NOTES OF MATH 807 - A COURSE ON WEAK KAM THEORY

SON N. T. TU

ABSTRACT. The notes on Math 807, a course on weak KAM theory, were based on the lecture notes given at UW Madison in Spring 2021 by Hung V. Tran. Basic references are: Fathi's book, Ishii's lecture notes on weak KAM theory, and Tran's book on Hamilton–Jacobi equations. This note is still under development and will be updated frequently. A complete references will be updated in the future.

CONTENTS

Date: August 4, 2021.
2010 *Mathematics Subject Classification.* 35B40, 35D40, 49J20, 49L25, 70H20 .
Key words and phrases. first-order Hamilton–Jacobi equations; state-constraint problems; optimal control theory; rate of convergence; viscosity solutions.

1.1. **Definition.** Let $H : \mathbb{R}^n \to \mathbb{R}$ be a given convex function, we want to study properties of H and its Legendre's transform deeply. Generally speaking, *convexity* is *one-sided linearity*. H is the supremum of all affince functions whose graphs stay below the graph of H. Basically, H is convex there is an index set $\mathcal{A}$ so that

$$H(p) = \sup\left\{v_\alpha \cdot p + a_\alpha : \alpha \in \mathcal{A}\right\}$$

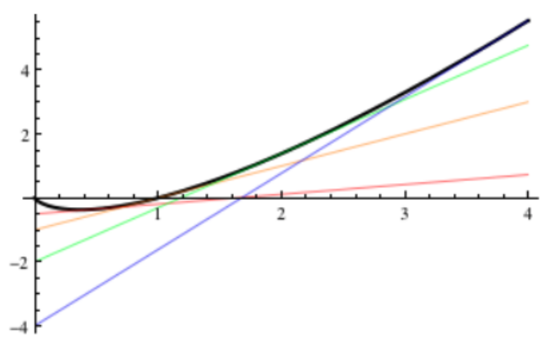
where $\{v_\alpha\}_{\alpha\in\mathcal{A}} \subset \mathbb{R}^n$,$\{a_\alpha\}_{\alpha\in\mathcal{A}} \subset \mathbb{R}$. Assume that H is convex and superlinear (H grows faster than linear speed), that is

$$\lim_{|p|\to\infty} \frac{H(p)}{|p|} = +\infty.$$

Definition 1 (Legendre's transform)**.** $L : \mathbb{R}^n \to \mathbb{R}$ *is*

$$L(v) = H^*(v) = \sup_{p\in\mathbb{R}^n}\Big(p\cdot v - H(p)\Big).$$

Example 1. *If* $H(p) = \frac{1}{2}|p|^2$ *for* $p \in \mathbb{R}^n$ *then* $L(v) = \frac{1}{2}|v|^2$ *for* $v \in \mathbb{R}^n$.



1.2. **Geometric meaning of Legendre's transform.** Consider all hyper-plane that touches the graph of H from below of the form $p\cdot v + c$. Basically when varying v, this give lines of slope v, and thus for $p \cdot v + c$ to touch H from below, we can see that at the touching point $H(p) = p \cdot v + c$, hence

$$L(v) = \sup_{p\in\mathbb{R}^n}\Big(p\cdot v - H(p)\Big) = -c.$$

Lemma 1.1. L *is finite, convex, and superlinear.*

If H is not superlinear then L is still defined, but could be infinite at some places. An example is $H(p) = |p|$ for $p \in \mathbb{R}^n$ that yields $L(v) = 0$ for $|v| \leqslant$ and $L(v) = +\infty$ otherwise.

Lemma 1.2. *If* H *is convex then* $L^* = H$, *i.e.*, $H^{**} = H$.

We refer the proofs of Lemmas 1.1 and 1.2 to [5]. The key ingredient is the existence of the so-called *subgradient*[1] $\partial^- f(x) \neq \emptyset$ at every point if f is a convex function. An important inequality arises from the proof of Lemma 1.2 is

$$H(p) + L(v) \geqslant p \cdot v \qquad \text{for all } p, v \in \mathbb{R}^n.$$

It is natural to ask when do we have equality, in short,

$$H(p) + L(v) = p \cdot v \qquad \Longleftrightarrow \qquad p \in \partial^- L(v) \qquad \Longleftrightarrow \qquad v \in \partial^- H(p).$$

In case H, L are C^1, they become $p = DL(v)$ iff $v = DH(p)$.

Question 1. *Show that if H is differentiable at p then one has $\partial H(p) = \{DH(p)\}$. Conversely, if H is convex and $\partial H(p) = \{\xi\}$ then H is differentiable at p with $\nabla H(p) = \xi$.*

Theorem 1.3. *Assume that H is convex and differentiable. Then, $H \in C^1$.*

Proof. Assume $p_k \to p_0$, we show $DH(p_k) \to \xi_0 = DH(p_0)$. Since $|p_k| \leqslant C$ for all k we have

$$H(p_k + h) \geqslant H(p_k) + DH(p_k) \cdot h \quad \text{for all } |h| \leqslant 1.$$

Thus $|DH(p_k)| \leqslant 2 \max_{|p| \leqslant C+1} |H(p)|$, hence up to subsequences $DH(p_k) \to \xi_0$ for some $\xi_0 \in \mathbb{R}^n$. By convexity

$$H(p) \geqslant H(p_k) + DH(p_k) \cdot (p - p_k)$$

and thus $H(p) \geqslant H(p_0) + \xi_0 \cdot (p - p_0)$, hence $\xi_0 \in D^- H(p_0)$ and thus $\xi_0 \equiv DH(p_0)$ since H is differetiable at p_0. The uniqueness of $\xi_0 = DH(p_0)$ enables us to have convergence of the whole sequence $DH(p_k) \to DH(p_0)$. $\square$

The proof also implies that:

Lemma 1.4. *If H is convex then there hold*

(i) *(Boundedness of subgradient)* $\partial H(B(0,R)) \subset B(0, C_R)$.
(ii) *(Stability) If $p_k \to p$ and $v_k \in \partial H(p_k)$ such that $v_k \to v$ then $v \in \partial H(p)$.*

About the existence of subgradient, there are some methods commonly used.

(1) Do convolution $H^\varepsilon = H * \eta_\varepsilon$, then H^ε is convex and smooth, thus at p one has $DH^\varepsilon(p) = v_\varepsilon$, and we get a subsequential limit $v_{\varepsilon_j} \to v \in \partial H(p)$.
(2) Proof by contradiction. H convex implies H is locally Lipschitz, thus by Radamacher's theorem H is differentiable a.e.. Assume H is differentiable at p_k and $p_k \to p$, then (by compactness) if $DH(p_k) \to v$ and $v \notin \partial H(p)$ we can derive a contradiction.
(3) Hahn-Banach theorem (supporting hyper-plane in finite dimensional spaces).

Question 2. *Show that if H is convex then $\partial H = D^- H$ is nonempty.*

Question 3. *If H is not convex, then what does the information H^{**} recover?*

[1] For convex functions, $\partial^- f$ and ∂f both mean the same thing as subgradients, even though the former one can be defined for general nonconvex functions.

1.3. **Strictly convex Hamiltonians.**

Theorem 1.5. *Assume that* H *is convex and super-linear, then the following are equivalent:*

(i) H *is strictly convex, that is for* $s \in (0,1)$ *then*

$$H(sp_1 + (1-s)p_2) < sH(p_1) + (1-s)H(p_2). \tag{1.1}$$

(ii) $\partial H(p_1) \cap \partial H(p_2) = \emptyset$ *if* $p_1 \neq p_2$.

(iii) $L = H^* \in C^1$.

Proof. For (i) implies (ii), if $\xi \in \partial H(p_1) \cap \partial H(p_2)$ for some $p_1 \neq p_2$ then by definition of subgradient we have

$$\begin{cases} H(sp_1 + (1-s)p_2) \geqslant H(p_1) + \xi \cdot (p_2 - p_1)(1-s), \\ H(sp_1 + (1-s)p_2) \geqslant H(p_2) + \xi \cdot (p_1 - p_2)s, \end{cases}$$

for $s \in (0,1)$. Multiplying the first equation with s, the second equation with $(1-s)$ and adding them we obtain a contradiction with (i).

For (ii) implies (iii), from Question 1 it suffices to show that $\partial L(v)$ is a singleton at any $v \in \mathbb{R}^n$. It is obvious since $p_1, p_2 \in \partial L(v)$ implies $v \in \partial H(p_1) \cap \partial H(p_2) = \emptyset$ is a contradiction.

For (iii) implies (i), if H is not strictly convex, i.e.,

$$H(s_0 p_1 + (1-s_0)p_2) = s_0 H(p_1) + (1-s_0)H(p_2).$$

for $s_0 \in (0,1)$ and $p_1 \neq p_2$, then for all $s \in (0,1)$ there holds

$$H(sp_1 + (1-s)p_2) = sH(p_1) + (1-s)H(p_2).$$

Take $v \in \partial H(p_s)$ where $p_s = sp_1 + (1-s)p_2$, then for $p \in \mathbb{R}^n$ we have $H(p) - H(p_s) \geqslant v \cdot (p - p_s)$, thus

$$\begin{aligned} H(p) - H(p_1) &\geqslant H(p_s) - H(p_1) + v \cdot (p - p_s) \\ &= (1-s)\big(H(p_2) - H(p_1)\big) + v \cdot (p - p_s) \\ &\geqslant (1-s)v \cdot (p_2 - p_1) + v \cdot (p - p_s) \\ &= v \cdot (p - p_1). \end{aligned}$$

Thus $v \in \partial H(p_1)$ and similarly $v \in \partial H(p_2)$ as well, which is a contradiction as it implies $p_1, p_2 \in \partial L(v) = \{\nabla L(v)\}$. □

Theorem 1.6. *Assume* $H \in C^k(\mathbb{R}^n)$ *with* $k \geqslant 2$, H *is convex, super-linear and is locally uniformly convex, i.e.,* $D^2 H(p) \succ 0$ *(positive definite) every where, then*

- $L \in C^k(\mathbb{R}^n)$.
- $DH : \mathbb{R}^n \to \mathbb{R}^n$ *is a* C^{k-1} *diffeormophism.*
- $DL(v) = (DH)^{-1}(v)$, $D^2 L(v) = \left[D^2 H(DL(v))\right]^{-1}$.
- $L(v) = v \cdot DL(v) - H(DL(v))$.

Proof. As H is locally uniformly convex, it is strictly convex and thus $L \in C^1$ and we have for all $v \in \mathbb{R}^n$ then

$$p = DL(v) \qquad \Longrightarrow \qquad v = DH(p).$$

Thus $(DL)^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is well-defined and $(DL)^{-1} = DH$, which is of class C^{k-1}. Since $D^2 H \succ 0$ everywhere in $\mathbb{R}^n$, the inverse function theorem says that $DH : \mathbb{R}^n \to \mathbb{R}^n$ is a local C^{k-1} diffeomorphism $\mathbb{R}^n \to \mathbb{R}^n$, thus DL is also a local C^{k-1} diffeomorphism and thus L is of class C^k. By definition $DH(DL(v)) = v$ for all $v \in \mathbb{R}^n$, thus

$$D^2H(DL(v)) \cdot D^2L(v) = \mathbb{I}_n$$

for all $v \in \mathbb{R}^n$ and therefore $D^2L(v) = \left[D^2H(DL(v))\right]^{-1}$. □

1.4. **Hamiltonians that depend on positions.** We consider Hamiltonians that depend also on position, generally $H = H(x,p)$ for $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$. Assume that

(H) $H(x,p) \in C(\mathbb{R}^n \times \mathbb{R}^n), p \mapsto H(x,p)$ is convex and

$$\lim_{|p|\to\infty} \left(\inf_{x\in\mathbb{R}^n} \frac{H(x,p)}{|p|} \right) = +\infty.$$

We define as usual the Lagrangian

$$L(x,v) = H^*(x,v) = \sup_{p\in\mathbb{R}^n} \big(p \cdot v - H(x,p)\big), \qquad (x,v) \in \mathbb{R}^n \times \mathbb{R}^n. \tag{1.2}$$

Theorem 1.7. *Assume* (H)*, then* L *is finite, convex and superlinear in* v *and* $L^* = H^{**} = H$*. The supremum in* (1.2) *is achieved and if* $|v| \leqslant R$ *then there exists* C_R *such that*

$$L(x,v) = \sup_{|p|\leqslant C_R} \big(p \cdot v - H(x,p)\big).$$

Also, $L \in C(\mathbb{R}^n \times \mathbb{R}^n)$ *and if* H *is strictly convex in* p *then* $D_vL(x,v)$ *exists and* $(x,v) \mapsto D_vL(x,v)$ *is continuous.*

Proof. The only new thing to prove here is the continuity of L. Assume $(x_k,v_k) \to (x_0,v_0)$ in $\mathbb{R}^n \times \mathbb{R}^n$, as $|v_k| \leqslant C$, we can find $p_k \in \mathbb{R}^n$ with $|p_k| \leqslant C$ such that

$$L(x_k,v_k) = p_k \cdot v_k - H(x_k,p_k), \qquad k \in \mathbb{N}.$$

Denote $\omega(k) = |H(x_0,p_k) - H(x_k,p_k)| + C|v_k - v_0| \to 0$ as $k \to \infty$ (here we use the fact that p_k is bounded and thus a local uniform modulus of continuity exists) then

$$L(x_k,v_k) = p_k \cdot v_k - H(x_k,p_k) \leqslant p_k \cdot v_0 - H(x_0,p_k) + \omega(k) \leqslant L(x_0,v_0) + \omega(k).$$

Therefore

$$\limsup_{k\to\infty} L(x_k,v_k) \leqslant L(x_0,v_0).$$

Take any $p \in \mathbb{R}^n$ then by definition

$$L(x_k,v_k) \geqslant p \cdot v_k - H(x_k,p), \qquad k \in \mathbb{N},$$

which gives us that, for all $p \in \mathbb{R}^n$ then

$$\liminf_{k\to\infty} L(x_k,v_k) \geqslant p \cdot v_0 - H(x_0,p) \qquad \Longrightarrow \qquad \liminf_{k\to\infty} L(x_k,v_k) \geqslant L(x_0,v_0).$$

Thus the proof is complete. □

Theorem 1.8. *Assume* (H), *then if* $H \in C^k(\mathbb{R}^n \times \mathbb{R}^n)$ *for* $k \geqslant 2$ *and* H *is locally uniformly convex in* p, *i.e.,* $D^2H(x,p) \succ 0$ *for all* $(x,p) \in \mathbb{R}^n \times \mathbb{R}^n$. *Then* $L \in C^k(\mathbb{R}^n \times \mathbb{R}^n)$ *and there exists a unique* $p(x,v)$ *such that*

$$\begin{aligned} p(x,v) &= D_vL(x,v) \\ D_xL(x,v) &= -D_xH(x,p(x,v)) \\ D^2_{vv}L(x,v) &= \left[D^2_{pp}H(x,p(x,v))\right]^{-1}. \end{aligned}$$

Also, $p(x,v) = D_vL(x,v)$ *implies* $v = D_pH(x,p(x,v))$.

Proof. We already known that for each fixed $x \in \mathbb{R}^n$ then $v \mapsto L(x,v)$ is of class $C^k(\mathbb{R}^n)$. By assumption, $D_pH : (x,p) \mapsto D_pH(x,p)$ is a $C^{k-1}(\mathbb{R}^n \times \mathbb{R}^n)$ diffeomorphism (by inverse function theorem with locally uniform convexity). The relations

$$p = D_vL(x,v) \qquad \Longleftrightarrow \qquad v = D_pH(x,p)$$

defines a map $\mathcal{L} : (x,v) \mapsto (x,p) = (x, D_vL(x,v))$ with its inverse $\mathcal{H} : (x,p) \mapsto (x, D_pH(x,p))$. Now since $\mathcal{H}$ is a $C^{k-1}(\mathbb{R}^n \times \mathbb{R}^n)$ diffeomorphism, $\mathcal{L}$ is also is a $C^{k-1}(\mathbb{R}^n \times \mathbb{R}^n)$ diffeomorphism, i.e.,

$$(x,v) \mapsto D_vL(x,v) \in C^{k-1}(\mathbb{R}^n \times \mathbb{R}^n).$$

We need to show that $(x,v) \mapsto D_xL(x,v)$ is $C^{k-1}(\mathbb{R}^n \times \mathbb{R}^n)$. Let us define[2]

$$p(x,v) = D_vL(x,v) \in C^{k-1}(\mathbb{R}^n \times \mathbb{R}^n).$$

From the identity

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

we deduce that $x \mapsto L(x,v)$ is differentiable in x for each v and thus, by differentiating with respect to x we have

$$D_xL(x,v) = -D_xH(x,p(x,v)) + \Big(v \cdot D_xp(x,v) - D_pH(x,p(x,v)) \cdot D_xp(x,v)\Big) = -D_xH(x,p(x,v))$$

since $v = D_pH(x,p(x,v))$. ☐

Definition 2. *Define*

$$\begin{aligned} \mathcal{H} : \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R}^n \times \mathbb{R}^n \\ (x,p) &\mapsto (x,v) = (x, D_pH(x,p)) \end{aligned}$$

and its dual

$$\begin{aligned} \mathcal{L} : \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R}^n \times \mathbb{R}^n \\ (x,v) &\mapsto (x,p) = (x, D_vL(x,v)). \end{aligned}$$

Under the assumption of Theorem 1.8, $\mathcal{H}, \mathcal{L}$ *are both local* C^{k-1} *diffeomorphisms.*

Remark 1. Sometimes we assume more that H is bounded in $\mathbb{R}^n \times \overline{B}(0,R)$ for each $R > 0$ (so is L) to get the boundedness of $|p(x,v)|$ given $|v| \leqslant C$. If $p(x,v) \in \partial_vL(x,v)$ then $L(x,v+h) \geqslant L(x,v) + p(x,v) \cdot h$, thus

$$|p(x,v)| = \max_{|h| \leqslant 1} p(x,v) \cdot h \leqslant |L(x,v)| + |L(x,v+h)|.$$

[2]This is significant as it says $x \mapsto p(x,v)$ is continuously differentiable.

Outline.

- We state the minimization for action functional problem and show that generally minimizers satisfy Euler-Lagrange equation with less and less regularity assumption.

$$\boxed{C^2 \longrightarrow C^1 \longrightarrow \text{piece-wise } C^1 \longrightarrow AC.}$$

- Existence and regularity of minimizers. We show that there exists an absolutely continuous minimizer and then show that the minimizer is indeed smooth (provided that the Lagrangian is smooth).

The main tool is the mechanism that allows us to go back and forth between Hamiltonian viewpoint and Lagrangian viewpoint.

2.1. **Action functional.** For a continuous, piece-wise C^1 curve $\gamma : [a,b] \to \mathbb{R}^n$, the action functional of γ for L is defined by

$$I[\gamma] = \int_a^b L(\gamma(s), \dot{\gamma}(s))ds.$$

We note that $\gamma \in AC([a,b];\mathbb{R}^n)$, the space of absolutely continuous curves, is enough to define the action functional here ($\dot{\gamma} \in L^1([a,b])$ is enough).

2.2. **Minimizers and the Euler–Lagrange equation.** There are many different notions of minimizers, we start with the following notion.

Definition 3. *Fix $y, z \in \mathbb{R}^n$. Denote the admissible set as*

$$\mathcal{A} = \Big\{\gamma \in C\left([a,b];\mathbb{R}^n\right) \text{ piece-wise } C^1, \gamma(a) = y, \gamma(b) = z\Big\}.$$

We say that $\gamma \in \mathcal{A}$ is a minimizer of class $\mathcal{A}$ if $I[\gamma] = \inf_{\eta \in \mathcal{A}} I[\eta]$.

We assume through out the chapter that

(L) $L \in C^k(\mathbb{R}^n \times \mathbb{R}^n)$ for $k \geqslant 2$, L is super-linear in v uniformly in x and $D^2_{vv}L(x,v) \succ 0$ for all $(x,v) \in \mathbb{R}^n \times \mathbb{R}^n$.

Theorem 2.1. *If $\gamma \in C^2([a,b];\mathbb{R}^n)$ is minimizer then γ satisfies the Euler-Lagrange equation*

$$\frac{d}{dt}\left(D_v L(\gamma(s), \dot{\gamma}(s))\right) = D_x L(\gamma(s), \dot{\gamma}(s)), \qquad a \leqslant s \leqslant b.$$

Proof. Let $\eta \in C^\infty([a,b];\mathbb{R}^n)$ with $\eta(a) = \eta(b) = 0$. For each $\tau \in \mathbb{R}$, $\gamma + \tau\eta \in \mathcal{A}$. Let

$$i(\tau) = I[\gamma + \tau\eta]$$

then $i : \mathbb{R}^n \to \mathbb{R}^n$ and thus $i'(0) = 0$. We recall that

$$i(\tau) = \int_a^b L\Big(\gamma(s) + \tau\eta(s), \dot{\gamma}(s) + \tau\dot{\eta}(s)\Big)\, ds.$$

We deduce that

$$i'(\tau) = \int_a^b \Big(D_x L\big(\gamma(s)+\tau\eta(s), \dot\gamma(s)+\tau\dot\eta(s)\big)\cdot\eta(s) + D_v L\big(\gamma(s)+\tau\eta(s), \dot\gamma(s)+\tau\dot\eta(s)\big)\cdot\dot\eta(s)\Big)\, ds.$$

Setting $\tau = 0$ we have

$$i'(0) = \int_a^b \Big(D_x L\big(\gamma(s), \dot\gamma(s)\big)\cdot\eta(s) + D_v L\big(\gamma(s), \dot\gamma(s)\big)\cdot\dot\eta(s)\Big)\, ds.$$

As $\dot\gamma(s)$ is C^1, by integration by parts we have

$$i'(0) = \int_a^b \left(D_x L(\gamma(s), \dot\gamma(s)) - \frac{d}{ds} D_v L(\gamma(s), \dot\gamma(s))\right)\eta(s)\, ds = 0$$

for all $\eta \in C^\infty([a,b];\mathbb{R}^n)$ with $\eta(a) = \eta(b) = 0$. This gives us the conclusion. □

Remark 2. If we only assume that $\gamma \in C^1([a,b];\mathbb{R}^n)$ then we cannot yet write down the Euler-Lagrange equation. What we have is

$$\int_a^b \left(D_v L(\gamma(t), \dot\gamma(t)) - \int_a^t D_x L(\gamma(s), \dot\gamma(s)) ds\right)\cdot\dot\eta(t) dt = 0$$

for all $\eta \in C^\infty([a,b];\mathbb{R}^n)$ with $\eta(a) = \eta(b) = 0$. In order words we only have

$$D_v L(\gamma(t), \dot\gamma(t)) = \xi + \int_a^t D_x L(\gamma(s), \dot\gamma(s)) ds \tag{2.1}$$

for some $\xi \in \mathbb{R}^n$. This says that $t \mapsto D_v L(\gamma(t), \dot\gamma(t))$ is C^1 provided that $s \mapsto (\gamma(s), \dot\gamma(s))$ is C^0 only.

