(s),s),\mathbf{x}(s))$$

with initial condition $\mathbf{x}(t) = x$. Then, since for $v = D_pH(p,x)$ we have $-H(p,x) = L(x,v) - p\cdot v$, we conclude that

$$\inf_{\mathbf{x}:\mathbf{x}(t)=x} \left(\int_{-T}^{t} e^{\alpha s}L(\mathbf{x}(s),\dot{\mathbf{x}}(s))ds + e^{-\alpha T}\varphi\big(\mathbf{x}(-T)\big) \right) \leq e^{\alpha t}\Phi(\mathbf{x}(t),t) +$$

$$\int_{-T}^{t} e^{\alpha s}\Big(-\alpha\Phi\big(\mathbf{x}(s),s\big) - \Phi_s\big(\mathbf{x}(s),s\big) - H\big(D_x\Phi(\mathbf{x}(s),s),\mathbf{x}(s)\big)\Big)ds =$$

$$e^{\alpha t}\Phi(\mathbf{x}(t),t),$$

which ends the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\Box$

In what follows n stands for d or $d+1$. So we can treat both space and time-space situation.

Let $F : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$, $F(p,z,x)$, be a continuous function. A bounded uniformly continuous function u is a viscosity solution to the Hamilton-Jacobi equation

$$F(Du, u, x) = 0, \tag{14}$$

if for any C^∞ function $\varphi : \mathbb{R}^n \to \mathbb{R}$, $\varphi(x)$, and any point x_0 which locally maximizes (respectively minimizes) the difference $u - \varphi$, the following inequality holds:

$$F(D\varphi(x_0), u(x_0), x_0) \leq 0 \text{ (resp. } \geq 0). \tag{15}$$

Note that for time dependent equations the definition also makes sense by taking $F(DV, V, x) = -V_t + \alpha V + H(D_xV, x)$.

By using the following trick we may assume without loss of generality that the maximum or minimum in the definition are strict: if $u - \varphi$ has

a maximum (resp. minimum) at a point x_0 then let $\tilde{\varphi}(x) = \varphi(x) - |x - x_0|^2$ (resp. $+$) then $u - \tilde{\varphi}$ has a strict maximum (resp. minimum) at x_0, furthermore the first derivative of $\tilde{\varphi}$ agrees with the derivative of φ at x_0.

In fact, although the definition of viscosity solution requires C^∞ test functions φ, the viscosity inequalities (15) also hold for C^1 functions as the next lemma shows.

Lemma 8. *If u is a viscosity solution of* (14) *then for any C^1 function $\varphi : \mathbb{R}^n \to \mathbb{R}$, $\varphi(x)$, and any point x_0 which is a strict local minimum (respectively maximum) of the difference $u - \varphi$ the inequalities* (15) *hold.*

Proof. Let η_ϵ be a standard mollifier, and consider the function $\varphi_\epsilon = \varphi * \eta_\epsilon$. Since $\varphi_\epsilon \to \varphi$ uniformly as $\epsilon \to 0$, for ϵ sufficiently small there exists a point x_ϵ which is a local minimum of $u - \varphi_\epsilon$, and $x_\epsilon \to x$. By the viscosity property

$$F(D\varphi_\epsilon(x_\epsilon), u(x_\epsilon), x_\epsilon) \leq 0 \text{ (resp. } \geq 0).$$

Since $D\varphi_\epsilon \to D\varphi$ uniformly, by passing to the limit we have the desired result. $\square$

Before proceeding, we need to recall some definitions. The *superdifferential* $D_x^+ \psi(x)$ of a function ψ at the point x is the set of values $p \in \mathbb{R}^n$ such that

$$\limsup_{|v| \to 0} \frac{\psi(x + v) - \psi(x) - p \cdot v}{|v|} \leq 0.$$

Similarly, the *subdifferential* $D_x^- \psi(x)$ of ψ at the point x is the set of values p such that

$$\liminf_{|v| \to 0} \frac{\psi(x + v) - \psi(x) - p \cdot v}{|v|} \geq 0.$$

From this definition it follows that both $D_x^\pm \psi(x)$ are convex sets (possibly empty).

Proposition 9. *If $D_x^- \psi(x), D_x^+ \psi(x) \neq \emptyset$ then $D_x^- \psi(x) = D_x^+ \psi(x) = \{p\}$, a single point in $\mathbb{R}^n$, and ψ is differentiable at x with $D_x \psi = p$. Conversely, if ψ is differentiable at x then*

$$D_x^- \psi(x) = D_x^+ \psi(x) = \{D_x \psi(x)\}.$$

Proof. First we claim that if $D_x^- \psi(x)$ and $D_x^+ \psi(x)$ are both non-empty they must coincide and have a single element denoted by p. Indeed, for any $p^- \in D_x^- \psi(x)$ and $p^+ \in D_x^+ \psi(x)$

$$\liminf_{|v| \to 0} \frac{\psi(x + v) - \psi(x) - p^- \cdot v}{|v|} \geq 0$$

$$\limsup_{|v|\to 0} \frac{\psi(x+v) - \psi(x) - p^+ \cdot v}{|v|} \leq 0.$$

By subtraction we conclude

$$\liminf_{|v|\to 0} \frac{(p^+ - p^-) \cdot v}{|v|} \geq 0.$$

In particular, by choosing $v = -\epsilon \frac{p^+ - p^-}{|p^- - p^+|}$, we get

$$-|p^- - p^+| \geq 0,$$

which implies $p^- = p^+ \equiv p$. This (unique) element p satisfies

$$\lim_{|v|\to 0} \frac{\psi(x+v) - \psi(x) - p \cdot v}{|v|} = 0,$$

and so $D_x \psi = p$.

To prove the converse, we just have to observe that if ψ is differentiable then we have $\psi(x+v) = \psi(x) + D_x\psi(x) \cdot v + o(v)$. $\qquad\square$

Proposition 10. *Let ψ be continuous. Then, if $p \in D_x^+ \psi(x_0)$ (resp. $p \in D_x^- \psi(x_0)$) there exists a C^1 function ϕ such that $\psi(x) - \phi(x)$ has a strict local maximum (resp. minimum) at x_0 and $p = D_x\phi(x_0)$. Conversely, if ϕ is a C^1 function such that $\psi(x) - \phi(x)$ has a local maximum (resp. minimum) at x_0 then $p = D_x\phi(x_0) \in D_x^+\psi(x_0)$ (resp. $D_x^-\psi(x_0)$).*

Proof. By subtracting $p \cdot (x - x_0) + \psi(x_0)$ to ψ we may assume $\psi(x_0) = 0$ and $p = 0$. By changing coordinates we can take $x_0 = 0$. Then $0 \in D_x^+\psi(0)$ and so

$$\limsup_{x\to 0} \frac{\psi(x)}{|x|} \leq 0.$$

Hence there exists a continuous function $\rho(x)$, with $\rho(0) = 0$ such that

$$\psi(x) \leq |x|\rho(x).$$

Let $\eta(r) = \max_{|x|\leq r}\{\rho(x)\}$. This function is continuous, non-decreasing and $\eta(0) = 0$. Define

$$\phi(x) = \int_{|x|}^{2|x|} \eta(r)dr + |x|^2.$$

Note that ϕ is C^1 and $\phi(0) = D_x\phi(0) = 0$. Moreover, for $x \neq 0$,

$$\psi(x) - \phi(x) \leq |x|\rho(x) - \int_{|x|}^{2|x|} \eta(r)dr - |x|^2 < 0.$$

Thus $\psi - \phi$ has a strict local maximum at 0.

Conversely, suppose $\psi(x) - \phi(x)$ has a local maximum at 0. Without loss of generality we may assume $\psi(0) - \phi(0) = 0$, and so $\phi(0) = 0$. Then $\psi(x) - \phi(x) \leq 0$ or equivalently

$$\psi(x) \leq p \cdot (x - x_0) + (\phi(x) - p \cdot (x - x_0)).$$

Thus, by choosing $p = D_x \phi(x_0)$, and using the fact that

$$\lim_{x \to x_0} \frac{\phi(x) - p \cdot (x - x_0)}{|x - x_0|} = 0,$$

we conclude that $D_x \phi(x_0) \in D_x^+ \psi(x_0)$. The case of a minimum is similar.
□

From the previous proposition we conclude that a function u is a viscosity solution of $F(Du, u, x) = 0$ if and only if for all x and any $p \in D^+ u(x)$ (resp. $p \in D^- u(x)$) we have $F(p, u(x), x) \leq 0$ (resp. ≥ 0).

Proposition 11. *Let u be a viscosity solution of $F(Du, u, x) = 0$. Then u solves the equation at all points of differentiability.*

Proof. It suffices to observe that at any point x of differentiability of u, by Proposition 9, $\{Du(x)\} = D^+ u(x) = D^- u(x)$, and therefore by the previous remark $F(Du, u, x) = 0$. □

Theorem 12. *Let u_α be the value function of the infinite horizon discounted cost problem (7). Then u_α is a viscosity solution to*

$$\alpha u_\alpha + H(Du_\alpha, x) = 0.$$

Similarly, let V be a solution to the initial value problem (8). Then V is a viscosity solution of

$$V_t + H(D_x V, x) = 0.$$

Proof. We present the proof only for the discounted cost infinite horizon as the other case is similar, and we refer the reader to [Eva98], for instance. Let $\varphi : \mathbb{T}^d \to \mathbb{R}$, $\varphi(x)$, be a C^∞ function, and let $x_0 \in \operatorname{argmin}(u_\alpha - \varphi)$. By adding a suitable constant to φ we may assume that $u(x_0) - \varphi(x_0) = 0$, and $u(x) - \varphi(x) \geq 0$ at all other points.

We must show that

$$\alpha \varphi(x_0) + H(D_x \varphi(x_0), x_0) \geq 0,$$

that is, there exists $v \in \mathbb{R}^d$ such that

$$\alpha \varphi(x_0) + v \cdot D_x \varphi(x_0) - L(x_0, v) \geq 0.$$

By contradiction assume that there exists $\theta > 0$ such that

$$\alpha \varphi(x_0) + v \cdot D_x \varphi(x_0) - L(x_0, v) < -\theta,$$

for all v. Because the mapping $v \mapsto L$ is superlinear and φ is C^1, there exists a $R > 0$ and $r_1 > 0$ such that for all $x \in B_{r_1}(x_0)$ and all $v \in B_R^c(0) = \mathbb{R}^d \setminus B_R(0)$ we have

$$\alpha\varphi(x) + v \cdot D_x\varphi(x) - L(x,v) < -\frac{\theta}{2}.$$

By continuity, for some $0 < r < r_1$ and all $x \in B_r(x_0)$ we have

$$\alpha\varphi(x) + v \cdot D_x\varphi(x) - L(x,v) < -\frac{\theta}{2},$$

for all $v \in B_R(0)$.

Therefore for any trajectory $\mathbf{x}$ with $\mathbf{x}(0) = x_0$ and any $T \geq 0$ such that the trajectory $\mathbf{x}$ stays near x_0 on $[-T,0]$, i.e., $\mathbf{x}(t) \in B_r(x_0)$ for $t \in [-T,0]$ we have

$$e^{-\alpha T}u(\mathbf{x}(-T)) - u(x_0) \geq e^{-\alpha T}\varphi(\mathbf{x}(-T)) - \varphi(x_0)$$

$$= -\int_{-T}^0 e^{\alpha t}\big(\alpha\varphi(\mathbf{x}(t)) + \dot{\mathbf{x}}(t) \cdot D_x\varphi(\mathbf{x}(t))\big)dt$$

$$\geq \frac{\theta}{2}\int_{-T}^0 e^{\alpha t}dt - \int_{-T}^0 e^{\alpha t}L(\mathbf{x},\dot{\mathbf{x}})dt.$$

This yields

$$u(x_0) \leq -\frac{\theta}{2}\int_{-T}^0 e^{\alpha t}dt + \int_{-T}^0 e^{\alpha t}L(\mathbf{x},\dot{\mathbf{x}})dt + e^{-\alpha T}u(\mathbf{x}(-T))$$

Since the infimum in (9) is, in fact, a minimum we can choose a time interval $[-T^*,0]$ and a trajectory $\mathbf{x}^*$ that minimizes (9):

$$u(x_0) = \int_{-T^*}^0 e^{\alpha t}L(\mathbf{x}^*,\dot{\mathbf{x}}^*)dt + e^{-\alpha T}u(\mathbf{x}^*(-T^*)).$$

A minimizing trajectory on $[-T^*,0]$ also minimizes on any sub interval: for any $T \in (0, T^*)$ we have

$$u(x_0) = \int_{-T}^0 e^{\alpha t}L(\mathbf{x}^*,\dot{\mathbf{x}}^*)dt + e^{-\alpha T}u(\mathbf{x}^*(-T)).$$

Taking T small enough we can insure that $\mathbf{x}^*$ stays near x_0 on $[-T,0]$. This yields a contradiction.

