Weak KAM Theorem in Lagrangian Dynamics Preliminary Version Number 10

Albert FATHI

Lyon, Version 15 June 2008

Contents

Preface

The project of this book started from my work published in the Comptes Rendus de l'Académie des Sciences, see [Fat97b, Fat97a, Fat98a, Fat98b].

I gave several courses and lectures on the material presented there.

The project went through several versions. The first version was in French. It was produced for the Graduate course "Systèmes lagrangiens et théorie d'Aubry-Mather", that I gave at the Ecole Normale Supérieure de Lyon during Spring Semester 1998. The French set of notes has circulated widely. Daniel Massart and Ezequiel Maderna caught up a large amount of mistakes in the French version. The first set of notes in english were a translated and improved version of lectures notes in French, and consited of versions of chapter 1 to 5. It was done while I was on sabbatical during Spring Semester 2000 at the University of Geneva. I wish to thank the Swiss National Scientific foundation for support during that time. This first version was distributed and used at the "Ecole d'été en géométrie" held at "Université de Savoie" June, 15-22, 2000. A certain number of typing mistakes were found by people attending the "Ecole d'été en géométrie"

After adding chapter 6, we incorporated some of the improvements suggested by Alain Chenciner and Richard Montgomery.

The subsequent versions, besides improvements, contained a couple of chapters on viscosity solutions of the Hamilton-Jacobi equation, especially the connection with the weak KAM theorem, and a last brief one making the connection with Mañé's point of view. The opportunity to teach a course of DEA in Lyon in 2001-2002 and 2002-2003 was instrumental in the expansions in this set of notes.

The seventh version was done in Pisa. I had the privilige of giving a seires of Lectures in Winter 2005 in the Centro di Giorgi at the Scuola Normale Superiore in Pisa. This seventh version is a major revision of the sixth.

In this version 8, we have incorporated several typing mistakes picked up by Maxime Zavidovique.

The present tenth version is prepared for a course given at the Summer School "Dynamical Systems: Theoretical and Applied Hamiltonian Dynamics" held at t Instituto Superior Técnico in Lisbon 16-20 June 2008. It has gone through a major revision of chapter 4. I have incorporated a new proof found in June 2008 of the Weak KAM Theorem that is more elementary than the previous ones in that it only uses the order properties for the Lax-Oleinik semi-group and some compacrness arguments. It avoids any appeal to a fixed point theorem. We hope that the simple arguments may be used in other contexts. We kept as a second proof the one using a fixed point theorem, since we consider it as much more natural and almost forced on us by the compactness obtained from Fleming's Lemma.

A lot of people have helped me for a better understanding of the subject, it is impossible to mention them all, among the ones that I can remember vividly in chronological order: John Mather, Michel Herman, Nicole Desolneux, Daniel Massart, Denis Serre (without whom, I would have never realized that there was a deep connection with viscosity solutions), Jean-Christophe Yoccoz, Francis Clarke, Gabriel & Miguel Paternain, Gonzalo Contreras, Renato Itturiaga, Guy Barles, Jean-Michel Roquejoffre, Ezequiel Maderna, Patrick Bernard, Italo Capuzzo-Dolcetta, Piermarco Cannarsa, Craig Evans. Special thanks to Alain Chenciner for his drive to understand and improve this subject. Last but not least Antonio Siconolfi, we have been enjoying now a long a solid collaboration, a large number of the improvements in these set of notes is due to the numerous conversation that we have specialy on the viscosity theory aspects.

Starting with the French notes, Claire Desecures helped a lot in the typing.

Lyon, 14 June 2008

Introduction

The object of this course is the study of the Dynamical System defined by a convex Lagrangian. Let M be a compact C^∞ manifold without boundary. We denote by TM the tangent bundle and by $\pi : TM \to M$ the canonical projection. A point of TM will be denoted by (x, v) with $x \in M$ and $v \in T_xM = \pi^{-1}(x)$. In the same way, a point of the cotangent bundle T^*M will be denoted by (x, p) with $x \in M$ and $p \in T_x^*M$ a linear form on the vector space T_xM.

We consider a function $L : TM \to \mathbb{R}$ of class at least C^3. We will call L the Lagrangian. As a typical case of L, we can think of $L(x, v) = \frac{1}{2}g_x(v, v)$ where g is a Riemannian metric on M. There is also the case of more general mechanical systems $L(x, v) = \frac{1}{2}g_x(v, v) - V(x)$, with g a Riemannian metric on M and $V : M \to \mathbb{R}$ a function.

The action functional $\mathbb{L}$ is defined on the set of continuous piecewise C^1 curves $\gamma : [a, b] \to M, a \leq b$ by

$$\mathbb{L}(\gamma) = \int_a^b L(\gamma(s), \dot{\gamma}(s))ds$$

We look for C^1 (or even continuous piecewise C^1) curves $\gamma : [a, b] \to M$ which minimize the action $\mathbb{L}(\gamma) = \int_a^b L(\gamma(s), \dot{\gamma}(s))ds$ among the C^1 curves (or continuous piecewise C^1) $\tilde{\gamma} : [a, b] \to M$ with the ends $\tilde{\gamma}(a)$ and $\tilde{\gamma}(b)$ fixed. We will also look for curves $\tilde{\gamma}$ which minimize the action among the curves homotopic to γ with fixed endpoints or even for curves which achieve a local minimum of the action among all curves homotopic with same endpoints.

The problem is tackled using differential calculus on a functional space. We first look for the critical points of the action

$\gamma \to \mathbb{L}(\gamma)$ on the space of curves

$\mathcal{C}^1_{x,y}([a,b],M) = \{\gamma : [a,b] \to M \mid \gamma \text{ of class } \mathrm{C}^1 \text{ and } \gamma(a) = x, \gamma(b) = y\}.$

Such a curve which is a critical point is called an extremal curve for the Lagrangian L. If an extremal curve γ is C^2, it is possible to show that the curve γ satisfies the Euler-Lagrange equation which, in a system of coordinates, is written as

$$\frac{\partial L}{\partial x}(\gamma(t),\dot\gamma(t)) - \frac{d}{dt}(\frac{\partial L}{\partial v}(\gamma(t),\dot\gamma(t)) = 0.$$

If the second partial vertical derivative $\frac{\partial^2 L}{\partial v^2}(x,v)$ is non-degenerate at each point of TM we then see that we can solve for $\ddot\gamma(t)$. It results that there is a vector field

$$(x,v) \mapsto X_L(x,v)$$

on TM such that the speed curves $t \mapsto (\gamma(t),\dot\gamma(t))$ of extremal curves γ for the Lagrangian are precisely the solutions of this vector field X_L. The (local) flow $\phi_s : TM \to TM$ of this vector field X_L is called the Euler-Lagrange flow of the Lagrangian L. By definition, a curve $\gamma : [a,b] \to M$ is an extremal curve if and only if $(\gamma(s),\dot\gamma(s)) = \phi_{s-a}(\gamma(a),\dot\gamma(a))$, for all $s \in [a,b]$.

As TM is not compact, it may happen that ϕ_s is not defined for all $s \in \mathbb{R}$, which would prevent us from making dynamics. It will be supposed that L verifies the two following conditions

(1) with x fixed $v \mapsto L(x,v)$ is C^2-strictly convex, i.e. the second partial vertical derivative $\frac{\partial^2 L}{\partial v^2}(x,v)$ is defined strictly positive, as a quadratic form;

(2) $L(x,v)$ is superlinear in v, i.e.

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

where $\|\cdot\|$ is a norm coming from a Riemannian metric on M.

Since all the Riemannian metrics are equivalent on a compact manifold, this condition (2) does not depend on the choice of the Riemannian metric.

Condition (2) implies that the continuous function $L : TM \to \mathbb{R}$ is proper, i.e. inverse images under L of compact sets are compact.

Conditions (1) and (2) are of course satisfied for the examples given above.

The function $H(x,v) = \frac{\partial L}{\partial v}(x,v)v - L(x,v)$ is called the Hamiltonian of the system. It is invariant by ϕ_s. Under the assumptions (1) and (2), this function $H : TM \to \mathbb{R}$ is also proper (in fact superlinear). The levels $H^{-1}(c), c \in \mathbb{R}$ are thus compact subsets of TM. As each trajectory of ϕ_s remains in such compact set, we conclude from it that ϕ_s is defined for all $s \in \mathbb{R}$, as soon as L satisfies conditions (1) and (2). We can, then, study the Euler-Lagrange flow using the theory of Dynamical Systems.

0.1 The Hamilton-Jacobi Method

A natural problem in dynamics is the search for subsets invariant by the flow ϕ_s. Within the framework which concerns us the Hamilton-Jacobi method makes it possible to find such invariant subsets.

To explain this method, it is better to think of the Hamiltonian H as a function on cotangent bundle T^*M. Indeed, under the assumptions (1) and (2) above, we see that the Legendre transform $\mathcal{L} : TM \to T^*M$, defined by

$$\mathcal{L}(x,v) = (x, \frac{\partial L}{\partial v}(x,v)),$$

is a diffeomorphism of TM onto T^*M. We can then regard H as a function on T^*M defined by

$$H(x,p) = p(v) - L(x,v), \text{ where } p = \frac{\partial L}{\partial v}(x,v).$$

As the Legendre transform $\mathcal{L}$ is a diffeomorphism, we can use it to transport the flow $\phi_t : TM \to TM$ to a flow $\phi_t^* : T^*M \to T^*M$ defined by $\phi_t^* = \mathcal{L}\phi_t\mathcal{L}^{-1}$.

Theorem 0.1.1 (Hamilton-Jacobi). *Let ω be a closed 1-form on M. If H is constant on the graph* $\mathrm{Graph}(\omega) = \{(x, \omega_n) \mid x \in M\}$, *then this graph is invariant by ϕ_t^*.*

We can then ask the following question:

Given a fixed closed 1-form ω_0, does there exist ω another closed 1-form cohomologous with ω_0 such that H is constant on the graph of ω?

The answer is in general negative if we require ω to be continuous. It is sometimes positive, this can be a way to formulate the Kolmogorov-Arnold-Moser theorem, see [Bos86].

However, there are always solutions in a generalized sense. In order to explain this phenomenon, we will first show how to reduce the problem to the 0 cohomology class. If ω_0 is a fixed closed 1-form, let us consider the Lagrangian $L_{\omega_0} = L - \omega_0$, defined by

$$L_{\omega_0}(x, v) = L(x, v) - \omega_{0,x}(v).$$

Since ω_0 is closed, if we consider only curves γ with the same fixed endpoints, the map $\gamma \mapsto \int_\gamma \omega_0$ is locally constant. It follows that L_{ω_0} and L have the same extremal curves. Thus they have also the same Euler-Lagrange flow. The Hamiltonian H_{ω_0} associated with L_{ω_0} verifies

$$H_{\omega_0}(x, p) = H(x, \omega_{0,x} + p).$$

By changing the Lagrangian in this way we see that we have only to consider the case $\omega_0 = 0$.

We can then try to solve the following problem:

Does there exist a constant $c \in \mathbb{R}$ and a differentiable function $u : M \to \mathbb{R}$ such that $H(x, d_x u) = c$, for all $x \in M$?

There is an "integrated" version of this question using the semigroup $T_t^- : \mathcal{C}^0(M, \mathbb{R}) \to \mathcal{C}^0(M, \mathbb{R})$, defined for $t \geq 0$ by

$$T_t^- u(x) = \inf\{\mathbb{L}(\gamma) + u(\gamma(0)) \mid \gamma : [0, t] \to M, \gamma(t) = x\}.$$

It can be checked that $T_{t+t'}^- = T_t^- \circ T_{t'}^-$, and thus T_t^- is a (nonlinear) semigroup on $\mathcal{C}^0(M, \mathbb{R})$.

A C^1 function $u : M \to \mathbb{R}$, and a constant $c \in \mathbb{R}$ satisfy $H(x, d_x u) = c$, for all $x \in M$, if and only if $T_t^- u = u - ct$, for each $t \geq 0$.

Theorem 0.1.2 (Weak KAM)**.** *We can always find a Lipschitz function $u : M \to \mathbb{R}$ and a constant $c \in \mathbb{R}$ such that $T_t^- u = u - ct$, for all $t \geq 0$.*

The case $M = \mathbb{T}^n$, in a slightly different form (viscosity solutions) is due to P.L. Lions, G. Papanicolaou and S.R.S. Varadharan 87, see [LPV87, Theorem 1, page 6]. This general version was obtained by the author in 96, see [Fat97b, Théorème1, page 1044]. Carlsson, Haurie and Leizarowitz also obtained a version of this theorem in 1992, see [CHL91, Theorem 5.9, page 115].

As u is a Lipschitz function, it is differentiable almost everywhere by Rademacher's Theorem. It can be shown that $H(x, d_x u) = c$ at each point where u is differentiable. Moreover, for such a function u we can find, for each $x \in M$, a C^1 curve $\gamma_x :]-\infty, 0] \to M$, with $\gamma_x(0) = x$, which is a solution of the multivalued vector field "$\mathrm{grad}_L\, u$"(x) defined on M by

$$\text{“}\mathrm{grad}_L\, u\text{”}(y) = \mathcal{L}^{-1}(y, d_y u).$$

These trajectories of $\mathrm{grad}_L\, u$ are minimizing extremal curves. The accumulation points of their speed curves in TM for $t \to -\infty$ define a compact subset of TM invariant under the Euler-Lagrange flow φ_t. This is an instance of the so-called Aubry and Mather sets found for twist maps independently by Aubry and Mather in 1982 and in this full generality by Mather in 1988.

We can of course vary the cohomology class replacing L by L_ω and thus obtain other extremal curves whose speed curves define compact sets in TM invariant under ϕ_t. The study of these extremal curves is important for the understanding of this type of Lagrangian Dynamical Systems.

Chapter 1

Convex Functions: Legendre and Fenchel

Besides some generalities in the first two sections, our main goal in this chapter is to look at the Legendre and Fenchel transforms. This is now standard material in Convex Analysis, Optimization, and Calculus of Variations. We have departed from the usual viewpoint in Convex Analysis by not allowing our convex functions to take the value $+\infty$. We think that this helps to grasp things on a first introduction; moreover, in our applications all functions have finite values. In the same way, we have not considered lower semi-continuous functions, since we will be mainly working with convex functions on finite dimensional spaces.

We will suppose known the theory of convex functions of one real variable, see for example [RV73, Chapter 1]or [Bou76, Chapitre 1].

1.1 Convex Functions: General Facts

Definition 1.1.1 (Convex Function). Let U be a convex set in the vector space E. A function $f : U \to \mathbb{R}$ is said to be convex if it satisfies the following condition

$$\forall x, y \in U, \forall t \in [0,1], f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

The function f is said to be strictly convex if it satisfies the following stronger condition

$$\forall x \neq y \in U, \forall t \in]0,1[, f(tx+(1-t)y) < tf(x)+(1-t)f(y).$$

It results from the definition that $f : U \to \mathbb{R}$ is convex if and only if for every line $D \subset E$ the restriction of f on $D \cap U$ is a convex function (of one real variable).

Proposition 1.1.2. (i) *An affine function is convex (an affine function is a function which is the sum of a linear function and a constant).*

(ii) *If $(f_i)_i \in I$ is a family of convex functions $: U \to \mathbb{R}$, and $\sup_{i\in I} f_i(x) < +\infty$ for each $x \in U$ then $\sup_{i\in I} f_i$ is convex.*

(iii) *Let U be an open convex subset of the normed space E. If $f : U \to \mathbb{R}$ is convex and twice differentiable at $x \in U$, then $D^2 f(x)$ is non-negative definite as a quadratic form. Conversely, if $g : U \to \mathbb{R}$ admits a second derivative $D^2 g(x)$ at every point $x \in U$, with $D^2 g(x)$ non-negative (resp. positive) definite as a quadratic form, then g is (resp. strictly) convex.*

Properties (i) and (ii) are immediate from the definitions. The property (iii) results from the case of the functions of a real variable by looking at the restrictions to each line of E.

Definition 1.1.3 (C^2-Strictly Convex Function)**.** Let be a in the vector space E. A function $f : U \to \mathbb{R}$, defined on the convex subset U of the normed vector space E, is said to be C^2-strictly convex if it is C^2, and its the second derivative $D^2 f(x)$ is positive definite as a quadratic form, for each $x \in U$.

Exercise 1.1.4. *Let U be an open convex subset of the normed space E, and let $f : U \to \mathbb{R}$ be a convex function.*

a) Show that f is not strictly convex if and only if there exists a pair of distinct points $x, y \in U$ such that f is affine on the segment $[x, y] = \{tx + (1-t)y \mid t \in [0,1]$.

b) If f is twice differentiable at every $x \in U$, show that it is strictly convex if and only if for every unit vector $v \in E$ the set $\{x \in U \mid D^2 f(x)(v,v) = 0\}$ does not contain a non trivial segment parallel to v. In particular, if $D^2 f(x)$ is non-negative definite as a quadratic form at every point $x \in U$, and the set of *points x where $D^2f(x)$ is not of maximal rank does not contain a non-trivial segment then f is strictly convex.*

Theorem 1.1.5. *Suppose that U is an open convex subset of the topological vector space E. Let $f : U \to \mathbb{R}$ be a convex function. If there exists an open non-empty subset $V \subset U$ with $\sup_{x\in V} f(x) < +\infty$, then f is continuous on U.*

Proof. Let us first show that for all $x \in U$, there exists an open neighborhood V_x of x, with $V_x \subset U$ and $\sup_{y\in V_x} f(y) < +\infty$. Indeed, if $x \notin V$, we choose $z_0 \in V$. The intersection of the open set U and the line containing x and z_0 is an open segment containing the compact segment $[x, z_0]$. We choose in this intersection, a point y near to x and such that $y \notin [x, z_0]$, thus $x \in]y, z_0[$, see figure 1.1. It follows that there exists t with $0 < t_0 < 1$ and such that $x = t_0y + (1-t_0)z_0$. The map $H : E \to E, z \mapsto x = t_0y + (1-t_0)z$ sends z_0 to x, is a homeomorphism of E, and, by convexity of U, it maps U into itself. The image of V by H is an open neighborhood V_x of x contained in U. Observe now that any point x' of V_x can be written as the form $x' = t_0y + (1-t_0)z$ with $z \in V$ for the same $t_0 \in]0, 1[$ as above, thus

$$\begin{aligned} f(x') &= f(t_0y + (1-t_0)z) \\ &\le t_0f(y) + (1-t_0)f(z) \\ &\le t_0f(y) + (1-t_0)\sup_{z\in V} f(z) < +\infty. \end{aligned}$$

This proves that f is bounded above on V_x.

Let us now show that f is continuous at $x \in U$. We can suppose by translation that $x = 0$. Let V_0 be an open subset of U containing 0 and such that $\sup_{y\in V_0} f(y) = M < +\infty$. Since E is a topological vector space, we can find an open set $\tilde{V}_0$ containing 0, and such that $t\tilde{V}_0 \subset V_0$, for all $t \in \mathbb{R}$ with $|t| \le 1$. Let us suppose that $y \in \epsilon\tilde{V}_0 \cap (-\epsilon\tilde{V}_0)$, with $\epsilon \le 1$. We can write $y = \epsilon z_+$ and $y = -\epsilon z_-$, with $z_+, z_- \in \tilde{V}_0$ (of course $z_- = -z_+$, but this is irrelevant in our argument). As $y = (1-\epsilon)0 + \epsilon z_+$, we obtain $f(y) \le (1-\epsilon)f(0) + \epsilon f(z_+)$, hence

$$\forall y \in \epsilon\tilde{V}_0 \cap (-\epsilon\tilde{V}_0),\ f(y) - f(0) \le \epsilon(M - f(0)).$$

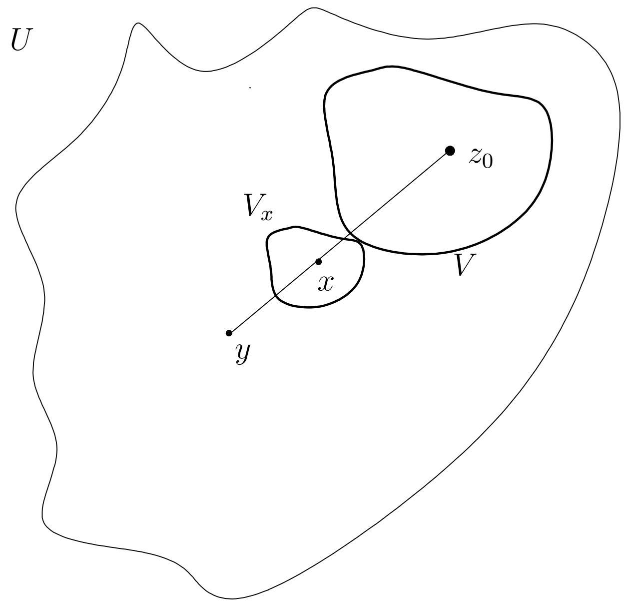

Figure 1.1:

We can also write $0 = \frac{1}{1+\epsilon}y + \frac{\epsilon}{1+\epsilon}z_-$, hence $f(0) \le \frac{1}{1+\epsilon}f(y) + \frac{\epsilon}{1+\epsilon}f(z_-)$ which gives $(1+\epsilon)f(0) \le f(y) + \epsilon f(z_-) \le f(y) + \epsilon M$. Consequently

$$\forall y \in \epsilon \tilde{V_0} \cap (-\epsilon \tilde{V_0}),\ f(y) - f(0) \ge -\epsilon M + \epsilon f(0).$$

Gathering the two inequalities we obtain

$$\forall y \in \epsilon \tilde{V_0} \cap (-\epsilon \tilde{V_0}),\ |f(y) - f(0)| \le \epsilon(M - f(0)). \qquad \square$$

Corollary 1.1.6. *A convex function $f : U \to \mathbb{R}$ defined on an open convex subset U of $\mathbb{R}^n$ is continuous.*

Proof. Let us consider $n+1$ affinely independent points $x_0, \cdots, x_n \in U$. The convex hull σ of $x_0, \cdots, x_n$ has a non-empty interior. By convexity, the map f is bounded by $\max_{i=0}^n f(x_i)$ on σ. $\square$

Most books treating convex functions from the point of view of Convex Analysis do emphasize the role of lower semi-continuous convex functions. When dealing with finite valued functions, the following exercise shows that this is not really necessary.

Exercise 1.1.7. *Let U be an open subset of the Banach space E. If $f : U \to \mathbb{R}$ is convex, and lower semi-continuous show that it is in fact continuous. [Indication: Consider the sequence of subsets $C_n = \{x \in U \mid f(x) \leq n\}, n \in \mathbb{N}$. Show that one of these subsets has non-empty interior.]*

We recall that a function $f : X \to Y$, between the metric spaces X, Y, is said to be locally Lipschitz if, for each $x \in X$, there exists a neighborhood V_x of x in X on which the restriction $f_{|V_x}$ is Lipschitz.

Theorem 1.1.8. *Let E be a normed space and $U \subset E$ an open convex subset. Any convex continuous function $f : U \to \mathbb{R}$ is a locally Lipschitz function.*

Proof. In fact, this follows from the end of the proof of Theorem 1.1.5. We now give a direct slightly modified proof.

We fix $x \in U$. Since f is continuous, there exists $r \in]0, +\infty[$ and $M < +\infty$ such that

$$\sup_{y \in \bar{B}(x,r)} |f(y)| \leq M.$$

We have used the usual notation $\bar{B}(x, r)$ to mean the closed ball of center x and radius r.

Let us fix $y, y' \in \bar{B}(x, r/2)$. We call z the intersection point of the boundary $\partial B(x, r) = \{x' \in E \mid \|x' - x\| = r\}$ of the closed ball $\bar{B}(x, r)$ with the line connecting y and y' such that y is in the segment $[z, y']$, see figure 1.2. We of course have $\|z - y'\| \geq r/2$. We write $y = tz + (1 - t)y'$, with $t \in [0, 1[$, from which it follows that $y - y' = t(z - y')$. By taking the norms and by using $\|z - y'\| \geq r/2$, we see that

$$t \leq \|y' - y\| \frac{2}{r}.$$

The convexity of f gives us $f(y) \leq tf(z) + (1 - t)f(y')$, from which we obtain the inequality $f(y) - f(y') \leq t(f(z) - f(y'))$. It results that

$$f(y) - f(y') \leq 2tM \leq \frac{4M}{r} \|y - y'\|,$$

and by symmetry

$$\forall y, y' \in \bar{B}(x, r/2),\ |f(y) - f(y')| \leq \frac{4M}{r} \|y - y'\|. \qquad \square$$

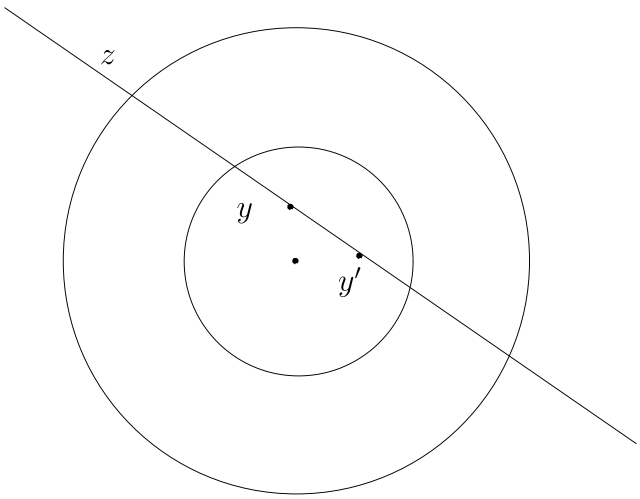

Figure 1.2:

Corollary 1.1.9. *If $f : U \to \mathbb{R}$ is convex with $U \subset \mathbb{R}^n$ open and convex, then f is a locally Lipschitz function.*

We recall Rademacher's Theorem, see [EG92, Theorem 2, page 81]or [Smi83, Theorem 5.1, page 388].

Theorem 1.1.10 (Rademacher)**.** *A locally Lipschitz function defined on open subset of $\mathbb{R}^n$ and with values in $\mathbb{R}^m$ is Lebesgue almost everywhere differentiable.*

Corollary 1.1.11. *A convex function $f : U \to \mathbb{R}$, where U is open convex of $\mathbb{R}^n$, is Lebesgue almost everywhere differentiable.*

It is possible to give a proof of this Corollary not relying on Rademacher's Theorem, see [RV73, Theorem D, page 116]. We conclude this section with a very useful lemma.

Lemma 1.1.12. *Let $f : V \to \mathbb{R}$ be a convex function defined on an open subset V of a topological vector space.*

(a) *A local minimum for f is a global minimum.*

(b) *If f is strictly convex, then f admits at most one minimum.*

Proof. (a) Let x_0 be a local minimum. For $y \in V$ and $t \in [0,1]$ and close to 1 we have

$$f(x_0) \leq f(tx_0 + (1-t)y) \leq tf(x_0) + (1-t)f(y),$$

thus $(1-t)f(x_0) \leq (1-t)f(y)$ for t close to 1. It follows that $f(y) \geq f(x_0)$.

(b) It results from the convexity of f that the subset $\{x \mid f(x) \leq \lambda\}$ is convex. If $\lambda = \inf f$, we have $\{x \mid f(x) = \inf f\} = \{x \mid f(x) \leq \inf f\}$. If f is strictly convex this convex set cannot contain more than one point. □

1.2 Linear Supporting Form and Derivative

As is usual, if E as a vector space (over $\mathbb{R}$) we will denote by $E^* = \mathrm{Hom}(E, \mathbb{R})$ its algebraic dual space. We will indifferently use both notations $p(v)$ or $\langle p, v \rangle$ to denote the value of $v \in E$ under the linear form $p \in E^*$.

Definition 1.2.1 (Supporting Linear Form)**.** We say that the linear form $p \in E^*$ is a supporting linear form at $x_0 \in U$ for the function $f : U \to \mathbb{R}$, defined on $U \subset E$, if we have

$$\forall x \in U,\ f(x) - f(x_0) \geq p(x - x_0) = \langle p, x - x_0 \rangle.$$

We will denote by $\mathrm{SLF}_x(f)$ the set of supporting linear form at x for f, and by $\mathrm{SLF}(f)$ the graph

$$\mathrm{SLF}(f) = \cup_{x \in U} \{x\} \times \mathrm{SLF}_x(f) \subset U \times E^*.$$

In the literature, the linear form p is also called subderivative of f at x_0 or even sometimes subgradient. We prefer to call it supporting linear form to avoid confusion with the notion of subdifferential that we will introduce in another chapter.

Example 1.2.2. a) If $f : \mathbb{R} \to \mathbb{R}, t \mapsto |t|$ then $\mathrm{SLF}_0(f) = [-1, 1]$, for $t > 0, \mathrm{SLF}_t(f) = \{1\}$, and for $t < 0, \mathrm{SLF}_t(f) = \{-1\}$.

b) If $g : \mathbb{R} \to \mathbb{R}, t \mapsto t^3$ then $\mathrm{SLF}_t(g) = \emptyset$, for every $t \in \mathbb{R}$.

The following Proposition is obvious.

Proposition 1.2.3. *The set* $\mathrm{SLF}_x(f)$ *is a convex subset of* E^*. *Moreover, if we endow* E^* *with the topology of simple convergence on* E *("weak topology") then* $\mathrm{SLF}_x(f)$ *is also closed.*

Here is the relation between supporting linear form and derivative.

Proposition 1.2.4. *Let* $f : U \to \mathbb{R}$ *be a function defined on an open subset* U *of the normed space* E.

a) If f *is differentiable at some given* $x \in U$ *then* $\mathrm{SLF}_x(f) \subset \{Df(x)\}$, *i.e. it is either empty or equal to the singleton* $\{Df(x)\}$.

b) If $E = \mathbb{R}^n$, *and all partial derivatives* $\partial f/\partial x_i(x), i = 1, \dots, n$, *exist at some given* $x \in U$, *then* $\mathrm{SLF}_x(f)$ *is either empty or reduced to the single linear form* $(a_1, \dots, a_n) \mapsto \sum_{i=1}^n a_i \partial f/\partial x_i(x)$.

Proof. a) If $\mathrm{SLF}_x(f) \neq \emptyset$, let p be a supporting linear form of f at x. If $v \in E$ is fixed, for all $\epsilon > 0$ small we have $x + \epsilon v \in U$ and thus $f(x + \epsilon v) - f(x) \geq \epsilon p(v)$. Dividing by ϵ and taking the limit as ϵ goes to 0 in this last inequality, we find $Df(x)(v) \geq p(v)$. For linear forms this implies equality, because a linear form which is ≥ 0 everywhere has to be 0.

b) We denote by $(e_1, \dots, e_n)$ the canonical base in $\mathbb{R}^n$. Let us consider a point $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ where all partial derivatives exist. This implies that the function of one variable $h \mapsto f(x_1, \dots, x_{i-1}, h, x_{i+1}, \dots, x_n)$ is differentiable at x_i, hence by part a), if $p \in \mathrm{SLF}_x(f)$, we have $p(e_i) = \partial f/\partial x_i(x)$. Since this is true for every $i = 1, \dots, n$, therefore the map p must be $(a_1, \dots, a_n) \mapsto \sum_{i=1}^n a_i \partial f/\partial x_i(x)$. □

We have not imposed any continuity in the definition of a supporting linear form for a function f. This is indeed the case under very mild conditions on f, as we will see presently.

Proposition 1.2.5. *Let* U *be an open subset of the topological vector space* E, *and let* $f : U \to \mathbb{R}$ *be a function. Suppose that* f *is bounded from above on a neighborhood of* $x_0 \in U$, *then any supporting linear form of* f *at* x_0 *is continuous.*

Proof. Let V be a neighborhood of 0 such that $V = -V$, and f is defined and bounded from above by $K < +\infty$ on $x_0 + V$. Since V

is symmetrical around 0, for each $v \in V$, we have

$$p(v) \leq f(x_0 + v) - f(x_0) \leq 2K$$
$$-p(v) = p(-v) \leq f(x_0 - v) - f(x_0) \leq 2K,$$

hence the linear form p is thus bounded on a nonempty open subset, it is therefore continuous. □

As is customary, if E is a topological vector space, we will denote by $E' \subset E^*$ the topological dual space of E, namely E' is the subset formed by the continuous linear forms. Of course $E' = E^*$ if E is finite-dimensional. If E is a normed space, with norm $\|\cdot\|$, then E' is also a normed space for the usual norm

$$\|p\| = \sup\{p(v) \mid v \in E, \|v\| \leq 1\}.$$

In the case of continuous map, we can improve Proposition 1.2.3.

Proposition 1.2.6. *Suppose that $f : U \to \mathbb{R}$ is a continuous function defined on the topological vector space E. If we endow E' with the topology of simple convergence on E ("weak topology"), then the graph $\mathrm{SLF}(f)$ is a closed subset of $U \times E'$.*

The proof of this Proposition is obvious.

Exercise 1.2.7. *Let $f : U \to \mathbb{R}$ be a locally bounded function defined on the open subset U of the normed space E. (Recall that locally bounded means that each point in U has a neighborhood on which the absolute value of f is bounded)*

a) Show that for every $x \in U$, we can find a constant K, and a neighborhood V such that for every $y \in V$ and every $p \in \mathrm{SLF}_y(f)$ we have $\|p\| \leq K$. [Indication: see the proof of Theorem 1.4.1]

b) If E is finite dimensional, and f is continuous, show the following continuity property: for every $x \in U$, and every neighborhood W of $\mathrm{SLF}_x(f)$ in $E' = E^*$, we can find a neighborhood V of x such that for every $y \in V$ we have $\mathrm{SLF}_y(f) \subset W$.

As we will see now the notion of linear supporting form is tailored for convex functions.

Proposition 1.2.8. *If the function $f : U \to \mathbb{R}$, defined on the convex subset U of the vector space E, admits a supporting linear form at every $x \in U$, then f is convex.*

Proof. Let us suppose that $x_0 = y + (1-t)z$ with $y, z \in U$ and $t \in [0,1]$. If p is a supporting linear form at x_0, we have

$$f(y) - f(x_0) \geq p(y - x_0) \text{ and } f(z) - f(x_0) \geq p(z - x_0),$$

hence

$$\begin{aligned} tf(y) + (1-t)f(z) - f(x_0) &\geq p(t(y - x_0) + (1-t)(z - x_0)) \\ &= p(ty + (1-t)z - x_0) = 0. \end{aligned} \qquad \square$$

The following theorem is essentially equivalent to the Hahn-Banach Theorem.

Theorem 1.2.9. *Let U be a convex open subset of the locally convex topological vector space E. If $f : U \to \mathbb{R}$ is continuous and convex, then we can find a supporting linear form for f at each $x \in U$.*

Proof. As f is continuous and convex, the set

$$O = \{(x,t) \mid x \in U, f(x) < t\}$$

is open, non-empty, and convex in $E \times \mathbb{R}$. Since $(x_0, f(x_0))$ is not in O, by the Hahn-Banach Theorem, see [RV73, Theorem C, page 84] or [Rud91, Theorem, 3.4, page 59], there exists a continuous and non identically zero linear form $\alpha : E \times \mathbb{R} \to \mathbb{R}$ and such that

$$\forall (x,t) \in O,\ \alpha(x,t) > \alpha(x_0, f(x_0)).$$

We can write $\alpha(x,t) = p_0(x) + k_0 t$, with $p_0 : E \to \mathbb{R}$ a continuous linear form and $k_0 \in \mathbb{R}$. Since $\alpha(x_0, t) > \alpha(x_0, f(x_0))$ for all $t > f(x_0)$, we see that $k_0 > 0$. If we define $\tilde{p}_0 = k_0^{-1} p_0$, we get $\tilde{p}_0(x) + t \geq \tilde{p}_0(x_0) + f(x_0)$, for all $t > f(x)$, therefore $f(x) - f(x_0) \geq (-\tilde{p}_0)(x - x_0)$. The linear form $-\tilde{p}_0$ is the supporting linear form we are looking for. $\square$

The following Proposition is a straightforward consequence of Theorem 1.2.9 and Proposition 1.2.4

Proposition 1.2.10. *Let $f : U \to \mathbb{R}$ be a continuous convex function defined on an open convex subset U of the normed space E. If f is differentiable at x_0 then the derivative $Df(x_0)$ is the only supporting linear form of f at x_0. In particular, we have*

$$\forall x \in U,\ f(x) - f(x_0) \geq Df(x_0)(x - x_0).$$

Corollary 1.2.11. *Let $f : U \to \mathbb{R}$ be a continuous convex function defined on an open convex subset U of a normed space. If f is differentiable at x_0, then x_0 is a global minimum if and only if $Df(x_0) = 0$.*

Proof. Of course, if the derivative exists at a minimum it must be 0, this is true even if f is not convex. The converse, which uses convexity, follows from the inequality

$$f(y) - f(x_0) \geq Df(x_0)(y - x_0) = 0$$

given by Proposition 1.2.10 above. □

Corollary 1.2.12. *If $U \subset \mathbb{R}^n$ is open and convex and $f : U \to \mathbb{R}$ is a convex function, then, for almost all x, the function f admits a unique supporting linear form at x.*

Proof. This is a consequence of Proposition 1.2.10 above and Rademacher's Theorem 1.1.10. □

Exercise 1.2.13. *Let U be an open and convex subset of $\mathbb{R}^n$. Suppose that $f : U \to \mathbb{R}$ is convex and continuous. Show that if f admits a unique supporting linear form p_0 at x_0 then $Df(x_0)$ exists, and is equal to p_0. [Indication: For each $x \in U \setminus 0$, choose $p_x \in \mathrm{SLF}_x(f)$, and prove that*

$$p_0(x - x_0) \leq f(x) - f(x_0) \leq p_x(x - x_0).$$

Conclude using exercise 1.2.7.

1.3 The Fenchel Transform

Recall that for a topological vector E, we denote its topological dual by E'.

Definition 1.3.1 (Fenchel Transform)**.** If $L : E \to \mathbb{R}$ is function, the Fenchel transform of L, denoted by H (or L^* if we want to refer explicitly to L), is the function $H : E' \to]-\infty, +\infty]$ defined by

$$H(p) = \sup_{v \in E} \langle p, v\rangle - L(v).$$

We will call Fenchel's formula the relation between H and L.

The everywhere satisfied inequality

$$\langle p, v\rangle \leq L(v) + H(p),$$

is called the Fenchel inequality.

It is easily seen that $H(0) = -\inf_{v\in E} L(v)$ and that $H(p) \geq -L(0)$, for all $p \in E'$.

We have not defined H on E^* because it is identically $+\infty$ on $E^* \setminus E'$ under a very mild hypothesis on L.

Exercise 1.3.2. *If $L : E \to \mathbb{R}$ is bounded on some non-empty open subset of the normed space E, show that if we extend the Fenchel H to E^*, using the same definition, then H is identically $+\infty$ on $E^* \setminus E'$.*

Usually H assumes the value $+\infty$ even on E'. To give a case where H is finite everywhere, we must introduce the following definition.

Definition 1.3.3 (Superlinear)**.** Let E be a normed space. A map $f : E \to]-\infty, +\infty]$ is said to be superlinear, if for all $K < +\infty$, there exists $C(K) > -\infty$ such that $f(x) \geq K\|x\| + C(K)$, for all $x \in E$.

When E is finite-dimensional, all norms are equivalent hence the notion of superlinearity does not depend on the choice of a norm.

Exercise 1.3.4. *1) Show that $f : E \to \mathbb{R}$, defined on the normed space E, is superlinear if and only if $\lim_{\|x\|\to\infty} \frac{f(x)}{\|x\|} = +\infty$ and f is bounded below.*

2) If $f : E \to \mathbb{R}$ is continuous on the finite dimensional vector space E, show that it is superlinear if and only if

$$\lim_{x\to\infty} \frac{f(x)}{\|x\|} = +\infty.$$

Proposition 1.3.5. *Let $L : E \to \mathbb{R}$ be a function, defined on the normed space E, and let H be its Fenchel transform.*

(1) *If L is superlinear, then H is finite everywhere. It is even bounded on bounded subsets of E.*

(2) *If H is finite everywhere, it is convex.*

(3) *If L is bounded on bounded subsets of E, then H is superlinear. In particular, if L is continuous, and E is finite-dimensional, then H is superlinear.*

Proof. Let us show (1). We know that H is bounded below by $-L(0)$. It remains to show it is finite an bounded from above on each subset $\{p \in E' \mid \|p\| \leq K\}$, for each $K < +\infty$. By the superlinearity of L, there exists $C(K) > -\infty$ such that $L(v) \geq K\|v\| + C(K)$, for all $v \in E$, and thus for $p \in E'$ such that $\|p\| \leq K$, we have

$$\langle p, v\rangle - L(v) \leq \|p\| \, \|x\| - K\|x\| - C(\|p\|) \leq -C(\|p\|) < +\infty.$$

From which follows $\sup_{\|p\| \leq K} H(p) \leq -C(\|p\|) < +\infty$.

Property (2) results from the fact that H is an upper bound of a family of functions affine in p.

Let us show (3). We have

$$H(p) \geq \sup_{\|v\|=K} \langle p, v\rangle - \sup_{\|v\|=K} L(v).$$

But $\sup_{\|v\|=K} \langle p, v\rangle = K\|p\|$, and $\sup_{\|v\|=K} L(v) < +\infty$ by the hypothesis, since the sphere $\{v \in E \mid \|v\| = K\}$ is bounded. If E is finite dimensional, bounded sets are compact, and therefore, if L is continuous, it is bounded on bounded subsets of E. ☐

Theorem 1.3.6 (Fenchel)**.** *Let us suppose that $L : E \to \mathbb{R}$ is superlinear on the normed space E.*

(i) *The equality $\langle p_0, v_0\rangle = H(p_0) + L(v_0)$ holds if and only if p_0 is a supporting linear form for L at v_0.*

(ii) *If L is convex and differentiable everywhere then $\langle p, v\rangle = H(p) + L(v)$ if and only if $p = DL(v)$. Moreover*

$$\forall v \in E, H \circ DL(v) = DL(v)(v) - L(v).$$

(iii) *If we have* $L(v) = \sup_{p\in E'}\langle p, v\rangle - H(p)$*, for each* $v \in E$*, then* L *is convex. Conversely, if* L *is convex and continuous then* $L(v) = \sup_{p\in E'}\langle p, v\rangle - H(p)$*, for each* $v \in E$*.*

Proof. Let us show (i). If $L(v) - L(v_0) \geq \langle p_0, v - v_0\rangle$, we find $\langle p_0, v_0\rangle - L(v_0) \geq \langle p_0, v\rangle - L(v)$, for all $v \in E$, and thus $H(p_0) = \langle p_0, v_0\rangle - L(v_0)$. Conversely, by Fenchel's inequality $\langle p_0, v\rangle \leq H(p_0) + L(v)$, for all $v \in E$. If we subtract from this inequality the equality $\langle p_0, v_0\rangle = H(p_0) + L(v_0)$, we obtain $\langle p_0, v - v_0\rangle \leq L(v) - L(v_0)$.

Part (ii) follows from part (i) since for a differentiable function the only possible supporting linear form is the derivative, see Proposition 1.2.4.

Let us show (iii). If $L(v) = \sup_{p\in E'}\langle p, v\rangle - H(p)$, then, the function L is convex as a supremum of affine functions. Conversely, by (i) we always have $L(v) \geq \langle p, v\rangle - H(p)$. Therefore $L(v) \geq \sup_{p\in E'}\langle p, v\rangle - H(p)$. If L is convex, let p_0 be a linear supporting form for L at v, by (ii), we obtain $L(v) = \langle p_0, v\rangle - H(p_0)$ and thus $L(v) = \sup_{p\in E'}\langle p, v\rangle - H(p)$. □

Exercise 1.3.7. *Let* $L : E \to \mathbb{R}$ *be superlinear on the normed space* E*, and let* H *be its Fenchel transform. Denote by* A_L *the set of affine continuous functions* $v \mapsto p(v) + c, p \in E', c \in \mathbb{R}$*, such that* $L(v) \geq p(v) + c$*, for each* $v \in E$*. If* $L^{**} : E \to \mathbb{R}$ *is defined by* $L^{**}(v) = \sup_{f\in A_L} f(v)$*, show that*

$$L^{**}(v) = \sup_{p\in E'}\langle p, v\rangle - H(p).$$

[Indication: An affine function $f = p + c, p \in E', c \in \mathbb{R}$*, is in* A_L *if and only if* $c \leq -H(p)$*.]*

Proposition 1.3.8. *Suppose that* $L : E \to \mathbb{R}$ *is continuous and superlinear on the finite-dimensional linear space* E*, and* $H : E^* \to \mathbb{R}$ *is its Fenchel transform.*

(i) H *is everywhere continuous, and superlinear.*

(ii) *For every* $p \in E^*$*, there exists* $v \in E$ *such that* $\langle p, v\rangle = H(p) + L(v)$*.*

(iii) *If* L *is convex, for every* $v \in E$*, there exists* $p \in E^*$ *such that* $\langle p, v\rangle = H(p) + L(v)$*.*

Proof. We are assuming that E is finite-dimensional, and that L is continuous. Therefore, in part (i), the continuity follows from the convexity of H, see 1.1.6, and the superlinearity follows from part (iii) of Theorem 1.3.6.

We now prove part (ii). Since $\lim_{\|v\|\to+\infty} L(x,v)/\|v\| = +\infty$, and $|p(v)| \leq \|p\|\|v\|$, we see that

$$\lim_{\|v\|\to+\infty} \frac{[p(v) - L(x,v)]}{\|v\|} = -\infty.$$

Hence the supremum $H(x,p)$ of the continuous function $p(\cdot) - L(x,\cdot)$ is the same as the supremum of its restriction to big enough bounded sets. Since bounded sets in E are compact, the supremum $H(x,p)$ is achieved.

For part (iii), we remark that $E = E^{**}$, and that L is the Fenchel transform of H, by part (ii) of Fenchel's Theorem 1.3.6, therefore we can apply part (ii) of the present Proposition. □

Corollary 1.3.9. *If E is finite-dimensional and $L : E \to \mathbb{R}$ is everywhere differentiable and superlinear, then $DL : E \to E^*$ is surjective.*

Proof. This follows from part (ii) of Fenchel's Theorem 1.3.6 together with part (ii) of Proposition 1.3.8 (note that L is continuous since it is differentiable everywhere). □

We will need some fibered version of the results in this section.

We will have to consider locally trivial finite-dimensional vector bundle $\pi : E \to X$, where X is a Hausdorff topological space. We will use the notation (x, v) for a point in E to mean $x \in X$ and $v \in E_x = \pi^{-1}(x)$, with this notation $\pi : E \to X$ is the projection on the first coordinate $(x, v) \mapsto x$.

We denote, as is customary by $\pi^* : E^* \to X$ the dual vector bundle.

We recall that a continuous norm on $\pi : E \to X$ is a continuous function $(x, v) \mapsto \|v\|_x$ such that $v \mapsto \|v\|_x$ is a norm on the fiber E_x, for each $x \in X$. Such a norm induces a dual norm on $\pi^* : E^* \to X$ defined, for $p \in E^*_x$, in the usual way by

$$\|p|_x = \sup\{p(v) \mid v \in E_x, \|v|_x \leq 1\}.$$

The following result is classical.

Proposition 1.3.10. *Let $\pi : E \to X$ be a locally trivial vector bundle with finite-dimensional fibers over the Hausdorff topological space X, then all continuous norms on this bundle are equivalent above compact subsets of X. This means that for each compact subset $C \subset X$, and each pair $\|\cdot\|, \|\cdot\|'$ of continuous norms, there exists constants α, β, with $\alpha > 0$, and such that*

$$\forall (x,v) \in E, x \in C \Rightarrow \alpha^{-1}\|v\|_x \leq \|v\|'_x \leq \alpha\|v\|_x. \qquad \square$$

Proof. We do it first for the case of the trivial bundle $x \times \mathbb{R}^n \to X$, with X compact. It is not difficult to see that it suffices to do it with $\|\cdot\|_x$ a fixed norm independent of x, for example the Euclidean norm on $\mathbb{R}^n$, which we simply denote by $\|\cdot\|$. The set $S = X \times \{v \in \mathbb{R}^n \mid \|v\| = 1\}$ is compact and disjoint from $\times \{0\}$, therefore by continuity the two bounds $\alpha = \inf_{(x,v)\in S}\|v\|_x, \beta = \sup_{(x,v)\in S}\|v\|'_x$ are attained, hence they are finite and $\neq 0$. It is not difficult to see by homogeneity that

$$\forall (x,v) \in X \times \mathbb{R}^n, \alpha\|v\| \leq \|v\|'_x \leq \|v\|.$$

For the case of a general bundle, if $C \subset X$ is compact, we can find a finite number $U_1, \ldots, U_n$ of open subsets of X such that the bundle is trivial over each U_i, and $C \subset U_1 \cup \cdots \cup U_n$. Since X is Hausdorff, we can write $C = C_1 \cup \cdots \cup C_n$, with C_i compact, and included in U_i. From the first part of the proof two norms on the bundle are equivalent above each C_i, hence this is also the case of their (finite) union C. $\square$

Definition 1.3.11 (Superlinear Above Compact subsets)**.** Suppose that $\pi : E \to X$ is a finite-dimensional locally trivial vector bundle over the topological space X. We say that a function $L : E \to X$ is superlinear above compact subsets if for every compact subset $C \subset X$, and every $K \geq 0$, we can find a constant $A(C,K) > -\infty$ such that

$$\forall (x,v) \in E, x \in C \Rightarrow L(x,v) \geq K\|v\|_x + A(C,K),$$

where $\|\cdot\|_x$ is a fixed continuous norm on the vector bundle E.

When X is compact we will say that L is superlinear instead of superlinear above compact subsets. Of course in that case, it suffices to verify the condition of superlinearity with $K = X$.

Of course, the condition above is independent of the choice of the continuous norm on the vector bundle, since all norms are equivalent by Proposition 1.3.10. We have not defined the concept of uniform superlinearity for a general X because it depends on the choice of the norm on the bundle, since if X is not compact not all norms are equivalent.

Theorem 1.3.12. *Suppose $L : E \to \mathbb{R}$ is a continuous function on the total space of the finite-dimensional locally trivial vector bundle $\pi : E \to X$. We consider $\pi^* : E^* \to X$, the dual vector bundle and define $H : E^* \to \mathbb{R}$ by*

$$H(x,p) = \sup_{v \in E_x} p(v) - L(x,v).$$

If L is superlinear above compact subsets of X, and X is a Hausdorff locally compact, topological space, then H is continuous and superlinear above compact subsets of X.

Proof. Since continuity is a local property, and X is Hausdorff locally compact, without loss of generality, we can assume X compact, and $\pi : E \to X$ trivial, therefore $E = X \times \mathbb{R}^n$. We choose a norm $\|\cdot\|$ on $\mathbb{R}^n$.

Fix $K \geq 0$, we can pick $C > -\infty$ such that

$$\forall (x,v) \in X \times \mathbb{R}^n, L(x,v) \geq (K+1)\|v\| + C.$$

If we choose $R > 0$ such that $R + C > \sup_{x \in X} L(x,0)$ (this is possible since the right hand side is finite by the compactness of X), we see that for each $x \in X, v \in \mathbb{R}^n$,and each $p \in \mathbb{R}^{n*}$ satisfying $\|p\| \leq K, \|v\| \geq R$, we have

$$\begin{aligned} p(v) - L(x,v) &\leq \|p\|v\|-\| - (K+1)\|v\| - C \\ &\leq -R - C < -\sup_{x \in X} L(x,0) \\ &\leq -L(x,0) = p(0) - L(x,0). \end{aligned}$$

Therefore for $\|p\| \leq K$, we have $H(x,p) = \sup_{\|v\| \leq R} p(v) - L(x,v)$. Since $\{v \in \mathbb{R}^n \mid \|v\| \leq R\}$ is compact, we see that H is continuous on the set $X \times \{p \in \mathbb{R}^{n*} \mid \|p\| \leq K\}$. But $K \geq 0$ is arbitrary, therefore the function H is continuous everywhere.

We prove superlinearity above compact subsets of X. Using the same argument as in final part the proof of Proposition 1.3.10 above, we can without loss of generality suppose that X is compact, and that the bundle is the trivial bundle $X \times \mathbb{R}^n \to X$. For a fixed K, remark that by compactness, and continuity $A = \sup\{L(x, v) \mid x \in X, v \in \mathbb{R}^n, \|v\| \leq K\}$ is finite. Therefore $H(x, p) \geq p(v) - A$, for each $v \in \mathbb{R}^n$, satisfying $\|v\| \leq K$. If we take the supremum over all such v's, since $K\|p\| = \sup\{p(v) \mid v \in \mathbb{R}^n, \|v\| \leq K\}$, we get $H(x, p) \geq K\|p\| - A$. □

Definition 1.3.13 (Convex in the Fibers)**.** Let $L : E \to \mathbb{R}$ be a continuous function on the total space of the finite-dimensional locally trivial vector bundle $\pi : E \to X$, where X is a Hausdorff space. We will say that a Lagrangian L on the manifold M is convex in the fibers, if the restriction $L_{|E_x}$ is convex for each $x \in X$.

In fact, for convex functions superlinearity above compact sets is not so difficult to have, because of the following not so well known theorem.

Theorem 1.3.14. *Suppose $L : E \to \mathbb{R}$ is a continuous function on the total space of the finite-dimensional locally trivial vector bundle $\pi : E \to X$, where X is a Hausdorff space. If L is convex in the fibers, then L is superlinear above each compact subsets of X if and only if $L_{|E_x}$ is superlinear, for each $x \in X$.*

Proof. Of course, it is obvious that if L is superlinear above each compact subset, then each restriction $L_{|E_x}$ is superlinear.

Suppose now that $L_{|E_x}$ is convex and superlinear for each $x \in X$, to prove that L is linear above compact subsets of X, again by the same argument as in final part the proof of Proposition 1.3.10 above, we can without loss of generality that X is compact, and that the bundle is the trivial bundle $X \times \mathbb{R}^n \to X$.

We choose a fixed norm $\|\cdot\|$ on $\mathbb{R}^n$. For given $x_0 \in X$, and $K \geq 0$, we will show that there exists a neighborhood V_{x_0} of x_0 and $C(x_0, K) > -\infty$ such that

$$\forall x \in V_{x_0}, \forall v \in \mathbb{R}^n, L(x, v) \geq K\|v\| + C(x_0, K). \qquad (*)$$

A compactness argument finishes the proof.

We now prove (*). We choose $C_1 > -\infty$ such that

$$\forall v \in \mathbb{R}^n, L(x_0, v) \geq (K+1)\|v\| + C_1.$$

We then pick $R > 0$ such that $R + C_1 \geq L(x_0, 0) + 1$. Now if $p \in \mathbb{R}^{n*}$, and $v \in \mathbb{R}^n$ satisfy respectively $\|p\|_{x_0} \leq K$, and $\|v\|_{x_0} = R$, we see that

$$\begin{aligned} L(x_0, v) - p(v) &\geq (K+1)\|v\| + C_1 - K\|v\| \\ &\geq R + C_1 \\ &\geq L(x_0, 0) + 2. \end{aligned}$$

Since the set $\{(v, p) \in \mathbb{R}^n \times \mathbb{R}^{n*} \mid \|v\|_{x_0} = R, \|p\|_{x_0} \leq K\}$ is compact, and L is continuous, we can find a neighborhood V_{x_0} of x_0 in X such that for each $x \in V_{x_0}, v \in \mathbb{R}^n$, and each $p \in \mathbb{R}^{n*}$, we have

$$\|v\| = R, \|p\| \leq K \Rightarrow L(x, v) - p(v) > L(x, 0).$$

This implies that for fixed $x \in V_{x_0}$, and $p \in \mathbb{R}^{n*}$ satisfying $\|p\| \leq K$, the convex function $L(x, \cdot) - p(\cdot)$ achieves its minimum on the compact set $\{v \in \mathbb{R}^n \mid \|v\| \leq R\}$ in the interior of that set. Therefore, the convex function $L(x, \cdot) - p(\cdot)$ has a local minimum attained in $\{v \in \mathbb{R}^n \mid \|v\| < R\}$. By convexity this local minimum must be global, see 1.1.12. Therefore, defining $C = \inf\{L(x, v) - p(v) \mid x \in X, \|v\|_x \leq R, \|p\|_x \leq K\}$, we observe that C is finite by compactness, and we have

$$\forall (x, v, p) \in V_{x_0} \times \mathbb{R}^n \times \mathbb{R}^{n*}, \|p\| \leq K \Rightarrow L(x, v) - p(v) \geq C.$$

Taking the infimum of the right hand side over all $\|p\|_x \leq K$, we get

$$\forall (x, v) \in V_{x_0} \times \mathbb{R}^n, L(x, v) - K\|v\| \geq C. \qquad \square$$

1.4 Differentiable Convex Functions and Legendre Transform

Theorem 1.4.1. *Let U be an open convex subset of $\mathbb{R}^n$. If $f : U \to \mathbb{R}$ is convex and differentiable at each point of U, then f is C^1.*

Proof. We fix $x \in U$. Let $r \in]0, \infty[$ be such that the closed ball $\bar{B}(x, r)$ is contained in U. Let us set $M = \sup_{y \in \bar{B}(x,r)} |f(y)| < +\infty$. For $h, k \in \bar{B}(0, \frac{r}{2})$, we have

$$f(x + h + k) - f(x + k) \geq Df(x + k)(h), \qquad (*)$$

taking the supremum over all h such that $\|h\| = r/2$, we obtain $\|Df(x + k)\| \leq 4M/r$. Since the ball $\{p \in \mathbb{R}^{n*} \mid \|p\| \leq 4M/r\}$ is compact, it is enough to see that if $k_n \to 0$ and $Df(x + k_n) \to p$, then $p = Df(x)$. But taking the limit in the inequality $(*)$, we get

$$\forall k \in \bar{B}(0, r/2),\ f(x + h) - f(x) \geq \langle p, h \rangle.$$

It results that $Df(x) = p$, since we have already seen that at a point where a function is differentiable only its derivative can be a supporting linear form, see Proposition 1.2.4. $\square$

Exercise 1.4.2. *Let K be a compact topological space and U an open convex subset of $\mathbb{R}^n$. If $L : K \times U \to \mathbb{R}$ is continuous and such that for each $k \in K$, the map $U \to \mathbb{R} : v \mapsto L(k, v)$ is convex and everywhere differentiable, then $\frac{\partial L}{\partial v} : K \times U \to (\mathbb{R}^n)^*, (k, v) \mapsto \frac{\partial L}{\partial v}(k, v)$ is continuous. [Indication: Adapt the proof of Theorem 1.4.1.]*

Definition 1.4.3 (Legendre Transform)**.** Let $L : U \to \mathbb{R}$ be a C^1 function, with $U \subset \mathbb{R}^n$ open. The Legendre transform associated with L is the map $\mathcal{L} : U \to \mathbb{R}^{n*}, v \mapsto DL(v)$.

We can rephrase part (ii) of Fenchel's Theorem 1.3.6 and Corollary 1.3.9 in the following way:

Proposition 1.4.4. *Let $L : \mathbb{R}^n \to \mathbb{R}$ be C^1, convex and superlinear, then its Legendre transform $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^{n*}$ is surjective. Moreover, if we denote by $H : \mathbb{R}^{n*} \to \mathbb{R}$ its Fenchel transform then $\langle p, v \rangle = H(p) + L(v)$ if and only if $p = DL(v)$, and we have*

$$\forall v \in \mathbb{R}^n, H \circ \mathcal{L}(v) = DL(v)(v) - L(x, v).$$

In particular, the surjectivity of $\mathcal{L}$ is a consequence of superlinearity of L.

We are interested in finding out, for a C^1 convex function $L : \mathbb{R}^n \to \mathbb{R}$, when its Legendre transform $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^{n*}$ is bijective. It is easy to understand when $\mathcal{L}$ is injective.

Theorem 1.4.5. *Suppose $L : U \to \mathbb{R}$ is a C^1 convex function, defined on an open subset U of $\mathbb{R}^n$. Its associated Legendre transform $\mathcal{L}$ is injective if and only if L is strictly convex.*

Proof. Let $p \in \mathbb{R}^{n*}$. We have $p = DL(x)$ if and only if $DL_p(x) = 0$ where $L_p(x) = L(x) - p(x)$. Hence x is a point where the function L_p reaches its minimum, see 1.2.11. If L is strictly convex so is L_p. However a strictly convex function can achieve its minimum at most at one point.

Conversely, if $\mathcal{L}$ is injective, the convex function $L(x) - DL(x_0)(x)$ has only x_0 as a critical point and hence, and again by Corollary 1.2.11, it reaches its minimum only at x_0. If $x_0 = tx + (1-t)y$ with $t \in]0,1[$, $x \neq x_0$ and $y \neq x_0$, we therefore have

$$\begin{aligned} L(x) - DL(x_0)(x) &> L(x_0) - DL(x_0)(x_0) \\ L(y) - DL(x_0)(y) &> L(x_0) - DL(x_0)(x_0). \end{aligned}$$

Since $t > 0$ and $(1-t) > 0$, we obtain

$$tL(x) + (1-t)L(y) - DL(x_0)(\underbrace{tx + (1-t)y}_{x_0}) > L(x_0) - DL(x_0)(x_0),$$

hence $tL(x) + (1-t)L(y) > L(x_0)$. $\square$

We would like now to prove the following theorem.

Theorem 1.4.6. *Let $L : \mathbb{R}^n \to \mathbb{R}$ be C^1, and convex. If $\mathcal{L}$ is its Legendre transform, then the following statements are equivalent:*

(1) *The function L is strictly convex, and superlinear.*

(2) *Its Legendre transform $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^{n*}$ is a homeomorphism.*

(3) *Its Legendre transform $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^{n*}$ is bijective.*

Proof. We first show that (1) implies (3). If (1) is true then from Proposition 1.4.4, we know that $\mathcal{L}$ is surjective, and from Theorem 1.4.5 it is injective.

The fact that (3) implies (2) follows from Brouwer's Theorem on the invariance of the domain see [Dug66, Theorem 3.1, page 358]. (Note that one can obtain a proof independent from Brouwer's Theorem by using Theorem 1.4.13 below.)

We now prove that (2) implies (1). Another application of Theorem 1.4.5 shows that L is strictly convex. It remains to show the superlinearity. Since $\mathcal{L}$ is a homeomorphism, the set $A_K = \{x \mid \|DL(x)\| = K\}$ is compact, and $\mathcal{L}(A_K) = \{p \in \mathbb{R}^{n*} \mid \|p\| = K\}$, thus

$$\forall v \in \mathbb{R}^n,\ K\|v\| = \sup_{x \in A_K} DL(x)(v).$$

As $L(v) \geq DL(x)(v) + L(x) - DL(x)(x)$ we see that

$$L(v) \geq K\|v\| + \inf_{x \in A_K} [L(x) - DL(x)(x)],$$

but $\inf_{x \in A_K}[L(x) - DL(x)(x)] > -\infty$, because A_K is compact and L is of class C^1. □

When it comes to Lagrangians, Analysts like to assume that they are superlinear, and Geometers prefer to assume that its associated Legendre transform is bijective. The following Corollary shows that for C^2-strictly convex Lagrangians, these hypothesis are equivalent.

Corollary 1.4.7. *Let $L : \mathbb{R}^n \to \mathbb{R}$ be a C^2 convex function. Its associated Legendre transform $\mathcal{L}$ is a C^1 diffeomorphism from $\mathbb{R}^n$ onto its dual space $\mathbb{R}^{n*}$ if and only if L is superlinear, and C^2-strictly convex.*

Proof. Suppose that $\mathcal{L} = DL$ is a C^1 diffeomorphism. By the previous Theorem 1.4.6, the map L is superlinear. Moreover, the derivative $D\mathcal{L}(v) = D^2L(v)$ is an isomorphism, for each $v \in \mathbb{R}^n$. Therefore $D^2L(v)$ is non degenerate as a bilinear form, for each $v \in \mathbb{R}^n$. Since, by the convexity of L, the second derivative $D^2L(v)$ is non negative definite as a quadratic form, it follows that $D^2L(v)$ is positive definite as a quadratic form, for each $v \in \mathbb{R}^n$.

Conversely, suppose L superlinear, and C^2-strictly convex . Then $DL : \mathbb{R}^n \to \mathbb{R}^{n*}$ is a homeomorphism by Theorem 1.4.6. Moreover, since $D\mathcal{L}(v) = D^2L(v)$, the derivative $D\mathcal{L}(v)$ is thus an isomorphism at every point $v \in \mathbb{R}^n$. By the Local Inversion Theorem, the inverse map $\mathcal{L}^{-1}$ is also C^1. □

In the sequel of this section, we will discuss some aspects of the Legendre transform that will not be used in this book. They nonetheless deserve to be better known.

We start with the notion of proper map.

Definition 1.4.8 (Proper Map)**.** A map $f : X \to Y$, between the topological spaces X and Y, is said to be proper if for every compact subset K of the target space Y, the inverse image $f^{-1}(K) \subset X$ is also compact.

The main properties of proper maps are recalled in the following exercise.

Exercise 1.4.9. *Let $f : X \to Y$ be a proper continuous map between metric spaces.*

1) Show that for each closed subset $F \subset X$, the image $f(F)$ is closed in Y. [Indication: Use the fact that if a sequence converges, then the subset formed by this sequence together with its limit is compact.

2) Conclude that f is a homeomorphism as soon as it is bijective.

3) Show that a continuous map $f : \mathbb{R}^n \to \mathbb{R}^m$ is proper if and only if

$$\lim_{\|x\|\to+\infty} \|f(x)\| = +\infty.$$

Theorem 1.4.10. *Let $L : U \to \mathbb{R}$ be a C^1 convex function, where U is an open convex subset of $\mathbb{R}^n$. If its associated Legendre transform $\mathcal{L} : U \to \mathbb{R}^{n*}$ is proper, then $\mathcal{L}$ is surjective.*

We need some preliminaries in order to prove the theorem.

Lemma 1.4.11 (Of the Minimum)**.** *Let $f : \bar{B}(x, r) \to \mathbb{R}$ be a function which has a derivative at each point of $\bar{B}(x, r)$. If f achieves its minimum at $x_0 \in \bar{B}(x, r)$, a closed ball in a normed space, then $Df(x_0)(x_0 - x) = -\|Df(x_0)\|r = -\|Df(x_0)\|\|x_0 - x\|$.*

Proof. Without loss of generality, we can suppose $x = 0$. For all $y \in \bar{B}(0, r)$ and for all $t \in [0, 1]$, we have

$$f(ty + (1 - t)x_0) \geq f(x_0),$$

thus, the function $\phi_y : [0, 1] \to \mathbb{R}, t \mapsto f(ty + (1 - t)x_0)$ has a minimum at $t = 0$, its derivative at 0, namely $Df(x_0)(y - x_0)$, is thus ≥ 0. Hence $Df(x_0)(y - x_0) \geq 0$, for each $y \in \bar{B}(0, r)$, and consequently $Df(x_0)(x_0) \leq Df(x_0)(y)$, for each $y \in \bar{B}(0,r)$. It follows that

$$Df(x_0)(x_0) = \inf_{y \in \bar{B}(0,r)} Df(x_0)(y) = -\|Df(x_0)\|r.$$

If $Df(x_0) = 0$, we also have the second part of the required equalities. If $Df(x_0) \neq 0$, then x must be on the boundary $\partial B(0,r)$ of $\bar{B}(0,r)$ and we again have the second part of the equalities. □

Corollary 1.4.12. *Let $f : U \to \mathbb{R}$ be a C^1 convex function defined on the open convex subset U of $\mathbb{R}^n$. If the derivative $Df(x)$ is never the 0 linear form, for $x \in U$, then for each compact subset $K \subset U$, we have*

$$\inf_{x \in U} \|Df(x)\| = \inf_{x \in U \setminus K} \|Df(x)\|.$$

Proof. The inequality $\inf_{x \in U} \|Df(x)\| \leq \inf_{x \in U \setminus K} \|Df(x)\|$ is obvious. If we do not have equality for some compact subset K, then

$$\inf_{x \in U} \|Df(x)\| < \inf_{x \in U \setminus K} \|Df(x)\|,$$

and therefore

$$\inf_{x \in U} \|Df(x)\| = \inf_{x \in K} \|Df(x)\|,$$

If we set $K_0 = \{x \in U \mid \|Df(x)\| = \inf_{z \in U} \|Df(z)\|\}$, it follows that K_0 is closed, non-empty, and contained in K, therefore K_0 is compact. Moreover

$$\forall x \in U \setminus K_0, \|Df(x)\| > \inf_{z \in U} \|Df(z)\|.$$

Since K_0 is compact there exists $r > 0$ such that the closed set $\bar{V}_r(K_0) = \{x \mid d(x, K_0) \leq r\}$ is contained in the open set U. As this set $\bar{V}_r(K_0)$ is also compact, there exists $x_0 \in \bar{V}_r(K_0)$ such that $f(x_0) = \inf_{x \in \bar{V}_r(K_0)} f(x)$. Necessarily x is on the boundary of $\bar{V}_r(K_0)$, because otherwise x_0 would be a local minimum of f and therefore $Df(x_0) = 0$, which is excluded. Hence $d(x_0, K_0) = r$. By compactness of K_0, we can find $x \in K_0$ such that $d(x_0, x) = r$. Since $\bar{B}(x,r) \subset \bar{V}_r(K_0)$ and $x_0 \in \bar{B}(x,r)$, we also have $f(x_0) =$

$\inf_{y\in \bar{B}(x,r)} f(y)$. By the previous Lemma 1.4.11, we must have $Df(x_0)(x_0 - x) = -\|Df(x_0)\| r$. The convexity of f gives

$$f(x) - f(x_0) \geq Df(x_0)(x - x_0)$$
$$f(x_0) - f(x) \geq Df(x)(x_0 - x),$$

hence $Df(x)(x - x_0) \geq Df(x_0)(x - x_0) = \|Df(x_0)\| r$. As $\|x - x_0\| = r$, we get

$$Df(x)(x - x_0) \leq \|Df(x)\| \|x - x_0\| = \|Df(x)\| r.$$

This implies $\|Df(x)\| \geq \|Df(x_0)\|$, which is absurd. In fact, we have $\|Df(x)\| = \inf_{z\in U} \|Df(z)\|$, since $x \in K_0$, and $\|Df(x_0)\| > \inf_{z\in U} \|Df(z)\|$, because $x_0 \notin K_0$. □

Proof of theorem 1.4.10. Fix $p \in \mathbb{R}^{n*}$, the Legendre transform of $L_p = L - p$ is $\mathcal{L} - p$, it is thus also proper. By the previous Corollary, it must vanish at some point in U, because $\inf_{x\notin \bar{B}(0,r)} \|DL_p(x)\| \to \infty$, when $r \to \infty$. □

Theorem 1.4.13. *Let $L : \mathbb{R}^n \to \mathbb{R}$ be a C^1 convex function. Its associated Legendre transform $\mathcal{L} : \mathbb{R}^n \to \mathbb{R}^{n*}$ is proper if and only if L is superlinear.*

Proof. Let us suppose L superlinear. By convexity we have $L(0) - L(x) \geq DL(x)(0 - x)$ and thus $DL(x)(x) \geq L(x) - L(0)$ from which we obtain

$$\|DL(x)\| \geq DL(x)\left(\frac{x}{\|x\|}\right) \geq \frac{L(x)}{\|x\|} - \frac{L(0)}{\|x\|},$$

by the superlinearity of L, we do have $\|DL(x)\| \to \infty$, when $\|x\| \to \infty$.

The proof of the converse is very close to the end of the proof of Theorem 1.4.6. If $\mathcal{L}$ is proper, the set $A_K = \{x \mid \|DL(x)\| = K\}$ is compact. Moreover, since $\mathcal{L}$ is necessarily surjective, see 1.4.10, we have $\mathcal{L}(A_K) = \{p \in \mathbb{R}^{n*} \mid \|p\| = K\}$, and thus

$$\forall v \in \mathbb{R}^n,\ K\|v\| = \sup_{x\in A_K} DL(x)(v).$$

As $L(v) \geq DL(x)(v) + L(x) - DL(x)(x)$ we see that

$$L(v) \geq K\|v\| + \inf_{x\in A_K} [L(x) - DL(x)(x)],$$

but $\inf_{x\in A_K}[L(x)-DL(x)(x)] > -\infty$, because A_K is compact and L is of class C^1. □

We would like to conclude this section with a very nice theorem due to Minty see [Min64, Min61] (see also [Gro90, 1.2. Convexity Theorem]). In order to give the best statement, we recall some notions about convex subsets of a finite-dimensional vector space E. If $C \subset F$ is a convex subset, we will denote by $\mathrm{Aff}(C)$ the affine subspace generated by C, the *relative interior* $\mathrm{relint}(C)$ of C is its interior as a subset of $\mathrm{Aff}(C)$.

Theorem 1.4.14 (Minty). *If $L : \mathbb{R}^n \to \mathbb{R}$ is a C^1 convex function, then the closure of the image $\mathcal{L}(\mathbb{R}^n) = DL(\mathbb{R}^n)$ of its associated Legendre transform $\mathcal{L}$ is convex. Moreover $\mathcal{L}$ contains the relative interior of its closure.*

In order to prove this theorem, we will need the following lemma.