Theorem 2.2. *If $\gamma \in C^1([a,b];\mathbb{R}^n)$ is minimizer then $\gamma \in C^2([a,b];\mathbb{R}^n)$ and in fact $\gamma \in C^k([a,b];\mathbb{R}^n)$.*

Proof. Fix $t_0 \in [a,b]$ and let $(x_0, v_0) = \big(\gamma(t_0), \dot\gamma(t_0)\big)$. Recall that $(x,v) \mapsto (x, D_v L(x,v))$ is a local C^{k-1} diffeomorphism. Let $\mathcal{H}$ be the local inverse that maps $(x,p) \to (x, D_p H(x,p))$ which is also a C^{k-1} diffeomorphism, we have

$$\begin{cases} \mathcal{H}\left(x_0, D_v L(x_0, v_0)\right) = (x_0, v_0), \\ \mathcal{H}\left(\gamma(t), D_v L(\gamma(t), \dot\gamma(t))\right) = (\gamma(t), \dot\gamma(t)), \end{cases}$$

for $t \approx t_0$. As a minimizer, we still have γ satisfies (2.1), thus $D_v L\left(\gamma(\cdot), \dot\gamma(\cdot)\right) \in C^1$. Together with $\mathcal{H}$ is C^{k-1} and

$$\mathcal{H}\left(\gamma(t), D_v L(\gamma(t), \dot\gamma(t))\right) = (\gamma(t), \dot\gamma(t))$$

we deduce that $t \mapsto (\gamma(t), \dot\gamma(t))$ is C^1, therefore $\gamma \in C^2$ and by induction $\gamma \in C^k$. □

Remark 3 (Hamiltonian and Lagrangian viewpoints)**.** If one thinks of $x(t) = \gamma(t)$ as the position of a particle and $v(t) = \dot\gamma(t)$ as the velocity then the momentum is defined as

$$p(t) = D_v L(\gamma(t), \dot\gamma(t)) \quad \Longrightarrow \quad v(t) = D_p H(x(t), p(t)).$$

We call $(\gamma(t), \dot\gamma(t))$ the Lagrangian coordinates, and the associated Hamiltonian system is defined by $(x(t), p(t))$ with

$$\begin{cases} \dot x(t) = \dot\gamma(t) = D_p H(x(t), p(t)), \\ \dot p(t) = D_x L(\gamma(t), \dot\gamma(t)) = -D_x H(x(t), p(t)). \end{cases}$$

We remark that if L is not *nice enough* then minimizers can have bad regularity or minimizers may not satisfy the Euler-Lagrange equation. To further relax the regularity of minimizers, we will need the following lemma.

Lemma 2.3. *If $\gamma : [a, b] \to \mathbb{R}^n$ is a continuous, piece-wise C^1 minimizer then $\gamma|_{[c,d]}$ is also a minimizer to*

$$\inf_{\eta \in \mathcal{A}'} \int_c^d L(\eta(s), \dot\eta(s))\, ds$$

where $\mathcal{A}'$ is the set of all η continuous and piece-wise C^1 from $[c, d]$ to $\mathbb{R}^n$ with $\eta(c) = \gamma(c)$ and $\eta(d) = \gamma(d)$.

Proof. For each $\eta \in \mathcal{A}'$, we define $\tilde\eta = \eta$ on $[c, d]$ and $\tilde\eta = \gamma$ on $[a, b] \backslash (c, d)$. It is clear that $\tilde\eta \in \mathcal{A}$, thus $I[\tilde\eta] \geqslant I[\gamma]$ and therefore

$$\inf_{\eta \in \mathcal{A}'} \int_c^d L(\eta(s), \dot\eta(s))\, ds \geqslant \int_c^d L(\gamma(s), \dot\gamma(s))\, ds$$

and thus γ is a minimizer on $[c, d]$. □

Theorem 2.4. *If $\gamma \in \mathcal{A}$, i.e., γ is a continuous, piece-wise C^1 minimizer then $\gamma \in C^k$.*

Proof. Let $a = a_0 < \ldots < a_m = b$ such that $\gamma \in C^1([a_i, a_{i+1}])$ for $i = 0, \ldots, m-1$. By the previous Lemma $\gamma|_{[a_i, a_{i+1}]}$ is a minimizer on the subinterval, therefore by Theorem 2.2 $\gamma \in C^k([a_i, a_{i+1}])$ for $i = 1, 2, \ldots, m-1$. We only have to show that γ is C^1 at a_i for $i = 1, \ldots, m-1$, then again Theorem 2.2 concludes. By calculus of variation we have

$$\int_a^b \Big(D_x L(\gamma, \dot\gamma) \cdot \eta + D_v L(\gamma, \dot\gamma) \cdot \dot\eta \Big)\, ds = \sum_{i=0}^{m-1} \int_{a_i}^{a_{i+1}} \Big(D_x L(\gamma, \dot\gamma) \cdot \eta + D_v L(\gamma, \dot\gamma) \cdot \dot\eta \Big)\, ds = 0$$

for all η smooth with $\eta(a) = \eta(b) = 0$. Using integration by parts we have

$$\sum_{i=0}^{m-1} \int_{a_i}^{a_{i+1}} \Big(D_x L(\gamma, \dot\gamma) - \frac{d}{ds} D_v L(\gamma, \dot\gamma) \Big) \cdot \eta\, ds \\ + \sum_{i=0}^{m} \Big[D_v L\big(\gamma(a_i^-), \dot\gamma(a_i^-)\big) - D_v L\big(\gamma(a_i^+), \dot\gamma(a_i^+)\big) \Big] \eta(a_i) = 0.$$

From the Euler-Lagrange equation on $[a_i, a_{i+1}]$ the integral term is zero, thus

$$\sum_{i=0}^{m} \Big[D_v L\big(\gamma(a_i^-), \dot\gamma(a_i^-)\big) - D_v L\big(\gamma(a_i^+), \dot\gamma(a_i^+)\big) \Big] \eta(a_i) = 0.$$

Since η can be chosen arbitrarily we conclude that

$$D_v L\big(\gamma(a_i^-), \dot\gamma(a_i^-)\big) = D_v L\big(\gamma(a_i^+), \dot\gamma(a_i^+)\big)$$

for $i = 0, 1, \ldots, m$. Now using the identity $p \cdot v = H(x,p) + L(x,v)$ iff $p = D_v L(x,v)$ iff $v = D_p H(x,p)$ we obtain the conclusion $\dot\gamma(a_i^-) = \dot\gamma(a_i^+)$ and hence γ is C^1 at a_i. □

Definition 4 (Extremal curves). *A continuous and piece-wise C^1 curve $\gamma : [a,b] \to \mathbb{R}^n$ is called extremal if it is a critical point of the action functional*

$$\frac{d}{ds} I[\gamma + s\eta]\Big|_{s=0} = 0 \quad \text{for all } \eta \in C_c^\infty([a,b];\mathbb{R}^n).$$

It is clear that any continuous, piece-wise C^1 extremal curve is C^k and satisfies the Euler-Lagrange equation.

2.3. **Absolutely continuous minimizers.** The space of continuous and piece-wise smooth curves with fixed endpoints is not compact (under a reasonable topology), therefore it is convenient to relax to a better space in which we have compactness (to construct minimizers).

Definition 5 (Absolutely continuous). $\gamma : [a,b] \to \mathbb{R}^n$ *is* absolutely continuous *if for each $\varepsilon > 0$ there exists $\delta > 0$ such that, if $\{(a_i, b_i)\}_{i\in\mathbb{N}}$ is a disjoint family of intervals in (a,b) then*

$$\sum_{i\in\mathbb{N}} |b_i - a_i| < \delta \quad \Longrightarrow \quad \sum_{i\in\mathbb{N}} |\gamma(b_i) - \gamma(a_i)| < \varepsilon.$$

Theorem 2.5 (Characterization of absolutely continuous curves). *γ is absolutely continuous if and only if all of the following hold*

(i) *$\dot\gamma$ exists a.e. in (a,b).*
(ii) *$\dot\gamma$ is Lebesgue integrable on (a,b).*
(iii) *$\gamma(t) - \gamma(a) = \int_a^t \dot\gamma(s)\,ds$ for each $t \in [a,b]$.*

Note that from our super-linearity assumption, there are two scenarios:

- $L(x,v) \geqslant C|v|^p - C$ for some $p > 1$, we can get some compactness in $L^p([a,b])$ of $\dot\gamma$ and the existence is simple (see Appendix of [13] for example).
- One cannot get any L^p bound with $p > 1$ for a minimizing sequence. The best we can do is $\dot\gamma \in L^1([a,b])$, which makes it harder as compactness in $L^1([a,b])$ requires some additional *tightness* condition. This theorem says that generally one may replaces absolutely continuous curves by curves γ with $\dot\gamma \in L^1$.

2.4. **Compactness of absolutely continuous curves.** The space $AC([a,b];\mathbb{R}^n)$ of absolutely continuous curves enjoy the following compactness (*tightness*) property.

Theorem 2.6. *Let $\{\gamma_k\}_{k\in\mathbb{N}} \subset AC([a,b];\mathbb{R}^n)$. Suppose that $\{\dot\gamma_k\}_{k\in\mathbb{N}}$ is uniformly integrable on $[a,b]$, that is for each $\varepsilon > 0$, there is $\delta > 0$ such that if $E \subset [a,b]$ is a Borel measurable set with measure $|E| < \delta$ then*

$$\sup_{k\in\mathbb{N}} \int_E |\dot\gamma(s)|ds < \varepsilon. \tag{2.2}$$

If there exists $t_0 \in [a,b]$ such that $\{\gamma_k(t_0)\}$ is bounded, then there exists a subsequence γ_{k_j} and $\gamma \in AC([a,b];\mathbb{R}^n)$ such that $\gamma_{k_j} \to \gamma$ uniformly on $[a,b]$ and $\dot\gamma_{k_j} \rightharpoonup \dot\gamma$ weakly in $L^1([a,b])$, that *is*

$$\lim_{k_j\to\infty}\int_a^b \dot\gamma_{k_j}(s)\cdot\phi(s)\ ds = \int_a^b \dot\gamma(s)\cdot\phi(s)\ ds$$

for all $\phi \in L^\infty([a,b]\ \mathbb{R}^n)$.

This theorem allows us to utilize the nice property of the space $AC([a,b];\mathbb{R}^n)$ (togeher with the lower semi-continuity of the action functional).

Proof of Theorem 2.6. We split the proof to several steps for clarity.

(1) $\{\gamma_k\}$ is equi-continuous. Since $\{\gamma_k(t_0)\}$ is bounded, for $|t_1-t_2|<\delta$ we have

$$|\gamma_k(t_1)-\gamma_k(t_2)| \leqslant \left|\int_{t_1}^{t_2}\dot\gamma_k(s)ds\right| < \varepsilon$$

for all $k\in\mathbb{N}$. By Arzela-Ascoli theorem, there exists $\gamma_{k_j}\to\gamma$ uniformly on $[a,b]$. By abusing of notation, we will write $\gamma_{k_j}\to\gamma$ uniformly on $[a,b]$.

(2) γ is absolutely continuous. Fix ε and pick $\delta>0$ as in the assumption of the Theorem. Let $\{(a_i,b_i)\}_{i\in\mathbb{N}}$ be a sequence of disjoint open intervals with $\sum_{i\in N}(b_i-a_i)<\delta$, then the tightness condition gives us that, for all $k\in\mathbb{N}$ then

$$\sum_{i\in\mathbb{N}}|\gamma_k(b_i)-\gamma_k(a_i)| \leqslant \sum_{i\in\mathbb{N}}\int_{a_i}^{b_i}|\dot\gamma_k(s)|ds<\varepsilon$$

since $\sum_{i\in\mathbb{N}}(b_i-a_i)<\delta$. Let $k\to\infty$ we deduce that $\gamma\in AC([a,b];\mathbb{R}^n)$.

(3) We show $\dot\gamma_k\to\dot\gamma$ weakly in L^1. To show that

$$\lim_{k\to\infty}\int_a^b\dot\gamma_k(s)\phi(s)ds = \int_a^b\dot\gamma(s)\phi(s)ds \tag{2.3}$$

for $\phi\in L^\infty([a,b];\mathbb{R}^n)$ we use approximation $\phi\in L^\infty$ by simple functions from $[a,b]$ to $\mathbb{R}^n$. First of all, any open set U in (a,b) can be written as a disjoint union of countably many open sub-intervals $\{(a_i,b_i)\}_{i\in\mathbb{N}}$. For $\varepsilon>0$ take $\delta>0$ as in (2.2) and choose m such that $E=U\backslash\bigcup_{i=1}^m(a_i,b_i)$ has $|E|<\delta$, we have

$$\sup_{k\in\mathbb{N}}\left|\int_U\dot\gamma_k(s)ds-\sum_{i=1}^m\int_{a_i}^{b_i}\dot\gamma_k(s)ds\right|<\varepsilon. \tag{2.4}$$

Now we have

$$\sum_{i=1}^m\int_{a_i}^{b_i}\dot\gamma_k(s)ds=\sum_{i=1}^m\big(\gamma_k(b_i)-\gamma_k(a_i)\big)\longrightarrow\sum_{i=1}^m\big(\gamma(b_i)-\gamma(a_i)\big)=\sum_{i=1}^m\int_{a_i}^{b_i}\dot\gamma(s)ds.$$

Taking the limit as $k\to\infty$ in (2.4) we obtain

$$\sum_{i=1}^m\int_{a_i}^{b_i}\dot\gamma(s)ds-\varepsilon\leqslant\liminf_{k\to\infty}\int_U\dot\gamma_k(s)ds\leqslant\limsup_{k\to\infty}\int_U\dot\gamma_k(s)ds\leqslant\sum_{i=1}^m\int_{a_i}^{b_i}\dot\gamma(s)ds+\varepsilon.$$

Taking $m\to\infty$, and since ε is arbitrary we deduce

$$\lim_{k\to\infty}\int_U\dot\gamma_k(s)ds=\int_U\dot\gamma(s)ds. \tag{2.5}$$

By approximation, (2.5) holds for all measurable set $A \subset [a,b]$, and again by approximation (2.3) follows. □

2.5. **Existence of *absolutely continuous* minimizers.** Now we define the new admissible for fixed $y, z \in \mathbb{R}^n$ set as

$$\mathcal{A}_{ac} = \{\gamma \in AC([a,b];\mathbb{R}^n) : \gamma(a) = y, \gamma(b) = z\}.$$

The general framework for the existence by *direct method* goes like this.

(1) $I[\gamma] \geqslant -C$ for all $\gamma \in \mathcal{A}_{ad}$, usually by the super-linearity of L that $L(x,v) \geqslant \theta|v| - C_\theta$.
(2) Taking a minimizing sequence $\{\gamma_k\} \subset \mathcal{A}_{ad}$, the compactness result gives (via subsequence) $\gamma_k \to \gamma$ uniformly on $[a,b]$ and $\dot\gamma_k \rightharpoonup \dot\gamma$ weakly in $L^1([a,b];\mathbb{R}^n)$.
(3) Show that $I[\gamma] \leqslant \liminf_{k\to\infty} I[\gamma_k]$, this is a key point.

Theorem 2.7. *Assume $L \in C^1(\mathbb{R}^n \times \mathbb{R}^n)$ such that $L(x,v)$, $D_vL(x,v)$ belong to $\mathrm{BUC}(\mathbb{R}^n \times \overline{B(0,R)})$ for each $R > 0$, $v \mapsto L(x,v)$ is convex and is super linear uniformly in x, i.e.,*

$$\lim_{|v|\to\infty}\left(\inf_{x\in\mathbb{R}^n} \frac{L(x,v)}{|v|}\right) = +\infty. \tag{2.6}$$

Then there exists a minimizer $\gamma \in \mathcal{A}_{ac}$ of the action functional.

Proof. Assume $\inf_{\gamma\in\mathcal{A}_{ac}} I[\gamma]$ is finite, we can take minimizing sequence $\{\gamma_k\}_{k\in\mathbb{N}}$. From the super-linearity of $L(x,v)$, for each $\theta > 0$ there exist $C_\theta > 0$ such that

$$L(x,v) \geqslant \theta|v| - C_\theta \qquad \text{for all } (x,v) \in \mathbb{R}^n \times \mathbb{R}^n.$$

As $I[\gamma_k] \leqslant C$, we deduce that (the tightness condition follows)

$$\int_a^b |\dot\gamma_k(s)|ds \leqslant \frac{C + C_\theta}{\theta} \qquad \text{for all } k \in \mathbb{N}.$$

Since $\gamma_k(a) = y$ is fixed, $\{\gamma_k\}$ satisfies the tightness (uniform integrability) condition due to super-linearity, thus there exists $\gamma \in \mathcal{A}_{ac}$ such that $\gamma_k \to \gamma$ uniformly and $\dot\gamma_k \rightharpoonup \dot\gamma$ weakly in $L^1([a,b];\mathbb{R}^n)$. We have left to show that

$$I[\gamma] \leqslant \liminf_{k\to\infty} I[\gamma_k].$$

To use the uniformly bound on modulus of continuity of L and D_vL we need a L^∞ bound on $|\dot\gamma_k|$, this is unfortunately cannot be obtained on the whole interval $[a,b]$. Nevertheless, for $m \in \mathbb{N}$ we have

$$\sup_{k\in\mathbb{N}} \left|\left\{s \in [a,b] : |\dot\gamma_k(s)| \geqslant m\right\}\right| \leqslant \frac{C}{m}. \tag{2.7}$$

As $\dot\gamma \in L^1([a,b];\mathbb{R}^n)$, we also have

$$\left|\left\{s \in [a,b] : |\dot\gamma(s)| \geqslant m\right\}\right| \leqslant \frac{\|\dot\gamma\|_{L^\infty}}{m}.$$

Let

$$E_m = \left\{s \in [a,b] : |\dot\gamma(s)| \leqslant m, |\dot\gamma_k(s)| \leqslant m \text{ for all } k \in \mathbb{N}\right\}$$

then we can choose m such that $|[a,b]\backslash E_m| < \delta$ for any given $\delta > 0$. On $\mathbb{R}^n \times \overline{B(0,m)}$ there exists a (uniform) modulus of continuity $\omega(x,v)$ of L, D_vL, we have for a.e. $s \in E_m$ that

$$\begin{aligned} L(\gamma_k(s),\dot{\gamma}_k(s)) &\geqslant L(\gamma(s),\dot{\gamma}_k(s)) - \omega\left(\|\gamma_k - \gamma\|_{L^\infty}\right) \\ &\geqslant L\big(\gamma(s),\dot{\gamma}(s)\big) - \omega\left(\|\gamma_k - \gamma\|_{L^\infty}\right) + D_vL\big(\gamma(s),\dot{\gamma}(s)\big)\cdot\big(\dot{\gamma}_k(s) - \dot{\gamma}(s)\big). \end{aligned}$$

On E_m we have $s \mapsto D_vL(\gamma(s),\dot{\gamma}(s)) \in L^\infty([a,b];\mathbb{R}^n)$, therefore after taking integration over $s \in E_m$ and let $k \to \infty$ we obtain

$$\int_a^b L(\gamma_k(s),\dot{\gamma}_k(s))ds \geqslant \int_{E_m} L(\gamma_k(s),\dot{\gamma}_k(s))ds - |[a,b]\backslash E_m|C_\theta.$$

Therefore

$$\liminf_{k\to\infty}\int_a^b L(\gamma_k(s),\dot{\gamma}_k(s))ds \geqslant \int_{E_m} L(\gamma(s),\dot{\gamma}(s))ds - C_\theta\delta \geqslant \int_a^b L(\gamma(s),\dot{\gamma}(s))ds - 2C_\theta\delta.$$

Let $m \to \infty$ and $\delta \to 0$ we obtain the conclusion. □

Assume $L \in C^k(\mathbb{R}^n \times \mathbb{R}^n)$ for some $k \geqslant 2$ and $D_v^2L(x,v) \succ 0$ for all (x,v) and the superlinearity (2.6). We recall that if γ is a continuous piece-wise C^1 minimizer (or an extremal curve) then γ is C^k and γ satisfies the Euler-Lagrange equation

$$\frac{d}{dt}\Big(D_vL\big(\gamma(t),\dot{\gamma}(t)\big)\Big) = D_xL\big(\gamma(t),\dot{\gamma}(t)\big), \qquad a \leqslant t \leqslant b. \tag{2.8}$$

2.6. **Hamiltonian and Lagrangian viewpoints.** Before showing similar result for absolutely continuous minimizers, let us give some remarks on Lagrangian and Hamiltonian viewpoints. For a continuous and piece-wise C^1 curve γ, denote by

$$\begin{cases} x(t) = \gamma(t) \\ p(t) = D_vL(\gamma(t),\dot{\gamma}(t)) \longrightarrow \dot{\gamma}(t) = D_pH(x(t),p(t)). \end{cases}$$

By Legendre's transform we have

$$H(x(t),p(t)) + L(\gamma(t),\dot{\gamma}(t)) = \dot{\gamma}(t)\cdot p(t).$$

If γ satisfies the Euler-Lagrange equation then

$$\dot{p}(t) = D_xL(\gamma(t),\dot{\gamma}(t)) = -D_xH(x(t),p(t)).$$

Thus the Euler-Lagrange equation for γ is the key to the Hamiltonian system of $2n$ variables

$$\begin{cases} \dot{x}(t) = \quad D_pH(x(t),p(t)) \\ \dot{p}(t) = -D_xH(x(t),p(t)). \end{cases} \tag{2.9}$$

The Lipchitz property of $D_pH(x,p)$ and $D_xH(x,p)$ are important here, as they ensure the existence and uniqueness of such a solution $(x(t),p(t))$ for all time. For now, let say $(x(t),p(t))$ exists on a domain.