Now consider $x_0 \in \operatorname{argmax}(u_\alpha - \varphi)$. Again, by adding a suitable constant to φ we may assume that $u(x_0) - \varphi(x_0) = 0$, and $u(x) - \varphi(x) \leq 0$ at all other points.

We must show that

$$\alpha\varphi(x_0) + H(D_x\varphi(x_0), x_0) \leq 0,$$

that is, for all $v \in \mathbb{R}^d$ we have

$$\alpha\varphi(x_0) + v \cdot D_x\varphi(x_0) - L(x_0, v) \le 0.$$

By contradiction assume that there exists $\theta > 0$ such that for some $\bar{v}$

$$\alpha\varphi(x_0) + \bar{v} \cdot D_x\varphi(x_0) - L(x_0, \bar{v}) > \theta.$$

By continuity, for some $r > 0$ and all $x \in B_r(x_0)$ we have

$$\alpha\varphi(x) + \bar{v} \cdot D_x\varphi(x) - L(x, \bar{v}) > \frac{\theta}{2}.$$

The trajectory $\mathbf{x}$, with $\mathbf{x}(0) = x_0$, $\dot{\mathbf{x}} = \bar{v}$ stays near x_0 for $t \in [-T, 0]$, provided $T > 0$ is sufficiently small. Therefore

$$e^{-\alpha T}u(\mathbf{x}(-T)) - u(x_0) \le e^{-\alpha T}\varphi(\mathbf{x}(-T)) - \varphi(x_0)$$

$$= -\int_{-T}^{0} e^{\alpha t}\big(\alpha\varphi(\mathbf{x}(t)) + \dot{\mathbf{x}}(t) \cdot D_x\varphi(\mathbf{x}(t))\big)dt$$

$$\le -\frac{\theta}{2}\int_{-T}^{0} e^{\alpha t}dt - \int_{-T}^{0} e^{\alpha t}L(\mathbf{x}, \dot{\mathbf{x}})dt.$$

This yields

$$u(x_0) \ge \frac{\theta}{2}\int_{-T}^{0} e^{\alpha t}dt + \int_{-T}^{0} e^{\alpha t}L(\mathbf{x}, \dot{\mathbf{x}})dt + e^{-\alpha T}u(\mathbf{x}(-T)).$$

But since by (9)

$$u(x_0) \le \int_{-T}^{0} e^{\alpha t}L(\mathbf{x}, \dot{\mathbf{x}})dt + e^{-\alpha T}u(\mathbf{x}(-T)),$$

this yields the contradiction $\frac{\theta}{2}\frac{1-e^{-\alpha T}}{\alpha} \le 0$ with $T > 0$. $\qquad\square$

Theorem 13. *Let u_α be a viscosity solution to*

$$\alpha u_\alpha + H(Du_\alpha, x) = 0.$$

Then αu_α is uniformly bounded and u_α is Lipschitz, uniformly in α.

Proof. First let x_M be the point where $u_\alpha(x)$ has a global maximum, and x_m a point of global minimum. Then, by the viscosity property, i.e., the definition of the viscosity solution, we have

$$\alpha u_\alpha(x_M) + H(0, x_M) \le 0, \quad \alpha u_\alpha(x_m) + H(0, x_m) \ge 0,$$

which yields that αu_α is uniformly bounded.

Now we establish the Lipschitz bound. Observe that if u_α is Lipschitz, then there exists $M > 0$ such that

$$u_\alpha(x) - u_\alpha(y) \le M|x - y|,$$

for all x, y. By contradiction, assume that for every $M > 0$ there exists x and y such that
$$u_\alpha(x) - u_\alpha(y) > M|x - y|.$$
Let $\varphi(x) = u_\alpha(y) + M|x - y|$. Then $u_\alpha(x) - \varphi(x)$ has a maximum at some point $x \neq y$. Therefore
$$\alpha u_\alpha(x) + H\big(M\tfrac{x-y}{|x-y|}, x\big) \leq 0,$$
which by the coercivity of H yields a contradiction if M is sufficiently large. $\qquad\square$

Theorem 14. *(Stability theorem for viscosity solutions) Assume that for $\alpha > 0$ function u^α is a viscosity solution for $H^\alpha(u, Du, x) = 0$. Let $H^\alpha \to H$ uniformly on compact sets, and $u^\alpha \to u$ uniformly. Then u is a viscosity solution for $H(u, Du, x) = 0$.*

Proof. Suppose $u - \varphi$ has a strict local maximum (resp. minimum) at a point x_0. Then there exists $x_\alpha \to x$ such that $u_\alpha - \varphi$ has a local maximum (resp. minimum) at x_α. Then
$$H^\alpha(u^\alpha(x_\alpha), D\varphi(x_\alpha), x_\alpha) \leq 0 \qquad (\text{resp. } \geq 0).$$
Letting $\alpha \to 0$ finishes the proof. $\qquad\square$

As demonstrated in context of homogenization of Hamilton-Jacobi equations, in the classic but unpublished paper by Lions, Papanicolaou and Varadhan [LPV88], it is possible to construct, using the previous result, viscosity solutions to the stationary Hamilton-Jacobi equation
$$H(Du, x) = \overline{H}. \tag{16}$$

Theorem 15 (Lions, Papanicolao, Varadhan)**.** *There exists a number $\overline{H}$ and a function $u(x)$, $\mathbb{Z}^d$ periodic in x, that solves (16) in the viscosity sense.*

Proof. Since $u_\alpha - \min u_\alpha$ is periodic, equicontinuous, and uniformly bounded, it converges, up to subsequences, to a function u. Moreover $u_\alpha \leq \frac{C}{\alpha}$, thus αu_α converges uniformly, up to subsequences, to a constant, which we denote by $-\overline{H}$. Then, the stability theorem for viscosity solutions, theorem 14, implies that u is a viscosity solution of
$$H(Du, x) = \overline{H}.$$
$\qquad\square$

Theorem 16. *Let $u : \mathbb{T}^d \to \mathbb{R}$ be a viscosity solution to*
$$H(Du, x) = C.$$
Then u is Lipschitz, and the Lipschitz constant does not depend on u.

Proof. First observe that from the fact that $u = u - 0$ achieves maximum and minimum in $\mathbb{T}^d$ we have

$$\min_{x \in \mathbb{T}^d} H(0, x) \leq C \leq \max_{x \in \mathbb{T}^d} H(0, x).$$

Then, it is enough to argue as in the proof of Theorem 13.   □

To establish uniqueness of viscosity solutions we need the following lemma:

Lemma 17. *Let V be a viscosity solution of*

$$V_t + H(D_x V, x) = 0$$

in $[-T, 0] \times \mathbb{R}^n$ and ϕ a C^1 function. If $V - \phi$ has a maximum (resp. minimum) at $(x_0, t_0) \in \mathbb{R}^d \times (-T, 0]$ then

$$\phi_t(x_0, t_0) + H(D_x\phi(x_0, t_0), x_0) \leq 0 \ (resp. \ \geq 0) \quad at \quad (x_0, t_0). \qquad (17)$$

REMARK: The important point is that the inequality is valid even for some non-interior points ($t_0 = 0$).

Proof. Only the case $t_0 = 0$ requires proof since in the other case the maximum is interior and then the viscosity property (the definition of viscosity solution) yields the inequality. Consider

$$\tilde{\phi} = \phi - \frac{\epsilon}{t}.$$

Then $V - \tilde{\phi}$ has an interior local maximum at (x_ϵ, t_ϵ) with $t_\epsilon < 0$. Furthermore, $(x_\epsilon, t_\epsilon) \to (x_0, 0)$, as $\epsilon \to 0$. At the point (x_ϵ, t_ϵ) we have

$$\phi_t(x_\epsilon, t_\epsilon) + \frac{\epsilon}{t_\epsilon^2} + H(D_x\phi(x_\epsilon, t_\epsilon), x_\epsilon) \leq 0,$$

that is, since $\frac{\epsilon}{t_\epsilon^2} \geq 0$,

$$\phi_t(x_0, 0) + H(D_x\phi(x_0, 0), x_0) \leq 0.$$

Analogously we obtain the opposite inequality, using $\tilde{\phi} = \phi + \frac{\epsilon}{t}$.   □

Next we prove the uniqueness of viscosity solutions:

Theorem 18 (Uniqueness)**.** *Suppose H satisfies*

$$|H(p, x) - H(q, x)| \leq C(|p| + |q|)|p - q|$$
$$|H(p, x) - H(p, y)| \leq C|x - y|(C + H(p, x))$$

Then the value function given by (8) is the unique viscosity solution to the Hamilton-Jacobi equation

$$V_t + H(D_x V, x) = 0$$

that satisfies $V(x, -T) = \psi(x)$.

Proof. Let V and $\tilde{V}$ be two viscosity solutions with

$$\sup_{-T \leq t \leq 0} V - \tilde{V} = \sigma > 0.$$

For $0 < \epsilon, \lambda < 1$ we define

$$\psi(x, y, t, s) = V(x, t) - \tilde{V}(y, s) - \lambda(t + s + 2T) - \frac{1}{\epsilon^2}(|x-y|^2 + |t-s|^2) - \epsilon(|x|^2 + |y|^2).$$

When ϵ, λ are sufficiently small we have

$$\max \psi(x, y, t, s) = \psi(x_{\epsilon,\lambda}, y_{\epsilon,\lambda}, t_{\epsilon,\lambda}, s_{\epsilon,\lambda}) > \frac{\sigma}{2}.$$

Since $\psi(x_{\epsilon,\lambda}, y_{\epsilon,\lambda}, t_{\epsilon,\lambda}, s_{\epsilon,\lambda}) \geq \psi(0, 0, -T, -T)$, and both V and $\tilde{V}$ are bounded, we have

$$|x_{\epsilon,\lambda} - y_{\epsilon,\lambda}|^2 + |t_{\epsilon,\lambda} - s_{\epsilon,\lambda}|^2 \leq C\epsilon^2$$

and

$$\epsilon(|x_{\epsilon,\lambda}|^2 + |y_{\epsilon,\lambda}|^2) \leq C.$$

From these estimates and the fact that V and $\tilde{V}$ are continuous, it then follows that

$$\frac{|x_{\epsilon,\lambda} - y_{\epsilon,\lambda}|^2 + |t_{\epsilon,\lambda} - s_{\epsilon,\lambda}|^2}{\epsilon^2} = o(1),$$

as $\epsilon \to 0$.

Denote by ω and $\tilde{\omega}$ the modulus of continuity of V and $\tilde{V}$. Then

$$\frac{\sigma}{2} \leq V(x_{\epsilon,\lambda}, t_{\epsilon,\lambda}) - \tilde{V}(y_{\epsilon,\lambda}, s_{\epsilon,\lambda})$$
$$= V(x_{\epsilon,\lambda}, t_{\epsilon,\lambda}) - V(x_{\epsilon,\lambda}, -T) + V(x_{\epsilon,\lambda}, -T) - \tilde{V}(x_{\epsilon,\lambda}, -T) +$$
$$+ \tilde{V}(x_{\epsilon,\lambda}, -T) - \tilde{V}(x_{\epsilon,\lambda}, s_{\epsilon,\lambda}) + \tilde{V}(x_{\epsilon,\lambda}, s_{\epsilon,\lambda}) - \tilde{V}(y_{\epsilon,\lambda}, s_{\epsilon,\lambda}) \leq$$
$$\leq \omega(T + t_{\epsilon,\lambda}) + \tilde{\omega}(T + s_{\epsilon,\lambda}) + \tilde{\omega}(o(\epsilon)).$$

Therefore, if ϵ is sufficiently small $T + t_{\epsilon,\lambda} > \mu > 0$, uniformly in ϵ.

Let ϕ be given by

$$\phi(x, t) = \tilde{V}(y_{\epsilon,\lambda}, s_{\epsilon,\lambda}) + \lambda(2T + t + s_{\epsilon,\lambda}) +$$
$$+ \frac{1}{\epsilon^2}(|x - y_{\epsilon,\lambda}|^2 + |t - s_{\epsilon,\lambda}|^2) + \epsilon(|x|^2 + |y_{\epsilon,\lambda}|^2).$$

Then, the difference

$$V(x, t) - \phi(x, t)$$

achieves a maximum at $(x_{\epsilon,\lambda}, t_{\epsilon,\lambda})$.