Lemma 1.4.15. *Let $L : \mathbb{R}^n \to \mathbb{R}$ be a C^1 convex function. If $p \notin \mathcal{L}(\mathbb{R}^n)$, then there exists $v \in \mathbb{R}^n \setminus \{0\}$ such that $\mathcal{L}(x)(v) \geq p(v)$ for all $x \in \mathbb{R}^n$. If p is not in the closure of $\mathcal{L}(\mathbb{R}^n)$, then, moreover, there exists $\epsilon > 0$ such that $\mathcal{L}(x)(v) \geq \epsilon + p(v)$ for all $x \in \mathbb{R}^n$.*

Proof of Theorem 1.4.14. To simplify notations, we call C the closure of $\mathcal{L}(\mathbb{R}^n)$. Observe that a linear form on $\mathbb{R}^{n*}$ is of the form $p \mapsto p(v)$, where $v \in \mathbb{R}^n$. Therefore the Lemma above 1.4.15 shows that a point in the complement of C, can be strictly separated by a hyperplane from $\mathcal{L}(\mathbb{R}^n)$, and hence from its closure C. This implies the convexity of C.

It remains to prove the second statement.

We first assume that the affine subspace generated by $\mathcal{L}(\mathbb{R}^n)$ is the whole of $\mathbb{R}^{n*}$. We have to prove that the interior of C is contained in $\mathcal{L}(\mathbb{R}^n)$. Suppose that $p_0 \in \mathring{C}$ is not contained in $\mathcal{L}(\mathbb{R}^n)$, by Lemma 1.4.15 above we can find $v \in \mathbb{R}^n \setminus \{0\}$ with $\mathcal{L}(z)(v) \geq p_0(v)$ for all $z \in \mathbb{R}^n$, therefore $p(v) \geq p_0(v)$ for all p in the closure C of $\mathcal{L}(\mathbb{R}^n)$. This is clearly impossible since $p_0 \in \mathring{C}$ and $v \neq 0$.

To do the general case, call E the affine subspace generated by $\mathcal{L}(\mathbb{R}^n)$. Replacing L by $L - DL(0)$, we can assume that $0 \in E$, and therefore E is a vector subspace of $\mathbb{R}^{n*}$. Changing bases, we can assume that $E = \mathbb{R}^{k*} \times \{0\} \subset \mathbb{R}^{n*}$. Since $DL(x) \in E = \mathbb{R}^{k*} \times \{0\}$, we see that $\partial L/\partial x_i$ is identically 0 for $i > 0$, and therefore $L(x_1, \dots, x_n)$ depends only on the first k variables, so we can write $L(x_1, \dots, x_n) = \tilde{L}(x_1, \dots, x_k)$, with $\tilde{L} : \mathbb{R}^k \to \mathbb{R}$ convex and C^1. It is obvious that DL and $D\tilde{L}$ have the same image in $\mathbb{R}^{k*} = \mathbb{R}^{k*} \times \{0\} \subset \mathbb{R}^{n*}$, therefore the image of $D\tilde{L}$ generates affinely $\mathbb{R}^{k*}$. We can therefore apply the first case treated above to finish the proof. □

Proof of Lemma 1.4.15. We fix $p_0 \notin \mathcal{L}(\mathbb{R}^n) = DL(\mathbb{R}^n)$. The function $L_{p_0} = L - p_0$ is convex. As its derivative is never the 0 linear map, it does not have a local minimum. To simplify the notations let us set $f = L - p_0$. For each integer $k \geq 1$, by Lemma 1.4.11 applied to f, and to the ball $\bar{B}(0, k)$, we can find x_k with $\|x_k\| = k$ such that

$$Df(x_k)(x_k) = -\|Df(x_k)\|\|x_k\|.$$

The convexity of f gives

$$\forall y \in \mathbb{R}^n,\ Df(y)(y - x_k) \geq f(y) - f(x_k) \geq Df(x_k)(y - x_k),$$

hence

$$\forall y \in \mathbb{R}^n,\ Df(y)(y - x_k) \geq Df(x_k)(y - x_k). \qquad (*)$$

In particular, we have $Df(0)(-x_k) \geq Df(x_k)(-x_k) = \|Df(x_k)\|\|x_k\|$, and thus $\|Df(0)\| \geq \|Df(x_k)\|$. Taking a subsequence, we can suppose that $\|x_k\| \to \infty$, that $Df(x_k) \to p_\infty$, and that $x_k/\|x_k\| \to v_\infty$, with v_∞ of norm 1. Dividing both sides the inequality $(*)$ above by $\|x_k\|$, and using the equality

$$Df(x_k)(x_k) = -\|Df(x_k)\|\|x_k\|,$$

and taking limits we obtain

$$\forall y \in \mathbb{R}^n,\ Df(y)(-v_\infty) \geq \|p_\infty\|,$$

we rewrite it as

$$\forall y \in \mathbb{R}^n DL(y)(-v_\infty) \geq p_0(-v_\infty) + \|p_\infty\|.$$

It then remains to observe that $p_\infty = 0$ implies that $Df(x_k) = DL(x_k) - p_0$ tends to 0 and thus p_0 is in the closure of $DL(\mathbb{R}^n)$. □

Exercise 1.4.16. *Suppose that $L : \mathbb{R}^n \to \mathbb{R}$ is C^1, and strictly convex. Show that the image $\mathcal{L}(\mathbb{R}^n)$ is a convex open subset of $\mathbb{R}^{n*}$.*

Does this result remain true, for $L : U \to \mathbb{R}$ C^1, and strictly convex, on the open subset $U \subset \mathbb{R}^n$?

1.5 Quasi-convex functions

At some point, we will need a class of functions more general than the convex ones.

Definition 1.5.1 (Quasi-convex)**.** Let $C \subset E$ be a convex subset of the vector space E. A function $f : C \to \mathbb{R}$ is said to be quasi-convex if for each $t \in \mathbb{R}$, the subset $f^{-1}(]-\infty, t])$ is convex.

Proposition 1.5.2. *Let $f : C \to \mathbb{R}$ be a function defined on the convex subset C of the vector space E.*

1 *The function f is quasi-convex if and only*

$$\forall x, y \in C, \forall \alpha \in [0,1], f(\alpha x + (1-\alpha)y) \leq \max(f(x), f(y)).$$

2 *If f is quasi-convex then for every $x_1, \cdots, x_n \in C$ and every $\alpha_1, \cdots, \alpha_n \in [0,1]$, with $\sum_{i=1}^n \alpha_i = 1$, we have*

$$f(\sum_{i=1}^n \alpha_i x_i) \leq \max_{1 \leq i \leq n} f(x_i).$$

Proof. We prove (2) first. Suppose f is quasi convex. Since $f^{-1}(]-\infty, \max_{\leq i \leq n} f(x_i)])$ is convex and contains $x_1, \cdots, x_n$ necessarily $\sum_{i=1}^n \alpha_i, x_i \in f^{-1}(-]\infty, \max_{\leq i \leq n} f(x_i)]$.

To finish proving (1), suppose conversely that

$$\forall x, y \in C, \forall \alpha \in [0,1], f(\alpha x + (1-\alpha)y) \leq \max(f(x), f(y)).$$

If x, y are in $f^{-1}(]-\infty, t])$, then $f(x)$ and $f(y)$ are $\leq t$. Therefore any convex combination $\alpha x + (1-\alpha)y$ satisfies $f(\alpha x + (1-\alpha)y) \leq \max(f(x), f(y)) \leq t$, and hence $\alpha x + (1-\alpha)y \in f^{-1}(]-\infty, t])$. □

Example 1.5.3. 1) Any convex function is quasi-convex.

2) Any monotonic function $\varphi : I \to \mathbb{R}$, where I is an interval in $\mathbb{R}$, is quasi-convex.

We need a slight generalization of property (2) of Proposition 1.5.2. We start with a lemma. Although we do not need it in its full generality, it is nice to have the general statement.

Lemma 1.5.4. *Suppose that $(X, \mathcal{A}, \mu)$ is a probability measure, and that $\varphi : X \to C$ is a measurable function with value in a convex subset C of a finite-dimensional normed space E. If $\int_X \|\varphi(x)\| \, d\mu(x) < +\infty$, then $\int_X \varphi(x) \, d\mu(x)$ is contained in C.*

Proof. We will do the proof by induction on $\dim C$. Recall that the dimension of a convex set is the dimension of the smallest affine subspace that contains it. If $\dim C = 0$, then by convexity C is reduced to one point and the result is trivial.

We assume now that the result is true for every $n < \dim C$. Replacing E by an affine subset we might assume that C has a non empty interior. Therefore the convex set C is contained in the closure of its interior $\mathring{C}$, see Lemma 1.5.5 below. Let us define $v_0 = \int_X \varphi(x) d\mu(x)$. If $v_0 \notin \mathring{C}$, since $\mathring{C}$ is open and convex, by Hahn-Banach Theorem there exists a linear form $\theta : E \to \mathbb{R}$ such that $\theta(v) > \theta(v_0)$, for each $v \in \mathring{C}$. Since C is contained in the closure of $\mathring{C}$, we obtain

$$\forall v \in C, \theta(v) \geq \theta(v_0).$$

Therefore, we have the inequality

$$\forall x \in X, \theta \circ \varphi(x) \geq \theta(v_0). \tag{†}$$

If we integrate this inequality we get $\int_X \theta \circ \varphi(x) \, d\mu(x) \geq \theta(v_0)$. By linearity of θ, the integral $\int_X \theta \circ \varphi(x) d\mu(x)$ is equal to $\theta(\int_X \varphi(x) \, d\mu(x))$, hence to $\theta(v_0)$, by the definition of v_0. This means that the integration of the inequality (†) leads to an equality, therefore we have $\theta \circ \varphi(x) = \theta(v_0)$, for μ-almost every $x \in X$. It follows that, on a set of full μ-measure $\varphi(x) \in \theta^{-1}(\theta(v_0)) \cap C$. But the subset $\theta^{-1}(\theta(v_0)) \cap C$ is convex and has a dimension $< \dim C = \dim E$, because it is contained in the affine hyperplane $\theta^{-1}(\theta(v_0))$. By induction $\int_X \varphi(x) \, d\mu(x) \in \theta^{-1}(\theta(v_0)) \cap C$. $\square$

Lemma 1.5.5. *If C is a convex subset of the topological vector space E, then its interior $\mathring{C}$ is convex. Moreover if $\mathring{C}$ is non-empty, then C is contained in the closure of $\mathring{C}$.*

Proof. Suppose $x \in \mathring{C}$, and $y \in C$. For $t > 0$, the map $H_t : E \to E, z \mapsto tz + (1-t)y$ is a homeomorphism of E. Moreover, for if $0 < t \leq 1$, by convexity of C, we have $H_t(C) \subset C$, therefore $tx + (1-t)y = H_t(x) \in H_t(\mathring{C}) \subset C$. Since $H_t(\mathring{C})$ is open, we obtain that $tx + (1-t)y \in \mathring{C}$, for $0 < t \leq 1$. This implies the convexity of $\mathring{C}$. Now, if $\mathring{C}$ is non-empty, we can find $x_0 \in \mathring{C}$, for $y \in C$, and $0 < t \leq 1$, we know that $tx_0 + (1-t)y \in \mathring{C}$. Since $y = \lim_{t\to 0} tx_0 + (1-t)y$. Therefore C is contained in the closure of $\mathring{C}$. □

Proposition 1.5.6. *Suppose that $f : C \to \mathbb{R}$ is a quasi-convex function defined on the convex subset C of the finite dimensional normed space E. If $(X, \mathcal{A}, \mu)$ is a probability space, and $\varphi : X \to C$ is a measurable function with $\int_X \|\varphi(x)\| \, d\mu(x) < +\infty$, then*

$$f\Big(\int_X \varphi(x)\, d\mu(x)\Big) \leq \sup_{x \in X} f(\varphi(x)).$$

Proof. The set $D = \{c \in C \mid f(c) \leq \sup_{x\in X} f(\varphi(x))\}$ is convex, and, by definition, contains $\varphi(x)$ for every $x \in X$. Therefore by Lemma 1.5.4, we obtain $\int_X \varphi(x)\, d\mu(x) \in D$. □

1.6 Exposed Points of a Convex Set

Let us recall the definition of an extremal point.

Definition 1.6.1 (Extremal Point)**.** A point p in a convex set C is said to be extremal if each time we can write $p = tx + (1-t)y$, with $x, y \in C$ and $t \in [0, 1]$, then $p = x$ or $p = y$.

Theorem 1.6.2 (Krein-Milman)**.** *If K is a convex compact subset of a normed space, then K is the closed convex envelope of its extremal points.*

The proof of the Krein-Milman Theorem can be found in most books on Functional Analysis, see [Bou81, Théorème 1, page II.59], [RV73, Theorem D, page 84] or [Rud91, Theorem 3.23, page 75]

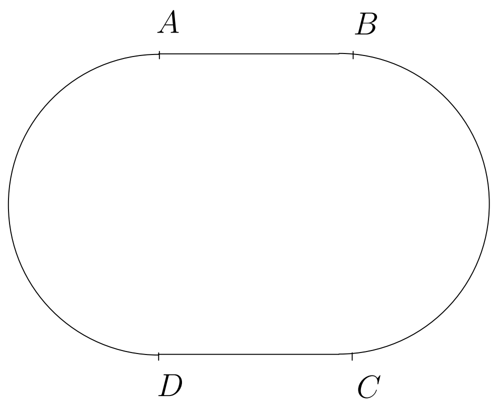

Figure 1.3: Stadium: The four points A, B, C, D are extremal but not exposed.

Let us recall that an affine hyperplane in a $\mathbb{R}$-vector space E, determines two open (resp. closed) half-spaces. If H is the set of points where the affine function $a : E \to \mathbb{R}$ is 0, then the two open (resp. closed) half-spaces are given by $a > 0$ and $a < 0$ (resp. $a \geq 0$ and $a \leq 0$). An hyperplane H is said to be an hyperplane of support of a subset $A \subset E$ if $A \cap H \neq \emptyset$ and A is entirely contained in one of the two closed half-spaces determined by H.

We will need a concept a little finer than that of extremal point, it is the concept of exposed point.

Definition 1.6.3 (Exposed Point)**.** Let C be a convex subset of a normed space. A point p of C is exposed, if there is a hyperplane H of support of C with $H \cap C = \{p\}$.

An exposed point is necessarily an extremal point (exercise). The converse is not necessarily true, as it can be seen on the example of a stadium, see figure 1.3.

Theorem 1.6.4 (Straszewicz)**.** *If C is a convex compact subset of $\mathbb{R}^n$, then C is the closed convex envelope of its exposed points.*

Proof. We will use the Euclidean norm on $\mathbb{R}^n$. Let us denote by C_1 the closure of the convex envelope of the set of the exposed points of C. Let us suppose that there exists $x \in C \setminus C_1$. As

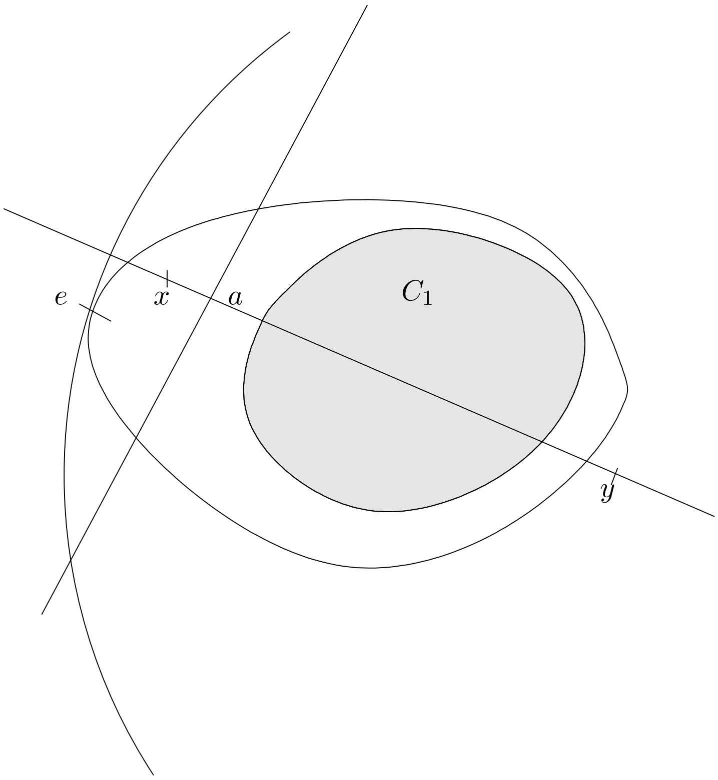

Figure 1.4: Proof of Straszewicz's Theorem.

closed subset of C, the set C_1 is also compact. By the Theorem of Hahn Banach, we can find a hyperplane H strictly separating x from C_1. We consider the line D orthogonal to H and passing through x. We denote by a the intersection $D \cap H$, see figure 1.4. If $c \in \mathbb{R}^n$, we call c_D the orthogonal projections of c on D. We set $d = \sup_{c \in C_1} \|c - c_D\|$. By compactness of C_1, and continuity of $c \mapsto c_D$, this sup d is finite.

Let us fix y a point in D on the same side of H as C_1, we can write

$$\begin{aligned} d(y,c) &= \sqrt{\|y-c_D\|^2 + \|c-c_D\|^2} \\ &\leq \sqrt{\|y-a\|^2 + d^2} \\ &\leq \|y-a\|\sqrt{1+\frac{d^2}{\|y-a\|^2}} \\ &\leq \|y-a\|\left(1+\frac{d^2}{2\|y-a\|^2}\right) \\ &= \|y-a\| + \frac{d^2}{2\|y-a\|}. \end{aligned}$$

Since $x \notin H$, we get $x \neq a$, and $0 < \|a-x\|$. Therefore, since d is finite, for y far away on D so that $d^2/[2\|y-a\|] < \|a-x\|$, we obtain

$$d(y,c) < \|y-a\| + \|a-x\|.$$

But x, a, y are all three on the line D, and a is between x and y, hence $\|y-a\| + \|a-x\| = \|y-x\|$. It follows that for y far away enough

$$\forall c \in C_1,\ d(y,c) < \|y-x\|. \tag{*}$$

Let us then set $R = \sup_{c\in C} d(y,c)$. We have $R \geq \|y-x\|$ because $x \in C$. This supremum R is attained at a point $e \in C$. By $(*)$ we must have $e \notin C_1$. The hyperplane $\tilde{H}$ tangent, at the point e, to the Euclidean sphere $S(y,R) = \{x \in \mathbb{R}^n \mid \|x-y\| = R\}$ is a hyperplane of support for C which cuts C only in e, since $C \subset \bar{B}(y,R)$. Therefore e is an exposed point of C and $e \notin C_1$. This is a contradiction since C_1 contains all the exposed points of C. □

Theorem 1.6.5. *Suppose $L : \mathbb{R}^n \to \mathbb{R}$ is convex and superlinear. We consider the graph of L*

$$\operatorname{Graph} L = \{(x, L(x)) \mid x \in R^n\} \subset \mathbb{R}^n \times \mathbb{R}.$$

Any point of Graph L *belongs to the closed convex envelope of the exposed points of the convex set*

$$\operatorname{Graph}_{\geq} L = \{(x,t) | t \geq L(x)\},$$

formed by the points of $\mathbb{R}^n \times \mathbb{R}$ *which are above* $\operatorname{Graph} L$.

In fact, for each $x \in \mathbb{R}^n$ *we can find a compact subset* $C \subset \mathbb{R}^n \times \mathbb{R}$ *such that* $(x, L(x))$ *is in the closed convex envelope of the exposed points of* $\operatorname{Graph}_{\geq} L$ *which are in* C.

Proof. Let p be a linear form of support for L at x_0. Since $L(x) - L(x_0) \geq p(x - x_0)$, the function $\tilde{L}$ defined by $\tilde{L}(v) = L(v + x_0) - p(v) - L(x_0)$ is ≥ 0 everywhere and takes the value 0 at 0. It is also superlinear, since L is so. Moreover $\operatorname{Graph} \tilde{L}$ is obtained starting from $\operatorname{Graph} L$ using the affine map $(v,t) \mapsto (v - x_0, t - p(v) - L(x_0))$, therefore the exposed points of $\operatorname{Graph} \tilde{L}$ are the images by this affine map of the exposed points of $\operatorname{Graph} L$. From what we have just done without loss of generality, we can assume that $x_0 = 0, L(0) = 0, L \geq 0$, and that we have to show that the point $(0,0)$ is in the closure of the convex envelope of the exposed points of $\operatorname{Graph}_{\geq} L$. For this, we consider the convex subset

$$C = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} \mid L(x) \leq t \leq 1\}.$$

The exposed points of C are either of the form $(x, 1)$ or of the form $(x, L(x))$ with $L(x) < 1$. These last points are also exposed points of $\operatorname{Graph}_{\geq} L$, see Lemma 1.6.6 below. By the superlinearity of L, the subset C is compact. We apply Straszewicz Theorem to conclude that $(0,0)$ is in the closure of the convex envelope of the exposed points of C. We can then gather the exposed points of C of the form $(x, 1)$ and replace them by their convex combination. This allows us to find, for each $n \geq 1$, exposed points of C of the form $(x_{i,n}, L(x_{i,n})), 1 \leq i \leq \ell_n$, with $L(x_{i,n}) < 1$, a point y_n with $(y_n, 1) \in C$, and positive numbers $\alpha_{1,n}, \ldots, \alpha_{\ell_n,n}$ and β_n such that $\beta_n + \sum_{i=1}^{\ell_n} \alpha_{i,n} = 1$ and

$$(0,0) = \lim_{n \to \infty} \beta_n (y_n, 1) + \sum_{i=1}^{\ell_n} \alpha_{i,n} (x_{i,n}, L(x_{i,n})).$$

As $L(x_{i,n}) \geq 0$ we see that $\beta_n \to 0$ and $\sum_{i=1}^{\ell_n} \alpha_{i,n} L(x_{i,n}) \to 0$. It follows that $\alpha_n = \sum_{i=1}^{\ell_n} \alpha_{i,n} \to 1$, since $\alpha_n + \beta_n = 1$. Moreover, since C is compact, the y_n are bounded in norm, therefore $\beta_n y_n \to 0$, because $\beta_n \to 0$. It results from what we obtained above that

$$(0,0) = \lim_{n \to \infty} \sum_{i=1}^{\ell_n} \frac{\alpha_{i,n}}{\alpha^n} (x_{i,n}, L(x_{i,n})).$$

This is the required conclusion, because $\sum_{i=1}^{\ell_n} \alpha_{i,n}/\alpha^n = 1$ and the $(x_{i,n}, L(x_{i,n}))$ are exposed points of $\mathrm{Graph}_{\geq} L$. $\square$

Lemma 1.6.6. *Let C be a convex subset in the topological vector space E. Suppose H is a hyperplane containing $x_0 \in C$. If there exists a neighborhood V of x_0 such that H is a hyperplane of support of $C \cap V$, then H is a hyperplane of support of C.*

Moreover, if $H \cap C \cap V = \{x_0\}$ then x_0 is an exposed point of C.

Proof. This is almost obvious. Call H_+ is a closed half-space determined by H and containing $C \cap V$. If $v \in E$, then the open ray $D_v^+ = \{x_0 + tv \mid t > 0\}$ is either entirely contained in H_+ or disjoint from it. Now if $x \in C$, by convexity of C, for $t \geq 0$ small enough $tx + (1-t)x_0 = x_0 + t(x - x_0) \in C \cap V \subset H_+$, therefore the open ray $D_{x-x_0}^+ \subset H_+$. But $x = x_0 + 1(x - x_0) \in D_{x-x_0}^+$.

Suppose $H \cap C \cap V = \{x_0\}$. If $y \neq x_0$ and $y \in C \cap H$ then the ray $D_{y-x_0}^+$ is contained in H, therefore for every $t \in [0,1]$ small enough $ty + (1-t)x_0 \in H \cap C \cap V$. This is impossible since $ty + (1-t)x_0 \neq x_0$ for $t > 0$. $\square$

Chapter 2

Calculus of Variations

Our treatment of the Calculus of Variations is essentially the classical treatment (up to Tonelli) of the one-dimensional setting in a modern setting. We have mainly used [Cla90], [Mn] and the appendix of [Mat91]. After most of it was typed we learned from Bernard Dacorogna the existence of an excellent recent introduction to the subject [BGH98].

In this chapter, we treat general Lagrangians (i.e. not necessarily convex in the fibers). In the second chapter, we will treat the Lagrangians convex in fibers, therefore all properties concerning existence of minimizing curves will be in next chapter.

2.1 Lagrangian, Action, Minimizers, and Extremal Curves

In this chapter (and the following ones) we will us the standard notations that we have already seen in the introduction, namely:

If M is a manifold (always assumed C^∞, and without boundary), we denote by TM its tangent bundle, and by $\pi : TM \to M$ the canonical projection. A point of TM is denoted by (x, v), where $x \in M$, and $v \in T_xM = \pi^{-1}(x)$. With this notation, we of course have $\pi(x, v) = x$. The cotangent bundle is $\pi^* : T^*M \to M$. A point of T^*M is denoted by (x, p), where $x \in M$, and $p \in T_x^*M = L(T_xM \to \mathbb{R})$.

Definition 2.1.1 (Lagrangian)**.** A Lagrangian L on the manifold M is a continuous function $L : TM \to \mathbb{R}$.

Notice that although L is a function on TM, we will nonetheless say that L is a Lagrangian on M.

Definition 2.1.2 (Action of a Curve)**.** If L is a Lagrangian on the manifold M, and $\gamma : [a,b] \to M$ is a continuous piecewise C^1 curve, with $a \leq b$, the action $\mathbb{L}(\gamma)$ of γ for L is

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

We are interested in curves that minimize the action.

Definition 2.1.3 (Minimizer)**.** Suppose L is a Lagrangian on M. If $\mathcal{C}$ is some set of (parametrized) continuous curves in M, we will say that $\gamma : [a,b] \to M$ is a minimizer for the class $\mathcal{C}$ if for every curve $\delta : [a,b] \to M$, with $\delta(a) = \gamma(a), \delta(b) = \gamma(b)$, and $\delta \in \mathcal{C}$, we have $\mathbb{L}(\gamma) \leq \mathbb{L}(\delta)$.

If $\mathcal{C}$ is the class of continuous piecewise C^1 curves, then minimizers for this class are simply called minimizers.

It should be noticed that to check that $\gamma : [a,b] \to M$ is a minimizer for some class, we only use curves parametrized by the same interval, and with the same endpoints.

In order to find minimizers, we will use differential calculus so that minimizers are to be found among citical points of the action functional $\mathbb{L}$. In section 2.2 we will first treat the linear case, i.e. the case where M is an open subset of $\mathbb{R}^n$. In section 2.3 we will treat the case of a general manifold.

We conclude this section with some definitions that will be used in the following sections.

Definition 2.1.4 (Non-degenerate Lagrangian)**.** If L is a C^2 Lagrangian on the manifold M, we say that L is non-degenerate if for each $(x,v) \in TM$ the second partial derivative $\partial^2 L/\partial v^2(x,v)$ is non-degenerate as a quadratic form.

Notice that the second partial derivative $\partial^2 L/\partial v^2(x,v)$ makes sense. In fact, this is the second derivative of the restriction of L to the vector space T_xM, and defines therefore a quadratic form on T_xM. In the same way, the first derivative $\partial L/\partial v(x,v)$ is a linear form on T_xM, and therefore $\partial L/\partial v(x,v) \in (T_xM)^* = T_x^*M$.

Definition 2.1.5 (Global Legendre Transform)**.** If L is a C^1 Lagrangian on the manifold M, we define the global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ associated to L by

$$\tilde{\mathcal{L}}(x,v) = (x, \frac{\partial L}{\partial v}(x,v)).$$

Of course, if L is C^r, then $\tilde{\mathcal{L}}$ is C^{r-1}.

Proposition 2.1.6. *If L is a C^r Lagrangian, with $r \geq 2$, on the manifold M, then the following statements are equivalent*

(1) *the Lagrangian L is non-degenerate;*

(2) *the global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ is a C^{r-1} local diffeomorphism;*

(3) *the global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ is a C^{r-1} local diffeomorphism.*

Moreover, the following statements are equivalent

(i) *the Lagrangian L is non-degenerate, and $\tilde{\mathcal{L}}$ is injective;*

(ii) *the global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ is a (global) C^{r-1} diffeomorphism onto its image;*

(iii) *the global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ is a (global) C^{r-1} diffeomorphism onto its image.*

Proof. Statements (1), (2), and (3) above are local in nature, it suffices to prove them when M is an open subset of $\mathbb{R}^n$.We use the canonical coordinates on $M \subset \mathbb{R}^n, TM = M \times \mathbb{R}^n$, and $T^*M = M \times \mathbb{R}^{n*}$. In these coordinates , at the point $(x,v) \in TM$, the derivative $D\tilde{\mathcal{L}}(x,v) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^{n*}$ of the global Legendre transform $\tilde{\mathcal{L}}$ has the following matrix form

$$D\tilde{\mathcal{L}}(x,v) = \begin{bmatrix} \mathrm{Id}_{\mathbb{R}^n} & \frac{\partial^2 L}{\partial x \partial v}(x,v) \\ 0 & \frac{\partial^2 L}{\partial v^2}(x,v) \end{bmatrix}$$

therefore $D\tilde{\mathcal{L}}(x,v)$ is invertible as a linear function $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^{n*}$ if and only if $\frac{\partial^2 L}{\partial v^2}(x,v)$ is non-degenerate as a quadratic form. The equivalence of (1), (2), and (3) (resp. (i), (ii), and (iii)) is now a consequence of the inverse function theorem. $\square$

Finally a last definition for this section.

Definition 2.1.7 (C^r Variation of a Curve)**.** Let M be an arbitrary differentiable manifold. Let us consider a C^r curve $\gamma : [a,b] \to M$. A variation of class C^r of γ is a map $\Gamma : [a,b] \times]-\epsilon, \epsilon[\to M$ of class C^r, where $\epsilon > 0$, such that $\Gamma(t,0) = \gamma(t)$, for all $t \in [a,b]$. For such a variation, we will denote by Γ_s the curve $t \mapsto \Gamma(t,s)$ which is also of class C^r.

2.2 Lagrangians on Open Subsets of $\mathbb{R}^n$

We suppose that M is an open subset contained in $\mathbb{R}^n$. In that case $TM = M \times \mathbb{R}^n$, and the canonical projection $\pi : TM \to M$ is the projection on the first factor.

We study the differentiability properties of $\mathbb{L}$, for this we have to assume that L is C^1.

Lemma 2.2.1. *Suppose that L is a* C^1 *Lagrangian the open subset M of $\mathbb{R}^n$. Let $\gamma, \gamma_1 : [a,b] \to \mathbb{R}^n$ be two continuous piecewise* C^1 *curves, with $\gamma([a,b]) \subset M$. The function $\to \mathbb{L}(\gamma + t\gamma_1)$ is defined for t small. It has a derivative at $t = 0$, which is given by*

$$\begin{aligned}\frac{d}{dt}\mathbb{L}(\gamma + t\gamma_1)_{|t=0} &= \int_a^b DL[\gamma(s), \dot\gamma(s)](\gamma_1(s), \dot\gamma_1(s))\, ds \\ &= \int_a^b \left[\frac{\partial L}{\partial x}[\gamma(s), \dot\gamma(s)](\gamma_1(s)) + \frac{\partial L}{\partial v}[\gamma(s), \dot\gamma(s)](\dot\gamma_1(s))\right] ds.\end{aligned}$$

Proof. Both γ, and γ_1 are continuous, hence the map $\Gamma : [a,b] \times \mathbb{R} \to \mathbb{R}^n (s,t) \mapsto \gamma(s) + t\gamma_1(s)$ is continuous, and therefore uniformly continuous on $[a,b] \times [-1,1]$. Since $\Gamma(s,0) = \gamma(s)$ is in the open subset M, for every $s \in [a,b]$, we conclude that there exists $\epsilon > 0$ such that $\Gamma([a,b] \times [-\epsilon, \epsilon]) \subset M$. Therefore the action of the curve $\Gamma(\cdot, t) = \gamma + t\gamma_1$ is defined for every $t \in [-\epsilon, \epsilon]$.

Pick F a finite subset of $[a,b]$ such that both γ, and γ_1 are differentiable at each point of $[a,b]\setminus F$. The function $\lambda : ([a,b]\setminus F)\times\mathbb{R}$ defined by

$$\lambda(s,t) = L(\gamma(s)+t\gamma_1(s), \dot\gamma(s)+t\dot\gamma_1(s)),$$

has a partial derivative with respect to t given by

$$\frac{\partial\lambda}{\partial t}(t,s) = DL[\gamma(s)+t\gamma_1(s), \dot\gamma(s)+t\dot\gamma_1(s)](\gamma_1(s), \dot\gamma_1(s)).$$

Moreover, this partial derivative is uniformly bounded on $([a,b]\setminus F)\times[-1,1]$, because DL is continuous, and the curves γ, γ_1 are continuous, and piecewise C^1. Therefore we can differentiate $\mathbb{L}(\gamma + t\gamma_1) = \int_a^b L(\gamma(s)\dot\gamma(s))\,dt$ under the integral sign to obtain the desired result. $\square$

Exercise 2.2.2. *Suppose that L is a C^1 Lagrangian on the open subset M of $\mathbb{R}^n$. If $\gamma : [a,b]\to M$ is a Lipschitz curve, then $\dot\gamma(s)$ exists almost everywhere. Show that the almost everywhere defined function $s \mapsto L(\gamma(s), \dot\gamma(s))$ is integrable. If $\gamma_1 : [a,b]\to M$ is also Lipschitz, show that $\mathbb{L}(\gamma+t\gamma_1)$ is well defined for t small, finite, and differentiable.*

Definition 2.2.3 (Extremal Curve)**.** An extremal curve for the Lagrangian L is a continuous piecewise C^1 curve $\gamma : [a,b]\to M$ such that $\frac{d}{dt}\mathbb{L}(\gamma+t\gamma_1)_{t=0} = 0$, for every C^∞ curve $\gamma_1 : [a,b]\to\mathbb{R}^n$ satisfying $\gamma_1 = 0$ in the neighborhood of a and b.

By lemma 2.2.1, it is equivalent to say that

$$\int_a^b \left[\frac{\partial L}{\partial x}(\gamma(s),\dot\gamma(s))(\gamma_1(s)) + \frac{\partial L}{\partial v}(\gamma(s),\dot\gamma(s))(\dot\gamma_1(s))\right] ds = 0,$$

for each curve $\gamma_1 : [a,b]\to M$ of class C^∞ which satisfies $\gamma_1 = 0$ in the neighborhood of a and b.

Remark 2.2.4. If γ is an extremal curve, then for all $a', b' \in [a,b]$, with $a' < b'$, the restriction $\gamma|[a',b']$ is also an extremal curve.

The relationship between minimizers and extremal curves is given by the following proposition.

Proposition 2.2.5. *If L is a Lagrangian on M, and $\gamma : [a,b] \to M$ is a C^r curve, with $r \geq 1$ (resp. continuous piecewise C^1) curve, which minimizes the action on the set of C^r (resp. continuous piecewise C^1), then γ is an extremal curve for L.*

Proposition 2.2.6 (Euler-Lagrange)**.** *Let us assume that L is a Lagrangian is of class C^2 on the open subset M of $\mathbb{R}^n$. If $\gamma : [a,b] \to M$ is a curve of class C^2, then γ is an extremal curve if and only if it satisfies the Euler-Lagrange equation*

$$\frac{d}{dt}\frac{\partial L}{\partial v}(\gamma(t),\dot\gamma(t)) = \frac{\partial L}{\partial x}(\gamma(t),\dot\gamma(t)), \tag{E-L}$$

for all $t \in [a,b]$.

Proof. Since L and γ are both C^2, if $\gamma_1 : [a,b] \to M$ is C^∞ and vanishes in the neighborhood of a and b, then the map

$$t \mapsto \frac{\partial L}{\partial v}[\gamma(t),\dot\gamma(t)](\gamma_1(t))$$

is C^1 and is 0 at a and b. It follows that

$$\int_a^b \frac{d}{dt}\left[\frac{\partial L}{\partial v}[\gamma(t),\dot\gamma(t)](\gamma_1(t))\right] dt = 0,$$

which implies

$$\int_a^b \frac{\partial L}{\partial v}[\gamma(t),\dot\gamma(t)](\dot\gamma_1(t))\,dt = -\int_a^b \frac{d}{dt}\left[\frac{\partial L}{\partial v}(\gamma(t),\dot\gamma(t))\right](\gamma_1(t))\,ds.$$

We thus obtain that γ is an extremal curve if and only if

$$\int_a^b \left\{\frac{\partial L}{\partial x}(\gamma(t),\dot\gamma(t)) - \frac{d}{dt}\left[\frac{\partial L}{\partial v}(\gamma(t),\dot\gamma(t))\right]\right\}(\gamma_1(t))\,ds = 0,$$

for every C^∞ curve $\gamma_1 : [a,b] \to M$ satisfying $\gamma_1 = 0$ in the neighborhood of a and b. It is then enough to apply the following lemma:

Lemma 2.2.7 (Dubois-Raymond)**.** *Let $A : [a,b] \to L(\mathbb{R}^n,\mathbb{R}) = \mathbb{R}^{n*}$ be a continuous map such that $\int_a^b A(t)(\gamma_1(t))dt = 0$, for each C^∞ curve $\gamma_1 : [a,b] \to \mathbb{R}^n$ which vanishes in the neighborhood of a and b, then $A(t) = 0$, for all $t \in [a,b]$.*

Proof. Suppose that there exists $t_0 \in]a, b[$ and $v_0 \in \mathbb{R}^n$ such that $A(t_0)(v_0) \neq 0$. Replacing v_0 by $-v_0$ if necessary, we can suppose that $A(t_0)(v_0) > 0$. We fix $\epsilon > 0$ small enough so that $A(t)(v_0) > 0$, for all $t \in [t_0 - \epsilon, t_0 + \epsilon] \subset]a, b[$. We then choose C^∞ curve $\phi : [a, b] \to [0, 1]$ with $\phi = 0$ outside of the interval $[t_0 - \epsilon, t_0 + \epsilon]$ and $\phi(t_0) = 1$. Of course $\int_a^b A(t)(\phi(t)v_0)dt = 0$, but $\int_a^b A(t)(\phi(t)v_0)dt = \int_{t_0-\epsilon}^{t_0+\epsilon} \phi(t)A(t)(v_0)\, dt$ and the function $\phi(t)A(t)(v_0)$ is continuous, non-negative on $[t_0 - \epsilon, t_0 + \epsilon]$ and $\phi(t_0)A(t_0)(v_0) > 0$, since $\phi(t_0) = 1$, hence its integral cannot vanish. This is a contradiction. □

In the remainder of this section, we show that, under natural assumptions on the Lagrangian L, the extremal curves which are C^1 or even continuous piecewise C^1 are necessarily of class C^2, and must thus verify the Euler-Lagrange equation.

Lemma 2.2.8. *Let L be a Lagrangian on the open subset M of $\mathbb{R}^n$, and let $\gamma : [a, b] \to M$ be an extremal curve of class C^1 for L, then there exists $p \in \mathbb{R}^{n*}$ such that*

$$\forall t \in [a, b],\ \frac{\partial L}{\partial v}(\gamma(t), \dot\gamma(t)) = p + \int_a^t \frac{\partial L}{\partial x}(\gamma(s), \dot\gamma(s))\, ds.$$

Proof. If $\gamma_1 : [a, b] \to M$ is C^∞ and vanishes in the neighborhood of a and b, then the map

$$t \mapsto \left[\int_a^t \frac{\partial L}{\partial x}(\gamma(s), \dot\gamma(s))\, ds\right](\gamma_1(t))$$

is C^1 and is 0 at a and b. It follows that

$$\int_a^b \frac{d}{dt}\left\{\left[\int_a^t \frac{\partial L}{\partial x}(\gamma(s), \dot\gamma(s))\, ds\right](\gamma_1(t))\right\} dt = 0,$$

which implies

$$\int_a^b \frac{\partial L}{\partial x}(\gamma(t), \dot\gamma(t))(\gamma_1(t))\, dt = -\int_a^b \left[\int_a^t \frac{\partial L}{\partial x}(\gamma(s), \dot\gamma(s))\, ds\right](\dot\gamma_1(t))\, dt.$$

Thus the condition

$$\int_a^b \left[\frac{\partial L}{\partial x}(\gamma(t), \dot\gamma(t))(\gamma_1(t)) + \frac{\partial L}{\partial v}(\gamma(t), \dot\gamma(t))(\dot\gamma_1(t))\right] dt = 0$$

is equivalent to

$$\int_a^b \left[\frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t)] - \int_a^t \frac{\partial L}{\partial x}(\gamma(s), \dot{\gamma}(s))\, ds\right] (\dot{\gamma}_1(t))\, dt = 0.$$

It is then enough to apply the following lemma:

Lemma 2.2.9 (Erdmann)**.** *If $A : [a,b] \to \mathbb{R}^{n*}$ is a continuous function such that $\int_a^b A(t)(\dot{\gamma}_1(t))\, dt = 0$, for every curve $\gamma_1 : [a,b] \to \mathbb{R}^n$ of C^∞ class and vanishing in the neighborhood of a and b, then, the function $A(t)$ is constant.*

Proof. Let us choose $\phi_0 : [a,b] \to \mathbb{R}$ of class C^∞ with $\int_a^b \phi_0(t)dt = 1$ and $\phi_0 = 0$ in a neighborhood of a and b. Let $\tilde{\gamma}_1 : [a,b] \to \mathbb{R}^n$ be a C^∞ curve which is 0 in the neighborhood of a and b, then, if we set $C = \int_a^b \tilde{\gamma}_1(s)\, ds$, the curve $\gamma_1 : [a,b] \to \mathbb{R}^n$, defined by

$$\gamma_1(t) = \int_a^t \tilde{\gamma}_1(s) - C\phi_0(s)\, ds$$

is C^∞ and is 0 in a neighborhood of a and b. Therefore, since $\dot{\gamma}_1(t) = \tilde{\gamma}_1(t) - C\phi_0(t)$, we have

$$\int_a^b A(t)[\tilde{\gamma}_1(t) - C\phi_0(t)]\, dt = 0,$$

consequently

$$\int_a^b A(t)[\tilde{\gamma}_1(t)]dt - \int_a^b [\phi_0(t)A(t)](C)dt = 0. \qquad (*)$$

If we define $p = \int_a^b \phi_0(t)A(t)dt \in \mathbb{R}^{n*}$, then $\int_a^b [\phi_0(t)A(t))](C)\, dt$ is nothing but $p(C)$. On the other hand by the definition of C and the linearity of p, we have that $p(C) = \int_a^b p(\tilde{\gamma}_1(t))dt$. We then can rewrite the equation $(*)$ as

$$\int_a^b [A(t) - p](\tilde{\gamma}_1(t))\, dt = 0.$$

Since $\tilde{\gamma}_1 : [a,b] \to \mathbb{R}^n$ is a map which is subject only to the two conditions of being C^∞ and equal to 0 in a neighborhood of a and b, Dubois-Raymond's Lemma 2.2.7 shows that $A(t) - p = 0$, for all $t \in [a,b]$. $\square$

Remark 2.2.10. Of course, the proofs of both the Dubois-Raymond and the Erdmann lemmas are reminiscent of now classical proofs of analogous statements in Laurent Schwartz's Theory of Distributions, but these statements are much older.

Corollary 2.2.11. *If L is a non-degenerate, C^r Lagrangian, with $r \geq 2$, on the open subset M of $\mathbb{R}^n$, then every extremal C^1 curve is C^r.*

Proof. Let $\gamma : [a, b] \to M$ be a C^1 extremal curve. Let us fix t_0 and consider $(x_0, v_0) = (\gamma(t_0), \dot\gamma(t_0)) \in TM$. From proposition 2.1.6, the Legendre transform $\tilde{\mathcal{L}} : (x, v) \to (x, \frac{\partial L}{\partial v}(x, v))$ is a local diffeomorphism Let us call $\mathcal{K}$ a local inverse of $\tilde{\mathcal{L}}$ with $\mathcal{K}(x_0, \frac{\partial L}{\partial v}(x_0, v_0)) = (x_0, v_0)$. The map $\mathcal{K}$ is of class C^{r-1}. By continuity of γ and $\dot\gamma$, for t near to t_0, we have

$$(\gamma(t), \dot\gamma(t)) = \mathcal{K}\big[\gamma(t), \frac{\partial L}{\partial v}(\gamma(t), \dot\gamma(t))\big]. \tag{$*$}$$

But, by lemma 2.2.8, we have

$$\frac{\partial L}{\partial v}(\gamma(t), \dot\gamma(t)) = p + \int_a^t \frac{\partial L}{\partial x}(\gamma(s), \dot\gamma(s))\, ds.$$

It is clear that the right-hand side of this equality is of class C^1. Referring to $(*)$ above, as $\mathcal{K}$ is C^{r-1}, we see that $(\gamma(t), \dot\gamma(t))$ is also of class C^1, for t near to t_0. We therefore conclude that γ is C^2. By induction, using again $(*)$, we see that γ is C^r. □

Corollary 2.2.12. *Suppose that the Lagrangian L on M is of class C^r, $r \geq 2$, and that its global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ is a diffeomorphism on its image $\tilde{\mathcal{L}}(TM)$), then every continuous piecewise C^1 extremal curve of L is in fact C^r, and must therefore satisfy the Euler-Lagrange equation.*

Proof. The assumption that $\tilde{\mathcal{L}}$ is a diffeomorphism implies by proposition 2.1.6 that L is non-degenerate. Therefore by corollary 2.2.11, we already know that the extremal C^1 curves are all C^r. Let $\gamma : [a, b] \to M$ be a continuous piecewise C^1 extremal curve. Let us consider a finite subdivision $a_0 = a < a_1 < \cdots < a_n = b$ such that the restriction $\gamma|[a_i, a_{i+1}]$ is C^1. Since $\gamma|[a_i, a_{i+1}]$ is also an extremal curve, we already know that $\gamma|[a_i, a_{i+1}]$ is C^r. Using that γ is an extremal curve, for every C^∞ curve $\gamma_1 : [a, b] \to M$ which is equal to 0 in a neighborhood of a and b, we have

$$\int_a^b \frac{\partial L}{\partial x}[\gamma(t), \dot{\gamma}(t)](\gamma_1(t)) + \frac{\partial L}{\partial v}[\gamma(t), \dot{\gamma}(t)](\dot{\gamma}_1(t))\, dt = 0. \qquad (*)$$

Since $\gamma|[a_i, a_{i+1}]$ is of class at least C^2, we can integrate by parts to obtain

$$\int_{a_i}^{a_{i+1}} \frac{\partial L}{\partial v}[\gamma(t), \dot{\gamma}(t)](\dot{\gamma}_1(t))dt = \frac{\partial L}{\partial v}[\gamma(a_{i+1}), \dot{\gamma}_-(a_{i+1})](\gamma_1(a_{i+1}))$$
$$-\frac{\partial L}{\partial v}[\gamma(a_i), \dot{\gamma}_+(a_i)](\gamma_1(a_i)) - \int_{a_i}^{a_{i+1}} \left\{ \frac{d}{dt}\left[\frac{\partial L}{\partial v}[\gamma(s), \dot{\gamma}(s)]\right]\right\}(\gamma_1(s))\, ds,$$

where $\dot{\gamma}_-(t)$ is the left derivative and $\dot{\gamma}_+(t)$ is the right derivative of γ at $t \in [a, b]$. Using this, we conclude that

$$\begin{aligned}
&\int_{a_i}^{a_{i+1}} \frac{\partial L}{\partial x}[\gamma(t), \dot{\gamma}(t)](\gamma_1(t)) + \frac{\partial L}{\partial v}[\gamma(t), \dot{\gamma}(t)](\dot{\gamma}_1(t))\, dt = \\
&\int_{a_i}^{a_{i+1}} \frac{\partial L}{\partial x}[\gamma(t), \dot{\gamma}(t)](\gamma_1(t)) - \left\{ \frac{d}{dt}\left[\frac{\partial L}{\partial v}[\gamma(s), \dot{\gamma}(s)]\right]\right\}(\gamma_1(s))\, ds \\
&\quad + \frac{\partial L}{\partial v}[\gamma(a_{i+1}), \dot{\gamma}_-(a_{i+1})](\gamma_1(a_i)) - \frac{\partial L}{\partial v}[\gamma(a_i), \dot{\gamma}_+(a_i)](\gamma_1(a_i)) \\
&\quad = \frac{\partial L}{\partial v}[\gamma(a_{i+1}), \dot{\gamma}_-(a_{i+1})](\gamma_1(a_i)) - \frac{\partial L}{\partial v}[\gamma(a_i), \dot{\gamma}_+(a_i)](\gamma_1(a_i))
\end{aligned}$$

where the last equality holds, because $\gamma|[a_i, a_{i+1}]$ is a C^2 extremal curve, and therefore must satisfy the Euler-Lagrange equation (E-L) on the interval $\gamma|[a_i, a_{i+1}]$. Summing on i, and using $(*)$, we get that

$$\sum_{i=1}^{n-1} \left[\frac{\partial L}{\partial v}[\gamma(a_{i+1}), \dot{\gamma}_-(a_{i+1})](\gamma_1(a_i)) - \frac{\partial L}{\partial v}[\gamma(a_i), \dot{\gamma}_+(a_i)](\gamma_1(a_i))\right] = 0,$$

for every C^∞ curve $\gamma_1 : [a, b] \to M$ which vanishes in a neighborhood of a and b. For $1 \le i \le n-1$, we can choose the C^∞ curve γ_1, vanishing in a neighborhood of the union of the two intervals $[a, a_{i-1}]$ and $[a_{i+1}, b]$, and taking at a_i an arbitrary value fixed in advance. This implies that

$$\forall i = 1, \ldots, N,\ \frac{\partial L}{\partial v}(\gamma(a_i), \dot{\gamma}_-(a_i)) = \frac{\partial L}{\partial v}(\gamma(a_i), \dot{\gamma}_+(a_i)).$$

The injectivity of the Legendre transform gives $\dot\gamma_-(a_i) = \dot\gamma_+(a_i)$. Hence, the curve γ is in fact of class C^1 on $[a,b]$ and, consequently, it is also of class C^r. □

The proof of following lemma is essentially the same as that of lemma 2.2.1.

Lemma 2.2.13. *If Γ is a C^2 variation of the C^2 curve $\gamma : [a,b] \to M$ with values in the open subset M of $\mathbb{R}^n$, then the map $s \mapsto \mathbb{L}(\Gamma_s)$ is differentiable and its derivative in 0 is*

$$\frac{d}{ds}\mathbb{L}(\Gamma_s)_{s=0} = \int_a^b DL[(\gamma(t),\dot\gamma(t))]\big(\frac{\partial\Gamma}{\partial s}(t,0), \frac{\partial^2\Gamma}{\partial s\partial t}(t,0)\big)\,dt.$$

We now obtain a characterization of extremal curves that does not use the fact that M is contained in an open subset of an Euclidean space.

Lemma 2.2.14. *A C^2 curve $\gamma : [a,b] \to M$ is an extremal curve of the Lagrangian L if and only if $\frac{d}{ds}\mathbb{L}(\Gamma_s)_{s=0} = 0$, for any C^2 variation Γ of γ such that $\Gamma(a,s) = \gamma(a), \Gamma(b,s) = \gamma(b)$ for s in a neighborhood of 0.*

Proof. The variations of the type $\gamma(t) + b\gamma_1(t)$ with γ_1 of C^∞ class are particular variations of class C^2.

It thus remains to be seen that, if γ is a C^2 extremal curve, then we have $\frac{d}{ds}\mathbb{L}(\Gamma_s)_{s=0} = 0$ for variations of class C^2 of γ such that $\Gamma(a,s) = \gamma(a)$ and $\Gamma(b,s) = \gamma(b)$. That results from the following theorem.

Theorem 2.2.15 (First Variation Formula)**.** *If $\gamma : [a,b] \to M$ is a C^2 extremal curve, then for any variation Γ of class C^2 of γ, we have*

$$\frac{d}{ds}\mathbb{L}(\Gamma_s)_{s=0} = \frac{\partial L}{\partial v}[\gamma(b),\dot\gamma(b)]\big(\frac{\partial\Gamma}{\partial s}(b,0)\big) - \frac{\partial L}{\partial v}[\gamma(a),\dot\gamma(a)]\big(\frac{\partial\Gamma}{\partial s}(a,0)\big).$$

Proof. We have

$$\begin{aligned}\frac{d}{ds}\mathbb{L}(\Gamma_s)_{s=0} &= \int_a^b DL[\gamma(t),\dot\gamma(t)]\big(\frac{\partial\Gamma}{\partial s}(t,0), \frac{\partial^2\Gamma}{\partial t\partial s}(t,0)\big)\,dt\\ &= \int_a^b \left[\frac{\partial L}{\partial x}[\gamma(t),\dot\gamma(t)]\big(\frac{\partial\Gamma}{\partial s}(t,0)\big) + \frac{\partial L}{\partial v}[\gamma(t),\dot\gamma(t)]\big(\frac{\partial^2\Gamma}{\partial t\partial s}(t,0)\big)\right] dt.\end{aligned}$$

As γ is a C^2 extremal curve, it satisfies the Euler-Lagrange equation

$$\frac{\partial L}{\partial x}[\gamma(t), \dot{\gamma}(t)] = \frac{d}{dt}\left[\frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t))\right],$$

plugging in, we find

$$\frac{d}{ds}\mathbb{L}(\Gamma_s)_{s=0} =$$
$$\int_a^b \left[\frac{d}{dt}\left[\frac{\partial L}{\partial v}\right](\gamma(t), \dot{\gamma}(t))\right]\big(\frac{\partial \Gamma}{\partial s}(t,0)\big) + \frac{\partial L}{\partial v}[\gamma(t), \dot{\gamma}(t)]\big(\frac{\partial^2 \Gamma}{\partial t \partial s}(t,0)\big)\,dt.$$

However the quantity under the last integral is nothing but the derivative of the function $t \mapsto \frac{\partial L}{\partial v}[\gamma(t), \dot{\gamma}(t)](\frac{\partial \Gamma}{\partial s}(t,0))$ which is of class C^1 thus

$$\frac{d}{ds}\mathbb{L}(\Gamma_s)_{s=0} = \frac{\partial L}{\partial v}[\gamma(b), \dot{\gamma}(b)]\big(\frac{\partial \Gamma}{\partial s}(b,0)\big) - \frac{\partial L}{\partial v}[\gamma(a), \dot{\gamma}(a)]\big(\frac{\partial \Gamma}{\partial s}(a,0)\big).$$

□

2.3 Lagrangians on Manifolds

We consider an arbitrary C^∞ manifold M endowed with a Lagrangian L of class C^r, with $r \geq 2$.

Lemma 2.3.1. *Consider* $\Gamma : [a,b] \times [c,d] \to M$ *of class* C^2*. Define* $\Gamma_s : [a,b] \to M$ *by* $\Gamma_s(t) = \Gamma(t,s)$*. The map* $s \mapsto \mathbb{L}(\Gamma_s)$ *is* C^1*.*

Proof. To simplify notation we assume that $0 \in [c,d]$ and we show that $s \mapsto \mathbb{L}(\Gamma_s)$ is C^1 on some interval $[-\eta, \eta]$, with $\eta > 0$. We can cover the compact set $\Gamma([a,b] \times \{0\})$ by a finite family of coordinate charts. We then find a subdivision $a_0 = a < a_1 < \cdots < a_n = b$ such that $\Gamma([a_i, a_{i+1}] \times \{0\}))$ is contained in a U_i the domain of definition of one of these charts. By compactness, for η small enough, we have $\Gamma([a_i, a_{i+1}] \times [-\eta, \eta]) \subset U_i$, for $i = 0, \dots, n-1$. Transporting the situation via the chart to an open set in $\mathbb{R}^n$, we can apply 2.2.13 to obtain that $s \mapsto \int_{a_i}^{a_{i+1}} L[\Gamma(t,s), \frac{\partial \Gamma}{\partial t}(t,s)]\,dt$ is C^1 on some interval $[-\eta, \eta]$. It is now enough to notice that

$$\mathbb{L}(\Gamma_s) = \sum_{i=0}^{n-1} \int_{a_i}^{a_{i+1}} L\big[\Gamma(s,t), \frac{\partial \Gamma}{\partial t}(s,t)\big]\,dt,$$

to be able to finish the proof. □

We can then introduce the concept of an extremal curve for the C^2 Lagrangian L in the case of C^2 curves $\gamma : [a, b] \to M$ with values in arbitrary manifold M.

Definition 2.3.2 (Extremal C^2 Curve)**.** A C^2 curve $\gamma : [a, b] \to M$ is an extremal curve for the C^2 Lagrangian L, if for each C^2 variation $\Gamma : [a, b] \times] - \epsilon, \epsilon[\to M$ of γ, with $\Gamma(t, s) = \gamma(t)$ in the neighborhood of $(a, 0)$ and $(b, 0)$, we have $\frac{d}{ds}\mathbb{L}(\Gamma_s)_{s=0} = 0$.

Remark 2.3.3. By lemma 2.2.14, if the curve and the Lagrangian are of class C^2, this definition coincides with the definition given for the case where the manifold is an open subset of $\mathbb{R}^n$.

Lemma 2.3.4. *If $\gamma : [a, b] \to M$ is a C^2 extremal curve and $[a', b'] \subset [a, b]$ then, the restriction $\gamma|[a', b']$ is also an extremal curve*

Proof. For any C^2 variation $\Gamma : [a', b'] \times] - \epsilon, \epsilon[\to M$ of $\gamma|[a', b']$, with $\Gamma(t, s) = \gamma(t)$ in the neighborhood of $(a', 0)$ and $(b', 0)$, we find ϵ' with $0 < \epsilon' \le \epsilon$, and $\delta > 0$ such that with $\Gamma(t, s) = \gamma(t)$ for every $(t, s) \in ([a', a' + \delta] \cup [b' - \delta, b']) \times [-\epsilon', \epsilon']$. We can therefore extend $\Gamma|[a', b'] \times [-\epsilon', \epsilon']$ to $\tilde{\Gamma}[a, b] \times [-\epsilon', \epsilon']$ by $\Gamma(t, s) = \gamma(t)$, for $t \notin [a', b'] \times [-\epsilon', \epsilon']$. It is clear that $\tilde{\Gamma}$ is a C^2 variation, with $\Gamma(t, s) = \gamma(t)$ in the neighborhood of $(a, 0)$ and $(b, 0)$. Moreover, for $s \in [-\epsilon', \epsilon']$, the difference $\mathbb{L}(\tilde{\Gamma}_s) - \mathbb{L}(\Gamma_s)$ is equal to $\mathbb{L}(\gamma|[a, a']) + \mathbb{L}(\gamma|[b', b])$. ◻

Theorem 2.3.5 (Euler-Lagrange)**.** *Suppose L is a C^2 Lagrangian on the manifold M. Let $\gamma : [a, b] \to M$, be a C^2 curve. If γ is extremal, then, for each subinterval $[a', b'] \subset [a, b]$ such that $\gamma([a', b'])$ is contained in a domain U of a coordinate chart, the restriction $\gamma|[a', b']$ satisfies (in coordinates) the Euler-Lagrange equation.*

Conversely, if for every $t_0 \in [a, b]$, we can find an $\epsilon > 0$ and a domain U of a coordinate chart such that $\gamma([t_0 - \epsilon, t_0 + \epsilon] \cap [a, b]) \subset U$ and $\gamma|[t_0 - \epsilon, t_0 + \epsilon] \cap [a, b]$ satisfies in the chart the Euler-Lagrange equation, then the curve γ is an extremal curve.

Proof. If γ is an extremal curve, then $\gamma|[a', b']$ is also an extremal curve. Since $\gamma([a', b']) \subset U$, where U is a domain of a coordinate chart, we can then transport, via the coordinate chart, the situation to an open subset of $\mathbb{R}^n$ hence $\gamma|[a', b']$ must verify the Euler-Lagrange equation.

To prove the second part, we remark that by compactness s, we can find a subdivision $a_0 = a < a_1 < \cdots < a_n = b$, and a sequence $U_0, \ldots, U_{n-1}$ of domains of coordinate charts such that $\gamma([a_i, a_{i+1}]) \subset U_i$. If Γ is a variation of class C^2 of γ, we can find $\eta > 0$ such that $\Gamma([a_i, a_{i+1}] \times [-\eta, \eta]) \subset U_i, i = 1, \cdots, n)$. The first variation formula 2.2.15 shows that

$$\frac{d}{ds}\mathbb{L}(\Gamma_s|[a_i, a_{i+1}])_{s=0} = \\ \frac{\partial L}{\partial v}[\gamma(a_{i+1}), \dot\gamma(a_{i+1})]\big(\frac{\partial \Gamma}{\partial s}(a_{i+1})\big) - \frac{\partial L}{\partial v}[\gamma(a_i), \dot\gamma(a_i)]\big(\frac{\partial \Gamma}{\partial s}(a_i)\big).$$

Adding these equalities, we find

$$\frac{d}{ds}\mathbb{L}(\Gamma_s)_{s=0} = \frac{\partial L}{\partial v}[\gamma(b), \dot\gamma(b)]\big(\frac{\partial \Gamma}{\partial s}(b,0)\big) - \frac{\partial L}{\partial v}[\gamma(a), \dot\gamma(a)]\big(\frac{\partial \Gamma}{\partial s}(a,0)\big).$$

If $\Gamma(a, s) = \gamma(a)$ and $\Gamma(s, b) = \gamma(b)$ in a neighborhood of $s = 0$, we get that both $\frac{\partial \Gamma}{\partial s}(a, 0)$, and $\frac{\partial \Gamma}{\partial s}(b, 0)$ are equal to 0,therefore the second member of the equality above is 0. □

The previous proof also shows that the first variation formula is valid in the case of arbitrary manifolds.

Theorem 2.3.6 (First Variation Formula)**.** *Let L be a C^2 Lagrangian on the manifold M. If $\gamma : [a, b] \to M$ is a C^2 extremal curve, for each C^2 variation $\Gamma : [a, b]\times] - \epsilon, \epsilon[\to M, (t, s) \mapsto \Gamma(t, s)$ of γ, we have*

$$\frac{d}{ds}\mathbb{L}(\Gamma_s)_{s=0} = \frac{\partial L}{\partial v}[\,(\gamma(b), \dot\gamma(b)]\big(\frac{\partial \Gamma}{\partial s}(b,0)\big) - \frac{\partial L}{\partial v}[\gamma(a), \dot\gamma(a)]\big(\frac{\partial \Gamma}{\partial s}(a,0)\big).$$

By the same type chart by chart argument, using proposition 2.2.12, we can show the following proposition.

Proposition 2.3.7. *Suppose the C^r Lagrangian L, with $r \geq 2$, on the manifold M is such that its global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ is a diffeomorphism onto its image. If $\gamma : [a, b] \to M$ is a C^k curve, with $k \geq 1$, (resp. a continuous piecewise C^1 curve) which is a minimizer for the class of C^k curves (resp. of continuous piecewise C^1 curves), then γ is an extremal of class at least C^r.*

2.4 The Euler-Lagrange Equation and its Flow

We will first consider an open subset M of $\mathbb{R}^n$ and a non-degenerate C^r Lagrangian $L : TM \to \mathbb{R}$, with $r \geq 2$. We will assume that for each $(x,v) \in TM$ the bilinear form $\partial^2 L/\partial v^2(x,v)$ is non-degenerate. It follows from 2.2.11 that every C^1 (locally) extremal curve $\gamma : [a,b] \to M$ is necessarily of class C^2, and satisfies the Euler-Lagrange equation

$$\frac{\partial L}{\partial x}(\gamma(t),\dot\gamma(t)) = \frac{d}{dt}\frac{\partial L}{\partial v}(\gamma(t),\dot\gamma(t)),$$

and hence by differentiation we obtain

$$\frac{\partial^2 L}{\partial v^2}[\gamma(t),\dot\gamma(t)](\ddot\gamma(t),\cdot) = \frac{\partial L}{\partial x}[\gamma(t),\dot\gamma(t)](\cdot) - \frac{\partial^2 L}{\partial x \partial v}[\gamma(t),\dot\gamma(t)](\dot\gamma(t),\cdot),$$

where this is to be understood as an equality between elements of $\mathbb{R}^{n*}$. Since $\partial^2 L/\partial v^2(x,v)$ is non-degenerate, we can in fact solve for $\ddot\gamma(t)$, and therefore we see that γ satisfies a second order differential equation. This suggests to define a vector field X_L on $TM = M \times \mathbb{R}^n$ by

$$X_L(x,v) = (v, \tilde X_L(x,v)) \in T_{(x,v)}(TM),$$

where, due to the non-degeneracy of $\partial^2 L/\partial v^2(x,v)$, the function $\tilde X_L$ is uniquely defined by

$$\frac{\partial^2 L}{\partial v^2}(x,v)[\tilde X_L(x,v),\cdot] = \frac{\partial L}{\partial x}(x,v)(\cdot) - \frac{\partial^2 L}{\partial x \partial v}(x,v)(v,\cdot).$$

This function $\tilde X_L$ is C^{r-2}, if L is C^r.

From our previous computation, if γ is a curve satisfying the Euler-Lagrange equation, its speed curve $t \mapsto (\gamma(t),\dot\gamma(t))$ is an integral curve of X_L. Conversely, since the first of the coordinates of $X_L(x,v), (v,\tilde X_L(x,v))$ is v, the solutions of this vector field are curves of the form $t \mapsto (\gamma(t),\dot\gamma(t))$ with $\gamma : [a,b] \to M$ of class C^2, and $\ddot\gamma(t) = \tilde X_L(\gamma(t),\dot\gamma(t))$, and therefore γ satisfies the Euler-Lagrange equation. Thus, these integral curves are the curves of the form $t \mapsto (\gamma(t),\dot\gamma(t))$ with $\gamma : [a,b] \to M$ of class C^2 which satisfy the Euler-Lagrange equation, in other words with γ an extremal curve.

Theorem 2.4.1. *Let L be a C^r Lagrangian on M, with $r \geq 2$. Assume that for each $(x, v) \in TM$ the bilinear form $\partial^2 L/\partial v^2(x, v)$ is non-degenerate. Then for every $(x, v) \in TM$, we can find an extremal $\gamma : [-\epsilon, \epsilon] \to M$ with $\gamma(0) = x$, and $\dot{\gamma}(0) = v$; moreover, if L is at least C^3, then any two such extremals coincide on their common domain of definition.*

Proof. Suppose $(x, v) \in TM$ is given. Since X_L is at least continuous, we can apply the Cauchy-Peano theorem, see [Bou76], to find an integral curve Γ of X_L defined on some interval $[-\epsilon, \epsilon]$, with $\epsilon > 0$ and passing through (x, v) at time $t = 0$. But as we have seen above such a solution if of the form $\Gamma(t) = (\gamma(t), \dot{\gamma}(t))$, with γ an extremal. This γ is obviously the required extremal.

If L is C^3, then X_L is C^1, and we therefore have uniqueness of solution by the Cauchy-Lipschitz theorem. Therefore, if $\gamma_1 : [[-\epsilon', \epsilon'] \to M$ is another extremal with $\gamma_1(0) = x$, and $\dot{\gamma}_1(0) = v$, then $\Gamma_1(t) = (\gamma_1(t), \dot{\gamma}_1(t))$ is another solution of X_L with the same initial condition as Γ, therefore $\Gamma = \Gamma'$ on the intersection of their domain of definition. □

We can, then, summarize what we obtained in the following theorem.

Theorem 2.4.2 (Euler-Lagrange). *Let M be an open subset of $\mathbb{R}^n$. If L is a Lagrangian on M of class C^r, with $r \geq 2$, and for every $(x, v) \in TM$ the quadratic form $\frac{\partial^2 L}{\partial v^2}(x, v))$ is non-degenerate, then there exists one and only one vector field X_L on TM such that the solutions of X_L are precisely the curves the form $t \mapsto (\gamma(t), \dot{\gamma}(t))$ where $\gamma : [a, b] \to M$ is an extremal curve of L. This vector field is of class C^{r-2}. The vector field X_L is called the Euler-Lagrange vector field of the Lagrangian L.*

Everything in this theorem was proved above, but maybe we should say a word about the uniqueness. In fact, we can obtain $X_L(x, v)$ from the extremals of L. In fact, if we choose an extremal $\gamma : [-\epsilon, \epsilon] \to M$ with $\gamma(0) = x$, and $\dot{\gamma}(0) = v$ (this is possible by theorem 2.4.1), then $X_L(x, v)$ is nothing but the speed at 0 of the curve $t \mapsto (\gamma(t), \dot{\gamma}(t))$.

Let us extend this result to the case of an arbitrary manifold M.

Theorem 2.4.3 (Euler-Lagrange)**.** *Let M be a differentiable manifold. If $L : TM \to \mathbb{R}$ is a C^r Lagrangian, with $r \geq 2$, and for every (x, v)*
inTM the quadratic form $\frac{\partial^2 L}{\partial v^2}(x, v))$ is non-degenerate, then there exists one and only one vector field X_L on TM such that the solutions of X_L are precisely of the form $t \mapsto (\gamma(t), \dot{\gamma}(t))$ where $\gamma : [a, b] \to M$ is an extremal curve of L. This vector field is of class C^{r-2}.

Proof. By theorem 2.4.2 above, for every open subset $U \subset M$ which is contained in the domain of a coordinate chart, we can find such a vector field X_L^U on TU for which the solutions are precisely of the form $t \mapsto (\gamma(t), \dot{\gamma}(t))$ where $\gamma : [a, b] \to U$ with γ an extremal curve having values in U. But if U and V are two such open subsets, for both restrictions $X_L^U | U \cap V, X_L^U | U \cap V$, the solution curves are the curves $t \mapsto (\gamma(t), \dot{\gamma}(t))$ where γ is an extremal curve of L whose image is contained in $U \cap V$, so they both coincide with $X_L^{U \cap V}$. □

Definition 2.4.4 (The Euler-Lagrange Vector Field and its Flow)**.** If $L : TM \to \mathbb{R}$ is a C^r Lagrangian, with $r \geq 2$, and for every $(x, v) \in TM$ the quadratic form $\frac{\partial^2 L}{\partial v^2}(x, v))$ is non-degenerate, the vector field X_L defined by theorem 2.4.3 above is called the Euler-Lagrange vector field of the Lagrangian L. If L is C^r with $r \geq 3$, then, by the Cauchy-Lipschitz theorem, the field X_L generates a partial flow on TM of class C^{r-2}. We will denote this partial flow by ϕ_t^L and we will call it the Euler-Lagrange flow of the Lagrangian L.

In fact, under the stronger hypothesis that the global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$, defined by

$$\tilde{\mathcal{L}}(x, v)(x, v) = \big(x, \frac{\partial L}{\partial v}(x, v)\big),$$

is a C^1 diffeomorphism onto its image, we will see in Theorem 2.6.5 that the partial flow is of class C^{r-1} and that, even if L is only of class C^2, the vector field X_L is uniquely integrable and generates a partial flow ϕ_t^L of class C^1 which is also called in that case the Euler-Lagrange flow of the Lagrangian L.

2.5 Symplectic Aspects

Let M be differentiable manifold of class C^∞, denote by $\pi^* : T^*M \to M$ the canonical projection of the cotangent space T^*M onto M and denote by $T\pi^* : TT^*M \to TM$ the derivative of π^*. On T^*M, we can define a canonical differential 1-form α called the Liouville form. Thus the value this form at a given $(x,p) \in T^*M$ is a linear map $\alpha_{(x,p)} : T_{(x,p)}(T^*M) \to \mathbb{R}$. To define it, we just remark that the two linear maps $T_{(x,p)}\pi^* : T_{(x,p)}(T^*M) \to T_xM$, and $p : T_xM \to \mathbb{R}$ can be composed, and thus we can define $\alpha_{(x,p)}$ by

$$\forall W \in T_{(x,p)}(T^*M), \alpha_{(x,p)}(W) = p[T_{(x,p)}\pi^*(W)].$$

To understand this differential 1-form α on the T^*M manifold, we take a chart $\theta : U \to \theta(U) \subset \mathbb{R}^n$, we can consider the associated chart $T^*\theta : T^*U \to T^*(\theta(U)) = \theta(U) \times \mathbb{R}^{n*}$. In these charts the canonical projection π^* is nothing but the projection $\theta(U) \times \mathbb{R}^{n*} \to \theta(U)$ on the first factor. This gives us coordinates $(x_1, \ldots, x_n)$ on U, and therefore coordinates $(x_1, \cdots, x_n, p_1, \cdots, p_n)$ on T^*U such that the projection π^* is nothing but $(x_1, \cdots, x_n, p_1, \cdots, p_n) \to (x_1, \cdots, x_n)$. A vector $W \in T_{(x,p)}(T^*M)$ has therefore coordinates $(X_1, \cdots, X_n, P_1, \cdots, P_n)$, and the coordinates of $T_{(x,p)}\pi^*(W) \in T_xU$ are $(X_1, \cdots, X_n)$. It follows that $\alpha_{(x,p)}(W) = \sum_{i=1}^n p_i X_i$. Since X_i is nothing but the differential form dx_i evaluated on W, we get that

$$\alpha | T^*U = \sum_{i=1}^n p_i \, dx_i.$$

We therefore conclude that α is of class C^∞.