Lemma 2.8 (Conservation of energy)**.** *On its domain we have* $t \mapsto H(x(t),p(t))$ *is constant.*

Proof. By definition

$$\frac{d}{dt}\Big(H(x(t),p(t))\Big) = D_xH\cdot\dot{x} + D_pH\cdot\dot{p} = 0$$

since $\dot{x} = D_pH$ and $\dot{p} = -D_xH$. □

Remark 4 (Boundedness of the Hamiltonian flow)**.** As long as $(x(t), p(t))$ exists (in its domain) then

$$H(x(t), p(t)) = H(x(0), p(0)) \leqslant C.$$

The super-linearity of H yields that $|p(t)| \leqslant C$. If we assume that $D_pH, D_xH \in \mathrm{Lip}(\mathbb{R}^n \times \overline{B(0,R)})$ for each $R > 0$ then this implies that $(x(t), p(t))$ is defined for all $t \in \mathbb{R}$. In summary, we have

$$\begin{aligned} \text{A priori knowledge} \quad &\longrightarrow \quad \text{Lipschitz vector field} \\ &\longrightarrow \quad \text{Wellposedness of ODEs.} \end{aligned}$$

Definition 6 (Hamiltonian flow and Lagrangian flow)**.** *Assume the wellposedness of* $(x(t), p(t))$ *for all* $t \in \mathbb{R}$*, we define*

$$\phi_t^H(x, p) = (x(t), p(t))$$

where $(x(t), p(t))$ *solves the Hamiltonian system* (2.9) *with initial condition* $(x(0), p(0)) = (x, p)$*. Similarly, assume the wellposedness of solution to the Euler-Lagrange equation for* $t \in \mathbb{R}$*, we define*

$$\phi_t^L(x, v) = (\gamma(t), \dot{\gamma}(t))$$

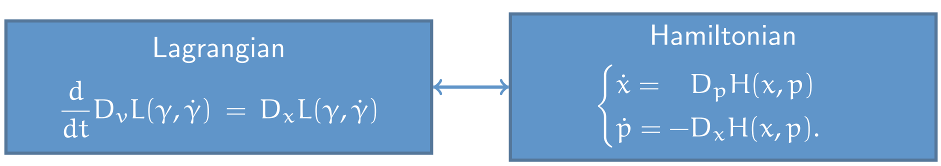
where γ *solves the Euler-Lagrange equation* (2.8) *with initial condition* $(\gamma(0), \dot{\gamma}(0)) = (x, v)$*.*

Recall that, $\mathcal{L}(x, v) = (x, D_vL(x, v))$ for $(x, v) \in \mathbb{R}^n \times \mathbb{R}^n$ is a local C^{k-1} diffeomorphism with its inverse

$$\mathcal{L}^{-1}(x, p) = (x, D_pH(x, p)), \qquad (x, p) \in \mathbb{R}^n \times \mathbb{R}^n.$$

The relation between Hamiltonian flow and Lagrangian flow is

$$\mathcal{L} \circ \phi_t^L \circ \mathcal{L}^{-1} = \phi_t^H. \tag{2.10}$$



Remark 5 (Integrable system)**.** For $\phi = \phi(x, p)$ is continuously differentiable, let us introduce the notation

$$\{H, \phi\} := D_pH \cdot D_x\phi - D_xH \cdot D_p\phi.$$

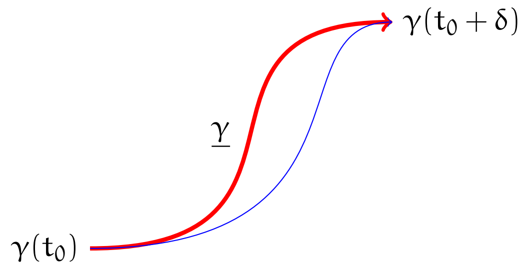
Then the conservation of energy is simply $\{H, H\} = 0$. Such an identity is called an invariant, as it says $t \mapsto \phi(x(t), p(t))$ is constant. If there exist $\phi_1, \dots, \phi_{n-1}$ linearly independent so that $\{H, \phi_i\} = 0$ then the Hamiltonian system (2.9) can be reduced to n unknowns instead of original $2n$ unknowns. Such a system is called *integrable system*.

2.7. **Regularity of absolutely continuous minimizers.** Assume $L \in C^k(\mathbb{R}^n \times \mathbb{R}^n)$ for some $k \geqslant 2$ and $D_v^2L(x, v) \succ 0$ for all (x, v), L is super-linear (2.6) and further that for each $R > 0$ then

$$L, \nabla L \in \mathrm{Lip}(\mathbb{R}^n \times B(0, R)) \cap \mathrm{Lip}(\mathbb{R}^n \times B(0, R)).$$

The ultimate goal is to show that, any minimizer $\gamma \in AC$ is also C^k. The idea of the proof is again, seemingly ad-hoc.

- Pick a point $t_0 \in (a, b)$, let $\underline{\gamma}$ be the solution to Euler-Lagrange equation with initial condition $(\gamma(t_0), v)$ where v is to be chosen such that $\underline{\gamma}(t_0 + \delta) = \gamma(t_0 + \delta)$ for some $\delta > 0$, then $\underline{\gamma} \in C^k$.



- We show $\underline{\gamma}$ is a minimizer on $[t_0, t_0 + \delta]$, thus by the uniqueness of minimizer (we have strict convexity) $\gamma = \underline{\gamma} \in C^k$.

To prepare for the proof, we recall that Lagrangian flow is denoted by $\phi_t^L : (x, v) \mapsto (\gamma(t), \dot{\gamma}(t))$ where

$$\begin{cases} \frac{d}{ds}\left(D_v L(\gamma(s), \dot{\gamma}(s))\right) = D_x L(\gamma(s), \dot{\gamma}(s)), & s > 0 \\ \left(\gamma(0), \dot{\gamma}(0)\right) = (x, v). \end{cases}$$

Lemma 2.9. *Let $x_0 \in \mathbb{R}^n$ and $\pi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ be the projection that maps $(x, v) \to x$, there exists $\delta > 0$ such that for a any fixed $x_0 \in \mathbb{R}^n$ then*

$$B(x_0, C|s|) \subset \pi \circ \phi_s^L\left(\{x_0\} \times B(0, 2C)\right) \qquad \textit{for all } |s| \leqslant \delta.$$

The key point is the same constant C on both sides.

Proof. We take it for granted the fact that, if we start the Lagrangian flow at (x, v) then

$$\phi_t^L(x, v) = \left(\gamma(t, x, v), \frac{\partial \gamma}{\partial t}(t, x, v)\right)$$

gives us $\gamma \in C^k$ defined for short time, as it satisfies the Euler-Lagrange equation. We want to show that

$$\gamma(t, x_0, v) \textit{ can be anywhere in } B(x_0, tC), \textit{ as } v \textit{ varies in } B(0, 2C).$$

This is equivalent to

$$\frac{\gamma(t, x_0, v) - \gamma(0, x_0, v)}{t} \textit{ can be anywhere in } B(0, C) \textit{ as } v \textit{ varies in } B(0, 2C).$$

Let us define

$$\begin{aligned} \Gamma :& (-\varepsilon, \varepsilon) \times \mathbb{R}^n \longrightarrow \mathbb{R}^n \\ & (t, v) \longmapsto \frac{\gamma(t, x_0, v) - \gamma(0, x_0, v)}{t} = \int_0^1 \frac{\partial \gamma}{\partial t}(st, x_0, v)\, ds. \end{aligned}$$

By the Lipschitz bound $|\gamma(y) - \gamma(x_0)| \leqslant C|x_0 - y|$ near x_0 we see that in fact $\Gamma : (-\varepsilon, \varepsilon) \times \mathbb{R}^n \to \overline{B(0, C)}$.

By ODE theory, $\gamma \in C^k$ and thus $(t,v) \mapsto \Gamma(t,v)$ is C^1. If $t=0$ then $\dot\gamma(0) = \frac{d\gamma}{ds}(0,x_0,v) = v$ for $v \in \mathbb{R}^n$, thus

$$\Gamma(0,v) = v \qquad \text{for all } v \in \mathbb{R}^n \qquad \text{and thus} \qquad \frac{\partial \Gamma}{\partial v}(0,\xi) = \mathrm{Id}_{\mathbb{R}^n} \qquad \text{for all } \xi \in \mathbb{R}^n.$$

To use the inverse function theorem, let us define

$$\begin{aligned} \tilde\Gamma : (-\varepsilon,\varepsilon) \times B(0,2C) &\longrightarrow \mathbb{R}^{n+1} \\ (t,v) &\longmapsto \big(t, \Gamma(t,v)\big). \end{aligned}$$

It is clear that $\tilde\Gamma$ is C^1 in $(-\varepsilon,\varepsilon) \times B(0,2C)$ into $\mathbb{R}^{n+1}$ and

$$\tilde\Gamma(0,v) = (0,v) \qquad \text{for all } v \in \mathbb{R}^n$$

and

$$D\tilde\Gamma(0,v) = \begin{pmatrix} 1 & 0 \\ \frac{\partial\Gamma}{\partial t}(0,v) & \mathrm{Id}_{\mathbb{R}^n} \end{pmatrix}$$

which is non-degenerate[3]. We want to show that there exists $\delta > 0$ such that

$$\tilde\Gamma : (-\delta,\delta) \times B(0,2C) \to (-\delta,\delta) \times B(0,C)$$

is a C^1 diffeomorphism. By inverse function theorem[4] for $(0,v) \in \overline{B(0,C)}$, there exists $\delta_v > 0, \tau_v, \kappa_v > 0$ such that

$$\tilde\Gamma : (-\delta_v,\delta_v) \times B(v,\tau_v) \to (-\delta_v,\delta_v) \times B(v,\kappa_v) \tag{2.11}$$

is a C^1 diffeomorphism. Denote its (injective) inverse by

$$\tilde\Gamma_v^{-1} : (-\delta_v,\delta_v) \times B(v,\kappa_v) \to (-\varepsilon,\varepsilon) \times B(0,2C).$$

- There exists $\delta > 0$ such that $\tilde\Gamma : (-\delta,\delta) \times \overline{B(0,C)} \to (-\varepsilon,\varepsilon) \times \overline{B(0,C)}$ is *injective* on its domain. To see that, assume the contrary that we can find $t_n \to 0$ and $v_{1n} \neq v_{2n}$ in $\overline{B(0,C)}$ such that $\tilde\Gamma(t_n,v_{1n}) = \tilde\Gamma(t_n,v_{2n})$. By compactness we can assume $v_{1n} \to v_1, v_{2n} \to v_2$ for $v_1, v_2 \in \overline{B(0,C)}$. In the limit as $t_n \to 0$ we have

  $$v_1 = \tilde\Gamma(0,v_1) = \tilde\Gamma(0,v_2) = v_2.$$

  By the previous argument, as $v_1 = v_2 = v$, there exists $\delta_v > 0, \tau_v > 0$ such that $\tilde\Gamma$ is invertible on $(-\delta_v,\delta_v) \times B(v,\delta_v)$, which means $v_{1n} = v_{2n}$ for n large, a contradiction.
- As a consequence, from (2.11) we obtain that the image $\tilde\Gamma(-\delta,\delta) \times \overline{B(0,C)}$ contains $(-\delta,\delta) \times B(0,C)$.

In other words, we have shown that $B(x_0, C|t|) \subset \pi \circ \phi_t^L\Big(\{x_0\} \times B(0,2C)\Big)$ for all $|t| < \delta$. □

Theorem 2.10. *Let $\gamma \in \mathcal{A}_{ac}$ be a minimizer. Then $\gamma \in C^k$ and γ satisfies the Euler-Lagrange equation.*

Proof. We divide the proof into several steps.

---

[3] A non-degenerate matrix in $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$.

[4] If we apply the inverse function theorem for $(0,0) \in \mathbb{R} \times \mathbb{R}^n$ then there exists $\delta > 0$ and $\tau > 0$ such that

$$\tilde\Gamma : (-\delta,\delta) \times B(0,C) \to (-\delta,\delta) \times B(0,\tau)$$

is a C^1 diffeomorphism. However, this does not give us the same constant C on both sides.

(1) As $\gamma \in AC([a,b];\mathbb{R}^n)$, $\dot\gamma \in L^1([a,b];\mathbb{R}^n)$ and γ is differentiable a.e. in $[a,b]$. We can pick $t_0 \in (a,b)$ such that γ is differentiable at t_0 and $|\dot\gamma(t)_0| \leqslant C_0$.

(2) There exists $\delta_1 > 0$ such that

$$|\gamma(t)-\gamma(t_0)| \leqslant 2C_0|t-t_0| \qquad \text{for } t \in (t_0-\delta_1, t_0+\delta_1) \subset (a,b).$$

Indeed, by definition of diffrentiability at t_0, there is a modulus ω such that

$$\gamma(t)-\gamma(t_0)-\dot\gamma(t_0)(t-t_0) = (t-t_0)\omega(|t-t_0|)$$

for $t \in (t_0-\delta_1, t_0+\delta_1)$, hence $|\gamma(t)-\gamma(t_0)| \leqslant 2C_0|t-t_0|$ for $t \in (t_0-\delta_1, t_0+\delta_1)$.

(3) (*Crucial step*) Using the previous lemma we can find $\nu \in B(0, 2C_0)$ such that

$$\pi \circ \phi^L_{t+\delta_1}(\gamma(t_0), \nu) = \gamma(t+\delta_1)$$

since $\gamma(t+\delta_1) \in B(\gamma(t_0), 2C_0\delta_1)$.

(4) Let $p = D_\nu L(x,\nu)$ and $\varphi \in C^2(\mathbb{R}^n)$ with $D\varphi(\gamma(t_0)) = p$. Using method of characteristics, there exists $\delta_2 > 0$ such that we have existence of a $C^2(\mathbb{R}^n \times (t_0-\delta_2, t_0+\delta_2))$ function that solves

$$\begin{cases} u_t(x,t) + H(x, Du(x,t)) = 0, \\ u(x,t_0) = \varphi(x), \qquad x \in \mathbb{R}^n. \end{cases}$$

Let us denote $\eta(s) = \pi \circ \phi^L_{t_0+s}(\gamma(t_0), \nu)$ for $s \in (-\delta, \delta)$ where $\delta = \min\{\delta_1, \delta_2\}$. By regularity of the Lagrangian flow, we have $\eta \in C^k(t_0, t_0+\delta)$, thus the goal is to show that $\gamma \equiv \eta$ on $(t_0, t_0+\delta)$. We do so by showing that

$$\tilde\gamma(t) = \begin{cases} \gamma(t), \qquad t \notin (t_0, t_0+\delta) \\ \eta(t), \qquad t \in (t_0, t_0+\delta) \end{cases}$$

is a minimizer of the action functional on $[a,b]$ and the result follows by uniqueness of minimizer. It is clear that $\tilde\gamma \in AC([a,b];\mathbb{R}^n)$. Take any generic absolutely continuous curve $\zeta : [t_0, t_0+\delta] \to \mathbb{R}^n$ with $\zeta(t_0) = \gamma(t_0)$ and $\zeta(t_0+\delta) = \gamma(t_0+\delta)$, we have

$$\begin{aligned} u\big(\zeta(t_0+\delta), t_0+\delta\big) - u\big(\zeta(t_0), t_0\big) &= \int_{t_0}^{t_0+\delta} \frac{d}{ds} u\big(\zeta(s), s\big)\,ds \\ &= \int_{t_0}^{t_0+\delta} \big(u_t(\zeta(s), s) + Du(\zeta(s)) \cdot \dot\zeta(s)\big)\,ds \\ &= \int_{t_0}^{t_0+\delta} \Big(Du(\eta(s)) \cdot \dot\zeta(s) - H\Big(\zeta(s), Du(\zeta(s))\Big)\Big)\,ds \leqslant \int_{t_0}^{t_0+\delta} L\Big(\zeta(s), \dot\zeta(s)\Big)\,ds. \end{aligned}$$

On the other hand, by definition of the Lagrangian flow, η satisfies

$$\dot\eta(s) = D_p H\Big(\eta(s), Du(\zeta(s))\Big)$$

and therefore

$$u\big(\eta(t_0+\delta), t_0+\delta\big) - u\big(\eta(t_0), t_0\big) = \int_{t_0}^{t_0+\delta} L\Big(\zeta(s), \dot\zeta(s)\Big)\,ds$$

Therefore η is a minimizer on $(t_0, t_0+\delta)$, hence $\gamma = \eta \in C^k$ locally. □

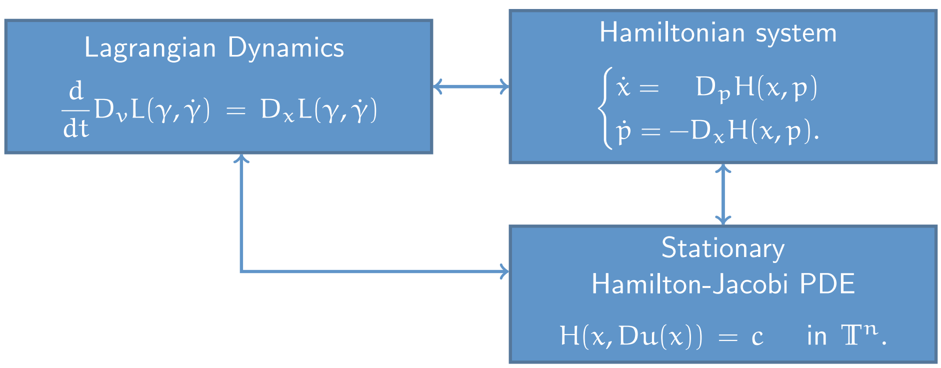
3.1. **Outline.** Our standing assumptions *through out this chapter* will be the following.

$$\begin{cases} L \in C^k(\mathbb{T}^n \times \mathbb{R}^n) \text{ for some } k \geqslant 2, \\ \lim_{|v|\to\infty} \left(\inf_{\mathbb{T}^n} \frac{L(x,v)}{|v|} \right) = +\infty, \\ D_v^2 L(x,v) \succ 0 \text{ for all } (x,v) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{L}$$

As usual, the natural corresponding assumptions on H follows.

$$\begin{cases} H \in C^k(\mathbb{T}^n \times \mathbb{R}^n) \text{ for some } k \geqslant 2, \\ \lim_{|p|\to\infty} \left(\inf_{\mathbb{T}^n} \frac{L(x,p)}{|p|} \right) = +\infty, \\ D_p^2 H(x,p) \succ 0 \text{ for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{H}$$

The connection between PDE and Hamiltonian dynamics can be summarized as follows.



The new object $H(x, Du(x)) = c$ in $\mathbb{T}^n$ will be our main object of study in this chapter. This arises in many areas like large time behavior, homogenization, canonical transformation and ergodic theory.

(1) Invariant under Hamiltonian and Lagrangian flow. Calibrated curves and weak KAM solutions.
(2) The ergodic constant, the existence of calibrated curves with respect to ergodic constant and the existence of weak KAM solutions. We use heavily the object following object

$$h_t(x,y) = \inf \left\{ \int_0^t L(\gamma(s), \dot\gamma(s)) ds : \gamma \in AC([0,t]; \mathbb{T}^n) : \gamma(0) = x, \gamma(y) = t \right\}.$$

(3) The Weak KAM theorem, which implies the existence of a (one-sided) calibrated curve that exists for all time $t \to -\infty$ (or $t \to \infty$).
(4) Properties of the weak KAM solution.

The ultimate goal is to describe solution to the ergodic problem $H(x, Du(x)) = c$ in $\mathbb{T}^n$. We will also describe how it can be applied to find the large time behavior of solution

$u(x,t)$ to the Cauchy problem

$$\begin{cases} u_t(x,t) + H(x, Du(x,t)) = 0 & (x,t) \in \mathbb{R}^n \times (0,T), \\ u(x,0) = g(x) & (x,t) \in \mathbb{R}^n \times \{0\}. \end{cases}$$

We note that solution $u(x,t)$ is strongly related to the following Hamiltonian system

$$\begin{cases} \dot{x}(t) = \quad D_pH(x(t), p(t)), \\ \dot{p}(t) = -D_xH(x(t), p(t)). \end{cases}$$

3.2. **Invariant under Hamiltonian and Lagrangian flow.** The new object $H(x, Du(x)) = c$ in $\mathbb{T}^n$ is connected with the Hamiltonian and Lagrangian flows as we have the following invariant of the graph of $Du(\cdot)$. We will show the invariant under weaker and weaker regularity assumptions as follows.

$$C^2(\mathbb{T}^n) \longrightarrow C^1(\mathbb{T}^n) \longrightarrow \mathrm{Lip}(\mathbb{T}^n) \longrightarrow C(\mathbb{T}^n).$$

Theorem 3.1 (Invariant under the Hamiltonian flow). *Let* $u \in C^2(\mathbb{T}^n)$ *solves* $H(x, Du(x)) = c$ *in* $\mathbb{T}^n$. *For each* $x_0 \in \mathbb{T}^n$, *let* $p_0 = Du(x_0)$ *and consider the Hamiltonian system*

$$\begin{cases} \dot{x}(t) = \quad D_pH(x(t), p(t)), \\ \dot{p}(t) = -D_xH(x(t), p(t)), \end{cases}$$

with initial condition $(x(0), p(0)) = (x_0, p_0)$, *then the system has solution for all time* $t \in \mathbb{R}$ *and furthermore* $p(t) = Du(x(t))$ *for all* $t \in \mathbb{R}$. *In particular, the Hamiltonian flow preserves the graph of* Du, *i.e.,*

$$\phi_t^H(\Gamma) \subset \Gamma \qquad \text{for } t \in \mathbb{R}$$

where $\Gamma = \{(x,p) \in \mathbb{T}^n \times \mathbb{R}^n : p = Du(x)\}$.

Proof. Let $(\mathbf{x}(t), \mathbf{p}(t))$ be solution to the following ODE

$$\begin{cases} \dot{\mathbf{x}}(t) = D_pH\Big(\mathbf{x}(t), Du\big(\mathbf{x}(t)\big)\Big), & t > 0, \\ \mathbf{x}(0) = x_0 \end{cases} \tag{3.1}$$

and $\mathbf{p}(t) = Du\big(\mathbf{x}(t)\big)$ for $t > 0$. First of all, $|Du(x)| \leqslant C$ in $\mathbb{T}^n$ by coercivity and thus $x \mapsto D_pH(x, Du(x))$ for $x \in \mathbb{T}^n$ is Lipschitz (since $u \in C^2(\mathbb{T}^n)$), therefore solution $\mathbf{x}(t)$ exists for all time $t \in \mathbb{R}$ and so is $\mathbf{p}(t)$. We will show that $(\mathbf{x}, \mathbf{p})$ satisfies the Hamiltonian system, then the result follows from the uniqueness of Hamiltonian ODEs. From $\mathbf{p}(t) = Du\big(\mathbf{x}(t)\big)$ we have

$$\dot{\mathbf{p}}(t) = D^2u(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t).$$

From $H(x, Du(x)) = c$ for all $t \in \mathbb{R}$ we obtain

$$D_xH(x, Du(x)) + D^2u(x) \cdot D_pH(x, Du(x)) = 0.$$

Plug in $x = \mathbf{x}(t)$, we obtain $\dot{\mathbf{p}}(t) = -D_xH\big(\mathbf{x}(t), \mathbf{p}(t)\big)$ and the result follows. □

Remark 6. The idea of canonical transformation in classical mechanics is that, to solve the Hamiltonian system of $2n$ variables, if one can find a solution for $H(x, Du(x)) = c$ in $\mathbb{T}^n$ (*cell problem*) then the system can be reduced to (3.1) which consists of n unknowns only.

We remark that the PDE can be defined if $u \in C^1(\mathbb{T}^n)$ only. Our next goal is proving the invariant of Γ under ϕ_t^H when $u \in C^1(\mathbb{T}^n)$ only. Recalling the relation between ϕ_t^H and ϕ_t^L as in (2.10) is given by

$$\mathcal{L} \circ \phi_t^L \circ \mathcal{L}^{-1} = \phi_t^H$$

where $\mathcal{L} : (x, v) \mapsto (x, D_v L(x, v))$ is a local C^{k-1} diffeomorphism. Define

$$\tilde{\Gamma} = \mathcal{L}^{-1}(\Gamma) = \{(x, D_p H(x, Du(x))) : x \in \mathbb{T}^n\}$$

we see that

$$\phi_t^H(\Gamma) \subset \Gamma \qquad \Longleftrightarrow \qquad \phi_t^L(\tilde{\Gamma}) \subset \tilde{\Gamma}.$$

In order words, it suffices to show $\tilde{\Gamma}$ is invariant under the Lagrangian flow. As we cannot differentiate Du, we take a different path of going through the Lagrangian characterization of $C^1(\mathbb{T}^n)$ solution of $H(x, Du(x)) = c$ in $\mathbb{T}^n$ using curves.