Similarly, for $\tilde{\phi}$ given by

$$\tilde{\phi}(y, s) = V(x_{\epsilon,\lambda}, t_{\epsilon,\lambda}) - \lambda(2T + t_{\epsilon,\lambda} + s) -$$
$$- \frac{1}{\epsilon^2}(|x_{\epsilon,\lambda} - y|^2 + |t_{\epsilon,\lambda} - s|^2) - \epsilon(|x_{\epsilon,\lambda}|^2 + |y|^2),$$

the difference

$$\tilde{V}(y, s) - \tilde{\phi}(y, s)$$

has a minimum at $(y_{\epsilon,\lambda}, s_{\epsilon,\lambda})$.

Therefore

$$\phi_t(x_{\epsilon,\lambda}, t_{\epsilon,\lambda}) + H(D_x\phi(x_{\epsilon,\lambda}, t_{\epsilon,\lambda}), x_{\epsilon,\lambda}) \le 0,$$

and

$$\tilde{\phi}_s(y_{\epsilon,\lambda}, s_{\epsilon,\lambda}) + H(D_y\tilde{\phi}(y_{\epsilon,\lambda}, s_{\epsilon,\lambda}), y_{\epsilon,\lambda}) \ge 0.$$

Simplifying, we have

$$\lambda + 2\frac{t_{\epsilon,\lambda} - s_{\epsilon,\lambda}}{\epsilon^2} + H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} + 2\epsilon x_{\epsilon,\lambda}, x_{\epsilon,\lambda}) \le 0, \qquad (18)$$

and

$$-\lambda + 2\frac{t_{\epsilon,\lambda} - s_{\epsilon,\lambda}}{\epsilon^2} + H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} - 2\epsilon y_{\epsilon,\lambda}, y_{\epsilon,\lambda}) \ge 0. \qquad (19)$$

From (18) we gather that

$$H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} + 2\epsilon x_{\epsilon,\lambda}, x_{\epsilon,\lambda}) \le -\lambda + \frac{o(1)}{\epsilon}. \qquad (20)$$

By subtracting (18) to (19) we have

$$2\lambda \le H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} - 2\epsilon y_{\epsilon,\lambda}, y_{\epsilon,\lambda}) - H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} + 2\epsilon x_{\epsilon,\lambda}, x_{\epsilon,\lambda})$$

$$\le H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} - 2\epsilon y_{\epsilon,\lambda}, y_{\epsilon,\lambda}) - H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} - 2\epsilon y_{\epsilon,\lambda}, x_{\epsilon,\lambda})$$

$$+ H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} - 2\epsilon y_{\epsilon,\lambda}, x_{\epsilon,\lambda}) - H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} + 2\epsilon x_{\epsilon,\lambda}, x_{\epsilon,\lambda})$$

$$\le \left(C + CH(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} + 2\epsilon x_{\epsilon,\lambda}, x_{\epsilon,\lambda})\right)|x_{\epsilon,\lambda} - y_{\epsilon,\lambda}|$$

$$+ C\epsilon\left(\left|2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} + 2\epsilon x_{\epsilon,\lambda}\right| + \left|2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} - 2\epsilon y_{\epsilon,\lambda}\right|\right)|x_{\epsilon,\lambda} - y_{\epsilon,\lambda}|$$

$$\le \left(\frac{o(1)}{\epsilon} + C\right)(|x_{\epsilon,\lambda} - y_{\epsilon,\lambda}| + |t_{\epsilon,\lambda} - s_{\epsilon,\lambda}|),$$

when $\epsilon \to 0$, which is a contradiction. $\square$

A function u is semiconcave if there exists a constant C such that
$$u(x+y) - 2u(x) + u(x-y) \leq C|y|^2.$$
We assume that $L(x, v)$ satisfies the following estimate
$$L(x+\theta y, v+\eta y) - 2L(x, v) + L(x-\theta y, v-\eta y) \leq (C+CL(x,v))(\theta^2+\eta^2)|y|^2. \tag{21}$$

Proposition 19. *Consider the solution V to the initial value problem* (8). *Then V is semiconcave in x for each fixed time t.*

Proof. We will do the proof for $t = 0$. Fix $\epsilon > 0$. Let $\mathbf{x}$ be a trajectory such that
$$V(x, 0) \geq \int_{-T}^{0} L(\mathbf{x}, \dot{\mathbf{x}})ds + \psi(\mathbf{x}(-T)) - \epsilon.$$
Then we have
$$\int_{-T}^{0} L(\mathbf{x}, \dot{\mathbf{x}})ds \leq C,$$
for some constant uniformly bounded as $\epsilon \to 0$.

Clearly
$$V(x \pm y, 0) \leq \int_{-T}^{0} L(\mathbf{x} \pm y\frac{s+T}{T}, \dot{\mathbf{x}} \pm \frac{y}{T})ds + \psi(\mathbf{x}(-T)).$$
Therefore
$$V(x+y, 0) - 2V(x, 0) + V(x-y, 0) \leq$$
$$\epsilon + \int_{-T}^{0} L(\mathbf{x} + y\frac{s+T}{T}, \dot{\mathbf{x}} + \frac{y}{T}) - 2L(\mathbf{x}, \dot{\mathbf{x}}) + L(\mathbf{x} - y\frac{T+s}{T}, \dot{\mathbf{x}} - \frac{y}{T}) \leq$$
$$C(1 + \int_{-T}^{0} L(\mathbf{x}, \dot{\mathbf{x}})ds)|y|^2 \leq C|y|^2.$$
$\square$

Proposition 20. *Let u be a viscosity solution of $H(D_x u, x) = 0$. Then u is semiconcave.*

Proof. Consider the Hamilton-Jacobi equation.
$$V_t + H(D_x V, x) = 0 \tag{22}$$
with $V(x, -T) = u(x)$. Then $V(x, t) = u(x)$ is a viscosity solution to (22). By the uniqueness result for viscosity solutions we have that $V = u$ is the value function for the initial value problem (8) with $\psi = u$. But then the previous proposition implies semiconcavity. $\square$

Corollary 21. *Let $u : \mathbb{T}^1 \to \mathbb{R}$ be a viscosity solution of* (16). *Then Du satisfies the following jump condition: $D_x u(x^-) - D_x u(x^+) > 0$.*

Proof. Since $f(x) = u - C|x|^2$ is concave, the derivative of f is decreasing. This implies that f' cannot have jump discontinuities upwards. ☐

Theorem 22. *Let H be convex in p. Let u be a viscosity sub-solution of $H(Du, x) = C$ and let $u^\epsilon = u * \eta_\epsilon$ be a standard smoothing. Then:*

$$H(Du^\epsilon(x), x) \leqslant C + O(\epsilon), \tag{23}$$

where $O(\epsilon) = \sup\left|\frac{\partial H}{\partial x}\right| \int_{\mathbb{R}^d} |y| \eta_\epsilon(y) dy$.

Proof. Since the viscosity solutions of $H(Du, x) = C$ are uniformly Lipschitz, we may assume for the purpose of this proof that $\frac{\partial H}{\partial x}$ is bounded.

For any $x \in \mathbb{T}^d$ and any $p, y \in \mathbb{R}^d$ we have $|H(p, x - y) - H(p, x)| \leqslant |y| \sup |D_x H|$.

$$C \geqslant \int \eta_\epsilon(y) H(Du(x - y), x - y) dy \geqslant \int \eta_\epsilon(y) H(Du(x - y), x) dy - O(\epsilon).$$

Now, Jensen's inequality yields

$$\int \eta_\epsilon(y) H(Du(x - y), x) dy \geqslant H\left(\int \eta_\epsilon(y) Du(x - y) dy, x\right) = H(Du^\epsilon(x), x),$$

which completes the proof. ☐

For the unbounded case, $x \in \mathbb{R}^d$, the problem $H(Du, x) = C$ (might) have a viscosity solution (or even a regular solution) for infinitely many C's. Indeed, Let $H(p) = |p|^2$, then for any $P \in \mathbb{R}^d$ the function $u(x) = P \cdot x$ solves $H(Du) = |P|^2$, i.e., $C = |P|^2$. However for the case $x \in \mathbb{T}^d$ the above number C is unique. We will give an elementary proof of the uniqueness of the number $\overline{H}$.

Theorem 23. *Let H be convex in p, $\mathbb{Z}^d$-periodic in x and $\frac{\partial H}{\partial x}$ is bounded. Let C be a real number, such that $H(Du, x) = C$ has a viscosity solution u. Then*

$$C = \inf_{\varphi:smooth} \sup_{x \in \mathbb{T}^d} H(D\varphi(x), x).$$

Proof. Let u be a viscosity solution. Inequality (23) implies

$$\inf_{\varphi:\text{smooth}} \sup_{x \in \mathbb{T}^d} H(D\varphi(x), x) \leqslant C.$$

To show the opposite inequality we take any smooth function $\varphi(x)$. Due to periodicity, the set of points where $u - \varphi$ achieves a local minimum is non empty. For example, one could consider a point of global minimum. Let

x_φ be a local minimum point for $u - \varphi$. The definition of viscosity solution implies $H(D\varphi(x_\varphi), x_\varphi) \geqslant C$. Thus, for any smooth function $\varphi(x)$, we have

$$\sup_x H(D\varphi(x), x) \geqslant C.$$

Taking infimum over φ completes the proof. □

The Theorem does not assert uniqueness of the viscosity solution u. Indeed, as the next examples illustrate, such viscosity solutions are not unique even up to constants, see for instance [Con95] for a detailed discussion and additional examples.

Example 1. A simple example in which there is no uniqueness of solution is the following: let $\psi : \mathbb{T}^d \to \mathbb{R}$ be an arbitrary C^1 function, and consider the Hamilton-Jacobi equation:

$$Du(Du - D\psi) = 0.$$

Clearly this equation admits two solutions, $u = 0$ and $u = \psi$ (modulo constants). ◄

Example 2. Let $P \in \mathbb{R}$. Consider the Hamiltonian

$$H^P(p, x) = \frac{|P + p|^2}{2} - \cos 2\pi x.$$

For $P = 0$ this corresponds to a one-dimensional pendulum with mass and length normalized to unity. We will look for explicit the solutions of

$$\frac{|P + D_x u|^2}{2} - \cos 2\pi x = \overline{H}(P). \tag{24}$$

Proposition 24. *Any solution $(u, \overline{H}(P))$ of (24) satisfies*

$$u(P, x) = \int_0^x -P + s(y)\sqrt{2(\overline{H}(P) + \cos 2\pi y)}\, dy + u(P, 0), \tag{25}$$

where $|s(y)| = 1$, with $\overline{H}(P) = 1$ for $|P| \leq \frac{4}{\pi}$ and

$$P = \pm \int_0^1 \sqrt{2(\overline{H}(P) + \cos 2\pi y)}\, dy, \tag{26}$$

otherwise.

Proof. For a.e. $x \in \mathbb{T}^1$, the solution $u(x)$ satisfies

$$\frac{(P + D_x u)^2}{2} = \overline{H}(P) + \cos 2\pi x.$$

This implies $\overline{H}(P) \geq 1$ and so,

$$D_x u = -P \pm \sqrt{2(\overline{H}(P) + \cos 2\pi x)}, \quad \text{a.e.} \ \ x \in \mathbb{R}.$$

Thus (25) holds, for $|s(y)| = 1$.

By Corollary 21 the only possible discontinuities in the derivative of u are the ones that satisfy $D_x u(x^-) - D_x u(x^+) > 0$. Therefore s can change sign from 1 to -1 at any point but jumps from -1 to 1 can happen only when $\sqrt{2(\overline{H}(P) + \cos 2\pi x)} = 0$.

If we require 1-periodicity there are two cases:

(i) if $\overline{H}(P) > 1$ the solution is C^1 since $\sqrt{2(\overline{H}(P) + \cos 2\pi y)}$ is never zero. In this case P and $\overline{H}(P)$ satisfy (26). It is easy to check that this equation has a solution $\overline{H}(P)$ whenever

$$|P| \geq \int_0^1 \sqrt{2(1 + \cos 2\pi y)} dy,$$

that is, whenever $|P| > \frac{4}{\pi}$.

(ii) Otherwise, when $|P| \leq 4\pi^{-1}$, $\overline{H}(P) = 1$ and $s(x)$ can have a discontinuity. Indeed, $s(x)$ jumps from -1 to 1 when $x = \frac{1}{2} + k$, with $k \in \mathbb{Z}$, and there is a point x_0 defined by the equation

$$\int_0^1 s(y) \sqrt{2(1 + \cos 2\pi y)} dy = P,$$

in which $s(x)$ jumps from 1 to -1.