Let ω be a differential 1-form defined on the open subset U of M. This 1-form is a section $\omega : U \to T^*U, x \mapsto \omega_x$. The graph of ω is the set

$$\mathrm{Graph}(\omega) = \{(x, \omega_x) \mid x \in U\} \subset T^*M.$$

Lemma 2.5.1. *Let ω be a differential 1-form defined on the open subset U of M. We have*

$$\omega^*\alpha = \omega,$$

where $\omega^*\alpha$ is the pull-back of the Liouville form α on T^*M by the map $\omega : U \to T^*U$

Proof. Using coordinates charts, it suffices to verify in the case where U is an open subset of $\mathbb{R}^n$. Using the canonical coordinates on $U \subset \mathbb{R}^n$, we can write $\omega_x = \sum_{i=1}^n \omega_i(x)dx_i$. As a map $\omega : U \to T^*U$ it is thus given in these coordinates by

$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, \omega_1(x), \dots, \omega_n(x)).$$

But in these coordinates, it is clear that the pull-back $\omega^*\alpha$ is $\sum_{i=1}^n \omega_i(x)dx_i = \omega$. □

By taking, the exterior derivative Ω of α we can define a symplectic structure on T^*M. To explain what that means, let us recall that a symplectic form on a vector space E is an alternate (or antisymmetric) bilinear form $a : E \times E \to \mathbb{R}$ which is non-degenerate as a bilinear form, i.e. the map $a^\sharp : E \to E^*, x \mapsto a(x, \cdot)$ is an isomorphism.

Lemma 2.5.2. *If the finite dimensional vector space E admits a symplectic form, then its dimension is even.*

Proof. We choose a basis on E. If A is the matrix of a in this base, its transpose ${}^tA$ is equal to $-A$ (this reflects the antisymmetry). Therefore taking determinants, we get $\det(A) = \det({}^tA) = \det(-A) = -1^{\dim E}\det(A)$. The matrix of $a^\sharp : E \to E^*$, using on E^* the dual basis, is also A. the non degeneracy of $a^\sharp$ gives $\det(A) \neq 0$. It follows that $-1^{\dim E} = 1$, and therefore $\dim E$ is even. □

Definition 2.5.3 (Symplectic Structure)**.** A symplectic structure on a C^∞ differentiable manifold V is a C^∞ *closed* differential 2-form Ω on V such that, for each $x \in V$, the bilinear form $\Omega_x : T_xV \times T_xV \to \mathbb{R}$ is a symplectic form on the vector space T_xV.

As an exterior derivative is closed, to check that $\Omega = -d\alpha$ is a symplectic form on T^*M, it is enough to check the non-degeneracy condition. We have to do it only in an open subset U of $\mathbb{R}^n$, with the notations introduced higher, we see that

$$\Omega = -d\alpha = -\sum_{i=1}^n dp_i \wedge dx_i = \sum_{i=1}^n dx_i \wedge dp_i,$$

it is, then, easy to check the non-degeneracy condition. In fact, using coordinates we can write a $W \in T_{(x,p)}(T^*M)$ as

$$W = \sum_{i=1}^{n} X_i \frac{\partial}{\partial x_i} + \sum_{i=1}^{n} P_i \frac{\partial}{\partial p_i},$$

Therefore $\Omega_{(x,p)}(W, \cdot) = \sum_{i=1}^{n} X_i dp_i - \sum_{i=1}^{n} P_i dx_i$, and $\Omega_{(x,p)}(W, \cdot) = 0$ implies $X_i = P_i = 0$, for $i = 1, \ldots, n$, by the independence of the family $(dx_1, \ldots, dx_n, dp_1, \ldots, dp_n)$.

In the following, we will suppose that V is a manifold provided with a symplectic structure Ω. If H is a C^r function defined on the open subset O of V, By the fact that $\Omega_x^\sharp$ is an isomorphism, we can associate to H a vector field X_H on O well defined by

$$\Omega_x(X_H(x), \cdot) = d_x H(\cdot).$$

Since Ω is non-degenerate, the vector field X_H is as smooth as the derivative of H, therefore it is C^{r-1}. In particular, if H is C^2, then the solutions of the vector field X_H define a partial flow $\phi_t^H : O \to O$.

Definition 2.5.4 (Hamiltonian Flow)**.** Suppose $H : O \to \mathbb{R}$ is a C^1 function, defined on the open subset O of the symplectic manifold V. The Hamiltonian vector of H is the vector field X_H on O, uniquely defined by

$$\Omega_x(X_H(x), \cdot) = d_x H(\cdot),$$

where Ω is the symplectic form on V. If, moreover, the function H is C^r, the vector X_H field is C^{r-1}. Therefore for $r \geq 2$, the partial flow ϕ_t^H generated by H exists, it is called the Hamitonian flow of H, and H is called the Hamiltonian of the flow ϕ_t^H.

Lemma 2.5.5. *$H : O \to \mathbb{R}$ is a C^2 function, defined on the open subset O of the symplectic manifold V, then H is constant on the orbits of its Hamiltonian flow ϕ_t^H.*

Proof. We must check that $d_x H(X_H(x))$ is 0, for all $x \in O$. But $d_x H(X_H(x))$ is $\Omega_x(X_H(x), X_H(x))$ vanishes because the bilinear form Ω_x is alternate. $\square$

Definition 2.5.6 (Lagrangian Subspace)**.** In a vector space E endowed with a symplectic bilinear form $a : E \times E \to \mathbb{R}$, a Lagrangian subspace is a vector subspace F of E with $\dim E = 2 \dim F$ and a is identically 0 on $F \times F$.

Lemma 2.5.7. *Let F be a subspace of the vector space E which is Lagrangian for the symplectic form a on E. If $x \in E$ is such that $a(x, y) = 0$, for all $y \in F$, then, the vector x is itself in F.*

Proof. Define $F^\perp = \{x \in E \mid \forall y \in F, a(x, y) = 0\}$. We have $F^\perp \supset F$, since a is 0 on $F \times F$. Since $a^\sharp : E \to E^*$ is an isomorphism, the dimension of $F^\perp$ is the same as that of its image $a^\sharp(F^\perp) = \{p \in E^* \mid p|F = 0\}$. This last subspace can be identified with the dual $(E/F)^*$ of the quotient of E by F. Therefore $\dim F^\perp = \dim E - \dim F = 2 \dim F - \dim F = \dim F$. Therefore $F^\perp = F$. □

Definition 2.5.8 (Lagrangian Submanifold)**.** If V is a symplectic manifold, a Lagrangian submanifold of V is a submanifold N of class at least C^1, and such that the subspace T_xN of T_xV is, for each $x \in N$, a Lagrangian subspace for the symplectic bilinear form Ω_x.

By the lemma 2.5.7 above, if $x \in N$, any vector $v \in T_xV$ such that $\Omega_x(v, v') = 0$, for all $v' \in T_xN$, is necessarily in T_xN.

Lemma 2.5.9. *If ω is a C^1 differential 1-form on the manifold M, then the graph* $\mathrm{Graph}(\omega)$ *of ω is a Lagrangian submanifold of T^*M if and only if ω is a closed form.*

Proof. Indeed,by lemma 2.5.1, we have $\omega = \omega^*\alpha$, and thus also $d\omega = \omega^* d\alpha = -\omega^*\Omega$. However, the form ω regarded as map of $M \to T^*M$ induces a diffeomorphism of C^1 class of M on $\mathrm{Graph}(\omega)$, consequently $d\omega = 0$ if and only if $\Omega|\,\mathrm{Graph}(\omega) = 0$. □

Theorem 2.5.10 (Hamilton-Jacobi)**.** *Let $H : O \to \mathbb{R}$ be a C^2 function defined on the open subset O of the symplectic manifold V. If $N \subset O$ is a connected C^1 Lagrangian submanifold of V, it is locally invariant by the partial flow ϕ_t^H if and only if H is constant on N.*

Proof. If H is constant on N, we have

$$\forall_x \in N,\ d_xH|T_xN = 0,$$

and thus $\Omega_x(X_H(x), v) = 0$, for all $v \in T_xN$, which implies that $X_H(x) \in T_xN$. By the theorem of Cauchy-Peano [Bou76], if N is of class C^1 (or Cauchy-Lipschitz, if N is of C^2 class), the restriction $X_H|N$ has solutions with values in N. By uniqueness of the solutions of X_H in O (which holds because X_H is C^1 on O), the solutions with values in N must be orbits of ϕ_t^H. We therefore conclude that N is invariant by ϕ_t^H as soon as H is constant on N. Conversely, if N is invariant by ϕ_t^H, the curves $t \mapsto \phi_t^H(x)$ have a speed $X_h(x)$ for $t = 0$ which must be in T_xN, therefore $X_h(x) \in T_xN$ and $d_xH|N = \Omega(X_H(x), \cdot)$ vanishes at every point of N, since N is a Lagrangian submanifold. By connectedness of N, the restriction $H|N$ is constant. □

2.6 Lagrangian and Hamiltonians

Definition 2.6.1 (Hamiltonian)**.** If L is a C^1 Lagrangian on the manifold M, its Hamiltonian $\hat{H} : TM \to \mathbb{R}$ is the function defined by

$$\hat{H}(x, v) = \frac{\partial L}{\partial v}(x, v)(v) - L(x, v).$$

Obviously, if L is C^k, with $k \geq 1$, then, its associated Hamiltonian $\hat{H}$ is of class C^{k-1}.

Proposition 2.6.2. *Suppose that L is a C^2 Lagrangian on the manifold M. If $\gamma : [a, b] \to M$ is a C^2 extremal curve then the Hamiltonian $\hat{H}$ is constant on its speed curve $t \mapsto (\gamma(t), \dot{\gamma}(t))$.*

In particular, the Hamiltonian $\hat{H}$ is invariant under the Euler-Lagrange flow ϕ_t^L when it exists.

Proof. We have

$$\hat{H}(\gamma(t), \dot{\gamma}(t)) = \frac{\partial L}{\partial v}(\gamma(t), \dot{\gamma}(t))(\dot{\gamma}(t)) - L(\gamma(t), \dot{\gamma}(t)),$$

We want to show that its derivative is zero. This is a local result, we can suppose that γ takes its values in a chart on M, and therefore use coordinates. Performing the differentiation with respect to t in coordinates, after simplifications, we get

$$\frac{d}{dt}\hat{H}(\gamma(t),\dot{\gamma}(t)) = \left[\frac{d}{dt}\Big(\frac{\partial L}{\partial v}(\gamma(t),\dot{\gamma}(t))\Big)\right](\dot{\gamma}(t)) - \frac{\partial L}{\partial x}(\gamma(t),\dot{\gamma}(t))(\dot{\gamma}(t)).$$

This last quantity is zero since γ satisfies the Euler-Lagrange equation, see 2.2.6. □

Suppose that the restriction of the global Legendre transform $\tilde{\mathcal{L}}$ to some open subset $O \subset M$ is a diffeomorphism onto its image $\tilde{O} = \tilde{\mathcal{L}}(O)$, we will define the function $H_{\tilde{O}} : \tilde{O} \to \mathbb{R}$ by $H_{\tilde{O}} = \hat{H} \circ (\tilde{\mathcal{L}}|O)^{-1}$, where $(\tilde{\mathcal{L}}|O)^{-1} : \tilde{O} \to O$ is the inverse of the restriction $\tilde{\mathcal{L}}|O$. This function $H_{\tilde{O}}$ is also called the (associated) Hamiltonian.

Proposition 2.6.3. *Let L be a C^k Lagrangian, with $k \geq 2$, on a manifold M. Suppose that the restriction $\tilde{\mathcal{L}}|O$ of the global Legendre transform to the open $O \subset M$ is a diffeomorphism onto its image $\tilde{O}$. Then the Hamiltonian $H_{\tilde{O}} = \hat{H} \circ (\tilde{\mathcal{L}}|O)^{-1}$ is also of class C^k on the open subset $\tilde{O} \subset T^*M$. If $M = U$ is an open subset of $\mathbb{R}^n$, then, in the natural coordinates on TU and T^*U, we have*

$$\begin{aligned}\frac{\partial H_{\tilde{O}}}{\partial p}(x,p) &= v\\ \frac{\partial H_{\tilde{O}}}{\partial x}(x,p) &= -\frac{\partial L}{\partial x}(x,v),\end{aligned}$$

with $\mathcal{L}(x,v) = p$.

Proof. To simplify notations we will set $H = H_{\tilde{O}}$. We know that $\tilde{\mathcal{L}}$ is C^{k-1}. It follows that the diffeomorphism $(\tilde{\mathcal{L}}|O)^{-1}$ is also C^{k-1}, and hence H is C^{k-1}, since obviously $\hat{H}$ is C^{k-1}. therefore H is at least C^1 We then take coordinates to reduce the proof to the case where $M = U$ is an open subset of $\mathbb{R}^n$. Using the canonical coordinates on $\mathbb{R}^n$, we write $(x,v) = (x_1,\cdots,x_n,v_1,\cdots,v_n)$. By definition of $\hat{H}$, and H, we have

$$H\big(x,\frac{\partial L}{\partial v}(x,v)\big) = -L(x,v) + \sum_{j=1}^{n}\frac{\partial L}{\partial v_j}(x,v)v_j. \tag{*}$$

If we differentiate both sides with respect to the variable v_i, we find

$$\sum_{j=1}^{n} \frac{\partial H}{\partial p_j}\big(x, \frac{\partial L}{\partial v}(x,v)\big)\frac{\partial^2 L}{\partial v_i \partial v_j}(x,v) = \sum_{j=1}^{n} v_j \frac{\partial^2 L}{\partial v_i \partial v_j}\big(x, \frac{\partial L}{\partial x}(x,v)\big),$$

for all $i = 1, \ldots, n$. Since we have

$$D\tilde{\mathcal{L}}(x,v) = \begin{bmatrix} \mathrm{Id}_{\mathbb{R}^n} & \frac{\partial^2 L}{\partial x \partial v}(x,v) \\ 0 & \frac{\partial^2 L}{\partial v^2}(x,v) \end{bmatrix}$$

the matrix $\left[\frac{\partial^2 L}{\partial v_i \partial v_j}\right]$ is invertible, for $(x,v) \in O$. Thereore we obtain

$$\frac{\partial H}{\partial p_j}\big(x, \frac{\partial L}{\partial v}(x,v)\big) = v_j.$$

If $p = \frac{\partial L}{\partial x}(x,v) = \mathcal{L}(x,v)$, we indeed found the first equation. If we differentiate both sides of the equality $(*)$ with respect to the variable x_i, using what we have just found, we obtain

$$\frac{\partial H}{\partial x_i}\big(x, \frac{\partial L}{\partial v}(x,v)\big) + \sum_{j=1}^{n} v_j \frac{\partial^2 L}{\partial x_i \partial v_j}(x,v) = \\ \sum_{j=1}^{n} \frac{\partial^2 L}{\partial x_i \partial v_j}(x,v) v_j - \frac{\partial L}{\partial x_i}(x,v),$$

hence

$$\frac{\partial H}{\partial x_i}[x, \frac{\partial L}{\partial v}(x,v)] = -\frac{\partial L}{\partial x_i}.$$

As $\tilde{\mathcal{L}}|O$ is a diffeomorphism of class C^{k-1} and L is of class C^k, writing the formulas just obtained as

$$\begin{aligned} \frac{\partial H}{\partial p}(x,p) =& p_2 \tilde{\mathcal{L}}^{-1}(x,p) \\ \frac{\partial H}{\partial x}(x,p) =& -\frac{\partial L}{\partial x}[\tilde{\mathcal{L}}^{-1}(x,p)] \end{aligned}$$

where p_2 is the projection of $TU = U \times \mathbb{R}^n$ on the second factor, we see that the derivative of H is C^{k-1}, and thus H is C^k. □

Let us recall that, for U an open subset of $\mathbb{R}^n$, the Euler-Lagrange flow is the flow of the vector field X_L on $TU = U \times \mathbb{R}^n$ defined by $X_L(x, v) = (x, v_1, v, \tilde{X}_L(x, v))$ with

$$\frac{\partial L}{\partial x}(x, v) = \frac{\partial^2 L}{\partial v^2}(x, v)(\tilde{X}_L(x, v), \cdot) + \frac{\partial^2 L}{\partial x \partial v}(x, v)(v, \cdot). \qquad (**)$$

Suppose now that $\tilde{\mathcal{L}}|O$ is a diiffeomorphism onto its image for the open subset $O \subset TM$. Since the diffeomorphism $\tilde{\mathcal{L}}$ is C^{k-1}, with $k \geq 2$, we can transport by this diffeomorphism the vector field $X_L|O$ to a vector field defined on $\tilde{O}$.

Theorem 2.6.4. *Let L be a C^k Lagrangian, with $k \geq 2$, on a manifold M. Suppose that the restriction $\tilde{\mathcal{L}}|O$ of the global Legendre transform to the open $O \subset M$ is a diffeomorphism onto its image $\tilde{O}$. If we transport on $\tilde{O} = \tilde{\mathcal{L}}(O)$ the Euler-Lagrange vector field X_L, using the diffeomorphism $\tilde{\mathcal{L}}|O$, we find on $\tilde{O}$ the Hamiltonian vector field X_{H_O} associated to $H_O = \hat{H} \circ (\tilde{\mathcal{L}}|O)^{-1}$. In particular, even if $k = 2$, the Euler-Lagrange vector field X_L is uniquely integrable on O, therefore the partial Euler-Lagrange flow ϕ_t^L is defined and C^1 on O. More generally, for every $r \geq 2$, the Euler-Lagrange flow ϕ_t^L on O is of class C^{r-1}.*

Proof. Let us fix $(x, v) \in O$. We set $p = \frac{\partial L}{\partial v}(x, v)$. As the Euler-Lagrange vector field is of the form $X_L(x, v) = (x, v, v, \tilde{X}_L(x, v))$, we have

$$T_{(x,v)}\tilde{\mathcal{L}}\big(X_L(x, v)\big) = \big(x, p, v, \frac{\partial \mathcal{L}}{\partial v}(x, v)(\tilde{X}_L(x, v)) + \frac{\partial \mathcal{L}}{\partial x}(x, v)(v)\big).$$

But $\frac{\partial \mathcal{L}}{\partial v}(x, v) = \frac{\partial^2 L}{\partial v^2}$ et $\frac{\partial \mathcal{L}}{\partial x} = \frac{\partial^2 L}{\partial x \partial v}$. Using equation $(**)$, we then find

$$T\tilde{\mathcal{L}}\big(\tilde{X}(x, v)\big) = \big(x, p, v, \frac{\partial L}{\partial x}(x, v)\big) = \big(x, p, \frac{\partial H}{\partial p}, -\frac{\partial H}{\partial x}\big),$$

this is precisely X_{H_O}, because

$$\begin{aligned}
&\Omega_{(x,p)}(\frac{\partial H_O}{\partial p}\frac{\partial}{\partial x} - \frac{\partial H_O}{\partial x}\frac{\partial}{\partial p}, \cdot)\\
&= (\sum_{i=1}^n dx_i \wedge dp_i)(\frac{\partial H_O}{\partial p}\frac{\partial}{\partial x} - \frac{\partial H_O}{\partial x}\frac{\partial}{\partial p}, \cdot)\\
&= \sum_{i=1}^n \frac{\partial H_O}{\partial p_i} dp_i + \frac{\partial H}{\partial x_i} dx_i\\
&= dH_O.
\end{aligned}$$

Since the Hamiltonian is C^k, the vector field X_{H_O} is of class C^{k-1}, its local flow $\phi_t^{H_O}$ is, by the theorem of Cauchy-Lipschitz, well defined and of class C^{k-1}. However, the diffeomorphism $\tilde{\mathcal{L}}|O$ is of class C^{k-1} and sends the vector field X_L on X_{H_O}, consequently the local flow of $X_L|O$ is well defined and equal to $(\tilde{\mathcal{L}}|O)^{-1}\phi_t^{H_O}\tilde{\mathcal{L}}$ which is C^{r-1}. □

The following theorem is clearly a consequence of 2.6.4, propositions 2.6.3 and 2.1.6.

Theorem 2.6.5. *Suppose that L is a non-degenerate* C^r *Lagrangian, with $r \geq 2$, on the manifold M. Then for $r = 2$, the the Euler-Lagrange vector field X_L is uniquely integrable and defines a local flow ϕ_t^L which is of class C^1. More generally, for every $r \geq 2$, the Euler-Lagrange flow ϕ_t^L is of class C^{r-1}.*

Remark 2.6.6. When L is non degenerate of class C^r, with $r \geq 2$, we can define $\Omega^L = \tilde{\mathcal{L}}*\Omega$, where Ω is the canonical symplectic form on T^*M. Since $\tilde{\mathcal{L}}$ is a local diffeomorphism, for each $(x,v) \in TM$, the bilinear form $\Omega^L_{(x,v)}$ is non-degenerate. Obviously, the 2-form is of class C^{r-2}. To be able to say that Ω^L is closed, we need to have $r \geq 3$. Under this condition Ω^L defines a symplectic structure of class C^{r-2} on TM, and we can interpret what we have done by saying that X_L is the Hamiltonian flow associated to $\hat{H}$ by Ω^L. Notice however that we cannot conclude from this that X_L is uniquely integrable because $\hat{H}$ is only C^{r-1}, and we do not gain one more degree of differentiabilty, like in proposition 2.6.3, for the Hamiltonian. Note that this also gives an explanation for the invariance of $\hat{H}$ by the Euler-Lagrange flow.

2.7 Existence of Local Extremal Curves

We will use the following form of the inverse function theorem.

Theorem 2.7.1 (Inverse Function)**.** *Let U be an open subset of $\mathbb{R}^m$ and K a compact space. We suppose that $\varphi : K \times U \to \mathbb{R}^m$ is a continuous map such that*

(1) *for each $k \in K$, the map $\varphi_k : U \to \mathbb{R}^m, x \mapsto \varphi(k,x)$ is C^1,*

(2) *The map $\frac{\partial\varphi}{\partial x} : K \times U \to \mathcal{L}(\mathbb{R}^m, \mathbb{R}^m), (k,x) \mapsto \frac{\partial\varphi}{\partial x}(k,x)$ is continuous.*

If $C \subset U$ is a compact subset such that

(i) *for each $k \in K$, and each $x \in C$, the derivative $\frac{\partial\varphi}{\partial x}(k,x)$ is an isomorphism,*

(ii) *for each $k \in K$, the map φ_k is injective on C,*

then there exists an open subset V such that

(a) *we have the inclusions $C \subset V \subset U$,*

(b) *for each $k \in K$, the map φ_k induces a C^1 diffeomorphism of V on an open subset of $\mathbb{R}^m$.*

Proof. Let $\|\cdot\|$ denote a norm on $\mathbb{R}^m$, and let d denote its associated distance. Let us show that there is an integer n such that, if we set

$$V_n = \{x \mid d(x,C) < \frac{1}{n}\},$$

then, the restriction $\varphi_k|V_n$ is injective, for each $k \in K$. We argue by contradiction. Suppose that for each positive integer n we can find $v_n, v_n' \in U$ and $k_n \in K$ with

$$v_n \neq v_n', d(v_n, C) < \frac{1}{n}, d(v_n', C) < \frac{1}{n} \text{ and } \varphi(k_n, v_n) = \varphi(k_n, v_n').$$

By the compactness of C and K, we can extract subsequences v_{n_i}, v_{n_i}' and k_{n_i} which converge respectively to $v_\infty, v_\infty' \in C$ and $k_\infty \in K$. By continuity of φ, we see that $\varphi(k_\infty, v_\infty) = \varphi(k_\infty, v_\infty')$. From (ii), it results that $v_\infty = v_\infty'$. Since $v_n \neq v_n'$, we can set $u_n =$

$\frac{v_n - v'_n}{\|v_n - v'_n\|}$. Extracting a subsequence if necessary, we can suppose that $u_{n_i} \to u_\infty$. This limit u_∞ is also of norm 1. As v_{n_i} and v'_{n_i} both converge to $v_\infty = v'_\infty$, for i large, the segment between v_{n_i} and v'_{n_i} is contained in the open set U. Hence for i big enough we can write

$$\begin{aligned} 0 &= \varphi(k_{n_i}, v_{n_i}) - \varphi(k_{n_i}, v'_{n_i}) \\ &= \int_0^1 \frac{\partial \varphi}{\partial v}(k_{n_i}, s v_{n_i} + (1-s) v'_{n_i})(v_{n_i} - v'_{n_i})\, ds, \end{aligned}$$

dividing by $\|v'_{n_i} - v_{n_i}\|$ and taking the limit as $n_i \to \infty$, we obtain

$$\begin{aligned} 0 &= \int_0^1 \frac{\partial \varphi}{\partial v}(k_\infty, v_\infty)(u_\infty)\, ds \\ &= \frac{\partial \varphi}{\partial v}(k_\infty, v_\infty)(u_\infty). \end{aligned}$$

However $\frac{\partial \varphi}{\partial v}(k_\infty, v_\infty)$ is an isomorphism, since $(k_\infty, v_\infty) \in K \times C$. But $\|u_\infty\| = 1$, this is a contradiction. We thus showed the existence of an integer n such that the restriction of φ_k on V_n is injective, for each $k \in K$. The continuity of $(k, v) \mapsto \frac{\partial \varphi}{\partial v}(k, v)$ and the fact that $\frac{\partial \varphi}{\partial v}(k, v)$ is an isomorphism for each (k, v) in the compact set $K \times C$, show that, taking n larger if necessary, we can suppose that $\frac{\partial \varphi}{\partial v}(k, v)$ is an isomorphism for each $(k, v) \in K \times V_n$. The usual inverse function theorem then shows that φ_k restricted to V_n is a local diffeomorphism for each $k \in K$. Since we have already shown that φ_k is injective on V_n, it is a diffeomorphism of V_n on an open subset of $\mathbb{R}^m$. □

The following lemma is a simple topological result that deserves to be better known because it simplifies many arguments.

Lemma 2.7.2. *Let X be a topological space, and let Y be a locally compact locally connected Hausdorff space. Suppose that $\varphi : X \times U \to Y$ is continuous, where U is an open subset of Y , and that, for each $x \in X$, the map $\varphi_x : U \to Y, y \mapsto \varphi(x, y)$ is a homeomorphism onto an open subset of Y. Then, the map $\Phi : X \times U \to X \times Y, (x, y) \mapsto (x, \varphi(x, y))$ is an open map, i.e. it maps open subsets of $X \times U$ to open subsets of $X \times Y$. It is thus a homeomorphism onto an open subset of $X \times Y$.*

Proof. It is enough to show that if V is open and relatively compact in Y, with $\bar{V} \subset U$, and $x_0 \in X, y_0 \in Y$ are such that $y_0 \in \varphi_{x_0}(V)$, then, there exists a neighborhood W of x_0 in X and a neighborhood N of y_0 in Y, such that $\varphi_x(V) \supset N$, for each $x \in W$. In fact, this will show the inclusion $W \times N \subset \Phi(W \times V)$. As $\varphi_{x_0}(V)$ is an open set containing y_0, there exists N, a compact and connected neighborhood of y_0 in Y, such that $N \subset \varphi_{x_0}(V)$. Since $\partial V = \bar{V} \setminus V$ is compact and $N \cap \varphi_{x_0}(\partial V) = \emptyset$, by continuity of φ, we can find a neighborhood W of x_0 such that

$$\forall x \in W, \varphi_x(\partial V) \cap N = \emptyset. \tag{*}$$

We now choose $\tilde{y}_0 \in V$, such that $\varphi_{x_0}(\tilde{y}_0) = y_0$. Since N is a neighborhood of y_0 and φ is continuous, taking W smaller if necessary we can assume that

$$\forall x \in W, \varphi_x(\tilde{y}_0) \in N. \tag{**}$$

By condition $(*)$, for $x \in W$, we have $\varphi_x(V) \cap N = \varphi_x(\bar{V}) \cap N$, therefore the intersection $\varphi_x(V) \cap N$ is both open and closed as a subset of the connected space N. This intersection is not empty because it contains $\varphi_x(\tilde{y}_0)$ by condition $(**)$. By the connectedness of N, this of course implies that $\varphi_x(V) \cap N = N$. □

Lemma 2.7.3 (Tilting)**.** *Let $\|\cdot\|$ be a norm on $\mathbb{R}^n$. We denote by $\mathring{B}_{\|\cdot\|}(0, R)$ (resp. $\bar{B}_{\|\cdot\|}(0, R)$) the open (resp. closed) ball of $\mathbb{R}^n$ of center 0 and radius R for this norm. We suppose that K is a compact space and that ϵ, η, C_1 and C_2 are fixed > 0 numbers, with $C_1 > C_2$.*

Let $\theta : K \times]-\epsilon, \epsilon[\times \mathring{B}_{\|\cdot\|}(0, C_1 + \eta) \to \mathbb{R}^n$ be continuous map such that

(1) *for each fixed $k \in K$, the map $(t, v) \mapsto \theta(k, t, v)$ has everywhere a partial derivative $\frac{\partial \theta}{\partial t}$, and this partial derivative is itself C^1;*

(2) *the map θ and its partial derivatives $\frac{\partial \theta}{\partial t}, \frac{\partial^2 \theta}{\partial t^2}, \frac{\partial^2 \theta}{\partial v \partial t} = \frac{\partial}{\partial v}\left[\frac{\partial \theta}{\partial t}\right]$ are continuous on the product space $K \times]-\epsilon, \epsilon[\times \mathring{B}_{\|\cdot\|}(0, C_1 + \eta)$;*

(3) *for each* $k \in K$ *and each* $v \in \bar{B}_{\|\cdot\|}(0, C_1 + \eta)$, *we have*

$$\frac{\partial \theta}{\partial t}(k, 0, v) = v;$$

(4) *for each* $(k, v) \in K \times B(0, C_1 + \eta)$, *we have*

$$\theta(k, 0, v) = \theta(k, 0, 0).$$

Then, there exists $\delta > 0$ *such that, for each* $t \in [-\delta, 0[\cup]0, \delta]$ *and each* $k \in K$, *the map* $v \mapsto \theta(k, t, v)$ *is a diffeomorphism of* $\mathring{B}_{\|\cdot\|}(0, C_1 + \eta/2)$ *onto an open subset of* $\mathbb{R}^n$, *and moreover*

$$\{\theta(k, t, v) \mid v \in \bar{B}_{\|\cdot\|}(0, C_1)\} \supset \{x \in \mathbb{R}^n \mid \|x - \theta(k, 0, 0)\| \leq C_2|t|\}.$$

Proof. Let us consider the map

$$\Theta(k, t, v) = \frac{\theta(k, t, v) - \theta(k, 0, v)}{t},$$

defined on $K \times ([-\epsilon, 0[\cup]0, \epsilon]) \times \mathring{B}_{\|\cdot\|}(0, C_1 + \eta)$. We can extend it by continuity at $t = 0$ because

$$\Theta(k, t, v) = \int_0^1 \frac{\partial \theta}{\partial t}(k, st, v)\, ds \tag{$*$}$$

The right-hand side is obviously well-defined for $t = 0$, and equal to $\partial\theta/\partial t(k, 0, v) = v$. Moreover, upon inspection of the right-hand side of $(*)$, the extension $\Theta : K \times]-\epsilon, \epsilon[\times \mathring{B}_{\|\cdot\|}(0, C_1 + \eta) \to \mathbb{R}^n$ is such that for each fixed $k \in K$, the map $(t, v) \to \Theta(k, t, v)$ is C^1, with

$$\frac{\partial \Theta}{\partial t}(k, t, v) = \int_0^1 \frac{\partial^2 \theta}{\partial t^2}(k, st, v) s\, ds,$$
$$\frac{\partial \Theta}{\partial v}(k, t, v) = \int_0^1 \frac{\partial^2 \theta}{\partial v \partial t}(k, st, v)\, ds.$$

Therefore both the partial derivatives $\partial\Theta/\partial t, \partial\Theta/\partial v$ are continuous on the product space $K \times]-\epsilon, \epsilon[\times \mathring{B}_{\|\cdot\|}(0, C_1 + \eta)$. Let us then define the map $\tilde{\Theta} : K \times]-\epsilon, \epsilon[\times \mathring{B}_{\|\cdot\|}(0, C_1 + \eta) \to \mathbb{R} \times \mathbb{R}^n$ by

$$\tilde{\Theta}(k, t, v) = (t, \Theta(k, t, v)).$$

To simplify, we will use the notation $x = (t, v)$ to indicate the point $(t, v) \in \mathbb{R} \times \mathbb{R}^n = \mathbb{R}^{n+1}$. The map $\tilde{\Theta}$ is obviously continuous, and the derivative $\partial\tilde{\Theta}/\partial x$ is also continuous on the product space $K \times]-\epsilon, \epsilon[\times \mathring{B}_{\|\cdot\|}(0, C_1 + \eta)$. Since $\Theta(k, 0, v) = v$, for each $(k, v) \in K \times \mathring{B}_{\|\cdot\|}(0, C_1 + \eta)$, we find that

$$\frac{\partial\tilde{\Theta}}{\partial x}(k, 0, v) = \begin{bmatrix} 1 & 0 \\ \frac{\partial\Theta}{\partial t}(k, 0, v) & \mathrm{Id}_{\mathbb{R}^n} \end{bmatrix},$$

where we used a block matrix to describe a linear map from the product $\mathbb{R} \times \mathbb{R}^n$ into itself. It follows that $\frac{\partial\tilde{\Theta}}{\partial x}(k, 0, v)$ is an isomorphism for each $(k, v) \in K \times \mathring{B}_{\|\cdot\|}(0, C_1 + \eta)$.

Since K and $\bar{B}_{\|\cdot\|}(0, C_1 + \eta/2)$ are compact, using the inverse function theorem 2.7.1, we can find $\delta_1 > 0$ and $\eta' \in]\eta/2, \eta[$ such that, for each $k \in K$, the map $(t, v) \mapsto \Theta(k, t, v)$ is a C^1 diffeomorphism from the open set $]-\delta_1, \delta_1[\times \mathring{B}_{\|\cdot\|}(0, C_1 + \eta')$ onto an open set $\mathbb{R} \times \mathbb{R}^n$. It follows that, for each $(k, t) \in K \times]-\delta_1, \delta_1[$, the map $v \mapsto \tilde{\Theta}(k, t, v)$ is a C^1 diffeomorphism $\mathring{B}_{\|\cdot\|}(0, C_1 + \eta/2)$ onto some open subset of $\mathbb{R}^n$. By lemma 2.7.2, we obtain that the image of $K \times]-\delta_1, \delta_1[\times \mathring{B}_{\|\cdot\|}(0, C_1)$ by the map $(k, s, v) \mapsto (k, \tilde{\Theta}(k, s, v))$ is an open subset of $K \times \mathbb{R} \times \mathbb{R}^n$. This open subset contains the compact subset $K \times \{0\} \times \bar{B}_{\|\cdot\|}(0, C_2)$, since $\tilde{\Theta}(k, 0, v) = (0, v)$. We conclude that there exists $\delta > 0$ such that the image of $K \times]-\delta_1, \delta_1[\times \mathring{B}_{\|\cdot\|}(0, C_1)$ by map the $(k, s, v) \mapsto (k, \tilde{\Theta}(k, s, v))$ contains $K \times [-\delta, \delta] \times \bar{B}_{\|\cdot\|}(0, C_2)$. Hence, for $(k, t) \in K \times [-\delta, \delta]$, the image of $\mathring{B}_{\|\cdot\|}(0, C_1)$ by the map $v \mapsto \Theta(k, t, v)$ contains $\bar{B}_{\|\cdot\|}(0, C_2)$.

Since we have

$$\theta(k, s, v) = s\Theta(k, s, v) + \theta(k, 0, v)$$
$$\theta(k, 0, v) = \theta(k, 0, 0),$$

we can translate the results obtained for Θ in terms of θ. This gives that, for $s \neq 0$, and $|s| \leq \delta$, the map $v \mapsto \theta(k, s, v)$ is also a diffeomorphism of $\mathring{B}_{\|\cdot\|}(0, C_1 + \eta/2)$ on an open subset of $\mathbb{R}^n$ and that the image of $\bar{B}_{\|\cdot\|}(0, C_1)$ by this map contains the ball $\bar{B}(\theta(k, 0, 0), C_2 s)$. □

Theorem 2.7.4 (Existence of local extremal curves)**.** *Let $L : TM \to M$ be a non-degenerate C^r Lagrangian, with $r \geq 2$. We* *fix a Riemannian metric g on M. For $x \in M$, we denote by $\|\cdot\|_x$ the norm induced on T_xM by g. We call d the distance on M associated with g.*

If $K \subset M$ is compact and $C \in [0, +\infty[$, then, there exists $\epsilon > 0$ such that for $x \in K$, and $t \in [-\epsilon, 0[\cup]0, \epsilon]$, the map $\pi \circ \phi_t^L$ is defined, and induces a diffeomorphism from an open neighborhood of $\{v \in T_xM \mid \|v\|_x \leq C\}$ onto an open subset of M. Moreover, we have

$$\pi \circ \phi_t^L(\{v \in T_xM \mid \|v\|_x \leq C\}) \supset \{y \in M \mid d(y, x) \leq C|t|/2\}.$$

To prove the theorem, it is enough to show that for each $x_0 \in M$, there exists a compact neighborhood K, such that the conclusion of the theorem is true for this compact neighborhood K. For such a local result we can assume that $M = U$ is an open subset of $\mathbb{R}^n$, with $x_0 \in U$. In the sequel, we identify the tangent space TU with $U \times \mathbb{R}^n$ and for $x \in U$, we identify $T_xU = \{x\} \times \mathbb{R}^n$ with $\mathbb{R}^n$. We provide $U \times \mathbb{R}^n$ with the natural coordinates $(x, v) = (x_1, \cdots, x_n, v_1, \cdots, v_n)$. We start with a lemma which makes it possible to replace the norm obtained from the Riemannian metric by a constant norm on $\mathbb{R}^n$.

Lemma 2.7.5 (Distance Estimates). *For each $\alpha > 0$, there exists an open neighborhood V of x_0 with $\bar{V}$ compact $\subset U$ and such that*

(1) *for each $v \in T_xU \cong \mathbb{R}^n$ and each $x \in \bar{V}$ we have*

$$(1 - \alpha)\|v\|_{x_0} \leq \|v\|_x \leq (1 + \alpha)\|v\|_{x_0};$$

(2) *for each $(x, x') \in \bar{V}$ we have*

$$(1 - \alpha)\|x - x'\|_{x_0} \leq d(x, x') \leq (1 + \alpha)\|x - x'\|_{x_0}.$$

Proof. For (1), we observe that, for $x \to x_0$, the norm $\|v\|_x$ tends uniformly to 1 on the compact set $\{v \mid \|v\|_{x_0} = 1\}$, by continuity of the Riemannian metric. Therefore for x near to x_0, we have

$$\forall v \in \mathbb{R}^n,\ (1 - \alpha) < \left\| \frac{v}{\|v\|_{x_0}} \right\|_x < (1 + \alpha).$$

For (2), we use the exponential map $\exp_x : T_xU \to U$, induced by the Riemannian metric. It is known that the map $\exp : TU =$

$U \times \mathbb{R}^n \to U \times U, (x,v) \mapsto (x, \exp_x v)$ is a local diffeomorphism on a neighborhood of $(x_0, 0)$, that $\exp_x(0) = x$, and $D[\exp_x](0) = \mathrm{Id}_{\mathbb{R}^n}$. Thus, there is a compact neighborhood $\bar{W}$ of x_0, such that any pair $(x, x') \in \bar{W} \times \bar{W}$ is of the form $(x, \exp_x[v(x,x')])$ with $v(x,x') \to 0$ if x and x' both tend to x_0. The map $(x,v) \mapsto \exp_x(v)$ is C^1, therefore, using again $\exp_x(0) = x, D[\exp_x](0) = \mathrm{Id}_{\mathbb{R}^n}$, we must have

$$\exp_x v = x + v + \|v\|_{x_0} k(x,v),$$

with $\lim_{v \to 0} k(x,v) = 0$, uniformly in $x \in \bar{W}$. Since $d(x, \exp_x v) = \|v\|_x$, for v small, it follows that for x, x' close to x_0

$$\frac{\|x - x'\|_x}{d(x,x')} = \frac{\|v(x,x') + \|v(x,x')\|_{x_0} k(x, v(x,x'))\|_x}{\|v(x,x')\|_x}.$$

We can therefore conclude that $\frac{\|x-x'\|_x}{d(x,x')} \to 1$, when $x, x' \to x_0$. But we also have $\frac{\|x-x'\|_x}{\|x-x'\|_{x_0}} \to 1$ when $x \to x_0$, we conclude that $\frac{d(x,x')}{\|x-x'\|_{x_0}}$ is close to 1, if x and x' are both in a small compact neighborhood of x_0. $\square$

Proof of the theorem 2.7.4. Let us give α and η two > 0 numbers, with α enough small to have

$$\frac{C}{1-\alpha} < \frac{C}{1+\alpha} + \frac{\eta}{2}$$
$$\frac{1}{2(1-\alpha)} < \frac{1}{1+\alpha}.$$

We set $C_1 = C/(1+\alpha)$. Let $\bar{W} \subset U$ be a compact neighborhood of x_0. Since $\bar{W} \times \bar{B}_{\|\cdot\|_{x_0}}(0, C_1 + \eta)$ is compact, there exists $\epsilon > 0$ such that ϕ_t^L is defined on $\bar{W} \times \bar{B}_{\|\cdot\|_{x_0}}(0, C_1 + \eta)$ for $t \in]-\epsilon, \epsilon[$. We then set $\theta(x,t,v) = \pi \circ \phi_t^L(x,v)$. The map θ is well defined on $\bar{W} \times [-\epsilon, \epsilon] \times \bar{B}_{\|\cdot\|_{x_0}}(0, C_1 + \eta)$. Moreover, since $t \mapsto \phi_t^L(x,v)$ is the speed curve of its projection $t \mapsto \theta(t,x,v)$, we have

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

and

$$\theta(x,0,v) = x, \text{ and } \frac{\partial \theta}{\partial t}(x,0,v) = v.$$

Since the flow ϕ_t^L is of class C^{r-1}, see theorem 2.6.5, both maps θ and $\frac{\partial\theta}{\partial t}$ are of class C^{r-1}, with respect to all variables. Since $r \geq 2$, we can then apply the tilting lemma 2.7.3, with $C_1 = C/(1+\alpha)$ and $C_2 = C/2(1-\alpha)$, to find $\delta > 0$ such that, for each $x \in \bar{W}$, and each $t \in [-\delta, 0[\cup]0, \delta]$, the map $(x, v) \mapsto \pi \circ \phi_t^L(x, v)$ induces a C^1 diffeomorphism from $\{v \mid \|v\|_{x_0} < \eta/2 + C/(1+\alpha)\}$ onto an open subset of $\mathbb{R}^n$ with

$$\{\pi \circ \phi_t^L(x, v) \mid \|v\|_{x_0} \leq \frac{C}{1+\alpha}\} \supset \{y \in \mathbb{R}^n \mid \|y - x\|_{x_0} \leq \frac{Ct}{2(1-\alpha)}\}.$$

Since $\pi \circ \phi_0^L(x, v) = x$, taking $\bar{W}$ and $\delta > 0$ smaller if necessary, we can assume that $\bar{W} \subset V$ and

$$\{\pi\phi_t^L(x, v) \mid t \in [-\delta, \delta], x \in \bar{W}, v \in \bar{B}(0, C/(1+\alpha) + \eta)\} \subset V,$$

where V is given by lemma 2.7.5. Since $\bar{W} \subset V$, by what we obtained in lemma 2.7.5, for $x \in \bar{W}$, for every $R \geq 0$, we have

$$\begin{aligned}\{v \in T_xU \mid \|v\|_{x_0} \leq \frac{R}{1+\alpha}\} &\subset \{v \in T_xU \mid \|v\|_x \leq R\}\\ &\subset \{v \in T_xU \mid \|v\|_{x_0} \leq \frac{R}{1-\alpha}\}.\end{aligned}$$

As $\bar{W}$ is compact and contained in the open set V, for $t > 0$ small and $x \in \bar{W}$, we have

$$\{y \in M \mid d(y, x) \leq \frac{Ct}{2}\} \subset V,$$

hence again by lemma 2.7.5

$$\{y \in M \mid d(y, x) \leq \frac{Ct}{2}\} \subset \{y \in V \mid \|y - x\|_{x_0} \leq \frac{Ct}{2(1-\alpha)}\}.$$

Therefore by the choices made, taking $\delta > 0$ smaller if necessary, for $t \in [-\delta, \delta]$, and $x \in \bar{W}$, the map $\pi \circ \phi_t^L$ is a diffeomorphism from a neighborhood of $\{v \in T_xU \mid \|v\|_x \leq C\}$ onto an open subset of U such that

$$\pi \circ \phi_t^L(\{v \in T_xU \mid \|v\|_x \leq C\}) \supset \{y \in V \mid d(y, x) \leq \frac{C|t|}{2}\}. \square$$

2.8 The Hamilton-Jacobi method

We already met an aspect of the theory of Hamilton and Jacobi, since we saw that a connected Lagrangian submanifold is invariant by a Hamiltonian flow if and only if the Hamiltonian is constant on this submanifold. We will need a little more general version for the case when functions depend on time.

We start with some algebraic preliminaries.

Let a be an alternate 2-form on a vector space E. By definition, the characteristic subspace of a is

$$\ker a = \{\xi \in E \mid a(\xi, \cdot) = 0\},$$

i.e. $\ker a$ is the kernel of the linear map $a^{\#} : E \to E^*, \xi \mapsto a(\xi, \cdot)$. If E is of finite dimension, then $\mathrm{Ker}(a) = 0$ if and only if a is a symplectic form. Since a space carrying a symplectic form is of even dimension, we obtain the following lemma.

Lemma 2.8.1. *Let a be an alternate 2-form on the vector space E is provided with the alternate bilinear 2-form. If the dimension of E is finite and odd, then $\mathrm{Ker}(a)$ is not reduced to $\{0\}$.*

We will need the following complement.

Lemma 2.8.2. *Let a be an alternate bilinear 2-form on the vector space E of finite odd dimension. We suppose that there is a codimension one subspace $E_0 \subset E$ such that the restriction $a|E_0$ is a symplectic form, then, the dimension of the characteristic subspace $\mathrm{Ker}(a)$ is 1.*

Proof. Indeed, we have $E_0 \cap \mathrm{Ker}(a) = 0$, since E_0 is symplectic. As E_0 is a hyperplane, the dimension of $\mathrm{Ker}(a)$ is ≤ 1. But we know by the previous lemma that $\mathrm{Ker}(a) \neq 0$. □

Definition 2.8.3 (Odd Lagrangian Subspace)**.** Let E be a vector space of dimension $2n+1$ provided with an alternate 2-form a, such that the dimension of $\mathrm{Ker}(a)$ is 1. A vector subspace F of E is said to be odd Lagrangian if $\dim F = n+1$ and the restriction $a|F$ is identically 0.

Lemma 2.8.4. *Let E be a vector space of dimension $2n+1$ provided with an alternate 2-form a, such that the dimension of $\mathrm{Ker}(a)$ is 1. If F is an odd Lagrangian subspace, then $\mathrm{Ker}(a) \subset F$.*

Proof. We set $F^\perp = \{\xi \in E \mid \forall f \in F, a(\xi, f) = 0\}$. We have $F^\perp \supset F + \mathrm{Ker}(a)$, therefore $\dim a^\#(F^\perp) = \dim F^\perp - 1$, since $\mathrm{Ker}(a)^\# = \mathrm{Ker}(a)$, where $a^\# : E \to E^*, \xi \mapsto a(\xi, \cdot)$. But $a^\#(F^\perp) \subset \{\varphi \in E^* \mid \varphi(F) = \{0\}\}$ which is of dimension n. It follows that $\dim F^\perp - 1 \leq n$ and thus $\dim F^\perp \leq n+1$. Since $F^\perp \supset F + \mathrm{Ker}(a)$ and $\dim F = n+1$, we must have $F^\perp = F$ and $\mathrm{Ker}(a) \subset F$. □

In the sequel of this section, we fix a manifold M and O an open subset of its cotangent space T^*M. We denote by $\pi^* : T^*M \to M$ the canonical projection.

We suppose that a C^2 Hamiltonian $H : O \to \mathbb{R}$ is given. We denote by X_H the Hamiltonian vector field associated to H, and by ϕ_t^H the local flow of X_H. We define the differential 1-form α_H on $O \times \mathbb{R}$ by

$$\alpha_H = \alpha - H dt,$$

where α is the Liouville form on T^*M. More precisely, we should write

$$\alpha_H = p_1^* \alpha - (H \circ p_1)\, dt,$$

where $p_1 : T^*M \times \mathbb{R} \to T^*M$ is the projection on the first factor and dt is the differential on $T^*M \times \mathbb{R}$ of the projection $T^*M \times \mathbb{R} \to \mathbb{R}$ on the second factor. The exterior derivative $\Omega_H = -d\alpha_H$ defines a differential 2-form which is closed on $O \times \mathbb{R}$. We have

$$\Omega_H = p_1^* \Omega + (p_1^* dH) \wedge dt,$$

where $\Omega = -d\alpha$ is the canonical symplectic form on T^*M. If $(x, p, t) \in (O \times \mathbb{R})$, then, the tangent space $T_{(x,p,t)}(O \times \mathbb{R}) = T_{(x,p)}(T^*M) \times \mathbb{R}$ is of odd dimension. Since, moreover, the restriction $(\Omega_H)_{(x,p,t)} | T_{(x,p)} T^*M$ is nothing but the symplectic form $\Omega = -d\alpha$, the lemmas above show that the characteristic space of $(\Omega_H)_{(x,p,t)}$ is of dimension 1, at each point $(x, p, t) \in O \times \mathbb{R}$.

Lemma 2.8.5. *At a point (x, p, t) in $O \times \mathbb{R}$, the characteristic subspace of Ω_H is generated by the vector $X_H + \frac{\partial}{\partial t}$, where X_H is the Hamiltonian vector field on O associated with H.*

Proof. The vector field $X_H + \frac{\partial}{\partial t}$ is never 0, because of the part $\frac{\partial}{\partial t}$, it is then enough to see that

$$\Omega_H(X_H + \frac{\partial}{\partial t}, \cdot) = 0.$$

But, we have

$$\begin{aligned}\Omega_H\big(X_H + \frac{\partial}{\partial t}, \cdot\big) &= \Omega(X_H, \cdot) + (dH \wedge dt)\big(X_H + \frac{\partial}{\partial t}, \cdot\big)\\ &= dH + dH(X_H)dt - dH\\ &= 0,\end{aligned}$$

since $dH(X_H) = \Omega(X_H, X_H) = 0$. □

Definition 2.8.6 (Odd Lagrangian Submanifold)**.** We say that a C^1 submanifold V of $O \times \mathbb{R}$ is odd Lagrangian for Ω_H, if $\dim V = \dim M + 1$ and the restriction $\Omega_H|V$ is identically 0. This last condition is equivalent to the fact that the restriction $\alpha_H|V$ is closed as differential 1-form.

Lemma 2.8.7. *If the* C^1 *submanifold* V *of* $O \times \mathbb{R}$ *is odd Lagrangian for* Ω_H*, then, the vector field* $X_H + \frac{\partial}{\partial t}$ *is tangent everywhere to* V*.*

It is not difficult to see that the local flow of $X_H + \frac{\partial}{\partial t}$ on $O \times \mathbb{R}$ is

$$\Phi_s^H(x, p, t) = (\phi_s^H(x, p), t + s).$$

Corollary 2.8.8. *If the* C^1 *submanifold* V *of* $O \times \mathbb{R}$ *is odd Lagrangian for* Ω_H*, then, it is invariant by the local flow* Φ_s^H*.*

Proof. Again since we are assuming that V is only C^1, the restriction of $X_H + \frac{\partial}{\partial t}$ to V which is tangent to V is only C^0, as a section $V \to TV$. We cannot apply the Cauchy-Lipschitz theorem. Instead as in the proof of 2.5.10, if $(x, p, t) \in V$, we apply the Cauchy-Peano [Bou76] to find a a curve in V which is a solution of the vector field. Then we apply the uniqueness in O where the vector field to conclude that this curve in V is a part of an orbit of the flow. □

Lemma 2.8.9. *the form* Ω_H *is preserved by the flow* Φ_s^H*.*

Proof. By the Cartan formula, a closed differential form β is preserved by the flow of the vector field X, if and only if the exterior derivative $d[\beta(X, \cdot)]$ of the form $\beta(X, \cdot)$ is identically 0. However in our case

$$\Omega_H(X_H + \frac{\partial}{\partial t}, \cdot) = 0.\square$$

Let us consider U, an open subset of M, and a local C^1 section $s : U\times]a, b[\to T^*M\times\mathbb{R}$ of projection $\pi^*\times Id_{\mathbb{R}} : T^*M\times\mathbb{R}\to M\times\mathbb{R}$ such that $s(x,t)\in O$, for each $(x,t)\in U\times]a,b[$. If we set $s(x,t) = (x, p(x,t), t)$, then the image of the section s is an odd Lagrangian submanifold for Ω_H if and only if the form: $s^*[\alpha - H(x, p(x,t))\, dt]$ is closed. If we choose coordinates $x_1,\ldots\ldots, x_n$ in a neighborhood of a point in U, we get

$$s^*[\alpha - H(x,p(x,t))\,dt] = -H(x,p(x,t))\,dt + \sum_{i=1}^{n} p_i(x,t)\,dx_i$$

and thus the image of the section s is an odd Lagrangian submanifold if and only if the differential 1-form $-H(x,p(x,t))\,dt + \sum_{i=1}^{n} p_i(x,t)\,dx_i$ is closed. If this is the case and U is simply connected, this form is then exact and there is a C^2 function $S : U\times]a,b[\to\mathbb{R}$ such that

$$dS = -H(x,p(x,t))\,dt + \sum_{i=1}^{n} p_i(x,t)\,dx_i,$$

which means that we have

$$\frac{\partial S}{\partial x} = p(x,t) \text{ and } \frac{\partial S}{\partial t} = -H(x,p(x,t)).$$

This brings us to the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} + H(x, \frac{\partial S}{\partial x}) = 0. \tag{H-J}$$

Conversely, any C^2 solution $S : U\times]a,b[\to\mathbb{R}$ of this equation gives us an invariant odd Lagrangian submanifold, namely the image of the section $s : U\times]a,b[\to T^*M\times\mathbb{R}$ defined by $s(x,t) = \big(x, \frac{\partial S}{\partial x}(x,t), t\big)$. Indeed, with this choice and using coordinates, we find that

$$s^*(\alpha_H) = \sum_{i=1}^{n} \frac{\partial S}{\partial x_i}(x,t)\,dx_i - H(x,p(x,t))\,dt.$$

Since S satisfies the Hamilton-Jacobi equation, we have

$$s^*(\alpha_H) = \frac{\partial S}{\partial t}(x,t)\,dt + \sum_{i=1}^{n} \frac{\partial S}{\partial x_i}(x,t)\,dx_i,$$

which implies that $s^*(\alpha_H) = dS$ is closed.

The following theorem partly summarizes what we obtained:

Theorem 2.8.10. *Let $H : O \to \mathbb{R}$ be a C^2 Hamiltonian, defined on the open subset O of the cotangent space T^*M of the manifold M. We denote by ϕ_t^H the local flow of the Hamiltonian vector field X_H associated with H. Let U be an open subset of M and $a, b \in \mathbb{R}$, with $a < b$. Suppose that the C^2 function $S : U \times]a, b[\to \mathbb{R}$ is such that*

(1) *for each $(x, t) \in U \times]a, b[$, we have $\big(x, \frac{\partial S}{\partial x}(x, t)\big) \in O$;*

(2) *the function S satisfies the Hamilton-Jacobi equation*

$$\frac{\partial S}{\partial t} + H(x, \frac{\partial S}{\partial x}) = 0.$$

We fix (x_0, t_0) a point in $U \times]a, b[$, and to simplify notations we denote the point $\phi_t^H\big(x_0, \frac{\partial S}{\partial x}(x_0, t_0)\big)$ by $(x(t), p(t))$. If $]\alpha, \beta[\subset]a, b[$ is the maximum open interval such that $\phi_t^H\big(x_0, \frac{\partial S}{\partial x}(x_0, t_0)\big)$ is defined, and its projection $x(t)$ is in U, for each $t \in]\alpha, \beta[$, then, we have

$$\forall t \in]\alpha, \beta[,\ p(t) = \frac{\partial S}{\partial x}(x(t), t + t_0).$$

Moreover, for t tending to α or β, the projection $x(t)$ leaves every compact subset of U.

Proof. By what we have already seen, the image of the section $s : U \times]a, b[\to T^*M \times \mathbb{R}, (x, t)$ defined by $s(x, t) = \big(x, \frac{\partial S}{\partial x}(x, t), t\big)$ is odd Lagrangian for Ω_H, it is thus invariant by the local flow Φ_t^H of $X_H + \frac{\partial}{\partial t}$. If we denote by $]\alpha_0, \beta_0[$ the maximum open interval such that $\Phi_t^H\big(x_0, \frac{\partial S}{\partial x}(x_0, t_0), t_0\big)$ is defined and in the image of the section s, it results from the invariance that

$$\forall t \in]\alpha_0, \beta_0[,\ p(t) = \frac{\partial S}{\partial x}(x(t), t + t_0).$$

We of course have $]\alpha_0, \beta_0[\subset]\alpha, \beta[$. If we suppose that $\beta_0 < \beta$, then $x(\beta_0) \in U$ and, by continuity of the section s, we also have $\Phi_{\beta_0}^H\big(x_0, \frac{\partial S}{\partial x}(x_0, t_0), t_0\big) = s(x(\beta_0))$. By the invariance of the image of s by Φ_t^H, we then find $\epsilon > 0$ such that $\Phi_t^H\big(x_0, \frac{\partial S}{\partial x}(x_0, t_0), t_0\big) =$

$\Phi^H_{t-\beta_0}\big(s(x(\beta_0))\big)$ is defined and in the image of the section s, for each $t \in [\beta_0, \beta_0 + \epsilon]$. This contradicts the definition of β_0.

It remains to see that $x(t)$ comes out of every compact subset of U, for example, when $t \to \beta$. Indeed if this would not be the case, we could find a sequence $t_i \to \beta$ such that $x(t_i)$ would converge to a point $x_\infty \in U$, but, by continuity of s, the sequence $\phi^H_{t_i}\big(x_0, \frac{\partial S}{\partial x}(x_0, t_0)\big) = s[x(t_i)]$ would converge to a point in T^*U, which would make it possible to show that $\phi^H_t\big(x_0, \frac{\partial S}{\partial x}(x_0, t_0)\big)$ would be defined and in T^*U for t near to β and $t > \beta$. This contradicts the definition of β. □

We know consider the problem of constructing (local) solutions of the Hamilton-Jacobi equation. We will include in this construction a parameter that will be useful in the sequel.

Theorem 2.8.11 (Method of Characteristics)**.** *Let $H : O \to \mathbb{R}$ be a C^2 Hamiltonian, defined on the open subset O of cotangent space T^*M of the manifold M. We denote by ϕ^H_t the local flow of the Hamiltonian vector field X_H associated with H. Let U be open in M and let K be a compact space. We suppose that $S_0 : K \times U \to \mathbb{R}$ is a function such that*

(1) *for each $k \in K$, the map $S_{0,k} : U \to \mathbb{R}$ is C^2;*

(2) *the map $(k, x) \mapsto (x, \frac{\partial S_0}{\partial x}(k, x))$ is continuous on the product $K \times U$ with values in T^*M, and its image is contained in the open subset $O \subset T^*M$;*

(3) *the derivative with respect to x of $(k, x) \mapsto (x, \frac{\partial S_0}{\partial x}(k, x))$ is also continuous on the product $K \times U$ (this is equivalent to the continuity of the $(k, x) \mapsto \frac{\partial^2 S_0}{\partial x^2}(k, x)$ in charts contained in U).*

Then, for each open simply connected subset W, with $\bar{W}$ compact and included in U, there exists $\delta > 0$ and a continuous map $S : K \times W \times]-\delta, \delta[$ satisfying

(i) *for each $(k, x) \in K \times W$, we have $S(k, x, 0) = S_0(k, x)$;*

(ii) *for each $k \in M$, the map $S_k : W \times]-\delta, \delta[\to \mathbb{R}, (x, t) \mapsto S(k, x, t)$ is C^2, and satisfies the Hamilton-Jacobi equation*

$$\frac{\partial S_k}{\partial t} + H(x, \frac{\partial S_k}{\partial x}) = 0;$$

(iii) *both maps* $(k,x,t) \mapsto DS_k(x,t), (k,x,t) \mapsto D^2S_k(x,t)$ *are continuous on the product* $K \times W \times]-\delta,\delta[$.

Proof. We want to find a function $S(k,x,t)$ such that $S(k,x,0) = S_0(k,x)$, and $S_k(x,t) = S(k,x,t)$ is a solution of the Hamilton-Jacobi equation, for each $k \in K$. As we already know the graph $\{(x, \frac{\partial S_k}{\partial x}, t) \mid x \in U\} \subset T^*M \times \mathbb{R}$ must be invariant by the local flow Φ_t^H. This suggests to obtain this graph like a part of the image of the map σ_k defined by

$$\sigma_k(x,t) = \Phi_t^H\big(x, \frac{\partial S_0}{\partial x}(k,x), 0\big) = (\phi_t^H[x, \frac{\partial S_0}{\partial x}(k,x)], t).$$

The map $\sigma(k,x,t) = \sigma_k(x,t)$ is well defined and continuous on an open neighborhood $\mathcal{U}$ of $K \times U \times \{0\}$ in $K \times O \times \mathbb{R}$. The values of σ are in $O \times \mathbb{R}$. Moreover, the map σ_k is C^1 on the open subset $\mathcal{U}_k = \{(x,t) \mid (k,x,t) \in \mathcal{U}\}$, and both maps $(k,x,t) \mapsto \frac{\partial \sigma_k}{\partial t}(k,x,t), (k,x,t) \mapsto \frac{\partial \sigma_k}{\partial x}(k,x,t)$ are continuous on $\mathcal{U}$.

Given its definition, it is not difficult to see that σ_k is injective. Let us show that the image of σ_k is a C^1 odd-Lagrangian submanifold of $O \times \mathbb{R}$. Since $\sigma_k(x,t) = \Phi_t^H\big(x, \frac{\partial S_0}{\partial x}(k,x), 0\big)$ and Φ_t^H is a local diffeomorphism which preserves Ω_H, it is enough to show that the image of the derivative of σ_k at a point of the form $(x_0, 0)$ is odd Lagrangian for Ω_H (and thus of dimension $n+1$ like $\mathcal{U}_k$). Since we have

$$\sigma_k(x,0) = (x, \frac{\partial S_0}{\partial x}(k,x), 0),$$
$$\sigma_k(x_0,t) = \Phi_t^H\big(\sigma_k(x_0,0)\big),$$

we see that the image of $D\sigma_k(x_0,0)$ is the sum of the subspace

$$E = T_{(x_0, \frac{\partial S_0}{\partial x}(k,x_0))} \operatorname{Graph}(d_xS_0) \times \{0\} \subset T_{\sigma_k(x_0,0)}(T_x^*M \times \mathbb{R})$$

and the subspace generated by $X_H(x_0, \frac{\partial S_0}{\partial x}(k,x_0)) + \frac{\partial}{\partial t}$. Note now that the restriction $\Omega_H | T_x^*M \times \{0\}$ is Ω, and that $\operatorname{Graph}(d_xS_0)$ is a Lagrangian subspace for Ω. Moreover the vector $X_H\big(x_0, \frac{\partial S_0}{\partial x}(k,x_0)\big) + \frac{\partial}{\partial t}$ is not in E and generates the characteristic subspace of Ω_H. Therefore the image of $D\sigma_k(x_0,0)$ is indeed odd Lagrangian for Ω_H. This image is the tangent space to the image σ_k at the point $\sigma_k(x_0,0)$. We thus have shown that the whole image of σ_k is an odd

Lagrangian submanifold for Ω_H. In the remainder of this proof, we denote by $\tilde\pi^*$ the projection $\pi^*\times \mathrm{Id}_{\mathbb R} : T^*M\times\mathbb R \to M\times\mathbb R$. Let us show that the derivative of $\tilde\pi^*\sigma_k : \mathcal U_k \to M\times\mathbb R$ is an isomorphism at each point $(x,0)\in U\times\{0\}$. Indeed, we have $\tilde\pi^*\sigma_k(x,0)=(x,0)$, for each $x\in U$, and $\tilde\pi^*\sigma_k(x,t) = (\pi^*\phi^H_t(x,\frac{\partial S_0}{\partial x}(k,x)),t)$, therefore, by writing the derivative $D(\tilde\pi^*\sigma_k)(x,0) : T_xM\times\mathbb R\to T_xM\times\mathbb R$ in matrix form, we find a matrix of the type

$$D(\tilde\pi^*\sigma_k)(x,0) = \begin{bmatrix} \mathrm{Id}_{T_xM} & \star \\ 0 & 1\end{bmatrix}$$

which is an isomorphism. As $K\times\bar W\times\{0\}$ is a compact subset of $\mathcal U$ and the derivative of $\pi^*\sigma_k$ is an isomorphism at each point $(x,0)\in U\times\{0\}$, we can apply the inverse function theorem 2.7.1 to find an open neighborhood $\tilde W$ of $\bar W$ in U and $\epsilon>0$ such that, for each $k\in K$, the map $\tilde\pi^*\sigma_k$ induces a C^1 diffeomorphism of $\tilde W\times]-\epsilon,\epsilon[$ onto an open subset of $M\times\mathbb R$. Since moreover $\pi^*\sigma_k(x,0)=x$, we can apply lemma 2.7.2 to obtain $\delta>0$ such that $\pi^*\sigma_k(\tilde W\times]-\epsilon,\epsilon[)\supset \bar W\times[-\delta,\delta]$, for each $k\in K$. Let us then define the C^1 section $\tilde\sigma_k : W\times]-\delta,\delta[\to T^*M\times\mathbb R$ of the projection $\tilde\pi^* : T^*M\times\mathbb R\to M\times\mathbb R$ by

$$\tilde\sigma_k(x,t) = \sigma_k\big[(\tilde\pi^*\sigma_k)^{-1}(x,t)\big].$$

Since $\tilde\sigma_k$ is a section of $\tilde\pi^*$, we have

$$\tilde\sigma_k(x,t) = (x,p_k(x,t),t),$$

with $p_k(x,t)\in T^*_xM$, and $p_k(x,0)=\frac{\partial S_0}{\partial x}(k,x)$. Moreover, the image of $\tilde\sigma_k$ is an odd Lagrangian submanifold for Ω_H, since it is contained in the image of σ_k, consequently, the differential 1-form

$$\tilde\sigma_k^*\alpha_H = p_k(x,t) - H(x,p_k(x,t))dt$$

is closed on $W\times]-\delta,\delta[$. However, the restriction of this form to $W\times\{0\}$ is $\frac{\partial S_0}{\partial x}(k,x) = dS_{0,k}$ which is exact, therefore, there exists a function $S_k : W\times]-\delta,\delta[\to\mathbb R$ such that $S_k(x,0)=S_{0,k}(x)$ and $dS_k = p(x,t) - H(x,p_k(x,t))dt$. We conclude that $\frac{\partial S_k}{\partial x} = p_k(x,t)$ and $\frac{\partial S_k}{\partial t} = -H(x,p_k(x,t))$. Consequently, the function S_k is a solution of the Hamilton-Jacobi equation and $S_k(x,t) = S_0(k,x) - \int_0^t H(x,p_k(x,s))\,ds$, which makes it possible to see that $(k,x,t)\mapsto S_k(x,t)$ is continuous. The property (iii) results from the two equalities $\frac{\partial S_k}{\partial x} = p_k(x,t)$ and $\frac{\partial S_k}{\partial t} = -H(x,p_k(x,t))$. $\square$

Corollary 2.8.12. *Let $H : O \to \mathbb{R}$ be a C^2 Hamiltonian, defined on the open subset O of the cotangent space T^*M of the manifold M. Call ϕ_t^H the local flow of the Hamiltonian vector field X_H associated to H. Denote by d a distance defining the topology of M. The open ball of center x and radius r for this distance will be denoted by $\mathring{B}_d(x, r)$.*

For every compact $C \subset O$, we can find $\delta, \epsilon > 0$ such that, for each $(x, p) \in C$, there exists a C^2 function $S_{(x,p)} : \mathring{B}_d(x, \epsilon) \times] -\delta, \delta[\to \mathbb{R}$, with $\frac{\partial S_{(x,p)}}{\partial x}(x, 0) = p$, and satisfying the Hamilton-Jacobi equation

$$\frac{\partial S_{(x,p)}}{\partial t} + H\big(x, \frac{\partial S_{(x,p)}}{\partial x}\big) = 0.$$

Proof. Since we do not ask that $S_{(x,p)}$ depends continuously on (x, p), it is enough to show that, for each point (x_0, p_0) in O, there is a compact neighborhood C of (x_0, p_0), contained in O, and satisfying the corollary. We can of course assume that M is an open subset of $\mathbb{R}^n$. In this case $T^*M = M \times \mathbb{R}^{n*}$. To begin with, let us choose a compact neighborhood of (x_0, p_0) of the form $\bar{U} \times K \subset O$, with $\bar{U} \subset M$ and $K \subset \mathbb{R}^{n*}$. The function $S_0 : K \times U \to \mathbb{R}$ defined by

$$S_0(p, x) = p(x),$$

is C^∞ and verifies $(x, \frac{\partial S_0}{\partial x}) = (x, p) \in U \times K \subset O$. Let us choose a neighborhood compact $\bar{W}$ of x_0 with $\bar{W} \subset U$. By the previous theorem, there exists $\delta > 0$, and a function $S : K \times W \times]-\delta, \delta[\to \mathbb{R}$ satisfying the following properties

- $S(p, x, 0) = S_0(p, x) = p(x)$, for each $(p, x) \in K \times W$;
- for each $p \in K$, the map $S_p : W \times] - \delta, \delta[\to \mathbb{R}$ is C^2, and is a solution of the Hamilton-Jacobi equation.

We now choose $\bar{V}$ a compact neighborhood of x_0 contained in W. By compactness of $\bar{V}$, we can find $\epsilon > 0$ such that $B_d(x, \epsilon) \subset W$, for each $x \in \bar{V}$. It then remains to take $C = \bar{V} \times K$ and to define, for $(x, p) \in C$, the function $S_{(x,p)} : B_d(x, \epsilon) \times] - \delta, \delta[\to \mathbb{R}$ by

$$S_{(x,p)}(y, t) = S_p(y, t).$$

It is not difficult to obtain the required properties of $S_{(x,p)}$ from those of S_p. $\square$

Chapter 3

Calculus of Variations for a Lagrangian Convex in the Fibers: Tonelli's Theory

The goal of this chapter is to prove Tonelli's theorem which establishes the existence of minimizing extremal curves. This theory requires the convexity of the Lagrangian in the fibers, and the use of absolutely continuous curves.

Another good reference for this chapter is [BGH98]. Again, we have mainly used [Cla90], [Mn] and the appendix of [Mat91]. However, we have departed from the usual way of showing minimizing properties using Mayer fields. Instead we have systematically used (local solutions) of the the Hamilton-Jacobi equation, since this is the main theme of this book.

3.1 Absolutely Continuous Curves.

Definition 3.1.1 (Absolutely Continuous Curve)**.** A curve $\gamma : [a, b] \to \mathbb{R}^n$ is said to be absolutely continuous, if for each $\epsilon > 0$, there exists $\delta > 0$ such that for each family $]a_i, b_i[_{i\in\mathbb{N}}$ of disjoint intervals included in $[a, b]$, and satisfying $\sum_{i\in N} b_i - a_i < \delta$, we have $\sum_{i\in\mathbb{N}} \|\gamma(b_i) - \gamma(a_i)\| < \epsilon$.

It is clear that such an absolutely continuous map is (uniformly) continuous.