Definition 7 (Dominated)**.** *We say* u *is* dominated *by* $L + c$ *in* $\mathbb{T}^n$ *and denote by* $u \prec L + c$ *if* $u \in C(\mathbb{T}^n)$ *satisfying*

$$u(\gamma(b)) - u(\gamma(a)) \leqslant \int_a^b L(\gamma(s), \dot{\gamma}(s))ds + c(b - a)$$

for every $\gamma \in AC([a, b]; \mathbb{T}^n)$*. The set of all* $u \in C(\mathbb{T}^n)$ *which are dominated by* $L + c$ *is denoted by* $\mathbb{D}^c(\mathbb{T}^n)$*.*

Remark 7. We only require a priori that $u \in C(\mathbb{T}^n)$ in the definition of $\mathbb{D}^c(\mathbb{T}^n)$.

Lemma 3.2. *We have the following:*

(i) *If* $u \in \mathbb{D}^c(\mathbb{T}^n)$ *then so is* $u + C$ *for any* $C \in \mathbb{R}$*,*
(ii) $\mathbb{D}^c(\mathbb{T}^n)$ *is a closed convex subset of* $C(\mathbb{T}^n)$*.*
(iii) *If* $u \in \mathbb{D}^c(\mathbb{T}^n)$ *then* u *is Lipschitz with a Lipschitz constant depending only on* L *and* c*.*

Proof. (i) and (ii) are obvious. For (iii), assume that $u \in \mathbb{D}^c(\mathbb{T}^n)$ we show that u is Lipschitz. Take $y, z \in \mathbb{T}^n$ and γ be the straight line connecting them with $\gamma(0) = y$ and $\gamma(1) = z$, let $\tau = |z - y|$ and $v = \frac{z-y}{|z-y|}$ then

$$\gamma(s) = y + sv, \qquad s \in [0, \tau].$$

Clearly $\gamma \in AC([0, \tau]; \mathbb{T}^n)$, thus since $u \in \mathbb{D}^c(\mathbb{T}^n)$ we obtain

$$u(z) - u(y) \leqslant \int_0^\tau L(\gamma(s), v)ds + c\tau \leqslant C|y - z|.$$

where $C = \max\{L(x, v) : x \in \mathbb{T}^n, |v| \leqslant 1\} + |c|$. Reversing the roles of y, z we obtain the Lipschitz property of u. □

Theorem 3.3 (Characterization of subsolutions)**.**

$$\begin{cases} u \in \mathrm{Lip}(\mathbb{T}^n) \\ H(x, Du(x)) \leqslant c \text{ a.e. } \mathbb{T}^n. \end{cases} \qquad \Longleftrightarrow \qquad u \in \mathbb{D}^c(\mathbb{T}^n).$$

Remark 8. If $u \in C^1(\mathbb{T}^n)$ then the equivalence comes from the fundamental theorem of calculus easily.

Proof. Let $u \in \mathrm{Lip}(\mathbb{T}^n)$ such that $H(x, Du(x)) \leqslant c$ a.e. in $\mathbb{T}^n$. Let $\{\eta_\varepsilon\}_{\varepsilon>0} \subset C_c^\infty(\mathbb{R}^n)$ be the standard mollifiers and denote $u^\varepsilon = \eta_\varepsilon * u$ then thanks to convexity and Jensen's inequality

$$H(x, Du^\varepsilon(x)) \leqslant c + \omega(\varepsilon) \qquad \text{in } \mathbb{T}^n.$$

By the fundamental theorem of calculus

$$\begin{aligned} u^\varepsilon(\gamma(b)) - u^\varepsilon(\gamma(a)) &= \int_a^b Du^\varepsilon(\gamma(s) \cdot \dot{\gamma}(s) ds \\ &\leqslant \int_a^b \Big(L(\gamma, \dot{\gamma}) + H(\gamma), Du^\varepsilon(\dot{\gamma}) \Big) ds \leqslant \int_a^b L(\gamma, \dot{\gamma}) ds + c(b-a) + \omega(\varepsilon)(b-a). \end{aligned}$$

Let $\varepsilon \to 0$ we obtain that $u \in \mathbb{D}^c(\mathbb{T}^n)$. Conversely, let $u \in \mathbb{D}^c(\mathbb{T}^n)$ then we already known that u is Lipschitz, thus it is differentiable a.e. in $\mathbb{T}^n$. At a point $x \in \mathbb{T}^n$ where u is differentiable, let $\gamma(s) = x + sv$ for $|v| \leqslant 1$ and $s \in [0, \varepsilon]$, we have

$$\frac{u(x + \varepsilon v) - u(x)}{\varepsilon} \leqslant \frac{1}{\varepsilon} \int_0^\varepsilon L(x + sv, v) ds + c.$$

Let $\varepsilon \to 0^+$, since u is differentiabele at x, we obtain

$$Du(x) \cdot v \leqslant L(x, v) + c$$

This is true for all $v \in \mathbb{R}^n$, thus $H(x, Du(x)) \leqslant c$. □

Remark 9. This characterization can be relaxed to quasi-convex Hamiltonians (convex level-sets), as there is a quasi-convex version of Jensen's inequality, see for example [18].

Theorem 3.4. *Let* $u \in C^1(\mathbb{T}^n)$ *such that* $H(x, Du(x)) = c$ *in* $\mathbb{T}^n$*, then the graph* Γ *of* Du *is invariant under the Hamiltonian flow* ϕ_t^H*.*

We will present a seemingly *ad-hoc* proof via the Lagrangian framework. The idea is choosing an optimal curve in Theorem 3.3 such that the inequality is actually equality, which renders subsolution into solution. We will need the following lemma, which illustrates the idea about equality.

Lemma 3.5. *Let* $u \in C^1(\mathbb{T}^n)$ *solves* $H(x, Du(x)) = c$ *in* $\mathbb{T}^n$*. If* $\gamma : [a, b] \to \mathbb{T}^n$ *is a solution to*

$$\dot{\gamma}(s) = D_p H\big(\gamma(s), Du(\gamma(s))\big), \qquad s \in (a, b),$$

then

$$u(\gamma(b)) - u(\gamma(a)) = \int_a^b L(\gamma(s), \dot{\gamma}(s)) ds + c(b-a).$$

It follows that γ *is a minimizer of the action over* $[a, b]$ *with fixed endpoints* $\gamma(a), \gamma(b)$*, thus* $\gamma \in C^k$ *and it satisfies the Euler-Lagrange equation.*

Proof. By the duality $Du(\gamma(s)) = D_v L(\gamma(s), \dot{\gamma}(s))$ and

$$Du(\gamma(s)) \cdot \dot{\gamma}(s) = H(\gamma(s), Du(\gamma(s))) + L(\gamma(s), \dot{\gamma}(s))$$

for $s \in (a, b)$. Integrating we obtain the conclusion. □

Proof of Theorem 3.4. To show that $\tilde{\Gamma}$ is invariant under ϕ_t^L, for $(x_0, p_0) \in \mathbb{T}^n \times \mathbb{R}^n$, let $(x_0, v_0) = (x_0, D_p H(x_0, p_0)) \in \tilde{\Gamma}$. We consider

$$\begin{cases} \dot{\gamma}(t) = D_p H(\gamma(t), Du(\gamma(t))), & t \in \mathbb{R}, \\ \gamma(0) = x_0. \end{cases}$$

As $Du \in C(\mathbb{T}^n)$ only, we have $x \mapsto F(x) = D_p H(x, Du(x))$ is in $C(\mathbb{T}^n)$, thus we have the existence for all time[5] (but not uniqueness) of $\gamma(t)$ by Cauchy–Peano existence theorem. Lemma 3.5 gives us $\gamma \in C^k$ and γ satisfies the Euler-Lagrange equation (necessary condition to apply the Lagrangian flow), which gives us that

$$\phi_t^L(x_0, v_0) = \big(\gamma(t), \dot{\gamma}(t)\big) = \big(\gamma(t), D_p H(\gamma(t), Du(\gamma(t)))\big) \subset \tilde{\Gamma}.$$

Therefore $\phi_t^L(\tilde{\Gamma}) \subset \tilde{\Gamma}$ and thus $\phi_t^H(\Gamma) \subset \Gamma$. □

Remark 10. The main idea of this ad-hoc proof, that is for any curve γ it is clear that

$$\int_a^b Du(\gamma(s)) \cdot \dot{\gamma}(s) ds \leqslant \int_a^b \Big(L(\gamma(s), \dot{\gamma}(s)) + H(\gamma(s), Du(\gamma(s)))\Big) ds.$$

To achieve the equality, i.e., γ is optimal, we must choose $\dot{\gamma}(s) = D_p H\big(\gamma(s), Du(\gamma(s))\big)$.

Corollary 3.6. *If* $u \in C^1(\mathbb{T}^n)$ *solves* $H(x, Du(x)) = c$ *in* $\mathbb{T}^n$ *then*

$$\tilde{\Gamma} = \{(x, D_p H(x, Du(x))) : x \in \mathbb{T}^n\}$$

is invariant under the Lagrangian flow ϕ_t^L.

3.3. **Calibrated curves and Weak KAM solutions.**

Definition 8. *Given a continuous function* $u \in C(\mathbb{T}^n)$ *and an open interal* $I \subset \mathbb{R}$, *we say a continuous and piece-wise* C^1 *curve* $\gamma : I \to \mathbb{T}^n$ *is* (L, u, c)*-calibrated if for any* $a < b$ *in* I *then*

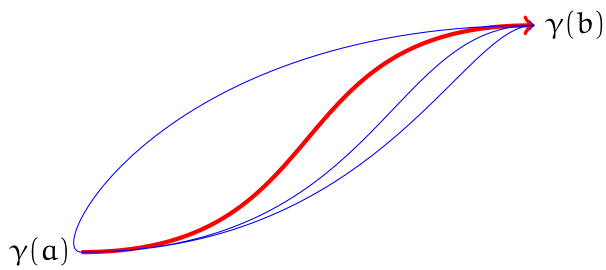
$$u(\gamma(b)) - u(\gamma(a)) = \int_a^b L(\gamma(s), \dot{\gamma}(s)) ds + c(b - a).$$

By definition, if γ is calibrated on $[a, b]$ then γ is calibrated on $[c, d]$ for any $[c, d] \subset [a, b]$.

Theorem 3.7. *If* $u \prec L + c$ *and* $\gamma : I \to \mathbb{T}^n$ *is* (L, u, c)*-calibrated then* $\gamma \in C^k(I; \mathbb{T}^n)$.

Proof. Take $[a, b] \subset I$, for any continuous and piece-wise C^1 curve $\eta : [a, b] \to \mathbb{T}^n$ with fixed endpoints $\eta(a) = \gamma(a)$ and $\eta(b) = \gamma(b)$ we have

$$u(\eta(b)) - u(\eta(a)) \leqslant \int_a^b L(\eta(s), \dot{\eta}(s)) ds + c(b - a).$$



[5]We have the existence on $[-\delta, \delta]$ for small $\delta > 0$ and we keep extending it to $[-n\delta, n\delta]$.

In other words, γ is a minimizer of $L + c$ in the class of continuous and piece-wise C^1 curves connecting $\gamma(a)$ and $\gamma(b)$, thus $\gamma \in C^k$ by Theorem 2.4. □

Theorem 3.8 (Characterization of C^1-solutions). *Let $u \in C^1(\mathbb{T}^n)$ and $c \in \mathbb{R}$. The followings are equivalent.*

(i) $H(x, Du(x)) = c$ *in* $\mathbb{T}^n$.

(ii) $u \prec L + c$ *and for each* $x \in \mathbb{T}^n$*, there exists a* (L, u, c)*-calibrated curve* $\gamma : [-\varepsilon, \varepsilon] \to \mathbb{T}^n$ *with* $\gamma(0) = x$.

(iii) $u \prec L + c$ *and for each* $x \in \mathbb{T}^n$*, there exists a* (L, u, c)*-calibrated curve* $\gamma : [-\varepsilon, 0] \to \mathbb{T}^n$ *with* $\gamma(0) = x$.

(iv) $u \prec L + c$ *and for each* $x \in \mathbb{T}^n$*, there exists a* (L, u, c)*-calibrated curve* $\gamma : [0, \varepsilon] \to \mathbb{T}^n$ *with* $\gamma(0) = x$.

Proof. For (i) implies (ii), it is clear that $u \prec L + c$. As $Du(\cdot) \in C(\mathbb{T}^n)$, we have $x \mapsto D_pH(x, Du(x))$ is continuous, hence by Cauchy-Peano theorem there exists $\varepsilon > 0$ such that the following ODE has a classical solution

$$\begin{cases} \dot{\gamma}(t) = D_pH(\gamma(t), Du(\gamma(t))), & t \in (-\varepsilon, \varepsilon), \\ \gamma(0) = x. \end{cases}$$

In fact one can extend this to $\gamma : \mathbb{R} \to \mathbb{T}^n$ since $D_pH(x, Du(x))$ is bounded, which means the number ε obtained in Cauchy-Peano construction is universal. Because of this choice, we have

$$Du(\gamma(t)) \cdot \dot{\gamma}(t) = H(\gamma(t), Du(\gamma(t))) + L(\gamma(t), \dot{\gamma}(t))$$

for $t \in (-\varepsilon, \varepsilon)$, thus taking integration we obtain that γ is calibrated on its domain. For (iv) implies (i), let us fix $x \in \mathbb{T}^n$. For any $t \in (0, \varepsilon)$ we have

$$u(\gamma(0)) - u(\gamma(-t)) = \int_{-t}^{0} L(\gamma(s), \dot{\gamma}(s))ds + ct.$$

Dividing both sides by t and let $t \to 0^+$ we deduce that

$$Du(\gamma(0)) \cdot \dot{\gamma}(0) = L(\gamma(0), \dot{\gamma}(0)) + c.$$

Therefore

$$H(\gamma(0), Du(\gamma(0))) \geqslant Du(\gamma(0)) \cdot \dot{\gamma}(0) - L(\gamma(0), \dot{\gamma}(0)) = c$$

and thus $H(x, Du(x)) = c$. □

Remark 11. The significant thing in this characterization is that (ii), (iii), (iv) do not require $u \in C^1(\mathbb{T}^n)$, thus one can generalized the notion of solution to the following.

Definition 9 (Weak KAM solutions of *negative type*). *A function $u \in C(\mathbb{T}^n)$ is a* weak KAM solution of negative type *to* $H(x, Du(x)) = 0$ *in* $\mathbb{T}^n$ *if*

- $u \prec L + c$, *and*
- *for each* $x \in \mathbb{T}^n$*, there exists a* (L, u, c)*-calibrated curve* $\gamma : (-\infty, 0] \to \mathbb{T}^n$ *with* $\gamma(0) = x$.

Remark 12. The set of all weak KAM solutions of negative type is denoted by $\mathcal{S}_-$, and such a calibrated curve γ is also called a *backward characteristic* (see [20]).

Question 4. *Let* $H(x,p) = \frac{1}{2}|p|^2$ *for* $(x,v) \in \mathbb{T} \times \mathbb{R}$. *Characterize all solutions in* $\mathbb{D}^c(\mathbb{T})$.

3.4. **Ergodic constant.** Recall that we have $H(x,p) \geqslant -C$ for all $(x,p) \in \mathbb{T}^n \times \mathbb{R}^n$ thanks to super-linearity, we can define the following *additive eigenvalue*, or *ergodic constant*.

Definition 10 (Additive eigenvalue).

$$c(0) := \inf\Big\{c \in \mathbb{R} : \exists\, u \in \mathrm{Lip}(\mathbb{T}^n) : H(x, Du(x)) \leqslant c \text{ a.e.}\Big\}.$$

The constant $c(0)$ *is called* Mañé *critical value (in the language of dynamical system) or* effective Hamiltonian *(in the language of PDEs via Homogenization), or* ergodic constant.

It is clear that $c(0)$ exists and is finite, since if we pick any $u \in \mathrm{Lip}(\mathbb{T}^n)$ and denote $c_u = \mathrm{esssup}_{x\in\mathbb{T}^n} H(x, Du(x))$ then $u \prec L + c_u$ and thus $-C \leqslant c(0) \leqslant c_u$. Roughly speaking, for each $P \in \mathbb{R}^n$ there exists a unique constant $c(P)$ such that $H(x, P + Du(x,P)) = c$ can be solved with a (reasonable) Lipschitz solutions u. If $u \in C^1$, then one obtain a *canonical transformation* in classical machanics that reduce the $2n$ unknowns Hamiltonian system to n unknowns only.

Theorem 3.9 (inf $-$ sup formula). *We have*

$$c(0) = \inf_{u\in \mathrm{Lip}(\mathbb{T}^n)} \big(\mathrm{esssup}_{x\in\mathbb{T}^n} H(x, Du(x))\big) = \inf_{u\in C^1(\mathbb{T}^n)} \max_{x\in\mathbb{T}^n} H(x, Du(x)).$$

Remark 13. In PDE, people often call $c(0)$ the *additive eigenvalue* of the PDE, and the second formula (inf-max formula) above is an analog of the inf formula for the principle eigenvalue of elliptic PDE $Lu = \lambda u$, where

$$\lambda = \inf_{\varphi\in H^1_0(\Omega)} \frac{\langle L\varphi, \varphi\rangle}{\langle \varphi, \varphi\rangle}$$

where the inner product is taken in $L^2(\Omega)$.

Example 2. *Let us see how to find* $c(0)$ *when* $H(x,p) = \frac{1}{2}|p|^2 - V(x)$ *for* $(x,p) \in \mathbb{T}^n \times \mathbb{R}^n$, *where* $\min_{\mathbb{T}^n} V = 0$. *Naively, we have*

$$c(0) = \inf_{\varphi\in C^1(\mathbb{T}^n)} \max_{x\in\mathbb{T}^n} H(x, D\varphi(x)) \geqslant \inf_{\varphi\in C^1(\mathbb{T}^n)} \left(\max_{x\in\mathbb{T}^n}(-V(x))\right) = 0.$$

Choose $\phi \equiv 0$ *then* $c(0) \leqslant \max_{x\in\mathbb{T}^n}(-V(x)) = 0$, *therefore* $c(0) = 0$.

Remark 14. In the language of homogenization, the effective Hamiltonian is defined by

$$\overline{H}(p) = \inf_{\varphi\in C^1(\mathbb{T}^n)} \max_{x\in\mathbb{T}^n} \Big(H(x, p + D\varphi(x))\Big).$$

We have $\overline{H}(0) = c(0)$, and we know that $\overline{H}$ is convex. One of the open problem is to understand deeply the behavior of $\overline{H}$. In particular, where does $\overline{H}$ behaves nicely? What is the set of singularities of $\overline{H}$?

Theorem 3.10 (Existence). *There exists* $u \in \mathrm{Lip}(\mathbb{T}^n)$ *such that* $u \prec L + c(0)$. *In other words,*

$$H(x, Du(x)) \leqslant c(0) \qquad \text{a.e. in } \mathbb{T}^n.$$

In fact, one can show that there exists $u \in \mathrm{Lip}(\mathbb{T}^n)$ such that $H(x, Du(x)) = c(0)$ a.e. in $\mathbb{T}^n$ (or even a *viscosity solution*). One way is using the *vanishing discount* procedure in PDE.

Proof. Let $(u_k, c_k) \in \mathrm{Lip}(\mathbb{T}^n) \times \mathbb{R}$ such that $u_k \prec L + c_k$ a.e. in $\mathbb{T}^n$ ($H(x, D(u_k) \leqslant c_k$) and $c_k \to c(0)$, the existence follows from the definition of $c(0)$. Let $\tilde{u}_k(x) = u_k(x) - u_k(0)$ then $\tilde{u}_k$ is bounded uniformly in $\mathbb{T}^n$, thus by Arzelà-Ascoli theorem $\tilde{u}_{k_j} \to u$ uniformly on $\mathbb{T}^n$ for some $u \in \mathrm{Lip}(\mathbb{T}^n)$. To show that $u \prec L + c(0)$, pick $\gamma \in AC([a,b]; \mathbb{T}^n)$ we have

$$u_k(\gamma(b)) - u_k(\gamma(a)) \leqslant \int_a^b L(\gamma(s), \dot{\gamma}(s))ds + c(0)(b-a).$$

Let $k_j \to \infty$ we obtain

$$u(\gamma(b)) - u(\gamma(a)) \leqslant \int_a^b L(\gamma(s), \dot{\gamma}(s))ds + c(0)(b-a)$$

and thus $u \prec L + c(0)$. $\square$

Remark 15. We use the Lagrangian framework here as it is clear and stable under uniform convergence. The PDE framework is harder a bit, as if we start with

$$\begin{cases} H(x, D\tilde{u}_k(x)) \leqslant c_k & \text{a.e. in } \mathbb{T}^n, \\ \tilde{u}_k \to u & \text{uniformly}, \end{cases}$$

then it is harder to prove that $H(x, Du(x)) \leqslant c(0)$ a.e., as $Du_k \rightharpoonup Du$ is an issue here. It can be resolved in the framework of viscosity solution.

Remark 16.

- If $\gamma : (-\infty, 0] \to \mathbb{T}^n$ is a calibrated curve with $\gamma(0) = x$, then some questions of interest are
  - Is there a rotation vector $q = \lim_{t \to -\infty} \frac{\gamma(t)}{t}$?
  - Ergodic behavior of γ.

$$\begin{cases} \dot{\gamma}(s) = D_pH(\gamma(s), Du(\gamma(s))), & s < 0, \\ \gamma(0) = x. \end{cases}$$

Proposition 3.11 (Stability of calibrated curves)**.** *The following claims hold.*

(a) *If* $I = \bigcup_{k \in \mathbb{N}} I_k$ *with* $I_i \subset I_2 \subset \ldots$ *and* $\gamma : I \to \mathbb{T}^n$ *such that* $\gamma|_{I_k}$ *is* (L, u, c)*-calibrated then* γ *is* (L, u, c)*-calibrated on* I*.*

(b) *Let* $\{\gamma_k\}_{k \in \mathbb{N}} \subset C^1([a,b]; \mathbb{T}^n)$ *such that* $\gamma_k \to \overline{\gamma}$ *in the topology of* $C^1([a,b]; \mathbb{T}^n)$*. If* γ_k *is* (L, u, c)*-calibrated for all* $k \in \mathbb{N}$ *then so is* $\overline{\gamma}$*.*

Remark 17. In general, additive eigenvalues occur in all kinds of nonlinear elliptic PDE that have a maximum principle. One particular example is

$$-\Delta u + |Du|^2 + V(x) = c(0) \qquad \text{in } \mathbb{T}^n$$

where $V \in C(\mathbb{T}^n)$. This came from the rate function in large deviation theory. The relation between the additive eigenvalue $c(0)$ and the principle eigenvalue to Laplace equation can be seen via a Hopf-Cole transform

$$\begin{aligned} \varphi(x) = e^{-u(x)} \quad &\Longrightarrow \quad \Delta\varphi = e^{-u}(-\Delta u + |Du|^2) \\ &\Longrightarrow \quad -\Delta\varphi - V(x)\varphi = -c(0)\varphi \end{aligned}$$

and thus $-c(0)$ is the principle eigenvalue of the operator $-(\Delta + V)$.

Remark 18. Open question: $-\varepsilon\Delta u^\varepsilon + |Du^\varepsilon|^2 + V(x) = c^\varepsilon(0)$ in $\mathbb{T}^n$. What is the behavior of $\{u^\varepsilon\}$ as $\varepsilon \to 0$ and fine details of expansion of $c^\varepsilon(0) - c(0)$? Does $u^\varepsilon(x) - u^\varepsilon(0)$ converges uniformly in the full sequence to u solving $|Du|^2 + V = c(0)$?