$\square$

This example also shows that (24) does not have a unique solution. Indeed, since $\cos 2\pi x$ is also 2-periodic, we may look for 2-periodic solutions. In this case we find out that for $|P|$ small, there are two points in which the derivative can be discontinuous and we can choose freely one of them because the only constraint is periodicity. Note, however, that the value of $\overline{H}$ is uniquely determined and is the same whether we look for 1 or 2 periodic solutions. ◀

5. Duality

Let

$$\mathcal{M}_1^+ = \left\{ \mu \in \mathcal{M} : \int_{\mathbb{T}^d \times \mathbb{R}^d} d\mu = 1, \mu \geq 0 \right\}$$

be the subset of the probability measures among all γ-weighted (signed) measures. In this section we show using Fenchel-Legendre-Rockafellar's

theorem that

$$
\inf_{\mu \in \mathcal{M}_1^+} \sup_{\varphi \in C^1(\mathbb{T}^d)} \int \big(L(x,v) - v\nabla\varphi\big)d\mu = \sup_{\varphi \in C^1(\mathbb{T}^d)} \inf_{\mu \in \mathcal{M}_1^+} \int \big(L(x,v) - v\nabla\varphi\big)d\mu.
$$
(27)

We define the set of holonomic measures as follows:

$$
\mathcal{M}_{\text{hol}} = \Big\{ \mu \in \mathcal{M} : \int_{\mathbb{T}^d \times \mathbb{R}^d} v \cdot D_x\varphi(x)d\mu = 0, \quad \forall\, \varphi \in C^1(\mathbb{T}^d) \Big\}.
$$

Consider the set

$$
\mathcal{C} = \Big\{ \psi \in C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d) : \psi(x,v) = v \cdot D_x\varphi(x), \quad \text{for some } \varphi \in C^1(\mathbb{T}^d) \Big\}.
$$
(28)

We observe that

$$
\mathcal{M}_{\text{hol}} = \Big\{ \mu \in \mathcal{M} : \int_{\mathbb{T}^d \times \mathbb{R}^d} \psi\, d\mu = 0, \quad \forall \psi \in \mathcal{C} \Big\}.
$$

Since $\varphi \mapsto vD_x\varphi$ is a linear mapping, $\mathcal{C}$ is a convex set. Using the fact that

$$
\inf_{\mu \in \mathcal{M}_1^+} \int \big(L(x,v) - v \cdot \nabla\varphi\big)d\mu = \min_{x,v}\big(L(x,v) - v \cdot \nabla\varphi\big),
$$

we observe that the R.H.S. of (27) can be written as

$$
\sup_{\varphi \in C^1(\mathbb{T}^d)} \inf_{\mu \in \mathcal{M}_1^+} \int \big(L(x,v) - v \cdot \nabla\varphi\big)d\mu = \sup_{\psi(x,v)} f(\psi),
$$

where $f : C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d) \to \mathbb{R}$ is given by

$$
f(\psi) = \begin{cases} \min_{x,v}\big(L(x,v) + \psi(x,v)\big), & \text{if } \psi \in \mathcal{C}, \\ -\infty, & \text{otherwise.} \end{cases}
$$

and we can further decompose f as $f(\psi) = g(\psi) - h(\psi)$, where

$$
g(\psi) = \min_{x,v}\big(L(x,v) + \psi(x,v)\big),
$$

and

$$
h(\psi) = \begin{cases} 0, & \text{if } \psi \in \mathcal{C}, \\ +\infty, & \text{otherwise.} \end{cases}
$$

Proposition 25. *The function h is convex. The function g is concave and continuous.*

Proof. Since $\mathcal{C}$ is a convex set, the function h is convex. The function g is concave as it is the infimum of affine functions. It remains to show the continuity of g.

Let $\psi_n \to \psi$ in C_0^γ. Then $\|\psi_n\|_\gamma$ and $\|\psi\|_\gamma$ are uniformly bounded by some constant C. The growth condition in v for the function L implies the existence of $R > 0$ such that

$$\min_{\mathbb{T}^d \times \mathbb{R}^d} (L + \hat{\psi}) = \min_{\mathbb{T}^d \times B_R} (L + \hat{\psi}),$$

for all $\hat{\psi}$ in $C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)$ with $\|\hat{\psi}\|_\gamma < C$. Here B_R is the ball in $\mathbb{R}^d$ of the radius R centered at the origin. In B_R, $\psi_n \to \psi$ uniformly, therefore

$$\min_{\mathbb{T}^d \times \mathbb{R}^d} (L + \psi_n) \to \min_{\mathbb{T}^d \times \mathbb{R}^d} (L + \psi).$$

$\square$

Proposition 26. *We have*

$$h^*(\mu) = \begin{cases} 0, & if \quad \mu \in \mathcal{M}_{\text{hol}} \\ +\infty & otherwise, \end{cases}$$

and

$$g^*(\mu) = \begin{cases} -\int L d\mu & if \quad \mu \in \mathcal{M}_1^+ \\ -\infty & otherwise. \end{cases}$$

Proof. By Proposition 25 the function h is convex. Therefore, using the definition of the Legendre-Fenchel transform for convex functions and then the definition of h we have

$$h^*(\mu) = \sup_{\psi \in C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)} \left(\int \psi d\mu - h(\psi) \right) = \sup_{\psi \in \mathcal{C}} \int \psi d\mu.$$

If $\mu \in \mathcal{M}_{\text{hol}}$ then $\int \psi d\mu = 0$, for all $\psi \in \mathcal{C}$. Hence $h^*(\mu) = 0$ if $\mu \in \mathcal{M}_{\text{hol}}$. If $\mu \notin \mathcal{M}_{\text{hol}}$ then there exists $\hat{\phi} \in \mathcal{C}$ such that $\int \hat{\phi} d\mu \neq 0$. Thus

$$h^*(\mu) = \sup_{\psi \in \mathcal{C}} \int \psi d\mu \geq \sup_{\alpha \in \mathbb{R}} \alpha \int \hat{\phi} d\mu = +\infty.$$

By Proposition 25 the function g is concave. Therefore, by the definition of the Legendre-Fenchel transform for concave functions, we have

$$g^*(\mu) = \inf_{\psi \in C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)} \left(\int \psi d\mu - g(\psi) \right).$$

First we prove that if μ is not a non-negative measure then $g^*(\mu) = -\infty$. Indeed, take $\psi_1 \geq 0$ such that $\int \psi_1 d\mu < 0$. Then, for $\psi_n = n\psi_1$, we have $\int \psi_n d\mu \to -\infty$ as $n \to \infty$. Observing that $g(\psi)$ is positive, if $\psi \geq 0$ (since $L \geq 1$), we have

$$g^*(\mu) \leq \inf_n \left\{ \int \psi_n d\mu - g(\psi_n) \right\} = -\infty.$$

Now, and for the rest of the proof, assume that $\mu \geqslant 0$. We will show that

$$g^*(\mu) \leqslant - \int L d\mu + \inf_{\psi \in C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)} \left(\int \psi d\mu - \min \psi \right). \qquad (29)$$

Note that (29) implies $g^*(\mu) \leq - \int L d\mu$ by taking $\psi = 0$.

To establish (29), let L_n be a sequence of functions in $C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)$ increasing pointwise to L. Any function ϕ in $C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)$ can be written as $\phi = \psi - L_n$, for some ψ also in $C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)$. Thus

$$g^*(\mu) = \inf_{\psi \in C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)} \left(\int (\psi - L_n) d\mu - g(\psi - L_n) \right) =$$

$$- \int L_n d\mu + \inf_{\psi \in C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)} \left(\int \psi d\mu - \min_{x,v}(\psi + L - L_n) \right).$$

Since $L - L_n \geqslant 0$ we obtain

$$g^*(\mu) \leqslant - \int L_n d\mu + \inf_{\psi \in C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)} \left(\int \psi d\mu - \min \psi \right).$$

By the Monotone Convergence Theorem $\int L_n d\mu \to \int L d\mu$ and we arrive at (29).

If $\int L d\mu = +\infty$ then $g^*(\mu) = -\infty$. On the other hand, if $\int d\mu \neq 1$ then

$$\inf_{\psi \in C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)} \left(\int \psi d\mu - \min \psi \right) \leq \inf_{\alpha \in \mathbb{R}} \alpha \left(\int d\mu - 1 \right) = -\infty,$$

by choosing $\psi = \alpha$, constant. Thus, if the mass of μ is not one, then (29) implies $g^*(\mu) = -\infty$.

If $\int L d\mu < \infty$ then adding and subtracting this term from the definition of $g^*(\mu)$ we can write

$$g^*(\mu) = - \int L d\mu + \inf_{\psi \in C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)} \left(\int (L + \psi) d\mu - \min_{x,v}(L + \psi) \right).$$

Additionally, if $\int d\mu = 1$, then we have $\int (L + \psi) d\mu \geqslant \min(L + \psi)$ for each ψ. Therefore $g^*(\mu) \geq - \int L d\mu$. $\qquad \square$

Observe that

$$\inf_{\mu \in \mathcal{M}} (h^*(\mu) - g^*(\mu)) = \inf_{\mu \in \mathcal{M}_1^+ \cap \mathcal{M}_{\text{hol}}} \int L d\mu = \text{L.H.S. of (27)} .$$

We are now in position to apply the Legendre-Fenchel-Rockafellar duality theorem with $E = C_0^\gamma$, $E' = \mathcal{M}$ and g and h as above. Applying it we arrive at (27).

5.1. **Closeness remarks.** In this sub section we show that $\mathcal{M}_1^+$, $\mathcal{M}_{\mathrm{hol}}$, and $\mathcal{C}$ are not only convex sets, but also they are closed sets.

Proposition 27. *Both $\mathcal{M}_1^+$ and $\mathcal{M}_{\mathrm{hol}}$ are closed subsets of $\mathcal{M}$. The set $\mathcal{C}$ is closed in $C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)$.*

Proof. The set $\mathcal{M}_1^+$ is the intersection of a cone of nonnegative measures (which is closed) with a kernel of a linear continuous function on $\mathcal{M}$. The set $\mathcal{M}_{\mathrm{hol}}$ is an intersection of kernels of a family of linear continuous functions on $\mathcal{M}$.

Now we will prove that $\mathcal{C}$ is closed in $C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)$. Let $\psi_n(x, v) = v \cdot \nabla\varphi_n(x)$ be a sequence in $\mathcal{C}$, $\varphi_n \in C^1(\mathbb{T}^d)$. Assume that the sequence ψ_n converges in $C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d)$ to a function ψ_∞. We have to prove that there is a function $\varphi_\infty \in C^1(\mathbb{T}^d)$ such that

$$\psi_\infty = v \cdot \nabla\varphi_\infty. \tag{30}$$

For any $k = 1, \ldots, d$ and $n \in \mathbb{N}$ define

$$\xi_{n,k}(x) = \psi_n(x, \hat{v}_k)$$

where $\hat{v}_k$ is the unit vector in k^{th} direction. We have $\xi_{n,k} = \frac{\partial}{\partial x_k}\varphi_n$. The sequence of the continuous vector fields

$$\{\xi_{n,1}, \xi_{n,2}, \ldots, \xi_{n,d}\}_{n=1}^\infty$$

satisfies the following two properties:

(1) the vector field the integral over every closed loop in $\mathbb{T}^d$ is zero for any n (because $\xi_{n,k} = \frac{\partial}{\partial x_k}\varphi_n$);
(2) it converges uniformly as $n \to \infty$.

Therefore the limit $\{\xi_{\infty,1}, \xi_{\infty,2}, \ldots, \xi_{\infty,d}\}$ is a continuous vector field such that the integral over every closed loop in $\mathbb{T}^d$ is zero and $\psi_\infty(x, v) = v \cdot \xi_\infty(x)$. Fix $x_0 \in \mathbb{T}^d$. The function

$$\varphi_\infty(x) = \int_{x_0}^x \xi_\infty$$

is well defined, is C^1 and satisfies (30). $\square$

As a Corollary we have that h is a convex lower semicontinuous function.

5.2. **First application of duality.** Using the fact that

$$\sup_{\mu \in \mathcal{M}_1^+} \int \big(v \cdot \nabla\varphi(x) - L(x,v)\big)d\mu = \sup_x \sup_v \big(v \cdot \nabla\varphi(x) - L(x,v)\big) =$$

$$\sup_x H(\nabla\varphi(x), x),$$

$$\inf_{\mu \in \mathcal{M}_1^+ \cap \mathcal{M}_{\text{hol}}} \int L(x,v)d\mu(x,v) = \inf_{\mu \in \mathcal{M}_1^+} \sup_{\varphi:\text{smooth}} \int \big(L(x,v) - v \cdot \nabla\varphi(x)\big)d\mu(x,v),$$

and the fact that $\inf\big(-(\ldots)\big) = -\sup(\ldots)$ we see that (27) can be restated as

$$\inf_{\mu \in \mathcal{M}_1^+ \cap \mathcal{M}_{\text{hol}}} \int L(x,v)d\mu(x,v) = -\inf_\varphi \sup_x H(\nabla\varphi(x), x) \qquad (31)$$

Definition. $\overline{H} = -\inf\limits_{\mu \in \mathcal{M}_1^+ \cap \mathcal{M}_{\text{hol}}} \int L(x,v)d\mu(x,v).$

Theorem 28. $C = \overline{H}$ *is the unique number, for which the equation*

$$H(Du, x) = C \qquad (32)$$

admits a periodic viscosity solution.