Theorem 3.1.2. *A curve $\gamma : [a, b] \to \mathbb{R}^n$ is absolutely continuous if and only if the following three conditions are satisfied*

(1) *the derivative $\dot{\gamma}(t)$ exists almost everywhere on $[a, b]$;*

(2) *the derivative $\dot{\gamma}$ is integrable for the Lebesgue measure on $[a, b]$.*

(3) *For each $t \in [a, b]$, we have $\gamma(t) - \gamma(a) = \int_a^t \dot{\gamma}(s)\, ds$.*

Proofs of this theorem can be found in books on Lebesgue's theory of integration, see for example [WZ77, Theorem 7.29, page 116]. A proof can also be found in [BGH98, Theorem 2.17]

Lemma 3.1.3. *Suppose that $\gamma : [a, b] \to R^k$ is an absolutely continuous curve. 1) If $f : U \to \mathbb{R}^m$ is a locally Lipschitz map, defined on the neighborhood U of the image $\gamma([a, b])$, then $f \circ \gamma : [a, b] \to R^m$ is also absolutely continuous.*

2) We have

$$\forall t, t' \in [a, b], t \leq t',\ \|\gamma(t') - \gamma(t)\| \leq \int_t^{t'} \|\dot{\gamma}(s)\|\, ds.$$

Proof. To prove the first statement, since $\gamma([a, b])$ is compact, we remark that, cutting down U if necessary, we can assume that f is (globally) Lipschitz. If we call K a Lipschitz constant for f, then for each family $]a_i, b_i[_{i\in\mathbb{N}}$ of disjoint intervals included in $[a, b]$, we have

$$\sum_{i\in\mathbb{N}} \|f \circ \gamma(b_i) - f \circ \gamma(a_i)\| \leq K \sum_{i\in\mathbb{N}} \|\gamma(b_i) - \gamma(a_i)\| \epsilon.$$

Therefore the absolute continuity of $f \circ \gamma$ follows from that of γ.

To prove the second statement, we choose some $p \in \mathbb{R}^k$ with $\|p\| = 1$. The curve $p \circ \gamma : [a, b] \to \mathbb{R}$ is absolutely continuous with derivative equal almost everywhere to $p(\dot{\gamma}(t))$. Therefore by theorem 3.1.2 above

$$\begin{aligned} p(\gamma(t') - \gamma(t)) &= p \circ \gamma(t') - p \circ \gamma(t)) \\ &= \int_t^{t'} p(\dot\gamma(s))\, ds \\ \|eq &\int_t^{t'} \|p\| \|\dot\gamma(s)\|\, ds \\ &= \int_t^{t'} \|\dot\gamma(s)\|\, ds. \end{aligned}$$

It now suffices to observe that

$$\|\gamma(t') - \gamma(t)\| = \sup_{\|p\|=1} p(\gamma(t') - \gamma(t))$$

to finish the proof. □

The following proposition will be useful in the sequel.

Proposition 3.1.4. *Let $\gamma_n : [a,b] \to \mathbb{R}^k$ be a sequence of absolutely continuous curves. We suppose that the sequence of derivatives $\dot\gamma_n : [a,b] \to \mathbb{R}^k$ (which exist a.e.) is uniformly integrable for the Lebesgue measure m on $[a,b]$, i.e. for each $\epsilon > 0$, there exists $\delta > 0$, such that if $A \subset [a,b]$ is a Borel subset with its Lebesgue measure $m(A) < \delta$ then $\int_A \|\dot\gamma_n(s)\|\, ds < \epsilon$, for each $n \in \mathbb{N}$.*

If for some $t_0 \in [a,b]$ the sequence $\gamma_n(t_0)$ is bounded in norm, then there is a subsequence $\gamma_{n_j} : [a,b] \to \mathbb{R}^k$ which converges uniformly to a curve $\gamma : [a,b] \to \mathbb{R}^k$. The map γ is absolutely continuous, and the sequence of derivatives $\dot\gamma_{n_j}$ converges to the derivative $\dot\gamma$ in the weak topology $\sigma(L^1, L^\infty)$, which means that for each function $\Phi : [a,b] \to \mathbb{R}^{k*}$ measurable and bounded, we have

$$\int_a^b \Phi(s)(\dot\gamma_n(s))\, ds \to_{n\to+\infty} \int_a^b \Phi(s)(\dot\gamma(s))\, ds.$$

Proof. We first show that the sequence γ_n is equicontinuous. If $\epsilon > 0$ is fixed, let $\delta > 0$ be the corresponding given by the condition that the sequence γ_n is uniformly integrable. If $t < t'$ are such that $t' - t < \delta$ then, for each $n \in \mathbb{N}$

$$\begin{aligned} \|\gamma_n(t') - \gamma_n(t)\| &\leq \int_{t'}^{t} \|\dot\gamma_n(s)\|\, ds \\ &< \epsilon. \end{aligned}$$

Since the $\gamma_n(t_0)$ form a bounded sequence, and the γ_n are equicontinuous, by Ascoli's theorem, we can find a subsequence γ_{n_j} such that γ_{n_j} converges uniformly to $\gamma : [a,b] \to \mathbb{R}^n$, which is therefore continuous. Let us show that γ is absolutely continuous. We fix $\epsilon > 0$, and pick the corresponding $\delta > 0$ given by the condition that the sequence γ_n is uniformly integrable. If $(]a_i, b_i[)_{i\in\mathbb{N}}$ is a sequence of disjoint open intervals, and such that $\sum_{i\in\mathbb{N}}(b_i - a_i) < \delta$, we have

$$\begin{aligned}\forall n \in \mathbb{N}, \sum_{i\in\mathbb{N}} \|\gamma_n(b_i) - \gamma_n(a_i)\|\| &\leq \sum_{i\in\mathbb{N}} \int_{a_i}^{b_i} \|\dot\gamma_n(s)\|\, ds \\ &= \int_{\bigcup_{i\in\mathbb{N}}]a_i, b_i[} \|\dot\gamma_n(s)\| ds \\ &< \epsilon,\end{aligned}$$

Taking the limit we also get

$$\sum_{i\in\mathbb{N}} \|\gamma(b_i) - \gamma(a_i)\| \leq \epsilon.$$

The curve γ is thus absolutely continuous. The derivative $\dot\gamma : [a,b] \to \mathbb{R}^n$ therefore exists, for Lebesgue almost any point of $[a,b]$, and we have

$$\forall t, \in [a,b], \gamma(t) - \gamma(t') = \int_{t'}^{t} \dot\gamma(s)\, ds.$$

It remains to show that $\dot\gamma_{n_j}$ tends to $\dot\gamma$ in the $\sigma(L^1, L^\infty)$ topology. The convergence of γ_{n_j} to γ shows that for any interval $[t,t'] \subset [a,b]$ we have

$$\int_t^{t'} \dot\gamma_{n_j}(s)\, ds = \gamma_{n_j}(t') - \gamma_{n_j}(t) \to \gamma(t') - \gamma(t) = \int_t^{t'} \dot\gamma(s)\, ds.$$

If U is an open subset of $[a,b]$ let us show that $\int_U \dot\gamma_{n_j}(s)\, ds \to \int_U \dot\gamma(s)\, ds$. We can write $U = \bigcup_{i\in I}]a_i, b_i[$ with the $]a_i, b_i[$ disjoint, where I is at most countable. If the set of the indices I is finite, we apply what precedes each interval $]a_i, b_i[$, and adding, we obtain the convergence

$$\int_U \dot\gamma_{n_j}(s)\, ds \to \int_U \dot\gamma(s)\, ds.$$

To deal with the case were I is infinite and countable, we fix $\epsilon > 0$ and choose the corresponding $\delta > 0$ given by the fact that the sequence γ_n is uniformly integrable. We can find a finite subset $I_0 \subset I$ such that $\sum_{i\in I\setminus I_0}(b_i - a_i) < \delta$. We have

$$\forall n \in \mathbb{N}, \sum_{i\in I\setminus I_0} \|\gamma_n(b_i) - \gamma_n(a_i)\| < \epsilon.$$

Taking limits, we then obtain

$$\sum_{i\in I\setminus I_0} \|\gamma(b_i) - \gamma(a_i)\| \le \epsilon.$$

We pose $U_0 = \bigcup_{i\in I_0}]a_i, b_i[$, so we can write

$$\begin{aligned}\|\int_{U\setminus U_0} \dot\gamma_n(s)\,ds\| &= \|\sum_{i\in I\setminus I_0} \dot\gamma_n(b_i) - \gamma_n(a_i)\| \\ &\le \sum_{i\in I\setminus I_0} \|\gamma_n(b_i) - \gamma_n(a_i)\| \\ &< \epsilon,\end{aligned}$$

and also

$$\|\int_{U\setminus U_0} \dot\gamma(s)\,ds\| \le \epsilon.$$

As I_0 is finite, we have

$$\lim_{j\to\infty} \int_{U_0} \dot\gamma_{n_j}(s)\,ds = \int_{U_0} \dot\gamma(s)\,ds.$$

We conclude that

$$\limsup_{j\to\infty} \|\int_U \dot\gamma_{n_j}(s)\,ds - \int_U \dot\gamma(s)\,ds\| \le 2\epsilon.$$

Since ϵ is arbitrary, we see that

$$\int_U \dot\gamma_{n_j}(s)\,ds \to \int_U \dot\gamma(s)\,ds.$$

If we take now an arbitrary measurable subset A of $[a, b]$, we can find a decreasing sequence $(U_\ell)_{\ell\in\mathbb{N}}$ of open subset with $A \subset$

$\bigcap_{\ell\in\mathbb{N}} U_\ell$ and the Lebesgue measure $m(U_\ell)$ decreasing to $m(A)$, where m is the Lebesgue measure. We fix $\epsilon > 0$, and we choose the corresponding $\delta > 0$ given by the fact that the sequence γ_n is uniformly integrable. As the U_ℓ form a decreasing sequence of sets of finite measure and $m(U_\ell) \searrow m(A)$, by Lebesgue's theorem of dominated convergence, we have

$$\int_{U_\ell} \dot\gamma(s)\,ds \to \int_A \dot\gamma(s)\,ds.$$

We now fix an integer ℓ big enough to satisfy

$$\|\int_{U_\ell} \dot\gamma(s)\,ds - \int_A \dot\gamma(s)ds\| \le \epsilon,$$

and $m(U_\ell \setminus A) < \delta$. By the choice of δ, we have

$$\forall n \in \mathbb{N}, \|\int_{U_\ell\setminus A} \dot\gamma_n(s)\,ds\| \le \int_{U_\ell\setminus A} \|\dot\gamma_n(s)\|ds \le \epsilon.$$

Since U_ℓ is a fixed open set, taking the limit for $j \to \infty$, we obtain

$$\int_{U_\ell} \dot\gamma_{n_j}(s)\,ds \to \int_{U_\ell} \dot\gamma(s)\,ds.$$

We conclude that

$$\limsup_{j\to\infty} \|\int_A \dot\gamma_{n_j}(s)\,ds - \int_A \dot\gamma(s)\,ds\| \le 2\epsilon,$$

and thus, since $\epsilon > 0$ is arbitrary

$$\int_A \dot\gamma_{n_j}(s)\,ds \to \int_A \dot\gamma(s)\,ds.$$

We then consider the vector space $L^\infty([a,b],\mathbb{R}^{k*})$, formed by the $\Phi : [a,b] \to \mathbb{R}^{k*}$ measurable bounded maps, and provided with the standard norm $\|\Phi\|_\infty = \sup_{t\in[a,b]}\|\Phi(t)\|$. The subset $\mathcal{E}$ formed by the characteristic functions χ_A, where $A \subset [a,b]$ is measurable, generates a vector subspace $\tilde{\mathcal{E}}$ which is dense in $L^\infty([a,b],\mathbb{R}^{k*})$, see exercise 3.1.5 below. The maps $\theta_n : L^\infty([a,b],\mathbb{R}^{k*}) \to \mathbb{R}$ defined by

$$\theta_n(\Phi) = \int_a^b \Phi(s)(\dot\gamma_n(s))\,ds$$

are linear and continuous on $L^\infty([a,b],\mathbb{R}^{k*},m)$, with

$$\|\theta_n\| \leq \int_a^b \|\dot\gamma_n(s)\| ds.$$

We also define $\theta : L^\infty([a,b],\mathbb{R}^{k*},m) \to \mathbb{R}$ by

$$\theta(\Phi) = \int_a^b \Phi(s)(\dot\gamma(s))\, ds,$$

which is also linear and continuous on $L^\infty([a,b],\mathbb{R}^{k*},m)$. We now show that the sequence of norms $\|\theta_n\|$ is bounded. Applying the fact that the sequence of derivatives $\dot\gamma_n$ is uniformly integrable with $\epsilon = 1$, we find the corresponding $\delta_1 > 0$. We can then cut the interval $[a,b]$ into $[(b-a)/\delta_1] + 1$ intervals of length $\leq \delta_1$, where $[x]$ indicates the integer part of the real number x. On each one of these intervals, the integral of $\|\dot\gamma_n\|$ is bounded by 1, hence

$$\forall n \in \mathbb{N}, \|\theta_n\| \leq \int_a^b \|\dot\gamma_n(s)\|\, ds \leq [(b-a)/\delta_1] + 1.$$

As $\theta_{n_j}(\Phi) \to \theta(\Phi)$, for $\Phi \in \mathcal{E}$, by linearity the same convergence is true for $\Phi \in \tilde{\mathcal{E}}$. As $\tilde{\mathcal{E}}$ is dense in $L^\infty([a,b],\mathbb{R}^{k*},m)$ and $\sup_{n\in\mathbb{N}}\|\theta_n\| < +\infty$, a well-known argument of approximation (see the exercise below 3.1.6) shows, then, that

$$\forall \Phi \in L^\infty([a,b],\mathbb{R}^{k*},m),\ \theta_{n_j}(\Phi) \to \theta(\Phi).\square$$

Exercise 3.1.5. *Consider the vector space $L^\infty([a,b],\mathbb{R}^{k*})$, formed by the measurable bounded maps $\Phi : [a,b] \to \mathbb{R}^{k*}$, and provided with the standard norm $\|\Phi\|_\infty = \sup_{t\in[a,b]}\|\Phi(t)\|$. The subset $\mathcal{E}$ formed by the characteristic functions χ_A, where $A \subset [a,b]$ is measurable, generates a vector subspace which we will call $\tilde{\mathcal{E}}$. Show that $\tilde{\mathcal{E}}$ which is dense in $L^\infty([a,b],\mathbb{R}^{k*})$. [Indication: Consider first $\varphi : [a,b] \to]-K,K[$, with $K \in \mathbb{R}_+$, and define $\varphi_n = \sum_{i=-n}^{n-1} \varphi(iK/n)\chi_{A_{i,n}}$, where $A_{i,n}$ is the set $\{t \in [a,b] \mid iK/n \leq \varphi(x) < (i+1)K/n\}$.]*

Exercise 3.1.6. *Let $\theta_n : E \to F$ be a sequence of continuous linear operators between normed spaces. Suppose that the sequence* *of the norms $\|\theta_n\|$ is bounded, and that there exists a continuous linear operator $\theta : E \to F$ and a subset $\mathcal{E} \subset E$ generating a dense subspace of E such that $\theta_n(x) \to \theta(x)$ for every $x \in \mathcal{E}$. Show then that $\theta_n(x) \to \theta(x)$ for every $x \in E$*

We can replace in the definition of an absolutely continuous curve, the norm by any distance Lipschitz-equivalent to a norm. This makes it possible to generalize the definition of absolutely continuous curve to a manifold.

Definition 3.1.7 (Absolutely continuous Curve)**.** Let M be a manifold, we denote by d the metric obtained on M from some fixed Riemannian metric.

A curve $\gamma : [a, b] \to M$ is said to be absolutely continuous if, for each $\epsilon > 0$, there exists $\delta > 0$ such that for any countable family of disjoint intervals $(]a_i, b_i[)_{i\in\mathbb{N}}$ all included in $[a, b]$ and satisfying $\sum_{i\in\mathbb{N}}(b_i - a_i) < \delta$, we have $\sum_{i\in\mathbb{N}} d(\gamma(b_i), \gamma(a_i)) < \epsilon$.

Definition 3.1.8. We denote by $\mathcal{C}^{ac}([a, b], M)$ the space of absolutely continuous curves defined on the compact interval $[a, b]$ with values in the manifold M. This space $\mathcal{C}^{ac}([a, b], M)$ is provided with the topology of uniform convergence.

Lemma 3.1.9. *If $\gamma : [a, b] \to M$ is absolutely continuous, then, the derivative $\dot\gamma(t) \in T_{\gamma(t)}M$ exists for almost every $t \in [a, b]$. If $\gamma([a', b']) \subset U$, where U is the domain of definition of the coordinate chart $\theta : U \to \mathbb{R}^n$, then, we have*

$$\forall t \in [a', b'],\ \theta \circ \gamma(t) - \theta \circ \gamma(a') = \int_{a'}^{t} \frac{d(\theta \circ \gamma)}{dt}(s)\, ds.$$

Exercise 3.1.10. *Suppose that the manifold M is provided with a Riemannian metric. We denote by d the distance induced by this Riemannian metric on M, and by $\|\cdot\|_x$ the norm induced on the fiber T_xM, for $x \in M$. If $\gamma : [a, b] \to M$ is an absolutely continuous curve, show that*

$$d(\gamma(b), \gamma(a)) \leq \int_a^b \|\dot\gamma(s)\|_{\gamma(s)}\, ds.$$

[Indication: Use lemmas 2.7.5, and 3.1.3.]

Definition 3.1.11 (Bounded Below)**.** Let $L : TM \to \mathbb{R}$ be a Lagrangian on the manifold M. We will say that L is bounded below above every compact subset of M, if for every compact $K \subset M$, we can find a constant $C_K \in \mathbb{R}$ such that

$$\forall (x,v) \in \pi^{-1}(K),\ L(x,v) \geq C_K,$$

where $\pi : TM \to M$ is the canonical projection.

If the Lagrangian $L : TM \to \mathbb{R}$ is bounded below above every compact subset of the manifold, we can define the action of the Lagrangian for an absolutely continuous curve admitting possibly $+\infty$ as a value. Indeed, if $\gamma \in \mathcal{C}^{ac}([a,b],M)$, since $\gamma([a,b])$ is compact, there exists a constant $C_\gamma \in \mathbb{R}$ such that

$$\forall t \in [a,b], \forall v \in T_{\gamma(t)}M,\ L(\gamma(t),v) \geq C_\gamma.$$

In particular, the function $t \mapsto L(\gamma(t),\dot\gamma(t)) - C_\gamma$ is well defined and positive almost everywhere for the Lebesgue measure. Therefore $\int_a^b [L(\gamma(t),\dot\gamma(T)) - C_\gamma]\,ds$ makes sense and belongs to $[0,+\infty]$. We can then set

$$\begin{aligned}\mathbb{L}(\gamma) &= \int_a^b L(\gamma(t),\dot\gamma(t))ds\\ &= \int_a^b [L(\gamma(t),\dot\gamma(t)) - C_\gamma]ds + C_\gamma(b-a).\end{aligned}$$

This quantity is clearly in $\mathbb{R}\cup\{+\infty\}$. It is not difficult to see that the action $\mathbb{L}(\gamma)$ is well defined (i.e. independent of the choice of lower bound C_γ).

3.2 Lagrangian Convex in the Fibers

In this section, we consider a manifold M provided with a Riemannian metric of reference. For $x \in M$, we denote by $\|\cdot\|_x$ the norm induced on T_xM by the Riemannian metric. We will denote by d the distance induced on M by the Riemannian metric. The canonical projection of TM on M is as usual $\pi : TM \to M$.

We will consider a C^1 Lagrangian L on M which is convex in the fibers, and superlinear above every compact subset of M, see definitions 1.3.11 and 1.3.13. Note that, for Lagrangians convex in the fibers, the superlinearity above compact subset is equivalent to the restriction $L|T_xM : T_xM \to \mathbb{R}$ is superlinear, for each $x \in M$, see 1.3.14. It is clear that a Lagrangian superlinear above every compact subset of M is also bounded below above every compact subset of M, therefore we can define the action for absolutely continuous curves.

Theorem 3.2.1. *Suppose that $L : TM \to \mathbb{R}$ is a C^1 Lagrangian convex in the fibers, and superlinear above compact subsets of M. If a sequence of curves $\gamma_n \in \mathcal{C}^{ac}([a,b], M)$ converges uniformly to the curve $\gamma : [a,b] \to M$ and*

$$\liminf_{n\to\infty} \mathbb{L}(\gamma_n) < +\infty,$$

then the curve γ is also absolutely continuous and

$$\liminf_{n\to\infty} \mathbb{L}(\gamma_n) \geq \mathbb{L}(\gamma).$$

Proof. Let us start by showing how we can reduce the proof to the case where M is an open subset of $\mathbb{R}^k$, where $k = \dim M$.

Consider the set $K = \gamma([a,b]) \cup \bigcup_n \gamma_n([a,b])$, it is compact because γ_n converges uniformly to γ. By superlinearity of L above each compact subset of M, we can find a constant C_0 such that

$$\forall x \in K, \forall v \in T_xM,\ L(x,v) \geq C_0.$$

If $[a',b']$ is a subinterval of $[a,b]$, taking C_0 as a lower bound of $L(\gamma_n(s), \dot\gamma_n(s))$ on $[a,b] \setminus [a',b']$, we see that

$$\forall n,\ \mathbb{L}(\gamma_n|[a',b']) \leq \mathbb{L}(\gamma_n) - C_0[(b-a) - (b'-a')].$$

It follows

$$\forall [a',b'] \subset [a,b],\ \liminf_{n\to\infty} \mathbb{L}(\gamma_n|[a',b']) < +\infty. \tag{*}$$

By continuity of $\gamma : [a,b] \to M$, we can find a finite sequence $a_0 = a < a_1 < \cdots < a_p = b$ and a sequence of domains of coordinate charts $U_1, \ldots, U_p$ such that $\gamma([a_{i-1}, a_i]) \subset U_i$, for $i = 1, \ldots, p$. Since γ_n converges uniformly to γ, forgetting the first curves γ_n if necessary, we can assume that $\gamma_n([a_{i-1}, a_i]) \subset U_i$, for $i = 1, \ldots, p$.

By condition $(*)$, we know that $\liminf_{n\to\infty} \mathbb{L}(\gamma_n|[a_{i-1}, a_i]) < \infty$, it is then enough to show that this condition implies that $\gamma|[a_{i-1}, a_i]$ is absolutely continuous and that

$$\mathbb{L}(\gamma|[a_{i-1}, a_i]) \leq \liminf_{n\to\infty} \mathbb{L}(\gamma_n|[a_{i-1}, a_i]),$$

because we have $\liminf(\alpha_n + \beta_n) \geq \liminf \alpha_n + \liminf \beta_n$, for sequences of real numbers α_n and β_n. As the U_i are domains of definition of coordinates charts, we do indeed conclude that it is enough to show the theorem in the case where M is an open subset of $\mathbb{R}^k$.

In the sequel of the proof, we will thus suppose that $M = U$ is an open subset of $\mathbb{R}^k$ and thus $TU = U \times \mathbb{R}^k$ and that $\gamma([a, b])$ and all the $\gamma_n([a, b])$ are contained in the same compact subset $K_0 \subset U$. Let us set $\ell = \liminf_{\gamma_n\to\infty} \mathbb{L}(\gamma_n)$. Extracting a subsequence such that $\mathbb{L}(\gamma_n) \to \ell < +\infty$ and forgetting some of the first curves γ_n, we can suppose that

$$\mathbb{L}(\gamma_n) \to \ell \text{ and } \forall n \in \mathbb{N}, \mathbb{L}(\gamma_n) \leq \ell + 1 < +\infty.$$

We denote by $\|\cdot\|$ a norm on $\mathbb{R}^k$.

Lemma 3.2.2. *If $C \geq 0$ is a constant, $K \subset U$ is compact, and $\epsilon > 0$, we can find $\eta > 0$ such that for each $x, y \in U$ with $x \in K$ and $\|y - x\| < \eta$, and for each $v, w \in \mathbb{R}^k$ with $\|v\| \leq C$, we have*

$$L(y, w) \geq L(x, v) + \frac{\partial L}{\partial v}(x, v)(w - v) - \epsilon.$$

Proof. Let us choose $\eta_0 > 0$ such that the set

$$\bar{V}_{\eta_0}(K) = \{y \in \mathbb{R}^k \mid \exists x \in K, \|y - x\| \leq \eta_0\}$$

is a compact subset of U.

We denote by A the finite constant

$$A = \sup\{\|\frac{\partial L}{\partial x}(x, v)\| \mid x \in K, \|v\| \leq C\}.$$

Since L is uniformly superlinear above every compact subset of $M = U$, we can find a constant $C_1 > -\infty$ such that

$$\forall y \in \bar{V}_{\eta_0}(K), \forall w \in \mathbb{R}^k, L(y, w) \geq (A + 1)\|w\| + C_1.$$

We then set

$$C_2 = \sup\{L(x,v) - \frac{\partial L}{\partial x}(x,v)(v) \mid x \in K, \|v\| \leq C\}.$$

By compactness the constant C_2 is finite. We remark that if $\|w\| \geq C_2 - C_1$, then for $y \in \bar{V}_{\eta_0}(K), x \in K$ and $\|v\| \leq C$, we have

$$\begin{aligned} L(y,w) &\geq (A+1)\|w\| + C_1 \\ &\geq A\|w\| + (C_2 - C_1) + C_1 \\ &= A\|w\| + C_2 \\ &\geq \frac{\partial L}{\partial v}(x,v)(w) + \big(L(x,v) - \frac{\partial L}{\partial v}(x,v)(v)\big) \\ &= L(x,v) + \frac{\partial L}{\partial v}(x,v)(w-v). \end{aligned}$$

It then remains to find $\eta \leq \eta_0$, so that we satisfy the sought inequality when $\|w\| \leq C_2 - C_1$. But the set

$$\{(x,v,w) \mid x \in K, \|v\| \leq C, \|w\| \leq C_2 - C_1\}$$

is compact and $L(x,w) \geq L(x,v) + \frac{\partial L}{\partial v}(x,v)(w-v)$ by convexity of L in the fibers of the tangent bundle TU. It follows that for $\epsilon > 0$ fixed, we can find $\eta > 0$ with $\eta \leq \eta_0$ and such that, if $x \in K, \|y - x\| \leq \eta, \|v\| \leq C$ et $\|w\| \leq C_2 - C_1$, we have

$$L(y,w) > L(x,v) + \frac{\partial L}{\partial x}(x,v)(w-v) - \epsilon. \qquad \square$$

We return to the sequence of curves $\gamma_n : [a,b] \to U$ which converges uniformly to $\gamma : [a,b] \to U$. We already have reduced the proof to the case where $\gamma([a,b]) \cup \bigcup_{n \in N} \gamma_n([a,b])$ is included in a compact subset $K_0 \subset U$ with U an open subset of $\mathbb{R}^k$. We now show that the derivatives $\dot{\gamma}_n$ are uniformly integrable. Since L is superlinear above each compact subset of U, we can find a constant $C(0) > -\infty$ such that

$$\forall x \in K_0, \forall v \in \mathbb{R}^k, L(x,v) \geq C(0).$$

We recall that

$$\forall n \in \mathbb{N}, \mathbb{L}(\gamma_n) \leq \ell + 1 < +\infty.$$

We then fix $\epsilon > 0$ and we take $A > 0$ such that

$$\frac{\ell + 1 - C(0)(b-a)}{A} < \epsilon/2.$$

Again by the superlinearity of L above compact subsets of the base M, there exists a constant $C(A) > -\infty$ such that

$$\forall x \in K_0, \forall v \in \mathbb{R}^k,\ L(x,v) \geq A\|v\| + C(A).$$

Let $E \subset [a,b]$ be a measurable subset, we have

$$C(A)m(E) + A\int_E \|\dot\gamma_n(s)\| ds \leq \int_E L(\gamma_n(s), \dot\gamma_n(s))\, ds$$

and also

$$C(0)(b-a-m(E)) \leq \int_{[a,b]\setminus E} L(\gamma_n(s), \dot\gamma_n(s))\, ds.$$

Adding the inequalities and using $\mathbb{L}(\gamma_n) \leq \ell + 1$, we find

$$[C(A) - C(0)]m(E) + C(0)(b-a) + A\int_E \|\dot\gamma_n(s)\| ds \leq \ell + 1,$$

this in turn yields

$$\begin{aligned}\int_E \|\dot\gamma_n(s)\| ds &\leq \frac{\ell + 1 - C(0)(b-a)}{A} + \frac{[C(0) - C(A)]}{A} m(E)\\ &\leq \epsilon/2 + \frac{[C(0) - C(A)]}{A} m(E).\end{aligned}$$

If we choose $\delta > 0$ such that $\frac{C(0)-C(A)}{A}\delta < \epsilon/2$, we see that

$$m(E) < \delta \Rightarrow \forall n \in \mathbb{N}, \int_E \|\dot\gamma_n(s)\|\, ds < \epsilon.$$

This finishes to proves the uniform integrability of the sequence $\dot\gamma_n$. We can then conclude by proposition 3.1.4 that $\dot\gamma_n$ converges to $\dot\gamma$ in the $\sigma(L^1, L^\infty)$ topology. We must show that $\lim_{n\to\infty} \mathbb{L}(\gamma_n) \geq \mathbb{L}(\gamma)$.

Let C be a fixed constant, we set

$$E_C = \{t \in [a,b] \mid \|\dot\gamma(t)\| \leq C\}.$$

We fix $\epsilon > 0$ and we apply lemma 3.2.2 with this ϵ, the constant C fixed above and the compact set $K = K_0 \supset \gamma([a,b]) \cup \bigcup_{n\in\mathbb{N}} \gamma_n[a,b]$ to find $\eta > 0$ as in the conclusion of the lemma 3.2.2. Since $\gamma_n \to \gamma$ uniformly, there exists an integer n_0 such that for each $n \geq n_0$, we have $\|\gamma_n(t) - \gamma(t)\| < \eta$, for each $t \in [a,b]$. Lemma 3.2.2 then shows that for each $n \geq n_0$ and almost all $t \in E_C$, we have

$$L(\gamma_n(t), \dot\gamma_n(t)) \geq L(\gamma(t), \dot\gamma(t)) + \frac{\partial L}{\partial v}(\gamma(t), \dot\gamma(t))(\dot\gamma_n(t) - \dot\gamma(t)) - \epsilon,$$

hence using this, together with the inequality $L(\gamma_n(t), \dot\gamma_n(t)) \geq C(0)$ which holds almost everywhere, we obtain

$$\begin{aligned}\mathbb{L}(\gamma_n) \geq &\int_{E_C} L(\gamma(t), \dot\gamma(t))\, dt + C(0)[(b-a) - m(E_C)] \\ &+ \int_{E_C} \frac{\partial L}{\partial v}(\gamma(t), \dot\gamma(t))(\dot\gamma_n(t) - \dot\gamma(t))\, dt - \epsilon m(E_C). \qquad (*)\end{aligned}$$

Since $\|\dot\gamma(t)\| \leq C$, for $t \in E_C$, the map $t \to \chi_{E_C}(t)\frac{\partial L}{\partial v}(\gamma(t), \dot\gamma(t))$ is bounded. But $\dot\gamma_n \to \dot\gamma$ for topology $\sigma(L^1, L^\infty)$, therefore

$$\int_{E_C} \frac{\partial L}{\partial v}(\gamma(t), \dot\gamma(t))(\dot\gamma_n(t) - \dot\gamma(t))\, dt \to 0, \text{ when } n \to \infty.$$

Going to the limit in the inequality (*), we find

$$\begin{aligned}\ell &= \lim_{n\to\infty} \mathbb{L}(\gamma_n) \\ &\geq \int_{E_C} L(\gamma(t), \dot\gamma(t))\, dt + C(0)[(b-a) - m(E_C)] - \epsilon m(E_C).\end{aligned}$$

We can then let $\epsilon \to 0$ to obtain

$$\ell \geq \int_{E_C} L(\gamma(t), \dot\gamma(t))\, dt + C(0)[(b-a) - m(E_C)]. \qquad (**)$$

Since the derivative $\dot\gamma(t)$ exists and is finite for almost all $t \in [a,b]$, we find $E_C \nearrow E_\infty$, when $C \nearrow +\infty$, with $[a,b] \setminus E_\infty$ of zero Lebesgue measure. Since $L(\gamma(t), \dot\gamma(t))$ is bounded below by $C(0)$, we have by the monotone convergence theorem

$$\int_{E_C} L(\gamma(t), \dot\gamma(t))\, dt \to \int_a^b L(\gamma(t), \dot\gamma(t))\, dt, \text{ when } C \to \infty.$$

If we let $C \nearrow +\infty$ in $(**)$, we find

$$\ell = \lim_{n\to\infty} \mathbb{L}(\gamma_n) \geq \int_a^b L(\gamma(t), \dot\gamma(t))\,dt = \mathbb{L}(\gamma).$$

□

Corollary 3.2.3. *Suppose $L : TM \to \mathbb{R}$ is a C^1 Lagrangian convex in the fibers, and superlinear above every compact subset of M. Then the action $\mathbb{L} : \mathcal{C}^{ac}([a,b],M) \to \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous for the topology of uniform convergence on $\mathcal{C}^{ac}([a,b],M)$. In particular, on any compact subset of $\mathcal{C}^{ac}([a,b],M)$, the action $\mathbb{L}$ achieves its infimum.*

Proof. It is enough to see that if $\gamma_n \to \gamma$ uniformly, with all the γ_n and γ in $\mathcal{C}^{ac}([a,b],M)$, then, we have

$$\liminf_{\gamma_n\to\infty} \mathbb{L}(\gamma_n) \geq \mathbb{L}(\gamma).$$

If $\liminf_{n\to\infty} \mathbb{L}(\gamma_n) = +\infty$, there is nothing to show. The case where we have $\liminf_{n\to\infty} \mathbb{L}(\gamma_n) < +\infty$ results from theorem 3.2.1. □

3.3 Tonelli's Theorem

Corollary 3.3.1 (Compact Tonelli)**.** *Let $L : TM \to \mathbb{R}$ be a C^1 Lagrangian convex in the fibers, and superlinear above every compact subset of the manifold M. If $K \subset M$ is compact, and $C \in \mathbb{R}$, then the subset*

$$\Sigma_{K,C} = \{\gamma \in \mathcal{C}^{ac}([a,b],M) \mid \gamma([a,b]) \subset K, \mathbb{L}(\gamma) \leq C\}$$

is a compact subset of $\mathcal{C}^{ac}([a,b],M)$ endowed with the topology of uniform convergence.

Proof. By the compactness of K, and theorem 3.2.1, the subset $\Sigma_{K,C}$ is closed in the space $\mathcal{C}([a,b],M)$ of all the continuous curves, for the topology of uniform convergence. Since $\gamma([a,b]) \subset K$, for each $\gamma \in \Sigma_{K,C}$, by Ascoli's theorem, it is enough to see that the family of the $\gamma \in \Sigma_{K,C}$ is equicontinuous. Let us then fix some Riemannian metric on M. We denote by d the distance induced on M, and by $\|\cdot\|_x$ the norm induced on T_xM, for $x \in M$. Using the superlinearity of L above the compact the compact set K, for each $A \geq 0$, we can find $C(A)$ such that

$$\forall q \in K, \forall v \in T_qM, C(A) + A\|v\|_x \leq L(q,v).$$

Applying this with $(q,v) = (\gamma(s), \dot{\gamma}(s))$, and integrating, we see that for each $\gamma \in \Sigma_{K,C}$, we have

$$C(A)(t'-t) + A\int_t^{t'} \|\dot{\gamma}(s)\|_{\gamma(s)}\,ds \leq \int_t^{t'} L(\gamma(s),\dot{\gamma}(s))\,ds$$
$$C(0)[(b-a)-(t'-t)] \leq \int_{[a,b]\setminus[t,t']} L(\gamma(s),\dot{\gamma}(s))\,ds.$$

Adding these two inequalities and using the following one

$$d(\gamma(t'),\gamma(t)) \leq \int_t^{t'} \|\dot{\gamma}(s)\|_{\gamma(s)}\,ds,$$

we obtain

$$Ad(\gamma(t'),\gamma(t)) \leq [\mathbb{L}(\gamma) - C(0)(b-a)] + (C(0)-C(A))(t'-t).$$

If $\epsilon > 0$ is given, we choose A such that $[C - C(0)(b-a)]/A \leq \epsilon/2$. It follows that for $\gamma \in \Sigma_{K,C}$ and $t,t' \in [a,b]$, we have

$$d(\gamma(t'),\gamma(t)) \leq \frac{\epsilon}{2} + \frac{C(0)-C(A)}{A}(t'-t).$$

We therefore conclude that the family of curves in $\Sigma_{K,C}$ is equicontinuous. $\square$

Definition 3.3.2. We will say that the Lagrangian $L : TM \to \mathbb{R}$ is bounded below by the Riemannian metric g on M, if we can find a constant $C \in \mathbb{R}$ such that

$$\forall (x,v) \in TM, L(x,v) \geq \|v\|_x + C,$$

where $\|\cdot\|_x$ is the norm on T_xM obtained from the Riemannian metric.

This is a global condition which is relevant only when M is not compact and the Riemannian metric g is complete.

Theorem 3.3.3 (Non Compact Tonelli). *Let L be a C^1 Lagrangian on the manifold M. Suppose that the Lagrangian L is convex in fibers, superlinear above every compact subset of M, and that it is bounded below by a complete Riemannian metric on M. If $K \subset M$ is compact and $C \in \mathbb{R}$, then the subset*

$$\tilde{\Sigma}_{K,C} = \{\gamma \in \mathcal{C}^{ac}([a,b], M) \mid \gamma([a,b]) \cap K \neq \emptyset, \mathbb{L}(\gamma) \leq C\}$$

is a compact subset of $\mathcal{C}^{ac}([a,b], M)$ for the topology of uniform convergence.

Proof. We denote by d the distance on M obtained from the Riemannian metric, and, for $x \in M$, by $\|\cdot\|_x$ the norm induced on T_xM by this same Riemannian metric. We first show that there exists a constant $r < +\infty$ such that

$$\forall \gamma \in \tilde{\Sigma}_{K,C}, \gamma([a,b]) \subset \bar{V}_r(K),$$

where $\bar{V}_r(K) = \{y \in M \mid \exists x \in K, d(y,x) \leq r\}$. Indeed, there exists a constant $C_0 \in \mathbb{R}$ such that

$$\forall (x,v) \in TM, L(x,v) \geq \|v\|_x + C_0.$$

Therefore for every absolutely continuous curve $\gamma : [a,b] \to M$, and every $t, t' \in [a,b]$, with $t \leq t'$, we have

$$C_0(t'-t) + \int_t^{t'} \|\dot{\gamma}(s)\|_{\gamma(s)} ds \leq \mathbb{L}(\gamma),$$

For $\gamma \in \Sigma_{K,C}$, it follows that

$$d(\gamma(t'), \gamma(t)) \leq C + |C_0|(b-a).$$

If we set $r = C + |C_0|$,we get

$$\forall \gamma \in \tilde{\Sigma}_{K,C}, \gamma([a,b]) \subset \bar{V}_r(K).$$

Since the Riemannian metric on M is complete, the d-balls of M of finite radius are compact, and thus $\bar{V}_r(K)$ is also a compact subset of M. By the compact case of Tonelli's Theorem 3.3.1, the set $\Sigma_{\bar{V}_r(K),C}$ is compact in $\mathcal{C}^{ac}([a,b], M)$. Since $\tilde{\Sigma}_{K,C}$ is closed, and contained in $\Sigma_{\bar{V}_r(K),C}$, it is thus also compact in $\mathcal{C}^{ac}([a,b], M)$. ∎

Corollary 3.3.4 (Tonelli Minimizers)**.** *Let M be a connected manifold. Suppose $L : TM \to \mathbb{R}$ is a C^1 Lagrangian convex in fibers, superlinear above every compact subset of M, and bounded below by a complete Riemannian metric on M. Then, for each $x, y \in M$, and each $a, b \in \mathbb{R}$, with $a < b$, there exists an absolutely continuous curve $\gamma : [a, b] \to M$ with $\gamma(a) = x, \gamma(b) = y$ which is a minimizer for $\mathcal{C}^{ac}([a, b], M)$.*

Proof. Let us set $C_{\inf} = \inf \mathbb{L}(\gamma_1)$, where the infimum is taken on the absolutely continuous curves $\gamma_1 : [a, b] \to M$ with $\gamma_1(a) = x$ and $\gamma_1(b) = y$, $C_{\inf} < +\infty$. This quantity makes sense since there exists a C^1 curve between x to y. For each integer $n \geq 1$, we define the subset $\mathcal{C}_n$ of $\mathcal{C}^{ac}([a, b], M)$ formed by the curves $\gamma : [a, b] \to M$ such that

$$\gamma(a) = x, \gamma(b) = y \text{ and } \mathbb{L}(\gamma) \leq C_{\inf} + \frac{1}{n}.$$

This set is by definition a nonempty subset of $C_{\inf}$. It is also compact in $\mathcal{C}^{ac}([a, b], M)$ by the previous theorem 3.3.3. Since the sequence $\mathcal{C}_n$ is decreasing, the intersection $\bigcap_{n \geq 1} \mathcal{C}_n$ is nonempty. Any curve $\gamma : [a, b] \to M$ in this intersection is such that $\gamma(a) = x, \gamma(b) = y$ and $\mathbb{L}(\gamma) = C_{\inf}$. $\square$

3.4 Tonelli Lagrangians

Definition 3.4.1 (Tonelli Lagrangian)**.** A Lagrangian L on the manifold M is called a Tonelli Lagrangian if it satisfies the following conditions:

(1) $L : TM \to \mathbb{R}$ is of class at least C^2.

(2) For each $(x, v) \in TM$, the second partial derivative $\partial^2 L / \partial v^2(x, v)$ is positive definite as a quadratic form.

(3) L is superlinear above compact subset of M.

Condition (2) is equivalent to:

(2') L non-degenerate and convex in the fibers.

Theorem 3.4.2. *If L is a C^r Tonelli Lagrangian on the manifold M, then we have the following properties:*

(1) *For each $x \in M$, the restriction $L|T_xM$ is strictly convex.*

(2) *The Legendre transform*

$$\tilde{\mathcal{L}} : TM \to T^*M, (x, v) \mapsto \big(x, \frac{\partial L}{\partial v}(x, v)\big)$$

is a diffeomorphism of class C^{r-1}.

(3) *The Euler-Lagrange vector field X_L on TM is well defined, of class C^{r-1} and uniquely integrable, and the(local) flow ϕ_t^L of X_L is of class C^{r-1}.*

(4) *The extremal curves are* C^r.

(5) *The (continuous) piecewise C^1 minimizing curves are extremal curves, and therefore of class C^r.*

(6) *The Hamiltonian associated with L, denoted $H : T^*M \to \mathbb{R}$, is well-defined by*

$$\forall (x, v) \in TM,\ H(\tilde{\mathcal{L}}(x, v)) = \frac{\partial L}{\partial v}(x, v)(v) - L(x, v).$$

It is of class C^r. We have the have the Fenchel inequality

$$p(v) \leq L(x, v) + H(x, p),$$

with equality if and only if $p = \partial L/\partial v(x, v)$, or equivalently $(x, p) = \tilde{\mathcal{L}}(x, v)$. Therefore

$$\forall (x, p) \in T^*M,\ H(x, p) = \sup_{v \in T_xM} p(v) - L(x, v).$$

(7) *The (local) Hamiltonian flow ϕ_t^H of H is conjugated by $\tilde{\mathcal{L}}$ to the Euler-Lagrange flow ϕ_t^L.*

Proof. (1) This is clear since $\partial^2 L/\partial v^2(x, v)$ is positive definite, see Proposition 1.1.2.

(2) Note that, by the superlinearity and the strict convexity of $L|T_xM$, for each $x \in M$ the Legendre transtorm $T_xM \to T_x^*M, v \mapsto \partial L/\partial v(x, v)$ is bijective, see 1.4.6. Therefore the global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ is bijective, and we obtain from Proposition 2.1.6 that it is a diffeomorphism.

(3) See Theorem 2.6.5.

(4) In fact, if γ is an extremal, its speed curve $t \mapsto (\gamma(t), \dot\gamma(t))$ is a piece of an orbit of ϕ_t^L, and is therefore of class C^{r-1}.

(5) See Proposition 2.3.7.

(6) See Proposition 2.6.3, and Fenchel's Theorem 1.3.6.

(7) See Theorem 2.6.4. □

Lemma 3.4.3. *Suppose that L is a Tonelli Lagrangian on the manifold M. Its associated Hamiltonian $H : T^*M \to \mathbb{R}$ is also superlinear above each compact subset of M. In particular, if $C \in \mathbb{R}$ and $K \subset M$ is compact, the set*

$$\{(x,p) \in T^*M \mid x \in K, H(x,p) \le C\},$$

is a compact subset of T^*M.

Proof. We know that in this case

$$H(x,p) = \sup_{v \in T_xM} \langle p, v\rangle - L(x,v).$$

Therefore we can apply theorem 1.3.12 to conclude that H is superlinear above compact subset. In particular, if $K \subset M$ is compact, and we fix Riemannian metric g on M, we can find a constant $C_1(K) > -\infty$ such that

$$\forall (x,p) \in T_x^*M, H(x,p) \ge \|p\|_x + C_1(K),$$

where $\|\cdot\|_x$ is the norm on T_x^*M obtained from g. Therefore the closed set $\{(x,p) \in T^*M \mid x \in K, H(x,p) \le C\}$ is contained in $\{(x,p) \in T^*M \mid x \in K, \|p\|_x \le C - C_1(K)\}$. But this last set is a compact subset of T^*M. □

Corollary 3.4.4. *Let L be a C^2 Lagrangian on the manifold M which is non-degenerate, convex in the fibers, and superlinear above every compact subset of M. If $(x,v) \in TM$, denote by $]\alpha_{(x,v)}, \beta_{(x,v)}[$ is the maximal interval on which $t \mapsto \phi_t^L(x,v)$ is defined. If $\beta_{(x,v)} < +\infty$ (resp. $\alpha_{(x,v)} > -\infty$), then $t \to \pi\phi_t^L(x,v)$ leaves every compact subset of M as $t \to \beta_{(x,v)}$ (resp. $t \to \alpha_{(x,v)}$). In particular, if M is a compact manifold, the Euler-Lagrange vector field X_L is complete, i.e. the flow $\phi_t^L : M \to M$ is defined for each $t \in \mathbb{R}$.*

Proof. We know that the global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ is a diffeomorphism, and it conjugates ϕ_t^L to the Hamiltonian flow ϕ_t^H of the Hamiltonian associated to L. Therefore $H\tilde{\mathcal{L}}(\phi_t(x,v)) = H\tilde{\mathcal{L}}(x,v)$ by conservation of the Hamiltonian. Let us suppose that there exists $\beta_i \to \beta_{(x,v)}$ such that $\pi\phi_{\beta_i}^L(x,v) \to x_\infty$. The set $K = \{x_\infty\} \cup \{\pi\phi_{\beta_i}^L(x,v) \mid i \in N\}$ is then a compact subset of M. The sequence $\phi_{\beta_i}(x,v)$ is contained in the subset $\{(y,w) \mid y \in K, H\tilde{\mathcal{L}}(y,w) \leq H\tilde{\mathcal{L}}(x,v)\}$ of TM. But this last set is compact by lemma 3.4.3, and the fact that $\tilde{\mathcal{L}}$ is a diffeomorphism. Therefore, extracting a subsequence if necessary, we can suppose that $\phi_{\beta_i}^L(x,v) \to (x_\infty, v_\infty)$. By the theory of differential equations, the solution $\phi_t^L(x,v)$ can be extended beyond $\beta_{(x,v)}$, which contradicts the maximality of the interval $]\alpha_{(x,v)}, \beta_{(x,v)}[$. ☐

Definition 3.4.5 (Lagrangian Gradient)**.** Let $L : TM \to \mathbb{R}$ be a C^r Tonelli Lagrangian, with $r \geq 2$, on the manifold M. Suppose $\varphi : U \to \mathbb{R}$ be a $\mathrm{C}^k, k \geq 1$ function, defined on the open subset U of M, we define the Lagrangian gradient of φ as the vector field $\operatorname{grad}_L \varphi$ on U given by

$$\forall x \in U, \frac{\partial L}{\partial v}(x, \operatorname{grad}_L \varphi(x)) = d_x\varphi.$$

Note that is well-defined, because the global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ is a C^{r-1} diffeomorphism, and $(x, \operatorname{grad}_L \varphi(x)) = \tilde{\mathcal{L}}^{-1}(x, d_x u)$. It follows that $\operatorname{grad}_L \varphi$ is of class $\mathrm{C}^{\min(r,k)-1}$.

More generally, for a function $S :]a,b[\times U \to \mathbb{R}$ of class $\mathrm{C}^k, k \geq 1$, where $]a,b[$ is an open interval in $\mathbb{R}$, we define its Lagrangian gradient as the vector field $\operatorname{grad}_L S_t$, where we $S_t(x) = S(t,x)$. It is a vector field $\operatorname{grad}_L S_t$ defined on U and dependent on time $t \in]a,b[$. As a function defined on $]a,b[\times U$, it is of class C^{k-1}.

3.5 Hamilton-Jacobi and Minimizers

Theorem 3.5.1 (Lagrangian Gradient and Hamilton-Jacobi)**.** *Let $L : TM \to \mathbb{R}$ be a C^r Tonelli Lagrangian. If $S :]a,b[\times U \to \mathbb{R}$ is a C^1 solution of the Hamilton-Jacobi equation*

$$\frac{\partial S}{\partial t} + H(x, \frac{\partial S}{\partial x}) = 0,$$

then, for every absolutely continuous curve $\gamma : [\alpha, \beta] \to U$*, with* $[\alpha, \beta] \subset]a, b[$*, we have*

$$\mathbb{L}(\gamma) = \int_\alpha^\beta L(\gamma(s), \dot\gamma(s))\, ds \geq S(\beta, \gamma(\beta)) - S(\alpha, \gamma(\alpha)).$$

The inequality above is an equality if and only if γ *is a solution of the time dependent vector field* $\operatorname{grad}_L S_t$*.*

The solutions of the vector field $\operatorname{grad}_L S_t$ *are necessarily extremal curves of* L*. If* $\gamma : [\alpha, \beta] \to U$ *is such a solution, then for every absolutely continuous curve* $\gamma_1 : \mathcal{C}^{ab}([\alpha, \beta], U)$*, with* $\gamma_1(\alpha) = \gamma(\alpha), \gamma_1(\beta) = \gamma(\beta)$*, and* $\gamma_1 \neq \gamma$*, we have* $\mathbb{L}(\gamma_1) > \mathbb{L}(\gamma)$*.*

The vector field $\operatorname{grad}_L S_t$ *is uniquely integrable, i.e. if* $\gamma_i : I_i \to U$ *are 2 solutions of* $\operatorname{grad}_L S_t$ *and* $\gamma_1(t_0) = \gamma_2(t_0)$*, for some* $t_0 \in I_1 \cap I_2$*, then* $\gamma_1 = \gamma_2$ *on* $I_1 \cap I_2$*.*

Proof. Let $\gamma : [\alpha, \beta] \to U$ be an absolutely continuous curve. Since S is C^1, by lemma 3.1.3 the map $[\alpha, \beta] \to \mathbb{R}, t \mapsto S(t, \gamma(t))$ is absolutely continuous and thus by theorem 3.1.2

$$S(\beta, \gamma(\beta)) - S(\alpha, \gamma(\alpha)) = \int_\alpha^\beta \left\{ \frac{\partial S}{\partial t}(t, \gamma(t)) + \frac{\partial S}{\partial x}(t, \gamma(t))[\dot\gamma(t)] \right\} dt.$$

By the Fenchel inequality, for each t where $\dot\gamma(t)$ exists, we have

$$\frac{\partial S}{\partial x}(t, \gamma(t))[\dot\gamma(t)] \leq H\big(\gamma(t), \frac{\partial S}{\partial x}(t, \gamma(t))\big) + L(\gamma(t), \dot\gamma(t)). \qquad (*)$$

Since S satisfies the Hamilton-Jacobi equation, and $\dot\gamma(t)$ exists for almost all $t \in [a, b]$, adding $\partial S/\partial t(t, \gamma(t))$ to both sides in $(*)$, we find that for almost all $t \in [\alpha, \beta]$

$$\frac{\partial S}{\partial t}(t, \gamma(t)) + \frac{\partial S}{\partial x}(t, \gamma(t))[\dot\gamma(t)] \leq L(\gamma(t), \dot\gamma(t))$$

By integration, we then conclude that

$$S(\beta, \gamma(\beta)) - S(\alpha, \gamma(\alpha)) \leq \int_\alpha^\beta L(\gamma(t), \dot\gamma(t))\, dt. \qquad (**)$$

We have equality in this last inequality if and only if we have equality almost everywhere in the Fenchel inequality $(*)$, therefore if and only if $\dot\gamma(t) = \operatorname{grad}_L S_t(\gamma(t))$.

But the right-hand side of this last equality is continuous and defined for each $t \in [\alpha, \beta]$. It follows that $\dot{\gamma}(t)$ can be extended by continuity to the whole interval $[\alpha, \beta]$, hence that γ is C^1 and that it is a solution of the vector field $\operatorname{grad}_L S_t$. To sum up we have shown that (**) is an equality if and only if γ is a (C^1) solution of the time dependent vector field $\operatorname{grad}_L S_t$.

If a curve $\gamma : [\alpha, \beta] \to U$ satisfies the equality

$$S(\alpha, \gamma(\alpha)) - S(\beta, \gamma(\beta)) = \mathbb{L}(\gamma),$$

we have already shown that γ is C^1. Moreover, for every curve $\gamma_1 : [\alpha, \beta] \to U$ with $\gamma_1(\alpha) = \gamma(\alpha)$ and $\gamma_1(\beta) = \gamma(\beta)$, we have

$$\mathbb{L}(\gamma_1) \geq S(\alpha, \gamma(\alpha)) - S(\beta, \gamma(\beta)) = \mathbb{L}(\gamma).$$

It follows that γ is an extremal curve (and thus of class C^r).

Suppose now that for γ_1, and γ as above we have $\mathbb{L}(\gamma_1) = \mathbb{L}(\gamma)$. Therefore $\mathbb{L}(\gamma_1) = S(\alpha, \gamma(\alpha)) - S(\beta, \gamma(\beta))$. By what we already know, the curve γ_1 is also an extremal curve and $\dot{\gamma}_1(\alpha) = \operatorname{grad}_L S_\alpha(\gamma(\alpha))$. However, we have $\dot{\gamma}(\alpha) = \operatorname{grad}_L S_\alpha(\gamma(\alpha))$. The two extremal curves γ and γ_1 go through the same point with the same speed at $t = \alpha$. Since the Euler-Lagrange vector field X_L is uniquely integrable, these two extremal curves are thus equal on their common interval of definition $[\alpha, \beta]$.

The last argument also shows the unique integrability of the vector field $\operatorname{grad}_L S_t$. $\square$

In fact, it is possible to show that, under the assumptions made above on L, a C^1 solution of the Hamilton-Jacobi equation has a derivative which is Lipschitzian, see [Lio82, Theorem 15.1, page 255] or [Fat03, Theorem 3.1]. The proof uses the Lagrangian gradient of the solution. Consequently, note that we are a posteriori in the situation of uniqueness of solutions given by the Cauchy-Lipschitz theorem, since $\operatorname{grad}_L S_t$ is Lipschitzian.

3.6 Small Extremal Curves Are Minimizers

In this section, we will suppose that our manifold M is provided with a Riemannian metric g. We denote by d the distance on M associated to g. If $x \in M$, the norm $\|\cdot\|_x$ on T_xM is the one induced by g. The projection $TM \to M$ is denoted by π. We suppose that $L : TM \to M$ is a C^2 bounded below, such that, for each $(x,v) \in TM$, the second vertical derivative $\frac{\partial^2 L}{\partial v^2}(x,v)$ is positive definite as a quadratic form, and that L is superlinear in each fiber of the tangent bundle $\pi : TM \to M$.

Theorem 3.6.1. *Suppose that L is a Tonelli Lagrangian on the manifold M, and that $\inf_{(x,v)\in TM} L(x,v)$ is finite. Then for each compact subset $K \subset TM$ there exists a constant $\delta_0 > 0$ such that*

- *For $(x,v) \in K$, the local flow $\phi_t(x,v)$ is defined for $|t| \leq \delta_0$.*
- *For each $(x,v) \in K$ and for each $\delta \in]0,\delta_0]$, the extremal curve $\gamma_{(x,v,\delta)} : [0,\delta] \to M, t \mapsto \pi\phi_t(x,v)$ is such that for any absolutely continuous curve $\gamma_1 : [0,\delta] \to M$, with $\gamma_1(0) = x, \gamma_1(\delta) = \pi\phi_\delta(x,v)$, and $\gamma_1 \neq \gamma$ we have $\mathbb{L}(\gamma_1) > \mathbb{L}(\gamma_{(x,v,\delta)})$.*

Proof. By the compactness of K, we can find a $\delta_1 > 0$, such that $\phi_t(x,v)$ is defined for $(x,v) \in K$ and $|t| \leq \delta_1$.

Since $\bigcup_{t\in[0,\delta_1]} \phi_t(K)$ is compact, we can find a constant C_0, which is an upper bound for L on the set $\bigcup_{t\in[0,\delta_1]} \phi_t(K)$. With the notations of the statement, we see that

$$\forall (x,v) \in K, \forall \delta \in [0,\delta_1], \mathbb{L}(\gamma_{(x,v,\delta)}) \leq C_0\delta.$$

In the sequel of the proof, we consider A_0 a compact neighborhood of $\pi\big(\bigcup_{t\in[0,\delta_1]} \phi_t(K)\big)$ in M. Since L is superlinear above compact subsets of M, we can find $C_1 > -\infty$ such that

$$\forall x \in A_0, \forall v \in T_xM, L(x,v) \geq \|v\|_x + C_1. \tag{*}$$

Therefore $d(\gamma_{(x,v,\delta)}(\delta), x) \leq \mathbb{L}(\gamma_{(x,v,\delta)}) - C_1\delta \leq (C_0 - C_1)\delta$, and consequently, applying this for each $\delta' \leq \delta$, we find that the diameter $\mathrm{diam}\big(\gamma_{(x,v,\delta)}([0,\delta])\big)$ of the image of $\gamma_{(x,v,\delta)}$ satisfies

$$\forall (x,v) \in K, \forall \delta \in [0,\delta], \mathrm{diam}\big(\gamma_{(x,v,\delta)}([0,\delta])\big) \leq 2(C_0 - C_1)\delta.$$

By the corollary 2.8.12 and the compactness of K, we can find $\delta_2 \in]0,\delta_1]$ and $\epsilon > 0$ such that for each $(x,v) \in K$

- we have $\bar{B}_d(x,\epsilon) \subset A_0$;

– there exists a function $S^{(x,v)} :]-\delta_2, \delta_2[\times \mathring{B}_d(x,\epsilon) \to \mathbb{R}$, which is a solution of the Hamilton-Jacobi equation

$$\frac{\partial S^{(x,v)}}{\partial t} + H(x, \frac{\partial S^{(x,v)}}{\partial x}) = 0,$$

with $\frac{\partial S^{(x,v)}}{\partial x}(x,0) = \frac{\partial L}{\partial v}(x,v)$.

In particular, we have $v = \operatorname{grad}_L S_0^{(x,v)}(x)$, where we set $S_t^{(x,v)}(x) = S^{(x,v)}(t,x)$. Consequently, the solution of the vector field $\operatorname{grad}_L S_t$, going through x at time $t = 0$, is $t \mapsto \pi\phi_t(x,v)$, see theorem 3.5.1. Since for $\delta_3 \leq \delta_2$, such that $(C_0 - C_1)\delta_3 < \epsilon$, the curve $\gamma_{(x,v,\delta)}$ takes its values in $\mathring{B}_d(x,\epsilon)$ for each $(x,v) \in K$ and all $\delta \in]0,\delta_3]$, by the same Theorem 3.5.1 we obtain that for every $\gamma_1 : [0,\delta] \to \mathring{B}_d(x,\epsilon)$ which is absolutely continuous and satisfies $\gamma_1(0) = x$ and $\gamma_1(\delta) = \gamma_{(x,v,\delta)}(\delta)$, we have $\mathbb{L}(\gamma_1) > \mathbb{L}(\gamma_{(x,v,\delta)})$.

It remains to check that a curve $\gamma_1 : [0,\delta] \to M$ with $\gamma_1(0) = x$ and $\gamma_1([0,\delta]) \not\subset \mathring{B}_d(x,\epsilon)$ has a much too large action. This is where we use that

$$C_2 = \inf\{L(x,v) \mid (x,v) \in TM\} > -\infty.$$

In fact, if $\gamma_1([0,\delta]) \not\subset \mathring{B}_d(x,\epsilon)$, there exists $\eta > 0$ such that $\gamma_1([0,\eta[) \subset \mathring{B}_d(x,\epsilon)$ and $d(\gamma_1(0),\gamma_1(\eta)) = \epsilon$. Since $\bar{B}_d(x,\epsilon) \subset A_0$, we then obtain from inequality $(*)$ above

$$\mathbb{L}(\gamma_1|[0,\eta]) \geq \epsilon + C_1\eta.$$

We can use C_2 as a lower bound of $L(\gamma_1(s),\dot\gamma_1(s))$ on $[\eta,\delta]$ to obtain

$$\mathbb{L}(\gamma_1) \geq \epsilon + C_1\eta + C_2(\delta - \eta),$$

from which, setting $C_3 = \min(C_1, C_2) > -\infty$, it follows that

$$\mathbb{L}(\gamma_1) \geq \epsilon + C_3\delta,$$

as soon as $\gamma_1(0) = x$ and $\gamma_1([0,\delta]) \not\subset \mathring{B}_d(\eta,\epsilon)$. Since $\mathbb{L}(\gamma_{(x,v,\delta)}) \leq C_0\delta$, to finish the proof of the theorem, we see that it is enough to choose δ_0 with $\delta_0 \leq \delta_3$, and $(C_0 - C_3)\delta_0 < \epsilon$. □

Theorem 3.6.2. *Suppose that L is a Tonelli Lagrangian on the manifold M, and that $\inf_{(x,v)\in TM} L(x,v)$ is finite. Fix d a distance on M coming from a Riemannian metric. If $K \subset M$ is compact, and C is a strictly positive constant, then there exists a constant $\delta_0 > 0$ such that, if $x \in K, y \in M$, and $\delta \in]0, \delta_0]$, satisfy $d(x,y) \leq C\delta$, then there exists an extremal curve $\gamma_{(x,y,\delta)} : [0,\delta] \to M$ with $\gamma_{(x,y,\delta)}(0) = x, \gamma_{(x,y,\delta)}(\delta) = y$, and for every curve absolutely continuous $\gamma : [0,\delta] \to M$ which satisfies $\gamma(0) = x, \gamma(\delta) = y$, and $\gamma \neq \gamma_{(x,y,\delta)}$, we have $\mathbb{L}(\gamma) > \mathbb{L}(\gamma_{(x,y,\delta)})$.*

Proof. By the theorem of existence of local extremal curves 2.7.4, we know that there exists a constant $\delta_1 > 0$ such that, if $d(x,y) \leq C\delta$ with $\delta \in]0, \delta_1]$, then there exists an extremal curve $\gamma_{(x,y,\delta)}$ with $\gamma_{(x,y,\delta)}(0) = x, \gamma_{(x,y,\delta)} = y$, and $\|\dot\gamma_{(x,y,\delta)}(0)\|_x \leq 2C$. However the set $\{(x,v) \in TM \mid x \in K, \|v\|_x \leq 2C\}$ is compact in TM and we can apply the previous theorem 3.6.1 to find $\delta_0 \leq \delta_1$ satisfying the conclusions of the theorem we are proving. □

3.7 Regularity of Minimizers

In this section we will assume that the Lagrangian $L : TM \to \mathbb{R}$ is C^r , with $\frac{\partial^2 L}{\partial v^2}(x,v)$ definite positive as a quadratic form, for each $(x,v) \in TM$, and L superlinear in the fibers of the tangent bundle TM. We still provide M with a Riemannian metric of reference.

Theorem 3.7.1 (Regularity)**.** *Suppose that L is a Tonelli Lagrangian on the manifold M. Let $\gamma : [a,b] \to M$ be an absolutely continuous curve such that $\mathbb{L}(\gamma) \leq \mathbb{L}(\gamma_1)$, for each other absolutely continuous curve $\gamma_1 : [a,b] \to M$, with $\gamma_1(a) = \gamma(a), \gamma_1(b) = \gamma(b)$, then the curve γ is an extremal curve, and is therefore C^r.*

Proof. By an argument already used many times previously, it suffices to consider the case $M = U$ is an open subset of of $\mathbb{R}^k$. Let $\bar W$ be a compact subset of U, containing $\gamma([a,b])$ in its interior W. By the compactness of $\bar W$ and the uniform superlinearity of L above compact subsets of M, we have

$$\inf\{L(x,v) \mid x \in \bar W, v \in \mathbb{R}^k\} > -\infty.$$

We can then apply theorem 3.6.2 to the compact subset $\gamma([a,b])$ contained in the manifold $W \subset \mathbb{R}^k$ (which then plays the role of the M of theorem 3.6.2). Let us first show that if the derivative $\dot\gamma(t_0)$ exists for some $t_0 \in [a,b]$, then γ coincides with an extremal curve in the neighborhood of t_0. We have

$$\lim_{t\to t_0} \|\frac{\gamma(t)-\gamma(t_0)}{t-t_0}\| = \|\dot\gamma(t_0)\|.$$

Choosing $C > \|\dot\gamma(t_0)\|$, we can then find $\eta > 0$ such that

$$0 < |t-t_0| \le \eta \Rightarrow \|\gamma(t)-\gamma(t_0)\| < C|t-t_0|. \qquad (*)$$

Let us apply theorem 3.6.2 with $\gamma([a,b])$ as the compact subset and C as the constant, to find the $\delta_0 > 0$ given by this theorem. We can assume that $\delta_0 \le \eta$. We will suppose $t_0 \in]a,b[$, and let the reader make the trivial changes in the cases $t_0 = a$ or $t_0 = b$. From $(*)$ we get

$$\|\gamma(t_0+\delta_0/2) - \gamma(t_0-\delta_0/2)\| < C\delta_0.$$

By theorem 3.6.2, the curve $\gamma_1 : [0,\delta_0] \to M$ which minimizes the action among the curves connecting $\gamma(t_0-\delta_0/2)$ to $\gamma(t_0+\delta_0/2)$ is an extremal curve. However, the curve $[0,\delta_0] \to M, s \mapsto \gamma(s+t_0-\delta_0/2)$ minimizes the action among the curves connecting $\gamma(t_0-\delta_0/2)$ with $\gamma(t_0+\delta_0/2)$, since the curve $\gamma : [a,b] \to M$ minimizes the action for the curves connecting $\gamma(a)$ to $\gamma(b)$. We conclude that the restriction of γ to the interval $[t_0-\delta_0/2, t_0+\delta_0/2]$ is an extremal curve.

Let then $O \subset [a,b]$ be the open subset formed by the points t_0 such that γ coincides with an extremal curve in the neighborhood of t_0. For every connected component I of O, the restriction $\gamma|I$ is an extremal because it coincides locally with an extremal, and the Euler-Lagrange flow is uniquely integrable. Notice also that $\gamma(I) \subset \gamma([a,b])$ which is compact therefore by corollary 3.4.4 the extremal curve $\gamma|I$ can be extended to the compact closure $\bar I \subset [a,b]$, therefore by continuity $\gamma|\bar I$ is an extremal. If $O \neq [a,b]$, let us consider a connected component I of O, then $\bar I \setminus I$ is non-empty and not contained in O. This component can be of one of the following types $]\alpha,\beta[, [a,\beta[,]\alpha,b]$. We will treat the case $I =]\alpha,\beta[\subset O$. We have $\alpha,\beta \notin O$. Since $[\alpha,\beta]$ is compact and $\gamma|[\alpha,\beta]$ is an extremal curve, its speed is bounded. Therefore there exists a finite constant C, such that

$$\forall s \in]\alpha, \beta[,\ \|\dot\gamma(s)\| < C.$$

Once again we apply theorem 3.6.2 to the compact subset $\gamma([a,b])$ and the constant C, to find $\delta_0 > 0$ given by that theorem. We have

$$\begin{aligned}\|\gamma(\beta) - \gamma(\beta - \delta_0/2)\| &\leq \int_{\beta-\delta_0/2}^{\beta} \|\dot\gamma(s)\|\, ds \\ &< C\frac{\delta_0}{2}.\end{aligned}$$

By continuity, for $t > \beta$ and close to β, we do also have

$$\|\gamma(t) - \gamma(\beta - \delta_0/2)\| < C[t - (\beta - \delta_0/2)].$$

We can take $t > \beta$ close enough to β so that $t - \beta < \delta_0/2$. By theorem 3.6.2, the curve γ coincides with an extremal curve on the interval $[\beta - \delta_0/2, t]$. This interval contains β in its interior, and thus $\beta \in O$, which is absurd. Therefore we necessarily have $O = [a,b]$, and γ is an extremal curve. □

Let us summarize the results obtained for M compact.

Theorem 3.7.2. *Let $L : TM \to \mathbb{R}$ be a C^r Tonelli Lagrangian, with $r \geq 2$, where M is a compact manifold. We have:*

- *the Euler-Lagrange flow is well-defined complete and C^{r-1};*
- *the extremal curves are all of class C^r;*
- *for each $x, y \in M$, each and $a, b \in \mathbb{R}$, with $a < b$, there exists an extremal curve $\gamma : [a,b] \to M$ with $\gamma : [a,b] \to M$ with $\gamma(a) = x, \gamma(b) = y$ and such that for every absolutely continuous curves $\gamma_1 : [a,b] \to M$, with $\gamma_1(a) = x, \gamma_1(b) = y$, and $\gamma_1 \neq \gamma$ we have $\mathbb{L}(\gamma_1) > \mathbb{L}(\gamma)$;*
- *if $\gamma : [a,b] \to M$ is an absolutely continuous curve which is a minimizer for the class $\mathcal{C}^{ac}([a,b], M$, then it is an extremal curve. In particular, it is of class C^r;*
- *if $C \in \mathbb{R}$, the set $\Sigma_C = \{\gamma \in \mathcal{C}^{ac}([a,b], M) \mid \mathbb{L}(\gamma) \leq C\}$ is compact for the topology of uniform convergence.*

Chapter 4

The Weak KAM Theorem

In this chapter, as usual we denote by M a *compact and connected* manifold. The projection of TM on M is denoted by $\pi : TM \to M$. We suppose given a C^r Lagrangian $L : TM \to \mathbb{R}$, with $r \geq 2$, such that, for each $(x, v) \in TM$, the second partial vertical derivative $\frac{\partial^2 L}{\partial v^2}(x, v)$ is definite > 0 as a quadratic form, and that L is superlinear in each fiber of the tangent bundle $\pi : TM \to M$. We will also suppose that M is provided with a fixed Riemannian metric. We denote by d the distance on M associated with this Riemannian metric. If $x \in M$, the norm $\|\cdot\|_x$ on T_xM is the one induced by the Riemannian metric.

4.1 The Hamilton-Jacobi Equation Revisited

In this section we will assume that we have a Tonelli Lagrangian L of class $C^r, r \geq 2$, on the manifold M. The global Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ is a C^{r-1} diffeomorphism, see Theorem 3.4.2. Its associated Hamiltonian $H : T^*M \to R$ given by

$$H \circ \tilde{\mathcal{L}}(x, v) = \frac{\partial L}{\partial v}(x, v)(v) - L(x, v),$$

is C^r, and satisfies the Fenchel inequality

$$p(v) \leq L(x, v) + H(x, p),$$

with equality if and only if $p = \partial L/\partial v(x,v)$, or equivalently $(x,p) = \tilde{\mathcal{L}}(x,v)$. The Hamiltonian flow ϕ_t^H of H is conjugated by $\tilde{\mathcal{L}}$ to the Euler-Lagrange flow ϕ_t^L of L.

Theorem 4.1.1 (Hamilton-Jacobi)**.** *Suppose that L is a C^r Tonelli Lagrangian, with $r \geq 2$ on the manifold M. Call $H : T^*M \to \mathbb{R}$ the Hamiltonian associated to the Lagrangian L.*

Let $u : M \to \mathbb{R}$ be a C^1 function. If for some constant $c \in \mathbb{R}$ it satisfies the Hamilton-Jacobi equation

$$\forall x \in M, H(x, d_x u) = c$$

then the graph of du, defined by $\mathrm{Graph}(du) = \{x, d_x u) \mid x \in M\}$, *is invariant under the Hamiltonian flow ϕ_t^H of H.*

Moreover, for each $x \in M$, the projection $t \mapsto \pi^* \phi_t^H(x, d_u)$ is minimizing for the class of absolutely continuous curves.

Of course, if u is of class C^2 the first part theorem follows from 2.5.10. The second part can be deduced from Theorem 3.5.1. In fact, as we will see below, the main argument in proof of Theorem 4.1.1 is just a mere repetition of the main argument in the proof of Theorem 3.5.1. Although this proof of Theorem 4.1.1 could be rather short, we will cut it down in several pieces, because on doing so we will be able to find a notion of C^0 solution of the Hamilton-Jacobi equation. With this notion we will prove, in contrast to the C^1 case, that such a C^0 solution does always exist, see 4.7.1 below, and we will explain its dynamical significance.

We start by studying the meaning for C^1 functions of the Hamilton-Jacobi inequality.