3.5. **Existence of calibrated curves.**

Theorem 3.12 (Relation between calibrated curve and ergodic constant)**.** *Assume* $u \prec L + c$ *and* $\gamma : I \to \mathbb{T}^n$ *is* (L, u, c)*-calibrated. If* I *is of infinite length then we must have* $c = c(0)$*.*

Proof. By Theorem 3.10 there exists $v \in \mathrm{Lip}(\mathbb{T}^n)$ such that $v \prec L + c(0)$, therefore for every $(a, b) \subset I$ we have

$$u(\gamma(b)) - u(\gamma(a)) = \int_a^b L\big(\gamma(s), \dot\gamma(s)\big)\,ds + c(b-a)$$
$$v(\gamma(b)) - v(\gamma(a)) \leqslant \int_a^b L\big(\gamma(s), \dot\gamma(s)\big)\,ds + c(0)(b-a).$$

We deduce that

$$0 \leqslant \big(c - c(0)\big)(b-a) \leqslant \sqrt{n}\Big(\mathrm{Lip}(u) + \mathrm{Lip}(v)\Big).$$

Therefore if I is unbounded we must have $c = c(0)$ (the critical value). □

Corollary 3.13. *In order to have a weak KAM solution of negative (or positive) type, we must have* $c = c(0)$*.*

Lemma 3.14. *Let* $u \prec L + c$ *and* $\gamma : [a, b] \to \mathbb{T}^n$ *be a* (L, u, c)*-calibrated curve. If* u *is differentiable at* $\gamma(t)$ *for* $t \in (a, b)$ *then the gradient* $Du(\gamma(t))$ *can be computed as*

$$\begin{cases} Du(\gamma(t)) = D_v L(\gamma(t), \dot\gamma(t)) \\ H(\gamma(t), Du(\gamma(t))) = c. \end{cases}$$

Proof. As a calibrated curve, $\gamma \in C^k(a, b)$. If u is differentiable at $t_0 \in (a, b)$ then

$$u(\gamma(t)) - u(\gamma(t_0)) = \int_{t_0}^t L(\gamma(s), \dot{\gamma(s)})\,ds + c(t - t_0).$$

Since u is differentiable at $\gamma(t_0)$, we obtain

$$Du(\gamma(u(t_0))) \cdot \dot\gamma(t_0) = L(\gamma(t_0), \dot\gamma(t_0)) + c.$$

Thus $c = H\left(\gamma(t_0), Du(\gamma(t_0))\right)$ and $Du(\gamma(t)) = D_v L(\gamma(t), \dot\gamma(t))$ for any $t \in (a, b)$. □

Theorem 3.15. *If* $u \prec L + c$ *and* $\gamma : [a, b] \to \mathbb{T}^n$ *is a* (L, u, c)*-calibrated curve then* u *is differentiable at* $\gamma(t)$ *for all* $t \in (a, b)$ *(however it may fail to be differentiable at the two end-points).*

Remark 19. The idea of the proof can be easy understood using the language of viscosity solution, or superdifferential. Basically, we show that at any point $\gamma(t)$, the supergradient and subgradient of u are the same and contain only one vector, which is $\nabla u(\gamma(t))$. To do so, we construct C^1 functions that touch u from above and below, and show that their gradients are the same.

Proof. Ley y near $x = \gamma(t)$, we can construct naturally curves that going from $\gamma(a)$ to $\gamma(b)$ passing y naturally by shifting γ. Let us define

$$\eta(s) = \gamma(s) + \left(\frac{s-a}{t-a}\right)(y-x), \qquad s \in [a,t].$$

Clearly $\eta(a) = \gamma(a)$ and $\eta(t) = y$, using $u \prec L + c$ we have

$$u(\eta(t)) - u(\eta(a)) \leqslant \int_a^t L(\eta(s), \dot{\eta}(s))ds.$$

Therefore

$$u(x) \leqslant \underbrace{u(\eta(a)) + \int_a^t L\left(\gamma(s) + \tfrac{s-a}{t-a}(y-x), \dot{\gamma}(s) + \tfrac{(y-x)}{t-a}\right) ds.}_{\varphi(y)}$$

It is crucial that

$$\boxed{\gamma \text{ is calibrated} \qquad \Longrightarrow \qquad \varphi(x) = u(x)}$$

and $u \leqslant \varphi$ for all y near x. It is clear that $y \mapsto \varphi$ is continuously differentiable, hence $D\varphi(x) \in D^+u(x)$. Similarly, we can define

$$\zeta(s) = \gamma(s) + \frac{b-s}{b-t}(y-x), \qquad s \in [t,b].$$

Clearly $\zeta(b) = \gamma(b)$ and $\zeta(t) = y$, using $u \prec L + c$ we have

$$u(\eta(b)) - u(\eta(t)) \leqslant \int_t^b L(\zeta(s), \dot{\zeta}(s))ds.$$

Therefore

$$u(x) \geqslant \underbrace{u(\zeta(b)) - \int_t^b L\left(\gamma(s) + \tfrac{b-s}{b-t}(y-x), \dot{\gamma}(s) + \tfrac{(y-x)}{b-t}\right) ds.}_{\psi(y)}$$

Again, $\psi(x) = u(x)$ and $\psi(y) \leqslant u(x)$ for y near x, therefore $D\psi(x) \in D^-u(x)$. However, $\varphi - \psi \geqslant 0$ everywhere and $(\varphi - \psi)(x) = 0$, thus $D\varphi(x) = D\psi(x)$, hence u is differentiable at $x = \gamma(t)$. □

Remark 20. We can actually show directly that

$$D\varphi(x) = D\psi(x) = D_v L(\gamma(t), \dot{\gamma}(t))$$

as follows. As γ satisfies the Euler-Lagrange equation, we have

$$\begin{aligned} D\varphi(x) &= \frac{1}{t-a}\int_a^t \left[(s-a)\frac{d}{ds}(D_v L(\gamma, \dot{\gamma})) + D_v L(\gamma, \dot{\gamma})\right] ds \\ &= \frac{1}{t-a}\int_a^t \frac{d}{ds}\Big[(s-a)D_v L(\gamma(s), \dot{\gamma}(s))\Big] ds = D_v L(\gamma(t), \dot{\gamma}(t)). \end{aligned}$$

The existence of a calibrated curve and the weak KAM theorem are strongly related.

Theorem 3.16 (Weak KAM theorem). *There exists $u \in \mathrm{Lip}(\mathbb{T}^n)$ such that $u \prec L + c(0)$ and for every $x \in \mathbb{T}^n$, we can find a calibrated curve $\gamma : (-\infty, 0] \to \mathbb{T}^n$ with $\gamma(0) = x$, i.e., for all $t, t' \geqslant 0$ then*

$$u(\gamma(t')) - u(\gamma(t)) = \int_t^{t'} \Big(L(\gamma(s), \dot{\gamma}(s)) + c(0) \Big) ds.$$

We will prove this Theorem after some more preparations.

Remark 21. We have shown that in such a situation, u is differentiable at $\gamma(t)$ for all $t \in (-\infty, 0)$. It might be the case that u is not differentiable at the end point $x = \gamma(0)$.

Remark 22. One may keep running the Lagragian flow passing $t = 0$ with the velociy $\gamma'(0^-)$ to have a nice, smooth curve defined for the whole $\mathbb{R} \to \mathbb{T}^n$. However, there is nothing to guarantee that this curve is *calibrated*, since a minimizer of the problem

$$u(\gamma(t)) - u(\gamma(t')) = \int_{t'}^{t} L(\gamma(s), \dot{\gamma}(s)) ds + (t - t') c(0)$$

is a solution to the Euler-Lagrange equation (a solution to the Lagrangian flow) but the inverse may not be true.

3.6. **Minimal action for a given time.** We define $h_t(x, y)$ to be the minimal cost it takes to travel from $x \to y$ in a fixed amount of time $t > 0$.

Definition 11. *For given $x, y \in \mathbb{T}^n$, denote by*

$$h_t(x, y) = \inf_{\gamma \in AC([0,t];\mathbb{R}^n)} \left\{ \int_0^t L(\gamma(s), \dot{\gamma}(s)) ds : \gamma(0) = x, \gamma(t) = y \right\}. \tag{3.2}$$

Remark 23. We note that $h_t(x, y)$ is not a new object, but in this section we focus on the dependence of h_t on (x, y) more systematically. We can think of $h_t(x, y)$ as some sort of distance from x to y.

Proposition 3.17 (Properties of h_t). *We have the following:*

1. $h_t(x, y) \geqslant t \left(\inf_{(x,v)} L(x, v) \right)$.
2. *(Dynamic programming principle) For $x, y \in \mathbb{T}^n$ and $t, t' > 0$ we have*

$$h_{t+t'}(x, y) = \inf_{z \in \mathbb{T}^n} \Big(h_t(x, z) + h_{t'}(z, y) \Big).$$

3. *If $u \in C(\mathbb{T}^n)$ with $u \prec L + c$ then $u(y) - u(x) \leqslant h_t(x, y) + ct$ for all $x, y \in \mathbb{T}^n$ and $t > 0$.*
4. *There exists an extremal curve (critical point of the action functional) $\gamma \in C^k([0, t])$ such that*

$$h_t(x, y) = \int_0^t L(\gamma(s), \dot{\gamma}(s)) ds.$$

5. *We have $h_t(x, x) + c(0)t \geqslant 0$.*
6. *For each $u \in \mathcal{S}_-$ and $t_0 > 0$, there exists a constant $C = C(u, t_0)$ such that for all $t \geqslant t_0$ there holds*

$$-2\|u\|_{L^\infty(\mathbb{T}^n)} \leqslant h_t(x, y) + c(0)t \leqslant 2\|u\|_{L^\infty(\mathbb{T}^n)} + C.$$

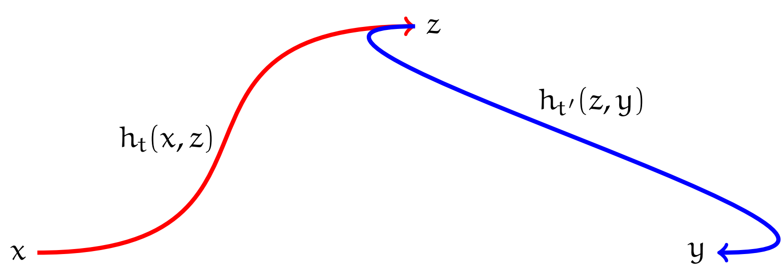
We postpone the uniform Lipschitz property of $h_t(\cdot, \cdot)$ for $t \geqslant \delta > 0$.

Proof. We observe that (1) and (3) are obvious, while (4) follows from the existence of a C^k minimizer, which is also an extremal curve. To show (2), let γ connecting x to y in time t and ζ connecting y to z in time t', connecting them we have a curve η that connects x to z in time $t+t'$, hence

$$\int_0^{t+t'} L(\eta,\dot\eta)ds = \int_0^t L(\gamma,\dot\gamma) + \int_t^{t+t'} L(\zeta(s-t),\dot\zeta(s-t))ds = \int_0^t L(\gamma,\dot\gamma)\,ds + \int_0^{t'} L\left(\zeta,\dot\zeta\right)ds.$$

Taking the infimum we have $h_{t+t'}(x,z) \leqslant h_t(x,y) + h_{t'}(y,z)$, therefore

$$h_{t+t'}(x,y) \leqslant \inf_{z\in\mathbb{T}^n}\Big(h_t(x,z)+h_{t'}(z,y)\Big).$$



Conversely, for every curve curve γ connecting x to y in time $t+t'$, we can pick $z=\gamma(t)$, then obviously the reverse inequality holds.
For (5), if $t>0$ and $x\in\mathbb{T}^n$, take a weak KAM solution $u\in\mathcal{S}_-$, then[6] as $u \prec L+c(0)$, we find that

$$0 = u(x)-u(x) \leqslant h_t(x,x)+c(0)t.$$

For (6), the lower bound is rather obvious. The upper bound is more important as we need it to be uniform for $t \geqslant t_0$. Let $\xi:(-\infty,0]\to\mathbb{T}^n$ be a calibrated curve ending at $\xi(0)=y$. For $t>t_0$, we pick $z=\xi(t_0-t)$ and connect $x\to z$ by finding a minimizer $\gamma:[0,t_0]\to\mathbb{T}^n$ with $\gamma(0)=$ and $\gamma(t_0)=z$ and

$$h_{t_0}(x,z) = \int_0^{t_0} L(\gamma(s),\dot\gamma(s))ds.$$

We have $|h_{t_0}(x,z)| \leqslant C(t_0)$ independent of z. Indeed, simply using the straight line $\eta(s)=x+(s/t_0)(z-x)$ then[7]

$$h_{t_0}(x,z) \leqslant \int_0^{t_0} L\left(x+\frac{s}{t_0}(z-x),\frac{z-x}{t_0}\right)ds \leqslant t_0\left(\sup_{\mathbb{T}^n\times\overline{B}(0,\sqrt{n}/t_0)} L(x,v)\right).$$

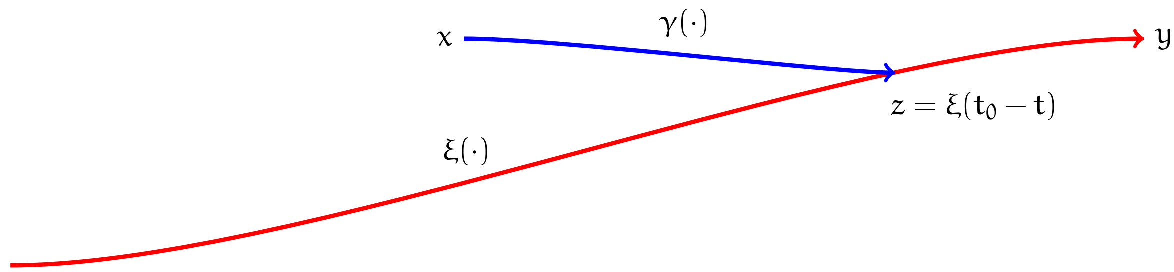
We have

$$\eta(s)=\begin{cases}\gamma(s) & s\in[0,t_0]\\ \xi(s-t) & s\in[t_0,t]\end{cases}$$

is a path connecting x and y.

[6]Actually here any Lipschitz function u that is a subsolution to $H(x,Du)\leqslant c(0)$ is enough, the existence of such a solution follows from Theorem 3.10.

[7]We could do better by using the fact that as a minimizer, the velocity $|\dot\gamma(\cdot)|$ is bounded uniformly, then we might not need the boundedness of the torus here.



The key is along the calibrated curve we have *optimality*

$$\begin{aligned} h_t(x,y) + c(0)t &\leqslant \int_0^t L(\eta(s),\dot{\eta}(s))\,ds + c(0)t \\ &= \underbrace{\int_0^{t_0} \Big(L(\gamma(s),\dot{\gamma}(s)) + c(0)\Big)\,ds}_{C(t_0)} + \underbrace{\int_{t_0}^{t-t_0} \Big(L(\xi(s),\dot{\xi}(s)) + c(0)\Big)\,ds}_{u(t-t_0)-u(t_0)} \\ &\leqslant C(t_0) + 2\|u\|_{L^\infty(\mathbb{T}^n)}. \end{aligned}$$

□

Remark 24. Fix $x \in \mathbb{T}^n$, $y \mapsto h_t(x,y)$ can be thought of as a fundamental solution to

$$\begin{cases} \Phi_t(y,t) + H(z, D\Phi(y,t)) = 0 & \text{in } \mathbb{T}^n \times (0,\infty), \\ \Phi(y,0) = \delta_x(y). \end{cases}$$

In other words, $h_t(x,y) = \Phi(y,t)$.

Lemma 3.18. *For each $t > 0$, there exists $C_t > 0$ such that*

$$h_t(x,y) \leqslant C_t \qquad \textit{for all } x,y \in \mathbb{T}^n.$$

Consequently, for each $\sigma > 0$, there exists $K_\sigma > 0$ such that if $t \geqslant \sigma$ then all minimizer γ satisfies $|\dot{\gamma}(s)| \leqslant K_\sigma$ for $s \in [0,t]$, i.e., if $|\dot{\gamma}(t_0)| \leqslant K_\sigma$ for some t_0 then $|\dot{\gamma}(t)| \leqslant K_\sigma$ for all $t \geqslant t_0$.

Proof. We connect x to y by a straight line $\gamma(s) = x + \frac{s}{t}(y-x)$ for $s \in [0,t]$, then obviously

$$h_t(x,y) \leqslant \int_0^t L(\gamma,\dot{\gamma})\,ds \leqslant t \max_{z\in\mathbb{T}^n} \left\{ L(x,v) : |v| \leqslant \frac{\sqrt{n}}{t} \right\}.$$

therefore we can choose

$$C_t = t \max_{z\in\mathbb{T}^n} \left\{ L(x,v) : |v| \leqslant \frac{\sqrt{n}}{t} \right\}.$$

Since $\gamma \in C^k([0,t];\mathbb{T}^n)$, the above equation implies that there exists $t_0 \in (0,t)$ such that

$$L(\gamma(t_0),\dot{\gamma}(t_0)) \leqslant \tilde{C}_t, \qquad \tilde{C}_t = \max_{z\in\mathbb{T}^n} \left\{ L(x,v) : |v| \leqslant \frac{\sqrt{n}}{t} \right\}.$$

It is clear that $t \mapsto \tilde{C}_t$ is decreasing, thus if $t \geqslant \sigma$ then $\tilde{C}_t \leqslant \tilde{C}_\sigma$. With that fixed σ, we proceed to get $t_0 \in (0,t)$ such that

$$L(\gamma(t_0),\dot{\gamma}(t_0)) \leqslant \tilde{C}_\sigma.$$

Since L is super-linear, there exists K_σ such that

$$|\dot\gamma(t_0)| \leqslant K_\sigma \quad \Longrightarrow \quad |p(t_0)| = |D_v L(\gamma(t_0), \dot\gamma(t_0))| \leqslant \tilde{K}_\sigma.$$

By conservation of energy, $H(\gamma(s), p(s)) = H(\gamma(t_0), p(t_0))$ for all $s \in (0, t)$, thus

$$H(\gamma(s), p(s)) \leqslant K_\sigma \quad \Longrightarrow \quad |p(s)| \leqslant K_\sigma \quad \forall\, s \in (0, t).$$

In turns we obtain that

$$|\dot\gamma(s)| = |D_p H(\gamma(s), p(s))| \leqslant K_\sigma$$

as well. □

Remark 25. The essence of this lemma is that, to go from x to y in a time t, if $t \to 0^+$ then the total cost blows up. For example, with a constant speed line segment then $v = \frac{y-x}{t} \to \infty$ as $t \to 0$, we have

$$\int_0^t L(\eta(s), \dot\eta(s))ds = \int_0^t L\left(\eta(s), \frac{y-x}{t}\right) ds \to \infty$$

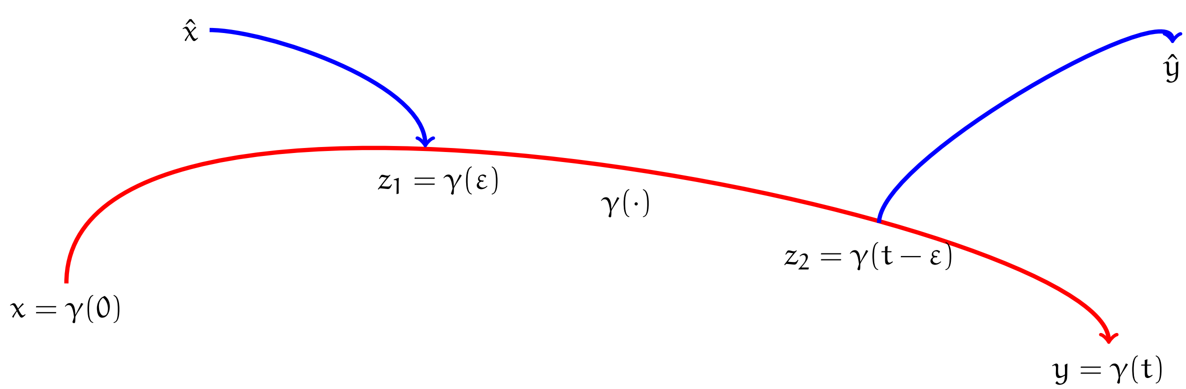
as $t \to 0^+$ since L is superlinear. However, if $t \geqslant \sigma > 0$ for some fixed $\sigma > 0$ then the cost remains bounded.

Theorem 3.19. *For each $\sigma > 0$, there exists $C_\sigma > 0$ such that $h_t : \mathbb{T}^n \times \mathbb{T}^n \to \mathbb{R}$ is Lipschitz with constant C_σ for all $t \geqslant \sigma$.*

Proof. Fix (x, y) and $(\hat{x}, \hat{y})$ in $\mathbb{T}^n \times \mathbb{T}^n$. Take a minimizer path $\gamma : [0, t] \to \mathbb{T}^n$ with $\gamma(0) = x, \gamma(t) = y$ and

$$h_t(x, y) = \int_0^t L(\gamma(s), \dot\gamma(s))ds$$

Fix $\varepsilon > 0$, let $z_1 = \gamma(\varepsilon)$ and $z_2 = \gamma(t - \varepsilon)$, we connect $\hat{x}, \hat{y}$ as following.



Let us define

$$\eta(s) = \begin{cases} \gamma(s) + \frac{\varepsilon - s}{\varepsilon}(\hat{x} - x) & s \in [0, \varepsilon] \\ \gamma(s) & s \in [\varepsilon, t - \varepsilon] \\ \gamma(s) + \frac{s - (t-\varepsilon)}{\varepsilon}(\hat{y} - y) & s \in [t - \varepsilon, t] \end{cases}$$

we obtain a curve connecting $\hat{x}$ to $\hat{y}$ in time t. We have

$$\begin{aligned}h_t(\hat{x},\hat{y}) - h_t(x,y) = &\int_0^{\varepsilon} \left[L\left(\gamma(s) + \frac{\varepsilon - s}{\varepsilon}(\hat{x}-x), \dot{\gamma}(s) - \frac{\hat{x}-x}{\varepsilon}\right) - L(\gamma(s),\dot{\gamma}(s))\right] ds \\ &+ \int_{t-\varepsilon}^{t} \left[L\left(\gamma(s) + \frac{s-(t-\varepsilon)}{\varepsilon}(\hat{y}-y), \dot{\gamma}(s) + \frac{\hat{y}-y}{\varepsilon}\right) - L(\gamma(s),\dot{\gamma}(s))\right] ds.\end{aligned}$$

Let us consider $|\hat{x}-x| + |\hat{y}-y| \leqslant \sigma$. Since $t \geqslant \sigma$, from Lemma 3.18 we have $|\dot{\gamma}(s)| \leqslant K_\sigma$ for all $s \in [0,t]$. We see that $|\dot{\eta}(s)| \leqslant |\dot{\gamma}(s)| + \frac{\sigma}{\varepsilon}$. Choose $\varepsilon = \frac{1}{4}\varepsilon$ we obtain that

$$|\dot{\eta}(s)| \leqslant \tilde{K}_\sigma = K_\sigma + 4 \qquad \text{for } s \in [0,t].$$

Thus there exists C_σ such that

$$|L(x_1,v_1) - L(x_2,v_2)| \leqslant C_\sigma\Big(|x_1 - x_2| + |v_1 - v_2|\Big) \qquad \text{for } |v_1|,|v_2| \leqslant \tilde{K}_\sigma.$$

We deduce that

$$h_t(\hat{x},\hat{y}) - h_t(x,y) \leqslant C_\sigma(|\hat{x}-x| + |\hat{y}-y|)$$

and by symmetry we obtain

$$|h_t(\hat{x},\hat{y}) - h_t(x,y)| \leqslant C_\sigma(|\hat{x}-x| + |\hat{y}-y|).$$

If $|\hat{x}-x| + |\hat{y}-y| > \sigma$, then since we are in $\mathbb{T}^n$, after a fixed finite $m_\sigma \in \mathbb{N}$ middle points we can obtain the same estimate with C_σ replaced by $m_\sigma C_\sigma$. □

Remark 26.