Remark. Recall, however, that the viscosity solutions to (32) are not unique, as was discussed in the previous section.

Proof. Apply Theorem 23 and (31). $\qquad\qquad\square$

Remark. The inequality $\sup_x H(D\varphi(x), x) \geqslant \overline{H}$ for any smooth function φ is elementary. Indeed,

$$\sup_x H(D\varphi(x), x) = \sup_{\mu:\text{prob}} \int_{x,v} H(D\varphi(x), x)d\mu \geqslant$$

$$\sup_{\substack{\mu:\text{prob} \\ \text{holonomy}}} \int_{x,v} H(D\varphi(x), x)d\mu(x,v).$$

Using the inequality $H(D\varphi(x), x) \geqslant v \cdot D\varphi(x) - L(x,v)$ we are done.

Remark. $\inf\limits_{\varphi:\text{smooth}} \sup\limits_x H(D\varphi(x), x) \leqslant \overline{H}.$

Take any viscosity solution u of $H(Du, x) = \overline{H}$ (it exists by Thm 28). Let u^ϵ be, as before, a standard smoothing. Applying Theorem 22 above we have $H(Du^\epsilon(x), x) \leqslant \overline{H} + o(\epsilon)$.

Theorem 29. *The graph from Theorem 3 can be chosen to be bounded.*

Proof. Let $u = u(x)$ be a viscosity solution of $H(Du, x) = \overline{H}$. From Theorem 22 we have : Du^ϵ is uniformly bounded for $\epsilon \in (0, 1)$, that is there exists K such that

$$|Du^\epsilon| < K.$$

We also have $L(x, v) - v \cdot Du^\epsilon \geqslant -H(Du^\epsilon(x), x) \geqslant -\overline{H} - o(\epsilon)$. Let μ be a minimizing measure. We have:

$$-\overline{H} = \int L(x, v) - v \cdot Du^\epsilon d\mu \geqslant \inf_{\mu \in \mathcal{M}_1^+} \int L(x, v) - v \cdot Du^\epsilon d\mu \geqslant -\overline{H} - o(\epsilon).$$

Let $M = \sup\{|v| : L(x, v) - |v|K - \max_x L(x, 0) \leqslant 0\}$, then $v(x)$ is bounded by M. Indeed, assuming the contrary we have, that some positive mass of the measure μ is in $\{x, v : |v| > M + \delta\}$. If we collapse (project) this part of measure to $v = 0$, then we obtain loss in action which is separated away from zero, and would be large than $o(\epsilon)$, provided ϵ is small enough. $\qquad\square$

6. **Regularity**

In this section we present (with small adaptations) the regularity results for viscosity solutions in the support of the Mather measures by [EG01]. We should point out that the proofs of Theorems 31–37 presented here appeared in [EG01]. For the setting of this survey, we had to add an elementary lemma, Lemma 30, for the presentation to be self-contained, as our definition of Mather measures differs from the one used in [EG01].

Lemma 30. *Let μ be a minimizing holonomic measure. Then*

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} D_x L(x, v) d\mu = 0.$$

Proof. Let $h \in \mathbb{R}^d$, consider the measure μ_h on $\mathbb{T}^d \times \mathbb{R}^d$ given by

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \phi(x, v) d\mu_h = \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi(x + h, v) d\mu,$$

for all continuous and compactly supported function $\phi : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$. Clearly, for every h, μ_h is holonomic. Since μ is minimizing, it follows

$$\frac{d}{d\epsilon} \int L(x + \epsilon h, v) d\mu \Big|_{\epsilon=0} = 0,$$

that is,

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} D_x L(x, v) h d\mu = 0.$$

Since $h \in \mathbb{R}$ is arbitrary, the statement of the Lemma follows. $\qquad\square$

It will be convenient to define the measure $\tilde{\mu}$ on $\mathbb{T}^d \times \mathbb{R}^d$ as the push forward measure of the measure μ with respect to the one to one map $(v, x) \mapsto (p, x)$, where $p = D_v L(v, x)$. In other words we define the measure $\tilde{\mu}$ on $\mathbb{T}^d \times \mathbb{R}^d$ to be

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \phi(x, p) d\tilde{\mu} = \int_{\mathbb{T}^d \times \mathbb{R}^d} \phi(x, D_v L(x, v)) d\mu.$$

We also define projection $\bar{\mu}$ in $\mathbb{T}^d$ of a measure μ in $\mathbb{T}^d \times \mathbb{R}^d$ as

$$\int_{\mathbb{T}^d} \varphi(x) d\bar{\mu}(x) = \int_{\mathbb{T}^d \times \mathbb{R}^d} \varphi(x) d\mu(x, v).$$

Note that, in similar way, $\bar{\mu}$ is also the projection of the measure $\tilde{\mu}$. Observe that for any smooth function $\varphi(x)$ we have that $\tilde{\mu}$ satisfies the following version of the holonomy condition:

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} D_p H(p, x) D_x \varphi(x) d\tilde{\mu} = 0,$$

because we can use identity (11) if $p = D_v(x, v)$.

Theorem 31. *Let u be any viscosity solution of* (16)*, and let μ be any minimizing holonomic measure. Then $\bar{\mu}$-almost everywhere, $D_x u(x)$ exists and $p = D_x u(x)$, $\tilde{\mu}$-almost everywhere.*

Proof. Let u be any viscosity solution of (16). Let η_ϵ be a standard mollifier, $u^\epsilon = \eta_\epsilon * u$. By strict uniform convexity there exists $\gamma > 0$ such that for any $p, q \in \mathbb{R}^d$ and any $x \in \mathbb{T}^d$ we have

$$H(p, x) \geqslant H(q, x) + D_p H(q, x)(p - q) + \frac{\gamma}{2} |p - q|^2.$$

By Theorem 16, any viscosity solution of (16), and in particular u, is Lipschitz.

Recall that, by Rademacher's theorem [Eva98], a locally Lipschitz function is differentiable Lebesgue almost everywhere. Using $p = D_x u(y)$ and $q = D_x u^\epsilon(x)$, conclude that for every point x and for Lebesgue almost every point y:

$$H(D_x u(y), x) \geq H(D_x u^\epsilon(x), x) + D_p H(D_x u^\epsilon(x), x)(D_x u(y) - D_x u^\epsilon(x)) + $$
$$\frac{\gamma}{2} |D_x u^\epsilon(x) - D_x u(y)|^2.$$

Multiplying the previous identity by $\eta_\epsilon(x - y)$ and integrating over $\mathbb{R}^d$ in y yields

$$H(D_x u^\epsilon(x), x) + \frac{\gamma}{2} \int_{\mathbb{R}^d} \eta_\epsilon(x - y) |D_x u^\epsilon(x) - D_x u(y)|^2 dy \leq$$

$$\int_{\mathbb{R}^d} \eta_\epsilon(x - y)H(D_x u(y), x)dy \leq \overline{H} + O(\epsilon).$$

Let

$$\beta_\epsilon(x) = \frac{\gamma}{2} \int_{\mathbb{R}^d} \eta_\epsilon(x - y)|D_x u^\epsilon(x) - D_x u(y)|^2 dy.$$

Now observe that

$$\frac{\gamma}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} |D_x u^\epsilon(x) - p|^2 d\tilde{\mu} \leq$$

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} [H(D_x u^\epsilon(x), x) - H(p, x) - D_p H(p, x)(D_x u^\epsilon(x) - p)]\, d\tilde{\mu} \leq$$

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} H(D_x u^\epsilon(x), x)d\tilde{\mu} - \overline{H},$$

because

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} D_p H(x, p)D_x u^\epsilon(x) = 0,$$

and

$$p D_p H(x, p) - H(x, p) = L(x, D_p H(x, p)),$$

and $\int_{\mathbb{T}^d \times \mathbb{R}^d} L(x, D_p H(x, p))d\tilde{\mu} = -\overline{H}$. Therefore,

$$\frac{\gamma}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} |D_x u^\epsilon(x) - p|^2 d\tilde{\mu} + \int_{\mathbb{T}^d} \beta_\epsilon(x)d\bar{\mu} \leq O(\epsilon).$$

Thus, for $\bar{\mu}$-almost every point x, $\beta_\epsilon(x) \to 0$. Therefore, $\bar{\mu}$-almost every point is a point of approximate continuity of $D_x u$ (see [EG92], p. 49). Since u is semiconcave (Proposition 20), it is differentiable at points of approximate continuity. Furthermore

$$D_x u^\epsilon \to D_x u$$

pointwise, $\bar{\mu}$-almost everywhere, and so $D_x u$ is $\bar{\mu}$ measurable. Also we have

$$p = Du(x), \quad \tilde{\mu} - \text{almost everywhere}.$$

$\square$

By looking at the proof the previous theorem we can also state the following useful result:

Corollary 32. *Let η_ϵ be a standard mollifier, $u^\epsilon = \eta_\epsilon * u$. Then*

$$\int_{\mathbb{T}^d} |D_x u^\epsilon - D_x u|^2 d\bar{\mu} \leq C\epsilon,$$

as $\epsilon \to 0$.

As a Corollary we formulate an equivalent form of Theorem 31.

Corollary 33. *Let u be any viscosity solution of* (16), *and let μ be any minimizing holonomic measure. Then μ-almost everywhere, $D_x u(x)$ exists and*

$$D_v L(v, x) = D_x u(x) \qquad \mu - almost\ everywhere. \tag{33}$$

and

$$D_x L(v, x) = -D_x H(D_x u(x), x) \qquad \mu - almost\ everywhere. \tag{34}$$

Proof. First we observe that the measure $\tilde{\mu}$ is the push forward measure of the measure μ with respect to the one to one map $(v, x) \mapsto (p, x)$, where $p = D_v L(v, x)$. Therefore an $\tilde{\mu}$ – almost everywhere identity

$$F_1(p, x) = F_2(p, x) \qquad (p, x)\text{-}\tilde{\mu}\ \text{almost everywhere}$$

implies the μ – almost everywhere identity

$$F_1(D_v L(v, x), x) = F_2(D_v L(v, x), x) \qquad (v, x)\text{-}\mu\ \text{almost everywhere}.$$

Thus (33) follows directly from Theorem 31.

Using (33) and the identity $D_x L(v, x) = -D_x H(D_v L(v, x), x)$, we arrive at (34). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We observe that from the previous corollary it also follows

$$\int_{\mathbb{T}^d} D_p H(D_x, x) D_x u \, d\bar{\mu} = 0.$$

Indeed,

$$\int_{\mathbb{T}^d} D_p H(D_x u, x) D_x u \, d\bar{\mu} = \int_{\mathbb{T}^d} D_p H(D_x, x) D_x u^\epsilon \, d\bar{\mu} \ +$$

$$\int_{\mathbb{T}^d} D_p H(D_x u, x) \left(D_x u - D_x u^\epsilon \right) d\bar{\mu}.$$

We have

$$\int_{\mathbb{T}^d} D_p H(D_x, x) D_x u^\epsilon \, d\bar{\mu} = 0.$$

To handle the second term, fix $\delta > 0$. Then

$$\left| \int_{\mathbb{T}^d} D_p H(D_x u, x) \left(D_x u - D_x u^\epsilon \right) \right| \leq \delta \int_{\mathbb{T}^d} |D_p H(D_x u, x)|^2 d\bar{\mu} \ +$$

$$\frac{1}{\delta} \int_{\mathbb{T}^d} \left| D_x u - D_x u^\epsilon \right|^2 d\bar{\mu}.$$

Note that since u is Lipschitz the term $D_p H(D_x u, x)$ is bounded, and so is $\int_{\mathbb{T}^d} |D_p H(D_x u, x)|^2 d\bar{\mu}$. Send $\epsilon \to 0$, and then let $\delta \to 0$.