Proposition 4.1.2. *Suppose that L is a Tonelli Lagrangian on the manifold M. Call $H : T^*M \to \mathbb{R}$ the Hamiltonian associated to the Lagrangian L.*

Let $c \in \mathbb{R}$ be a constant, and let $u : U \to \mathbb{R}$ be a C^1 function defined on the open subset $U \subset M$. If u satisfies the inequality

$$\forall x \in V, H(x, d_x u) \leq c \qquad (*)$$

then for every absolutely continuous curve $\gamma : [a,b] \to U$, with $a < b$, we have

$$u(\gamma(b)) - u(\gamma(a)) \leq \int_a^b L(\gamma(s), \dot\gamma(s))\, ds + c(b-a). \qquad (**)$$

Conversely, if inequality $(**)$ *holds for every* C^∞ *curve* $\gamma : [a,b] \to M$ *then* $(*)$ *holds.*

Proof. If $\gamma : [a,b] \to U$ is absolutely continuous then by Lemma 3.1.3 the function $u \circ \gamma$ is also absolutely continuous, and therefore

$$u(\gamma(b)) - u(\gamma(a)) \le \int_a^b d_{\gamma(s)}u(\dot\gamma(s)), ds. \qquad (***)$$

By Fenchel's inequality, at each point $s \in [a,b]$ where $\dot\gamma(s)$ exists we can write

$$d_{\gamma(s)}u(\dot\gamma(s)) \le L(\gamma(s),\dot\gamma(s)) + H(\gamma(s), d_{\gamma(s)}u).$$

Suppose that $(*)$ holds, we get

$$d_{\gamma(s)}u(\dot\gamma(s)) \le L(\gamma(s),\dot\gamma(s)) + c.$$

Integrating, and comparing with $(***)$, yields $(**)$.

Conversely, suppose that $(**)$ for every C^∞ curve $\gamma : [a,b] \to U$. Fix $x \in U$. For a given $v \in T_xM$ we can find a C^∞ curve $\gamma : [-\epsilon,\epsilon] \to U$, with $\epsilon > 0, \gamma(0) = x$, and $\dot\gamma(0) = v$. Then writing condition $(**)$ for every restriction $\gamma|[0,t], t \in]0,\epsilon]$, we obtain

$$u(\gamma(t)) - u(\gamma(0)) \le \int_0^t d_{\gamma(s)}u(\dot\gamma(s)), ds + c(t-0).$$

Dividing both sides by $t > 0$, and letting $t \to 0$ yields

$$d_xu(v) \le L(x,v) + c.$$

Since this is true for every $v \in T_xM$, we conclude that $H(x,d_xu) = \sup_{v\in T_xM} d_xu(v) - L(x,v) \le c$. □

This suggests the following definition.

Definition 4.1.3 (Dominated Function)**.** Let $u : U \to \mathbb{R}$ be a function defined on the open subset $U \subset M$. If $c \in \mathbb{R}$, we say that u is dominated by $L + c$ on U, which we denote by $u \prec L + c$, if for each continuous piecewise C^1 curve $\gamma : [a,b] \to U$ we have

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

If $U = M$, we will simply say that u is dominated by $L + c$.

We will denote by $\mathbb{D}^c(U)$ the set of functions $u : U \to \mathbb{R}$ dominated by $L + c$

In the definition above, we have used continuous piecewise C^1 curves in (D) instead of C^∞ curves because if condition (D) holds for C^∞ curves it also holds for continuous piecewise C^1 curves, and even for absolutely continuous curves, see the following exercise.

Exercise 4.1.4. *Suppose L is a Tonelli Lagrangian on the manifold M. Let $u : U \to \mathbb{R}$ be a function defined on the open subset U of M such that the inequality* (D) *of Definition 4.1.3 holds for every C^∞ curve γ.*

1) Show that (D) *holds for C^1 curves. [Indication: Use a density argument.]*

2) Show that (D) *holds for a continuous piecewise C^1 curve. [Indication: If $\gamma : [a, b] \to U$, there exists $a_0 < a_1 < \cdots < a_n = b$ such that $\gamma|[a_i, a_{i+1}]$ is C^1.]*

3) Show that (D) *holds for absolutely continuous curves. For this fix such a curve $\gamma : [a, b] \to U$, if $\mathbb{L}(\gamma) = +\infty$, there is nothing to prove. Therefore we can assume that*

$$\int_a^b L(\gamma(s), \dot\gamma(s))\, ds < +\infty.$$

We set

$$\omega(\eta) = \sup\{\int_{t'}^t L(\gamma(s), \dot\gamma(s))\, ds \mid t' \leq t, t - t' \leq \eta\}.$$

a) Show that $\omega(\eta) \to 0$, when $\eta \to 0$.

Fix $K, K' \subset U$ compact neighborhoods of $\gamma([a, b])$ with $K \subset \mathring{K}'$.

b) Show that there exists η_0 such that any absolutely continuous curve $\delta : [c, d] \to \mathring{K}'$, with $a \leq c \leq d \leq b, c - d \leq \eta_0, \delta(c) = \gamma(c), \delta(d) = \gamma(d)$, and $\mathbb{L}(\delta) \leq \omega(\eta_0)$, takes values only in K. [Indication: See the proof of Theorem 3.6.1. Notice that L is bounded below on the subset $\{(x, v) \in TM \mid x \in K'\}$.]

c) Show that for for every $c, d \in [a, b]$, with $c \leq d$ and $d-c \leq \eta_0$, there exists an absolutely continuous curve $\delta : [c, d] \to M$ which satisfies:

- $\delta(c) = \gamma(c), \delta(d) = \gamma(d)$.
- $\delta([c, d]) \subset K$.

– *δ is a minimizer for the class of absolutely continuous curves with values in $\mathring{K}'$.*

d) Conclude.

Of course, the notion of dominated function does not use any differentiability assumption on the function. Therefore we can use it as a notion of subsolution of the Hamilton-Jacobi equation. This notion is equivalent to the notion of viscosity subsolution as we will see in chapters 7 and 8.

The next step necessary to prove Theorem 4.1.1 is the introduction of the Lagrangian gradient. Let us recall, see definition 3.4.5, that for $u : U \to \mathbb{R}$ is a C^1 function, its Lagrangian gradient $\operatorname{grad}_L u$ is the vector field defined on U by

$$(x, d_x u) = \tilde{\mathcal{L}}(x, \operatorname{grad}_L u(x)).$$

It follows that

$$\operatorname{Graph}(du) = \tilde{\mathcal{L}}[\operatorname{Graph}(\operatorname{grad}_L u)],$$

where $\operatorname{Graph}(\operatorname{grad}_L u) = \{(x, \operatorname{grad}_L u(x)) \mid x \in U\}$. Since ϕ_t^H and ϕ_t^L are conjugated by $\tilde{\mathcal{L}}$, invariance of $\operatorname{Graph}(du)$ under ϕ_t^H is equivalent to invariance of $\operatorname{Graph}(\operatorname{grad}_L u)$ under ϕ_t^L. Therefore the following proposition finishes the proof of Theorem 4.1.1

Proposition 4.1.5. *Let L be a Tonelli Lagrangian on the manifold M. If $u : U \to \mathbb{R}$ is a C^1 function which satisfies on U the Hamilton-Jacobi equation*

$$H(x, d_x u) = c,$$

for some fixed $c \in \mathbb{R}$, then every solution $\gamma : [a, b] \to U$ of the vector field $\operatorname{grad}_L u$ satisfies

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

It follows that solutions of $\operatorname{grad}_L u$ are minimizing for the class of absolutely continuous curves with values in U, and that it must be an extremal of class C^2.

Moreover, the graph of $\operatorname{grad}_L u$ is locally invariant by the Euler-Lagrange flow. If $U = M$ and M is compact then the graph $\operatorname{Graph}(\operatorname{grad}_L u)$ is invariant by the Euler-Lagrange flow.

Proof. Since $(y, d_y u) = \tilde{\mathcal{L}}(y, \operatorname{grad}_L u(y))$, we have the following equality in the Fenchel inequality

$$d_y u(\operatorname{grad}_L u(y)) = L(y, \operatorname{grad}_L u(y)) + H(y, d_y u).$$

Taking into account that $H(y, d_y u) = c$, and using the equality along a solution $\gamma : [a, b] \to U$ of the vector field $\operatorname{grad}_L u$, we get

$$d_{\gamma(t)} u(\dot\gamma(t)) = L(\gamma(t), \dot\gamma(t)) + c.$$

If we integrate we get

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

This implies that γ is a minimizer for absolutely continuous curves, and is therefore a C^2 extremal. In fact, if $\delta : [a, b] \to U$ is a curve, by Proposition 4.1.2, we have $u(\delta(b)) - u(\delta(a)) \leq \mathbb{L}(\delta)$. If $\delta(a) = \gamma(a)$ and $\delta(b) = \gamma(a)$, we obtain

$$\begin{aligned}\mathbb{L}(\gamma) + c(b - a) &= u(\gamma(b)) - u(\gamma(a))\\ &= u(\delta(b)) - u(\delta(a))\\ &\leq \mathbb{L}(\delta).\end{aligned}$$

We now show that the $\operatorname{Graph}(\operatorname{grad}_L u)$ is locally invariant by the Euler-Lagrange flow. Given $x \in U$, since the vector field $\operatorname{grad}_L u$ is continuous, we can apply the Cauchy-Peano Theorem, see [Bou76], to find a map $\Gamma : [-\epsilon, \epsilon] \times V$, with $\epsilon > 0$ and V an open neighborhood of x, such that for every $y \in V$ the curve $t \mapsto \Gamma_y(t) = \Gamma(y, t)$ is a solution of $\operatorname{grad}_L u$ with $\Gamma_y(0) = y$. Therefore

$$\dot\Gamma_y(t) = \operatorname{grad}_L u(\Gamma_y(t)).$$

But we know that Γ_y is an extremal, hence its speed curve is $t \mapsto \phi_t^L(y, \dot\Gamma_y(0))$. Therefore, for every $(t, y) \in [-\epsilon, \epsilon] \times V$, we have

$$\phi_t^L(y, \dot\Gamma_y(0)) = (\Gamma_y(t), \operatorname{grad}_L u(\Gamma_y(t))) \in \operatorname{Graph}(\operatorname{grad}_L u). \qquad (*)$$

If $U = M$ is compact we can find a finite family of $[-\epsilon_i, \epsilon_i] \times V_i, i = 1, \ldots, \ell$ satisfying $(*)$, and such that $M = \cup_{i=1}^{\ell} V_i$. Setting $\epsilon = \min_{i=1}^{\ell} \epsilon_i > 0$, we obtain

$$\forall t \in [-\epsilon, \epsilon], \phi_t^L(\operatorname{Graph}(\operatorname{grad}_L u)) \subset \operatorname{Graph}(\operatorname{grad}_L u).$$

Since ϕ_t^L is a flow defined for all $t \in \mathbb{R}$, we obtain

$$\forall t \in \mathbb{R}, \phi_t^L(\mathrm{Graph}(\mathrm{grad}_L u)) \subset \mathrm{Graph}(\mathrm{grad}_L u).$$

Using the inclusion above for t and $-t$ yields $\phi_t^L(\mathrm{Graph}(\mathrm{grad}_L u)) = \mathrm{Graph}(\mathrm{grad}_L u)$, for all $t \in \mathbb{R}$. □

Exercise 4.1.6. *Under the hypothesis of Proposition 4.1.5, if $x \in U$, show that the orbit $\phi_t^L(x, \mathrm{grad}_L u(x))$ is defined at least on an interval $]\alpha, \beta[$ such that $\gamma(t) = \pi\phi_t^L(x, \mathrm{grad}_L u(x)) \in U$ for every $t \in]\alpha, \beta[$, and $t \mapsto \gamma(t)$ leaves every compact subset of U as t tends to either α or β. Show that this curve γ is a solution of* $\mathrm{grad}_L u$ *on $]\alpha, \beta[$. Show that, any other solution $\tilde{\gamma} : I \to U$ of* , $\mathrm{grad}_L u$, *with $\tilde{\gamma}(0) = x$, satisfies $I \subset]\alpha, \beta[$ and $\tilde{\gamma} = \gamma$ on I.*

Proposition 4.1.5 suggests the following definition

Definition 4.1.7 (Calibrated Curve)**.** Let $u : U \to \mathbb{R}$ be a function and let $c \in \mathbb{R}$ be a constant, where U is an open subset of M. We say that the (continuous) piecewise C^1 curve $\gamma : I \to U$, defined on the interval $I \subset \mathbb{R}$ is (u, L, c)-calibrated, if for every $t \leq t' \in I$, with $t \leq t'$, we have

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

Although the following proposition is an immediate consequence of the definition of calibrated curve, it will be used often.

Proposition 4.1.8. *Let L be a Tonelli Lagrangian defined on the manifold M. Suppose $u : U \to \mathbb{R}$ is a C^1 function defined on the open subset $U \subset M$, and $c \in \mathbb{R}$. If the curve $\gamma : I \to U$ is (u, L, c)-calibrated, then for any subinterval $I' \subset I$ the restriction $\gamma|I'$ is also (u, L, c)-calibrated.*

The following theorem explains why calibrated curves are special.

Theorem 4.1.9. *Suppose that L is a Tonelli Lagrangian on the manifold M. Let $u : U \to \mathbb{R}$ be a function defined on the open subset $U \subset M$. Assume that $u \prec L + c$, where $c \in \mathbb{R}$. Then any continuous piecewise C^1 (u, L, c)-calibrated curve $\gamma : I \to U$ is* *necessarily a minimizing curve for the class of continuous piecewise* C^1 *on* U*. Therefore it is an extremal curve and it is as smooth as the Lagrangian* L*.*

Proof. We fix a compact interval $[t,t'] \subset I$, with $t \leq t'$. If $\delta : [t,t'] \to U$ is a (continuous) piecewise C^1, from $u \prec L + c$, it follows that

$$u(\delta(t')) - u(\delta(t)) \leq \mathbb{L}(\delta) + c(t' - t).$$

Moreover, since γ is (u, L, c)-calibrated, we have equality when $\delta = \gamma|[t,t']$. If $\delta(t') = \gamma(t')$ and $\delta(t) = \gamma(t)$, we obtain

$$\mathbb{L}(\gamma|[t,t']) + c(t' - t) = u(\gamma(t')) - u(\gamma(t)) \leq \mathbb{L}(\delta) + c(t' - t),$$

hence $\mathbb{L}(\gamma|[t,t']) \leq \mathbb{L}(\delta)$, and γ is therefore a minimizing curve. This implies that γ is an extremal and is as smooth as L, see Proposition 2.3.7. □

We now can give a characterization of C^1 solutions of the Hamilton-Jacobi equation which does not involve the derivative.

Proposition 4.1.10. *Let* L *be a Tonelli Lagrangian defined on the manifold* M*. If* $u : U \to \mathbb{R}$ *is a* C^1 *function defined on the open subset* $U \subset M$*, and* $c \in \mathbb{R}$*, the following conditions are equivalent:*

(i) *The function* u *satisfies Hamilton-Jacobi equation*

$$\forall x \in U, H(x, d_x u) = c.$$

(ii) *The function* u *is dominated by* $L + c$*, and for every* $x \in U$ *we can find* $\epsilon > 0$ *and a* C^1 *curve* $\gamma : [-\epsilon, \epsilon] \to U$ *which is* (u, L, c)*-calibrated, and satisfies* $\gamma(0) = x$*.*

(iii) *The function* u *is dominated by* $L + c$*, and for every* $x \in U$ *we can find* $\epsilon > 0$ *and a* C^1 *curve* $\gamma : [-\epsilon, 0] \to U$ *which is* (u, L, c)*-calibrated, and satisfies* $\gamma(0) = x$*.*

(iv) *The function* u *is dominated by* $L + c$*, and for every* $x \in U$ *we can find* $\epsilon > 0$ *and a* C^1 *curve* $\gamma : [0, \epsilon] \to U$ *which is* (u, L, c)*-calibrated, and satisfies* $\gamma(0) = x$*.*

Proof. We prove that (i) implies (ii). The fact that $u \prec L + c$ follows from Proposition 4.1.2. By Proposition 4.1.5, we can take for γ any solution of $\operatorname{grad}_L u$ with $\gamma(0) = x$. Again such a solution exists by the Cauchy-Peano Theorem, see [Bou76], since $\operatorname{grad}_L u$ is continuous.

Obviously (ii) implies (iii) and (iv).

It remains to prove that either (iii) or (iv) implies (i). We show that (iii) implies (i), the other implication being similar. From Proposition 4.1.2, we know that

$$\forall x \in U, H(x, d_x u) \leq c.$$

To show the reversed inequality, we pick a (u, L, c)-calibrated C^1 curve $\gamma : [-\epsilon, 0] \to M$ with $\gamma(0) = x$. For every $t \in [0, \epsilon]$, we have

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

If we divide by $t > 0$, after changing signs in the numerator and denominator of the left hand side, we get

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

If we let $t \to 0$ taking into account that $\gamma(0) = x$, and that both u and γ are C^1, we obtain

$$d_x u(\dot\gamma(0)) = L(x, \dot\gamma(0)) + c.$$

But by Fenchel's inequality $H(x, d_x u) \geq d_x u(\dot\gamma(0)) - L(x, \dot\gamma(0))$, therefore $H(x, d_x u)) \geq c$. □

Therefore we could take any one of condition (ii), (iii) or (iv) above as a definition of a continuous solution of the Hamilton-Jacobi equation. In fact, a continuous function satisfying (ii) is necessarily C^1, see **??** below. Condition (iii) and (iv) lead both to the notion of continuous solutions of the Hamilton-Jacobi equation. The two sets of solutions that we obtain are in general different, and they both have a dynamical meaning as we will see later.

We will be mainly using continuous solution of the Hamilton-Jacobi equation which are defined on a compact manifold M. Notice that for C^1 solutions of the Hamilton-Jacobi equation defined on M, by the last part of Proposition 4.1.5, in condition (ii), (iii), and (iv) of Proposition 4.1.10 we can impose $\epsilon = +\infty$. This justifies the following definition (see also **??** below).

Definition 4.1.11. Let L be a Tonelli Lagrangian on the compact manifold M. A weak KAM solution of negative type (resp. of positive type) is a function $u : M \to \mathbb{R}$ for which there exists $c \in \mathbb{R}$ such that

(1) The function u is dominated by $L + c$.

(2) For every $x \in M$ we can find a (u, L, c)-calibrated C^1 curve $\gamma :]-\infty, 0] \to M$ (resp. $\gamma : [0, +\infty[\to M$) with $\gamma(0) = x$.

We denote by $\mathcal{S}_-$ (resp. $\mathcal{S}_+$) the set of weak KAM solutions of negative (resp. positive) type.

We will usually use the notation u_- (resp. u_+) to denote an element of $\mathcal{S}_-$ (resp. $\mathcal{S}_+$).

4.2 Dominated Functions and the Mañé Critical Value

We now establish some properties of dominated function. Before doing that let us recall that the notion of locally Lipschitz function makes perfect sense in a manifold M. In fact, a function $u : M \to \mathbb{R}$ is said to be locally Lipschitz if for every coordinate chart $\varphi : U \to M$, the function $u \circ \varphi$ is locally Lipschitz on the open subset U of some Euclidean space. Since all Riemannian metrics are equivalent above compact subsets, it is equivalent to say that u is locally Lipschitz for one distance (or for all distances) d coming from a Riemannian metric. If $u : X \to Y$ is a Lipschitz map between the metric spaces X, Y we will denote by $\mathrm{Lip}(u)$ its smallest Lipschitz constant

$$\mathrm{Lip}(u) = \sup \frac{d(u(x), u(x'))}{d(x, x')}$$

where the supremum is taken over all $x, x' \in X$, with $x \neq x'$.

Proposition 4.2.1. *Suppose L is a Tonelli Lagrangian on the manifold M. Endow M with a distance d and TM a norm $(x,v)\mapsto \|V\|_x$ both coming from the same Riemannian metric on M (Note that d well defined and finite on each connected component of M).*

Let U be an open subset of M, and $c\in\mathbb{R}$. We have the following properties:

(i) *The set $\mathbb{D}^c(U)$ of functions $u: U\to\mathbb{R}$ dominated by $L+c$ is a closed convex subset of the set of functions $U\to\mathbb{R}$ for the topology of point-wise convergence. Moreover, if $k\in\mathbb{R}$, we have $u\in\mathbb{D}^c(U)$ if and only if $u+k\in\mathbb{D}^c(U)$.*

(ii) *Every function in $\mathbb{D}^c(U)$ is locally Lipschitz. More precisely, for every $x_0\in U$, we can find a compact neighborhood V_{x_0} such that for every $u\in\mathbb{D}^c(U)$ the Lipschitz constant of $u|V_{x_0}$ is $\leq A_{V_{x_0}}+c$, where*

$$A_{V_{x_0}}=\sup\{L(x,v)\mid (x,v)\in TM, x\in V_{x_0}, \|v\|_x=1\}<+\infty.$$

In particular the family of functions in $\mathbb{D}^c(U)$ is locally equi-Lipschitzian.

(iii) *If M is compact and connected, and $u: M\to\mathbb{R}$ is defined on the whole of M and is dominated by $L+c$, then u is Lipschitz. More precisely, then Lipschitz constant $\operatorname{Lip}(u)$ of u is $\leq A+c$, where*

$$A=\sup\{L(x,v)\mid (x,v)\in TM, \|v\|_x=1\}.$$

In particular the family of functions in $\mathbb{D}^c(M)$ is l equi-Lipschitzian.

(iv) *Moreover, if M is compact and connected, then every Lipschitz function $u: M\to\mathbb{R}$ is dominated by $L+c$ for some $c\in\mathbb{R}$. More precisely, given $K\in[0,+\infty[$, we can find c_K such that every $u: M\to\mathbb{R}$, with $\operatorname{Lip}(u)\leq K$, satisfies $u\prec L+c_K$.*

(v) *Suppose that $c,k\in\mathbb{R}$, and that U is a connected open subset of M. If $x_0\in U$ is fixed, then the subset $\{u\in\mathbb{D}^c(U)\mid |u(x_0)|\leq k\}$ is a compact convex subset of $\mathcal{C}^0(U,\mathbb{R})$ for the topology of uniform convergence on compact subsets.*

Proof. Its is obvious from the definition of domination that $\mathbb{D}^c(U)$ is convex and that $u \in \mathbb{D}^c(U)$ if and only if $u + k \in \mathbb{D}^c(U)$. For a fixed continuous piecewise C^1 curve $\gamma : [a, b] \to U$ the set $\mathcal{F}_{\gamma,c}$ of function $u : U \to \mathbb{R}$ such that

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

is clearly closed in the topology of point-wise convergence. Since $\mathbb{D}^c(U)$ is the intersection of the $\mathcal{F}_{\gamma,c}$ for all γ's with values in U, it is also closed.

To prove (ii), let $u \in \mathbb{D}^c(U)$. Fix $x_0 \in U$. We can find a compact neighborhood $V_{x_0} \subset U$ of x_0, such that for every $x, y \in \mathring{V}_{x_0}$ we can find a geodesic $\gamma : [a, b] \to M$ parametrized by unit length such that $\gamma(a) = x, \gamma(b) = y$ and $\operatorname{length}(\gamma) = b - a = d(x, y)$. Since V_{x_0} is compact, the constant

$$A_{V_{x_0}} = \sup\{L(x, v) \mid (x, v) \in TM, x \in V_{x_0}, \|v\|_x = 1\}$$

is finite. Since $\|\dot\gamma(s)\| = 1$, for every $s \in [a, b]$, we have $L(\gamma(s), \dot\gamma(s)) \leq A_{V_{x_0}}$, therefore we obtain

$$\begin{aligned} u(\gamma(b)) - u(\gamma(a)) &\leq \int_a^b L(\gamma(s), \dot\gamma(s))\, ds + c(b - a) \\ &\leq \int_a^b A_{V_{x_0}}\, ds + c(b - a) \\ &= (A_{V_{x_0}} + c) \operatorname{length}(\gamma) \\ &= (A_{V_{x_0}} + c) d(x, y). \end{aligned}$$

To prove (iii), it suffices to observe that, when M is compact and connected, we can take $V_{x_0} = M$ in the argument above.

Suppose now that M is compact and connected. If we fix $K \geq 0$, by the superlinearity of L we can find $A(K) > -\infty$ such that

$$\forall (x, v) \in TM, L(x, v) \geq K\|v\|_x + A(K).$$

therefore if $\gamma : [a, b] \to M$ is a continuous piecewise C^1 curve, applying this inequality for $(x, v) = (\gamma(s), \dot\gamma(s))$ and integrating we get

$$K \operatorname{length}(\gamma) \leq \mathbb{L}(\gamma) - A(K)(b - a),$$

and hence

$$Kd((\gamma(b),\gamma(a)) \leq \mathbb{L}(\gamma) - A(K)(b-a).$$

If $u : M \to \mathbb{R}$ has Lipschitz constant $\leq K$, we therefore obtain

$$u(\gamma(b)) - u(\gamma(b)) \leq \mathbb{L}(\gamma) - A(K)(b-a).$$

This proves (iv) with $c_K = -A(K) < +\infty$.

It remains to prove (v). Set $\mathcal{E} = \{u \in \mathbb{D}^c(U) \mid |u(x_0)| \leq k\}$. By (1) this set is clearly closed for the topology of point-wise convergence. It suffices to show that for each compact subset $K \subset U$, the set of restrictions $\mathcal{E}|K = \{u|K \mid u \in \mathcal{E}\}$ is relatively compact in $\mathcal{C}^0(K,\mathbb{R})$. We apply Ascoli's Theorem, since this the family $\mathcal{E}|K$ is locally equi-Lipschitz by (ii), it suffices to check that

$$\sup\{|u(x)| \mid x \in K, u \in \mathcal{E}\} < +\infty.$$

Since U is connected locally compact and locally connected, enlarging K if necessary, we can assume that K is connected and contains x_0. Again by (ii), we can cover K by a finite number of open sets $V_1, \dots, V_n$, and find finite numbers $k_i \geq 0, i = 1, \dots, n$ such $\operatorname{Lip}(u|V_i) \leq k_i$ for every $u \in \mathbb{D}^c(U)$, and every $i = 1, \dots, n$. If $x \in K$, by connectedness of K, we can find $i_1, \dots, i_\ell \in \{1, \dots, n\}$ such that $V_{i_j} \cap V_{i_{j+1}} \cap K \neq \emptyset, j = 1, \dots, \ell, x_0 \in V_{i_1}, x \in V_{i_\ell}$. By assuming ℓ minimal with this properties, we get that the i_j are all distinct, therefore $\ell \leq n$. We can choose $x_j \in V_{i_j} \cap V_{i_{j+1}} \cap K$, for $i = 1, \dots, \ell - 1$. Therefore setting $x_\ell =$, for $u \in \mathbb{D}^c(U)$, we obtain

$$\begin{aligned}
|u(x) - u(x_0)| &\leq \sum_{j=0}^{\ell-1} |u(x_{j+1}) - u(x_j)| \\
&\leq \sum_{j=0}^{\ell-1} k_{i_j} d(x_{j+1}, x_j) \\
&\leq \ell \operatorname{diam}(K) \max_{i=1}^{n} k_i \\
&\leq n \operatorname{diam}(K) \max_{i=1}^{n} k_i.
\end{aligned}$$

Therefore $|u(x)| \leq k + n \operatorname{diam}(K) \max_{i=1}^n k_i$, for every $x \in K$ and every $u \in \mathcal{E}$. $\square$

Proposition 4.2.2. *Let L be a Tonelli Lagrangian on the manifold M. Suppose that $u : U \to \mathbb{R}$ is defined on the open subset $U \subset M$ and that it is dominated by $L + c$. Then at each point $x \in U$ where the derivative $d_x u$ exists, we have*

$$H(x, d_x u) \leq c.$$

By Rademacher's Theorem 1.1.10, the derivative $d_x u$ exists almost everywhere on U, the above inequality is therefore satisfied almost everywhere.

Proof. Suppose that $d_x u$ exists. We fix $v \in T_x(M)$. Let $\gamma : [0, 1] \to U$ be a C^1 curve such that $\gamma(0) = x$ and $\dot\gamma(0) = v$. Since $u \prec L+c$, we have

$$\forall t \in [0, 1], u(\gamma(t)) - u(\gamma(0)) \leq \int_0^t L(\gamma(s), \dot\gamma(s))\, ds + ct.$$

By dividing this inequality by $t > 0$ and letting t tend to 0, we find $d_x u(v) \leq L(x, v) + c$ and hence

$$H(x, d_x u) = \sup_{v \in T_x M} d_x u(v) - L(x, v) \leq c. \qquad \square$$

We now prove the converse of Proposition 4.2.2.

Proposition 4.2.3. *Let L be a Tonelli Lagrangian on the manifold M. Call H the Hamiltonian associated to L. Suppose that $u : U \to \mathbb{R}$ is a locally Lipschitz function defined on the open subset U of M. By Rademacher's Theorem 1.1.10, the derivative $d_x u$ exists for almost all $x \in U$. If there exists a c such that $H(x, d_x u) \leq c$, for almost all $x \in U$, then $u \prec L + c$.*

Proof. Using a covering of a curve by coordinates charts, it is not difficult to see that we can assume that U is an open convex set in $\mathbb{R}^k$. We call R the set of points $x \in U$ were $d_x u$ exists and $H(x, d_x u) \leq c$. By assumption $U \setminus R$ is negligible for Lebesgue measure.

We show that first that $u(\gamma(b)) - u(\gamma(a)) \leq \mathbb{L}(\gamma) + c(b - a)$, for an affine segment $\gamma : [a, b] \to U$. To treat the case where γ is constant we have to show that $L(x, 0) + c \geq 0$ for every $x \in U$.

If $x \in R$, this is true since $L(x, 0) + c \geq L(x, 0) + H(x, d_x u) \geq d_x u(0) = 0$, where the second inequality is a consequence of part (i) of Fenchel's Theorem 1.3.6. Since R is dense in U, the inequality $L(x, 0) + c \geq 0$ is therefore true on the whole of U. We now assume that the affine segment is not constant. We can then write $\gamma(s) = x + (t - a)v$, with $\|v\| = r > 0$. We call S the set of vectors w such that $\|w\| = r$, and the line $D_w = \{x + tw \mid t \in \mathbb{R}\}$ intersects R in a set of full linear measure in $U \cap D_w$. By Fubini's Theorem, the set S itself is of full Lebesgue measure in the sphere $\{w \in \mathbb{R}^k \mid \|w\| = r\}$. Hence we can find a sequence $v_n \in S$ with $v_n \to v$, when $n \to \infty$. Dropping the first n's, if necessary, we can assume that the affine curve $\gamma_n : [a, b] \to \mathbb{R}^k, t \mapsto x + (t - a)v_n$ is contained in fact in U. By the definition of the set S, for each n, the derivative $d_{\gamma_n(t)} u$ exists and verifies $H(\gamma_n(t), d_{\gamma_n(t)} u) \leq c$ for almost every $t \in [a, b]$. It follows that the derivative of $u \circ \gamma_n$ is equal to $d_{\gamma_n(t)} u(\dot{\gamma}_n(t))$ at almost every $t \in [a, b]$. Using again part (i) of Fenchel's Theorem 1.3.6, we see that

$$\begin{aligned}
\frac{du \circ \gamma_n}{dt}(t) &= d_{\gamma_n(t)} u(\dot{\gamma}_n(t)) \\
&\leq H(\gamma_n(t), d_{\gamma_n(t)} u) + L(\gamma_n(t), \dot{\gamma}_n(t)) \\
&\leq c + L(\gamma_n(t), \dot{\gamma}_n(t)),
\end{aligned}$$

Since $u \circ \gamma_n$ is Lipschitz, we obtain

$$\begin{aligned}
u(\gamma_n(b)) - u(\gamma_n(a)) &= \int_a^b \frac{du \circ \gamma_n}{dt}(t)\, dt \\
&\leq \int_a^b c + L(\gamma_n(t), \dot{\gamma}_n(t))\, dt \\
&= \mathbb{L}(\gamma_n) + c(b - a).
\end{aligned}$$

Since γ_n converges in the C^1 topology to γ we obtain

$$u(\gamma(b)) - u(\gamma(a)) \leq \mathbb{L}(\gamma) + c(b - a).$$

Of course, we now have the same inequality for any continuous piecewise affine segment $\gamma : [a, b] \to U$. To show the inequality $u(\gamma(b)) - u(\gamma(a)) \leq \mathbb{L}(\gamma) + c(b - a)$ for an arbitrary C^1 curve $\gamma : [a, b] \to U$, we introduce piecewise affine approximation γ_n :

$[a,b] \to U$ in the usual way. For each integer $n \geq 1$, the curve γ_n is affine on each of the intervals $[a+i(b-a)/n, a+(i+1)(b-a)/n]$, for $i = 0, \ldots, n-1$, and $\gamma_n(a+i(b-a)/n) = \gamma(a+i(b-a)/n)$, for $i = 0, \ldots, n$. The sequence γ_n converges uniformly on $[a,b]$ to γ. Since $\gamma_n(a) = \gamma(a)$ and $\gamma_n(b) = \gamma(b)$, we obtain from what we just proved

$$u(\gamma(b)) - u(\gamma(a)) \leq \mathbb{L}(\gamma_n) + c(b-a). \tag{$*$}$$

The derivative $\dot\gamma_n(t)$ exists for each n at each t in the complement A of the countable set $\{a+i(b-a)/n \mid n \geq 1, i = 0, \ldots, n\}$. Moreover, using the Mean Value Theorem, the sequence $\dot\gamma_n|A$ converges uniformly to $\dot\gamma|A$, as $n \to \infty$, since γ is C^1. Using the fact the the set $\{(\gamma(t), \dot\gamma(t)) \mid t \in [a,b]\}$ is compact and the continuity of L, it follows that the sequence of maps $A \to \mathbb{R}, t \mapsto L(\gamma_n(t), \dot\gamma_n(t))$ converges uniformly to $A \to \mathbb{R}, t \mapsto L(\gamma(t), \dot\gamma(t))$, therefore $\mathbb{L}(\gamma_n) \to \mathbb{L}(\gamma)$, since $[a,b] \setminus A$ is countable. passing to the limit in the above inequality $(*)$ we indeed obtain

$$u(\gamma(b)) - u(\gamma(a)) \leq \mathbb{L}(\gamma) + c(b-a).\square$$

Definition 4.2.4 (Hamiltonian constant of a function)**.** If $u : U \to \mathbb{R}$ is a locally Lipschitz function, we define $\mathbb{H}_U(u)$ as the essential supremum on U of the almost everywhere defined function $x \mapsto H(x, d_x u)$.

We summarize the last couples of Propositions 4.2.2 and 4.2.3 in the following theorem.

Theorem 4.2.5. *Suppose that L is a Tonelli Lagrangian on the manifold M. Let U be an open subset of M. A function $u : U \to \mathbb{R}$ is dominated by $L + c$ on U, for some $c \in \mathbb{R}$, if and only if it is locally Lipschitz and $c \geq \mathbb{H}_U(u)$.*

Definition 4.2.6 (Mañé's Critical Value)**.** If L is a Tonelli Lagrangian on the connected compact manifold M, the Mañé critical value of L is the constant $c[0]$ (or $c_L[0]$, if we need to precise the Lagrangian) defined by

$$c[0] = \inf\{c \in \mathbb{R} \mid \exists u : M \to \mathbb{R}, u \prec L + c\}.$$

Theorem 4.2.7. *The Mañé critical value $c[0]$ of L is finite. In fact we have $c[0] \geq -\inf_{x\in M} L(x,0)$.*

If $u : M \to \mathbb{R}$ is dominated by $L + c$ for some $c \in \mathbb{R}$, then $c \geq c[0]$. Moreover, there exists a function $u_0 : M \to \mathbb{R}$ such that $u_0 \prec L + c$.

Proof. For a given x, the constant curve $\gamma_x : [0,1] \mapsto x$ has action equal to $L(x,0)$. Therefore, if $u \prec L+c$,we have $0 = u(x)-u(x) \leq L(x,0)+c\cdot 1$ hence $c \leq -L(x,0)$. Since this is true for every $x \in M$, we obtain $c \geq -\inf_{x\in M} L(x,0)$, for every c for which we can find $u : M \to \mathbb{R}$ with $u \prec L + c$.taking the infimum over all such c yields $c[0] \geq -\inf_{x\in M} L(x,0)$, which is of course finite by the compactness of M.

By definition of $c[0]$, if c is such that there exists $u : M \to \mathbb{R}$ with $u \prec L + c$ we have $c \geq c[0]$.

It remains to find $u : M \to \mathbb{R}$ such that $u \prec L + c[0]$. By definition of $c[0]$ we can find a sequence $c_n \to c[0]$ of numbers, and a sequence of functions $u_n : M \to \mathbb{R}$, with $u_n \prec L + c_n$, for each n. We now fix $x_0 \in M$. By (i) of Proposition 4.2.1, the function $u_n - u_n(x_0)$ is also dominated by $L + c_n$. therefore we can assume $u_n(x_0) = 0$, for every n. We could apply part (v) of Proposition 4.2.1, to finish the proof by having a subsequence of u_n converge. In fact, it is easier to argue directly. We define $u : M \to [-\infty, +\infty]$ by

$$\forall x \in M, u(x) = \liminf_{n\to\infty} u_n(x).$$

Since $u_n(x_0) = 0$, we have $u(x_0) = 0$. Given a continuous piecewise C^1 curve $\gamma : [a,b] \to M$, since $u_n \prec L + c_n$, we have

$$u_n(\gamma(b)) \leq u_n(\gamma(a)) + \int_a^b L(\gamma(s), \dot\gamma(s))\, ds + c_n(b-a).$$

Since $c_n \to c[0]$, by taking the lim inf in the equality above, we obtain

$$u(\gamma(b)) \leq u(\gamma(a)) + \int_a^b L(\gamma(s), \dot\gamma(s))\, ds + c[0](b-a). \qquad ((*))$$

Since $u(x_0) = 0$, using a continuous piecewise C^1 curve starting at a given point $x \in M$ and ending at x_0, we obtain from (*) that $u(x) > -\infty$. Using a curve starting at x_0 and ending at x

we conclude $u(x) < +\infty$. Finally, the function u is finite valued everywhere. since (*) is true for an arbitrary continuous piecewise C^1 curve $\gamma : [a, b] \to M$. We obtain $u \prec L + c[0]$. □

4.3 Defect and Calibration of Curves

We now study the properties of calibrated curves. It is convenient to introduce the notion of defect of a curve on an interval with respect to a function $u \prec L + c$.

Definition 4.3.1 (Defect of a Curve)**.** Let U be an open subset of M, and $c \in \mathbb{R}$. If $u : U \to \mathbb{R}$ is dominated by $L + c$ and $\gamma : I \to U$ is a continuous piecewise C^1 curve, for $[\alpha, \beta] \subset I$, with $\alpha \leq \beta$, we define the defect $\mathcal{D}(\gamma, u, c; \alpha, \beta)$ of the curve γ on the interval $[\alpha, \beta]$ for the $(L + c)$-dominated function u by

$$\mathcal{D}(\gamma, u, c; \alpha, \beta) = \int_\alpha^\beta L(\gamma(s), \dot\gamma(s))\, ds - c(\beta - \alpha) - (u(\gamma(\beta) - u(\gamma(\alpha)).$$

‘ Of course, in the definition of defect, there is no need to assume $u \prec L + c$. However, this definition is useful only when $u \prec L + c$, as we will presently see.

Proposition 4.3.2. *Let L be a Tonelli Lagarangian on the manifold M. Suppose that $u : U \to \mathbb{R}$ is a continuous function defined on the open subset $U \subset M$, with $u \prec L + c$, and that $\gamma : I \to U$ is a continuous piecewise C^1 curve. We have the following properties:*

(1) *If $[\alpha, \beta] \subset I$, with $\alpha \leq \beta$, then*

$$\mathcal{D}(\gamma, u, c; \alpha, \beta) \geq 0.$$

(2) *For every $k \in \mathbb{R}$, we have $\mathcal{D}(\gamma, u, c; \alpha, \beta) = \mathcal{D}(\gamma, u + k, c; \alpha, \beta)$.*

(3) *If we define the curve γ_{t_0} by $\gamma_{t_0}(s) = \gamma(t_0 + s)$, then its interval of definition is $I - t_0 = \{s - t_0 \mid s \in\}$, and for every $[\alpha, \beta] \subset I - t_0$, with $\alpha \leq \beta$, then*

$$\mathcal{D}(\gamma_{t_0}, u, c; \alpha, \beta) = \mathcal{D}(\gamma, u, c; \alpha + t_0, \beta + t_0).$$

(4) **[Chasles Property]** *If* $a_,a_2, a_3 \in I$*, with* $a_1 \leq a_2 \leq a_3$*, then*

$$\mathcal{D}(\gamma, u, c; a_1, a_3) = \mathcal{D}(\gamma, u, c; a_1, a_2) + \mathcal{D}(\gamma, u, c; a_2, a_3).$$

(5) *If* $\alpha_1, \beta_1, \alpha, \beta \in I$*, with* $\alpha \leq \alpha_1 \leq \beta_1 \leq \beta$*, then*

$$\mathcal{D}(\gamma, u, c; \alpha, \beta) \geq \mathcal{D}(\gamma, u, c; \alpha_1, \beta_1) \geq 0.$$

(6) *The function* $(\alpha, \beta) \mapsto \mathcal{D}(\gamma, u, c; \alpha, \beta)$ *is continuous on the set* $\{(\alpha, \beta) \in I \times I \mid \alpha \leq \beta\}$*.*

(7) *If* $\gamma_n : [a, b] \to U$ *is a sequence of* C^1 *curves which converges to the* C^1 *curve* $\gamma_\infty : [a, b] \to U$ *in the* C^1 *topology then* $\mathcal{D}(\gamma_n, u, c; a, b) \to \mathcal{D}(\gamma_\infty, u, c; a, b)$*.*

(8) *Let* $u_n : U \to \mathbb{R}$ *is a sequence of functions, with* $u_n \prec L + c_n$*, where* $c_n \in \mathbb{R}$*. If* $c_n \to c$ *and* $u_n(x) \to u(x)$ *at every point* $x \in U$*, then* $\mathcal{D}(\gamma, u_n, c_n; a, b) \to \mathcal{D}(\gamma, u, c; a, b)$*.*

Proof. Claim (1) is easy since $U \prec L + c$ on U implies

$$u(\gamma(\beta) - u(\gamma(\alpha) \leq \int_\alpha^\beta L(\gamma(s), \dot{\gamma}(s))\, ds + c(\beta - \alpha).$$

Claims (2,3,4) follow easily from the definition of the defect.

For claim (5), by Chasles Property (4), we have

$$\mathcal{D}(\gamma, u, c; \alpha, \beta) = \mathcal{D}(\gamma, u, c; \alpha, \alpha_1) + \mathcal{D}(\gamma, u, c; \alpha_1, \beta_1) + \mathcal{D}(\gamma, u, c; \beta_1, \beta).$$

But by claim (1), we have $\mathcal{D}(\gamma, u, c; \alpha, \alpha_1) \geq 0$ and $\mathcal{D}(\gamma, u, c; \beta_1, \beta) \geq 0$.

Claims (6,7) and (8) can also be obtained from the definition of the defect since a dominated function is locally Lipschitz, and the action of the Lagrangian for C^1 curves is continuous in the C^1 topology on the space of C^1 curves. □

The first corollary we obtain is a simplification of the definition of a (u, L, c)-calibrated curve on a compact interval when $u \prec L+c$.

Corollary 4.3.3. *Let L be a Tonelli Lagarangian on the manifold M. Suppose that $u : U \to \mathbb{R}$ is a continuous function defined on the open subset $U \subset M$, with $u \prec L + c$, and that $\gamma : [a, b] \to U$ is a continuous piecewise C^1 curve. the following properties are equivalent*

(i) *The curve γ is (u, L, c)-calibrated.*

(ii) *We have*

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

(iii) *The defect $\mathcal{D}(\gamma, u, c; a, b)$ is equal to 0.*

Proof. It is clear that (ii) and (iii) are equivalent. Obviously if γ is (u, L, c)-calibrated we have

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

It remains to prove that (iii) implies (i). If $\mathcal{D}(\gamma, u, c; a, b) = 0$, from claims (1) and (5) of Proposition 4.3.2 above, we get that $\mathcal{D}(\gamma, u, c; \alpha, \beta) = 0$, for every subinterval $[\alpha, \beta] \subset [a, b]$, and hence

$$u(\gamma(\beta) - u(\gamma(\alpha) = \int_\alpha^\beta L(\gamma(s), \dot\gamma(s))\, ds + c(\beta - \alpha),$$

hence γ is (u, L, c)-calibrated on the interval $[a, b]$. □

We now state some of the properties of calibrated curves.

Corollary 4.3.4. *Let L be a Tonelli Lagarangian on the manifold M. Suppose that $u : U \to \mathbb{R}$ is a continuous function defined on the open subset $U \subset M$, that $\gamma : I \to U$ is a continuous piecewise C^1 curve. We have the following properties:*

(1) *If $\gamma : I \to M$ is (u, L, c)-calibrated, then for every subinterval $I' \subset I$ the restriction $\gamma|I'$ is also (u, Lc)-calibrated.*

(2) *If I' is a subinterval of I and the restriction $\gamma|I'$ is (u, L, c)-calibrated, then γ is (u, L, c)-calibrated on the interval $\bar{I}' \cap I$.*

(3) *Suppose that I is a finite union of subintervals $I_1, \dots, I_n$. If $\gamma|I_i$ is (u, L, c)-calibrated, for $i = 1, \dots, n$, then γ is (u, L, c)-calibrated (on I).*

(4) *Suppose that $I = \cup_{n\in\mathbb{N}} I_i$, where each I_i is an interval and $I_i \subset I_{i+1}$. If each $\gamma|I_i$ is (u, L, c)-calibrated, then γ is (u, L, c)-calibrated (on I).*

(5) *For every $t_0 \in I$, there exists a largest subinterval $I_{t_0} \subset I$ containing x_0 on which γ is (u, L, c)-calibrated. Moreover $I_{t_0} = \bar{I}_{t_0} \cap I$.*

(6) *If $k \in \mathbb{R}$, then γ is (u, L, c)-calibrated if and only if it is $(u + k, L, c)$-calibrated.*

(7) *If $t_0 \in \mathbb{R}$, then γ is (u, L, c)-calibrated on I if and only if the curve $s \mapsto \gamma(s + t_0)$ is (u, L, c)-calibrated on the interval $I - t_0 = \{t - t_0 \mid t \in I\}$.*

(8) *If $\gamma_n : [a, b] \to U$ is a sequence of C^1 curves which converges to the C^1 curve $\gamma_\infty : [a, b] \to U$ in the C^1 topology, and γ_n is (u, L, c)-calibrated for every n then the curve γ is also (u, L, c)-calibrated.*

Proof. Claim (1) was already given as Proposition 4.1.8. It is a simple consequence of the definition of a calibrated curve.

To prove claim (2), consider a compact subinterval $[a, b] \subset I \cap \bar{I}'$, with $a < b$. For n large enough we have $a + 1/n < b - 1/n$, and $[a+1/n, b-1/n] \subset I'$, therefore by by claim (1) and Corollary 4.3.3 above we get $\mathcal{D}(\gamma, u, c; a + 1/n, b - 1/n) = 0$. By claim (6) of Proposition 4.3.2 above $\mathcal{D}(\gamma, u, c; a, b) = 0$. Therefore $\gamma|[a, b]$ is (u, L, c)-calibrated for every compact subinterval $[a, b]$ of $\bar{I}'$. Hence $\gamma|\bar{I}'$ is (u, L, c)-calibrated.

To prove (3), we fix $[a, b] \subset I$. By (2) we can assume that each $I_i \cap [a, b]$ is a compact interval $[a_i, b_i]$. Extracting a minimal cover and reindexing, we can assume

$$\begin{aligned} a = a_1 \le a_2 \le b_1 \le a_3 \le b_2 \le \cdots \le \\ \le a_{i+1} \le b_i \le \cdots \le a_n \le b_{n-1} \le b_n = b. \end{aligned} \tag{$*$}$$

It suffices to prove by induction on i that $\mathcal{D}(\gamma, u, c; a_1, b_i) = 0$ For $i = 1$, this follows from the hypothesis that $\gamma|I_1$ is (u, L, c)-calibrated. Let us now do the induction step from i to $i + 1$. By $(*)$, we have $[b_i, b_{i+1}] \subset [a_{i+1}, b_{i+1}] \subset I_{i+1}$. Since $\gamma|I_{i+1}$ is (u, L, c)-calibrated, we therefore obtain $\mathcal{D}(\gamma, u, c; b_i, b_{i+1}) = 0$. From Chasles Property, claim (4) of Proposition 4.3.2, it follows that $\mathcal{D}(\gamma, u, c; a_1, b_{i+1}) = 0$.

To prove (4), it suffices to observe that if $[a, b] \in I$ is a compact subinterval, then for n large enough we have $[a, b] \subset I_n$.

To prove (5), call Λ the family of subinterval $J \subset I$ such that $t_0 \in J$ and $\gamma|J$ is (u, L, c)-calibrated. Notice that Λ is not empty since $[t_0, t_0] \in \Lambda$. Since $t_0 \in \bigcap_{J\in\Lambda} J$, the union $I_{t_0} \bigcup_{J\in\Lambda} J$ is an interval. We have to show that γ is (u, L, c)-calibrated on I_{t_0}. Let $[a, b] \subset I_{t_0}$. We can find $J_1, J_2 \in \Lambda$ with $a \in J_1, b \in J_2$. the union $J_3 = J_1 \cup J_2$ is an interval because $t_0 \in J_1 \cap J_2$. Therefore by (3), the restriction $\gamma|J_3$ is (u, L, c)-calibrated. This finishes the proof since $[a, b] \subset J_3$.

Claim (6), follows from Corollary 4.3.3 above and claim (2) of Proposition 4.3.2. In the same way claim (7), follows from Corollary 4.3.3 above and claim (3) of Proposition 4.3.2. Claim (8) follows from Corollary 4.3.3 above and claim (7) of Proposition 4.3.2. □

Exercise 4.3.5. *Let L be a Tonelli Lagarangian on the manifold M. Suppose that $u : U \to \mathbb{R}$ is a continuous function defined on the open subset $U \subset M$, that $\gamma : I \to U$ is a continuous piecewise C^1 curve. Suppose that $I = \bigcup_{n\in\mathbb{N}} I_n$, with I_n a subinterval I of I on which γ is (u, L, c)-calibrated. (Do not assume that the family $I_n, n \in \mathbb{N}$ is increasing.) Show that γ is (u, L, c)-calibrated on I. [Indication: Reduce to the case $I = [a, b]$ and each I_n compact. Show, using part (5) of Proposition 4.3.4 above, that you can assume that the I_n are pairwise disjoint. Call F the complement in $[a, b]$ of the union $\bigcup_{n\in\mathbb{N}} \mathring{I}_n$. Show that F is countable. Use Baire's Category Theorem to show that if $F \setminus \{a, b\}$ is not empty then it must has an isolated point.]*

Let us now show that infinite calibrated curves can only exist for the Ma né critical value $c[0]$.

Proposition 4.3.6. *Suppose $u : M \to \mathbb{R}$ is $\prec L + c$. If it admits a (u, L, c)-calibrated curve $\gamma : I \to M$, with I an infinite interval then necessarily c is equal to the Ma né critical value $c[0]$.*

Proof. By the definition of the Ma né critical value, since $u \prec L+c$ and u is defined on the whole of M, we have $c \geq c[0]$. To prove the converse inequality, we pick a continuous function $u_0 : M \to \mathbb{R}$ with $u_0 \prec L + c[0]$. For any $a, b \in I$ with $a \leq b$, we have

$$u(\gamma(b)) - u(\gamma(a)) = \int_a^b L(\gamma(s), \dot{\gamma}(s))\, ds + c(b - a)$$
$$u_0(\gamma(b)) - u_0(\gamma(a)) \leq \int_a^b L(\gamma(s), \dot{\gamma}(s))\, ds + c[0](b - a).$$

Subtracting the first equality from the second inequality, we get

$$u_0(\gamma(b)) - u_0(\gamma(b)) - u(\gamma(b)) + u(\gamma(a)) \leq (c[0] - c)(b - a).$$

Since both u and u_0 are continuous functions on the compact space M, the constant

$$K = \sup x \in M|u_0(x)| + \sup x \in M|u(x)|$$

is finite, and we obtain from the inequality above

$$-2K \leq (c[0] - c)(b - a),$$

for all $a, b \in I$ with $a \leq b$. Since I is an infinite interval, we can find sequences $a_n, b_n \in I$ with $a_n < lb_n$ such that $b_n - a_n \to \infty$, as $n \to \infty$. This yields $c[0] - c \geq -2L/(b_n - a_n) \to 0$. Hence we obtain $c[0] \geq c$. □

The following corollary is now a consequence of Definition 4.1.11 of a weak KAM solution and Proposition 4.3.6 above.

Corollary 4.3.7. *If $u : M \to \mathbb{R}$ is a negative (resp. positive) weak KAM solution with the constant c, on the compact manifold M, then c is necessarily c is equal to the Ma né critical value $c[0]$.*

Theorem 4.3.8. *Suppose that L is a Tonelli Lagrangian on the manifold M. Let $u : U \to \mathbb{R}$ be a function defined on the open subset $U \subset M$. Assume that $u \prec L + c$, where $c \in \mathbb{R}$, and that $\gamma : [a, b] \to M$ is a (u, L, c)-calibrated curve, with $a < b$, then we have:*

(i) *If for some* t*, the derivative of* u *at* $\gamma(t)$ *exists then*

$$d_{\gamma(t)}u = \frac{\partial L}{\partial v}(\gamma(t),\dot\gamma(t)) \text{ and } H(\gamma(t), d_{\gamma(t)}u) = c,$$

where H *is the Hamiltonian associated to* L*.*

(ii) *For every* $t \in]a,b[$ *the derivative of* u *at* $\gamma(t)$ *exists.*

Proof. We prove (i). We will assume $t < b$ (for the case $t = b$ use a $t' < t$ in the argument). For $t' \in [a,b]$ satisfying $t' > t$, the calibration condition implies

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

Dividing by $t' - t$ and letting $t' \to t$, we obtain

$$d_{\gamma(t)}u(\dot\gamma(t)) = L(\gamma(t),\dot\gamma(s)) + c.$$

Combining with the Fenchel Inequality 1.3.1 we get

$$c = d_{\gamma(t)}u(\dot\gamma(t)) - L(\gamma(t),\dot\gamma(s)) \leq H(\gamma(t), d_{\gamma(t)}u). \qquad (*)$$

But by Proposition 4.2.2, we know that $H(\gamma(t), d_{\gamma(t)}u) \leq c$. This yields equality in $(*)$. Therefore $H(\gamma(t), d_{\gamma(t)}u) = c$. But also the equality

$$d_{\gamma(t)}u(\dot\gamma(t)) - L(\gamma(t),\dot\gamma(s)) = H(\gamma(t), d_{\gamma(t)}u)$$

means that we have equality in the Fenchel inequality, therefore we conclude

$$d_{\gamma(t)}u = \frac{\partial L}{\partial v}(\gamma(t),\dot\gamma(t))$$

To prove (ii), we choose a open C^∞ chart $\varphi : U' \to \mathbb{R}^k$ on M, such that $\varphi(U') = \mathbb{R}^k$ and $x = \gamma(t) \in U' \subset U$. We can find a', b' such that $a \leq a' < t < b' \leq b$ and $\gamma([a',b']) \subset U$. To simplify notations, we identify U' and $\mathbb{R}^k$ via φ. For every $y \in U' = \mathbb{R}^k$, we define the curve $\gamma_y : [a',t] \to U'$ by

$$\gamma_y(s) = \gamma(s) + \frac{s-a'}{t-a'}(y-x).$$

We have $\gamma_y(a') = \gamma(a')$, and $\gamma_y(t) = y$, since $x = \gamma(t)$. Moreover, we also have $\gamma_x = \gamma|[a', t]$. Since $u \prec L + c$, we obtain

$$u(y) \leq u(\gamma(a')) + \int_{a'}^{t} L(\gamma_y(s), \dot{\gamma}_y(s))\, ds + c(t - a'),$$

with equality at $x = \gamma(t)$ since $\gamma_x = \gamma$ is (u, L, c)-calibrated. If we define $\psi_+ : U' \to \mathbb{R}$ by

$$\begin{aligned}\psi_+(y) &= u(\gamma(a')) + \int_{a'}^{t} L(\gamma_y(s), \dot{\gamma}_y(s))\, ds + c(t - a') \\ &= u(\gamma(a'))+ \int_{a'}^{t} L(\gamma(s) + \frac{s - a'}{t - a'}(y - x), \dot{\gamma}(s) + \frac{y - x}{t - a'})\, ds \\ &\qquad + c(t - a'),\end{aligned}$$

we easily see that ψ_+ is as smooth as L (note that γ is a minimizer, and is therefore as smooth as L). Moreover we have $u(y) \leq \psi_+(y)$, with equality at x.

We now will find a function $\psi_- : U' \to \mathbb{R}$ satisfying $\psi_- \leq u$ with equality at $x = \gamma(t)$. For this, given $y \in U' = \mathbb{R}^k$ we define $\tilde{\gamma}_y : [t, b'] \to U'$ by

$$\tilde{\gamma}_y(s) = \gamma(s) + \frac{b' - s}{b' - t}(y - x).$$

Again we get $\tilde{\gamma}_y(b') = \gamma(b'), \tilde{\gamma}_y(t) = y$ and $\tilde{\gamma}_x = \gamma|[t, b']$. Since $u \prec L + c$, we obtain

$$u(\gamma(b')) - u(y) \leq \int_{t}^{b'} L(\tilde{\gamma}_y(s), \dot{\tilde{\gamma}}_y(s))\, ds + c(b' - t),$$

with equality at x since $\tilde{\gamma}_x = \gamma$ is (u, L, c)-calibrated. If we define $\psi_- : U' \to \mathbb{R}$ by

$$\begin{aligned}\psi_-(y) &= u(\gamma(b')) - \int_{t}^{b'} L(\tilde{\gamma}_y(s), \dot{\tilde{\gamma}}_y(s))\, ds - c(b' - t) \\ &= u(\gamma(b'))) - \int_{t}^{b'} L(\gamma(s) + \frac{b' - s}{b' - t}(y - x), \dot{\gamma}(s) - \frac{y - x}{b' - t})\, ds \\ &\qquad - c(b' - t).\end{aligned}$$

Again we easily see that ψ_- is as smooth as L. Moreover we have $u(y) \geq \psi_-(y)$, with equality at x.

Since $\psi_-(y) \leq u(y) \leq \psi_+(y)$, with equality at 0, the C^1 function $\psi_+ - \psi_-$ is non negative and is equal to 0 at x therefore its derivative at x is 0. Call p the common value $d_x\psi_+ = d_x\psi_-$. By the definition of the derivative, using $\psi_-(x) = u(x) = \psi_+(x)$, we can write $\psi_\pm(y) = u(x) + p(y-x) + \|y-x\|\beta_\pm(y-x)$, with $\lim_{h\to 0}\beta_\pm(h) = 0$. The inequality $\psi_-(y) \leq u(y) \leq \psi_+(y)$ now gives

$$\begin{aligned} u(x) + p(y-x) + \|y-x\|\beta_-(y-x) \leq u(y) \leq \\ \leq u(x) + p(y-x) + \|y-x\|\beta_+(y-x). \end{aligned}$$

This obviously implies that p is the derivative of u at $x = \gamma(t)$. □

Another important property of calibrated curves and dominated functions is given in the following theorem. We will call this result the Lyapunov property for reasons that will become clear later, see ????? below.

Theorem 4.3.9 (Lyapunov Property)**.** *Let L be a Tonelli Lagrangian on the manifold M. Suppose that $\gamma : [a,b] \to M$ is a continuous piecewise C^1 curve, and that u_1, u_2 are two real-valued functions defined on a neighborhood of $\gamma([a,b])$. If, for some $c \in \mathbb{R}$, the curve γ is (u_1, L, c)-calibrated and $u_2 \prec L + c$ on a neighborhood of $\gamma([a,b])$, then the function $t \mapsto u_2(\gamma(t)) - u_1(\gamma(t))$ is non-increasing on $[a,b]$.*

Proof. Since γ is (u_1, L, c)-calibrated, for $t, t' \in [a,b]$ with $t \leq t'$ we have

$$u_1(\gamma(t')) - u_1(\gamma(t)) = \int_t^{t'} L(\gamma(s), \dot\gamma(s))\, ds + c(t'-t).$$

Using that $u_2 \prec L + c$ on a neighborhood of $\gamma([a,b])$, we get

$$u_2(\gamma(t')) - u_2(\gamma(t)) \leq \int_t^{t'} L(\gamma(s), \dot\gamma(s))\, ds + c(t'-t).$$

Comparing we obtain $u_2(\gamma(t')) - u_2(\gamma(t)) \leq u_1(\gamma(t')) - u_1(\gamma(t))$, therefore $u_2(\gamma(t')) - u_1(\gamma(t')) \leq u_2(\gamma(t)) - u_1(\gamma(t))$. □

4.4 Minimal Action for a Given Time

As we said minimizers are the important object of the theory. Therefore if we fix a pair of points x, y and a time $t > 0$, the minimal action of a curve joining x to y in time t will enjoy some special properties.

Definition 4.4.1 (Minimal Action). If L is a Tonelli Lagrangian on the compact connected manifold M, for $t > 0$ fixed, we define the function $h_t : M \times M \to \mathbb{R}$ by

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

where the infimum is taken over all the (continuous) piecewise C^1 curves $\gamma : [0, t] \to M$ with $\gamma(0) = x$ and $\gamma(t) = y$.

The quantity $h_t(x, y)$ is called the minimal action to go from x to y in time t.

Note that h_t is well-defined since we are assuming that M is connected, therefore any pair of points in M can be joined by a smooth path. Moreover, the function h_t is finite valued since L is bounded from below, by superlinearity and compactness of M.

Of course, in the definition of $h_t(x, y)$, we could have taken the infimum on all absolute continuous paths. this would have note changed the value of $h_t(x, y)$, since minimizers for a given positive time between two points do always exist by Tonelli's theorem 3.3.4 and are in fact as smooth as the Lagragian by the regularity theorem 3.7.1.

Here are some of the important properties of h_t.

Proposition 4.4.2 (Properties of h_t). *The properties of h_t are*

(1) *For each $x, y \in M$, and each $t > 0$, we have*

$$h_t(x, y) \geq t \inf_{TM} L.$$

(2) *For each $x, z \in M$ and each $t, t' > 0$, we have*

$$h_{t+t'}(x, z) = \inf_{y \in M} h_t(x, y) + h_{t'}(y, z).$$

(3) *If $u : M \to \mathbb{R}$ is a function defined on the whole of M, then $u \prec L + c$ if and only if*

$$\forall x, y \in M, \forall t > 0, u(y) - u(x) \leq h_t(x, y) + ct.$$

(4) *A continuous piecewise C^1 curve $\gamma : [a, b] \to M$, with $a < b$, is minimizing if and only if*

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

(5) *For each $t > 0$ and each $x, y \in M$, there exists an extremal curve $\gamma : [0, t] \to M$ with $\gamma(0) = x, \gamma(t) = y$ and $h_t(x, y) = \int_0^t L(\gamma(s), \dot{\gamma}(s))\, ds$.*

Proof. Property (1) is obvious since for any continuous piecewise C^1 curve $\gamma : [0, t] \to M$, with $t > 0$, we have

$$\int_0^t L(\gamma(s), \dot{\gamma}(s))\, ds \geq \int_0^t \inf_{TM} L\, ds = t \inf_{TM} L.$$

To prove property (2), let us consider two continuous piecewise linear curves $\gamma_1 : [0, t] \to M$, with $\gamma_1(0) = x, \gamma_1(t) = y$, and $\gamma_2 : [0, t'] \to M$, with $\gamma_2(0) = y, \gamma_2(t') = y$. We can define the curve $\gamma_3 = \gamma_1 * \gamma_2 : [0, t + t'] \to M$ by

$$\begin{aligned} \gamma_3(s) &= \gamma_1(s), \text{ for } s \in [0, t] \\ &= \gamma_2(s - t), \text{ for } s \in [t, t + t']. \end{aligned}$$

The curve γ_3 is continuous piecewise C^1, with $\gamma_3(0) = x$ and $\gamma_3(t+t') = z$. Moreover, its action $\mathbb{L}(\gamma_3)$ is equal to $\mathbb{L}(\gamma_1) + \mathbb{L}(\gamma_2)$. Therefore, we have

$$h_{t+t'}(x, z) \leq \mathbb{L}(\gamma_1) + \mathbb{L}(\gamma_2).$$

Taking the infimum over all possible γ_1 and γ_2 gives

$$h_{t+t'}(x, z) \leq h_t(x, y) + h_{t'}(y, z).$$

Taking now the infimum over $y \in M$, we obtain

$$h_{t+t'}(x, z) \leq \inf_{y \in M} h_t(x, y) + h_{t'}(y, z).$$

We now prove that the inequality above is an equality. Given $\epsilon > 0$, we can find a continuous piecewise C^1 curve $\gamma : [0, t+t'] \to M$, with $\gamma(0) = x, \gamma(t+t') = z$, and

$$\int_0^{t+t'} L(\gamma(s), \dot\gamma(s))\, ds \le h_{t+t'}(x,z) + \epsilon.$$

Using the curve $\gamma|[0,t]$, we obtain that

$$h_t(x, \gamma(t)) \le \int_0^t L(\gamma(s), \dot\gamma(s))\, ds.$$

Using a reparametrization of $\gamma|[t, t+t']$ by $[0, t']$, we obtain

$$h_{t'}(\gamma(t), z) \le \int_t^{t+t'} L(\gamma(s), \dot\gamma(s))\, ds.$$

Adding these inequalities, we get

$$h_t(x, \gamma(t)) + h_{t'}(\gamma(t), z) \le \int_0^{t+t'} L(\gamma(s), \dot\gamma(s))\, ds \le h_{t+t'}(x,z) + \epsilon.$$

Therefore $\inf_{y \in M} h_t(x,y) + h_{t'}(y,z) \le h_{t+t'}(x,z) + \epsilon$, for every $\epsilon > 0$. We conclude by letting $\epsilon \to 0$.

For property (3), we observe that if $\gamma : [a,b] \to M$ is a continuous piecewise C^1 curve, then the reparametrized curve $\tilde\gamma : [0, b-a] \to M$ defined by

$$\tilde\gamma(s) = \gamma(a+s),$$

has the same endpoints as γ, and also the same action, since L is time-independent. In particular we could have defined h_t by

$$h_t(x,y) = \inf_\gamma \int_a^b L(\gamma(s), \dot\gamma(s))\, ds,$$

where the infimum is taken over all the continuous piecewise C^1 curves $\gamma : [a,b] \to M$ with $\gamma(a) = x, \gamma(b) = y$, and $b - a = t$. With this observation, we get that for a given $t > 0$, the inequality

$$u(y) - u(x) \le h_t(x,y) + ct$$

is equivalent to

$$u(\gamma(b)) - u(\gamma(a)) \leq \int_a^b L(\gamma(s), \dot\gamma(s))\, ds,$$

for every continuous piecewise C^1 curves $\gamma : [a, b] \to M$ with $\gamma(a) = x, \gamma(b) = y$, and $b - a = t$.

For property (4), we observe that, by definition of h_t, we have

$$h_t(\gamma(0), \gamma(t)) = \int_0^t L(\gamma(s), \dot\gamma(s))\, ds,$$

for a curve $\gamma : [0, t] \to M$ if and only if γ is a minimizer. It remains to observe as indicated above that we can reparametrize any curve by an interval starting at 0 without changing neither its endpoints nor its action.

Property (5) results from Tonelli's Theorem 3.3.1. □

It is probably difficult to find out when the next theorem appeared, in some of its forms, for the first time in the literature. It has certainly been known for some time now, at least in its equivalent form given as Lemma 4.6.3 below, see for example [Fle69, Theorem1, page 518]. It has of course been, in a form or another, been rediscovered by several people, including the author himself, for whom it started weak KAM Theory since it has as an "obvious" consequence the existence of a fixed point (up to a constant) for the Lax–Oleinik semi-group, see section 4 below.

Theorem 4.4.3 (Fleming's Lemma)**.** *For each $t_0 > 0$, there exists a constant $\kappa_{t_0} \in [0, +\infty[$ such that, for each $t \geq t_0$ the function $h_t : M \times M \to \mathbb{R}$ is Lipschitzian with a Lipschitz constant $\leq \kappa_{t_0}$.*

Before proving the theorem, we need to prove some preliminary results.

Proposition 4.4.4. *Let L be a Tonelli Lagrangian on the compact connected manifold M. For every given $t > 0$, there exists a constant $C_t < +\infty$, such that, for each $x, y \in M$, we can find a C^∞ curve $\gamma : [0, t] \to M$ with $\gamma(0) = x, \gamma(t) = y$ and*

$$\mathbb{L}(\gamma) = \int_0^t L(\gamma(s), \dot\gamma(s))\, ds \leq C_t.$$

Proof. Choose a Riemannian metric g on M. By the compactness of M, we can find a geodesic (for the metric g) between x and y and whose length is $d(x,y)$. Let us parametrize this geodesic by the interval $[0,t]$ with a speed of constant norm and denote by $\gamma : [0,t] \to M$ this parametrization, with $\gamma(0) = x, \gamma(t) = y$. As the length of this curve is $d(x,y)$, we find that

$$\forall s \in [0,t],\ \|\dot\gamma(s)\|_{\gamma(s)} = \frac{d(x,y)}{t}.$$

Since the manifold M is compact, the diameter $\operatorname{diam}(M)$ of M for the metric d is finite, consequently, the set

$$A_t = \{(x,v) \in TM \mid \|v\|_x \leq \frac{\operatorname{diam}(M)}{t}\}$$

is compact. We have $(\gamma(s), \dot\gamma(s)) \in A_t$, for all $s \in [0,t]$. By compactness of A_t, we can find a constant $\tilde{C}_t < +\infty$ such that

$$\forall (x,v) \in A_t,\ L(x,v) \leq \tilde{C}_t.$$

If we set $C_t = t\tilde{C}_t$, we do indeed have $\mathbb{L}(\gamma) \leq C_t$. □

Corollary 4.4.5 (A Priori Compactness)**.** *Let L be a Tonelli Lagrangian on the compact manifold M. If $t > 0$ is fixed, there exists a compact subset $K_t \subset TM$ such that for every minimizing extremal curve $\gamma : [a,b] \to M$, with $b - a \geq t$, we have*

$$\forall s \in [a,b],\ (\gamma(s), \dot\gamma(s)) \in K_t.$$

Proof. Let us recall that we are assuming in this chapter that M is compact and *connected*. We first observe that it is enough to show the corollary if $[a,b] = [0,t]$. Indeed, if $t_0 \in [a,b]$, we can find an interval of the form $[c, c+t]$, with $t_0 \in [c, c+t] \subset [a,b]$. The curve $\gamma_c : [0,t] \to M, s \mapsto \gamma(c+s)$ satisfies the assumptions of the corollary with $[0,t]$ in place of $[a,b]$.

Thus let us give the proof of the corollary with $[a,b] = [0,t]$. With the notations of the previous Proposition 4.4.4, we necessarily have

$$\mathbb{L}(\gamma) = \int_0^t L(\gamma(s), \dot\gamma(s))\, ds \leq C_t.$$

Since $s \mapsto L(\gamma(s), \dot\gamma(s))$ is continuous on $[0,t]$, by the Mean Value Theorem, we can find $s_0 \in [0,t]$ such that

$$L(\gamma(s_0), \dot\gamma(s_0)) \leq \frac{C_t}{t}. \tag{*}$$

The set $B = \{(x,v) \in TM \mid L(x,v) \leq \frac{C_t}{t}\}$ is a compact subset of TM. By continuity of the flow ϕ_t, the set $K_t = \bigcup_{|s|\leq t} \phi_s(B)$ is also compact subset of TM. As γ is an extremal curve, the inequality $(*)$ shows that

$$\forall s \in [0,t],\ (\gamma(s), \dot\gamma(s)) \in \phi_{s-s_0}(B) \subset K_t.$$

□

Proof of Theorem 4.4.3. We fix some $t_0 > 0$, and we will study h_t only for $t \geq t_0$.