1. It is important to note that even as $t \to \infty$, the Lipchistz constant of $(x,y) \mapsto h_t(x,y)$ remains C_σ.
2. As $t \to 0^+$ howerver, the Lipschitz constant blows up and $C_\sigma \to \infty$ as $\sigma \to 0$. Since

$$|\dot{\gamma}| \leqslant \tilde{C}_t = \max\left\{|L(x,v)| : x \in \mathbb{T}^n, |v| \leqslant \frac{\sqrt{n}}{t}\right\}.$$

3.7. **The Lax–Oleinik semigroup (Optimal control formula).** Given $g \in C(\mathbb{T}^n)$, we can define

$$u(x,t) = \inf_{y\in\mathbb{T}^n} \Big\{h_t(y,x) + g(y)\Big\}.$$

From the viewpoint of Bellman, for $t > 0$ we have

1. $u(x,t) = g(y_0) + \int_0^t L(\gamma(s),\dot{\gamma}(s))ds$ for some $y_0 \in \mathbb{T}^n$ and $\gamma \in C^k([0,t];\mathbb{T}^n)$ minimizer with $\gamma(0) = y_0$ and $\gamma(t) = x$.
2. $u(x,0) = g(x)$ on $\mathbb{T}^n$ and a regularizing effect, even though we start with continuous only datum $u(x)$, instantaneously for $t > 0$ then $u(x,t)$ is Lipschitz with constant at most C_t.
3. Dynamic programming principle:

$$\begin{aligned}u(x,t+\sigma) &= \min_{y\in\mathbb{T}^n} \Big[u_0(y) + h_{t+\sigma}(y,x)\Big] \\ &= \min_{y\in\mathbb{T}^n} \left[u_0(y) + \left(\min_{z\in\mathbb{T}^n} h_t(y,z) + t_\sigma(z,x)\right)\right] = \min_{z\in\mathbb{T}^n} \Big[u(z,t) + h_\sigma(z,x)\Big].\end{aligned}$$

We can show $u(x,t)$ is a viscosity solution to the Hamilton–Jacobi equation

$$\begin{cases} u_t(x,t) + H(x, Du(x,t)) = 0 & \text{in } \mathbb{T}^n \times (0,\infty), \\ u(x,0) = u_0(x). \end{cases}$$

We introduce the following new subject, which is motivated from the formulation above.

Definition 12 (Lax-Oleinik semigroup). $T_t^- : C(\mathbb{T}^n) \to C(\mathbb{T}^n)$ *is defined by*

$$\begin{aligned} T_t^- u(x) &= \inf_{y \in \mathbb{T}^n} \left\{ h_t(y,x) + u(y) \right\} \\ &= \inf \left\{ \int_0^t L(\gamma(s), \dot{\gamma}(s))ds + u(\gamma(0)) : \gamma \in AC([0,t];\mathbb{T}^n), \gamma(t) = x \right\} \end{aligned}$$

for $t > 0$, *and* $T_0^- u = u$.

Remark 27. We can define the Lax-Oleinik semigroup for all function $u : \mathbb{T}^n \to \mathbb{R}$ but we restrict ourselve to $C(\mathbb{T}^n)$ to avoid technicalities.

This object is well-defined from the following proposition.

Proposition 3.20 (Preliminaries properties of T_t^- for $t > 0$)**.** *We have the following*

(i) $\min_{\mathbb{T}^n} u + t \min_{\mathbb{T}^n \times \mathbb{R}^n} L \leqslant T_t^- u \leqslant \min_{\mathbb{T}^n} u + \max_{\mathbb{T}^n \times \mathbb{T}^n} h_t(\cdot,\cdot)$.

(ii) *If* $t \geqslant \sigma > 0$ *then* $x \mapsto T_t^- u(x)$ *is Lipschitz with constant* K_σ.

Proof. The property (i) is rather obvious. For (ii), let $x, z \in \mathbb{T}^n$, we can find $\overline{x} \in \mathbb{T}^n$ so that $T_t^- u(x) = u(\overline{x}) + h_t(\overline{x}, x)$ and $T_t^- u(z) \leqslant u(\overline{x}) + h_t(\overline{x}, z)$, thus

$$T_t^- u(z) - T_t^- u(x) \leqslant h_t(\overline{x}, z) - h_t(\overline{x}, x) \leqslant C_\sigma |x - z|$$

by Theorem 3.19. By symmetry we have the conclusion. □

Proposition 3.21 (Semi-group properties of T_t^- for $t > 0$)**.** *We have the following*

(i) $T_{t+s}^- = T_t^- \circ T_s^-$ *and* $T_t^-(u + c) = T_t^- u + c$ *for* $c \in \mathbb{R}$.

(ii) *If* $u, v \in C(\mathbb{T}^n)$ *and* $u \leqslant v$ *in* $T_t^- u \leqslant T_t^- v$ *(monotonicity).*

(iii) *If* $u = \inf_{i \in I} u_i$ *for a family* $\{u_i\}_{i \in I} \subset C(\mathbb{T}^n)$ *then* $T_t^- u = \inf_{i \in I} T_t^- u_i$.

Proof. From the Dynamic Programming Principle of h_t (Theorem 3.17) we have

$$\begin{aligned} T_{t+s}^- u(x) &= \inf_{y \in \mathbb{T}^n} [u(y) + h_{t+s}(y,x)] \\ &= \inf_{y \in \mathbb{T}^n} \left[u(y) + \inf_{z \in \mathbb{T}^n} \Big(h_t(y,z) + h_s(z,x) \Big) \right] \\ &= \inf_{z \in \mathbb{T}^n} \left[h_s(z,x) + \inf_{y \in \mathbb{T}^n} \Big(u(y) + h_t(y,z) \Big) \right] \\ &= \inf_{z \in \mathbb{T}^n} \Big(h_s(z,x) + T_t^- u(z) \Big) \\ &= T_s^- \left(T_t^- u(x) \right). \end{aligned}$$

The identity $T_t^-(u + c) = T_t^- u$ and the monotonicity are obvious. Lastly, if $u = \inf_{i \in I} u_i$ then it is clear from the monotonicity that $T_t^- u \leqslant \inf_{i \in I} T_t^- u_i$. Conversely, fix $x \in \mathbb{T}^n$ we show that $\inf_{i\in I} T_t^- u_i(x) \leqslant T_t^- u(x)$. For any $\varepsilon > 0$ there exists $i \in I$ so that $u_i(x) - \varepsilon < u(x)$, then

$$\inf_{i\in I} T_t^- u_i(x) - \varepsilon \leqslant T_t^- u_i(x) - \varepsilon = T_t^- \big(u_i(x) - \varepsilon\big) \leqslant T_t^- u(x)$$

Let $\varepsilon \to 0$ we obtain the conclusion. □

Remark 28 (Hopf-Lax formula)**.** In case $H(x,p) = H(p)$, which implies $L(x,v) = L(v)$ then the Hopf-Lax formula can be deduced directly from the Lax-Oleinik semi-group. Let $u_0 \in C(\mathbb{T}^n)$ be the initial condition, we have

$$T_t^- u_0(x) = \inf_{y\in\mathbb{T}^n} \Big[u_0(y) + h_t(y,x)\Big] = \inf_{y\in\mathbb{T}^n} \left[u_0(y) + \inf_{\gamma(0)=y,\gamma(t)=x} \int_0^t L(\dot\gamma(s))ds\right].$$

By Jensen's inequality we have

$$\frac{1}{t}\int_0^t L(\dot\gamma(s))ds \geqslant L\left(\frac{1}{t}\int_0^t \dot\gamma(s)ds\right) = L\left(\frac{x-y}{t}\right)$$

and the inequality can be achieved by choosing γ as a straight line, thus

$$T_t^- u(x) = \inf_{y\in\mathbb{T}^n} \left[u_0(y) + tL\left(\frac{x-y}{t}\right)\right].$$

Corollary 3.22 (Non-expansive property of the semigroup)**.** *For $u, v \in C(\mathbb{T}^n)$ and $t > 0$ then $\|T_t^- u - T_t^- v\|_{L^\infty(\mathbb{T}^n)} \leqslant \|u - v\|_{L^\infty(\mathbb{T}^n)}$. As a consequence $t \mapsto \|T_t^- u - T_t^- v\|_{L^\infty(\mathbb{T}^n)}$ is non-increasing.*

Remark 29. Because of this non-expansive property, we can approximate solution $T_t^- u$ by nice initial data $u_k \to u$ uniformly where $u_k \in \mathrm{Lip}(\mathbb{T}^n)$ instead.

Proposition 3.23. *For a given $u \in C(\mathbb{T}^n)$ we have $\lim_{t\to 0^+} T_t^- u = u$ and $t \mapsto T_t^- u$ is uniformly continuous.*

Proof. We have

$$T_t^- u(x) = \inf_{y\in\mathbb{T}^n} \Big(u(y) + h_t(y,x)\Big) \leqslant u(x) + h_t(x,x) \leqslant u(x) + tL(x,0)$$

by choosing $\gamma(s) = x$ for $x \in [0,t]$, thus $\limsup_{t\to 0^+} T_t^- u(x) \leqslant u(x)$. We can reduce the problem to showing for $u \in \mathrm{Lip}(\mathbb{T}^n)$. For any $\gamma \in AC([0,t],\mathbb{T}^n)$ with $\gamma(t) = x$ we observe that by superlinearity, $L(x,v) \geqslant C|v| - C$ for $v \in \mathbb{R}^n$, thus

$$\begin{aligned} u(\gamma(0)) + \int_0^t L(\gamma(s),\dot\gamma(s))ds &\geqslant u(\gamma(0)) + C\int_0^t |\dot\gamma(s)|ds - Ct \\ &\geqslant u(x) - C|\gamma(t)-\gamma(0)| + C\left|\int_0^t \dot\gamma(s)ds\right| - Ct \\ &= u(x) - C|\gamma(t)-\gamma(0)| + C\,|\gamma(t)-\gamma(0)| - Ct. \end{aligned}$$

Thus $\liminf_{t\to 0^+} T_t^- u(x) \geqslant u(x)$. We also have that $|T_t^- u(x) - u(x)| \leqslant Ct$. □

Corollary 3.24. *Fix $\sigma > 0$, then the family of functions $\{T_t^- u : u \in C(\mathbb{T}^n)\}$ is equi-Lipschitz on $\mathbb{T}^n \times [\sigma, \infty)$. As a consequence, $T_t^-\big(C(\mathbb{T}^n) \cap \overline{B}(0,R)\big)$ is pre-compact in $C(\mathbb{T}^n)$ for $t \geqslant \sigma$ and for any $R > 0$.*

Remark 30. Here we have the regularizing effect, that is $T_t^- u$ is Lipschitz both in space and time immediately when $t > 0$, even though we only started with $u \in C(\mathbb{T}^n)$.

Question 5. *Show that if $u \in \mathrm{Lip}(\mathbb{T}^n)$ then $|T_t^- u(x) - T_t^- u(y)| \leqslant C|x-y|$ for all $x, y \in \mathbb{T}^n$ and for all $t \geqslant 0$.*

Now we are ready to prove Theorem 3.16.

3.8. The weak KAM theorem proof via dynamical system.

Theorem 3.25. *There exists $u_- \in C(\mathbb{T}^n)$ such that*

$$T_t^- u_- + c(0)t = u_- \qquad \text{for all } t \geqslant 0.$$

Proof. Let $u \in C(\mathbb{T}^n)$ such that $u \prec L + c(0)$, i.e., $H(x, Du(x)) \leqslant c(0)$ a.e. in $\mathbb{T}^n$. We show[8]

$$t \mapsto \Big(T_t^- u + c(0)t\Big)$$

is increasing (as a function in $C(\mathbb{T}^n)$).

- We show $u(x) \leqslant T_t^- u(x) + c(0)t$ for all $t \geqslant 0$. Take any $\gamma \in AC([0,t];\mathbb{T}^n)$ with $\gamma(0) = y$ and $\gamma(t) = x$, we have $u \prec L + c(0)$, thus

$$u(x) \leqslant u(y) + \int_0^t L(\gamma(s), \dot{\gamma}(s))ds + c(0)t.$$

  Taking infimum over all γ connecting y to x in time t and then over all $y \in \mathbb{T}^n$ we obtain $u(x) \leqslant T_t^- u(x) + c(0)t$.
- If $0 < t < t'$ we have $u \leqslant T_{t'-t}^- u + c(0)(t'-t)$. Using the semigroup property

$$T_t^- u \leqslant T_t^- \circ T_{t'-t}^- u + c(0)(t'-t) = T_{t'}^- u + c(0)(t'-t)$$

  which gives is the desired property.

Now we show that there exists $C > 0$ such that

$$|T_t^- u(x) + c(0)t| \leqslant C \qquad \text{for all } x \in \mathbb{T}^n, t \geqslant 0.$$

As $u \in \mathrm{Lip}(\mathbb{T}^n)$, $(x,t) \mapsto T_t^- u(x)$ is globally Lipschitz. We observe that, if for any $t \geqslant 0$ there exists $x_t \in \mathbb{T}^n$ such that $T_t^- u(x_t) + c(0)t \leqslant u(x_t)$ then by the Lipschitz property

$$T_t^- u(x) + c(0)t \leqslant u(x_t) + C|x - x_t| \leqslant C$$

and hence we have boundedness. Thus assume the contrary that, there exists $r > 0$ such that, there exists $\delta > 0$ and

$$T_r^- u(x) + c(0)r \geqslant u(x) + \delta$$

for all $x \in \mathbb{T}^n$, we will derive a contradiction. Using the semigroup property we have

$$T_{mr}^- u(x) + c(0)mr \geqslant u(x) + m\delta \qquad \Longrightarrow \qquad T_{mr}^- u(x) + \zeta mr \geqslant u(x)$$

[8]Universal PDE phenomenon: take a subsolution and run the PDE, then we increasing sequence.

for all $x \in \mathbb{T}^n$ and $m \in \mathbb{N}$ and $\zeta = c(0) - \frac{\delta}{r} > 0$. This allows us to define

$$w(x) = \inf_{t \geqslant 0} \Big(T_t^- u(x) + \zeta t \Big) = \inf_{0 \leqslant t \leqslant r} \Big(T_t^- u(x) + \zeta t \Big).$$

It is clear that $w(x) \leqslant u(x)$ and $w \in \mathrm{Lip}(\mathbb{T}^n)$. By the property of the semigroup we have

$$T_s^- (w + \zeta s) = T_s^- \left(\inf_{t \geqslant 0} \Big(T_t^- u(x) + \zeta t \Big) + \zeta s \right) = \inf_{t \geqslant 0} \Big(T_{t+s}^- u(x) + \zeta(t+s) \Big) \geqslant w(x).$$

Take any path $\gamma \in AC([0,s]; \mathbb{T}^n)$ then

$$w(\gamma(s)) \leqslant T_s^- (w + \zeta s) = T_s^- w + \zeta s \leqslant w(\gamma(0)) + \int_0^s L(\gamma(\tau), \dot{\gamma}(s)) d\tau + \zeta s.$$

In other words, we have $w \prec L + \zeta$, thus $H(x, Dw(x)) \leqslant \zeta < c(0)$, which is a contradiction to the definition of $c(0)$. Now $|T_t^- u(x) + c(0)t|$ and is bounded, equi-Lipschitz for $t \geqslant \sigma$ and increasing, thus we can define

$$u_-(x) = \lim_{t \to \infty} \Big(T_t^- u(x) + c(0)t \Big).$$

We show that v is the function that satisfies $T_t^- u_- + c(0)t = u_-$. By continuity ($u_k \to u$ then $T_t^- u_k \to T_t^- u$) we have

$$\begin{aligned} T_s^- u_-(x) + c(0)s &= T_s^- \left[\lim_{t \to \infty} \Big(T_t^- u(x) + c(0)t \Big) \right] + c(0)s \\ &= \lim_{t \to \infty} T_t^- \Big(T_s^- u(x) + c(0)t + c(0)s \Big) = \lim_{t \to \infty} \Big(T_{t+s}^- u(x) + c(0)(t+s) \Big) = u_-(x). \end{aligned}$$

□

Remark 31. In general u_- is not unique, in the next part of the note we will characterize solutions to $H(x, Du(x)) = c(0) \in \mathbb{T}^n$ in terms of minimizing measures.

3.9. The weak KAM theorem proof via fixed point theorem.

Theorem 3.26 (Using Schauder's fixed point theorem)**.** *There exists* $u_- \in C(\mathbb{T}^n)$ *such that*

$$T_t^- u_- + c(0)t = u_- \qquad \textit{for all } t \geqslant 0.$$

Proof using Schauder's fixed point theorem. Let $E = C(\mathbb{T}^n)/\mathbb{R}$, we view each element of E as $[u]$ and $[u_1] = [u_2]$ if $u_1 = u_2 + C$ for some constant $C \in \mathbb{R}$. Also $\|[u]\|_E = \inf_{c \in \mathbb{R}} \|u + c\|_{L^\infty(\mathbb{T}^n)}$.

- As $T_t^-(u + c) = T_t^- u + c$ we can view $T_t^- : E \to E$.
- We recall that for each $\sigma > 0$, the family $\{T_t^- u(x) : (x,t) \in \mathbb{T}^n \times [\sigma, \infty)\}$ is equi-Lipschitz. In other words, for each fixed $\sigma > 0$ we see that $T_\sigma^-(E)$ is equi-Lipschitz in $\mathbb{T}^n$ with constant C_σ and thus for all $[\varphi] \in E$ then $\|[\varphi]\|_E \leqslant C_\sigma \sqrt{n}$.
- By Arzelà–Ascoli theorem, $T_\sigma^-(E)$ is compact in E, thus by Schauder's fixed point theorem there exists $[u_\sigma] \in E$ such that

$$T_\sigma^- ([u_\sigma]) = [u_\sigma] \qquad \Longrightarrow \qquad T_{k\sigma}^- ([u_\sigma]) = [u_\sigma]$$

  for all $k \in \mathbb{N}$.

- Let $\sigma = 2^{-j}$ and $[u_j] \in E$ be the fixed point of $T^-_{2^{-j}}\left([u_j]\right) = [u_j]$, then

$$T^-_{k2^{-j}}\left([u_j]\right) = [u_j] \qquad \text{for all } k \in \mathbb{N}.$$

  By choosing different values of k we obtain that

$$T^-_t\left([u_j]\right) = [u_j] \qquad \text{for all } t > 0.$$

  Consequently, we deduce that

$$T^-_t\left([u]\right) = [u] \qquad \text{for all } t > 0.$$

- For each $t > 0$ we can find $c_t \in \mathbb{R}$ such that $T^-_t u = u + c_t$, which means $t \mapsto c_t$ is additive

$$c_{t+s} = c_t + c_s \qquad \text{for all } s, t \geqslant 0.$$

  It is also clear that $t \mapsto c_t$ is continuous since $t \mapsto T^-_t u$ is continuous, thus $c_t = (-c)t$ for some constant $c \in \mathbb{R}$, then $T^-_t u + ct = u$ for $t \geqslant 0$.

The fact that $c = c(0)$ come indirectly later, as we will show that if $T^-_t u + ct = u$ for all t then u is a weak KAM solution of the negative type, that is for each $x \in \mathbb{T}^n$ there exists a backward characteristic $\gamma_x : (-\infty, 0] \to \mathbb{T}^n$ with $\gamma(0) = x$. Theorem 3.12 gives us that $c = c(0)$. □

Proof of Theorem 3.16. Now we show that the existence of $u \in C(\mathbb{T}^n)$ with $T^-_t u + ct = u$ for all $t \geqslant 0$ implies the existence of a backward characteristic curve ending at x for any $x \in \mathbb{T}^n$.

- Fix $x \in \mathbb{T}^n$, note that $u(x) = T^-_1 u(x) + c$. There exists a minimizer $\gamma : [-1, 0] \to \mathbb{T}^n$ with $\gamma(0) = x$ that realizes $T^-_1 u(x)$, that is

$$u((x) = \left(u(\gamma(-1)) + \int_{-1}^{0} L(\gamma(s), \dot\gamma(s))\,ds \right) + c.$$

- We have $u(\gamma(-1)) = T^-_{-1} u(\gamma(-1)) + c$, thus we can choose $\gamma : [-2, -1] \to \mathbb{T}^n$ that realizes $T^-1u(\gamma(-1))$, i.e.,

$$u(\gamma(-1)) = \left(u(\gamma(-2)) + \int_{-2}^{-1} L(\gamma(s), \dot\gamma(s))\,ds \right) + c.$$

We can connect continuously so that $\gamma : [-2, 0] \to \mathbb{T}^n$ is absolutely continuous with $\gamma(0) = x$ and

$$u(x) = u(\gamma(-2)) + \int_{-2}^{0} L(\gamma(s), \dot\gamma(s))\,ds + 2c.$$

We can repeat this procedure to obtain $\gamma : (-\infty, 0] \to \mathbb{T}^n$ as a calibrated curve for u with $\gamma(0) = x$. By Theorem 3.12 we have $c = c(0)$ and furthermore $\gamma \in C^k((-\infty, 0])$ satisfying the Euler-Lagrange equation (minimizer on each interval $[a, b] \subset (-\infty, 0]$). □

Remark 32. For each $t > 0$ we define μ_t as a probability measure on $\mathbb{T}^n \times \mathbb{R}^n$ that is supported in $\{(\gamma(s), \dot\gamma(s)) : s \in [-t, 0]\}$. Then the behavior of $\gamma(t)/t$ as $t \to -\infty$ can be studied via the limiting measure $\mu_t \rightharpoonup \mu$. These limiting measures are called *Mather measures*.

Remark 33. The weak KAM theorem gives us the existence of $u_0 \in C(\mathbb{T}^n)$ such that

$$T_t^- u_0(x) + c(0)t = u_0(x) \qquad \text{for } x \in \mathbb{T}^n.$$

In other words, if we run the Hamilton–Jacobi equation

$$\begin{cases} u_t(x,t) + H(x, Du(x,t)) = 0 & \text{in } \mathbb{T}^n \times (0,\infty), \\ u(x,0) = u_0(x) \end{cases}$$

with this special initial datum u_0 then solution is $u(x,t) = u_0(x) - c(0)t$, a separable solution. In this way, it is rather clear that we have another proof for the weak KAM solution (solution to the cell problem) via PDE method, the vanishing discount problem (see [20]).