Theorem 34. *Let u be any viscosity solution of (16), and let μ be any minimizing holonomic measure. Then*

$$\int_{\mathbb{T}^d} |D_x u(x+h) - D_x u(x)|^2 d\bar{\mu} \le C|h|^2.$$

Proof. Applying Theorem 22 we have

$$H(D_x u^\epsilon(x+h), x+h) \le \overline{H} + C\epsilon.$$

By Theorem 31 the derivative $D_x u(x)$ exists $\bar{\mu}$ almost everywhere. By proposition 11 viscosity solution satisfies equation (16) in classical sense at all points of differentiability. Thus $H(D_x u(x), x) = \overline{H}$ for $\bar{\mu}$ almost all points x. Now observe that

$$C\epsilon \ge H(D_x u^\epsilon(x+h), x+h) - H(D_x u(x), x)$$
$$= H(D_x u^\epsilon(x+h), x+h) - H(D_x u^\epsilon(x+h), x) + H(D_x u^\epsilon(x+h), x) -$$
$$H(D_x u(x), x)$$

The term

$$H(D_x u^\epsilon(x+h), x+h) - H(D_x u^\epsilon(x+h), x) =$$
$$D_x H(D_x u^\epsilon(x+h), x)h + O(h^2) =$$
$$D_x H(D_x u(x), x)h + O(h^2 + h|D_x u^\epsilon(x+h) - D_x u(x)|) \ge$$
$$D_x H(D_x u(x), x)h + O(h^2) - \frac{\gamma}{4}|D_x u^\epsilon(x+h) - D_x u(x)|^2.$$

Therefore, for $\bar{\mu}$ almost every x, we have

$$H(D_x u^\epsilon(x+h), x) - H(D_x u, x) \le C\epsilon - D_x H(D_x u(x), x)h +$$
$$\frac{\gamma}{4}|D_x u^\epsilon(x+h) - D_x u(x)|^2 + Ch^2.$$

Since

$$H(D_x u^\epsilon(x+h), x) - H(D_x u, x) \ge \frac{\gamma}{2}|D_x u^\epsilon(x+h) - D_x u(x)|^2 +$$
$$D_p H(D_x u, x)(D_x u^\epsilon(x+h) - D_x u(x))$$

we have

$$\frac{\gamma}{4}\int |D_x u^\epsilon(x+h) - D_x u(x)|^2 d\bar{\mu} \le C\epsilon + C|h|^2 - \int D_x H(D_x u(x), x)h d\bar{\mu}.$$

By (34) and Lemma 30 it follows

$$\int D_x H(D_x u(x), x)h d\bar{\mu} = -\int D_x L(v, x)h d\mu = 0.$$

As $\epsilon \to 0$, through a suitable subsequence (since $D_x u^\epsilon(x+h)$ is bounded in $L^2_{\bar\mu}$), we may assume that $D_x u^\epsilon(x+h) \rightharpoonup \xi(x)$ in $L^2_{\bar\mu}$, for some function $\xi \in L^2_{\bar\mu}$, and

$$\int |\xi - D_x u|^2 d\bar\mu \leq C|h|^2.$$

Finally, we claim that $\xi(x) = D_x u(x+h)$ for $\bar\mu$ almost all x. This follows from Theorem 31 and the fact that for $\bar\mu$ almost all x we have $\xi(x) \in D_x^- u(x+h)$, where D_x^- stands for the subdifferential. To see this, observe that by Proposition 20 u is semiconcave, therefore u^ϵ are uniformly semiconcave, that is

$$u^\epsilon(y+h) \leq u^\epsilon(x+h) + D_x u^\epsilon(x+h)(y-x) + C|y-x|^2,$$

where C is independent of ϵ. Fixing y and integrating against a nonnegative function $\varphi(x) \in L^2_{\bar\mu}$ yields

$$\int_{\mathbb{T}^d} \left(u^\epsilon(y+h) - u^\epsilon(x+h) - D_x u^\epsilon(x+h)(y-x) - C|y-x|^2 \right) \varphi(x) d\bar\mu \leq 0$$

By passing to the limit we have that

$$u(y+h) \leq u(x+h) + \xi(x)(y-x) + C|y-x|^2 \qquad \text{for all } y \text{ and } \bar\mu\text{-almost all } x,$$

that is, $\xi(x) \in D_x^- u(x+h)$ for $\bar\mu$-almost all x. $\qquad\square$

Lemma 35. *Let u be any viscosity solution of (16), and let μ be any minimizing holonomic measure. Let $\psi : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}$ be a smooth function. Then*

$$\int_{\mathbb{T}^d} D_p H(D_x u, x) D_x [\psi(x, u(x))] \, d\bar\mu = 0$$

Proof. Clearly we have

$$\int_{\mathbb{T}^d} D_p H(D_x u, x) D_x [\psi(x, u^\epsilon(x))] \, d\bar\mu = 0.$$

By the uniform convergence of u^ϵ to u, and $L^2_{\bar\mu}$ convergence of $D_x u^\epsilon$ to $D_x u$, see Corollary 32, we get the result. $\qquad\square$

Theorem 36. *Let u be any viscosity solution of (16), and let μ be any minimizing holonomic measure. Then, for $\bar\mu$ almost every x and all $h \in \mathbb{R}^d$,*

$$|u(x+h) - 2u(x) + u(x-h)| \leq C|h|^2.$$

Proof. Let $h \neq 0$ and define

$$\tilde{u}(x) = u(x+h), \qquad \hat{u}(x) = u(x-h).$$

Consider the mollified functions $\tilde{u}^\epsilon, \hat{u}^\epsilon$, where we take

$$0 < \epsilon \leq \eta |h|^2, \tag{35}$$

for small $\eta > 0$. We have

$$H(D\tilde{u}^\epsilon, x + h) \leq \overline{H} + C\epsilon, \qquad H(D\hat{u}^\epsilon, x - h) \leq \overline{H} + C\epsilon.$$

For $\bar{\mu}$-almost every point x, (for which $Du(x)$ exists and therefore $H(Du(x), x) = \overline{H}$) we have

$$H(D\tilde{u}^\epsilon, x) - 2H(Du, x) + H(D\hat{u}^\epsilon, x) \leqslant$$

$$2C\epsilon + H(D\tilde{u}^\epsilon, x) - H(D\tilde{u}^\epsilon, x + h) + H(D\hat{u}^\epsilon, x) - H(D\hat{u}^\epsilon, x - h).$$

Hence

$$\frac{\gamma}{2}(|D\tilde{u}^\epsilon - Du|^2 + |D\hat{u}^\epsilon - Du|^2) + D_pH(Du, x) \cdot (D\tilde{u}^\epsilon - 2Du + D\hat{u}^\epsilon)$$

$$\leqslant C(\epsilon + |h|^2) + (D_xH(D\hat{u}^\epsilon, x) - D_xH(D\tilde{u}^\epsilon, x)) \cdot h.$$

Using the inequality

$$\left|\left(D_xH(p, x) - D_xH(q, x)\right) \cdot h\right| \leqslant \left\|\frac{\partial^2 H}{\partial p \partial x}\right\| |p - q| \, |h| \leqslant \frac{\gamma}{4}|p-q|^2 + \frac{1}{\gamma}\left\|\frac{\partial^2 H}{\partial p \partial x}\right\|^2 |h|^2,$$

where $\left\|\frac{\partial^2 H}{\partial p \partial x}\right\| = \sup\limits_{p,x} \sup\limits_{|z|=1, |h|=1} \sum\limits_{i,j}\left|z_j h_i \frac{\partial^2 H}{\partial p_j \partial x_i}(p, x)\right|$, we arrive at

$$\frac{\gamma}{4}(|D\tilde{u}^\epsilon - Du|^2 + |D\hat{u}^\epsilon - Du|^2) + D_pH(Du, x) \cdot (D\tilde{u}^\epsilon - 2Du + D\hat{u}^\epsilon) \leqslant$$

$$C(\epsilon + |h|^2).$$

Fix now a smooth, nondecreasing, function $\Phi : \mathbb{R} \to \mathbb{R}$, and write $\phi := \Phi' \geqslant 0$. Multiply the last inequality above by $\phi\left(\frac{\tilde{u}^\epsilon - 2u + \hat{u}^\epsilon}{|h|^2}\right)$, and integrate with respect to $\bar{\mu}$:

$$\frac{\gamma}{4}\int_{\mathbb{T}^d}(|D\tilde{u}^\epsilon - Du|^2 + |D\hat{u}^\epsilon - Du|^2)\phi\left(\frac{\tilde{u}^\epsilon - 2u + \hat{u}^\epsilon}{|h|^2}\right)d\bar{\mu} \tag{36}$$

$$+ \int_{\mathbb{T}^d} D_pH(Du, x) \cdot (D\tilde{u}^\epsilon - 2Du + D\hat{u}^\epsilon)\phi(\cdots)\,d\bar{\mu}$$

$$\leqslant C(\epsilon + |h|^2)\int_{\mathbb{T}^d}\phi(\cdots)\,d\bar{\mu}.$$

Now the second term on the left hand side of (36) equals

$$|h|^2 \int_{\mathbb{R}^d}\int_{\mathbb{T}^d} D_pH(p, x) \cdot D_x\left[\Phi\left(\frac{\tilde{u}^\epsilon - 2u + \hat{u}^\epsilon}{|h|^2}\right)\right]d\bar{\mu} \tag{37}$$

and thus, by Lemma 35 it vanishes. So now dropping the above term from (36) and rewriting, we deduce

$$\int_{\mathbb{T}^d} |Du^\epsilon(x+h) - Du^\epsilon(x-h)|^2 \phi\left(\frac{u^\epsilon(x+h) - 2u(x) + u^\epsilon(x-h)}{|h|^2}\right) d\bar{\mu}$$

$$\tag{38}$$

$$\leqslant C(\epsilon + |h|^2) \int_{\mathbb{T}^d} \phi\left(\frac{u^\epsilon(x+h) - 2u(x) + u^\epsilon(x-h)}{|h|^2}\right) d\bar{\mu}.$$

We confront now a technical problem, as (38) entails a mixture of first-order difference quotients for Du^ϵ and second-order difference quotients for u, u^ϵ. We can however relate these expressions, since u is semiconcave.

To see this, first of all define

$$E_\epsilon := \{x \in \mathrm{supp}(\mu) \mid u^\epsilon(x+h) - 2u(x) + u^\epsilon(x-h) \leqslant -\kappa|h|^2\}, \tag{39}$$

the large constant $\kappa > 0$ to be fixed below. The functions

$$\bar{u}(x) := u(x) - \frac{\alpha}{2}|x|^2, \quad \bar{u}^\epsilon(x) := u^\epsilon(x) - \frac{\alpha}{2}|x|^2 \tag{40}$$

are concave. Also a point $x \in \mathrm{supp}(\bar{\mu})$ belongs to E_ϵ if and only if

$$\bar{u}^\epsilon(x+h) - 2\bar{u}(x) + \bar{u}^\epsilon(x-h) \leqslant -(\kappa + \alpha)|h|^2. \tag{41}$$

Set

$$f^\epsilon(s) := \bar{u}^\epsilon\left(x + s\frac{h}{|h|}\right) \qquad (-|h| \leqslant s \leqslant |h|). \tag{42}$$

Then f is concave, and

$$\bar{u}^\epsilon(x+h) - 2\bar{u}^\epsilon(x) + \bar{u}^\epsilon(x-h) = f^\epsilon(|h|) - 2f^\epsilon(0) + f^\epsilon(-|h|)$$

$$= \int_{-|h|}^{|h|} f^{\epsilon''}(x)(|h| - |s|)\,ds$$

$$\geqslant |h| \int_{-|h|}^{|h|} f^{\epsilon''}(s)\,ds \quad (\text{since } f^{\epsilon''} \leqslant 0)$$

$$= |h|\left[f^{\epsilon'}(|h|) - f^{\epsilon'}(-|h|)\right]$$

$$= (D\bar{u}^\epsilon(x+h) - D\bar{u}^\epsilon(x-h)) \cdot h.$$

Consequently, if $x \in E_\epsilon$, this inequality and (41) together imply

$$2|\bar{u}^\epsilon(x) - \bar{u}(x)| + |D\bar{u}^\epsilon(x+h) - D\bar{u}^\epsilon(x-h)||h| \geqslant (\kappa + \alpha)|h|^2.$$

Now $|\bar{u}^\epsilon(x) - \bar{u}(x)| \leqslant C\epsilon$ on $\mathbb{T}^d$, since u is Lipschitz continuous. We may therefore take η in (35) small enough to deduce from the foregoing that

$$|D\bar{u}^\epsilon(x+h) - D\bar{u}^\epsilon(x-h)| \geqslant (\frac{\kappa}{2} + \alpha)|h|. \tag{43}$$

But then

$$|Du^\epsilon(x+h) - Du^\epsilon(x-h)| \geqslant (\frac{\kappa}{2} - \alpha)|h|. \tag{44}$$

Return now to (38). Taking $\kappa > 2\alpha$ and

$$\phi(z) = \begin{cases} 1 & \text{if } z \leqslant -\kappa \\ 0 & \text{if } z > -\kappa. \end{cases}$$

The inequality (38) was derived for smooth functions ϕ. However, by replacing ϕ in (38) by a sequence ϕ_n of smooth functions increasing pointwise to ϕ, and using the monotone convergence theorem, we conclude that (38) holds for this function ϕ. Then we discover from (38) that

$$(\frac{\kappa}{2} - \alpha)^2 |h|^2 \bar{\mu}(E_\epsilon) \leqslant C(\epsilon + |h|^2)\bar{\mu}(E_\epsilon).$$

We fix κ so large that

$$(\frac{\kappa}{2} - \alpha)^2 \geqslant C + 1,$$

to deduce

$$(|h|^2 - C\epsilon)\bar{\mu}(E_\epsilon) \leqslant 0.$$

Thus $\bar{\mu}(E_\epsilon) = 0$ if η in (35) is small enough, and this means

$$u^\epsilon(x+h) - 2u(x) + u^\epsilon(x-h) \geqslant -\kappa|h|^2$$

for $\bar{\mu}$-almost every point x. Now let $\epsilon \to 0$:

$$u(x+h) - 2u(x) + u(x-h) \geqslant -\kappa|h|^2$$

$\bar{\mu}$-almost everywhere Since

$$u(x+h) - 2u(x) + u(x-h) \leqslant \alpha|h|^2$$

owing to the semiconcavity, we have

$$|u(x+h) - 2u(x) + u(x-h)| \leq C|h|^2$$

for $\bar{\mu}$-almost every point x. As u is continuous, the same inequality obtains for all $x \in \text{supp}(\bar{\mu})$. $\qquad\square$

Now we state and prove the main result of this section.