Let us consider $\bar B(0,3)$ the closed ball of center 0 and radius 3 in the Euclidean space $\mathbb{R}^k$, where k is the dimension of M. By compactness of M, we can find a finite number of coordinate charts $\varphi_i : \mathbb{R}^k \to M, i = 1,\dots,p$, such that $M = \bigcup_{i=1}^p \varphi_i(\mathring{B}(0,1))$. We denote by $\eta > 0$, a constant such that

$$\begin{aligned}\forall i = 1,\dots,p, \forall x,x' \in M,\ d(x,x') \leq \eta \text{ and } x \in \varphi_i\big(\bar B(0,1)\big) \Rightarrow \\ x' \in \varphi_i\big(\mathring{B}(0,2)\big) \text{ and } \|\varphi_i^{-1}(x') - \varphi_i^{-1}(x)\| \leq 1,\end{aligned}$$

where the norm $\|\cdot\|$ is the Euclidean norm. Let us denote by K_{t_0} the compact set obtained from corollary 4.4.5. We can find a constant $A < +\infty$ such that

$$\forall (x,v) \in K_{t_0}, \|v\|_x \leq A.$$

In the remaining part of the proof, we set

$$\epsilon = \min\big(t/2, \frac{\eta}{A}\big).$$

In the same way by the compactness of K_{t_0}, we can find a constant $B < +\infty$ such that

$$\forall x \in \bar B(0,3), \forall v \in \mathbb{R}^k, \forall i = 1,\dots,p,\ T\varphi_i(x,v) \in K_{t_0} \Rightarrow \|v\| \leq B.$$

Consider two points $x, y \in M$, and suppose that $i, j \in \{1, \ldots, p\}$ are such that $x \in \varphi_i\big(\mathring{B}(0,1)\big)$ and $y \in \varphi_j\big(\mathring{B}(0,1)\big)$. For a fixed $t \geq t_0$, we can find a minimizing extremal curve $\gamma : [0,t] \to M$ such that $\gamma(0) = x, \gamma(t) = y$ and

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

By the a priori compactness given by Corollary 4.4.5, since $t \geq t_0$, we have $(\gamma(s), \dot\gamma(s)) \in K_{t_0}$, for each $s \in [0,t]$. Consequently, by the choice of A, we obtain $d(\gamma(s), \gamma(s')) \leq A|s - s'|$, for all $s, s' \in [0,t]$. In particular, since $x = \gamma(0) \in \varphi_i\big(\mathring{B}(0,1)\big)$ and $y = \gamma(t) \in \varphi_j\big(\mathring{B}(0,1)\big)$by the choice of ϵ, we find that

$$\begin{aligned}\gamma([0,\epsilon]) &\subset \varphi_i\big(\mathring{B}(0,2)\big)\\ \gamma([t-\epsilon,t]) &\subset \varphi_j\big(\mathring{B}(0,2)\big).\end{aligned}$$

We can then define the two curves $\tilde\gamma^0 : [0,\epsilon] \to \mathring{B}(0,2)$ and $\tilde\gamma^1 : [t-\epsilon, t] \to \mathring{B}(0,2)$ by $\varphi_i(\tilde\gamma^0(s)) = \gamma(s)$ and $\varphi_j(\tilde\gamma^1(s)) = \gamma(s)$. If $d(x,x') \leq \eta$ and $d(y,y') \leq \eta$, there are unique $\tilde x', \tilde y' \in \mathring{B}(0,2)$ such that $\varphi_i(\tilde x') = x'$ and $\varphi_j(\tilde y') = y'$. By the definition of η, we also have $\|\tilde x' - \tilde x\| \leq 1$ and $\|\tilde y' - \tilde y\| \leq 1$. Let us define curves $\tilde\gamma^0_{x',y'} : [0,\epsilon] \to \mathring{B}(0,3)$ and $\tilde\gamma^1_{x',y'} : [t-\epsilon,t] \to \mathring{B}(0,3)$ by

$$\begin{aligned}\tilde\gamma^0_{x',y'}(s) &= \frac{\epsilon - s}{\epsilon}(\tilde x' - \tilde x) + \tilde\gamma(s), \text{ for } s \in 0,\epsilon],\\ \tilde\gamma^1_{x',y'}(s) &= \frac{s - (t-\epsilon)}{\epsilon}(\tilde y' - \tilde y) + \tilde\gamma(s), \text{ for } s \in [t-\epsilon, t].\end{aligned}$$

The curve $\tilde\gamma^0_{x',y'}|[0,\epsilon]$ connects the point $\tilde x'$ to the point $\tilde\gamma(\epsilon)$, and the curve $\tilde\gamma^1_{x',y'}|[t-\epsilon,t]$ connects the point $\tilde\gamma(t-\epsilon)$ to the point $\tilde y'$. Since $\epsilon \leq T/2$ we can, then, define the curve $\gamma_{x',y'} : [0,t] \to M$ by $\gamma_{x',y'} = \gamma$ on $[\epsilon, t-\epsilon]$, $\gamma_{x',y'} = \varphi_i \circ \tilde\gamma^0_{x,'y'}$ on $[0,\epsilon]$, and $\gamma_{x',y'} = \varphi_j \circ \tilde\gamma^1_{x,'y'}$ on $[t-\epsilon,t]$. The curve $\gamma_{x',y'}$ is continuous on the interval $[0,t]$, moreover, it of class C^1 on each of each the intervals $[0,\epsilon], [\epsilon, t-\epsilon]$ and $[t-\epsilon, t]$. We of course have $\gamma_{x,y} = \gamma$. As $\gamma_{x',y'}(s)$ is equal to x', for $s = 0$, and to y', for $s = t$, we have

$$h_t(x',y') \leq \int_0^t L(\gamma_{x',y'}(s), \dot\gamma_{x',y'}(s))\, ds.$$

Subtracting the equality $(*)$ above from this inequality and using the fact that $\gamma_{x',y'} = \gamma$ on $[\epsilon, t-\epsilon]$, we find

$$
\begin{aligned}
h_t(x', y') - h_t(x, y) \leq \int_0^\epsilon L(\gamma_{x',y'}(s), \dot{\gamma}_{x',y'}(s))\, ds - \int_0^\epsilon L(\gamma(s), \dot{\gamma}(s))\, ds \\
+ \int_{t-\epsilon}^t L(\gamma_{x',y'}(s), \dot{\gamma}_{x',y'}(s))\, ds - \int_{t-\epsilon}^t L(\gamma(s), \dot{\gamma}(s))\, ds.
\end{aligned}
$$

We will use the coordinate charts φ_i and φ_j to estimate the right-hand side of this inequality. For that, it is convenient, for $\ell = 1, \dots, p$, to consider the Lagrangian $\tilde{L}_\ell : \bar{B}(0,3) \times \mathbb{R}^k \to \mathbb{R}$ defined by

$$\tilde{L}_\ell(z, w) = L\big(\varphi_\ell(z), D\varphi_\ell(w)[w]\big).$$

This Lagrangian $\tilde{L}_\ell$ is of class C^r on $\bar{B}(0,3) \times \mathbb{R}^k$. Using these Lagrangian for $\ell = i, j$, we have

$$
\begin{aligned}
h_t(x', y') - h_t(x, y) \leq \\
\int_0^\epsilon \tilde{L}_i(\tilde{\gamma}^0_{x',y'}(s), \dot{\tilde{\gamma}}^0_{x',y'}(s))\, ds - \int_0^\epsilon \tilde{L}_i(\tilde{\gamma}^0(s), \dot{\tilde{\gamma}}^0(s))\, ds \\
+ \int_{t-\epsilon}^t \tilde{L}(\tilde{\gamma}^1_{x',y'}(s), \dot{\tilde{\gamma}}^1_{x',y'}(s))\, ds - \int_{t-\epsilon}^t L(\tilde{\gamma}^1(s), \dot{\tilde{\gamma}}^1(s))\, ds. \qquad (*)
\end{aligned}
$$

By the choice of B, we have $\|\dot{\tilde{\gamma}}(s)\| \leq B$, for each $s \in [0, \epsilon]$. By the definition of $\tilde{\gamma}^0_{x',y'}$ and the fact that $\|\tilde{x}' - \tilde{x}\| \leq 1$, we obtain

$$
\begin{aligned}
\forall s \in [0, \epsilon],\ \|\dot{\tilde{\gamma}}^0_{x',y'}(s)\| &\leq \frac{\|\tilde{x}' - \tilde{x}\|}{\epsilon} + B \\
&\leq B + \epsilon^{-1}.
\end{aligned}
$$

Since $\tilde{L}_i$ is C^1 and the set

$$E_{B,\epsilon} = \{(z, v) \in \bar{B}(0,3) \times \mathbb{R}^k \mid \|v\| \leq B + \epsilon^{-1}\}$$

is compact, we see that there exists a constant C_i, which depends only on the restriction of the derivative of $\tilde{L}_i$ on this set $E_{B,\epsilon}$, such that

$$
\begin{gathered}
\forall (z, v), (z', v') \in E_{B,\epsilon}, \\
|\tilde{L}_i(z, v) - \tilde{L}_i(z', v')| \leq C_i \max(\|z - z'\|, \|v - v'\|).
\end{gathered}
$$

Since the two points $(\tilde\gamma^0_{x',y'}(s),\dot{\tilde\gamma}^0_{x',y'}(s)),(\tilde\gamma^0(s),\dot{\tilde\gamma}^0(s))$ are in $E_{B,\epsilon}$, and $\tilde\gamma^0_{x',y'}(s)-\tilde\gamma^0(s)=\frac{\epsilon-s}{\epsilon}(\tilde x'-\tilde x)$, we find

$$\begin{aligned}&|\tilde L_i(\tilde\gamma^0_{x',y'}(s),\dot{\tilde\gamma}^0_{x',y'}(s))-\tilde L(\tilde\gamma^0(s),\dot{\tilde\gamma}^0(s))|\\&\qquad\le C_i\max\big[\|\frac{\epsilon-s}{\epsilon}(\tilde y-\tilde x)\|,\|\frac{1}{\epsilon}(\tilde y-\tilde x)\|\big]\\&\qquad\le C_i\max(1,\frac{1}{\epsilon})\|\tilde y-\tilde x\|.\end{aligned}$$

By integration on the interval $[0,\epsilon]$, it follows that

$$\begin{aligned}\int_0^\epsilon \tilde L_i(\tilde\gamma^0_{x',y'}(s),\dot{\tilde\gamma}^0_{x',y'}(s))\,ds-\int_0^\epsilon \tilde L_i(\tilde\gamma^0(s),\dot{\tilde\gamma}^0(s))\,ds&\\ \le C_i\max(\epsilon,1)\|\tilde y-\tilde x\|.&\end{aligned}$$

Since φ_i is a diffeomorphism of class C^∞, its inverse is Lipschitzian on the compact subset $\varphi_i\big(\bar B(0,3)\big)$. We then see that there exists a constant $\tilde C_i$, independent of x,x',y,y' and $t\ge t_0$, and such that

$$\begin{aligned}&x\in\varphi_i\big(\mathring{B}(0,1)\big)\text{ and }d(x',x)\le\eta\Rightarrow\\&\int_0^\epsilon \tilde L_i(\tilde\gamma^0_{x',y'}(s),\dot{\tilde\gamma}^0_{x',y'}(s))\,ds-\int_0^\epsilon \tilde L_i(\tilde\gamma^0(s),\dot{\tilde\gamma}^0(s))\,ds\le\tilde C_i d(x',x).\end{aligned}$$

In the same way we can prove the existence of a constant $\tilde C'_j$, independent of x,x',y,y' and $t\ge t_0$, and such that

$$\begin{aligned}&y\in\varphi_j\big(\mathring{B}(0,1)\big)\text{ and }d(y',y)\le\eta\Rightarrow\\&\int_{t-\epsilon}^t \tilde L_j(\tilde\gamma^1_{x',y'}(s),\dot{\tilde\gamma}^1_{x',y'}(s))\,ds-\int_{t-\epsilon}^t \tilde L_j(\tilde\gamma^1(s),\dot{\tilde\gamma}^1(s))\,ds\le\tilde C'_j d(y',y).\end{aligned}$$

Therefore by the inequality (*) above, we obtain

$$\begin{aligned}&x\in\varphi_i\big(\mathring{B}(0,1)\big),y\in\varphi_j\big(\mathring{B}(0,1)\big),d(x',x)\le\eta\text{ and }d(y',y)\le\eta\Rightarrow\\&\qquad h_t(x',y')-h_t(x,y)\le\tilde C_i d(x',x)+\tilde C'_j d(y',y).\end{aligned}$$

Setting $\kappa_{t_0}=\max_{i=1}^p\max(\tilde C_i,\tilde C'_i)$, we find that we have

$$\begin{aligned}&d(x',x)\le\eta,d(y',y)\le\eta\text{ and }t\ge t_0\Rightarrow\\&\qquad h_t(x',y')-h_t(x,y)\le\kappa_{t_0}[d(x',x)+d(y',y)].\end{aligned}$$

If x, x', y, y' are arbitrary points, and $\gamma_0 : [0,1] \to M, \gamma_1 : [0,1] \to M$ are geodesics of length respectively $d(x,y)$ and $d(x',y')$, parameterized proportionally arclength and connecting respectively x to y and x' to y' , we can find a finite sequence $t_0 = 0 \leq t_1 \leq \cdots \leq t_\ell = 1$ such that $d(\gamma^0(t_{i+1}), \gamma^0(t_i)) \leq \eta$ and $d(\gamma^1(t_{i+1}), \gamma^1(t_i)) \leq \eta$. Applying what we did above, we obtain

$$\begin{aligned}\forall i \in \{0,1,\ldots,\ell-1\},\ & h_t(\gamma^0(t_{i+1}), \gamma^1(t_{i+1})) - h_t(\gamma^0(t_i), \gamma^1(t_i)) \\ & \leq \kappa_{t_0}[d(\gamma^0(t_{i+1}), \gamma^0(t_i)) + d(\gamma^1(t_{i+1}), \gamma^1(t_i))].\end{aligned}$$

Adding these inequalities, we find

$$h_t(x',y') - h_t(x,y) \leq \kappa_{t_0}[d(x',x) + d(y'y)].$$

We finish the proof by exchanging the roles of (x,y) and (x',y') $\square$

The following theorem is a consequence of Corollary 4.4.5 and

Theorem 4.4.6. *Let L be a C^r Tonelli Lagrangian, with $r \geq 2$, on the compact connected manifold M. Given $a, b \in \mathbb{R}$ with $a < b$, call $\mathbb{M}_{a,b}$ the set of curves $\gamma : [a,b] \to M$ which are minimizers for the class of continuous piecewise C^1 curves. Then $\mathbb{M}_{a,b}$ is a compact subset of $\mathcal{C}^r([a,b], M)$ for the C^r topology.*

Moreover, For every $t_0 \in [a,b]$ the map $\mathbb{M}_{a,b} \to TM, \gamma \mapsto (\gamma(t_0), \dot\gamma(t_0))$ is a homeomorphism on its image (which is therefore a compact subset of TM.)

4.5 The Lax-Oleinik Semi-group.

We call $\mathcal{F}(M, [-\infty, +\infty])$ the set of arbitrary functions from the manifold M to the set $[-\infty, +\infty]$ of extended real numbers. We will also use the notation $\mathcal{F}(M, \mathbb{R})$ for the set of arbitrary functions $M :\to \mathbb{R}$.

4.6 The Lax-Oleinik semi-group

We introduce a semi-group of non-linear operators $(T_t^-)_{t\geq 0}$ from $\mathcal{F}(M, [-\infty, +\infty])$ into itself. This semigroup is well-known in PDE and in Calculus of Variations, it is called the Lax-Oleinik semi-group. Again it has been rediscovered many times in different forms; For the author, it came as a natural by-product of the proof of the hamilton-Jacobi Theorem for C^1 functions, see the proof of Theorem 4.1.1 and the discussion in section 1.

Definition 4.6.1 (Lax-Oleinik semi-group)**.** Fix $u \in \mathcal{F}^0(M, [-\infty, +\infty])$ and $t > 0$. The function $T_t^- u : M \to [-\infty, +\infty]$ is defined at $x \in M$ by

$$T_t^- u(x) = \inf_\gamma \{u(\gamma(0)) + \int_0^t L(\gamma(s), \dot\gamma(s))\, ds\},$$

where the infimum is taken over all the absolutely continuous curves $\gamma : [0, t] \to M$ such that $\gamma(t) = x$. By the definition of the minimal action h_t, see Definition 4.4.1, for $t > 0$, we have

$$T_t^- u(x) = \inf_{y \in M} u(y) + h_t(y, x).$$

We will also set $T_0^- u = u$. The family of maps $T_t^- : \mathcal{F}(M, [-\infty, +\infty]) \to \mathcal{F}(M, [-\infty, +\infty]), t \in [0, +\infty[$ is called the Lax-Oleinik semi-group.

Here are some properties of T_t^- on $\mathcal{F}(M, [-\infty, +\infty])$.

Proposition 4.6.2. *Consider $u \in \mathcal{F}(M, [-\infty, +\infty])$.*
(1) *For $x \in M$ and $t > 0$, we have*

$$\inf_M u + t \inf_{TM} L \leq T_t u(x) \leq \inf_M u + \max_{M \times M} h_t.$$

Since $\inf_{TM} L > -\infty$ by the superlinearity of L and h_t is continuous on the compact space $M \times M$, we obtain that the following properties are equivalent:

(a) *there exists a $t > 0$ and and an $x \in M$ such that $T_t^- u(x)$ is finite;*

(b) *we have $\inf_M u \in] - \infty, +\infty[$;*

(c) *for every $t > 0$, the function $T_t^- u$ is finite, valued.*

(2) **(Semigroup Property)** *We have $T_{t+t'}^- = T_t^- \circ T_{t'}^-$, for each $t, t' \leq 0$.*
(3) *for every $c \in \mathbb{R}$, we have $T_t^-(c + u) = c + T_t^- u$.*

(4) **(Inf Commutativity)** *If $u_i, i \in I$ is a family of functions in $u \in \mathcal{F}(M, [-\infty, +\infty])$, we have*

$$T_t^-(\inf_{i\in I} u_i) = \inf_{i\in I} T_t^-(u_i).$$

(5) **(Monotony)** *For each $u, v \in \mathcal{F}(M, [-\infty, +\infty]))$ and all $t > 0$, we have*

$$u \leq v \Rightarrow T_t^- u \leq T_t^- v.$$

(6) *If $c \in \mathbb{R}$, the function $u \in \mathcal{F}(M, [-\infty, +\infty])$ satisfies $u \leq T_t^- u + ct$ if and only if one of the following three things happens:*

(i) *the function u is identically $-\infty$;*

(ii) *the function u is identically $+\infty$;*

(iii) *the function u is finite everywhere and $u \prec L + c$.*

(7) *Suppose that $c \in \mathbb{R}$, and $u : M \to \mathbb{R}$ are such that $u \prec L + c$, then, for every $t \geq 0$, the function $T_t^- u$ is finite valued, and $T_t^- u \prec L + c$.*

Proof. To prove assertion (1), we juste notice that $T_t^- u(x) = \inf_{y\in M} u(y) + h_t(y, x) \leq \inf_{y\in M} u(y) + \max_{M\times M} h_t = \inf_M u + \max_{M\times M} h_t$. Moreover by part (1) of Proposition 5.3.2, we have $h_t(y, x) \geq t \inf_{TM} L$, from which it follows that

$$\begin{aligned} T_t^- u(x) &\geq \inf_{y\in M} u(y) + t \inf_{TM} L \\ &= t \inf_{TM} L + \inf_M u. \end{aligned}$$

The rest of assertion (1) follows easily from this double inequality.

Assertion (2) follows from part (2) of Proposition 5.3.2, which states that $h_{t'+t}(y,x) = \inf_{z\in M} h_{t'}(y,z) + h_t(z,x)$. Therefore

$$\begin{aligned}
T^-_{t+t'}u(x) &= \inf_{y\in M} u(y) + h_{t'+t}(x,y)\\
&= \inf_{y\in M}[u(y) + \inf_{z\in M} h_{t'}y,z) + h_t(z,x)]\\
&= \inf_{y\in M}\inf_{z\in M}[u(y) + h_{t'}(y,z) + h_t(z,x)]\\
&= \inf_{z\in M}\inf_{y\in M}[u(y) + h_{t'}(y,z) + h_t(z,x)]\\
&= \inf_{z\in M}[\inf_{y\in M}[u(y) + h_{t'}(y,z)] + h_t(z,x)\\
&= \inf_{z\in M} T^-_{t'}u(z) + h_t(z,x)\\
&= T^-_t[T^-_{t'}u].
\end{aligned}$$

Assertion (3) is obvious from the definition of T^-_t.

For assertion (4), we notice that

$$\begin{aligned}
T^-_t(\inf_{i\in I} u_i)(x) &= \inf_{y\in M}\inf_{i\in I} u_i(y) + h_t(y,x)\\
&= \inf_{i\in I}\inf_{y\in M} u_i(y) + h_t(y,x)\\
&= \inf_{i\in I} T^-_t(u_i)(x).
\end{aligned}$$

Assertion (5) is also an immediate consequence of the definition of T^-_t. It also easily follows from assertion (4), since we have $u = \inf(u,v)$, which yields $T^-_t(u) = T^-_t(\inf(u,v)) = \inf(T^-_t(u), T^-_t(v)) \leq T^-_t(v)$.

For assertion (6), assume that $u \leq T^-_t u + ct$ for every t. Obviously, by assertion (1), if $\inf_M u = -\infty$, then $T^-_t u \equiv -\infty$ hence $u \leq T^-_t u + ct$ is also $\equiv -\infty$. If $\inf_M u = +\infty$, of course $u \equiv +\infty$. In the remaining case $\inf_M u \in]-\infty, +\infty[$, we obtain from (1) above that $T^-_t u$ is finite everywhere therefore u, which satisfies $u \leq T^-_t u + ct$ and $u \geq \inf_M u > -\infty$, is also finite valued. The condition $u \leq T^-_t u + ct$ yields $u(x) \leq \inf_{y\in M} u(y) + h_t(y,x) + ct$. Therefore $u(x) \leq u(y) + h_t(y,x) + ct$, for every $x, y \in M$ and $t > 0$. Since u is finite valued this is equivalent to $u(x) - u(y) \leq h_t(y,x) + ct$, for every $x, y \in M$ and $t > 0$. By Assertion (3) of Proposition 5.3.2 this is equivalent to $u \prec L + c$. Note that reversing the reasoning just done we can show that if u is finite valued and $u \prec L + c$ then $u \leq T^-_t u + ct$, for every $t \geq 0$. Moreover, if $u \equiv -\infty$ (resp.

$u \equiv +\infty$) we have $T_t^-(u) \equiv -\infty$ (resp. $T_t^-(u) \equiv +\infty$) which of course yields $u = T_t^- u + ct$.

For assertion (7), we first observe that by part (iii) of Proposition 4.2.1, then function u is continuous (it is even Lipschitz) on the compact manifold M. Therefore it is bounded from below and by Assertion (1, the function $T_t^- u$ is finite everywhere. By Assertion (6), we have

$$u \leq T_{t'}^- u + ct',$$

for every $t' \geq 0$. Using Assertions (2), (3), and (5), we get

$$\begin{aligned} T_t^- u &\leq T_t^-[T_{t'}^- u + ct'] \\ &= T_t^-[T_{t'}^- u] + ct' \\ &= T_{t+t'}^- u \qquad\qquad\qquad = T_{t'}^-[T_t^- u] + ct' \end{aligned}$$

Therefore $T_t^- u \leq T_{t'}^-[T_t^- u] + ct'$, for every $t' \geq 0$. Since $T_t^- u$ is finite, from Assertion (6), we get $T_t^- u \prec L + c$. □

Let us now introduce the space $\mathcal{B}(M, \to \mathbb{R})$ of bounded functions $u : M \to \mathbb{R}$. As usual, we endow this space of the norm $\|\cdot\|$ defined by

$$\|u\|_\infty = \sup_{x \in M} |u(x)\|.$$

By assertion (1) of Proposition 4.6.2, if $u \in \mathcal{F}(M, \mathbb{R})$ satisfies $\inf_M u > -\infty$, then $T_t^- u \in \mathcal{B}(M, \mathbb{R})$. Fleming's Lemma 4.4.3 yields the following much stronger property on the Lax-Oleinik semi-group.

Lemma 4.6.3. *For each $t_0 > 0$, there exists a constant κ_{t_0} such that for every $u \in \tilde{\mathcal{F}}(M, \mathbb{R})$ with $\inf_M u > -\infty$, and every $t \geq t_0$ the function$T_t^- u : M \to \mathbb{R}$ is κ_{t_0}-Lipschitzian.*

In particular for every $t > 0$, we have $T_t^-(\mathcal{B}(M, \mathbb{R}) \subset \mathcal{C}^0(M, \mathbb{R})$.

Proof. By Fleming's Lemma 4.4.3, we can find a constant κ_{t_0} such that for every $t \geq t_0$, the function $h_t : M \times M : to\mathbb{R}$ is Lipschitz with Lipschitz constant κ_{t_0}. If follows that for any $u\ in\mathcal{C}^0(M, \mathbb{R})$ and any $t \geq t_0$, the family of function $u(y) + h(y, \cdot), y \in M$ is equi-Lipschitzian with Lipschitz constant κ_{t_0}. Since, by the condition $\inf_M u > -\infty$, its infimum $T_t^- u(x) = \inf_{y \in M} u(y) + h(y, x)$ is finite everywhere, it is also Lipschitz with Lipschitz constant κ_{t_0}. □

Before giving further properties of the semi-group T_t^-, we recall the following well-known definition

Definition 4.6.4 (Non-expansive Map)**.** A map $\varphi : X \to Y$, between the metric spaces X and Y, is said to be non-expansive if it is Lipschitzian with a Lipschitz constant ≤ 1.

Proposition 4.6.5 (Non-expansiveness of the Lax Oleinik semi–group)**.** *The maps T_t^- : are non-expansive for the norm $\|\cdot\|_\infty$, i.e.*

$$\forall u, v \in \mathcal{B}(M, \mathbb{R}), \forall t \geq 0,\ \|T_t^- u - T_t^- v\|_\infty \leq \|u - v\|_\infty.$$

Proof. If $u, v \in \mathcal{B}(M, \mathbb{R})$, we have

$$-\|u - v\|_\infty + v \leq u \leq \|u - v\|_\infty + v.$$

By parts (5) and (3) of Proposition 4.6.2, we get

$$-\|u - v\|_\infty + T_t^- v \leq T_t^- u \leq \|u - v\|_\infty + T_t^- v.$$

This clearly implies $\|T_t^- u - T_t^- v\|_\infty \leq \|u - v\|_\infty$. □

We now turn to the properties of the semi-group on the space $\mathcal{C}^0(M, \mathbb{R})$ of continuous functions, endowed with the topology of uniform convergence, i.e. the topology induced by the norm $\|\cdot\|$.

Proposition 4.6.6. *The semi-group T_t^- sends $\mathcal{C}^0(M, \mathbb{R})$ to itself. It satisfies the following properties:*

(1) *For each $u \in \mathcal{C}^0(M, \mathbb{R})$, we have $\lim_{t\to 0} T_t^- u = u$.*

(2) *For each $u \in \mathcal{C}^0(M, \mathbb{R})$, the map $t \mapsto T_t^- u$ is uniformly continuous.*

(3) *For each $u \in \mathcal{C}^0(M, \mathbb{R})$, the function $(t, x) \mapsto T_t^- u(x)$ is continuous on $[0, +\infty[\times M$ and locally Lipschitz on $]0, +\infty[\times M$. In fact, for each $t_0 > 0$, the family of functions $(t, x) \mapsto T_t^- u(x), u \in \mathcal{C}^0(M, \mathbb{R})$, is equi-Lipschitzian on $[t_0, +\infty[\times M$.*

(4) *For each $u \in \mathcal{C}^0(M, \mathbb{R})$, each $x \in M$, and each $t > 0$ we can find a a continuous piecewise C^1 curve $\gamma_{x,t} : [0, t] \to M$ with $\gamma_{x,t}(t) = x$ and*

$$T_t^-(x) = u(\gamma_{x,t}(0)) + h_t(\gamma_{x,t}(0), x)) = u(\gamma_{x,t}(0)) + \int_0^t L(\gamma_{x,t}(s), \dot\gamma_{x,t}(s))\, ds.$$

Proof. As the C^1 maps form a dense subset of $\mathcal{C}^0(M,\mathbb{R})$ in the topology of uniform convergence, it is not difficult to see, using Proposition 4.6.5 above, that it is enough to show property (1) when u is Lipschitz. We denote by K the Lipschitz constant of u. By the compactness of M and the superlinearity of L, there is a constant C_K such that

$$\forall (x,v) \in TM,\ L(x,v) \geq K\|v\|_x + C_K.$$

It follows that for every curve $\gamma : [0,t] \to M$, we have

$$\int_0^t L(\gamma(s),\dot\gamma(s))\,ds \geq Kd(\gamma(0),\gamma(t)) + C_K t.$$

Since the Lipschitz constant of u is K, we conclude that

$$\int_0^t L(\gamma(s),\dot\gamma(s))\,ds + u(\gamma(0)) \geq u(\gamma(t)) + C_K t,$$

which gives

$$T_t^- u(x) \geq u(x) + C_K t.$$

Moreover, using the constant curve $\gamma_x : [0,t] \to M, s \mapsto x$, we obtain

$$T_t^- u(x) \leq u(x) + L(x,0)t.$$

Finally, if we set $A_0 = \max_{x\in M} L(x,0)$, we have obtained

$$\|T_t^- u - u\|_\infty \leq t\max(C_K\, A_0),$$

which does tend to 0, when t tends to 0.

To show (2), we notice that by the semi-group property of T_t^-, see part (2) of Proposition 4.6.2, we have

$$\|T_{t'}^- u - T_t^- u\|_\infty \leq \|T_{|t'-t|}^- u - u\|_\infty,$$

and we apply (1) above.

To prove (3), we remark that the continuity of $(t,x) \mapsto T_t^- u(x)$ follows from (1) and the fact that $T_t^- u$ is continuous for ever $t \geq 0$. To prove the equi-Lipschitzianity, we fix $t_0 > 0$. By Lemma 4.6.3, we know that there exists a finite constant $K(t_0)$ such that $T_t^- u$ is Lipschitz with Lipschitz constant $\leq K(t_0)$, for each $u \in \mathcal{C}^0(M,\mathbb{R})$

and each $t \geq t_0$. It follows from the semi-group property and the proof of (1) that

$$\|T_{t'}^- u - T_t^- u\|_\infty \leq (t' - t) \max(C_{K(t_0)}, A_0),$$

for all $t' \geq t \geq t_0$. It is then easy to check that $(t, x) \mapsto T_t^- u(x)$ is Lipschitz on $[t_0, +\infty[\times M$, with Lipschitz constant $\leq \max(C_{K(t_0)}, A_0)+ K(t_0)$, for anyone of the standard metrics on the product $\mathbb{R} \times M$. To prove (4), we recall that $T_t^-(x) = \inf_{y\in M} u(y) + h_t(y, x)$. since the function $y \mapsto u(y)+h_t(y, x)$ is continuous on the compact space M, we can find $y_x \in M$ such that $T_t^-(x) = u(y_x) + h_t(y_x, x)$. We can apply part (5) of Proposition 5.3.2 to find a continuous piecewise C^1 curve $\gamma_{x,t} : [0, t] \to M$ with $\gamma_{x,t}(0) = y_x, \gamma_{x,t}(t) = x$ and $h_t(y_x, x) = \int_0^t L(\gamma_{x,t}(s), \dot\gamma_{x,t}(s))\, ds$. Therefore $T_t^-(x) = u(y_x) + h_t(y_x, x)) = u(y_x) + \int_0^t L(\gamma_{x,t}(s), \dot\gamma_{x,t}(s))\, ds$. □

We now give the connection of the semi-group T_t^- with weal KAM solutions of the negative type.

Proposition 4.6.7. *Suppose that $u : M \to \mathbb{R}$ s a function and $c \in \mathbb{R}$. We have $T_t^- u + ct = u$, for each $t \in [0, +\infty[$, if and only if u is a negative weak KAM solution, i.e.we have*

(i) $u \prec L + c$;

(ii) *for each $x \in M$, there exists a (u, L, c)-calibrated curve $\gamma_-^x :]-\infty, 0] \to M$ such that $\gamma_-^x(0) = x$.*

Proof. We suppose that $T_t^- u + ct = u$, for each $t \in [0, +\infty[$. In particular, we have By part (6) of Proposition 4.6.2 above, since u is finite-valued, we obtain $u \prec L+c$. In particular, the function u is continuous. It remains to show the existence of $\gamma_-^x :]-\infty, 0] \to M$, for a given $x \in M$. We already know by part (4) of Proposition 4.6.6 that, for each $t > 0$, there exists a continuous piecewise C^1 curve $\gamma_t : [0, t] \to M$, with $\gamma_t(t) = x$ and

$$u_-(x)-ct = T_t u_-(x) = u_-(\gamma_t(0))+h_t(\gamma_t(0), \gamma_t(t)) = u_-(\gamma_t(0))+\int_0^t L(\gamma_t(s), \dot\gamma_t(s))\, ds.$$

We set $\bar\gamma_t(s) = \gamma_t(s+t)$, this curve is parametrized by the interval $[-t, 0]$, is equal to x at 0, and satisfies

$$u_-(\bar\gamma_t(0))-u_-(\bar\gamma_t(-t)) = h_t(\bar\gamma_t(-t), \bar\gamma_t(0))+ct \int_{-t}^0 L(\bar\gamma_t(s), \dot{\bar\gamma}_t(s))\, ds+ct.$$

It follows from Proposition **??** above that $\bar\gamma_t$ is (u, L, c)-calibrated. In particular, the curve $\bar\gamma_t$, and we have

$$\forall t' \in [-t, 0], u_-(x) - u_-(\bar\gamma_t(t')) = -ct' + \int_{t'}^0 L(\bar\gamma_t(s), \dot{\bar\gamma}_t(s))\, ds. \quad (*)$$

As the $\bar\gamma_t$ are minimizing extremal curves, by the a priori compactness given by corollary 4.4.5, there exists a compact subset $K_1 \subset TM$, such that

$$\forall t \geq 1, \forall s \in [-t, 0], (\bar\gamma_t(s), \dot{\bar\gamma}_t(s)) \in K_1.$$

Since the $\bar\gamma_t$ are extremal curves, we have $(\bar\gamma_t(s), \dot{\bar\gamma}_t(s)) = \phi_s(\bar\gamma_t(0), \dot{\bar\gamma}_t(0))$. The points $(\bar\gamma_t(0), \dot{\bar\gamma}_t(0))$ are all in the compact subset K_1, we can find a sequence $t_n \nearrow +\infty$ such that the sequence $(\bar\gamma_{t_n}(0), \dot{\bar\gamma}_{t_n}(0)) = (x, \dot{\bar\gamma}_{t_n}(0))$ tends to (x, v_∞), when $n \to +\infty$. The negative orbit $\phi_s(x, v_\infty), s \leq 0$ is of the form $(\gamma_-^x(s), \dot\gamma_-^x(s))$, where $\gamma_-^x :]-\infty, 0] \to M$ is an extremal curve with $\gamma_x(0) = x$. If $t' \in]-\infty, 0]$ is fixed, for n large enough, the function $s \mapsto (\bar\gamma_{t_n}(s), \dot{\bar\gamma}_{t_n}(s)) = \phi_s(\bar\gamma_{t_n}(0), \dot{\bar\gamma}_{t_n}(0))$ is defined on $[t', 0]$, and, by continuity of the Euler-Lagrange flow, this sequence converges uniformly on the compact interval $[t', 0]$ to the map $s \mapsto \phi_s(x, v_\infty) = (\gamma_x(s), \dot\gamma_x(s))$. We can then pass to the limit in the equality $(*)$ to obtain

$$u_-(x) - u_-(\gamma_-(t')) = -ct' + \int_{t'}^0 L(\gamma_-^x(s), \dot\gamma_-^x(s))\, ds.$$

Conversely, let us suppose that $u \prec L+c$ and that, for each $x \in M$, there exists a curve $\gamma_-^x :]-\infty, 0] \to M$, with $\gamma_-^x(0) = x$, and such that for each $t \in [0, +\infty[$

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

Let us show that $T_t^- u + ct = u$, for each $t \in [0, +\infty[$. If $x \in M$ and $t > 0$, we define the curve $\gamma : [0, t] \to M$ by $\gamma(s) = \gamma_-^x(s - t)$. It is not difficult to see that $\gamma(t) = x$ and that

$$u_-(x) = ct + \int_0^t L(\gamma(s), \dot\gamma(s))\, ds + u_-(\gamma(0)).$$

It follows that $T_t^- u(x) + ct \leq u(x)$ and thus $T_t^- u + ct \leq u$. The inequality $u \leq T_t^- u + ct$ results from $u \prec L + c$. $\square$

4.7 Existence of Negative Weak KAM Solutions

Our goal in this section is to prove the existence of negative weak KAM solutions.

Theorem 4.7.1 (Weak KAM)**.** *There exists a function $u_- : M \to \mathbb{R}$ which is a negative weak KAM solution with constant $c[0]$.*

Note that we already know by Corollary 4.3.7 that a weak KAM solution can only have c[0] as a constant. By Proposition the weak KAM theorem above is equivalent to

Theorem 4.7.2. *There exists a function $u_- : M \to \mathbb{R}$ such that $T_t^- u_- + tc[0] = u_-$, for each $t \in [0, +\infty[$.*

We will give two proofs of this theorem.

For the first proof we need some lemmas.

Lemma 4.7.3. *Let $u : M \to \mathbb{R}$ be a function (not necessarily continuous or even bounded), and $c \in \mathbb{R}$. If the function $\tilde{u}$ defined on M by*

$$\tilde{u}(x) = \inf_{t \geq 0} T_t^- u(x) + ct$$

is finite at some point, then it is finite everywhere and $\tilde{u} \prec L + c$.

Proof. We use parts (4), (3), (2) of Proposition 4.6.2 to obtain

$$\begin{aligned} T_{t'}^- \tilde{u} &= T_{t'}^- [\inf_{t \geq 0} T_t^- u + ct] \\ &= \inf_{t \geq 0} T_{t'}^- [T_t^- u + ct] \\ &= \inf_{t \geq 0} T_{t'}^- [T_t^- u] + ct \\ &= \inf_{t \geq 0} T_{t'+t}^- T_t^- u + ct. \end{aligned}$$

Therefore

$$\begin{aligned} T_{t'}^- \tilde{u} + ct' &= ct' + \inf_{t \geq 0} T_{t'+t}^- T_t^- u + ct \\ &= \inf_{t \geq 0} T_{t'+t}^- T_t^- u + ct + ct' \\ &= \inf_{t \geq 0} T_{t'+t}^- T_t^- u + c(t' + t) \\ &geq \inf_{t \geq 0} T_t^- u + ct = \tilde{u}. \end{aligned}$$

Since $\tilde{u}$ is finite at some point it follows from part (6) of Proposition 4.6.2 that it is finite everywhere and $u \prec L + c$. □

Lemma 4.7.4. *Let $u : M \to \mathbb{R}$ be a function with $\inf_M u > -\infty$, but not necessarily continuous or even bounded. Suppose $t > 0$ and $c \in \mathbb{R}$ are such that $u \leq T_t^- u + c$, then $c/t \geq c[0]$.*

Proof. Using parts (2), (3) and (5) of Proposition 4.6.2, we obtain for $n \in \mathbb{N}$

$$u \leq T_t^- u + c \leq T_{2t}^- u + 2c \leq \cdots \leq T_{nt}^- u + nc.$$

This implies that for $s \geq 0$

$$T_s^- u \leq T_{nt+s}^- u + nc.$$

If we set $\tilde{c} = c/t$, this yields

$$T_s^- u + s\tilde{c} \leq T_{nt+s}^- u + (nt + s)\tilde{c},$$

for all $s \geq 0$, and all $n \in \mathbb{N}$. In particular, we have

$$\tilde{u} = \inf_{s \geq 0} T_s^- u + s\tilde{c} = \inf_{0 \leq s \leq t} T_s^- u + s\tilde{c}.$$

We now note that by part (1) of Proposition 4.6.2 we have

$$T_s u \geq \inf_M u + s \inf_{TM} L \geq \inf_M u - s|\inf_{TM} L|.$$

Therefore

$$\begin{aligned}
\tilde{u} &= \inf_{0 \leq s \leq t} T_s^- u + s\tilde{c} \\
&\geq \inf_{0 \leq s \leq t} \inf_M u - s|\inf_{TM} L| + s\tilde{c} \\
&\geq \inf_M u - s(|\inf_{TM} L| + |\tilde{c}| > -\infty.
\end{aligned}$$

Hence we can apply Lemma 4.7.3 to obtain that $\tilde{u} \prec L + \tilde{c}$. Hence $\tilde{c} = c/t \geq c[0]$. □

Lemma 4.7.5. *Let $u : M \to \mathbb{R}$ be dominated by $L + c[0]$ on the compact manifold M. then for every $t \geq 0$, we can find a point $x_t \in M$ such that $u(x_t) = T_t^- u(x_t) + tc[0]$.*

In particular, we have $\sup_{t \geq 0} \|T_t^- u + tc[0]\|_\infty < +\infty$.

Proof. Let us observe that, part (iii) of Proposition 4.2.1 and part (7) of Proposition 4.6.2, the family of functions $T_t^- u, t \geq 0$, is equi-Lipschitzian. We call $L < +\infty$ a common Lipschitz constant for $T_t^- u, t \geq 0$.

To prove the first part we can assume $t > 0$. Suppose that no such x_t exists, since $u \leq T_t^- u + tc[0]$, we obtain that $T_t^- u + tc[0] - u > 0$ everywhere. Since both u and $T_t^- u$ are continuous, the compactness of M now implies that there exists $\epsilon > 0$ such that $T_t^- u + tc[0] - u \geq \epsilon$ everywhere. This can be rewritten as $u \leq T_t^- u + tc[0] - \epsilon$. By Lemma 4.7.4 above, this implies $(tc[0] - \epsilon)/t \geq c[0]$. Which is obviously false. This proves the existence of x_t.

For each $t \geq 0$ the function $T_t^- u + tc[0] - u$ is Lipschitzian with Lipschitz constant $\geq 2L$. Since it vanishes at some point $x_t \in M$, we have for every $x \in M$

$$\begin{aligned} |T_t^- u(x) + tc[0] - u(x)| &= |(T_t^- u + tc[0] - u(x)) - (T_t^- u(x_t) + tc[0] - u(x))| \\ &\leq 2Ld(x, x_t) \\ &\leq 2L \operatorname{diam}(M). \end{aligned}$$

Therefore we, obtain $\sup_{t \geq 0} \|T_t^- u + tc[0]\|_\infty \leq \|u\|_\infty + 2L \operatorname{diam}(M) < +\infty$. □

We can now prove the following theorem which yields immediately the weak KAM Theorem 4.7.1.

Theorem 4.7.6. *Let L be a Tonelli Lagrangian on the compact connected manifold M. Suppose uIf $u : M \to \mathbb{R}$ is dominated by $L + c[0]$. Then $T_t^- u + tc[0]$ converges uniformly to a continuous function $u_- : M \to \mathbb{R}$ which is a negative weak KAM solution.*

Proof. We note that by parts (6), (5) and (2) of Proposition 4.6.2 we have for $s, t \geq 0$

$$\begin{aligned} T_s^- u &\leq T_s^- (T_t^- u + tc[0]) \\ &= T_{t+s}^- u + tc[0]. \end{aligned}$$

Hence $T_s^- u + sc[0] \leq T_{t+s}^- u + (t+s)c[0]$, this implies that $T_t^- u + tc[0] \leq T_{t'}^- u + t'c[0]$, for every $t, t' \geq 0$, with $t \leq t'$. Since by Lemma 4.7.5, the family $T_t^- u + tc[0], t \geq 0$ is equibounded, it follows that the point-wise limit $u_-(x) = \lim_{t\to+\infty} T_t^- u(x) + tc[0]$ exists everywhere and is finite. We now remark as at the beginning of the proof of Lemma 4.7.5 above that the family $T_t^- u + tc[0, t \geq 0$ is equicontinuous, to conclude that the limit $u_- = \lim_{t\to+\infty} T_t^- u + tc[0]$ is uniform.

It remains to prove that $u_- : M \to \mathbb{R}$ which is a negative weak KAM solution. By Proposition 4.6.7, we have to check that $T_s^- u + sc[0] = u$, for each $s \geq$. Using the non-expansiveness of the Lax Oleinik semi-group Proposition 4.6.5, we obtain $T_s^- u_- = \lim_{t\to+\infty} T_s^-[T_t^- u + tc[0]] = \lim_{t\to+\infty} T_{t+s}^- u + tc[0]$. Therefore $T_s^- u_- + sc[0] = \lim_{t\to+\infty} T_{t+s}^- u + (t+s)c[0] = u_-$. $\square$

Before giving a second proof of the weak KAM Theorem 4.7.1, we will need to recall some fixed point theorems. We leave this as an exercise, see also [GK90, Theorem 3.1, page 28].

Exercise 4.7.7. *1) Let E be a normed space and $K \subset E$ a compact convex subset. We suppose that the map $\varphi : K \to K$ is non-expansive. Show that φ has a fixed point. [Hint: Reduce first to the case when $0 \in K$, and then consider $x \mapsto \lambda\varphi(x)$, with $\lambda \in]0, 1[$.]*

2) Let E be a Banach space and let $C \subset E$ be a compact subset. Show that the closed convex envelope of C in E is itself compact. [Hint: It is enough to show that, for each $\epsilon > 0$, we can cover the closed convex envelope of C by a finite number of balls of radius ϵ.]

3) Let E be a Banach space. If $\varphi : E \to E$ is a non-expansive map such that $\varphi(E)$ has a relatively compact image in E, then the map φ admits a fixed point. [Hint: Take a compact convex subset containing the image of φ.]

4) Let E be a Banach space and $\varphi_t : E \to E$ be a family of maps defined for $t \in [0, \infty[$. We suppose that the following conditions are satisfied

- *For each $t, t' \in [0, \infty[$, we have $\varphi_{t+t'} = \varphi_t \circ \varphi_{t'}$.*
- *For each $t \in [0, \infty[$, the map φ_t is non-expansive.*
- *For each $t > 0$, the image $\varphi_t(E)$ is relatively compact in E.*
- *For each $x \in E$, the map $t \mapsto \varphi_t(x)$ is continuous on $[0, +\infty[$.*

Show then that the maps φ_t *have a common fixed point. [Hint: A fixed point of* φ_t *is a fixed point of* φ_{kt} *for each integer* $k \geq 1$*. Show then that the maps* $\varphi_{1/2^n}$*, with* $n \in \mathbb{N}$*, admit a common fixed point.]*

We notice that we can use Brouwer's Fixed Point Theorem (instead of Banach's Fixed Point Theorem) and an approximation technique to show that the result established in part 1) of the exercise above remains true if φ is merely continuous. This is the Schauder-Tykhonov Theorem, see [Dug66, Theorems 2.2 and 3.2, pages 414 and 415] or [DG82, Theorem 2.3, page 74]. It follows that the statements in parts 3) and 4) are also valid when the involved maps are merely continuous.

Second proof of Weak KAM Theorem 4.7.1. Let us denote by $\mathbf{1}$ the constant function equal to 1 everywhere on M and consider the quotient $E = \mathcal{C}^0(M,\mathbb{R})/\mathbb{R}.\mathbf{1}$. This quotient space E is a Banach space for the quotient norm

$$\|[u]\| = \inf_{a\in\mathbb{R}} \|u + a\mathbf{1}\|_\infty,$$

where $[u]$ is the class in E of $u \in \mathcal{C}^0(M,\mathbb{R})$. Since $T_t^-(u + a\mathbf{1}) = T_t^-(u) + a\mathbf{1}$, if $a \in \mathbb{R}$, the maps T_t^- pass to the quotient to a semigroup $\bar{T}_t^- : E \to E$ consisting of non-expansive maps. Since, for each $t > 0$, the image of T_t^- is an equi-Lipschitzian family of maps, Ascoli's Theorem, see for example [Dug66, Theorem 6.4 page 267], then shows that the image of $\bar{T}_t^-$ is relatively compact in E (exercise). Using part 4) in the exercise above, we find a common fixed point for all the $\bar{T}_t^-$. We then deduce that there exists $u_- \in \mathcal{C}^0(M,\mathbb{R})$ such that $T_t^- u_- = u_- + c_t$, where c_t is a constant. The semigroup property gives $c_{t+t'} = c_t + c_{t'}$; since $t \mapsto T_t^- u$ is continuous, we obtain $c_t = -tc$ with $c = -c_1$. We thus have $T_t^- u_- + ct = u_-$. $\square$

FROM HERE ON NOTES HAVE TO REVISED.

4.8 Invariant Measures and Ma né's Critical Value

Corollary 4.8.1. *If $T_t^- u_- = u_- - ct$, then, we have*

$$-c = \inf_{\mu} \int_{TM} L \, d\mu,$$

where μ varies among Borel probability measures on TM invariant by the Euler-Lagrange flow ϕ_t. This lower bound is in fact achieved by a measure with compact support. In particular, the constant c is unique.

Proof. If $(x, v) \in TM$, then $\gamma : [0, +\infty[\to M$, defined by $\gamma(s) = \pi\phi_s(x, v)$, satisfies $(\gamma(s), \dot\gamma(s)) = \phi_s(x, v)$. Since $u_- \prec L + c$, we find

$$u_-(\pi\phi_1(x, v)) - u_-(\pi(x, v)) \leq \int_0^1 L(\phi_s(x, v)) \, ds + c.$$

If μ is a probability measure invariant by ϕ_t, the function $u_- \circ \pi$ is integrable since it is bounded. Invariance of the measure by ϕ_t gives

$$\int_{TM} [u_-(\pi\phi_1(x, v)) - u_-(\pi(x, v))] \, d\mu(x, v) = 0,$$

from where, by integration of the inequality above, we obtain

$$0 \leq \int_{TM} [\int_0^1 L(\phi_s(x, v)) \, ds + c] \, d\mu(x, v).$$

Since L is bounded below and μ is a probability measure, we can apply Fubini Theorem to obtain

$$0 \leq \int_0^1 \big[\int_{TM} (L(\phi_s(x, v)) + c) \, d\mu(x, v)\big] \, ds.$$

By the invariance of μ under ϕ_s, we find that $0 \leq \int (L + c) d\mu$. Since μ is a probability measure, this yields

$$-c \leq \int_{TM} L \, d\mu.$$

It remains to see that the value $-c$ is attained. For that, we fix $x \in M$, and we take a curve $\gamma_x^- :]-\infty, 0] \to M$ with $\gamma_x^-(0) = x$ and such that

$$\forall t \leq 0, u_-(\gamma_x^-(0)) - u_-(\gamma_x^-(t)) = \int_t^0 L(\gamma_x^-(s), \dot\gamma_x^-(s))\, ds - ct.$$

The curve γ_x is a minimizing extremal curve with $\gamma_x^-(0) = x$, therefore, we have

$$\phi_s(x, \dot\gamma_x^-(0)) = (\gamma_x^-(s), \dot\gamma_x^-(s))$$

and the curve $(\gamma_x^-(s), \dot\gamma_x^-(s)), s \leq 0$ is entirely contained in a compact subset K_1 of TM as given by corollary 4.4.5. Using Riesz Representation Theorem [Rud87, Theorem 2.14, page 40], for $t > 0$, we define a Borel probability measure μ_t on TM by

$$\mu_t(\theta) = \frac{1}{t} \int_{-t}^0 \theta(\phi_s(x, \dot\gamma_x^-(0)))\, ds,$$

for $\theta : TM \to \mathbb{R}$ a continuous function. All these probability measures have their supports contained in the compact subset K_1, consequently, we can extract a sequence $t_n \nearrow +\infty$ such that μ_{t_n} converges weakly to a probability measure μ with support in K_1. Weak convergence means that for each continuous function $\theta : TM \to \mathbb{R}$, we have

$$\int_{TM} \theta\, d\mu = \lim_{n\to\infty} \frac{1}{t_n} \int_{-t_n}^0 \theta(\phi_s(x, \dot\gamma_x^-(0))\, ds.$$

We have $\int_{TM} (L + c)\, d\mu = 0$, because

$$\int_{TM} (L + c)\, d\mu = \lim_{n\to\infty} \frac{1}{t_n} \int_{-t_n}^0 \big(L(\gamma_x^-(s), \dot\gamma_x^-(s)) + c\big)\, ds$$

and

$$\int_{-t_n}^0 L(\gamma_n^-(s), \dot\gamma_x^-(s))\, ds + ct_n = u_-(\gamma_x^-(0)) - u_-(\gamma_x^-(-t_n))$$

which is bounded by $2\|u_-\|_\infty$. This does indeed show that for the limit measure μ, we have $\int_{TM} L\, d\mu = -c$. $\square$

In the second part of the previous proof, we have in fact shown the following proposition.

Proposition 4.8.2. *If $T_t^- u_- = u_- + ct$, for every $t \geq 0$, and $\gamma_x^- :]-\infty, 0] \to M$ is an extremal curve, with $\gamma_x^-(0) = x$, and such that*

$$\forall t \geq 0, u_-(x) - u_-(\gamma_x^-(-t)) = \int_{-t}^{0} L(\gamma_x^-(s), \dot{\gamma}_x^-(s))\, ds. + ct,$$

the following properties are satisfied

- *for each $s \geq 0$, we have $\phi_{-s}(x, \dot{\gamma}_x^-(0)) = (\gamma_x^-(-s), \dot{\gamma}_x^-(-s))$;*
- *the α-limit set of the orbit of $(x, \dot{\gamma}_x^-(0))$ for ϕ_s is compact;*
- *there exists a Borel probability measure μ on TM, invariant by ϕ_t, carried by the α-limit set of the orbit of $(x, \dot{\gamma}_x^-(0))$, and such that $\int L\, d\mu = -c$.*

We define $c[0] \in \mathbb{R}$ by

$$c[0] = -\inf_{\mu} \int_{TM} L\, d\mu,$$

where the lower bound is taken with respect to all Borel probability measures on TM invariant by the Euler-Lagrange flow. We will use the notation $c_L[0]$, if we want to specify the Lagrangian.

Definition 4.8.3 (Minimizing Measure)**.** A measure μ on TM is said to be minimizing if it is a Borel probability measure μ, invariant by the Euler-Lagrange flow, which satisfies

$$c[0] = -\int_{TM} L\, d\mu.$$

Exercise 4.8.4. *Show that each Borel probability measure μ, invariant by the Euler-Lagrange flow, and whose support is contained in the α-limit set of a trajectory of the form $t \mapsto (\gamma_-^x(t), \dot{\gamma}_-^x(t))$, must be minimizing.*

Corollary 4.8.5. *If a weak KAM solution u_- is differentiable at $x \in M$, we have*

$$H(x, d_x u_-) = c[0].$$

Proof. If u_- is a weak KAM solution, we have $u_- \prec L + c[0]$ and thus

$$H(x, d_x u_-) = \sup_{v \in T_x M} d_x u_-(v) - L(x, v) \leq c[0].$$

It remains to find $v_0 \in T_x M$ such that $d_x u_-(v_0) = L(x, v_0) + c[0]$. For that, we pick extremal curves $\gamma_-^x :]-\infty, 0]$ such that $\gamma_-^x(0) = x$ and

$$\forall t \geq 0, u_-(\gamma_-^x(0)) - u_-(\gamma_-^x(-t)) = \int_{-t}^{0} L(\gamma_-^x(s), \dot{\gamma}_-^x(s))\, ds + c[0]t.$$

By dividing this equality by $t > 0$ and letting t tend to 0, we find

$$d_x u_-(\dot{\gamma}_-^x(0)) = L(x, \dot{\gamma}_-^x(0)) + c[0].$$

□

4.9 The Symmetrical Lagrangian

Definition 4.9.1 (Symmetrical Lagrangian)**.** If $L : TM \to \mathbb{R}$ is a Lagrangian we define its symmetrical Lagrangian $\check{L} : TM \to \mathbb{R}$ by

$$\check{L}(x, v) = L(x, -v).$$

If $\gamma : [a, b] \to M$ is a curve, we define the curve $\check{\gamma} : [a, b] \to M$ by $\check{\gamma}(s) = \gamma(a + b - s)$. It is immediate to check that $\dot{\check{\gamma}}(s) = -\dot{\gamma}(a + b - s)$ and thus

$$\check{\mathbb{L}}(\check{\gamma}) = \mathbb{L}(\gamma),$$

where $\check{\mathbb{L}}$ is the action associated with $\check{\mathbb{L}}$, i.e.

$$\check{\mathbb{L}}(\check{\gamma}) = \int_a^b \check{L}(\check{\gamma}(s), \dot{\check{\gamma}}(s))\, ds.$$

It clearly results that γ is an extremal curve of L if and only if $\check{\gamma}$ is an extremal curve of $\check{L}$. We want to express the Lax-Oleinik semigroup $\check{T}_t^- : \mathcal{C}^0(M, \mathbb{R}) \to \mathcal{C}^0(M, \mathbb{R})$ associated with $\check{L}$ in terms of L only. For that, we notice that when $\gamma : [0, t] \to M$ varies among all the curves such that $\gamma(0) = x$, then $\check{\gamma}$ varies among all

all the curves such that $\check{\gamma}(t) = x$, and we have $\check{\gamma}(0) = \gamma(t)$. We thus find

$$\begin{aligned}\check{T}_t^- u(x) &= \inf_\gamma \{\int_0^t \check{L}(\check{\gamma}(s), \dot{\check{\gamma}}(s))\, ds + u(\check{\gamma}(0))\} \\ &= \inf_\gamma \{\int_0^t L(\gamma(s), \dot{\gamma}(s))\, ds + u(\gamma(t))\},\end{aligned}$$

where the lower bound is taken all the piecewise C^1 curves $\gamma : [0,t] \to M$ such that $\gamma(0) = x$.

We then introduce the semigroup $T_t^+ : \mathcal{C}^0(M,\mathbb{R}) \to \mathcal{C}^0(M,\mathbb{R})$ defined by $T_t^+(u) = -\check{T}_t^-(-u)$. We find

$$T_t^+ u(x) = \sup_\gamma \{u(\gamma(t)) - \int_0^t L(\gamma(s), \dot{\gamma}(s))\, ds\},$$

where the upper bound is taken on the (continuous) piecewise C^1 curves $\gamma : [0,t] \to M$ such that $\gamma(0) = x$.

The following lemma is easily verified.

Lemma 4.9.2. *We have $u \prec L + c$ if and only if $T_t^+ u - ct \leq u$.*

By the Weak KAM Theorem 4.7.1, we can find $\check{u}_- \in \mathcal{C}^0(M,R)$ and $\check{c}$ such that $\check{T}_t^- \check{u} + \check{c}t = \check{u}_-$. If we set $u_+ = -\check{u}_-$, and we find $T_t^+ u_+ = u_+ + \check{c}t$. If $\gamma : [a,b] \to M$ is an arbitrary (continuous) piecewise C^1 curve of C^1, we see that

$$u_+(\gamma(a)) + \check{c}(b-a) = T_{(b-a)}^+ u_+(\gamma(a)) \geq u_+(\gamma(b)) - \int_a^b L(\gamma(s), \dot{\gamma}(s))ds,$$

which gives $u_+ \prec L + \check{c}$.

By arguments similar to the ones we made for u_-, we obtain:

Theorem 4.9.3 (Weak KAM)**.** *There exists a Lipschitzian function $u_+ : M \to \mathbb{R}$ and a constant c such that $T_t^+ u_+ - ct = u_+$. This function u_+ satisfies the following properties*

(a) $u_+ \prec L + c$.

(b) *For each $x \in M$, there exists a minimizing extremal curve $\gamma_+^x : [0, +\infty[\to M$ with $\gamma_+^x(0) = x$ and such that*

$$\forall t \in [0, +\infty[, u_+(\gamma_+^x(t)) - u_+(x) = \int_0^t L(\gamma_+^x(s), \dot{\gamma}_+^x(s))ds + ct.$$

Conversely, if $u_+ \in \mathcal{C}^0(M,R)$ satisfies the properties (a) *and* (b) *above, then, we have $T_t^+ u_+ - ct = u_+$.*

We also have

$$-c = \inf_\mu \int L\, d\mu,$$

where the lower bound is taken over all the Borel probability measures μ on TM invariant by the Euler-Lagrange flow ϕ_t. It follows that $c = c[0]$. For a curve γ_+^x of the type above, we have

$$\forall s \geq 0, (\gamma_+^x(s), \dot\gamma_+^x(s)) = \phi_s(x, \dot\gamma_+^x(0)),$$

since it is a minimizing extremal curve. This implies that $\phi_s(x, \dot\gamma_+^x(0)), s \geq 0$ is relatively compact in TM. Moreover, we can find a Borel probability measure μ on TM, invariant by the Euler-Lagrange flow ϕ_t, such that $-c[0] = \int L\, d\mu$, and whose support is contained in the ω-limit set of the orbit $\phi_s(x, \dot\gamma_+^x(0))$. At each point $x \in M$, where u_+ has a derivative, we have $H(x, d_x u_+) = c[0]$.

4.10 The Mather Function on Cohomology.

Let us first give several characterizations of $c[0]$.

Theorem 4.10.1. *Suppose that $u \in \mathcal{C}^0(M,\mathbb{R})$, and that $c \in \mathbb{R}$.*

If $c > c[0]$ then $T_t^- u + ct$ tends uniformly to $+\infty$, as $t \to +\infty$, and $T_t^+ u - ct$ tends uniformly to $-\infty$, as $t \to +\infty$.

If $c < c[0]$, then $T_t^- u + ct$ tends uniformly to $-\infty$, when $t \to +\infty$, and $T_t^+ u - ct$ tends uniformly to $+\infty$, as $t \to +\infty$.

Moreover, we have $\sup_{t\geq 0} \|T_t^- u + c[0]t\| < +\infty$ and $\sup_{t \geq 0} \|T_t^+ u - c[0]t\| < +\infty$.

Proof. By the Weak KAM Theorem 4.7.1, there exists $u_- \in \mathcal{C}^0(M,\mathbb{R})$ with $T_t^- u_- + c[0]t = u_-$, for each $t \geq 0$. As the T_t^- are non-expansive maps, we have

$$\|T_t^- u - T_t^- u_-\| \leq \|u - u_-\|_\infty$$

and thus

$$-\|u - u_-\|_\infty - \|u_-\|_\infty \leq T_t^- u + c[0]t \leq \|u - u_-\|_\infty + \|u_-\|_\infty.$$

□

Theorem 4.10.2. *We have the following characterization of the constant* $c[0]$*:*

- *The constant* $c[0]$ *is the only constant* c *such that the semi-group* $u \mapsto T_t^- u + ct$ *(resp.* $u \mapsto T_t^+ u - ct$*) has a fixed point in* $\mathcal{C}^0(M, \mathbb{R})$*.*
- *The constant* $c[0]$ *is the greatest lower bound of the set of the numbers* $c \in \mathbb{R}$ *for which there exists* $u \in \mathcal{C}^0(M, \mathbb{R})$ *with* $u \prec L + c$*.*
- *The constant* $c[0]$ *is the only constant* $c \in \mathbb{R}$ *such that there exists* $u \in \mathcal{C}^0(M, \mathbb{R})$ *with* $\sup_{t \geq 0} \|T_t^- u + ct\|_\infty < +\infty$ *(resp.* $\sup_{t \geq 0} \|T_t^+ u - ct\|_\infty < +\infty$*.)*

Proof. The first point results from the Weak KAM Theorem. The last point is a consequence of the previous Theorem 4.10.1. The second point also results from the previous theorem, because we have $u \prec L + c$ if and only if $u \leq T_t^- u + ct$ and in addition, by the weak KAM theorem, there exists u_- with $u_- = T_t u_- + c[0]t$. □

If ω is a C^∞ differential 1-form, it is not difficult to check the Lagrangian $L_\omega : TM \to \mathbb{R}$, defined by

$$L_\omega(x, v) = L(x, v) - \omega_x(v),$$

is C^r like L, with $r \geq 2$, that $\frac{\partial^2 L_\omega}{\partial v^2} = \frac{\partial^2 L}{\partial v^2}$ is thus also > 0 definite as a quadratic form, and that L_ω is superlinear in the fibers of the tangent bundle TM. We then set

$$c[\omega] = c_L[\omega] = c_{L-\omega}[0].$$

Proposition 4.10.3. *If* $\theta : M \to \mathbb{R}$ *is a differentiable function, then we have*

$$c[\omega + d\theta] = c[\omega].$$

In particular, for closed forms ω*, the constant* $c[\omega]$ *depends only on the cohomology class.*

Proof. We have $u \prec (L - [\omega + d\theta]) + c$ if and only if $u + \theta \prec (L - \omega) + c$. □

The following definition is due to Mather, see [Mat91, page 177].

Definition 4.10.4 (Mather's α Function)**.** The function α of Mather is the function $\alpha : H^1(M, \mathbb{R}) \to \mathbb{R}$ defined by

$$\alpha(\Omega) = c[\omega],$$

where ω is a class C^∞ differential 1-form representing the class of cohomology Ω.

The next theorem is due to Mather, see [Mat91, Theorem 1, page 178].

Theorem 4.10.5 (Mather)**.** *The function α is convex and superlinear on the first cohomology group $H^1(M, \mathbb{R})$.*

Proof. Let ω_1 and ω_2 be two differential 1-forms of class C^∞. By the weak KAM Theorem applied to L_{ω_1} and L_{ω_2}, we can find $u_1, u_2 \in \mathcal{C}^0(M, \mathbb{R})$ such that

$$u_i \prec (L - \omega_i) + c[\omega_i].$$

If $t \in [0, 1]$, it is not difficult to conclude that

$$tu_1 + (1-t)u_2 \prec (L - [t\omega_1 + (1-t)\omega_2]) + (tc[\omega_1] + (1-t)c[\omega_2].$$

It follows that

$$c[t\omega_1 + (1-t)\omega_2] \leq tc[\omega_1] + (1-t)c[\omega_2].$$

Let us show the superlinearity of α. Let us recall that by compactness of M, the first group of homology is a vector space of finite dimension. Let us fix a finite family $\gamma_1, \cdots, \gamma_n : [0, 1] \to M$ of C^∞ closed (i.e. $\gamma_i(0) = \gamma_i(1)$ curves such that the homology $H_1(M, \mathbb{R})$ is generated by the homology classes of $\gamma_1, \cdots, \gamma_\ell$. We can then define a norm on $H^1(M, \mathbb{R})$ by

$$\|\Omega\| = \max(|\int_{\gamma_1} \omega|, \ldots, |\int_{\gamma_\ell} \omega|),$$

where ω is a C^∞ closed differential 1-form representing the cohomology class Ω. Let k be an integer. Let us note by γ_i^k the closed curve $\gamma_i^k : [0,1] \to M$ obtained by going k times through γ_i in the direction of the increasing t and reparametrizing it by the interval $[0,1]$. Let us also note by $\check{\gamma}_i^k : [0,1] \to M$ the curve opposite to γ_i, i.e. $\check{\gamma}_i^k(s) = \gamma_i^k(1-s)$. We then have

$$\int_{\gamma_i^k} \omega = k \int_{\gamma_i} \omega$$
$$\int_{\check{\gamma}_i^k} \omega = -k \int_{\gamma_i} \omega,$$

from which we obtain the equality

$$k\|\Omega\| = \max\Big(\Big|\int_{\gamma_1^k} \omega\Big|, \dots, \Big|\int_{\gamma_\ell^k} \omega\Big|, \Big|\int_{\check{\gamma}_1^k} \omega\Big|, \dots, \Big|\int_{\check{\gamma}_\ell^k} \omega\Big|\Big),$$

where ω is a C^∞ closed differential 1-form representing Ω. By the Weak KAM Theorem 4.7.1, there exists $u_\omega \in \mathcal{C}^0(M, \mathbb{R})$ such that $u_\omega \prec (L - \omega) + c[\omega]$. We deduce from it that for every closed (i.e. $\gamma(b) = \gamma(a)$) curve $\gamma : [a,b] \to M$, we have

$$\mathbb{L}(\gamma) - \int_\gamma \omega + c[\omega](b-a) \geq 0.$$

In particular, we find

$$c[\omega] \geq \int_{\gamma_k^i} \omega - \mathbb{L}(\gamma_i^k),$$
$$c[\omega] \geq \int_{\check{\gamma}_k^i} \omega - \mathbb{L}(\check{\gamma}_i^k).$$

Hence, if we set $C_k = \max(\mathbb{L}(\gamma_1^k), \dots, \mathbb{L}(\gamma_\ell^k), \mathbb{L}(\check{\gamma}_1^k), \dots, \mathbb{L}(\check{\gamma}_\ell^k))$, which is a constant which depends only on k, we find

$$\alpha(\Omega) = c[\omega] \geq k\|\Omega\| - C_k. \square$$

Theorem 4.10.6. *If ω is a closed 1-form, the Euler-Lagrange flows $\phi_t^{L-\omega}$ and ϕ_t^L coincide.*

Proof. Indeed, if $\gamma_1, \gamma_2 : [a,b] \to M$ are two curves with the same ends and close enough, they are homotopic with fixed ends and thus $\int_{\gamma_1} \omega = \int_{\gamma_2} \omega$. It follows that the actions for L and L_ω are related by

$$\mathbb{L}_\omega(\gamma_i) = \mathbb{L}(\gamma_i) - \int_{\gamma_1} \omega.$$

Hence, the critical points of $\mathbb{L}_\omega$ and of $\mathbb{L}$ on the space of curves with fixed endpoints are the same ones. Consequently, the Lagrangians L and L_ω have same the extremal curves. $\square$

Corollary 4.10.7. *We have $c[\omega] = -\inf_\mu \int_{TM}(L-\omega)d\mu$, where μ varies among the Borel probability measures on TM invariant under the Euler-Lagrange flow ϕ_t of L. Moreover, there exists such a measure μ with compact support and satisfying $c[\omega] = \int_{TM}(L-\omega)d\mu$.*

4.11 Differentiability of Dominated Functions

In the sequel we denote by $\mathring{B}(0,r)$ (resp. $\bar{B}(0,r)$) the open ball (resp. closed) of center 0 and radius r in the Euclidean space $\mathbb{R}^k$, where k is the dimension of M.

Proposition 4.11.1. *Let $\varphi : \mathring{B}(0,5) \to M$ be a coordinate chart and $t_0 > 0$ be given. There is a constant $K \geq 0$ such that for each function $u \in \mathcal{C}^0(M,\mathbb{R})$, for each $x \in \bar{B}(0,1)$, and for each $t \geq t_0$, we have:*

(1) *For each $y \in \bar{B}(0,1)$ and each extremal curve $\gamma : [0,t] \to M$ with $\gamma(t) = \varphi(x)$ and*

$$T_t^- u[\varphi(x)] = u(\gamma(0)) + \int_0^t L(\gamma(s),\dot\gamma(s))ds,$$

we have

$$T_t^- u[\varphi(y)] - T_t^- u[\varphi(x)] \leq \frac{\partial L}{\partial v}(\varphi(x),\dot\gamma(t))(D\varphi(x)[y-x]) + K\|y-x\|^2.$$

In particular, if $T_t^- u$ is differentiable at $\varphi(x)$ then

$$d_{\varphi(x)} T_t^- u = \partial L/\partial v(\varphi(x),\dot\gamma(t))$$

and the curve γ is unique.

(2) *For each $y \in \bar{B}(0,1)$ and each curve $\gamma : [0,t] \to M$ with $\gamma(0) = \varphi(x)$ and*

$$T_t^+ u[\varphi(x)] = u(\gamma(t)) - \int_0^t L(\gamma(s), \dot\gamma(s))ds,$$

we have

$$T_t^+ u[\varphi(y)] - T_t^+ u[\varphi(x)] \geq \frac{\partial L}{\partial v}(\varphi(x), \dot\gamma(0))(D\varphi(x)[y-x]) - K\|y-x\|^2.$$

In particular, if $T_t^+ u$ is differentiable at $\varphi(x)$ then

$$d_{\varphi(x)} T_t^+ u = \partial L / \partial v(\varphi(x), \dot\gamma(0))$$

and the curve γ is unique.

Proof. We use some auxiliary Riemannian metric on M to have a norm on tangent space and a distance on M.