Our standing assumptions *through out this chapter* will be the following.

$$\begin{cases} L \in C^k(\mathbb{T}^n \times \mathbb{R}^n) \text{ for some } k \geqslant 2, \\ \lim_{|v|\to\infty} \left(\inf_{\mathbb{T}^n} \frac{L(x,v)}{|v|} \right) = +\infty, \\ D_v^2 L(x,v) \succ 0 \text{ for all } (x,v) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{L}$$

As usual, the natural corresponding assumptions on H follows.

$$\begin{cases} H \in C^k(\mathbb{T}^n \times \mathbb{R}^n) \text{ for some } k \geqslant 2, \\ \lim_{|p|\to\infty} \left(\inf_{\mathbb{T}^n} \frac{L(x,p)}{|p|} \right) = +\infty, \\ D_p^2 H(x,p) \succ 0 \text{ for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{H}$$

4.1. **Outline.**

(1) Flow invariant measure (under the Lagrangian flow), and the new representation formula of $c(0)$ by minimizing over measures.
(2) Mather measures (minimizing $\langle \mu, L \rangle$ among flow invariant measures).
(3) Mather set $\widetilde{\mathcal{M}}_0$ (closure of the union of supports of all Mather measures) and the projected Mather set.
(4) Important property of point $(x,v) \in \widetilde{\mathcal{M}}_0$: values of u_- (the weak KAM solution) along the flow at two endpoint is exactly the total cost (of L) along the flow.
(5) Compactness of $\widetilde{\mathcal{M}}_0$, uniqueness set for weak KAM solutions and the Lipschitz graph theorem.

The weak KAM theorem gives us the existence of $u_0 \in C(\mathbb{T}^n)$ such that, for each $x \in \mathbb{T}^n$ there exists a calibrated curve $\gamma : (-\infty, 0] \to \mathbb{T}^n$ absolutely continuous (actually it is C^k) such that $\gamma_0(0) = x$. We recall that u_0 is differentiable at $\gamma(t)$ for all $t < 0$ and

$$Du_0(\gamma(t)) = D_v L(\gamma(t), \dot{\gamma}(t))$$

for al $t < 0$. The goal is now to study the behavior of $\gamma(t)$ as $t \to -\infty$, in particular any rotation vector $\gamma(t)/t \to \xi \in \mathbb{R}^n$ as $t \to -\infty$.

4.2. **Invariant measures under Euler-Lagrange flow.** Let us recall that the Lagrangian flow is defined by $\Phi_t^L(x,v) = (\gamma(t), \dot{\gamma}(t))$ where

$$\begin{cases} \frac{d}{ds}\left(DvL(\gamma(s), \dot{\gamma}(s))\right) = D_x L(\gamma(s), \dot{\gamma}(s)) ds, \qquad s \neq 0, \\ \big(\gamma(0), \dot{\gamma}(0)\big) = (x,v). \end{cases}$$

Definition 13. *A probability Radon measure* $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ *is an* invariant measure *or* invariant under the Euler-Lagrange flow *if*

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \psi\left(\phi_t^L(x,v)\right) \, d\mu(x,v) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x,v) \, d\mu(x,v)$$

for all $t \geqslant 0$ *and* $\psi \in BC(T^n \times \mathbb{R}^n; \mathbb{R})$, *the space of bounded continuous functions.*

With this definition, we have a new formula to compute $c(0)$.

Theorem 4.1. *Let $\mathcal{P}_L$ be the set of flow invariant Radon measures $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$, we have*

$$c(0) = -\inf_{\mu \in \mathcal{P}_L} \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v) d\mu(x,v).$$

Proof. Take a flow invariant measure $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$. For each $(x,v) \in \mathbb{T}^n \times \mathbb{R}^n$, we run the Lagrangian flow $\phi_t^L(x,v) = (\zeta_{x,v}(t), \dot{\zeta}_{x,v}(t))$. Let $u \in C(\mathbb{T}^n)$ be the weak KAM solution, since $u \prec L + c(0)$ we have

$$u\left(\pi \circ \Phi_0^L(x,v)\right) - u\left(\pi \circ \Phi_{-1}^L(x,v)\right) \leqslant \int_{-1}^0 L\left(\Phi_t^L(x,v)\right) dt + c(0) \tag{4.1}$$

where $\pi : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{T}^n$ be the natural projection. Denote $\tilde{u}(x,\xi) = u(x)$ on $\mathbb{T}^n \times \mathbb{R}^n$, then

$$u\left(\pi \circ \phi_t^L(x,v)\right) = \tilde{u}\left(\phi_t^L(x,v)\right).$$

Integrating (4.1) with respect to μ, since μ is invariant we deduce that

$$\begin{aligned}
0 &= \int_{\mathbb{T}^n \times \mathbb{R}^n} \left(\tilde{u}\left(\phi_0^L(x,v)\right) - \tilde{u}\left(\phi_{-1}^L(x,v)\right)\right) d\mu(x,v) \\
&\leqslant \int_{-1}^0 \left(\int_{\mathbb{T}^n \times \mathbb{R}^n} L\left(\phi_t^L(x,v)\right) d\mu(x,v)\right) dt + c(0) \\
&= \int_{-1}^0 \left(\int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v) d\mu(x,v)\right) dt + c(0) = \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v) d\mu(x,v) + c(0).
\end{aligned}$$

We deduce that

$$-c(0) \leqslant \inf_{\mu \in \mathcal{P}_L} \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v) d\mu(x,v).$$

For the other direction, fix $x_0 \in \mathbb{T}^n$ let $\gamma : (-\infty, 0] \to \mathbb{T}^n$ be a calibrated curve from the weak KAM theorem, then for all $t > 0$ we have

$$u_-(\gamma(0)) - u_-(\gamma(-t)) = \int_{-t}^0 L(\gamma(s), \dot{\gamma}(s)) ds + c(0)t.$$

Define $\mu_t \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ as

$$\langle \mu_t, \psi \rangle = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x,v)\ d\mu_t(x,v) = \frac{1}{t} \int_{-t}^0 \psi(\gamma(s), \dot{\gamma}(s)) ds \qquad \text{for } \psi \in BC(\mathbb{T}^n \times \mathbb{R}^n).$$

We have

$$\frac{u_-(\gamma(0)) - u_-(\gamma(-t))}{t} = \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v) d\mu_t(x,v) + c(0).$$

Since $\|\dot{\gamma}\|_{L^\infty((-\infty,0])} \leqslant C$ we see that

$$\mathrm{supp}(\mu_t) \subset \mathbb{T}^n \times \overline{B(0,C)} \qquad \text{for all } t < 0.$$

Assume $u_t \rightharpoonup u$ weakly for some $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ as $t_j \to \infty$, we deduce

$$-c_0 = \int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v) d\mu(x,v).$$

We have left to show that μ is a flow invariant measure. Take $\psi \in BC(\mathbb{T}^n \times \mathbb{R}^n)$ we need to show that

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \psi\left(\phi^L_\kappa(x,v)\right) d\mu(x,v) = \int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,v)d\mu(x,v)$$

for all $\kappa > 0$. By definition of μ_t we have

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \psi\left(\phi^L_\kappa(x,v)\right) d\mu_t(x,v) = \frac{1}{t}\int_{-t}^{0} \psi(\gamma(\kappa+s),\dot\gamma(\kappa+s))ds$$

since γ satisfies the Euler-Lagrange equation. Thus we need to show

$$\lim_{t_j\to\infty} \frac{1}{t_j}\int_{-t_j}^{0} \psi(\gamma(\kappa+s),\dot\gamma(\kappa+s))ds = \int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,v)d\mu(x,v).$$

By definition of μ_t we already have

$$\lim_{t_j\to\infty} \frac{1}{t_j}\int_{-t_j}^{0} \psi(\gamma(s),\dot\gamma(s))ds = \int_{\mathbb{T}^n\times\mathbb{R}^n} \psi(x,v)d\mu(x,v).$$

Therefore, we just need to compare the difference

$$\begin{aligned}\frac{1}{t_j}&\left|\int_{-t_j}^{0} \psi(\gamma(s),\dot\gamma(s))ds - \int_{-t_j}^{0} \psi(\gamma(\kappa+s),\dot\gamma(\kappa+s))ds\right| \\ &= \frac{1}{t_j}\left|\int_{-t_j}^{0} \psi(\gamma(s),\dot\gamma(s))ds - \int_{-(t_j+\kappa)}^{\kappa} \psi(\gamma(s),\dot\gamma(s))ds\right| \to 0\end{aligned}$$

as $t_j \to \infty$ since κ is fixed. □

Example 3. *Let* $\eta(t) = x + vt$ *for some irrational vector* v*, then* $\{\eta(t) : t \leqslant 0\}$ *is dense in* $\mathbb{T}^n$*. If we define the measure* μ_t *as above, i.e.,*

$$\langle \mu_t, \phi\rangle = \frac{1}{t}\int_{-t}^{0} \phi(\eta(s),\dot\eta(s))ds$$

then if $\mu_t \rightharpoonup \mu$ *it is not hard to see that* $\mu \equiv dx \times \delta_v$*, where* dx *is the Lebesgue measure on* $\mathbb{T}^n$*. Indeed, let* $\psi(x) = \phi(x,v)$ *then*

$$\langle \mu_t, \phi\rangle = \frac{1}{t}\int_{-t}^{0} \phi\left(x+sv,v\right) ds = \int_{-1}^{0} \psi(x+tv\xi)d\xi.$$

Note that

$$\left|\int_{-1}^{0} \psi(x+tv\xi)d\xi - \int_{-1}^{0} \psi(x+v\xi)d\xi\right| \leqslant \frac{C}{t}$$

we deduce that

$$\langle \mu, \phi\rangle = \int_{-1}^{0} \phi(x+vs,v)ds = \int_{\mathbb{T}^n} \phi(y,v)dy.$$

Therefore it is clear that $\mu \equiv dx \times \delta_v$*.*

4.3. **Mather measures and Mather set.** A flow invariant measure $\mu \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ that minimizes $\langle \mu, L \rangle$ is called a Mather measure, i.e.,

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

Definition 14. *The Mather set is defined by*

$$\widetilde{\mathcal{M}}_0 := \overline{\bigcup_{\mu \text{ Mather}} \text{supp}(\mu)}$$

and the projected Mather set *is defined by*

$$\mathcal{M}_0 = \pi\left(\widetilde{\mathcal{M}}_0\right)$$

where $\pi : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{T}^n$ *is the natural projection* $\pi(x,v) = x$*.*

Recall that if $\gamma : (-\infty, 0] \to \mathbb{T}^n$ is $(L, u, c(0))$-calibrated then

$$u(\gamma(t')) - u(\gamma(t)) = \int_t^{t'} \Big(L(\gamma(s), \dot{\gamma}(s)) + c(0)\Big) ds$$

for $t, t' \leqslant 0$. This also holds for any Lagrangian flows started at any point $(x,v) \in \mathcal{M}_0$, as in the following Proposition.

Proposition 4.2. *For any* $(x,v) \in \widetilde{\mathcal{M}}_0$ *there holds*

$$u\left(\pi \circ \phi^L_{t'}(x,v)\right) - u\left(\pi \circ \phi^L_t(x,v)\right) = \int_t^{t'} \left(L\left(\phi^L_s(x,v)\right) + c(0)\right) ds$$

where u *is a weak KAM solution.*

Proof. Since $u \prec L + c(0)$, we have

$$u\left(\pi \circ \phi^L_{t'}(x,v)\right) - u\left(\pi \circ \phi^L_t(x,v)\right) \leqslant \int_t^{t'} \left(L\left(\phi^L_s(x,v)\right) + c(0)\right) ds. \tag{4.2}$$

This holds for all $(x,v) \in \mathbb{T}^n \times \mathbb{R}^n$, thus taking integration against the Mather measure μ we obtain

$$\begin{aligned} 0 &= \int_{\mathbb{T}^n \times \mathbb{R}^n} \left(u\left(\pi \circ \phi^L_{t'}(x,v)\right) - u\left(\pi \circ \phi^L_t(x,v)\right)\right) d\mu(x,v) \\ &\leqslant \int_t^{t'} \left(\int_{\mathbb{T}^n \times \mathbb{R}^n} L(x,v) d\mu(x,v) + c(0)\right) ds = 0. \end{aligned}$$

Therefore the equality in (4.2) must happen for any (x,v) in the support of μ. By continuity we can extend the equality to $(x,v) \in \widetilde{\mathcal{M}}_0$. □

Remark 34. If $(x,v) \in \widetilde{\mathcal{M}}_0$ and $(x(t), \dot{x}(t)) = \phi^L_t(x,v)$ for $t \in \mathbb{R}$ then by shifting the time forward and backward we obtain that u is differentiable at all $x(t)$ for $t \in \mathbb{R}$, and

$$Du(x(t)) = D_v L(x(t), \dot{x}(t))$$

which gives us

$$\dot{x}(t) = D_p H(x(t), Du(x,t)).$$

Therefore from the fact that $H(x, Du(x)) = c(0)$ we have $Du(x)$ is bounded, which implies $|\dot{x}|$ is bounded.

Corollary 4.3. *Following Theorem 3.15, if $(x, v) \in \widetilde{\mathcal{M}}_0$ then by choosing $t < 0 < s$ and applying Theorem 3.15 to the calibrated curve $\gamma(t) = \pi \circ \phi_t^L(x, v)$ in (t, s) we deduce that any weak KAM soltuion u_- is differentiable at x and $Du_-(x) = D_v L(x, v)$. Thus $H(x, Du_-(x)) = c(0)$ and*

$$\widetilde{\mathcal{M}}_0 \subset \Big\{(x, v) : H\big(x, D_v L(x, v)\big) = c(0)\Big\}.$$

Hence $\widetilde{\mathcal{M}}_0$ is compact.

4.4. **Uniqueness of weak KAM solution of negative type (uniqueness set).**

Theorem 4.4. *If $u_1, u_2 \in C(\mathbb{T}^n)$ are weak KAM solutions of negative type such that $u_1 \equiv u_2$ on $\mathcal{M}_0$ then $u_1 \equiv u_2$ on $\mathbb{T}^n$. We say $\mathcal{M}_0$ is a set of uniqueness.*

Proof. Let $\gamma : (-\infty, 0] \to \mathbb{T}^n$ be a calibrated curve with respect to u_1 with $\gamma(0) = x_0$. For $t < 0$ we have

$$u_1(\gamma(0)) - u_1(\gamma(t)) = \int_t^0 L(\gamma(s), \dot{\gamma}(s))ds + c(0)$$
$$u_2(\gamma(0)) - u_2(\gamma(t)) \leqslant \int_t^0 L(\gamma(s), \dot{\gamma}(s))ds + c(0).$$

For any $t < 0$ we have

$$u_2(x_0) - u_1(x_0) \leqslant u_2(\gamma(t)) - u_1(\gamma(t)) = \frac{1}{t_k}\int_{t_k}^0 \Big(u_2(\gamma(s)) - u_1(\gamma(s))\Big)ds \tag{4.3}$$

for any $t_k < 0$. Define $\mu_t \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ as

$$\langle \mu_t, \psi \rangle = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, v)\; d\mu_t(x, v) = \frac{1}{t}\int_{-t}^0 \psi(\gamma(s), \dot{\gamma}(s))ds$$

for all $\psi \in BC(\mathbb{T}^n \times \mathbb{R}^n)$. Let $\tilde{u}(x, v) = u \circ \pi(x, v)$ where $\pi : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{T}^n$ is the natural projection, we have

$$\frac{1}{t_k}\int_{t_k}^0 \Big(u_2(\gamma(s)) - u_1(\gamma(s))\Big)ds = \int_{\mathbb{T}^n \times \mathbb{R}^n} (\tilde{u}_2 - \tilde{u}_1)\, d\mu_t.$$

Assume $\mu_t \rightharpoonup u$ for some Mather measure μ (according to Theorem 4.1) as $t_k \to \infty$, we obtain from (4.3) that

$$u_2(x_0) - u_1(x_0) \leqslant \int_{\mathbb{T}^n \times \mathbb{R}^n} (\tilde{u}_2 - \tilde{u}_1)\, d\mu = 0$$

if $u_1 = u_2$ on $\mathrm{supp}(\mu)$. By symmetry $u_1(x_0) = u_2(x_0)$. ☐

4.5. **The Lipschitz graph theorem.** We show that a weak KAM solution $u \in C(\mathbb{T}^n)$ of $H(x, Du(x)) = c(0)$ can be $C^{1,1}$ in the projected Mather set. We start by showing such a weak KAM solution is semi-concave locally in $x \in \mathcal{M}_0$.

Theorem 4.5. *If u is a weak KAM solution then there exists $C > 0$ such that for all $x \in \mathcal{M}_0$ and $h \in \mathbb{R}^n$ we have*

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

Proof. For $(x, v) \in \widetilde{\mathcal{M}}_0$, we write $(x(t), \dot{x}(t)) = \phi_t^L(x, v)$ for $t \in \mathbb{R}$. By Proposition 4.2 we have

$$u(x(1)) - u(x(0)) = \int_0^1 L(x(s), \dot{x}(s))ds + c(0)$$
$$u(x(-1)) - u(x(0)) = \int_{-1}^0 L(x(s), \dot{x}(s))ds + c(0)$$

The main idea is the path $x(\cdot)$ going from $x(1)$ to $x(0)$ or $x(-1)$ to $x(0)$ as above are optimal, while if we perturb a bit to path going from $x+h$ for h small then we have sub-optimality instead. We build a path going from $x+h$ to $x(1)$ and a path going from $x-h$ to $x(-1)$, respectively, by

$$x^+(s) = x(s) + (1-s)h, \qquad 0 \leqslant s \leqslant 1,$$
$$x^-(s) = x(s) - (1-s)h, \qquad 0 \leqslant s \leqslant 1.$$

We obtain from $u \prec L + c(0)$ that

$$u(x(1)) - u(x+h) \leqslant \int_0^1 L\Big(x(s) + (1-s)h, \dot{x}(s) - h\Big)ds + c(0)$$
$$u(x(-1)) - u(x-h) \leqslant \int_0^1 L\Big(x(s) - (1-s)h, \dot{x}(s) + h\Big)ds + c(0).$$

Combining these equations and Taylor expansion we obtain

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

where C depends only on $\|D^2L\|_{L^\infty(\mathbb{T}^n \times \overline{B}(0,R))}$ and $R = \|\dot{x}\|_{L^\infty}$ is bounded (Lemma 3.18). $\square$

5.1. **Outline.** Our standing assumptions *through out this chapter* will be the following.

$$\begin{cases} L \in C^k(\mathbb{T}^n \times \mathbb{R}^n) \text{ for some } k \geqslant 2, \\ \lim_{|v|\to\infty}\left(\inf_{\mathbb{T}^n} \frac{L(x,v)}{|v|}\right) = +\infty, \\ D_v^2 L(x,v) \succ 0 \text{ for all } (x,v) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{L}$$

As usual, the natural corresponding assumptions on H follows.

$$\begin{cases} H \in C^k(\mathbb{T}^n \times \mathbb{R}^n) \text{ for some } k \geqslant 2, \\ \lim_{|p|\to\infty}\left(\inf_{\mathbb{T}^n} \frac{L(x,p)}{|p|}\right) = +\infty, \\ D_p^2 H(x,p) \succ 0 \text{ for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{H}$$

Outline.

(1) Introduction to this notion (introduced by Mather around 1993) and its basic properties. Heuristically it is the cost of going from x to y in an infinite amount of time.

5.2. **Introduction.** Following the minimal action for a given time $h_t(x,y)$ as in (3.2), it is natural to ask what is the cost going from x to y in an *infinite* amount of time?

Definition 15 (The Peierls barrier). *We define* $h : \mathbb{T}^n \times \mathbb{T}^n \to R$ *is defined as*

$$h(x,y) = \liminf_{t\to\infty} \Big(h_t(x,y) + c(0)t\Big).$$

Some of the properties of the map $(x,y) \mapsto h(x,y)$ can be derived from properties of the minimal action $h_t(x,y)$ for $t > 0$.

Lemma 5.1 (Properties of $h(x,y)$).

1. $h(x,y)$ *is uniformly bounded and* $(x,y) \mapsto h(x,y)$ *is uniformly Lipschitz.*
2. *If* $u \prec L + c(0)$ *then* $u(y) - u(x) \leqslant h(x,y)$, *consequently* $h(x,x) \geqslant 0$.
3. *(Triangle inequality)* $h(x,y) + h(y,z) \geqslant h(x,z)$, *consequently* $h(x,y) + h(y,x) \geqslant 0$.

Proof.

1. The boundedness of h follows from Proposition 3.17. For $t \geqslant 1$ we have $h_t(\cdot,\cdot)$ is Lipschitz with constant at most C_1, hence as $t \to \infty$ we have h is Lipschitz with constant at most C_1.
2. If $u \prec L + c(0)$ then
$$u(y) - u(x) \leqslant \inf\left\{\int_0^t L(\gamma(s),\dot\gamma(s))ds + c(0)t : \gamma(0) = x, \gamma(t) = y\right\}.$$
Let $t \to \infty$ we deduce that $u(y) - u(x) \leqslant h(x,y)$. Pick $u \in \mathcal{S}_-$ then the result follows.
3. It follows from the fact that $h_t(x,y) = \inf_{z\in\mathbb{T}^n}(h_t(x,z) + h_t(z,y))$. The claim follows from $h(x,x) \geqslant 0$ and $h(x,y) + h(y,x) \geqslant h(x,x)$.

$\square$

5.3. **Connection between the projected Mather set and the Peierls barrier.** We now study deeper properties of the Peierls barrier.

Theorem 5.2. *If* $x \in \mathcal{M}_0$ *then* $h(x,x) = 0$.

Proof.

- Take $x_0 \in \mathcal{M}_0$, then there is $v \in \mathbb{R}^n$ such that $(x,v) \in \widetilde{\mathcal{M}}_0$. Pick μ to be a Mather measure such that $(x,v) \in \operatorname{supp}(\mu)$.
- By Poincaré's recurrence theorem, the current points of ϕ_t^L contained in $\operatorname{supp}(\mu)$ form a dense set in $\operatorname{supp}(\mu)$.
- By the continuity of h, we can assume (x,v) is a recurrent point of ϕ_t^L, then for any $r > 0$ there exists $t_k \to \infty$ such that $\phi_{t_k}^L(x,v) \in B((x,v),r)$ for all $k \in \mathbb{N}$.