Theorem 37. *Let u be any viscosity solution of* (16)*, and let μ be any minimizing holonomic measure. Then for $\bar{\mu}$-almost every x, $D_x u(x)$ exists and for Lebesgue almost every y*

$$|D_x u(x) - D_x u(y)| \leq C|x - y|. \tag{45}$$

Proof. First we show that

$$|u(y) - u(x) - (y - x) \cdot D_x u(x)| \leqslant C|x - y|^2. \tag{46}$$

Fix $y \in \mathbb{R}^d$ and take any point $x \in \operatorname{supp}(\bar{\mu})$ at which u is differentiable.

According to Theorem 36 with $h := y - x$, we have

$$|u(y) - 2u(x) + u(2x - y)| \leqslant C|x - y|^2. \tag{47}$$

By semiconcavity, we have

$$u(y) - u(x) - Du(x) \cdot (y - x) \leqslant C|x - y|^2, \tag{48}$$

and also

$$u(2x - y) - u(x) - Du(x) \cdot (2x - y - x) \leqslant C|x - y|^2. \tag{49}$$

Use (49) in (47):

$$u(y) - u(x) - Du(x) \cdot (y - x) \geqslant -C|x - y|^2.$$

This and (48) establish (46).

Estimate (45) follows from (46), as follows. Take x, y as above. Let z be a point to be selected later, with $|x - z| \leqslant 2|x - y|$. The semiconcavity of u implies that

$$u(z) \leqslant u(y) + Du(y) \cdot (z - y) + C|z - y|^2. \tag{50}$$

Also,

$$u(z) = u(x) + Du(x) \cdot (z - x) + O(|x - z|^2),$$
$$u(y) = u(x) + Du(x) \cdot (y - x) + O(|x - y|^2),$$

according to (46). Insert these indentities into (50) and simplify:

$$(Du(x) - Du(y)) \cdot (z - y) \leqslant C|x - y|^2.$$

Now take

$$z := y + |x - y| \frac{Du(x) - Du(y)}{|Du(x) - Du(y)|}$$

to deduce (45).

Now take any point $x \in \operatorname{supp}(\bar{\mu})$, and fix y. There exist points $x_k \in \operatorname{supp}(\bar{\mu})$ $(k = 1, \dots)$ such that $x_k \to x$ and u is differentiable at x_k. According to estimate (46)

$$|u(y) - u(x_k) - Du(x_k) \cdot (y - x_k)| \leqslant C|x_k - y|^2 \quad (k = 1, \dots).$$

The constant C does not depend on k or y. Now let $k \to \infty$. Owing to (45) we see that $\{Du(x_k)\}$ converges to some vector η, for which

$$|u(y) - u(x) - \eta \cdot (y - x)| \leqslant C|x - y|^2.$$

Consequently u is differentiable at x and $Du(x) = \eta$. $\qquad\square$

It follows from Theorem 37 that function $\boldsymbol{v}$ defined by Theorem 3 is Lipschitz on a set of full measure $\bar{\mu}$. Indeed, by substituting the L.H.S. and the R.H.S. of (33) into $H_p(p, x) = H_p(p, x)$ in place of p's and using (11) we have

$$\boldsymbol{v}(x) = D_p H(Du(x), x) \quad \bar{\mu} \text{ almost everywhere.}$$

We can then extend $\boldsymbol{v}$ as a Lipschitz function to the support of μ, which is contained in the closure of this set of full measure. Note that any Lipschitz function φ defined on a closed set K can be extended to a globally defined Lipschitz function $\hat{\varphi}$ in the following way: without loss of generality assume that $Lip(\varphi) = 1$; define

$$\hat{\varphi}(x) = \inf_{y \in K} \varphi(y) + 2d(x, y).$$

An easy exercise then shows that $\hat{\varphi} = \varphi$ in K and that $\hat{\varphi}$ is Lipschitz. Therefore we may assume that $\boldsymbol{v}$ is globally defined and Lipschitz.

7. Holonomy variations

In this section we study a class of variations that preserve the holonomy constraint. These variations will be used in the last section of this paper to establish the invariance under the Euler-Lagrange flow of minimizing holonomic measures.

Let $\xi : \mathbb{T}^d \to \mathbb{R}^d$, $\xi(x)$ be a C^1 vector field on $\mathbb{T}^d$. Let $\Phi(t, x)$ be the flow by ξ, i.e.,

$$\Phi(0, x) = x, \qquad \text{and} \qquad \tfrac{\partial}{\partial t} \Phi(t, x) = \xi\big(\Phi(t, x)\big).$$

Consider the prolongation of ξ to $\mathbb{T}^d \times \mathbb{R}^d$, which is the vector field on $\mathbb{T}^d \times \mathbb{R}^d$ given by

$$\dot{x}_k(x, v) = \xi_k(x), \quad \dot{v}_k(x, v) = v_i \frac{\partial \xi_k}{\partial x_i}(x). \tag{51}$$

Lemma 38. *The flow of* (51) *is given by*

$$X_k(t, x, v) = \Phi_k(t, x), \qquad V_k(t, x, v) = v_s \frac{\partial \Phi_k}{\partial x_s}(t, x). \tag{52}$$

Proof. Since the X-part of the flow coincides with the Φ-flow, it only remains to show that

$$V(0, x, v) = v, \quad \text{and} \quad \tfrac{\partial}{\partial t} V(t, x, v) = \dot{v}\big(X(t, x, v), V(t, x, v)\big).$$

The first statement ($V(0, x, v) = v$) is clear since the map $x \mapsto \Phi(0, x)$ is the identity map. The second statement can be rewritten as

$$\tfrac{\partial}{\partial t} V_k(t, x, v) = V_i(t, x, v) \left.\frac{\partial \xi_k}{\partial x_i}\right|_{\Phi(t,x)}.$$

A simple computations yields

$$\tfrac{\partial}{\partial t} V_k(t, x, v) = v_s \tfrac{\partial}{\partial x_s} \big(\tfrac{\partial}{\partial t} \Phi_k(t, x)\big) = v_s \tfrac{\partial}{\partial x_s} \Big(\xi_k\big(\Phi(t, x)\big)\Big) =$$

$$v_s \left.\frac{\partial \xi_k}{\partial x_i}\right|_{\Phi(t,x)} \left.\frac{\partial \Phi_i}{\partial x_s}\right|_{(t,x)} = V_i(t, x, v) \left.\frac{\partial \xi_k}{\partial x_i}\right|_{\Phi(t,x)},$$

which is the desired identity. □

For any real number t and any function $\psi(x, v)$, define a new function ψ_t as follows

$$\psi_t(x, v) = \psi\big(X(t, x, v), V(t, x, v)\big). \tag{53}$$

Thus the flow (52) generates the flow on space of functions $\psi(x, v)$ given by (53).

Lemma 39. *The set $\mathcal{C}$, defined in* (28)*, is invariant under the flow given by* (53)*.*

Proof. Let $g \in C^1(\mathbb{T}^d)$ be such that $\psi(x, v) = v_i \frac{\partial}{\partial x_i} g(x)$. Let g_t denote the flow by Φ of the function g, i.e., $g_t(x) = g\big(\Phi(t, x)\big)$. We claim that for any real number t we have

$$\psi_t(x, v) = v_i \frac{\partial}{\partial x_i} g_t(x),$$

where ψ_t is given by (53). Indeed,

$$\psi_t(x, v) = V_k(t, x, v) \left.\frac{\partial g}{\partial x_k}\right|_{X(t,x,v)} = v_s \left.\frac{\partial g}{\partial x_k}\right|_{\Phi(t,x)} \left.\frac{\partial \Phi_k}{\partial x_s}\right|_{(t,x)} =$$

$$v_s \frac{\partial}{\partial x_s} \Big(g\big(\Phi(t, x)\big)\Big) = v_s \frac{\partial}{\partial x_s} g_t(x),$$

and so the Lemma is proven. □

The flow on functions (53) generates the flow on measures: $(t, \mu) \mapsto \mu_t$, where

$$\int \psi d\mu_t = \int \psi_t d\mu. \tag{54}$$

Lemma 40. *The flow* (54) *preserves the holonomy constraint.*

Proof. Let μ be a holonomic measure. We have to prove that μ_t is also a holonomic, i.e., $\int \psi d\mu_t = 0$ for any $\psi \in \mathcal{C}$. This is clear since the flow (53) preserves the set $\mathcal{C}$. $\qquad \square$

Theorem 41. *Let μ be a minimizing measure for the action* (3), *subject to the holonomy constraint. Then for any C^1 vector field $\xi : \mathbb{T}^d \to \mathbb{R}^d$ we have*

$$\int \frac{\partial L}{\partial x_s}\xi_s + \frac{\partial L}{\partial v_s}v_k\frac{\partial}{\partial x_k}\xi_s d\mu = 0. \tag{55}$$

Proof. Let μ_t be the flow generated from μ by (54). Relation (55) expesses the fact $\frac{d}{dt}\left(\int L(x,v)d\mu_t\right)\big|_{t=0} = 0$. $\qquad \square$

8. Invariance

In this section we present a new proof of the invariance under the Euler-Lagrange flow of minimal holonomic measures.

Lemma 42. *Let μ be a measure on a manifold M. Let χ be a smooth vector field on M. The measure μ is invariant with respect to the flow generated by the vector field χ iff for any smooth compactly supported function $\xi : M \to \mathbb{R}$ we have*

$$\int_M \nabla\xi \cdot \chi d\mu = 0.$$

Proof. Let Φ_t be the flow, generated by the vector field χ. Then if μ is invariant under Φ_t, for any smooth compactly supported function $\xi(x)$ and any $t > 0$ we have

$$\int \xi\big(\Phi_t(x)\big) - \xi(x)d\mu = 0.$$

By differentiating with respect to t, and setting $t = 0$, we obtain the "only if" part of the theorem.

To establish the converse, we have to prove that for any t the measure μ_t is well-defined as

$$\mu_t(S) = \mu\big((\Phi_t)^{-1}(S)\big).$$

and coincides with μ.

By the Riesz representation theorem it is sufficient to check that the identity

$$\int \xi d\mu = \int \xi d\mu_t$$

holds for any continuous function ξ (vanishing at ∞). Any continuous function can be uniformly approximated by smooth functions. Therefore it is sufficient to prove the above identity for smooth functions ξ with compact support.