Since $T_t^- u = T_{t_0}^- T_{t-t_0}^- u$ by the semigroup property, we have only to consider the case $t = t_0$. By corollary 4.4.5, we can find a finite constant A, such that any minimizer defined on an interval of time at least $t_0 > 0$ has speed uniformly bounded in norm by A.This means that such curves are all A-Lipschitz. We then pick $\epsilon > 0$ such that for each ball $\bar{B}(y, A\epsilon)$, for $y \in \varphi(\bar{B}(0,1))$ is contained in $\varphi(\mathring{B}(0,2))$. Notice that this ϵ does not depend on u. Since any curve γ, as in part (1) of the proposition is necessary a minimizer, we therefore have

$$\gamma([t_0 - \epsilon, t_0]) \subset \varphi(\mathring{B}(0,2)).$$

We then set $\tilde\gamma = \varphi^{-1} \circ \gamma$ and $\tilde{L}(x,w) = L(\varphi(x), D\varphi(x)w)$ for $(x,w) \in \mathring{B}(0,5) \times \mathbb{R}^k$. Taking derivatives, we obtain

$$\frac{\partial \tilde{L}(x,w)}{\partial w} = \frac{\partial L}{\partial v}(\varphi(x), D\varphi(x)(w))[D\varphi(x)(\cdot)].$$

The norm of the vector $h = y - x$ is ≤ 2, hence if we define $\tilde\gamma_h(s) = \frac{s-(t_0-\epsilon)}{\epsilon} h + \tilde\gamma(s)$, for $s \in [t_0 - \epsilon, t_0]$, we have $\tilde\gamma_h(s) \in \mathring{B}(0,4)$, and

$$T_{t_0}^- u[\varphi(x+h)] - T_{t_0}^- u[\varphi(x)] \leq \int_{t_0-\epsilon}^{t_0} [\tilde{L}(\tilde\gamma_h(s), \dot{\tilde\gamma}_h(s)) - L(\tilde\gamma(s), \dot{\tilde\gamma}(s))]\, ds.$$

Since the speed of γ is bounded in norm by A, we see that $(\tilde\gamma(s),\dot{\tilde\gamma}(s))$ is contained in a compact subset of $\bar B(0,2)\times\mathbb{R}^k$ independent of u,x and γ. Moreover, we have $\tilde\gamma_h(s)-\tilde\gamma(s)=\frac{s-(t-\epsilon)}{\epsilon}h$, which is of norm ≤ 2, and $\dot{\tilde\gamma}_h(s)-\dot{\tilde\gamma}(s)=\frac{1}{\epsilon}h$, which is itself of norm $\leq 2/\epsilon$. We then conclude that there exists a compact convex subset $C\subset\bar B(0,4)\times\mathbb{R}^k$, independent of u,x and γ, and containing $(\tilde\gamma(s),\dot{\tilde\gamma}(s))$ and $(\tilde\gamma_h(s),\dot{\tilde\gamma}_h(s))$, for each $s\in[t_0-\epsilon,t_0]$. Since we can bound uniformly on the compact subset C the norm of the second derivative of $\tilde L$, by Taylor's formula at order 2 applied to $\tilde L$, we see that there exists a constant $\tilde K$ independent of u,x and of γ such that for each $s\in[t_0-\epsilon,t_0]$

$$\begin{aligned}
|\tilde L(\tilde\gamma_h(s),\dot{\tilde\gamma}_h(s))-\tilde L(\tilde\gamma(s),\dot\gamma(s))&\\
-\frac{\partial\tilde L}{\partial x}(\tilde\gamma(s),\dot\gamma(s))\big(\frac{s-(t-\epsilon)}{\epsilon}h\big)&-\frac{\partial\tilde L}{\partial w}(\tilde\gamma(s),\dot{\tilde\gamma}(s))\big(\frac{h}{\epsilon}\big)|\\
&\leq\tilde K\max\big(\frac{|s-(t-\epsilon)|}{\epsilon}\|h\|,\frac{\|h\|}{\epsilon}\big)^2\\
&\leq\frac{\tilde K}{\epsilon^2}\|h\|^2,
\end{aligned}$$

supposing that $\epsilon<1$. As $\tilde\gamma$ is an extremal curve for the Lagrangian $\tilde L$, it satisfies the Euler-Lagrange equation

$$\frac{\partial\tilde L}{\partial x}(\tilde\gamma(s),\dot\gamma(s))=\frac{d}{ds}\big\{\frac{\partial\tilde L}{\partial w}(\gamma(s),\dot\gamma(s))\big\}$$

hence

$$\begin{aligned}
|\tilde L(\tilde\gamma_h(s),\dot{\tilde\gamma}_h(s))-\tilde L(\tilde\gamma(s),\dot\gamma(s))-\frac{1}{\epsilon}\frac{d}{ds}\big\{[s-(t-\epsilon)]\frac{\partial\tilde L}{\partial w}(\tilde\gamma(s),\dot\gamma(s))(h)\big\}|&\\
\leq\frac{\tilde K}{\epsilon^2}\|h\|^2.&
\end{aligned}$$

If we integrate on the interval $[t_0-\epsilon,t_0]$, we obtain

$$T^-_{t_0}u[\varphi(x+h)]-T^-_{t_0}u[\varphi(x)]\leq\frac{\partial\tilde L}{\partial w}(\varphi(x),\dot{\tilde\gamma}(t_0))(h)+\frac{\tilde K}{\epsilon}\|h\|^2.$$

We just have to take $K\geq\tilde K/\epsilon$.

If $T^-_t u$ is differentiable at $\varphi(x)$, for $v\in\mathbb{R}^n$ and δ small enough, we can write

$$T^-_t u[\varphi(x+\delta v)]-T^-_t u[\varphi(x)]\leq\frac{\partial L}{\partial v}(\varphi(x),\dot\gamma(t))(D\varphi(x)[\delta v])+K\|\delta v\|^2.$$

If we divide by $\delta > 0$ and we let δ go to 0, we obtain

$$\forall v \in \mathbb{R}^n d_{\varphi(x)} T_t^- u D\varphi(x)[v] \leq \frac{\partial L}{\partial v}(\varphi(x), \dot{\gamma}(t))(D\varphi(x)[v]).$$

Since we can also apply the inequality above with $-v$ instead of v, we conclude that If we divide by $\delta > 0$ and we let δ go to 0, we obtain

$$\forall v \in \mathbb{R}^n d_{\varphi(x)} T_t^- u D\varphi(x)[v] = \frac{\partial L}{\partial v}(\varphi(x), \dot{\gamma}(t))(D\varphi(x)[v]),$$

by the linearity in v of the involved maps. Since φ is a diffeomorphism this shows that $d_{\varphi(x)} T_t^- u \partial L/\partial v(\varphi(x), \dot{\gamma}(t))$. By the bijectivity of the Legendre Transform the tangent vector $\dot{\gamma}(t)$ is unique. Since γ is necessarily a minimizing extremal, it is also unique, since both its position x and its speed $\dot{\gamma}(t)$ at t are uniquely determined.

To prove (2), we can make a similar argument for T_t^+, or, more simply, apply what we have just done to the symmetrical Lagrangian $\check{L}$ of L. □

Exercise 4.11.2. *1) Let $u_- : M \to \mathbb{R}$ be a weak KAM solution. Show that u_- has a derivative at x if and only if there is one and only one curve $\gamma_-^x :]-\infty, 0] \to M$ such that $\gamma_-^x(0) = x$ and*

$$\forall t \geq 0, u_-(\gamma_-^x(0)) - u_-(\gamma_-^x(-t)) = \int_{-t}^0 L(\gamma_-^x(s), \dot{\gamma}_-^x(s))\, ds + c[0]t.$$

In that case we have $d_x u_- = \frac{\partial L}{\partial v}(x, \dot{\gamma}_-^x(0))$.

2) Suppose that $x \in M$, and that $\gamma_-^x :]-\infty, 0] \to M$ satisfies $\gamma_-^x(0) = x$ and

$$\forall t \geq 0, u_-(\gamma_-^x(0)) - u_-(\gamma_-^x(-t)) = \int_{-t}^0 L(\gamma_-^x(s), \dot{\gamma}_-^x(s))\, ds + c[0]t.$$

Show that necessarily

$$H(x, \frac{\partial L}{\partial v}(x, \dot{\gamma}_-^x(0)) = c[0].$$

We will need a criterion to show that a map is differentiable with a Lipschitz derivative. This criterion has appeared in different forms either implicitly or explicitly in the literature, see

[CC95, Proposition 1.2 page 8], [Her89, Proof of 8.14, pages 63–65], [Kni86], [Lio82, Proof of Theorem 15.1, pages 258-259], and also [Kis92] for far reaching generalizations. The simple proof given below evolved from discussions with Bruno Sevennec.

Proposition 4.11.3 (Criterion for a Lipschitz Derivative). *Let $\mathring{B}$ be the open unit ball in the normed space E. Fix a map $u : \mathring{B} \to \mathbb{R}$. If $K \geq 0$ is a constant, denote by $A_{K,u}$ the set of points $x \in \mathring{B}$, for which there exists $\varphi_x : E \to \mathbb{R}$ a continuous linear form such that*

$$\forall y \in \mathring{B}, |u(y) - u(x) - \varphi_x(y - x)| \leq K\|y - x\|^2.$$

Then the map u has a derivative at each point $x \in A_{K,u}$, and $d_x u = \varphi_x$. Moreover, the restriction of the map $x \mapsto d_x u$ to $\{x \in A_{K,u} \mid \|x\| < \frac{1}{3}\}$ is Lipschitzian with Lipschitz constant $\leq 6K$.

More precisely

$$\forall x, x' \in A_{K,u}, \|x - x'\| < \min(1 - \|x\|, 1 - \|x'\|) \Rightarrow \|d_x u - d_{x'} u\| \leq 6K\|x - x'\|.$$

Proof. The fact that $d_x u = \varphi_x$, for $x \in A_{K,u}$, is clear. Let us fix $x, x' \in A_{K,u}$, with $\|x - x'\| < \min(1 - \|x\|, 1 - \|x'\|)$. If $x = x'$ there is nothing to show, we can then suppose that $\|x - x'\| > 0$. If h is such that $\|h\| = \|x - x'\|$, then the two points $x + h$ et $x' + h$ are in $\mathring{B}$. This allows us to write

$$\begin{aligned} |u(x+h) - u(x) - \varphi_x(h)| &\leq K\|h\|^2 \\ |u(x) - u(x') - \varphi_{x'}(x - x') &\leq K\|x - x'\|^2 \\ |u(x+h) - u(x') - \varphi_{x'}(x - x' + h)| &\leq K\|x - x' + h\|^2. \end{aligned}$$

As $\|h\| = \|x - x'\|$, we obtain from the last inequality

$$|u(x') - u(x+h) + \varphi_{x'}(x - x' + h)| \leq 4K\|x - x'\|^2.$$

Adding this last inequality with the first two above, we find that, for each h such that $\|h\| = \|x - x'\|$, we have

$$|\varphi_{x'}(h) - \varphi_x(h)| \leq 6K\|x - x'\|^2,$$

hence

$$\|\varphi_{x'} - \varphi_x\| = \sup_{\|h\| = \|x - x'\|} \frac{|\varphi_{x'}(h) - \varphi_x(h)|}{\|x - x'\|} \leq 6K\|x - x'\|.$$

$\square$

Exercise 4.11.4. *If $A_{k,u}$ is convex, for example if $A_{K,U} = \mathring{B}$, show that the derivative is Lipschitzian on $A_{K,u}$ with Lipschitz constant $\leq 6K$.*

Theorem 4.11.5. *If $\epsilon > 0$ is given, then there are constants $A \geq 0$ and $\eta > 0$, such that any map $u : M \to \mathbb{R}$, with $u \prec L + c$, is differentiable at every point of the set $\mathcal{A}_{\epsilon,u}$ formed by the $x \in M$ for which there exists a (continuous) piecewise C^1 curve $\gamma : [-\epsilon, \epsilon] \to M$ with $\gamma(0) = x$ and*

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

Moreover we have

(1) *Such a curve γ is a minimizing extremal and*

$$d_x u = \frac{\partial L}{\partial x}(x, \dot{\gamma}(0));$$

(2) *the set $\mathcal{A}_{\epsilon,u}$ is closed;*

(3) *the derivative map $\mathcal{A}_{\epsilon,u} \to T^*M, x \mapsto (x, d_x u)$ is Lipschitzian with Lipschitz constant $\leq A$ on each subset $\mathcal{A}_{\epsilon,u}$ with diameter $\leq \eta$.*

Proof. The fact that γ is a minimizing extremal curve results from $u \prec L + c$. This last condition does also imply that

$$\begin{aligned} u(\gamma(\epsilon)) - u(x) &= \int_{0}^{\epsilon} L(\gamma(s), \dot{\gamma}(s))\, ds + c\epsilon, \\ u(x) - u(\gamma(-\epsilon)) &= \int_{-\epsilon}^{0} L(\gamma(s), \dot{\gamma}(s))\, ds + c\epsilon. \end{aligned} \tag{*}$$

Since $\epsilon > 0$ is fixed, by the Corollary of A Priori Compactness, we can find a compact subset $K_\epsilon \subset TM$ such that for each minimizing extremal curve $\gamma : [-\epsilon, \epsilon] \to M$, we have $(\gamma(s), \dot{\gamma}(s)) \in K_\epsilon$. It is not then difficult to deduce that $\mathcal{A}_{\epsilon,u}$ is closed. It also results from $(*)$ that

$$\begin{aligned} T_\epsilon^+ u(x) &\geq u(x) + c\epsilon \\ T_\epsilon^- u(x) &\leq u(x) - c\epsilon. \end{aligned} \tag{**}$$

As $u \prec L + c$, we have

$$T_\epsilon^+ u \leq u + c\epsilon,$$
$$T_\epsilon^- u \geq u - c\epsilon.$$

We then obtain equality in (**). Subtracting this equality from the inequality above, we find

$$\forall y \in M,\ T_\epsilon^+ u(y) - T_\epsilon^+ u(x) \leq u(y) - u(x) \leq T_\epsilon^- u(y) - T_\epsilon^- u(x). \tag{***}$$

Let us then cover the compact manifold M by a finite number of open subsets of the form $\varphi_1(\mathring{B}(0,1/3)), \cdots, \varphi_\ell(\mathring{B}(0,1/3))$, where $\varphi_i : \mathring{B}(0,5) \to M, i = 1,\ldots,\ell$, is a C^∞ coordinate chart. By Proposition 4.11.1, there exists a constant K, which depends only on ϵ and the fixed $\varphi_p, p = 1,\ldots,\ell$ such that, if $x \in \varphi_i(\mathring{B}(0,1/3))$, setting $x = \varphi_i(\tilde{x})$, for each $y \in \mathring{B}(0,1)$, we have

$$T_\epsilon^- u(\varphi_i(y)) - T_\epsilon^- u(\varphi_i(\tilde{x})) \leq \frac{\partial L}{\partial v}(x, \dot\gamma(0)) \circ D\varphi_i(\tilde{x})(y - \tilde{x}) + K\|y - \tilde{x}\|^2$$
$$T_\epsilon^+ u(\varphi_i(y) - T_\epsilon^+ u(\varphi_i(\tilde{x})) \geq \frac{\partial L}{\partial v}(x, \dot\gamma(0)) \circ D\varphi_i(\tilde{x})(y - \tilde{x}) - K\|y - \tilde{x}\|^2,$$

Using the inequalities (***) we get

$$|u(\varphi_i(y)) - u(\varphi_i(\tilde{x})) - \frac{\partial L}{\partial v}(x, \dot\gamma(0)) \circ D\varphi_i(\tilde{x})(y - \tilde{x})| \leq K\|y - \tilde{x}\|^2.$$

By the Criterion for a Lipschitz Derivative 4.11.3, we find that $d_x u$ exists and is equal to $\frac{\partial L}{\partial v}(x, \dot\gamma(0))$. Moreover the restriction of $x \mapsto d_x u$ on $\mathcal{A}_{\epsilon,U} \cap \varphi_i(\mathring{B}(0,1/3))$ is Lipschitzian with a constant of Lipschitz which depends only on ϵ. It is then enough to choose for $\eta > 0$ a Lebesgue number for the open cover $(\varphi_i(\mathring{B}(0,1/3))_{i=1,\cdots,\ell}$ of the compact manifold M. □

Definition 4.11.6 (The sets $\mathcal{S}_-$ and $\mathcal{S}_+$)**.** We denote by $\mathcal{S}_-$ (resp. $\mathcal{S}_+$) the set of weak KAM solutions of the type u_- (resp. u_+), i.e. the continuous functions $u : M \to \mathbb{R}$ such that $T_t^- u + c[0]t = u$ (resp. $T_t^+ u - c[0]t = u$).

Exercise 4.11.7. *If $c \in \mathbb{R}$, show that $u \in \mathcal{S}_-$ (resp. $u \in \mathcal{S}_+$) if and only if $u + c \in \mathcal{S}_-$ (resp. $u + c \in \mathcal{S}_+$). If $x_0 \in M$ is fixed, show that the set $\{u \in \mathcal{S}_- \mid u(x_0) = 0\}$ (resp. $\{u \in \mathcal{S}_+ \mid u(x_0) = 0\}$) is compact for the topology of uniform convergence.*

Theorem 4.11.8. *Let $u : M \to \mathbb{R}$ be a continuous function. The following properties are equivalent*

(1) *the function u is C^1 and belongs to $\mathcal{S}_-$,*

(2) *the function u is C^1 and belongs to $\mathcal{S}_+$.*

(3) *the function u belongs to the intersection $\mathcal{S}_- \cap \mathcal{S}_+$. (4)] the function u is C^1 and there exists $c \in \mathbb{R}$ such that $H(x, d_x u) = c$, for each $x \in M$.*

In all the cases above, the derivative of u is (locally) Lipschitzian.

Proof. Conditions (1) or (2) imply (4). It is enough, then to show that (4) implies (1) and (2) and that (3) implies that u is C^1, and that its derivative is (locally) Lipschitzian. Thus let us suppose condition (3) satisfied. Indeed, in this case, if $x \in M$, we can find extremal curves $\gamma_-^x :]-\infty, 0] \to M$ and $\gamma_+^x : [0, \infty[\to M$, with $\gamma_-^x(0) = \gamma_+^x(0) = x$ and

$$\begin{aligned}\forall t \geq 0, u(\gamma_+^x(t)) - u(x) &= \int_0^t L(\gamma_+^x(s), \dot\gamma_+^x(s))\, ds + c[0]t,\\ u(x) - u(\gamma_-^x(-t)) &= \int_{-t}^0 L(\gamma_-^x(s), \dot\gamma_-^x(s))\, ds + c[0]t.\end{aligned}$$

The curve $\gamma : [-1, 1] \to M$, defined by $\gamma|[-1, 0] = \gamma_-^x$ and $\gamma|[0, 1] = \gamma_+^x$, shows that $x \in \mathcal{A}_{1,u}$. By the previous theorem, the function u is of class C^1 and its derivative is (locally) Lipschitzian.

Let us suppose that u satisfies condition (4). By the Fenchel inequality, we have

$$\begin{aligned}\forall (x, v) \in TM, d_x u(v) &\leq H(x, d_x u) + L(x, v)\\ &= c + L(x, v).\end{aligned}$$

Consequently, if $\gamma : [a, b] \to M$ is a C^1 curve, we obtain

$$\forall s \in [a, b], d_{\gamma(s)} u(\dot\gamma(s)) \leq c + L(\gamma(s), \dot\gamma(s)), \tag{*}$$

and by integration on the interval $[a, b]$

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

which gives us $u \prec L + c$. If $\gamma : [a,b] \to M$ is an integral curve of $\operatorname{grad}_L u$, we have in fact equality in $(*)$ and thus in $(**)$. Since $u \prec L + c$, it follows that γ is a minimizing extremal curve. If two solutions go through the same point x at time t_0, they are, in fact, equal on their common interval of definition, since they are two extremal curves which have the same tangent vector $(x, \operatorname{grad}_L u(x))$ at time t_0. As we can find local solutions by the Cauchy-Peano Theorem, we see that $\operatorname{grad}_L u$ is uniquely integrable. Since M is compact, for any point $x \in M$, we can find an integral curve $\gamma^x :]-\infty, +\infty[\to M$ of $\operatorname{grad}_L u$ with $\gamma^x(0) = x$. This curve gives at the same time a curve of the type γ_-^x and one of the type γ_+^x for u. This establishes that $u \in \mathcal{S}_- \cap \mathcal{S}_+$. $\square$

4.12 Mather's Set.

The definition below is due to Mather, see [Mat91, page 184].

Definition 4.12.1 (Mather Set)**.** The Mather set is

$$\tilde{\mathcal{M}}_0 = \overline{\bigcup_\mu \operatorname{supp}(\mu)} \subset TM,$$

where $\operatorname{supp}(\mu)$ is the support of the measure μ, and the union is taken over the set of all Borel probability measures on TM invariant under the Euler-Lagrange flow ϕ_t, and such that $\int_{TM} L\, d_\mu = -c[0]$.

The projection $\mathcal{M}_0 = \pi(\tilde{\mathcal{M}}_0) \subset M$ is called the projected Mather set.

As the support of an invariant measure is itself invariant under the flow, the set $\tilde{\mathcal{M}}_0$ is invariant by ϕ_t.

Lemma 4.12.2. *If $(x,v) \in \tilde{\mathcal{M}}_0$ and $u \prec L + c[0]$, then, for each $t, t' \in \mathbb{R}$, with $t \leq t'$, we have*

$$u(\pi \circ \phi_{t'}(x,v)) - u(\pi \circ \phi_t(x,v)) = \int_t^{t'} L(\phi_s(x,v))\, ds + c[0](t'-t).$$

Proof. By continuity, it is enough to see it when $(x,v) \in \operatorname{supp}(\mu)$ with μ a Borel probability measure on TM, invariant by ϕ_t and such that $\int_{TM} L\, d_\mu = -c[0]$ Since $u \prec L + c[0]$, for each $(x,v) \in TM$, we have

$$u(\pi \circ \phi_{t'}(x,v)) - u(\pi \circ \phi_t(x,v)) \leq \int_t^{t'} L(\phi_s(y,w))\, ds + c[0](t'-t). \tag{$*$}$$

If we integrate this inequality with respect to μ, we find by the invariance of μ

$$\int_{TM} u \circ \pi\, d\mu - \int_{TM} u \circ \pi\, d\mu \leq (t'-t)\Big(\int L_{TM}\, d\mu + c[0]\Big),$$

which is in fact the equality $0 = 0$. It follows that the inequality $(*)$ is an equality at any point (x,v) contained in $\mathrm{supp}(\mu)$. $\square$

Theorem 4.12.3. *A function $u \in \mathcal{C}^0(M,\mathbb{R})$, such that $u \prec L + c[0]$, is differentiable at every point of the projected Mather set $\mathcal{M}_0 = \pi(\tilde{\mathcal{M}}_0)$. Moreover, if $(x,v) \in \tilde{\mathcal{M}}_0$, we have*

$$d_x u = \frac{\partial L}{\partial v}(x,v)$$

and the map $\mathcal{M}_0 \to T^*M, x \mapsto (x, d_x u)$ is locally Lipschitzian with a (local) Lipschitz constant independent of u.

Proof. If $(x,v) \in \tilde{\mathcal{M}}_0$, we set $\gamma_x(s) = \pi \circ \phi_s(x,v)$. We then have $\gamma_x(0) = x, \dot\gamma_x(0) = v$, and $(\gamma_x(s), \dot\gamma_x(s)) = \phi_s(x,v)$. In particular, by Lemma 4.12.2 above

$$u(\gamma_x(1)) - u(\gamma_x(-1)) = \int_{-1}^{1} L(\gamma_x(s), \dot\gamma_x(s))\, ds + 2c[0],$$

thus $x \in \mathcal{A}_{1,u}$, where $\mathcal{A}_{1,u}$ is the set introduced in 4.11.5. Consequently, the derivative $d_x u$ exists and is equal to $\frac{\partial L}{\partial v}(x,v)$. Moreover, the map $x \mapsto (x, d_x u)$ is locally Lipschitzian with a Lipschitz constant independent of u. $\square$

Corollary 4.12.4 (Mather)**.** *The map $\pi|\tilde{\mathcal{M}}_0 : \tilde{\mathcal{M}}_0 \to \mathcal{M}_0$ is injective. Its inverse is Lipschitzian.*

Proof. Let $u \prec L + c[0]$ be fixed, for example a weak KAM solution. By the previous Theorem **??**, the inverse of π on $\mathcal{M}_0$ is $x \mapsto \tilde{\mathcal{L}}^{-1}(x, d_x u)$, which is Lipschitz as a composition of Lipschitz functions. $\square$

The following corollary is due to Carneiro, see [Car95, Theorem 1, page 1078].

Corollary 4.12.5 (Carneiro)**.** *The Mather set $\tilde{\mathcal{M}}_0$ is contained in the energy level $c[0]$, i.e.*

$$\forall (x,v) \in \tilde{\mathcal{M}}_0, H(x, \frac{\partial L}{\partial v}(x,v)) = c[0].$$

Proof. Let u_- be a weak KAM solution. It is known that for $(x,v) \in \tilde{\mathcal{M}}_0$, the function u_- is differentiable at x and $d_x u_- = \frac{\partial L}{\partial v}(x,v)$. □

The functions of $\mathcal{S}_-$ and $\mathcal{S}_+$ are completely determined by their values on $\mathcal{M}_0$ as we show in the following theorem.

Theorem 4.12.6 (Uniqueness)**.** *Suppose that $u_-, \tilde{u}_-$ are both in $\mathcal{S}_-$ (resp. $u_+, \tilde{u}_+$ are both in $\mathcal{S}_+$). If $u_- = \tilde{u}_-$ (resp. $u_+ = \tilde{u}_+$) on $\mathcal{M}_0$, then, we have $u_- = \tilde{u}_-$ (resp. $u_+ = \tilde{u}_+$) everywhere on M.*

Proof. Let us fix $x \in M$, we can find an extremal curve $\gamma_-^x :]-\infty, 0] \to M$, with $\gamma_-^x(0) = x$ and such that

$$\forall t \geq 0,\, u_-(x) - u_-(\gamma_-^x(-t)) = \int_{-t}^0 L(\gamma_-^x(s), \dot{\gamma}_-^x(s))\, ds + c[0]t.$$

Since $\tilde{u}_- \prec L + c[0]$, we have

$$\forall t \geq 0,\, \tilde{u}_-(x) - \tilde{u}_-(\gamma_-^x(-t)) \leq \int_{-t}^0 L(\gamma_-^x(s), \dot{\gamma}_-^x(s))\, ds + c[0]t.$$

It follows that

$$\forall t \geq 0,\, \tilde{u}_-(x) - \tilde{u}_-(\gamma_-^x(-t)) \leq u_-(x) - u_-(\gamma_-^x(-t)). \qquad (*)$$

In addition, we know that $s \mapsto (\gamma_-^x(s), \dot{\gamma}_-^x(s))$ is a trajectory of the Euler-Lagrange flow ϕ_t and that the α-limit set of this trajectory carries a Borel probability measure μ invariant by ϕ_t and such that $\int_{TM} L\, d\mu = -c[0]$. The support of this measure is thus contained in $\tilde{\mathcal{M}}_0$. We conclude from it, that there exists a sequence $t_n \searrow +\infty$ such that $(\gamma_-^x(-t_n), \dot{\gamma}_-^x(-t_n))$ converges to a point $(x_\infty, v_\infty) \in \tilde{\mathcal{M}}_0$, in particular, we have $\gamma_-^x(-t_n) \to x_\infty$ which is in $\mathcal{M}_0$. It follows that $\tilde{u}_-(\gamma_-^x(-t_n)) - u_-(\gamma_-^x(-t_n))$ tends to $\tilde{u}_-(x_\infty) - u_-(x_\infty)$, which is 0 since $x_\infty \in \mathcal{M}_0$. From $(*)$, we then obtain the inequality $\tilde{u}_-(x) - u_-(x) \leq 0$. We can of course exchange the role of $\tilde{u}_-$ and that of u_- to conclude. □

4.13 Complements

If $u : M \to \mathbb{R}$ is a Lipschitz function, we will denote by $\mathrm{dom}(du)$ the domain of definition of du, i.e. the set of the points x where the derivative $d_x u$ exists. The graph of du is

$$\mathrm{Graph}(du) = \{(x, d_x u) \mid x \in \mathrm{dom}(u)\} \subset T^*M.$$

Let us recall that Rademacher's theorem 1.1.10 says that $M \setminus \mathrm{dom}(u)$ is negligible (for the Lebesgue class of measures). Since $\|d_x u\|_x$ is bounded by the Lipschitz constant of u, it is not difficult to use a compactness argument to show that the projection $\pi^*\big(\overline{\mathrm{Graph}(du)}\big)$ is the whole of M.

Lemma 4.13.1. *Suppose that $u_- \in \mathcal{S}_-$. If $x \in M$ and $\gamma_-^x :]-\infty, 0] \to M$ is such that $\gamma_-^x(0) = x$ and*

$$\forall t \geq 0,\; u_-(x) - u_-(\gamma_-^x(-t)) = \int_{-t}^{0} L(\gamma_-^x(s), \dot\gamma_-^x(s))\, ds + c[0]t,$$

then, the function u_- has a derivative at each point $\gamma_-^x(-t)$ with $t > 0$, and we have

$$\forall t > 0,\; d_{\gamma_-^x(-t)} u_- = \frac{\partial L}{\partial v}(\gamma_-^x(-t), \dot\gamma_-^x(-t)).$$

It follows that

$$\forall t, s > 0,\; (\gamma_-^x(-t-s), d_{\gamma_-^x(-t-s)} u_-) = \phi^*_{-s}(\gamma_-^x(-t), d_{\gamma_-^x(-t)} u_-).$$

We also have

$$\forall t \geq 0, H\big[\gamma_-^x(-t), \frac{\partial L}{\partial v}(\gamma_-^x(-t), \dot\gamma_-^x(-t))\big] = c[0].$$

Moreover, if u_- has a derivative at x, we have

$$d_x u_- = \frac{\partial L}{\partial v}(x, \dot\gamma_-^x(0))$$

and $d_{\gamma_-^x(-t)} u_- = \phi_t^*(x, d_x u_-)$, for each $t \geq 0$.

There is a similar statement for the functions $u_+ \in \mathcal{S}_+$.

Proof. For the first part, we notice that the curve $\gamma : [-t,t] \to M$ defined by $\gamma(s) = \gamma_-^x(s-t)$ shows that $\gamma(0) = \gamma_-^x(-t)$ is in $\mathcal{A}_{t,u_-}$, for each $t > 0$. It follows that $d_{\gamma_-^x(-t)}u_- = \frac{\partial L}{\partial v}(\gamma_-^x(-t), \dot\gamma_-^x(-t))$. The fact that $(\gamma_-^x(-t-s), d_{\gamma_-^x(-t-s)}u_-) = \phi_s^*(\gamma_-^x(-t), d_{\gamma_-^x(-t)}u_-)$ is, by Legendre transform, equivalent to $\phi_{-s}(\gamma_-^x(-t), \dot\gamma_-^x(-t)) = (\gamma_-^x(-t-s), \dot\gamma_-^x(-t-s))$, which is true since γ_-^x is an extremal curve. Let us suppose that u_- has a derivative at x. The relation $u_- \prec L + c[0]$ implies that

$$\forall v \in T_xM,\ d_xu_-(v) \leq c[0] + L(x,v).$$

Moreover, taking the derivative at $t = 0$ of the equality

$$\forall t \geq 0,\ u_-(x) - u_-(\gamma_-^x(-t) = \int_{-t}^0 L(\gamma_-^x(s), \dot\gamma_-^x(s))\, ds + c[0]t,$$

we obtain the equality $d_xu_-(\dot\gamma_-^x(0)) = c[0] + L(x, \dot\gamma_-^x(0))$. We then conclude that $H(x, d_xu_-) = c[0]$ and that $d_xu_- = \frac{\partial L}{\partial v}(x, \dot\gamma_-^x(0))$. In particular, it follows that we have $H(y, d_yu_-) = c[0]$ at any point $y \in M$ where u_- has a derivative. By the first part, without any assumption on the differentiability of u_- at x, we find that $H\big[\gamma_-^x(-t), \frac{\partial L}{\partial v}(\gamma_-^x(-t), \dot\gamma_-^x(-t))\big] = c[0]$, for each $t > 0$. By continuity this equality is also true for $t = 0$. $\square$

Theorem 4.13.2. *Let $u_- \in \mathcal{S}_-$. The derivative map $x \mapsto (x, d_xu_-)$ is continuous on its domain of definition $\operatorname{dom}(du_-)$. The sets $\operatorname{Graph}(du_-)$ and $\overline{\operatorname{Graph}(du_-)}$ are invariant by ϕ_{-t}^*, for each $t \geq 0$. Moreover, for each $(x,p) \in \overline{\operatorname{Graph}(du_-)}$, we have $H(x,p) = c[0]$.*

The closure $\overline{\operatorname{Graph}(du_-)}$ is the image by the Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ of the subset of TM formed by the $(x,v) \in TM$ such that

$$\forall t \geq 0,\ u_-(x) - u_-[\pi(\phi_{-t}(x,v))] = \int_{-t}^0 L(\phi_s(x,v))\, ds + c[0]t,$$

i.e. the set of the (x,v) such that the extremal curve $\gamma :]-\infty, 0] \to M$, with $\gamma(0) = x$ and $\dot\gamma(0) = v$, is a curve of the type γ_-^x for u_-.

Proof. The above Lemma 4.13.1 shows that $\operatorname{Graph}(du_-)$ is invariant by ϕ_{-t}^*, for each $t \geq 0$ and that $\operatorname{Graph}(du_-) \subset H^{-1}(c[0])$. Since the flow ϕ_t^* is continuous, the closure $\overline{\operatorname{Graph}(du_-)}$ is also invariant by ϕ^*_{-t}, for each $t \geq 0$. In the same way, the inclusion $\overline{\mathrm{Graph}(du_-)} \subset H^{-1}(c[0])$ results from the continuity of H. If we denote by D_- the subset of TM defined in the last part of the theorem, Lemma **??** also shows that $\mathrm{Graph}(du_-) \subset \tilde{\mathcal{L}}(D_-) \subset \overline{\mathrm{Graph}(du_-)}$. If x_n is a sequence in $\mathrm{dom}(du_-)$ and $(x_n, d_{x_n}u_-) \to (x, p)$, let us then show that

$$\forall t \geq 0,\ u_-(x) - u_-[\pi(\phi_{-t}(x,v))] = \int_{-t}^{0} L(\phi_s(x,v))\,ds + c[0]t, \quad (*)$$

where $v \in T_xM$ is defined by $p = \frac{\partial L}{\partial v}(x,v)$. For that we define $v_n \in T_{x_n}M$ by $d_{x_n}u_- = \frac{\partial L}{\partial v}(x_n, v_n)$. We have $(x_n, v_n) \to (x, v)$. By Lemma **??**, the extremal curve $\gamma^{x_n} :]-\infty, 0] \to M$ is $s \mapsto \pi(\phi_s(x_n, v_n))$, hence we obtain

$$\forall t \geq 0,\ u_-(x_n) - u_-[\pi(\phi_{-t}(x_n, v_n))] = \int_{-t}^{0} L(\phi_s(x_n, v_n))\,ds + c[0]t.$$

When we let n tend to $+\infty$, we find $(*)$. We conclude that $(x, v) \in D_-$ and thus $\tilde{\mathcal{L}}(D_-) = \overline{\mathrm{Graph}(du_-)}$. Moreover by Lemma **??**, if $x \in \mathrm{dom}(du_-)$ we necessarily have $d_xu_- = \frac{\partial L}{\partial v}(x,v) = p$. As $\mathrm{Graph}(du_-)$ is contained in the compact subset $H^{-1}(c[0])$, we then obtain the continuity of $x \mapsto (x, d_xu_-)$ on $\mathrm{dom}(du_-)$. □

We have of course a similar statement for the functions in $\mathcal{S}_+$.

Theorem 4.13.3. *If $u_+ \in \mathcal{S}_+$, the derivative map $x \mapsto (x, d_xu_+)$ is continuous on its domain of definition* $\mathrm{dom}(du_+)$.

The sets $\mathrm{Graph}(du_+)$ *and* $\overline{\mathrm{Graph}(du_+)}$ *are invariant by ϕ^*_t, for each $t \geq 0$. Moreover, for each $(x,p) \in \overline{\mathrm{Graph}(du_+)}$, we have $H(x,p) = c[0]$. The closure $\overline{\mathrm{Graph}(du_+)}$ is the image by the Legendre transform $\tilde{\mathcal{L}} : TM \to T^*M$ of the subset of TM formed by the $(x,v) \in TM$ such that*

$$\forall t \geq 0,\ u_+ \circ \pi(\phi_t(x,v)) - u_+(x) = \int_0^t L(\phi_s(x,v))\,ds + c[0]t,$$

i.e. the set of (x,v) such that the extremal curve $\gamma : [0, +\infty[\to M$, with $\gamma(0) = x$ and $\dot\gamma(0) = v$, is a curve of the type γ^x_+ for u_+.

4.14 Examples

Definition 4.14.1 (Reversible Lagrangian)**.** The Lagrangian L is said to be reversible if it satisfies $L(x,-v) = L(x,v)$, for each $(x,v) \in TM$.

Example 4.14.2. Let g be a Riemannian metric on M, we denote by $\|\cdot\|_x$ the norm deduced from g on T_xM. If $V : M \to \mathbb{R}$ is C^2, the Lagrangian L defined by $L(x,v) = \frac{1}{2}\|v\|_x^2 - V(x)$ is reversible.

Proposition 4.14.3. *For a reversible Lagrangian L, we have*

$$-c[0] = \inf_{x\in M} L(x,0) = \inf_{(x,v)\in TM} L(x,v).$$

Moreover $\tilde{\mathcal{M}}_0 = \{(x,0) \mid L(x,0) = -c[0]\}$.

Proof. By the strict convexity and the superlinearity of L in the fibers of the tangent bundle TM, we have $L(x,0) = \inf_{v\in T_xM} L(x,v)$, for all $x \in M$. Let us set $k = \inf_{x\in M} L(x,0) = \inf_{(x,v)\in TM} L(x,v)$. Since $-c[0] = \inf \int L\,d\mu$, where the infimum is taken over all Borel probability measures on TM invariant under the flow ϕ_t, we obtain $k \leq -c[0]$. Let then $x_0 \in M$ be such that $L(x_0,0) = k$, the constant curve $]-\infty,+\infty[\to M, t \mapsto x_0$ is a minimizing extremal curve. Consequently $\phi_t(x_0,0) = (x_0,0)$ and the Dirac mass $\delta_{(x_0,0)}$ is invariant by ϕ_t, but $\int L\,d\delta_{(x_0,0)} = k$. Therefore $-c[0] = k$ and $(x_0,0) \in \tilde{\mathcal{M}}_0$. Let μ be a Borel probability measure on TM such that $\int_{TM} L\,d\mu = -c[0]$. Since $-c[0] = \inf_{TM} L$, we necessarily have $L(x,v) = \inf_{TM} L$ on the support of μ. It follows that $\operatorname{supp}(\mu) \subset \{(x,0) \mid L(x,0) = -c[0]\}$. □

We then consider the case where M is the circle $\mathbb{T} = \mathbb{R}/\mathbb{Z}$. We identify the tangent bundle $T\mathbb{T}$ with $\mathbb{T}\times\mathbb{R}$. As a Lagrangian L we take one defined by $L(x,v) = \frac{1}{2}v^2 - V(x)$, where $V : \mathbb{T} \to \mathbb{R}$ is C^2. We thus have $-c[0] = \inf_{\mathbb{T}\times\mathbb{R}} L = -\sup V$, hence $c[0] = \sup V$. Let us identify $T^*\mathbb{T}$ with $\mathbb{T}\times\mathbb{R}$. The Hamiltonian H is given by $H(x,p) = \frac{1}{2}p^2 + V(x)$. The differential equation on $T^*\mathbb{T}$ which defines the flow ϕ_t is given by $\dot{x} = p$ and $\dot{p} = -V'(x)$. If $u_- \in \mathcal{S}_-$, the compact subset $\overline{\operatorname{Graph}(du_-)}$ is contained in the set $H^{-1}(c[0]) = \{(x,p) \mid p = \pm\sqrt{\sup V - V(x)}\}$. We strongly encourage the reader to do some drawings **FAIRE**

DES DESSINS of the situation in $\mathbb{R}\times\mathbb{R}$, the universal cover of $\mathbb{T}\times\mathbb{R}$. To describe u_- completely let us consider the case where V reaches its maximum only at 0. In this case the set $H^{-1}(c[0])$ consists of three orbits of ϕ_t^*, namely the fixed point $(0,0)$, the orbit $\mathcal{O}_+ = \{(x, \sqrt{\sup V - V(x)}) \mid x \neq 0\}$ and the orbit $\mathcal{O}_- = \{(x, -\sqrt{\sup V - V(x)}) \mid x \neq 0\}$. On $\mathcal{O}_+$ the direction of the increasing t is that of the increasing x (we identify in a natural way $\mathbb{T}\setminus 0$ with $]0,1[$). On $\mathcal{O}_-$ the direction of the increasing t is that of the decreasing x. Since $\overline{\mathrm{Graph}(du_-)}$ is invariant by the maps ϕ^*_{-t}, for $t \geq 0$, if $(x, \sqrt{\sup V - V(x)}) \in \overline{\mathrm{Graph}(du_-)}$, then we must have $(y, \sqrt{\sup V - V(y)}) \in \overline{\mathrm{Graph}(du_-)}$, for each $y \in]0,x]$. By symmetry we get $(y, -\sqrt{\sup V - V(y)}) \in \overline{\mathrm{Graph}(du_-)}$, for each $y \in [x,1[$. It follows that there is a point x_0 such that $\overline{\mathrm{Graph}(du_-)}$ is the union of $(0,0)$ and the two sets $\{(y, \sqrt{\sup V - V(y)}) \mid y \in]0,x_0]\}$ and $\{(y, -\sqrt{\sup V - V(y)}) \mid y \in [x_0,1[\}$. Moreover, since the function u_- is defined on $\mathbb{T}$, we have $\lim_{x\to 1} u_-(x) = u_-(0)$ and thus the integral on $]0,1[$ of the derivative of u_- must be 0. This gives the relation

$$\int_0^{x_0} \sqrt{\sup V - V(x)}\, dx = \int_{x_0}^1 \sqrt{\sup V - V(x)}\, dx.$$

This equality determines completely a unique point x_0, since $\sup V - V(x) > 0$ for $x \in]0,1[$. In this case, we see that u_- is unique up to an additive constant and that

$$u_-(x) = \begin{cases} u_-(0) + \int_0^x \sqrt{\sup V - V(x)}\, dx, & \text{if } x \in [0,x_0]; \\ u_-(0) + \int_x^1 \sqrt{\sup V - V(x)}\, dx, & \text{if } x \in [x_0,1]. \end{cases}$$

Exercise 4.14.4. *1) If $V : \mathbb{T} \to \mathbb{R}$ reaches its maximum exactly n times, show that the solutions u_- depend on n real parameters, one of these parameters being an additive constant.*

2) Describe the Mather function $\alpha : H^1(\mathbb{T},\mathbb{R}) \to \mathbb{R}, \Omega \mapsto c[\Omega]$.

3) If ω is a closed differential 1-form on $\mathbb{T}$, describe the function u_-^ω for the Lagrangian L_ω defined by $L(x,v) = \frac{1}{2}v^2 - V(x) - \omega_x(v)$.

Chapter 5

Conjugate Weak KAM Solutions

In this chapter, as in the previous ones, we denote by M a **compact and connected** manifold. The projection of TM on M is denoted by $\pi : TM \to M$. We suppose given a C^r Lagrangian $L : TM \to \mathbb{R}$, with $r \geq 2$, such that, for each $(x, v) \in TM$, the second vertical derivative $\frac{\partial^2 L}{\partial v^2}(x, v)$ is definite > 0 as a quadratic form, and that L is superlinear in each fiber of the tangent bundle $\pi : TM \to M$. We will also endow M with a fixed Riemannian metric. We denote by d the distance on M associated with this Riemannian metric. If $x \in M$, the norm $\|\cdot\|_x$ on T_xM is the one induced by this same Riemannian metric.

5.1 Conjugate Weak KAM Solutions

We start with the following lemma

Lemma 5.1.1. *If $u \prec L + c[0]$, then we have*

$$\forall x \in \mathcal{M}_0, \forall t \geq 0,\ u(x) = T_t^- u(x) + c[0]t = T_t^+ u(x) - c[0]t.$$

Proof. Since $u \prec L + c[0]$, we have $u \leq T_t^- u + c[0]t$ and $u \geq T_t^+ u - c[0]t$. We consider the point $(x, v) \in \tilde{\mathcal{M}}_0$ above x. Let us note by $\gamma :]-\infty, +\infty[\to M$ the extremal curve $s \mapsto \pi(\phi_s(x, v))$.

By lemma 4.12.2, for each $t \geq 0$, we have

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

Since $\gamma(0) = x$, we obtain the inequalities $u(x) \geq T_t^- u(x) + c[0]t$ and $u(x) \leq T_t^+ u(x) - c[0]t$. □

Theorem 5.1.2 (Existence of Conjugate Pairs)**.** *If $u : M \to \mathbb{R}$ is a function such that $u \prec L + c[0]$, then, there exists a unique function $u_- \in \mathcal{S}_-$ (resp. $u_+ \in \mathcal{S}_+$) with $u = u_-$ (resp. $u = u_+$) on the projected Mather set $\mathcal{M}_0$. These functions verify the following properties*

(1) *we have $u_+ \leq u \leq u_-$;*

(2) *if $u_-^1 \in \mathcal{S}_-$ (resp. $u_+^1 \in \mathcal{S}_+$) verifies $u \leq u_-^1$ (resp. $u_+^1 \leq u$), then $u_- \leq u_-^1$ (resp. $u_+^1 \leq u_+$);*

(3) *We have $u_- = \lim_{t\to+\infty} T_t^- u + c[0]t$ and $u_+ = \lim_{t\to+\infty} T_t^+ u - c[0]t$, the convergence being uniform on M.*

Proof. It will be simpler to consider the modified semigroup $\hat{T}_t^- v = T_t^- v + c[0]t$. The elements of $\mathcal{S}_-$ are precisely the fixed points of the semigroup $\hat{T}_t^-$. The condition $u \prec L + c[0]$ is equivalent to $u \leq \hat{T}_t^- u$. As $\hat{T}_t^-$ preserves the order, we see that $\hat{T}_t^- u \leq u_-^1$ for each $u_-^1 \in \mathcal{S}_-$ satisfying $u \leq u_-^1$. As $\hat{T}_t^- u = u$ on the projected Mather set $\mathcal{M}_0$, it then remains to show that $\hat{T}_t^- u$ is uniformly convergent for $t \to \infty$. However, we have $\hat{T}_t^- u \leq \hat{T}_{t+s}^- u$, if $s \geq 0$, because this is true for $t = 0$ and the semigroup $\hat{T}_t^-$ preserves the order. Since for $t \geq 1$ the family of maps $\hat{T}_t^- u$ is equi-Lipschitzian, it is enough to see that this family of maps is uniformly bounded. To show this uniform boundedness, we fix $u_-^0 \in \mathcal{S}_-$, by compactness of M, there exists $k \in \mathbb{R}$ such that $u \leq u_-^0 + k$. By what was already shown, we have $\hat{T}_t^- u \leq u_-^0 + k$. □

Corollary 5.1.3. *For any function $u_- \in \mathcal{S}_-$ (resp. $u_+ \in \mathcal{S}_+$), there exists one and only one function of $u_+ \in \mathcal{S}_+$ (resp. $u_- \in \mathcal{S}_-$) satisfying $u_+ = u_-$ on $\mathcal{M}_0$. Moreover, we have $u_+ \leq u_-$ on all M.*

Definition 5.1.4 (Conjugate Functions)**.** A pair of functions (u_-, u_+) is said to be conjugate if $u_- \in \mathcal{S}_-, u_+ \in \mathcal{S}_+$ and $u_- = u_+$ on $\mathcal{M}_0$. We will denote by $\mathcal{D}$ the set formed by the differences $u_- - u_+$ of pairs (u_-, u_+) of conjugate functions.

The following lemma will be useful in the sequel.

Lemma 5.1.5 (Compactness of the Differences)**.** *All the functions in $\mathcal{D}$ are ≥ 0. Moreover, the subset $\mathcal{D}$ is compact in $\mathcal{C}^0(M, \mathbb{R})$ for the topology of uniform convergence.*

Proof. If u_- and u_+ are conjugate, we then know that $u_+ \leq u_-$ and thus $u_- - u_+ \geq 0$. If we fix $x_0 \in M$, the set $\mathcal{S}_-^{x_0} = \{u_- \mid u_-(x_0) = 0\}$ (resp. $\mathcal{S}_+^{x_0} = \{u_+ \mid u_+(x_0) = 0\}$) is compact, since it is a family of equi-Lipschitzian functions on the compact manifold M which all vanishes at the point x_0. However, for $c \in \mathbb{R}$, it is obvious that the pair (u_-, u_+) is conjugate if and only if the pair $(u_- + c, u_+ + c)$ is conjugate. We conclude that $\mathcal{D}$ is the subset of the compact subset $\mathcal{S}_-^{x_0} - \mathcal{S}_+^{x_0}$ formed by the functions which vanish on $\mathcal{M}_0$. □

Corollary 5.1.6. *Let us suppose that all the functions $u_- \in \mathcal{S}_-$ are C^1 (what is equivalent to $\mathcal{S}_- = \mathcal{S}_+$). Then, two arbitrary functions in $\mathcal{S}_-$ differ by a constant.*

Proof. Conjugate functions are then equal, because the C^1 functions contained in $\mathcal{S}_-$ or $\mathcal{S}_+$ are also in $\mathcal{S}_- \cap \mathcal{S}_+$ by 4.11.8. Suppose then that u_-^1 and u_-^2 are two functions in $\mathcal{S}_-$. We of course do have $u = (u_-^1 + u_-^2)/2 \prec L + c[0]$. By the Theorem of Existence of Conjugate Pairs 5.1.2, we can find a pair of conjugate functions (u_-, u_+) with $u_+ \leq u \leq u_-$. As conjugate functions are equal, we have $u = (u_-^1 + u_-^2)/2 \in \mathcal{S}_-$. The three functions u, u_-^1 and u_-^2 are C^1 and in $\mathcal{S}_-$, we must then have $H(x, d_x(u_-^1 + u_-^2)/2) = H(x, d_x u_-^1) = H(x, d_x u_-^2) = c[0]$, for each $x \in M$. This is compatible with the strict convexity of H in fibers of T^*M only if $d_x u_-^1 = d_x u_-^2$, for each $x \in M$. □

5.2 Aubry Set and Mañé Set.

Definition 5.2.1 (The Set $\mathcal{I}_{(u_-,u_+)}$)**.** Let us consider a pair (u_-, u_+) of conjugate functions. We denote by $\mathcal{I}_{(u_-,u_+)}$, the set

$$\mathcal{I}_{(u_-,u_+)} = \{x \in M \mid u_-(x) = u_+(x)\}.$$

We have $\mathcal{I}_{(u_-,u_+)} \supset \mathcal{M}_0$.

Theorem 5.2.2. *For each $x \in \mathcal{I}_{(u_-,u_+)}$, there exists an extremal curve a $\gamma^x :]-\infty, +\infty[\to M$, with $\gamma^x(0) = x, \dot\gamma^x(0) = v$ and such that, for each $t \in \mathbb{R}$, we have $u_-(\pi[\phi_t(x,v)]) = u_+(\pi[\phi_t(x,v)])$ and*

$$\forall t \leq t' \in \mathbb{R},\ u_\pm(\gamma^x(t')) - u_\pm(\gamma^x(t)) = \int_t^{t'} L(\gamma^x(s), \dot\gamma^x(s))\, ds + c[0](t'-t).$$

It follows that the functions u_- and u_+ are differentiable at every point of $\mathcal{I}_{(u_-,u_+)}$ with the same derivative. Moreover, there exists a constant K which depends only on L and such that the section $\mathcal{I}_{(u_-,u_+)} \to T^*M, x \mapsto d_x u_- = d_x u_+$ is Lipschitzian with Lipschitz constant $\leq K$.

Proof. Let us fix $x \in \mathcal{I}_{(u_-,u_+)}$. There exists extremal curves $\gamma_-^x :]-\infty, 0] \to M$ and $\gamma_+^x : [0, +\infty[\to M$ with $\gamma_-^x(0) = \gamma_+^x(0) = x$ and for each $t \in [0, +\infty[$

$$u_+(\gamma_+^x(t)) - u_+(x) = c[0]t + \int_0^t L(\gamma_+^x(s), \dot\gamma_+^x(s))\, ds,$$

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

But, since $u_- \prec L + c[0], u_+ \prec L + c[0], u_+ \leq u_-$ and $u_-(x) = u_+(x)$, we have

$$u_+(\gamma_+^x(t)) - u_+(x) \leq u_-(\gamma_+^x(t)) - u_-(x)$$

$$\leq c[0]t + \int_0^t L(\gamma_+^x(s), \dot\gamma_+^x(s))\, ds,$$

and

$$\begin{aligned} u_-(x) - u_-(\gamma_-^x(-t)) &\leq u_+(x) - u_+(\gamma_-^x(-t)) \\ &\leq c[0]t + \int_{-t}^0 L(\gamma_-^x(s), \dot\gamma_-^x(s))\, ds. \end{aligned}$$

We thus have equality everywhere. It is not difficult to deduce that the curve γ^x which is equal to γ_-^x on $]-\infty, 0]$ and to γ_+^x on $[0, +\infty[$ satisfies

$$\forall t \leq t' \in \mathbb{R},\ u_\pm(\gamma^x t') - u_\pm(\gamma^x t) = \int_t^{t'} L(\gamma^x(s), \dot\gamma^x(s))\, ds + c[0](t'-t).$$

It follows that γ^x is the sought extremal curve. The existence of γ^x shows that $x \in \mathcal{A}_{1,u_-}$ and $x \in \mathcal{A}_{1,u_+}$ and thus u_- and u_+ are differentiable at x with $d_x u_- = d_x u_+ = \frac{\partial L}{\partial v}(x, \dot\gamma^x(0))$. The existence of K also results from $x \in \mathcal{A}_{1,u_-}$ (or $x \in \mathcal{A}_{1,u_+}$). $\square$

Definition 5.2.3 (The set $\tilde{\mathcal{I}}_{(u_-,u_+)}$)**.** If (u_-, u_+) is a pair of conjugate functions, we define the set $\tilde{\mathcal{I}}_{(u_-,u_+)}$ by

$$\tilde{\mathcal{I}}_{(u_-,u_+)} = \{(x, v) \mid x \in \mathcal{I}_{(u_-,u_+)}, d_x u_- = d_x u_+ = \frac{\partial L}{\partial v}(x, v)\}.$$

Theorem 5.2.4. *If (u_-, u_+) is a pair of conjugate functions, the projection $\pi : TM \to$ induces a bi-Lipschitzian homeomorphism $\tilde{\mathcal{I}}_{(u_-,u_+)}$ on $\mathcal{I}_{(u_-,u_+)}$. The set $\tilde{\mathcal{I}}_{(u_-,u_+)}$ is compact and invariant by the Euler-Lagrange flow ϕ_t. It contains $\tilde{\mathcal{M}}_0$. If $(x, v) \in \tilde{\mathcal{I}}_{(u_-,u_+)}$, for each $t \in \mathbb{R}$, we have $u_-(\pi[\phi_t(x, v)]) = u_+(\pi[\phi_t(x, v)])$ and for all $t \leq t' \in \mathbb{R}$*

$$u_\pm(\pi[\phi_{t'}(x, v)]) - u_\pm(\pi[\phi_t(x, v)]) = \int_t^{t'} L(\phi_s(x, v))\, ds + c[0](t'-t).$$

Proof. By the previous theorem 5.2.2 the projection π restricted to $\tilde{\mathcal{I}}_{(u_-,u_+)}$ is surjective onto $\mathcal{I}_{(u_-,u_+)}$ with a Lipschitzian inverse. In particular, the set $\tilde{\mathcal{I}}_{(u_-,u_+)}$ is compact. Moreover, if $(x, v) \in \tilde{\mathcal{I}}_{(u_-,u_+)}$ and γ_x is the extremal curve given by the previous theorem, we have $v = \dot\gamma^x(0)$, and, for each $s \in \mathbb{R}$, we have $(\gamma^x(s), \dot\gamma^x(s)) \in \tilde{\mathcal{I}}_{(u_-,u_+)}$, because $\gamma^x(s) \in \mathcal{I}_{(u_-,u_+)}$ and the extremal curve $t \mapsto \gamma^x(s+t)$ can be used as $\gamma^{\gamma^x(s)}$. Since $(\gamma^x(s), \dot\gamma^x(s)) = \phi_s(x, v)$, it follows that $\tilde{\mathcal{I}}_{(u_-,u_+)}$ is invariant by the flow ϕ_t. From theorem 4.12.3, if $x \in \mathcal{M}_0$, we have $d_x u\pm = \partial L/\partial v(x, v)$, where (x, v) is the point $\tilde{\mathcal{M}}_0$ above x. Since $\mathcal{M}_0 \subset \mathcal{I}_{(u_-,u_+)}$, the definition of $\tilde{\mathcal{I}}_{(u_-,u_+)}$ implies $(x, v) \in \tilde{\mathcal{I}}_{(u_-,u_+)}$. $\square$

The Mañé set was introduced in [Mn97, page 144], where it is denoted $\Sigma(L)$.

Definition 5.2.5 (Mañé Set)**.** The Mañé set is

$$\tilde{\mathcal{N}}_0 = \bigcup \tilde{\mathcal{I}}_{(u_-,u_+)},$$

where the union is taken over all pairs (u_-, u_+) of conjugate functions.

Proposition 5.2.6. *The Mañé set $\tilde{\mathcal{N}}_0$ is a compact subset of TM which is invariant by ϕ_t. It contains $\tilde{\mathcal{M}}_0$.*

Definition 5.2.7 (Aubry Set)**.** The Aubry Set in TM is

$$\tilde{\mathcal{A}}_0 = \bigcap \tilde{\mathcal{I}}_{(u_-,u_+)},$$

where the intersection is taken on the pairs (u_-, u_+) of conjugate functions. The projected Aubry set in M is $\mathcal{A}_0 = \pi(\tilde{\mathcal{A}}_0)$.

Theorem 5.2.8. *The Aubry sets $\mathcal{A}_0$ and $\tilde{\mathcal{A}}_0$ are both compact, and satisfy $\tilde{\mathcal{M}}_0 \subset \tilde{\mathcal{A}}_0$ and $\mathcal{M}_0 \subset \mathcal{A}_0$. The compact set $\tilde{\mathcal{A}}_0 \subset TM$ is invariant by the Euler-Lagrange flow ϕ_t.*

Moreover, there is a pair (u_-, u_+) of conjugate functions such that $\mathcal{A}_0 = \mathcal{I}_{(u_-,u_+)}$ and $\tilde{\mathcal{A}}_0 = \tilde{\mathcal{I}}_{(u_-,u_+)}$.

Therefore the projection $\pi : TM \to M$ induces a bi-Lipschitz homeomorphism $\tilde{\mathcal{A}}_0$ on $\mathcal{A}_0 = \pi(\tilde{\mathcal{A}}_0)$.

Proof. The first part of the theorem is a consequence of the same properties which hold true for $\tilde{\mathcal{I}}_{(u_-,u_+)}$ and $\mathcal{I}_{(u_-,u_+)}$ which are true by theorem 5.2.4. The last part of the theorem is, again by 5.2.4, a consequence of the second part.

It remains to prove the second part. We fix a base point $x_0 \in \mathcal{M}_0$, any pair of conjugate function is of the form $(u_- + c, u_+ + c)$, where (u_-, u_+) is a pair of conjugate functions with $u_-(x_0) = u_+(x_0) = 0$, and $c \in \mathbb{R}$. Using the fact that $\mathcal{C}(M, \mathbb{R})$ is metric and separable (i.e. contains a dense sequence) for the topology of uniform convergence, we can find a sequence of pairs of conjugate functions $(u_-^n + c_n, u_+^n c_n)$, dense in the set of pairs of conjugate functions, and such that $u_-^n(x_0) = u_+^n(x_0) = 0, c_n \in \mathbb{R}$. Since the sets $\mathcal{S}_-$ and $\mathcal{S}_+$ form equi-Lipschitzian families of functions on the compact space M, and $u_-^n(x_0) = u_+^n(x_0) = 0$, we can find a constant $C < +\infty$ such that $\|u_-^n\|_\infty \leq C$ and $\|u_+^n\|_\infty \leq C$, for each $n \geq 0$. It follows that the series $\sum_{n\geq 0} 2^{-n-1} u_-^n$ converges to a continuous function. The sum is dominated by $L + c[0]$, because this is the case for each u_-^n and $\sum_{n\geq 0} 2^{-n-1} = 1$. By theorem 5.1.2, we can thus find $u_- \in \mathcal{S}_-$ with $u_- \geq \sum_{n\geq 0} 2^{-n} u_-^n$ and $u_- = \sum_{n\geq 0} 2^{-n} u_-^n$ on $\mathcal{M}_0$. In the same way, we can find $u_+ \in \mathcal{S}_+$ with $u_+ \leq \sum_{n\geq 0} 2^{-n} u_+^n$ and $u_+ = \sum_{n\geq 0} 2^{-n} u_+^n$ on $\mathcal{M}_0$. Since $u_+^n \leq u_-^n$ with equality on $\mathcal{M}_0$, we see that

$$u_+ \leq \sum_{n\geq 0} 2^{-n} u_+^n \leq \sum_{n\geq 0} 2^{-n} u_-^n \leq u_-,$$

with equalities on $\mathcal{M}_0$. It follows that functions u_- and u_+ are conjugate. Moreover, if $u_-(x) = u_+(x)$, we necessarily have $u_-^n(x) = u_+^n(x)$ for each $n \geq 0$. By density of the sequence $(u_-^n + c_n, u_+^n c_n)$ we conclude that for each pair $v_-, v_+)$ of conjugate functions, we have $\mathcal{I}_{(u_-,u_+)} \subset \mathcal{I}_{(v_-,v_+)}$. Therefore shows that $\mathcal{I}_{(u_-,u_+)} = \mathcal{A}_0$.

If $(x, v) \in \tilde{\mathcal{I}}_{(u_-,u_+)}$, the curve $\gamma(s) = \pi(\phi_s(x, v))$ is contained in $\mathcal{I}_{(u_-,u_+)} = \mathcal{A}_0$, and $(u_\pm, L, c[0])$-calibrated. Therefore, for example

$$\forall t \leq t' \in \mathbb{R}, u_-(\gamma(t')) - u_-(\gamma(t)) = \int_t^{t'} L(\gamma(s), \dot\gamma(s))\, ds + c[0](t'-t).. \tag{*}$$

Since $u_n \prec L + c[0]$, for each $n \geq 1$, we also have

$$\forall t \leq t' \in \mathbb{R}, u_-^n(\gamma(t')) - u_-^n(\gamma(t)) \leq \ int_t^{t'} L(\gamma(s), \dot\gamma(s))\, ds + c[0](t'-t).. \tag{**}$$

From what we established we have

$$\forall y \in \mathcal{I}_{(u_-,u_+)}, u_-(y) = \sum_{n\geq 1} \frac{u_-^n(y)}{2^n}.. \tag{***}$$

The equalities (*) and (***), taken with the fact that the image of γ is contained in $\mathcal{I}_{(u_-,u_+)}$, do imply that the inequality (**) is in fact an equality, which means that γ is $(u_-^n, L, c[0])$-calibrated, for every $n \geq 1$. By denseness of the sequence $(u_-^n + c_n)$ in $\mathcal{S}_-$, we obtain that γ is $(v_-, L, c[0])$-calibrated, for every $v_- \in \mathcal{S}_-$. Therefore $d_x v_-$ is the Legendre transform of $(x, v) = (\gamma(0), \dot\gamma(0))$. Since $x \in \mathcal{A}_0 \subset \mathcal{I}_{(v_-,v_+)}$, this implies that $(x, v) \in \tilde{\mathcal{I}}_{(v_-,v_+)}$. It follows easily that $\tilde{\mathcal{I}}_{(u_-,u_+)} = \tilde{\mathcal{A}}_0$. $\square$

5.3 The Peierls barrier.

This definition of the Peierls barrier is due to Mather, see [Mat93, §7, page 1372].

Definition 5.3.1 (Peierls Barrier)**.** The Peierls barrier is the function $h : M \to \mathbb{R}$ defined by

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

It is not completely clear that h is finite nor that it is continuous. We start by showing these two points.

Lemma 5.3.2 (Properties of h_t)**.** *The properties of h_t are*

(1) *for each $x, y, z \in M$ and each $t, t' > 0$, we have*

$$h_t(x, y) + h_{t'}(y, z) \geq h_{t+t'}(x, z);$$

(2) *if $u \prec L + c$, we have $h_t(x, y) + ct \geq u(y) - u(x)$;*

(3) *for each $t > 0$ and each $x \in M$, we have $h_t(x, x) + c[0]t \geq 0$;*

(4) *for each $t_0 > 0$ and each $u_- \in \mathcal{S}_-$, there exists a constant C_{t_0,u_-} such that*

$$\forall t \geq t_0, \forall x, y \in M,\ -2\|u_-\|_\infty \leq h_t(x, y)+c[0]t \leq 2\|u_-\|_\infty+C_{t_0,u_-};$$

(5) *for each $t > 0$ and each $x, y \in M$, there exists an extremal curve $\gamma : [0, t] \to M$ with $\gamma(0) = x, \gamma(t) = y$ and $h_t(x, y) = \int_0^t L(\gamma(s), \dot\gamma(s))\,ds$. Moreover, an extremal curve $\gamma : [0, t] \to M$ is minimizing if and only if $h_t(\gamma(0), \gamma(t)) = \int_0^t L(\gamma(s), \dot\gamma(s))\,ds$;*

(6) *for each $t_0 > 0$, there exists a constant $K_{t_0} \in [0, +\infty[$ such that, for each $t \geq t_0$ the function $h_t : M \times M \to \mathbb{R}$ is Lipschitzian with a Lipschitz constant $\leq K_{t_0}$.*

Proof. Properties (1) and (2) are immediate, and property (3) results from (2) taking for u a function in $\mathcal{S}_-$.

To prove property (4), we first remark that the inequality $-2\|u_-\|_\infty \leq h_t(x, y) + c[0]t$ also results from (2). By compactness of M, we can find a constant C_{t_0} such that for each $x, z \in M$, there exists a C^1 curve $\gamma_{x,z} : [0, t_0] \to M$ with $\gamma_{x,z}(0) = x, \gamma_{x,z}(t_0) = z$, and $\int_0^{t_0} L(\gamma_{x,z}(s), \dot\gamma_{x,z}(s))\, ds \le C_{t_0}$. By the properties of u_-, we can find an extremal curve $\gamma_-^y :]-\infty, 0] \to M$, with $\gamma_-^y(0) = y$, and

$$\forall t \ge 0, u_-(y) - u_-(\gamma_-^y(-t)) = \int_{-t}^0 L(\gamma_-^y(s), \dot\gamma_-^y(s))\, ds + c[0]t.$$

If $t \ge t_0$, we can define a (continuous) piecewise C^1 curve $\gamma : [0, t]$ by $\gamma(s) = \gamma_{x, \gamma_-^y(t_0 - t)}(s)$, for $s \in [0, t_0]$, and $\gamma(s) = \gamma_-^y(s - t)$, for $s \in [t_0, t]$. This curve γ joins x with y, and we have

$$\int_0^t L(\gamma(s), \dot\gamma(s))\, ds + c[0]t \le C_{t_0} + c[0]t_0 + u_-(y) - u_-(\gamma_-^y(t_0 - t)).$$

It is then enough to set $C_{t_0, u_-} = C_{t_0} + c[0]t_0$ to finish the proof of (4).

The first part of the property (5) results from Tonelli's Theorem 3.3.1. The second part is immediate starting from the definitions.

To prove property (6), suppose that $\gamma : [0, t] \to M$ is an extremal curve such that $\gamma(0) = x, \gamma(t) = y$, and $h_t(x, y) = \int_0^t L(\gamma(s), \dot\gamma(s))\, ds$. Since $t \ge t_0$, we know by the Compactness Lemma that there exists a compact subset K of TM with $(\gamma(s), \dot\gamma(s)) \in K$ for each $x, y \in M$, each $t \ge t_0$ and each $s \in [0, t]$. It is then enough to adapt the ideas which made it possible to show that the family $\{T_t^- u \mid t \ge t_0, u \in \mathcal{C}^0(M, \mathbb{R})\}$ is equi-Lipschitzian. $\square$

Corollary 5.3.3 (Properties of h). *The values of the map h are finite. Moreover, the following properties hold*

(1) *the map h is Lipschitzian;*

(2) *if $u \prec L + c[0]$, we have $h(x, y) \ge u(y) - u(x)$;*

(3) *for each $x \in M$, we have $h(x, x) \ge 0$;*

(4) *$h(x, y) + h(y, z) \ge h(x, z)$;*

(5) *$h(x, y) + h(y, x) \ge 0$;*

(6) *for* $x \in \mathcal{M}_0$, *we have* $h(x,x) = 0$;

(7) *for each* $x, y \in M$, *there exists a sequence of minimizing extremal curves* $\gamma_n : [0, t_n] \to M$ *with* $t_n \to \infty, \gamma_n(0) = x, \gamma_n(t_n) = y$ *and*

$$h(x,y) = \lim_{n\to+\infty} \int_0^{t_n} L(\gamma_n(s), \dot\gamma_n(s))\, ds + c[0]t_n;$$

(8) *if* $\gamma_n : [0, t_n] \to M$ *is a sequence of (continuous) piecewise* C^1 *curves with* $t_n \to \infty, \gamma_n(0) \to x$, *and* $\gamma_n(t_n) \to y$, *then we have*

$$h(x,y) \le \liminf_{n\to+\infty} \int_0^{t_n} L(\gamma_n(s), \dot\gamma_n(s))\, ds + c[0]t_n.$$

Proof. Properties (1) to (5) are easy consequences of the lemma giving the properties of h_t 5.3.2. Let us show the property (6). By the continuity of h, it is enough to show that if μ is a Borel probability measure on TM, invariant by ϕ_t and such that $\int_T ML\, d\mu = -c[0]$, then, for each $(x,v) \in \operatorname{supp}(\mu)$, the support of μ, we have $h(x,x) = 0$. By Poincaré's Recurrence Theorem, the recurrent points for ϕ_t contained in $\operatorname{supp}(\mu)$ form a dense set in $\operatorname{supp}(\mu)$. By continuity of h, we can thus assume that (x,v) is a recurrent point for ϕ_t. Let us fix $u_- \in \mathcal{S}_-$. We have

$$u_-(\pi\phi_t(x,v)) - u_-(x) = \int_0^t L(\phi_s(x,v))\, ds + c[0]t$$

By the definition of a recurrent point, there exists a sequence $t_n \to \infty$ with $\phi_{t_n}(x,v) \to (x,v)$, it is not difficult, for each $\epsilon > 0$ and each $t' \ge 0$, to find a (continuous) piecewise C^1 curve $\gamma : [0,t] \to M$, with $t \ge t', \gamma(0) = \gamma(t) = x$, and such that $\int_0^t L(\gamma(s), \dot\gamma(s))\, ds + c[0]t \le \epsilon$. Consequently, we obtain $h(x,x) \le 0$. The inequality $h(x,x) \ge 0$ is true for each $x \in M$.

Property (7) results from part (5) of the lemma giving the properties of h_t, since there is a sequence $t_n \to +\infty$ such that $h(x,y) = \lim_{t_n\to+\infty} h_{t_n}(x,y) + c[0]t_n$.

Let us show property (8). By the previous lemma, there is a constant K_1such that

$$\forall t \ge 1, \forall x, x', y, y' \in M,\ |h_t(x,y) - h_t(x',y')| \le K_1(d(x,x') + d(y,y')).$$

In addition, we also have

$$h_{t_n}(\gamma_n(0),\gamma_n(t_n)) \le \int_0^{t_n} L(\gamma_n(s),\dot\gamma_n(s))\,ds + c[0]t_n.$$

For n large, it follows that

$$h_{t_n}(x,y)+c[0]t_n \le h_{t_n}(\gamma_n(0),\gamma_n(t_n))+c[0]t_n+K_1(d(x,\gamma_n(0))+d(y,\gamma_n(t_n)))$$

$$\le c[0]t_n+K_1(d(x,\gamma_n(0))+d(y,\gamma_n(t_n)))+\int_0^{t_n} L(\gamma_n(s),\dot\gamma_n(s))\,ds+c[0]t_n.$$

Since $d(x,\gamma_n(0))+d(y,\gamma_n(t_n))\to 0$, we obtain the sought inequality. □

The following lemma is useful.

Lemma 5.3.4. *Let V be an open neighborhood of $\tilde{\mathcal{M}}_0$ in TM. There exists $t(V)>0$ with the following property:*

If $\gamma:[0,t]\to M$ is a minimizing extremal curve, with $t\ge t(V)$, then, we can find $s\in[0,t]$ with $(\gamma(s),\dot\gamma(s))\in V$.