Fix $u \in \mathcal{S}_-$, we have

$$u\left(\pi \circ \Phi_t^L(x,v)\right) - u(x) = \int_0^t L\left(\phi_s^L(x,v)\right) ds + c(0)t.$$

As (x,v) a recurrent point, there exists a sequence $t_k \to \infty$ such that $\pi \circ \Phi_{t_k}^L(x,v) \to x$ as $t_k \to \infty$. Let $t_k \to \infty$ we deduce that Therefore

$$\lim_{t_k \to \infty} \left[\int_0^t L\left(\phi_s^L(x,v)\right) ds + c(0)t\right] = 0.$$

Thus the cost of connecting $x_k = \pi \circ \phi_{t_k}^L(x,v)$ to x vanishes as $t_k \to \infty$, hence $h(x,x) = 0$ if $x \in \mathcal{M}_0$. □

Theorem 5.3 (Stability and approximation)**.**

1. *For* $x, y \in \mathbb{T}^n$, *there exists sequence of minimizing extremal curves* $\gamma_k : [0,t_k] \to \mathbb{T}^n$ *with* $t_k \to \infty$ *such that* $\gamma_k(0) = x, \gamma_k(t_k) = y$ *and*
$$h(x,y) = \lim_{t_k \to \infty} \left(\int_0^{t_k} L(\gamma_k(s), \dot{\gamma}(s)) ds + c(0)t_k\right).$$
2. *If* $\gamma_k : [0,t_k] \to \mathbb{T}^n$ *is a sequence of continuous piece-wise* C^1 *curves with* $t_k \to \infty$ *such that* $\gamma_k(0) \to x, \gamma_k(t_k) \to y$ *then*
$$h(x,y) \leqslant \liminf_{k \to \infty} \left(\int_0^{t_k} L(\gamma_k(s), \dot{\gamma}(s)) ds + c(0)t_k\right).$$

Proof. The proof follows from the definition $h(x,y) = \lim_{t \to \infty} \big(h_t(x,y) + c(0)t\big)$.

□

5.4. **Weak attractor.**

Lemma 5.4. *Let* V *be an open neighborhood of* $\widetilde{\mathcal{M}}_0$ *in* $\mathbb{T}^n \times \mathbb{R}^n$. *Then, there exists* $T = T(V)$ *such that if* $\gamma : [0,t] \to \mathbb{T}^n$ *is a minimizing curve with* $t \geqslant T$ *then there exists* $s \in [0,t]$ *such that* $\big(\gamma(s), \dot{\gamma}(s)\big) \in V$.

Proof. Assume the contrary, then we can find $t_k \to \infty$ and $\gamma_k : [0,t_k] \to \mathbb{T}^n$ minimizing curves such that

$$\left\{\left(\gamma_k(s), \dot{\gamma}_k(s)\right) : 0 \leqslant s \leqslant t_k\right\} \cap \overline{V} = \emptyset.$$

We can always assume $t_k \geqslant 1$. From Lemma 3.18 there exists a compact set K such that

$$\left\{\left(\gamma_k(s), \dot{\gamma}_k(s)\right) : 0 \leqslant s \leqslant t_k\right\} \subset \mathbb{T}^n \times K.$$

We now construct a Mather measure from $\{\gamma_k\}$ to get a contradiction. Let $\mu_k \in \mathcal{P}(\mathbb{T}^n \times \mathbb{R}^n)$ be such that

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x, v)\, d\mu_k(x, v) = \frac{1}{t_k} \int_0^{t_k} \psi(\gamma_k(s), \dot{\gamma}_k(s))\, ds$$

for all bounded continuous ψ. We see that $\mathrm{supp}(\mu_k) \subset \mathbb{T}^n \times K$, thus we can find a weak convergent (in measure) subsequence $\mu_k \rightharpoonup \mu$. Clearly $\mathrm{supp}(\mu) \in \mathbb{T}^n \times K$ and similar to Theorem 4.1 we find that μ is invariant under $\phi^L_t(\cdot, \cdot)$. We have

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} L(x, v)\, d\mu_k(x, v) = \frac{1}{t_k} \int_0^{t_k} L(\gamma_k(s), \dot{\gamma}_k(s))\, ds = \frac{1}{t_k} h_{t_k}\left(\gamma_k(0), \gamma_k(t_k)\right)$$

Using the boundedness of $(x, t) \mapsto h_t(x, y)$ uniformly for $t \geqslant 1$ in Proposition 3.17, as $t_k \to \infty$ we deduce that

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

Thus μ is a Mather measure and $\mathrm{supp}\mu \cap V =$ which is a contradiction. □

5.5. The Aubry set.

Definition 16 (The Aubry set $\mathcal{A}_0$)**.** *The Aubry set $\mathcal{A}_0$ is defined by*

$$\mathcal{A}_0 = \{x \in \mathbb{T}^n : h(x, x) = 0\}.$$

Remark 35. It is clear that $\mathcal{A}_0 \neq \emptyset$ as $\emptyset \neq \mathcal{M}_0 \subset \mathcal{A}_0 \subset \mathbb{T}^n$.

We have the following properties (characterization) of the Aubry set $\mathcal{A}_0$.

Proposition 5.5. *The followings are equivalent.*

(i) $x \in \mathcal{A}_0$, *i.e.*, $h(x, x) = 0$.

(ii) *There exists a sequence $\{\gamma_k\}$ of continuous, piece-wise C^1 curves $\gamma_k : [0, t_k] \to \mathbb{T}^n$ with $\gamma_k(0) = \gamma_k(t_k) = x$ and $t_k \to \infty$ such that*

$$\lim_{t_k \to \infty} \left(\int_0^{t_k} L\left(\gamma_k(s), \dot{\gamma}_k(s)\right) + c(0) t_k\right) = 0.$$

(iii) *There exists a sequence $\{\gamma_k\}$ of minimizing extremal curves $\gamma_k : [0, t_k] \to \mathbb{T}^n$ with $\gamma_k(0) = \gamma_k(t_k) = x$ and $t_k \to \infty$ such that*

$$\lim_{t_k \to \infty} \left(\int_0^{t_k} L\left(\gamma_k(s), \dot{\gamma}_k(s)\right) + c(0) t_k\right) = 0.$$

6. Viscosity solutions

6.1. Outline. As usual, our standing assumptions on H are

$$\begin{cases} H \in C^k(\mathbb{T}^n \times \mathbb{R}^n) \text{ for some } k \geqslant 2, \\ \lim_{|p|\to\infty} \left(\inf_{\mathbb{T}^n} \frac{L(x,p)}{|p|} \right) = +\infty, \\ D_p^2 H(x,p) \succ 0 \text{ for all } (x,p) \in \mathbb{T}^n \times \mathbb{R}^n. \end{cases} \tag{H}$$

6.2. Vanishing viscosity process. To find a solution for

$$\begin{cases} u_t(x,t) + H(Du(x,t)) = 0 & (x,t) \in \mathbb{R}^n \times (0,T), \\ u(x,0) = g(x) & (x,t) \in \mathbb{R}^n \times \{0\}, \end{cases}$$

we look at the unique solution u^ε of the second-order problem with small diffusion

$$\begin{cases} u_t^\varepsilon(x,t) + H(Du^\varepsilon(x,t)) = \varepsilon \Delta u^\varepsilon(x,t) & (x,t) \in \mathbb{R}^n \times (0,T), \\ u(x,0) = g(x) & (x,t) \in \mathbb{R}^n \times \{0\} \end{cases}$$

and passing $\varepsilon \to 0$, using maximum principle to select a weak solution. The idea was originally introduced by Fleming, Kruzkov in deriving Euler equation from Navier–Stoke equation, and was done for Hamilton–Jacobi equation by Crandall–Lions and Evans (1960–1980). The main idea is using maximum principle to kick the derivative to test function, resembling L^∞-integration by parts.

6.3. Large time behavior of solutions. Let us consider the equation

$$\begin{cases} u_t(x,t) + H(x, Du(x,t)) = 0 & (x,t) \in \mathbb{R}^n \times (0,T), \\ u(x,0) = g(x) & (x,t) \in \mathbb{R}^n \times \{0\}, \end{cases} \tag{6.1}$$

The optimal control formula reads

$$u(x,t) = \inf \left\{ \int_0^t L(\gamma(s), \dot{\gamma}(s)) ds + g(\gamma(0)) : \gamma \in AC([0,t]; \mathbb{T}^n), \gamma(t) = x \right\}.$$

From the existence of a minimizer, there exists $z \in \mathbb{T}^n$ and a C^k curve $\xi : [0,t] \to \mathbb{T}^n$ with $\xi(0) = z$, $\xi(t) = x$ such that $u(x,t) = \int_0^t L(\xi(s), \dot{\xi}(s)) ds + g(z)$. The fact that $\xi \in C^k$ follows from the Euler-Lagrange equation which ξ solves

$$\frac{d}{ds}\Big(D_v L(\xi(s), \dot{\xi}(s)) \Big) = D_x L(\xi(s), \dot{\xi}(s)), \qquad 0 \leqslant s \leqslant t.$$

We recall that the ergodic problem

$$H(x, Dv(x)) = c(0) \qquad \text{in } \mathbb{T}^n \tag{6.2}$$

has a lot of solutions and if v is such a solution to (6.2) then $u(x,t) = v(x) - c(0)t$ is a solution to (6.1) without the initial condition. We thus hope for $u(x,t) \approx v(x) - c(0)t$ as t large for some v solves (6.2).

Theorem 6.1. *Let u be a viscosity solution to (6.1), then as $t \to \infty$ we have*

$$\lim_{t\to\infty} \Big(u(x,t) + c(0)t \Big) = v(x) \qquad \text{in } \mathbb{T}^n$$

where v is a solution to the ergodic problem (6.2).

Heuristic proof following Fathi. The key ingredients are the followings.

1. There exists a minimizing curve $\xi \in C^k$ with $\xi(t) = x$ such that
$$u(x,t) = \int_0^t L(\xi(s), \dot{\xi}(s)) + g(\xi(0))$$
and we have the conservation of energy along $\xi(\cdot)$ as $s \mapsto H(\xi(s), \dot{\xi}(s))$ is a constant for all $s \in (0,t)$.
2. The Mather set $\mathcal{M}$ is a weak attractor such that $\mathcal{M} \subset \{x \in \mathbb{T}^n : H(x, Du(x)) = c(0)\}$. For $\varepsilon > 0$, let $W_\varepsilon = \left\{x \in \mathbb{T}^n : H(x, Du(x)) \in (c(0)-\varepsilon, c(0)+\varepsilon)\right\}$ is an open set in $\mathbb{T}^n$, then $\mathcal{M} \subset W_\varepsilon$. There exists T_ε such that if $t \geqslant T_\varepsilon$ then there exists $\bar{s} \in [0, T_\varepsilon]$ such that
$$H(\xi(\bar{s}), Du(\xi(\bar{s}))) \in (c(0)-\varepsilon, c(0)+\varepsilon)$$
which implies that
$$H(\xi(s), Du(\xi(s))) \in (c(0)-\varepsilon, c(0)+\varepsilon)$$
for all s by conservation of energy, this holds for all s instead of $\bar{s}$ only, which is remarkable. Therefore using the equation we obtain $\boxed{u_t \approx -c(0) \pm \varepsilon}$ when t is large.
3. Concerning the ergodic problem (6.2), any weak KAM solution of negative type $v \in \mathcal{S}_-$ is a solution to (6.2). For such a weak KAM solution, given any $x \in \mathbb{T}^n$ there exists a calibrated curve $\gamma : (-\infty, 0] \to \mathbb{T}^n$ with $\gamma(0) = x$ so that if $-\infty < s < t \leqslant 0$ then
$$v(\gamma(t)) - v(\gamma(s)) = \int_s^t L(\gamma(s), \dot{\gamma}(s))ds + c(0)(t-s).$$
In other words, $v \prec L + c(0)$ with the exact equality. Using convexity it preserves the $\prec$ property onto any limiting solution u_∞ we may get. This is vague but we will see in the proof.

$\square$

Remark 36. This is a very nice framework but it does not cover every important direction in large time behavior of (6.1) (e.g., H is singular or the problem is set in a non-compact domains: forced mean curvature flow, coagulation-fragmentation, ...).

To make the proof of Theorem 6.1 clearer. We state the following Lemma on the weak attractor of $\mathcal{M}$ and how it relates to solution $u(x,t)$ of (6.1) independently.

Lemma 6.2. *For $\varepsilon > 0$, there exists $T_\varepsilon > 0$ such that for each $t > T_\varepsilon$, if $Du(x,t)$ exists then* $H(x, Du(x,t)) \in \big(c(0)-\varepsilon, c(0)+\varepsilon\big)$.

Proof. It is a simple consequence of the weak attractor property of the Mather set. As we assume $Du(x,t)$ exists, we can find a minimizer (run the Lagrangian flow with the initial data known) curve $\gamma : [0,t] \to \mathbb{T}^n$ with $\gamma \in C^k$, $\gamma(t) = x$ such that
$$u(x,t) = \int_0^t L(\gamma(s), \dot{\gamma}(s))ds + g(\gamma(0)) \qquad \text{and} \qquad Du(x,t) = D_v L(\gamma(t), \dot{\gamma}(t)).$$
We note that the second condition is simply $\dot{\gamma}(t) = D_p H(x, Du(x,t))$.

Recall that $\mathcal{L} : (x, v) \mapsto (x, p)$ maps $\mathcal{L}(x, v) = (x, D_v L(x, v))$ is a local C^{k-1} diffeomorphism (Definition 2) with its inverse $\mathcal{H} = \mathcal{L}^{-1}(x, p) = (x, D_p H(x, p))$. Let us define

$$W_\varepsilon = \Big\{(x, v) \in \mathbb{T}^n \times \mathbb{R}^n : H \circ \mathcal{L}(x, v) \in \big(c(0) - \varepsilon, c(0) + \varepsilon\big)\Big\}$$

then it is an open neighborhood of $\widetilde{\mathcal{M}}_0$, thus by the local attractor property there exists $T_\varepsilon > 0$ such that for any $t > T_\varepsilon$ there exists $\bar{s} \in [0, t]$ such that

$$\big(\gamma(\bar{s}), \dot{\gamma}(\bar{s})\big) \in W_\varepsilon \qquad \Longrightarrow \qquad \mathcal{L}\big(\gamma(\bar{s}), \dot{\gamma}(\bar{s})\big) \in \big(c(0) - \varepsilon, c(0) + \varepsilon\big).$$

By conservation of energy we obtain

$$H \circ \mathcal{L}\big(\gamma(s), \dot{\gamma}(s)\big) \in \big(c(0) - \varepsilon, c(0) + \varepsilon\big).$$

for all $s \in [0, t]$, which implies that $H(x, Du(x, t) \in (c(0) - \varepsilon, c(0) + \varepsilon)$. □

Proof of Theorem 6.1. Without loss of generality, let us assume $c(0) = 0$ by adding a constant to H. Let $v \in C(\mathbb{T}^n)$ solves (6.2) then $\tilde{v}(x, t) = v(x) - c(0)t = v(x)$ solves (6.1). Let C be large enough so that $v(x) - C \leqslant g(x) \leqslant v(x) + C$ for $x \in \mathbb{T}^n$. Run the Hamiltonian flows, i.e., by comparison principle

$$v(x) - C \leqslant u(x, t) \leqslant v(x) + C \qquad \text{for all } (x, t) \in \mathbb{T}^n \times (0, \infty).$$

We also have a priori estimate $\|u_t\|_{L^\infty(\mathbb{T}^n)} + \|Du\|_{L^\infty(\mathbb{T}^n)} \leqslant C$. In the space $C(\mathbb{T}^n)$, the family $\{u(\cdot, t) : t \geqslant 0\}$ is uniformly equi-continuous (following from $\|u_t\|_{L^\infty(\mathbb{T}^n)} \leqslant C$), thus by Arzelà–Ascoli Theorem we can find a subsequence $t_k \to \infty$ and a function $u_\infty \in C(\mathbb{T}^n)$ such that

$$T_{t_k} g(x) = u(x, t_k) \to u_\infty(x) \qquad \text{uniformly as } t_k \to \infty.$$

Here $T_t g(x) = u(x, t)$ is the solution map of (6.1). We now show u_∞ solves (6.2).

1. u_∞ inherits the subsolution property beautifully. Let $\varepsilon > 0$, from Lemma 6.2 we find $T_\varepsilon > 0$ such that for $t_k > T_\varepsilon$

$$H(x, Du(x, t_k)) \leqslant c(0) + \varepsilon = \varepsilon \qquad \text{for a.e. } x \in \mathbb{T}^n.$$

   By convexity (and Jensen's result),

$$H(x, Du(x, t_k)) \leqslant c(0) + \varepsilon = \varepsilon \qquad \text{for } x \in \mathbb{T}^n \text{ in the viscosity sense.}$$

   Let $t_k \to \infty$ and use stability of viscosity solution we obtain that, in the viscosity sense $H(x, Du_\infty(x)) \leqslant \varepsilon$ in $\mathbb{T}^n$ and thus u_∞ is a subsolution to (6.2) by sending $\varepsilon \to 0$.

2. The supersolution property is trickier. Using the fact that $u_\infty(\cdot)$ is a subsolution to (6.2), we have $T_t u_\infty(\cdot)$ is a subsolution to

$$\begin{cases} w_t(x, t) + H(x, Dw(x, t)) = 0 & (x, t) \in \mathbb{R}^n \times (0, T), \\ w(x, 0) = u_\infty(x) & (x, t) \in \mathbb{R}^n \times \{0\}, \end{cases} \tag{6.3}$$

   while $\tilde{u}(x, t) = u_\infty(x)$ is a viscosity solution, hence by comparison principle

$$T_t u_\infty(\cdot) \leqslant u_\infty(\cdot) \qquad \Longrightarrow \qquad T_{s+t} u_\infty(\cdot) \leqslant T_s u_\infty(\cdot)$$

   and thus $s \mapsto T_s u_\infty(\cdot)$ is non-increasing for $s \geqslant 0$. We claim that

$$T_t u_\infty(\cdot) \equiv u_\infty(\cdot) \qquad \text{for all } t \geqslant 0. \tag{6.4}$$

Assume that $s_k = t_{k+1} - t_k \to \infty$ we show $T_{s_k} u_\infty(\cdot) \to u_\infty(\cdot)$ uniformly as $s_k \to \infty$. The ingredients are $T_{t_k} g(\cdot) \to u_\infty(\cdot)$ and the contraction property

$$\|T_t g_1(\cdot) - T_t g_2(\cdot)\|_{L^\infty(\mathbb{T}^n)} \leqslant \|g_1 - g_2\|_{L^\infty(\mathbb{T}^n)}.$$

We have $T_{t_{k+1}} = T_{s_k} \circ T_{t_k}$, thus

$$\|T_{s_k} u_\infty - u_\infty\|_{L^\infty} \leqslant \underbrace{\|T_{s_k} u_\infty - T_{t_{k+1}} g\|_{L^\infty}}_{\|u_\infty - T_{t_k} g\|_{L^\infty}} + \|T_{t_{k+1}} g - u_\infty\|_{L^\infty} \to 0$$

as $t_k \to \infty$. Together with the fact that $s \mapsto T_s u_\infty$ is non-decreasing, we have $T_s u_\infty(\cdot) \to u_\infty(\cdot)$ uniformly as $s \to \infty$ and further that (6.4) holds, hence u_∞ is a solution to the ergodic problem (6.2).

Finally, write $t = s + t_k$ we have

$$\|T_t g - u_\infty(x)\|_{L^\infty} \leqslant \underbrace{\|T_{s+t_k} g - T_s u_\infty(x)\|_{L^\infty}}_{\|T_{t_k} g - u_\infty\|_{L^\infty}} + \|T_s u_\infty - u_\infty\|_{L^\infty} \to 0$$

as $t \to \infty$. □

Remark 37. Some open questions:

1. Can we quantify T_ε in the attractor property?
2. Rate of convergence of $u(x,t) \to u_\infty(x) - c(0)t$?

Some other proofs are available ([8, 19], ...).

Literature

Beside the one we cited earlier, the materials follow also some of the following sources [1, 2, 3, 4, 6, 7, 9, 10, 11, 12, 14, 15, 16]. The rate of convergence for homogenization using the tools developed from rotation vector is studied in [17]. The author will update more references in the future.

References

[1] Aubry, S. The twist map, the extended Frenkel-Kontorova model and the devil's staircase. vol. 7. 1983, pp. 240–258. Order in chaos (Los Alamos, N.M., 1982).

[2] Bardi, M., and Capuzzo-Dolcetta, I. *Optimal Control and Viscosity Solutions of Hamilton–Jacobi–Bellman Equations*. Modern Birkhäuser Classics. Birkhäuser Basel, 1997.

[3] Biryuk, A., and Gomes, D. A. An introduction to the Aubry-Mather theory. *São Paulo Journal of Mathematical Sciences 4*, 1 (Apr. 2010), 17–63.

[4] E, W. Aubry-Mather theory and periodic solutions of the forced Burgers equation. *Comm. Pure Appl. Math. 52*, 7 (1999), 811–828.

[5] Evans, L. C. *Partial differential equations*. American Mathematical Society, Providence, R.I., 2010.

[6] Evans, L. C., and Gomes, D. Effective hamiltonians and averaging for hamiltonian dynamics i. *I. Arch. Ration. Mech. Anal 157* (2001), 200–1.

[7] Fathi, A. Weak kam theorem in lagrangian dynamics. *to appear in Cambridge Studies in Advanced Mathematics.*.

[8] Fathi, A. Sur la convergence du semi-groupe de Lax-Oleinik. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics 327*, 3 (Aug. 1998), 267–270.

[9] Fathi, A., and Siconolfi, A. PDE aspects of Aubry-Mather theory for quasiconvex Hamiltonians. *Calculus of Variations and Partial Differential Equations 22*, 2 (Feb. 2005), 185–228.

[10] Ishii, H. *Lecture notes on the weak KAM theorem*. Lecture notes available on Ishii's website.

[11] Ishii, H., Mitake, H., and Tran, H. V. The vanishing discount problem and viscosity Mather measures. Part 1: The problem on a torus. *Journal de Mathématiques Pures et Appliquées 108*, 2 (2017), 125–149.

[12] Ishii, H., Mitake, H., and Tran, H. V. The vanishing discount problem and viscosity Mather measures. Part 2: Boundary value problems. *Journal de Mathématiques Pures et Appliquées 108*, 3 (2017), 261 – 305.

[13] Kim, Y., Tran, H. V., and Tu, S. N. State-Constraint Static Hamilton–Jacobi Equations in Nested Domains. *SIAM Journal on Mathematical Analysis 52*, 5 (Jan. 2020), 4161–4184. Publisher: Society for Industrial and Applied Mathematics.

[14] Le, N. Q., Mitake, H., and Tran, H. V. *Dynamical and geometric aspects of Hamilton-Jacobi and linearized Monge-Ampère equations—VIASM 2016*, vol. 2183 of *Lecture Notes in Mathematics*. Springer, Cham, 2017. Edited by Mitake and Tran.

[15] Mañé, R. Generic properties and problems of minimizing measures of Lagrangian systems. *Nonlinearity 9*, 2 (1996), 273–310.

[16] Mather, J. N. Action minimizing invariant measures for positive definite lagrangian systems. *Mathematische Zeitschrift 207*, 2 (1991), 169–208.

[17] Mitake, H., Tran, H. V., and Yu, Y. Rate of convergence in periodic homogenization of Hamilton-Jacobi equations: the convex setting. *Arch. Ration. Mech. Anal. 233*, 2 (2019), 901–934.

[18] Nakayasu, A. Two approaches to minimax formula of the additive eigenvalue for quasiconvex Hamiltonians. *ArXiv e-prints* (Dec. 2014).

[19] Namah, G., and Roquejoffre, J.-M. Remarks on the long time behaviour of the solutions of hamilton-jacobi equations. *Communications in Partial Differential Equations 24*, 5-6 (Jan. 1999), 883–893. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/03605309908821451.

[20] Tran, H. *Hamilton-Jacobi Equations: Theory and Applications*. Graduate studies in mathematics. American Mathematical Society, 2021.

Department of Mathematics, University of Wisconsin Madison, 480 Lincoln Drive, Madison, WI 53706, USA