Assume, without loss of generality, that $\xi(x)$ is a C^2-smooth function. Fix $t > 0$. We have to prove that

$$\int \xi\big(\Phi_t(x)\big) - \xi(x)d\mu = 0.$$

We have

$$\int \xi\big(\Phi_t(x)\big) - \xi(x)d\mu = \sum_{k=0}^{N-1} \int \xi\big(\Phi_{t(k+1)/N}(x)\big) - \xi\big(\Phi_{tk/N}(x)\big)d\mu =$$

$$\sum_{k=0}^{N-1} \int \xi_k\big(\Phi_{t/N}(x)\big) - \xi_k(x)d\mu \ ,$$

where $\xi_k(x) = \xi\big(\Phi_{tk/N}(x)\big)$

$$\sum_{k=0}^{N-1} \int \xi_k\big(\Phi_{t/N}(x)\big) - \xi_k(x)d\mu = \sum_{k=0}^{N-1} \int \nabla \xi_k(x) \cdot \big(\Phi_{t/N}(x) - x\big) + O(\tfrac{t}{N^2})d\mu =$$

$$= \sum_{k=0}^{N-1} \int \nabla \xi_k(x) \cdot \big(\tfrac{t}{N}\chi(x) + O(\tfrac{t}{N^2})\big) + O(\tfrac{t}{N^2})d\mu =$$

$$\tfrac{t}{N} \sum_{k=0}^{N-1} \int \nabla \xi_k(x) \cdot \chi(x)d\mu + O(\tfrac{t}{N}) = O(\tfrac{t}{N}).$$

Taking the limit $N \to \infty$ we complete the proof. $\qquad\square$

Now turn to the case where the manifold M is $\mathbb{T}^d \times \mathbb{R}^d$ with coordinates (x, v). In what follows $(\ )_{js}^{-1}$ denotes the j, s entry of the inverse matrix. We will only use this notation for symmetric matrices, thus, this notation will not lead to any ambiguity. Before stating and proving the main Theorem of this section, we will prove an auxiliary lemma.

Lemma 43. *Let μ be a minimal holonomic measure. Let $v^\epsilon(x)$ be any smooth function. Let $\phi(x,v)$ be any smooth compactly supported function. Then*

$$
\int v_k \frac{\partial \phi}{\partial x_k}\big(x, v^\epsilon(x)\big) +
$$

$$
\frac{\partial \phi}{\partial v_j}\big(x, v^\epsilon(x)\big) \left(\frac{\partial^2 L}{\partial^2 v}\right)^{-1}_{js}\big(x, v^\epsilon(x)\big)\left(\frac{\partial L}{\partial x_s}(x,v) - v_k \frac{\partial^2 L}{\partial x_k \partial v_s}\big(x, v^\epsilon(x)\big)\right) d\mu =
$$

$$
\int v_k \frac{\partial}{\partial x_k}\Big(\phi\big(x, v^\epsilon(x)\big)\Big) d\mu - \int v_k \frac{\partial}{\partial x_k}\Big(\frac{\partial L}{\partial v_s}\big(x, v^\epsilon(x)\big)\dot{X}^\epsilon_s\Big) d\mu +
$$

$$
\int v_k \Big(\frac{\partial L}{\partial v_s}\big(x, v^\epsilon(x)\big) - \frac{\partial L}{\partial v_s}(x,v)\Big)\frac{\partial}{\partial x_k}\big(\dot{X}^\epsilon_s\big) d\mu, \quad (56)
$$

where $\dot{X}^\epsilon_s$ is a function of x only (does not depend on v), and is defined as follows:

$$
\dot{X}^\epsilon_s(x) = \frac{\partial \phi}{\partial v_j}\big(x, v^\epsilon(x)\big)\left(\frac{\partial^2 L}{\partial^2 v}\right)^{-1}_{js}\big(x, v^\epsilon(x)\big).
$$

Remark. We will only use this lemma for the case when v^ϵ is the standard smoothing of the function $\boldsymbol{v}(x)$, that is, $v^\epsilon = \eta_\epsilon * \boldsymbol{v}$, where η_ϵ is a standard mollifier. The function $\boldsymbol{v}(x)$ is the function whose graph contains the support of μ, given in Theorem 3. This explains the notation v^ϵ.

Proof. This Lemma is based on Theorem 41. In this proof and bellow v^ϵ stands for the function $v^\epsilon(x)$. We have:

$$
v_k \frac{\partial \phi}{\partial x_k}\big(x, v^\epsilon(x)\big) = v_k \frac{\partial}{\partial x_k}\Big(\phi\big(x, v^\epsilon(x)\big)\Big) - v_k \frac{\partial \phi}{\partial v_j}\big(x, v^\epsilon(x)\big)\frac{\partial v^\epsilon_j}{\partial x_k}(x).
$$

Rewrite the last term:

$$
\begin{aligned}
v_k \frac{\partial \phi}{\partial v_j}\big(x, v^\epsilon(x)\big)\frac{\partial v^\epsilon_j}{\partial x_k}(x) &= v_k \frac{\partial \phi}{\partial v_j}(x, v^\epsilon)\left(\frac{\partial^2 L}{\partial^2 v}\right)^{-1}_{js}(x, v^\epsilon)\frac{\partial^2 L}{\partial v_s \partial v_q}(x, v^\epsilon)\frac{\partial v^\epsilon_q}{\partial x_k}(x) \\
&= v_k \dot{X}^\epsilon_s(x)\frac{\partial^2 L}{\partial v_s \partial v_q}(x, v^\epsilon)\frac{\partial v^\epsilon_q}{\partial x_k}(x).
\end{aligned}
$$

Plug these two lines into (56). And therefore we reduce (56) to

$$
\int \dot{X}^\epsilon_s(x)\left(\frac{\partial L}{\partial x_s}(x,v) - v_k\Big(\frac{\partial^2 L}{\partial x_k \partial v_s}(x, v^\epsilon) + \frac{\partial^2 L}{\partial v_s \partial v_q}(x, v^\epsilon)\frac{\partial v^\epsilon_q}{\partial x_k}\Big)\right) d\mu =
$$

$$
-\int v_k \frac{\partial}{\partial x_k}\Big(\frac{\partial L}{\partial v_s}(x, v^\epsilon)\dot{X}^\epsilon_s\Big) d\mu + \int v_k \Big(\frac{\partial L}{\partial v_s}(x, v^\epsilon) - \frac{\partial L}{\partial v_s}(x,v)\Big)\frac{\partial}{\partial x_k}\big(\dot{X}^\epsilon_s\big) d\mu.
$$

$$
(57)
$$

Using the chain rule in the LHS and the Leibniz rule in the RHS we further reduce (57) to

$$\int \dot{X}_s^\epsilon \left(\frac{\partial L}{\partial x_s}(x,v) - v_k \frac{\partial}{\partial x_k} \left(\frac{\partial L}{\partial v_s}(x,v^\epsilon) \right) \right) d\mu =$$

$$-\int v_k \dot{X}_s^\epsilon \frac{\partial}{\partial x_k} \left(\frac{\partial L}{\partial v_s}(x,v^\epsilon) \right) d\mu - \int v_k \frac{\partial L}{\partial v_s}(x,v) \frac{\partial}{\partial x_k} \left(\dot{X}_s^\epsilon \right) d\mu.$$

Noting the cancellation of the term $\int v_k \dot{X}_s^\epsilon \frac{\partial}{\partial x_k} \left(\frac{\partial L}{\partial v_s}(x,v^\epsilon) \right) d\mu$, we see that the last identity is equivalent to (55) with $\xi_s(x) = \dot{X}_s^\epsilon(x)$. $\qquad\square$

Theorem 44. *Let μ be a minimizing holonomic measure. Then μ is invariant under the Euler-Lagrange flow.*

Proof. By Lemma 42 we have to prove that for any smooth compactly supported function $\phi(x,v)$

$$\int v_k \frac{\partial \phi}{\partial x_k} + \frac{\partial \phi}{\partial v_j} \left(\frac{\partial^2 L}{\partial^2 v} \right)_{js}^{-1} \left[\frac{\partial L}{\partial x_s} - v_k \frac{\partial^2 L}{\partial x_k \partial v_s} \right] d\mu = 0, \qquad (58)$$

where $(\ )_{js}^{-1}$ stands for the j,s entry of the inverse matrix.

The idea of the proof is first to rewrite (58) in an equivalent form and then apply an approximation argument. Since μ is supported by the graph $v = \boldsymbol{v}(x)$ we will change the x,v arguments with $x, \boldsymbol{v}(x)$ for the following four types of functions $\frac{\partial \phi}{\partial x_k}$, $\frac{\partial \phi}{\partial v_j}$, $\left(\frac{\partial^2 L}{\partial^2 v} \right)_{js}^{-1}$, and $\frac{\partial^2 L}{\partial x_k \partial v_s}$, occuring in (58):

$$\int v_k \frac{\partial \phi}{\partial x_k} (x, \boldsymbol{v}(x)) +$$

$$\frac{\partial \phi}{\partial v_j}(x, \boldsymbol{v}(x)) \left(\frac{\partial^2 L}{\partial^2 v} \right)_{js}^{-1}(x, \boldsymbol{v}(x)) \left(\frac{\partial L}{\partial x_s}(x,v) - v_k \frac{\partial^2 L}{\partial x_k \partial v_s}(x, \boldsymbol{v}(x)) \right) d\mu = 0.$$

$$(59)$$

To complete the proof of the theorem, we use Lemma 43. The first and second integrals in the RHS of (56) are zero due to the holonomy constraint. The third integral in the RHS of (56) tends to zero as $\epsilon \to 0$, because $|v^\epsilon(x) - \boldsymbol{v}(x)| < c\epsilon$ and therefore $|v^\epsilon(x) - v| < c\epsilon$ μ-a.e., and because $\dot{X}_s^\epsilon$ is uniformly Lipschitz and hence $\partial_{x_k} \dot{X}_s^\epsilon$ is uniformly bounded. Therefore the LHS of (56) tends to zero as $\epsilon \to 0$.

But the LHS of (56) also tends to the LHS of (59) as $\epsilon \to 0$. Indeed, since $\boldsymbol{v}(x)$ is a Lipschitz vector field we have

$$v^\epsilon(x) \to \boldsymbol{v}(x) \text{ (uniformly)} \qquad \text{and} \qquad \frac{\partial v^\epsilon(x)}{\partial x} \text{ is uniformly bounded.}$$

Moreover for any smooth function $\Psi(x, v)$ we have

$$\Psi\big(x, v^\epsilon(x)\big) \to \Psi\big(x, \boldsymbol{v}(x)\big) \text{ (uniformly)}$$

and

$$\frac{\partial}{\partial x}\Big(\Psi\big(x, v^\epsilon(x)\big)\Big) \text{ is uniformly bounded.}$$

Also note that for μ almost all (x, v) we have $v = \boldsymbol{v}(x)$. Therefore the Theorem is proven. $\qquad\square$

References

[BCD97]   M. Bardi and I. Capuzzo-Dolcetta. Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations. Birkhauser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.

[Con90]   John B. Conway. A course in functional analysis, volume 96 of Graduate Texts in Mathematics. Springer-Verlag, New York, second edition, 1990.

[Con95]   M. Concordel. Periodic homogenization of Hamilton-Jacobi equations. PhD thesis, UC Berkeley, 1995.

[EG92]   L. C. Evans and R. F. Gariepy. Measure theory and fine properties of functions. CRC Press, Boca Raton, FL, 1992.

[EG01]   L. C. Evans and D. Gomes. Effective Hamiltonians and averaging for Hamiltonian dynamics. I. Arch. Ration. Mech. Anal., 157(1):133, 2001.

[Eva90]   L. C. Evans. Weak convergence methods for nonlinear partial differential equations. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1990.

[Eva98]   L. C. Evans. Partial differential equations. American Mathematical Society, Providence, RI, 1998.

[FS93]   Wendell H. Fleming and H. Mete Soner. Controlled Markov processes and viscosity solutions. Springer- Verlag, New York, 1993.

[FS04]   A. Fathi and A. Siconolfi. Existence of C1 critical subsolutions of the Hamilton-Jacobi equation. Invent. Math., 155(2):363388, 2004.

[Gom02]   D. Gomes. A stochastic analogue of Aubry-Mather theory. Nonlinearity, 15(3):581603, 2002.

[KF75]   A. N. Kolmogorov and S. V. Fomn. Introductory real analysis. Dover Publications Inc., New York, 1975. Translated from the second Russian edition and edited by Richard A. Silverman, Corrected reprinting.

[LPV88]   P. L. Lions, G. Papanicolao, and S. R. S. Varadhan. Homogeneization of Hamilton-Jacobi equations. Preliminary Version, 1988.

[Mat]   J. N. Mather. Personal communication.

[Mat91]   J. N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. Math. Z., 207(2):169207, 1991.

[Mn96]   Ricardo Mane. Generic properties and problems of minimizing measures of Lagrangian systems. Nonlinearity, 9(2):273310, 1996.

[Vil03]   Cedric Villani. Topics in optimal transportation, volume 58 of Graduate Stud-
            ies in Mathematics. American Mathematical Society, Providence, RI, 2003.