Proof. If the lemma were not true, we could find a sequence of extremal minimizing curves $\gamma_i:[0,t_i]\to M$, with $t_i\to\infty$, and such that $(\gamma_i(s),\dot\gamma_i(s))\notin V$, for each $s\in[0,t_i]$. Since $t_i\to+\infty$, by corollary 4.4.5, there exists a compact subset $K\subset TM$ with $(\gamma_i(s),\dot\gamma_i(s))\in K$, for each $s\in[0,t_i]$ and each $i\ge 0$. We then consider the sequence of probability measures μ_n on TM defined by

$$\int_{TM}\theta\,d\mu_n=\frac{1}{t_n}\int_0^{t_n}\theta(\gamma_n(s),\dot\gamma_n(s))\,ds,$$

for $\theta: TM\to\mathbb{R}$ continuous. All the supports of these measures are contained in the compact subset K of TM. Extracting a subsequence, we can assume that μ_n converge weakly to a probability measure μ. Since $(\gamma_n(s),\dot\gamma_n(s)), s\in[0,t_n]$ are pieces of orbits of the flow ϕ_t, and since $t_n\to+\infty$, the measure μ is invariant by ϕ_t. Moreover, its support $\operatorname{supp}(\mu)$ is contained in $TM\setminus V$, because this is the case for all $\operatorname{supp}(\mu_n)=\{(\gamma_n(s),\dot\gamma_n(s))\mid s\in[0,t_n\}$. Since the γ_n are minimizing extremals, we have $\int L\,d\mu_n=h_{t_n}(\gamma_n(0),\gamma(t_n))/t_n$. By the lemma giving the properties of h_t 5.3.2, if $u_-\in\mathcal{S}_-$, we can find a constant C_1 such that

$$\forall t\ge 1,\forall x,y\in M, -2\|u_-\|_0\le h_t(x,y)+c[0]t\le 2\|u_-\|_0+C_1.$$

It follows that $\lim_{n\to+\infty}\int L\,d\mu_n = -c[0]$. Hence $\int_{TM} L\,d\mu = -c[0]$ and the support of μ is included in Mather's set $\tilde{\mathcal{M}}_0$. This is a contradiction, since we have already observed that $\mathrm{supp}(\mu)$ is disjoint from the open set V which contains $\tilde{\mathcal{M}}_0$. $\square$

Corollary 5.3.5. *For each pair $u_- \in \mathcal{S}_-, u_+ \in \mathcal{S}_+$ of conjugate functions, we have*

$$\forall x, y \in M,\ u_-(y) - u_+(x) \le h(x, y).$$

Proof. We pick a sequence of extremals $\gamma_n : [0, t_n] \to M$ joining x to y, and such that

$$h(x, y) = \lim_{n\to\infty} \int_0^{t_n} L(\gamma_n(s), \dot\gamma_n(s))\,ds + c[0]t_n.$$

By the previous lemma 5.3.4, extracting a subsequence if necessary, we can find a sequence $t'_n \in [0, t_n]$ such that $\gamma_n(t'_n) \to z \in \mathcal{M}_0$. If $u_- \in \mathcal{S}_-$, and $u_+ \in \mathcal{S}_+$, we have

$$\begin{aligned} u_+(\gamma_n(t'_n)) - u_+(x) &\le \int_0^{t'_n} L(\gamma_n(s), \dot\gamma_n(s))\,ds + c[0]t'_n, \\ u_-(y) - u_-(\gamma_n(t'_n)) &\le \int_{t'_n}^{t_n} L(\gamma_n(s), \dot\gamma_n(s))\,ds + c[0](t_n - t'_n). \end{aligned}$$

If we add these inequalities, and we let n go to $+\infty$, we find

$$u_-(y) - u_-(z) + u_+(z) - u_+(x) \le h(x, y).$$

But the functions u_- and u_+ being conjugate, we have $u_+(z) = u_-(z)$, since $z \in \mathcal{M}_0$. $\square$

Theorem 5.3.6. *For $x \in M$, we define the function $h^x : M \to \mathbb{R}$ (resp. $h_x : M \to \mathbb{R}$) by $h^x(y) = h(x, y)$ (resp. $h_x(y) = h(y, x)$). For each $x \in M$, the function $h^x : M \to \mathbb{R}$ (resp. $-h_x$) is in $\mathcal{S}_-$ (resp. $\mathcal{S}_+$). Moreover, its conjugate function $u_+^x \in \mathcal{S}_+$ (resp. $u_-^x \in \mathcal{S}_-$) vanishes at x.*

Proof. We first show that the function h^x is dominated by $L + c[0]$. If $\gamma : [0, t] \to M$ is a (continuous) piecewise C^1 curve, we have

$h_t(\gamma(0), \gamma(t)) \leq \int_0^t L(\gamma(s), \dot\gamma(s))\, ds$ and thus, by part (1) of the lemma giving the properties of h_t 5.3.2, we obtain

$$h_{t'+t}(x, \gamma(t)) \leq h_{t'}(x, \gamma(0)) + \int_0^t L(\gamma(s), \dot\gamma(s))\, ds,$$

which gives by adding $c[0](t + t')$ to the two members

$$h_{t'+t}(x, \gamma(t)) + c[0](t+t') \leq h_{t'}(x, \gamma(0)) + c[0]t' + \int_0^t L(\gamma(s), \dot\gamma(s))\, ds + c[0]t.$$

By taking the liminf for $t' \to +\infty$, we find

$$h(x, \gamma(t)) \leq h(x, \gamma(0)) + \int_0^t L(\gamma(s), \dot\gamma(s))\, ds + c[0]t,$$

which we can write as

$$h^x(\gamma(t)) - h^x(\gamma(0)) \leq \int_0^t L(\gamma(s), \dot\gamma(s))\, ds + c[0]t.$$

To finish showing that $h^x \in \mathcal{S}_-$, it is enough to show that for $y \in M$, we can find an extremal curve $\gamma_- :]-\infty, 0]$ such that $\gamma_-(0) = y$ and

$$\forall t \leq 0, h(x, y) \geq h(x, \gamma_-(t)) + \int_t^0 L(\gamma_-(s), \dot\gamma_-(s))\, ds - c[0]t.$$

We take a sequence of extremal curves $\gamma_n : [0, t_n] \to M$ connecting x to y, and such that

$$h(x, y) = \lim_{n\to\infty} \int_0^{t_n} L(\gamma_n(s), \dot\gamma_n(s))\, ds + c[0]t_n.$$

Since $t_n \to \infty$ and the γ_n are all minimizing extremal curves, by extracting a subsequence if necessary, we can suppose that the sequence of extremal curves $\gamma'_n : [-t_n, 0] \to M, t \mapsto \gamma_n(t_n + t)$ converges to an extremal curve $\gamma_- :]-\infty, 0] \to M$. We have $\gamma_-(0) = \lim_{n\to\infty} \gamma_n(t_n) = y$. Let us fix $t \in]-\infty, 0]$, for n big enough, we have $t_n + t \geq 0$ and we can write

$$\int_0^{t_n} L(\gamma_n(s), \dot\gamma_n(s))\, ds + c[0]t_n = \int_0^{t_n+t} L(\gamma_n(s), \dot\gamma_n(s))\, ds + c[0](t_n+t)$$
$$\int_0^{t_n+t} L(\gamma_n(s), \dot\gamma_n(s))\, ds + c[0](t_n+t) + \int_t^0 L(\gamma'_n(s), \dot\gamma'_n(s))\, ds - c[0]t. \tag{$*$}$$

By convergence of the γ'_n, we have

$$\int_t^0 L(\gamma'_n(s), \dot{\gamma}'_n(s))\, ds \to \int_t^0 L(\gamma_-(s), \dot{\gamma}_-(s))\, ds.$$

Since $\lim_n t_n + t = \infty$, and $\lim_{n\to\infty} \gamma_n(t_n + t) = \lim_{n\to\infty} \gamma'_n(t) = \gamma_-(t)$, by part (8) of the corollary giving the properties of h, we obtain $h(x, \gamma_-(t)) \leq \liminf_{n\to\infty} \int_0^{t_n+t} L(\gamma_n(s), \dot{\gamma}_n(s))\, ds + c[0](t_n + t)$. By taking the liminf in the equality $(*)$, we do indeed find

$$h(x, y) \geq h(x, \gamma_-(t)) + \int_t^0 L(\gamma_-(s), \dot{\gamma}_-(s))\, ds - c[0]t.$$

It remains to be seen that $u_+^x \in \mathcal{S}_+$, the conjugate function of h^x, vanishes at x. For that, we define

$$a(t) = -c[0]t + \sup_\gamma h(x, \gamma(t)) - \int_0^t L(\gamma(s), \dot{\gamma}(s))\, ds,$$

where $\gamma : [0, t] \to M$ varies among C^1 curves with $\gamma(0) = x$. This quantity $a(t)$is nothing but $T_t^+(h^x)(x) - c[0]t$, and thus $u_+^x(x) = \lim_{t\to\infty} a(t)$. For each $t > 0$, we can choose an extremal curve $\gamma_t : [0, t] \to M$, with $\gamma_t(0) = x$ and

$$a(t) = -c[0]t + h(x, \gamma_t(t)) - \int_0^t L(\gamma_t(s), \dot{\gamma}_t(s))\, ds.$$

We then choose a sequence $t_n \to +\infty$ such that $\gamma_{t_n}(t_n)$ converges to a point of M which we will call y. By continuity of h we have $h(x, y) = \lim_{n\to+\infty} h(x, \gamma_{t_n}(t_n))$. Moreover, by part (8) of the corollary giving the properties of h, we have

$$h(x, y) \leq \liminf_{n\to+\infty} \int_0^t L(\gamma_t(s), \dot{\gamma}_t(s))\, ds + c[0]t.$$

It follows that $u_+^x(x) = \lim a(t_n) \leq 0$. Since we already showed the inequality $h(x, y) \geq u_-(y) - u_+(x)$, for any pair of conjugate functions $u_- \in \mathcal{S}_-, u_+ \in \mathcal{S}_+$, we have $h(x, x) \geq h^x(x) - u_+^x(x)$. However $h(x, x) = h^x(x)$, which gives $u_+^x(x) \geq 0$. □

Corollary 5.3.7. *For each $x, y \in M$, we have the equality*

$$h(x, y) = \sup_{(u_-, u_+)} u_-(y) - u_+(x),$$

the supremum being taken on pairs (u_-, u_+) of conjugate functions $u_- \in \mathcal{S}_-, u_+ \in \mathcal{S}_+$.

We can also give the following characterization for the Aubry set $\mathcal{A}_0$.

Proposition 5.3.8. *If $x \in M$, the following conditions are equivalent*

(1) *$x \in \mathcal{A}_0$;*

(2) *the Peierls barrier $h(x, x)$ vanishes;*

(3) *there exists a sequence $\gamma_n : [0, t_n] \to M$ of (continuous) piecewise C^1 curves such that*

–for each n, we have $\gamma_n(0) = \gamma_n(t_n) = x$;

–the sequence t_n tends to $+\infty$, when $n \to \infty$;

–for $n \to \infty$, we have $\int_0^{t_n} L(\gamma_n(s), \dot{\gamma}_n(s))\, ds + c[0]t_n \to 0$;

(4) *there exists a sequence $\gamma_n : [0, t_n] \to M$ minimizing extremal curves such that*

–for each n, we have $\gamma_n(0) = \gamma_n(t_n) = x$;

–the sequence t_n tends to $+\infty$, when $n \to \infty$;

–for $n \to \infty$, we have $\int_0^{t_n} L(\gamma_n(s), \dot{\gamma}_n(s))\, ds + c[0]t_n \to 0$.

Proof. Equivalence of conditions (1) and (2) results from the previous corollary. Equivalence of (2), (3) and (4) results from the definition of h. □

5.4 Chain Transitivity

Proposition 5.4.1. *Let (u_-, u_+) be a given pair of conjugate functions. If $t_0 > 0$ given, then for each $\epsilon > 0$, there exists $\delta > 0$ such that if $\gamma : [0, t] \to M$ is an extremal curve, with $t \geq t_0, u_-(\gamma(0)) \leq u_+(\gamma(0)) + \delta$ and $\int_0^t L(\gamma(s), \dot{\gamma}(s))\, ds + c[0]t \leq u_+(\gamma(t)) - u_+(\gamma(0)) + \delta$, then for each $s \in [0, t]$, we can find a point in $\tilde{\mathcal{I}}_{(u_-, u_+)}$ at distance at most ϵ from $(\gamma(s), \dot{\gamma}(s))$.*

Proof. Since $u_+ \prec L + c[0]$, for $0 \leq a \leq b \leq t$, we have

$$u_+(\gamma(a)) - u_+(\gamma(0)) \leq \int_0^a L(\gamma(s), \dot{\gamma}(s))\, ds + c[0]a,$$

$$u_+(\gamma(t)) - u_+(\gamma(b)) \leq \int_b^t L(\gamma(s), \dot{\gamma}(s))\, ds + c[0](t - b).$$

Subtracting the last two inequalities from the inequality $\int_0^t L(\gamma(s),\dot\gamma(s))\,ds + c[0]t \le u_+(\gamma(t)) - u_+(\gamma(0)) + \delta$, we find

$$\int_a^b L(\gamma(s),\dot\gamma(s))\,ds + c[0](b-a) \le u_+(\gamma(b)) - u_+(\gamma(a)) + \delta. \quad (*)$$

Moreover, since $u_- \prec L + c[0]$, we have

$$u_-(\gamma(a)) - u_-(\gamma(0)) \le \int_0^a L(\gamma(s),\dot\gamma(s))\,ds + c[0]a.$$

Since, by the inequality $(*)$, this last quantity is not larger than $u_+(\gamma(a)) - u_+(\gamma(0)) + \delta$, we obtain $u_-(\gamma(a)) - u_-(\gamma(0)) \le u_+(\gamma(a)) - u_+(\gamma(0)) + \delta$. The condition $u_-(\gamma(0)) \le u_+(\gamma(0)) + \delta$ gives then $u_-(\gamma(a)) \le u_+(\gamma(a)) + 2\delta$, for each $a \in [0,t]$. We conclude that it is enough to show the lemma with $t = t_0$, taking δ smaller if necessary. Let us argue by contradiction. We suppose that there exists a sequence of extremal curves $\gamma_n : [0,t_0] \to M$ and a sequence δ_n, such that the following conditions are satisfied

(1) $\delta_n \to 0$;
(2) $u_-(\gamma_n(0)) \le u_+(\gamma_n(0)) + \delta_n$;
(3) $\int_0^{t_0} L(\gamma_n(s),\dot\gamma_n(s))\,ds + c[0]t_0 \le u_+(\gamma_n(t_0)) - u_+(\gamma_n(0)) + \delta_n$;
(4) there exists $s_n \in [0,t_0]$ such that the distance from $(\gamma_n(s_n),\dot\gamma_n(s_n))$ with $\tilde{\mathcal{I}}_{(u_-,u_+)}$ is bigger than ϵ.

By conditions (1) and (3) above, there exists a constant $C < +\infty$ such that $\int_0^{t_0} L(\gamma_n(s),\dot\gamma_n(s))\,ds \le C$, for each $n \ge 0$. It follows that there exists $s'_n \in [0,t_0]$ such that $L(\gamma_n(s'_n),\dot\gamma_n(s'_n)) \le C/t_0$. Therefore the $(\gamma_n(s'_n),\dot\gamma_n(s'_n))$ are all in the compact subset $K = \{(x,v) \mid L(x,v) \le C/t_0\} \subset TM$. Since the γ_n are extremal curves, we have $(\gamma_n(t),\dot\gamma_n(t)) = \phi_{(t-s'_n)}(\gamma_n(s'_n),\dot\gamma_n(s'_n))$, and thus the point $(\gamma_n(t),\dot\gamma_n(t))$ is in the compact subset $\bigcup_{s\in[0,t_0]}\phi_s(K)$, for each $n \ge 0$ and each $t \in [0,t_0]$.

Extracting a subsequence if necessary, we can thus suppose that the sequence of extremal curves γ_n converges in the C^1 topology to the extremal curve $\gamma_\infty : [0,t_0] \to M$. As we saw above, we have $u_-(\gamma_n(t)) \le u_+(\gamma_n(t)) + 2\delta_n$, for each $t \in [0,t_0]$. Going to the limit and taking s_∞ a value of adherence of the sequence s_n, we thus obtain

(1) for each $t \in [0,t_0]$, we have $u_-(\gamma_\infty(t)) \le u_+(\gamma_\infty(t))$;
(2) $\int_0^{t_0} L(\gamma_\infty(s),\dot\gamma_\infty(s))\,ds + c[0]t_0 \le u_+(\gamma_\infty(t_0)) - u_+(\gamma_\infty(0))$;

(3) there exists a number $s_\infty \in [0,t_0]$ such that the distance from the point $(\gamma_\infty(s_\infty),\dot\gamma_\infty(s_\infty))$ to the set $\tilde{\mathcal{I}}_{(u_-,u_+)}$ is at least ϵ.

However $u_- \geq u_+$ thus $u_-(\gamma_\infty(t)) = u_+(\gamma_\infty(t))$, for each $t \in [0,t_0]$, which gives $\gamma_\infty(t) \in \mathcal{I}_{(u_-,u_+)}$. In the same way, the fact that $u_+ \prec L+c[0]$, forces the equality in condition (2) above. This gives $\int_0^{t_0} L(\gamma_\infty(s),\dot\gamma_\infty(s))\,ds + c[0]t_0 = u_+(\gamma_\infty(t_0)) - u_+(\gamma_\infty(0))$. In particular, the derivative of u_+ at $\gamma_\infty(s)$, for $s \in]0,t_0[$ is the Legendre transform of $(\gamma_\infty(s),\dot\gamma_\infty(s))$. It follows that $(\gamma_\infty(s_\infty),\dot\gamma_\infty(s_\infty)) \in \tilde{\mathcal{I}}_{(u_-,u_+)}$, which contradicts condition (3) above. $\square$

Corollary 5.4.2. *Let $u_- \in \mathcal{S}_-$ and $u_+ \in \mathcal{S}_+$ be a pair of conjugate functions. If $x,y \in M$ is such that $h(x,y) = u_+(y) - u_-(x)$, then $x,y \in \mathcal{I}_{(u_-,u_+)}$. Moreover, if (x,v_x) and (y,v_y) are the points of $\tilde{\mathcal{I}}_{(u_-,u_+)}$ above x and y, then for each $\epsilon > 0$, we can find a sequence of points $(x_i,v_i) \in \tilde{\mathcal{I}}_{(u_-,u_+)}, i = 0,1,\ldots,k$, with $k \geq 1$, such that $(x_0,v_0) = (x,v_x), (x_k,v_k) = (y,v_y)$, that there exists $t \in [1,2]$ with the distance from $\phi_t(x_{k-1},v_{k-1})$ to $(x_k,v_k) = (y,v_y)$ is less than ϵ and that, for $i = 0,\ldots,k-2$, the distance in TM from $\phi_1(x_i,v_i)$ to (x_{i+1},v_{i+1}) is also less than ϵ.*

Proof. We know that $h(x,y) \geq u_-(y) - u_+(x)$. Since $u_- \geq u_+$, we see that $u_-(x) = u_+(x)$ and $u_-(y) = u_+(y)$. By the properties of h, there exist a sequence of extremals $\gamma_n : [0,t_n] \to M$, with $t_n \to \infty$, such that $\gamma_n(0) = x, \gamma_n(t_n) = y$ and $\int_0^{t_n} L(\gamma_n(s),\dot\gamma_n(s))\,ds + c[0]t_0 \leq u_+(\gamma_n(t_n)) - u_+(\gamma_n(0)) + \delta_n$ with $\delta_n \to 0$.

Let us fix $\epsilon > 0$. The compactness of $\tilde{\mathcal{I}}_{(u_-,u_+)}$ gives the existence of ϵ' such that if $(a,v) \in \tilde{\mathcal{I}}_{(u_-,u_+)}$ and $(b,w) \in TM$ are at distance less than ϵ', then, for each $s \in [0,2]$, the distance from $\phi_s(a,v)$ with $\phi_s(b,w)$ is smaller than $\epsilon/2$. By the previous proposition, for n large enough, $(\gamma_n(s),\dot\gamma_n(s)), s \in [0,t_n]$ is at distance $\leq \min(\epsilon',\epsilon/2)$ from a point of $\tilde{\mathcal{I}}_{(u_-,u_+)}$. Let us fix such an integer n with $t_n \geq 1$ and call k the greatest integer $\leq t$. We thus have $k \geq 1$ and $t = t_n - (k-1) \in [1,2[$. For $i = 1,2,\ldots,k-2$, let us choose $(x_i,v_i) \in \tilde{\mathcal{I}}_{(u_-,u_+)}$ at distance $\leq \min(\epsilon',\epsilon/2)$ from $(\gamma_n(i),\dot\gamma_n(i))$ and let us set $(x_0,v_0) = (x,v_x), (x_k,v_k) = (y,v_y)$. For $i = 0,\ldots,k-2$, by the choice of ϵ', the distance between $\phi_1(x_i,v_i)$ and $\phi_1(\gamma_n(i),\dot\gamma_n(i)) = (\gamma_n(i+1),\dot\gamma_n(i+1))$ is $\leq \epsilon/2$, consequently the distance between $\phi_t(x_i,v_i)$ and (x_{i+1},v_{i+1}) is less than ϵ. In the same way, as $t \in [1,2[$ the distance between

$\phi_t(x_{k-1}, v_{k-1})$ and $\phi_t(\gamma_n(k-1), \dot\gamma_n(k-1)) = (\gamma_n(k), \dot\gamma_n(k))) = (y, v_y)$ is less than $\epsilon/2$. □

The following theorem is due to Mañé, see [Mn97, Theorem V, page 144]:

Theorem 5.4.3 (Mañé)**.** *The Mañé set $\tilde{\mathcal{N}}_0$ is chain transitive for the flow ϕ_t. In particular, it is connected.*

Proof. We recall that $\tilde{\mathcal{N}}_0 = \bigcup \tilde{\mathcal{I}}_{(u_-,u_+)}$, where the union is taken on all pairs of conjugate functions. Let us notice that $\tilde{\mathcal{M}}_0 \subset \tilde{\mathcal{I}}_{(u_-,u_+)}$, by definition of conjugate functions. If $x, y \in \mathcal{M}_0$, then by the compactness of the set of differences $u_- - u_+$ and the characterization of h, there exists a pair (u_-, u_+) of conjugate functions such that $h(x, y) = u_-(y) - u_+(x)$, since $x, y \in \mathcal{M}_0$, we have also $h(x, y) = u_+(y) - u_-(x)$. By the corollary above, we see that the points $(x, v_x) \in \tilde{\mathcal{M}}_0$ and $(y, v_y) \in \tilde{\mathcal{M}}_0$ above x and y can, for each $\epsilon > 0$, be connected by an ϵ-chain of points, for ϕ_t, in $\tilde{\mathcal{I}}_{(u_-,u_+)}$. It follows that the set $\tilde{\mathcal{M}}_0$ is contained in a single chain recurrent component of the union $\tilde{\mathcal{N}}_0 = \bigcup \tilde{\mathcal{I}}_{(u_-,u_+)}$. To finish showing that $\tilde{\mathcal{N}}_0$ is chain transitive, it is enough to notice that if $(x, v) \in \mathcal{I}_{(u_-,u_+)}$, then the α-limit and ω-limit sets, for ϕ_t, of (x, v) both contain points of $\tilde{\mathcal{M}}_0$. □

Chapter 6

A Closer Look at the Lax-Oleinik semi-group

6.1 Semi-convex Functions

6.1.1 The Case of Open subsets of $\mathbb{R}^n$

Proposition 6.1.1. *Let U be an open convex subset of $\mathbb{R}^n$, and let $u : U \to \mathbb{R}$ be a function. The following conditions are equivalent:*

(i) *There exists a C^2 function $\varphi : U \to \mathbb{R}$ with bounded second derivative and such that $u + \varphi$ is convex.*

(ii) *There exists a C^∞ function $\varphi : U \to \mathbb{R}$ with bounded second derivative and such that $u + \varphi$ is convex.*

(iii) *There exists a finite constant K and for each $x \in U$ there exists a linear form $\theta_x : \mathbb{R}^n \to \mathbb{R}$ such that*

$$\forall y \in U, u(y) - u(x) \geq \theta_x(y - x) - K\|y - x\|^2.$$

Proof. Obviously (ii) implies (i). To prove that (i) implies (iii), we denote by $2K$ an upper bound on U of the norm of the second derivative of φ, using Taylor's formula, we see that

$$\forall x, y \in U, \varphi(y) - \varphi(x) \leq d_x\varphi(y - x) + K\|y - x\|^2.$$

We now use theorem 1.2.9 to obtain a supporting linear form θ_1 at x for the convex function $u + \varphi$. This gives

$$u(y) - u(x) \geq \theta_1(y - x) - \varphi(y) + \varphi(x).$$

Combining the two inequalities, we get

$$u(y) - u(x) \geq \theta_1(y - x) - d_x\varphi(y - x) - K\|y - x\|^2,$$

but the map $v \mapsto \theta_1(v) - d_x\varphi(v)$ is linear.

It remains to prove that (iii) implies (ii). We consider $\varphi(y) = K\|y\|^2$, where K is the constant given by (iii). A simple computation gives

$$\varphi(y) - \varphi(x) = K\|y - x\|^2 + 2K\langle y - x, x\rangle.$$

Adding the inequality given by (iii) and the equality above, we obtain

$$(u + \varphi)(y) - (u + \varphi)(y) \geq \theta_x(y - x) + 2K\langle y - x, x\rangle.$$

This shows that $u+\varphi$ admits the linear map $v \mapsto \theta_x(v)+2K\langle v, x\rangle$ as a supporting linear form at x. It follows from proposition 1.2.8 that $u + \varphi$ is convex. □

Definition 6.1.2 (Semi-convex)**.** A function $u : U \to \mathbb{R}$, defined on the open convex subset U of $\mathbb{R}^n$, is said to be semi-convex if it satisfies one of (and hence all) the three equivalent conditions of proposition 6.1.1.

We will say that u is K-semi-convex if it satisfies condition (iii) of 6.1.1 with K as a constant.

A function u is said to be semi-concave (resp. K-semi-concave) is $-u$ semi-convex (resp. K-semi-convex).

We will say that a function $u : V \to \mathbb{R}$, defined on an open subset of $\mathbb{R}^n$ is locally semi-convex (resp. semi-concave) if for each point $x \in V$ there exists an open convex neighborhood U_x of x in V such that the restriction $u|U_x$ is semi-convex (resp. semi-concave).

Here are some properties of locally semi-convex or semi-concave functions:

Proposition 6.1.3. (1) *A locally semi-convex (resp. semi-concave) function is locally Lipschitz.*

(2) *If u is locally semi-convex (resp. semi-concave) then u is differentiable almost everywhere.*

(3) *If $u : V \to \mathbb{R}$ is locally semi-convex (resp. semi-concave) and $f : W \to V$ is a C^2 map then $u \circ f : W \to \mathbb{R}$ is also is locally semi-convex (resp. semi-concave).*

Proof. It suffices to prove these properties for a semi-convex function $u : U \to \mathbb{R}$ defined on the open convex subset U of $\mathbb{R}^n$. But u is the sum of a C^2 function and a convex function. A C^2 function is obviously locally Lipschitz and differentiable everywhere. Moreover, a convex function on an open subset of $\mathbb{R}^n$ is locally Lipschitz by corollary 1.1.9 and differentiable almost everywhere by corollary 1.1.11 (or by Rademacher's theorem 1.1.10). This proves (1) and (2).

It suffices to prove (3) for $u : V \to \mathbb{R}$ convex. We fix some $z \in W$, and we pick $r > 0$ such that the closed ball $\bar{B}(f(z), 2r) \subset V$. We first show that there is a constant K_1 such that each supporting linear form p of u at $y \in \bar{B}(f(z), r)$ satisfies $\|p\| \le K_1$. In fact, for v such that $\|v\| \le r$, the point $y + v$ is in V since it belongs to $\bar{B}(f(z), 2r)$, hence we can write

$$u(y+v) - u(y) \ge p(v).$$

We set $M = \max\{|u(x)| \mid x \in \bar{B}(f(z), 2r)$. The constant M is finite because $\bar{B}(f(z), 2r)$ is compact and u is continuous by (1). from the inequality above we obtain

$$\forall v \in \bar{B}(0, r), p(v) \le 2M.$$

It is not difficult to conclude that $\|p\| \le K_1$. We now pick an open convex set O which contains x, such that its closure $\bar{O}$ is compact and contained in W and $f(O) \subset B(f(z), r)$. Since f is C^2 and O is convex with compact closure, by Taylor's formula, we can find a constant K_2 such that

$$\forall z_1, z_2 \in O, |f(z_2) - f(z_1) - Df(z_1)[z_2 - z_1]| \le K_2 \|z_2 - z_1\|^2. \quad (*)$$

If z_1, z_2 are both in O, then $f(z_2), f(z_1)$ are both in $B(f(z), r)$. If we call p_1 a supporting linear form of u at $f(z_1)$, we have

$$\|p_2\| \le K_1 \text{ and } u \circ f(z_2) - u \circ f(z_1) \ge p_1(f(z_2) - f(z_1)).$$

Combining with $(*)$, we obtain

$$\forall z_1, z_2 \in O, u \circ f(z_2) - u \circ f(z_1) \ge p_1 \circ Df(z_1)[z_2 - z_1] - K_1 K_2 \|z_2 - z_1\|^2.$$

Thus $u \circ f$ is semi-convex on O. $\square$

Part (3) of the last proposition 6.1.3 implies that the notion of locally semi-convex (or semi-concave) is well-defined on a differentiable manifold (of class at least C^2).

Definition 6.1.4 (Semi-convex)**.** A function $u : M \to \mathbb{R}$ defined on the C^2 differentiable manifold M is locally semi-convex (resp. semi-concave), if for each $x \in M$ there is a C^2 coordinate chart $\varphi : U \to \mathbb{R}^n$, with $x \in U$, such that $u \circ \varphi^{-1} : \varphi(U) \to \mathbb{R}$ is semi-convex (resp. semi-concave).

In that case for each C^2 coordinate chart $\theta : V \to \mathbb{R}^n$, the map $u \circ \theta^{-1} : \theta(U) \to \mathbb{R}$ is semi-convex (resp. semi-concave).

Theorem 6.1.5. *A function $u : M \to \mathbb{R}$, defined on the C^2 differentiable manifold M, is both locally semi-convex and locally semi-concave if and only if it is $C^{1,1}$.*

Proof. Suppose that u is $C^{1,1}$. Since the result is by nature local, we can suppose that M is in fact the open ball subset $\mathring{B}(0,r)$ of $\mathbb{R}^n$ and that the derivative $du : \mathring{B}(0,r) \to \mathbb{R}^{n*}, x \mapsto d_x u$ is Lipschitz with Lipschitz constant $\leq K$. If $x, y \in \mathring{B}(0,r)$, we can write

$$u(y) - u(x) = \int_0^1 d_{ty+(1-t)x} u(y-x)\, dt.$$

Moreover, we have

$$\|d_{ty+(1-t)x} u - d_x u\| \leq Kt\|y-x\|.$$

Combining these two inequalities, we obtain

$$\begin{aligned}
|u(y) - u(x) - d_x u(y-x)| &= \Big|\int_0^1 d_{ty+(1-t)x} u(y-x) - d_x u(y-x)\, dt\Big| \\
&\leq \int_0^1 \|d_{ty+(1-t)x} u - d_x u\| \|y-x\|\, dt \\
&\leq \int_0^1 K\|ty + (1-t)x - x\| \|y-x\|\, dt \\
&= \int_0^1 Kt\|y-x\|^2\, dt \\
&= \frac{K}{2}\|y-x\|^2.
\end{aligned}$$

This implies that u is both semi-convex and semi-concave on the convex set $\mathring{B}(0,r)$.

Suppose now that u is both locally semi-convex and locally semi-concave. Again due to the local nature of the result, we can assume $M = \mathring{B}(0,r)$ and that u is both K_1-semi-convex and K_2-semi-concave. Given $x \in \mathring{B}(0,r)$, we can find $\theta_x^1, \theta_x^2 \in \mathbb{R}^{n*}$ such that

$$\forall y \in \mathring{B}(0,r), u(y) - u(x) \geq \theta_x^1(y-x) - K_1\|y-x\|^2,$$

and

$$\forall y \in \mathring{B}(0,r), u(y) - u(x) \leq \theta_x^2(y-x) + K_2\|y-x\|^2.$$

For a fixed $v \in \mathbb{R}^n$, and for all ϵ small enough, the point $x + \epsilon v \in \mathring{B}(0,r)$, thus combining the two inequalities above we obtain

$$\theta_x^1(\epsilon v) - K_1\|\epsilon v\|^2 \leq \theta_x^2(\epsilon v) + K_2\|\epsilon v\|^2,$$

for all ϵ small enough. Dividing by ϵ and letting ϵ go to 0, we obtain

$$\forall v \in \mathbb{R}^n, \theta_x^1(v) \leq \theta_x^2(v).$$

Changing v into $-v$, and using the linearity of $\theta_x^1, \theta_x^2 \in \mathbb{R}^{n*}$, we see that $\theta_x^1 = \theta_x^2$. Thus we have

$$\forall y \in \mathring{B}(0,r), |u(y) - u(x) - \theta_x^1(y-x)| \leq \max(K_1, K_2)\|y-x\|^2.$$

It follows from 4.11.3 that u is C^1 on $\mathring{B}(0,r)$ with a Lipschitz derivative. □

6.2 The Lax-Oleinik Semi-group and Semi-convex Functions

In this section we suppose that the compact manifold M is endowed with a C^2 Lagrangian $L : TM \to \mathbb{R}$ which is superlinear and C^2 strictly convex in the fibers of the tangent bundle $\pi : TM \to M$. We can then define the two Lax-Oleinik semi-groups $T_t^-, T_t^+ : \mathcal{C}^0(M,\mathbb{R}) \to \mathcal{C}^0(M,\mathbb{R})$. An immediate consequence of proposition 4.11.1 is the following proposition:

Proposition 6.2.1. *For each $u \in \mathcal{C}^0(M,\mathbb{R})$ and each $t > 0$, the function $T_t^-(u)$ (resp. $T_t^+(u)$) is locally semi-concave (resp. semi-convex)*

Theorem 6.2.2. *Suppose that $T_t^-(u)$ (resp. $T_t^+(u)$) is C^1, where $u \in \mathcal{C}^0(M,\mathbb{R})$ and $t > 0$, then we have $u = T_t^+T_t^-(u)$ (resp. $u = T_t^-T_t^+(u)$.*

In particular, the function u must be locally semi-convex. Moreover, for each $t' \in]0,t[$ the function $T_{t'}^-(u)$ (resp. $T_t^+(u)$) is $\mathrm{C}^{1,1}$.

Before embarking in the proof of this theorem, we will need the following lemma, whose proof results easily from the definitions of T_t^- and T_t^+:

Lemma 6.2.3. *If $u, v \in \mathcal{C}^0(M,\mathbb{R})$ and $t \geq 0$, the following three conditions are equivalent:*

(1) *$v \leq T_t^- u$;*
(2) *$T_t^+ v \leq u$;*
(3) *for each C^1 curve $\gamma : [0,t] \to M$, we have*

$$v(\gamma(t)) - u(\gamma(0)) \leq \mathbb{L}(\gamma) = \int_0^t L(\gamma(s), \dot\gamma(s))\, ds.$$

If anyone of these conditions is satisfied and $\gamma : [0,t] \to M$ is a C^1 curve with

$$v(\gamma(t)) - u(\gamma(0)) = \mathbb{L}(\gamma) = \int_0^t L(\gamma(s), \dot\gamma(s))\, ds,$$

then $v(\gamma(t)) = T_t^- u(\gamma(t))$ and $u(\gamma(0)) = T_t^+ v(\gamma(0))$.

Proof of theorem 6.2.2. We set $v = T_t^- u$. For each $x \in M$, we can find a minimizing C^2 extremal curve $\gamma_x : [0,t] \to M$ with $\gamma_x(t) = x$ and

$$v(\gamma_x(t)) = \int_0^t L(\gamma(s), \dot\gamma(s))\, ds + u(\gamma_x(0)).$$

From lemma 6.2.3, we have $u(\gamma_x(0)) = T_t^+ v(\gamma(0))$. To show that $u = T_t^+ v$, it then remains to see that the set $\{\gamma_x(0) \mid x \in M\}$ is the whole of M. Since v is C^1, it follows from proposition 4.11.1 that $d_x v = \partial L/\partial v(x, \dot\gamma_x(t))$ or $\dot\gamma_x(t) = \operatorname{grad}_L v(x)$. Since

γ_x is an extremal we conclude that $(\gamma_x(s), \dot\gamma_x(s)) = \phi_{s-t}(x, \dot\gamma_x(t))$, where ϕ_t is the Euler-Lagrange flow of L. In particular, the set $\{\gamma_x(0) \mid x \in M\}$ is nothing but the image of the map $f : M \to M, x \mapsto \pi\phi_{-t}(x, \operatorname{grad}_L v(x))$. Since $v = T_t^- u$ is C^1, the map f is continuous and homotopic to the identity, a homotopy being given by $(x, s) \mapsto \pi\phi_{-s}(x, \operatorname{grad}_L v(x))$, with $s \in [0, t]$. Since M is a compact manifold without boundary, it follows from degree theory $\mod 2$ that f is surjective. This finishes the proof of $u = T_t^+ T_t^-(u)$.

Using $t > 0$, we obtain the local semi-convexity of u from proposition 6.2.1.

If $t' \in]0, t[$, we also have $v = T_t^- u = T_{t-t'}^- T_{t'}^- u$. Applying the first part with $T_{t'}^- u$ instead of u and $t - t' > 0$ instead of t, we see that $T_{t'}^- u$ is locally semi-convex. It is also locally semi-concave by 6.2.1, since $t' > 0$. It follows from theorem 6.1.5 that $T_{t'}^- u$ is $\mathrm{C}^{1,1}$. □

6.3 Convergence of the Lax-Oleinik Semi-group

Up to now in this book, all the statements given above do hold for periodic time-dependent Lagrangians which satisfy the hypothesis imposed by Mather in [Mat91, Pages 170–172]. The results in this section depend heavily on the invariance of the energy by the Euler-Lagrange flow. Some of these results do not hold for the time-dependent case, see [FM00].

The main goal of this section is to prove the following theorem:

Theorem 6.3.1 (Convergence of the Lax-Oleinik Semi-group). *Let $L : TM \to \mathbb{R}$ be a C^2 Lagrangian, defined on the the compact manifold M, which is superlinear and C^2 strictly convex in the fibers of the tangent bundle $\pi : TM \to M$. If $T_t^-, T_t^+ : \mathcal{C}^0(M, \mathbb{R}) \to \mathcal{C}^0(M, \mathbb{R})$ are the two Lax-Oleinik semi-groups associated with L, then for each $u \in \mathcal{C}^0(M, \mathbb{R})$, the limits, for $t \to +\infty$, of $T_t^- u + c[0]t$ and $T_t^+ u - c[0]t$ exist. The limit of $T_t^- u + c[0]t$ is in $\mathcal{S}_-$, and the limit of $T_t^+ u - c[0]t$ is in $\mathcal{S}_+$.*

Particular cases of the above theorem are due to Namah and Roquejoffre, see [NR97b, NR97a, NR99] and [Roq98a]. The first proof of the general case was given in [Fat98b]. There now exists other proofs of the general case due to Barles-Souganidis and to Roquejoffre, see [BS00] and [Roq98b].

The main ingredient in the proof is the following lemma:

Lemma 6.3.2. *Under the hypothesis of theorem 6.3.1 above, for each $\epsilon > 0$, there exists a $t(\epsilon) > 0$ such that for each $u \in \mathcal{C}^0(M, \mathbb{R})$ and each $t \geq t(\epsilon)$, if $T_t^- u$ has a derivative at $x \in M$, then $c[0] - \epsilon \leq H(x, d_x T_t^- u) \leq c[0] + \epsilon$. Consequently $\lim_{t\to\infty} \mathbb{H}_M(T_t^- u) \to c[0]$, for each $u \in \mathcal{C}^0(M, \mathbb{R})$.*

Proof. By Carneiro's theorem 4.12.5, the set

$$W_\epsilon = \{(x, v) \mid c[0] - \epsilon \leq H \circ \tilde{\mathcal{L}}(x, v) \leq c[0] + \epsilon\}$$

is a neighborhood of the Mather set $\tilde{\mathcal{M}}_0$. We can now apply lemma 5.3.4 with W_ϵ as neighborhood of $\tilde{\mathcal{M}}_0$, to find $t(\epsilon) > 0$ such that for minimizing extremal curve $\gamma : [0, t] \to M$ with $t \geq t(\epsilon)$, there exists a $t' \in [0, t]$ with $(\gamma(t'), \dot\gamma(t')) \in W_\epsilon$, which means that $H \circ \tilde{\mathcal{L}}(\gamma(t'), \dot\gamma(t'))$ is in $[c[0] - \epsilon, c[0] + \epsilon]$. This implies that for such a minimizing curve, we have

$$\forall s \in [0, t], c[0] - \epsilon \leq H \circ \tilde{\mathcal{L}}(\gamma(s), \dot\gamma(s)) \leq c[0] + \epsilon.$$

In fact, since γ is a minimizing curve, its speed curve $s \mapsto (\gamma(s), \dot\gamma(s))$ is a piece of an orbit of the Euler-Lagrange flow. Since the energy $H \circ \tilde{\mathcal{L}}$ is invariant by the Euler-Lagrange flow, it follows that $H \circ \tilde{\mathcal{L}}(\gamma(s), \dot\gamma(s))$ does not depend on $s \in [0, t]$, but for $s = t'$, we know that this quantity is in $[c[0] - \epsilon, c[0] + \epsilon]$.

If we suppose that $T_t^- u$ is differentiable at $x \in M$, and we pick a curve $\gamma : [0, t] \to M$, with $\gamma(t) = x$, such that

$$T_t^- u(x) = u(\gamma(0)) + \int_0^t L(\gamma(s), \dot\gamma(s))\, ds,$$

then we know that γ is minimizing, and from proposition 4.11.1 we also know that $d_x T_t^- u = \mathcal{L}(x, \dot\gamma(t))$. From what we saw above, we indeed conclude that when $t \geq t(\epsilon)$ we must have $H(x, d_x T_t^- u) \in [c[0] - \epsilon, c[0] + \epsilon]$. It follows that the Hamiltonian constant $\mathbb{H}_M(T_t^- u)$ does converge to $c[0]$ when $t \to +\infty$. $\square$

Proof of Theorem 6.3.1. It is convenient to introduce $\tilde{T}_t^- : \mathcal{C}^0(M,\mathbb{R}) \to \mathcal{C}^0(M,\mathbb{R})$ defined by $\tilde{T}_t^- u = T_t^- u + c[0]t$. It is clear that $\tilde{T}_t^-$ is itself a semi-group of non-expansive maps whose fixed points are precisely the weak KAM solutions in $\mathcal{S}_-$. We fix some $u_-^0 \in \mathcal{S}_-$. If $u \in \mathcal{C}^0(M,\mathbb{R})$, since $\tilde{T}_t^-$ is non-expansive, and $u_-^0 \in \mathcal{S}_-$ is a fixed point of $\tilde{T}_t^-$, we obtain

$$\|\tilde{T}_t u - u_-^0\|_0 = \|\tilde{T}_t u - \tilde{T}_t u_-^0\|_0 \leq \|u - u_-^0\|.$$

It follows that

$$\|\tilde{T}_t u\| \leq \|u - u_-^0\| + \|u_-^0\|.$$

The family of functions $\tilde{T}_t u = T_t^- u + ct$, with $t \geq 1$, is equi-Lipschitzian by lemma **??**, hence there exists a sequence $t_n \nearrow +\infty$ such that $\tilde{T}_{t_n} u \to u_\infty$ uniformly. Using the Lemma above 6.3.2 and Theorem 4.2.5, we see that $u_\infty \prec L + c[0]$, and hence $u_\infty \leq \tilde{T}_t u_\infty$, for each $t \geq 0$. Since $\tilde{T}_t$ is order preserving, we conclude that

$$\forall t' \geq t \geq 0, u_\infty \leq \tilde{T}_t u_\infty \leq \tilde{T}_{t'} u_\infty.$$

To show that u_∞ is a fixed point for $\tilde{T}_t$, it then remains to find a sequence $s_n \nearrow +\infty$ such that $\tilde{T}_{s_n} u_\infty \to u_\infty$.

Extracting if necessary, we can assume that $t_{n+1} - t_n \nearrow +\infty$. We will show that the choice $s_n = t_{n+1} - t_n$ does work. We have $\tilde{T}_{s_n} \circ \tilde{T}_{t_n} u = \tilde{T}_{t_{n+1}} u$ therefore

$$\begin{aligned}\|\tilde{T}_{s_n} u_\infty - u_\infty\|_0 &\leq \|\tilde{T}_{s_n} u_\infty - \tilde{T}_{s_n} \circ \tilde{T}_{t_n} u\|_0 + \|\tilde{T}_{t_{n+1}} u - u_\infty\|_0 \\ &\leq \|u_\infty - \tilde{T}_{t_n} u\|_0 + \|\tilde{T} u_{t_{n+1}} - u_\infty\|_0,\end{aligned}$$

where we used the inequality $\|\tilde{T}_{s_n} u_\infty - \tilde{T}_{s_n} \circ \tilde{T}_{t_n} u\|_0 \leq \|u_\infty - \tilde{T}_{t_n} u\|_0$, which is valid since the maps $\tilde{T}_t$ are non-expansive. Since $\|u_\infty - \tilde{T}_{t_n} u\|_0 \to 0$, this indeed shows that $\tilde{T}_t u_\infty = u_\infty$.

We still have to see that $\tilde{T}_t u \to u_\infty$, when $t \to +\infty$. If $t \geq t_n$, we can write

$$\|\tilde{T}_t u - u_\infty\|_0 = \|\tilde{T}_{t-t_n} \circ \tilde{T}_{t_n} u - \tilde{T}_{t-t_n} u_\infty\|_0 \leq \|\tilde{T}_{t_n} u - u_\infty\|_0.$$

This finishes the proof, since $\tilde{T}_{t_n} u \to u_\infty$. $\square$

A corollary is that the lim inf in the definition of the Peierls barrier is indeed a limit.

Corollary 6.3.3. *For each $x, y \in M$, we have $h(x,y) = \lim_{t\to+\infty} h_t(x,y) + c[0]t$. Moreover, the convergence is uniform on $M \times M$.*

Proof. If we fix y in M, we can define continuous functions $h_t^y M \to \mathbb{R}$ by $h_t^y(x) = h_t(x,y)$. It is not difficult to check that $T_{t'}^- h_t^y = h_{t'+t}^y$. It follows from theorem 6.3.1 that the limit of $h_{t+1}(x,y) + c[0]t = T_t^- h_1^y(x) + c[0]t$ exists, it must of course coincide with $\liminf_{t\to+\infty} h_{t+1}(x,y) + c[0]t$. By the definition of the Peierls barrier, this last quantity is $h(x,y) - c[0]$.

The fact that the limit is uniform on M follows from the fact that the $h_t, t \geq 1$ are equi-Lipschitzian by part (6) of lemma 5.3.2. □

6.4 Invariant Lagrangian Graphs

Theorem 6.4.1. *Suppose that $N \subset T^*M$ is a compact Lagrangian submanifold of T^*M which is everywhere transverse to the fibers of the canonical projection $\pi^* : T^*M \to M$. If the image $\phi_t^*(N)$ is still transversal to the fibers of the canonical projection $\pi^* : T^*M \to M$, then the same is true for $\phi_s^*(N)$, for any $s \in [0,t]$. Moreover, if there exists $t_n \to \infty$ such that $\phi_{t_n}^*(N)$ is still transversal to the fibers of the canonical projection $\pi^* : T^*M \to M$, for each n, then N is in fact invariant by the whole flow $\phi_t^*, t \in \mathbb{R}$.*

Proof. We first treat the case where N is the graph $\mathrm{Graph}(du)$ of the derivative of a C^1 function $u : M \to \mathbb{R}$. Using proposition 4.11.1, we see that the derivative of the Lipschitz function $T_t^- u$ wherever it exists is contained in $\phi_t^*(N)$. Since this last set is a (continuous) graph over the base, the derivative of $T_t^- u$ can be extended by continuity, hence $T_t^- u$ is also C^1 and $\phi_t^*(N) = \mathrm{Graph}(dT_t^- u)$. By theorem 6.2.2, it follows that $T_s^- u$ is C^1 and $\phi_t^* s(N) = \mathrm{Graph}(dT_s^- u)$, for each $s \in [0,t]$. If moreover, there exists a sequence $t_n \to \infty$ with $\phi_{t_n}^*(\mathrm{Graph}(du))$ transversal to the fibers of $\pi^* : T^*M \to M$, then $T_{t_n}^- u$ is everywhere smooth. By lemma 6.3.2, if $\epsilon > 0$ is given, then the graph $\mathrm{Graph}(dT_{t_n}^- u)$ is contained in $H^{-1}([c[0]-\epsilon, c[0]+\epsilon])$, for n large enough. Since H is invariant by the flow ϕ_t^* and $\mathrm{Graph}(dT_{t_n}^- u) = \phi_{t_n}^*(\mathrm{Graph}(du))$, it follows that $\mathrm{Graph}(du)$ is contained in $H^{-1}([c[0]-\epsilon, c[0]+\epsilon])$, for all $\epsilon > 0$. Hence H is constant on Graph(du), with value $c[0]$. By lemma 2.5.10, the graph of du is invariant under ϕ_t^*.

To treat the general case, we observe that $\pi^*|N : N \to M$ is a covering since N is compact and transversal to the fibers of the submersion π. We will now assume that $\pi^*|N$ is a diffeomorphism. Since N is Lagrangian, it is the graph of a closed form $\tilde{\omega}$ on M. We choose some C^∞ form ω which is cohomologous to $\tilde{\omega}$. If we introduce the Lagrangian $L_\omega : TM \to \mathbb{R}, (x, v) \mapsto L(x, v) - \omega_x(v)$, its Hamiltonian H_ω is $(x, p) \mapsto H(x, p + \omega_x)$. This means that $H_\omega = H \circ \bar{\omega}$, where $\bar{\omega} : T^*M \to T^*M, (x, p) \mapsto (x, p+\omega_x)$. Since ω is closed, the diffeomorphism $\bar{\omega}$ preserves the canonical symplectic form on T^*M, hence $\bar{\omega}$ conjugates $\phi_t^{\omega *}$, the Hamiltonian flow associated to H_ω and with ϕ_t^*, the Hamiltonian flow associated to H. Moreover $\bar{\omega}$ sends fibers of π^* onto fibers of π^*. It follows that $\phi_{t_n}^{\omega *}(\bar{\omega}^{-1}(N))$ is transversal to the fibers of π^n. Since N is the graph of $\tilde{\omega}$, its inverse $\bar{\omega}^{-1}(N)$, is the graph of $\tilde{\omega} - \omega$, which is exact by the choice of ω. Hence we can apply what we already proved to conclude that $\bar{\omega}^{-1}(N)$ is invariant under $\phi_t^{\omega *}$, from which we obtain that N is invariant under ϕ_t^*.

It remain to consider the case where the covering map $\pi^*|N : N \to M$ is not necessarily injective. To simplify notations, we set $p = \pi^*|N$. If we consider the tangent map $Tp : TN \to TM, (x, v) \mapsto (p(x), T_x p(v)$, we obtain a covering, which reduces to p on the 0-section identified to N. Since $T_x p$ is an isomorphism, for each $x \in N$, we can also define $(Tp)^* : T^*N \to T^*M, (x, p) \mapsto (p(x), p \circ T_x p^{-1}$. We define the Lagrangian $\tilde{L} = L \circ Tp : TN \to \mathbb{R}$. It is easy to see that $\tilde{L}$ is as differentiable as L, is superlinear in each fiber of TN, and for each $(x, v) \in TN$ $\partial^2 \tilde{L}/\partial v^2(x, v)$ is positive definite. Moreover, the conjugate Hamiltonian of $\tilde{L}$ is $\tilde{H} = H \circ (Tp)^*$. It is not difficult to check that the pullback of the Liouville form α_M on T^*M by $(Tp)^*$ is the Liouville form α_N on T^*N. If we call $\tilde{\phi}_t^*$, the Hamiltonian flow associated with $\tilde{H}$, it follows that $(Tp)^* \circ \tilde{\phi}_t^* = \phi_t^*$. We conclude that $\tilde{N} = [(Tp)^*]^{-1}(N)$ is a Lagrangian submanifold of T^*N, which is transversal to the fibers of the projection $\pi_N^* : T^*N \to N$, and $\tilde{\phi}_{t_n}(\tilde{N})$ is transversal to the fibers of π_N^*. Using the identity map of N, we see that we can find a section σ_0 of the covering map $\pi_N^* : \tilde{N} \to N$. If we call N_0 the image of σ_0, we see that we can apply the previous case to conclude that N_0 is invariant by $\tilde{\phi}_t$. Hence $N = (Tp)^*(N_0)$ is invariant by ϕ_t. $\square$

Corollary 6.4.2. *Suppose that $N \subset T^*M$ is a compact Lagrangian submanifold of T^*M which is everywhere transverse to the fibers of the canonical projection $\pi^* : T^*M \to M$. If $\phi^*_{t_0}(N) = N$ for some $t_0 \neq 0$ then N is in fact invariant by the whole flow $\phi^*_t, t \in \mathbb{R}$.*

Chapter 7

Viscosity Solutions

In this chapter, we will study the notion of viscosity solutions which was introduced by Crandall and Lions, see [CL83]. There are two excellent books on the subject by Guy Barles [Bar94] and another one by Martino Bardi and Italo Capuzzo-Dolceta [BCD97]. A first introduction to viscosity solutions can be found in Craig Evans book [Eva98]. Our treatment has been extremely influenced by the content of these three books. Besides introducing viscosity solutions, the main goal of this chapter is to show that, at least, for the Hamiltonians introduced in the previous chapters, the viscosity solutions and the weak KAM solutions are the same.

7.1 The different forms of Hamilton-Jacobi Equation

We will suppose that M is a fixed manifold, and that $H : T^*M \to \mathbb{R}$ is a continuous function , which we will call the Hamiltonian.

Definition 7.1.1 (Stationary HJE)**.** The Hamilton-Jacobi associated to H is the equation

$$H(x, d_x u) = c,$$

where c is some constant.

A classical solution of the Hamilton-Jacobi equation $H(x, d_x u) = c$ (HJE in short) on the open subset U of M is a C^1 map $u : U \to \mathbb{R}$ such that $H(x, d_x u) = c$, for each $x \in U$.

We will deal usually only with the case $H(x, d_x u) = 0$, since we can reduce the general case to that case if we replace the Hamiltonian H by H_c defined by $H_c(x,p) = H(x,p) - c$.

Definition 7.1.2 (Evolutionary HJE)**.** The evolutionary Hamilton-Jacobi associated equation to the Hamiltonian H is the equation

$$\frac{\partial u}{\partial t}(t,x) + H\big(x, \frac{\partial u}{\partial x}(t,x)\big) = 0.$$

A classical solution to this evolutionary Hamilton-Jacobi equation on the open subset W of $\mathbb{R} \times T^*M$ is a C^1 map $u : W \to \mathbb{R}$ such that $\frac{\partial u}{\partial t}(t,x) + H\big(x, \frac{\partial u}{\partial x}(t,x)\big) = 0$, for each $(t,x) \in W$.

The evolutionary form can be reduced to the stationary form by introducing the Hamiltonian $\tilde{H} : T^*(\mathbb{R} \times M)$ defined by

$$\tilde{H}(t,x,s,p) = s + H(x,p),$$

where $(t,x) \in \mathbb{R} \times M$, and $(s,p) \in T^*_{t,x}(\mathbb{R} \times M) = \mathbb{R} \times T^*_x M$.

It is also possible to consider a time dependent Hamiltonian defined on an open subset of M. Consider for example a Hamiltonian $H : \mathbb{R} \times TM^* \to \mathbb{R}$, the evolutionary form of the HJE for that Hamiltonian is

$$\frac{\partial u}{\partial t}(t,x) + H\big(t, x, \frac{\partial u}{\partial x}(t,x)\big) = 0.$$

A classical solution of that equation on the open subset W of $\mathbb{R} \times M$ is, of course, a C^1 map $u : W \to \mathbb{R}$ such that $\frac{\partial u}{\partial t}(t,x) + H\big(t, x, \frac{\partial u}{\partial x}(t,x)\big) = 0$, for each $(t,x) \in W$. This form of the Hamilton-Jacobi equation can also be reduced to the stationary form by introducing the Hamiltonian $\tilde{H} : T^*(\mathbb{R} \times M) \to \mathbb{R}$ defined by

$$\tilde{H}(t,x,s,p) = s + H(t,x,p).$$

7.2 Viscosity Solutions

We will suppose in this section that M is a manifold and $H : T^*M \to M$ is a Hamiltonian.

As we said in the introduction of this book, it is usually impossible to find global C^1 solutions of the Hamilton-Jacobi equation $H(x, d_x u) = c$. One has to admit more general functions. A first attempt is to consider Lipschitz functions.

Definition 7.2.1 (Very Weak Solution)**.** We will say that $u : M \to \mathbb{R}$ is a very weak solution of $H(x, d_x u) = c$, if it is Lipschitz, and $H(x, d_x u) = c$ almost everywhere (this makes sense since the derivative of u exists almost everywhere by Rademacher's theorem).

This is too general because it gives too many solutions. A notion of weak solution is useful if it gives a unique, or at least a small number of solutions. This is not satisfied by this notion of very weak solution as can be seen in the following example.

Example 7.2.2. We suppose $M = \mathbb{R}$, so $T^*M = \mathbb{R} \times \mathbb{R}$, and we take $H(x, p) = p^2 - 1$. Then any continuous piecewise C^1 function u with derivative taking only the values ± 1 is a very weak solution of $H(x, d_x u) = 0$. This is already too huge, but there are even more very weak solutions. In fact, if A is any measurable subset of $\mathbb{R}$, then the function

$$f_A(x) = \int_0^x 2\chi_A(t) - 1\, dt,$$

where χ_A is the characteristic function of A, is Lipschitz with derivative ± 1 almost everywhere.

Therefore we have to define a more stringent notion of solutions. Crandall and Lions have introduced the notion of viscosity solutions, see [CL83] and [CEL84].

Definition 7.2.3 (Viscosity solution)**.** A function $u : V \to \mathbb{R}$ is a viscosity *subsolution* of $H(x, d_x u) = c$ on the open subset $V \subset M$, if for every C^1 function $\phi : V \to \mathbb{R}$ and every point $x_0 \in V$ such that $u - \phi$ has a *maximum* at x_0, we have $H(x_0, d_{x_0}\phi) \leq c$.

A function $u : V \to \mathbb{R}$ is a viscosity *supersolution* of $H(x, d_x u) = c$ on the open subset $V \subset M$, if for every C^1 function $\psi : V \to \mathbb{R}$ and every point $y_0 \in V$ such that $u - \psi$ has a *minimum* at y_0, we have $H(y_0, d_{y_0}\psi) \geq c$.

A function $u : V \to \mathbb{R}$ is a viscosity *solution* of $H(x, d_x u) = c$ on the open subset $V \subset M$, if it is *both* a subsolution and a supersolution.

This definition is reminiscent of the definition of distributions: since we cannot restrict to differentiable functions, we use *test* functions (namely ϕ or ψ) which are smooth and on which we can test the condition. We first see that this is indeed a generalization of classical solutions.

Theorem 7.2.4. *A* C^1 *function* $u : V \to \mathbb{R}$ *is a viscosity solution of* $H(x, d_x u) = c$ *on* V *if and only if it is a classical solution.*

In fact, the C^1 *function* u *is a viscosity subsolution (resp. supersolution) of* $H(x, d_x u) = c$ *on* V *if and only* $H(x, d_x u) \leq c$ *(resp.* $H(x, d_x u) \geq c$*), for each* $x \in V$.

Proof. We will prove the statement about the subsolution case. Suppose that the C^1 function u is a viscosity subsolution. Since u is C^1, we can use it as a test function. But $u - u = 0$, therefore every $x \in V$ is a maximum, hence $H(x, d_x u) \leq c$ for each $x \in V$.

Conversely, suppose $H(x, d_x u) \leq c$ for each $x \in V$. If $\phi : V \to \mathbb{R}$ is C^1 and $u - \phi$ has a maximum at x_0, then the differentiable function $u - \phi$ must have derivative 0 at the maximum x_0. Therefore $d_{x_0}\phi = d_{x_0}u$, and $H(x, d_{x_0}\phi) = H(x, d_{x_0}u) \leq c$. □

To get a feeling for these viscosity notions, it is better to restate slightly the definitions. We first remark that the condition imposed on the test functions (ϕ or ψ) in the definition above is on the derivative, therefore, to check the condition, we can change our test function by a constant. Suppose now that ϕ (resp. ψ) is C^1 and $u - \phi$ (resp. $u - \psi$) has a maximum (resp. minimum) at x_0 (resp. y_0), this means that $u(x_0) - \phi(x_0) \geq u(x) - \phi(x)$ (resp. $u(y_0) - \psi(y_0) \leq u(x) - \phi(x)$). As we said, since we can add to ϕ (resp. ψ) the constant $u(x_0) - \phi(x_0)$ (resp. $u(y_0) - \psi(y_0)$), these conditions can be replaced by $\phi \geq u$ (resp. $\psi \leq u$) and $u(x_0) = \phi(x_0)$ (resp. $u(y_0) = \psi(y_0)$). Therefore we obtain an equivalent definition for subsolution and supersolution.

Definition 7.2.5 (Viscosity Solution)**.** A function $u : V \to \mathbb{R}$ is a *subsolution* of $H(x, d_x u) = c$ on the open subset $V \subset M$, if for every C^1 function $\phi : V \to \mathbb{R}$, with $\phi \geq u$ everywhere, at every point $x_0 \in V$ where $u(x_0) = \phi(x_0)$ we have $H(x_0, d_{x_0}\phi) \leq c$, see figure 7.1.

A function $u : V \to \mathbb{R}$ is a *supersolution* of $H(x, d_x u) = c$ on the open subset $V \subset M$, if for every C^1 function $\psi : V \to \mathbb{R}$, with $u \geq \psi$ everywhere, at every point $y_0 \in V$ where $u(y_0) = \psi(y_0)$ we have $H(y_0, d_{y_0}\psi) \geq c$, see figure 7.2.

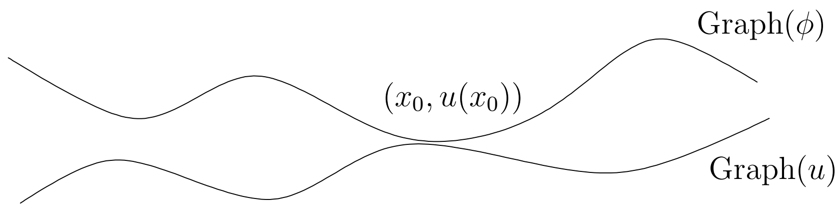

Figure 7.1: Subsolution: $\phi \geq u, u(x_0) = \phi(x_0) \Rightarrow H(x_0, d_{x_0}\phi) \leq c$

To see what the viscosity conditions mean we test them on the example 7.2.2 given above.

Example 7.2.6. We suppose $M = \mathbb{R}$, so $T^*M = \mathbb{R} \times \mathbb{R}$, and we take $H(x,p) = p^2 - 1$. Any Lipschitz function $u : \mathbb{R} \to \mathbb{R}$ with Lipschitz constant ≤ 1 is in fact a viscosity subsolution of $H(x, d_x u) = 0$. To check this consider ϕ a C^1 function and $x_0 \in \mathbb{R}$ such that $\phi(x_0) = u(x_0)$ and $\phi(x) \geq u(x)$, for $x \in \mathbb{R}$. We can write

$$\phi(x) - \phi(x_0) \geq u(x) - u(x_0) \geq -|x - x_0|.$$

For $x > x_0$, this gives

$$\frac{\phi(x) - \phi(x_0)}{x - x_0} \geq -1,$$

hence passing to the limit $\phi'(x_0) \geq -1$. On the other hand, if $(x - x_0) < 0$ we obtain

$$\frac{\phi(x) - \phi(x_0)}{x - x_0} \leq 1,$$

hence $\phi'(x_0) \leq 1$.This yields $|\phi'(x_0)| \leq 1$, and therefore

$$H(x_0, \phi'(x_0)) = |\phi'(x_0)|^2 - 1 \leq 0.$$

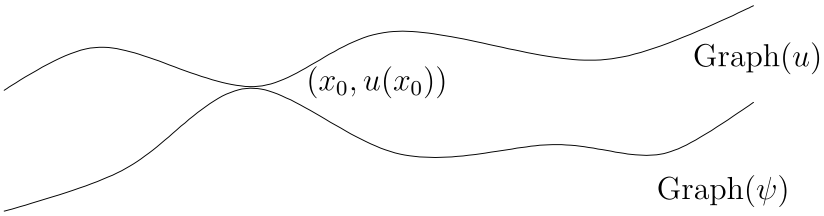

Figure 7.2: Supersolution: $\psi \leq u, u(x_0) = \psi(x_0) \Rightarrow H(x_0, d_{x_0}\psi) \geq c$

So in fact, any very weak subsolution (i.e. a Lipschitz function u such that $H(x, d_x, u) \leq 0$ almost everywhere) is a viscosity subsolution. This is due to the fact that, in this example, the Hamiltonian is convex in p, see 8.3.4 below.

Of course, the two smooth functions $x \mapsto x$, and $x \mapsto -x$ are the only two classical solutions in that example. It is easy to check that the absolute value function $x \mapsto |x|$, which is a subsolution and even a solution on $\mathbb{R} \setminus \{0\}$ (where it is smooth and a classical solution), is not a viscosity solution on the whole of $\mathbb{R}$. In fact the constant function equal to 0 is less than the absolute value everywhere with equality at 0, but we have $H(0, 0) = -1 < 0$, and this violates the supersolution condition.

The function $x \mapsto -|x|$ is a viscosity solution. It is smooth and a classical solution on $\mathbb{R} \setminus \{0\}$. It is a subsolution everywhere. Moreover, any function ϕ with $\phi(0) = 0$ and $\phi(x) \leq -|x|$ everywhere cannot be differentiable at 0. This is obvious on a picture of the graphs, see figure 7.3. Formally it results from the fact that both $\phi(x) - x$ and $\phi(x) + x$ have a maximum at 0.

We now establish part of the relationship between viscosity solutions and weak KAM solutions.

Proposition 7.2.7. *Let $L : TM \to \mathbb{R}$ be a Tonelli Lagrangian on the compact manifold M. If the function $u : V \to \mathbb{R}$, defined on the open subset $V \subset M$, is dominated by $L+c$, then u is a viscosity*

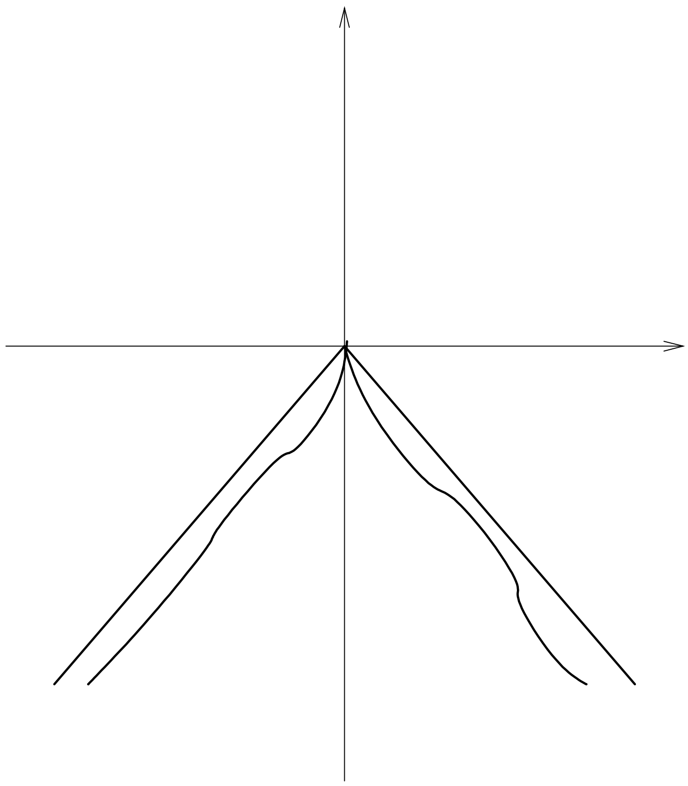

Figure 7.3: Graphs of $\psi(x) \leq -|x|$ with $\psi(0) = 0$.

subsolution of $H(x, d_x u) = c$ *on* V*, where* H *is the Hamiltonian associated to* L*, i.e.* $H(x,p) = \sup_{v \in T_x M} \langle p, v\rangle - L(x,v)$*.*

Moreover, any weak KAM solution $u_- \in \mathcal{S}_-$ *is a viscosity solution of* $H(x, d_x u) = c[0]$*.*

Proof. To prove the first part, let $\phi : V \to \mathbb{R}$ be C^1, and such that $u \leq \phi$ with equality at x_0. This implies $\phi(x_0) - \phi(x) \leq u(x_0) - u(x)$. Fix $v \in T_{x_0}M$ and choose $\gamma :]-\delta, \delta[\to M$, a C^1 path with $\gamma(0) = x_0$, and $\dot\gamma(0) = v$. For $t \in]-\delta, 0[$, we obtain

$$\begin{aligned} \phi(\gamma(0)) - \phi(\gamma(t)) &\leq u(\gamma(0)) - u(\gamma(t)) \\ &\leq \int_t^0 L(\gamma(s), \dot\gamma(s))\, ds - ct. \end{aligned}$$

Dividing by $-t > 0$ yields

$$\frac{\phi(\gamma(t)) - \phi(\gamma(0))}{t} \leq \frac{1}{-t} \int_t^0 L(\gamma(s), \dot\gamma(s))\, ds + c.$$

If we let $t \to 0$, we obtain $d_{x_0}\phi(v) \leq L(x_0, v) + c$, hence

$$H(x_0, d_{x_0}\phi) = \sup_{v \in T_{x_0}M} d_{x_0}\phi(v) - L(x_0, v) \leq c.$$

This shows that u is a viscosity subsolution.

To prove that $u_- \in \mathcal{S}_-$ is a viscosity solution, it remains to show that it is a supersolution of $H(x, d_x u_-) = c[0]$. Suppose that $\psi : M \to \mathbb{R}$ is C^1, and that $u_- \geq \psi$ everywhere with $u_-(x_0) = \psi(x_0)$. We have $\psi(x_0) - \psi(x) \geq u_-(x_0) - u_-(x)$, for each $x \in M$. We pick a C^1 path $\gamma :]-\infty, 0] \to M$, with $\gamma(0) = x_0$, and such that

$$\forall t \leq 0, u_-(\gamma(0)) - u_-(\gamma(t)) = \int_t^0 L(\gamma(s), \dot\gamma(s))\, ds - c[0]t.$$

Therefore

$$\psi(\gamma(0)) - \psi(\gamma(t)) \geq \int_t^0 L(\gamma(s), \dot\gamma(s))\, ds - c[0]t.$$

If, for $t < 0$, we divide both sides by $-t > 0$, we obtain

$$\frac{\psi(\gamma(t)) - \psi(\gamma(0))}{t} \geq \frac{1}{-t} \int_t^0 L(\gamma(s), \dot\gamma(s))\, ds + c[0].$$

If we let t tend to 0, this yields $d_{x_0}\psi(\dot\gamma(0)) \geq L(x_0, \dot\gamma(0)) + c[0]$; hence $H(x_0, d_{x_0}\psi) \geq d_{x_0}\psi(\dot\gamma(0)) - L(x_0, \dot\gamma(0)) \geq c[0]$. □

The following theorem will be needed to prove the converse of proposition 7.2.7.

Theorem 7.2.8. *Suppose $L : TM \to \mathbb{R}$ is a Tonelli Lagrangian on the compact manifold M. Let T_t^- be the associated Lax-Oleinik semi-group. If $u \in \mathcal{C}^0(M, \mathbb{R})$, then the continuous function $U : [0, +\infty[\times M \to \mathbb{R}$ defined by $U(t, x) = T_t^- u(x)$ is a viscosity solution of*

$$\frac{\partial U}{\partial t}(t, x) + H(x, \frac{\partial U}{\partial x}(t, x)) = 0$$

on the open set $]0, +\infty[\times M$, where $H : T^*M \to \mathbb{R}$ is the Hamiltonian associated to L, i.e. $H(x, p) = \sup_{v \in T_x M} p(v) - L(x, v)$.

Proof. Suppose that $\gamma : [a, b] \to M$. Since $T_b^- u = T_{b-a}^-[T_a^-(u)]$, using the definition of T_{b-a}^-, we get

$$\begin{aligned} T_b^- u(\gamma(a)) &= T_{b-a}^-[T_a^-(u)](\gamma(a) \\ &\leq T_a^- u(\gamma(a)) + \int_a^b L(\gamma(s), \dot\gamma(s))\, ds, \end{aligned}$$

therefore

$$U(b, \gamma(b)) - U(a, \gamma(a)) \leq \int_a^b L(\gamma(s), \dot\gamma(s))\, ds, \qquad (*)$$

We now show that U is a viscosity subsolution. Suppose $\phi \geq U$, with ϕ of class C^1 and $(t_0, x_0) = U(t_0, x_0)$, where $t_0 > 0$. Fix $v \in T_{x_0}M$, and pick a C^1 curve $\gamma : [0, t_0] \to M$ such that $(\gamma(t_0), \dot\gamma(t_0)) = (x, v)$.

If $0 \leq t \leq t_0$, we have by $(*)$ and therefore

$$U(t_0, \gamma(t_0)) - U(t, \gamma(t)) \leq \int_t^{t_0} L(\gamma(s), \dot\gamma(s))\, ds. \qquad (**)$$

Since $\phi \geq U$, with equality at (t_0, x_0), noticing that $\gamma(t_0) = x_0$, we obtain from $(**)$

$$\forall t \in]0, t_0[, \phi(t_0, \gamma(t_0)) - \phi(t, \gamma(t)) \leq \int_t^{t_0} L(\gamma(s), \dot\gamma(s))\, ds.$$

Dividing by $t_0 - t > 0$, and letting $t \to t_0$, we get

$$\forall v \in T_{x_0}M, \frac{\partial \phi}{\partial t}(t_0, x_0) + \frac{\partial \phi}{\partial x}(t_0, x_0)(v) \leq L(x_0, v).$$

By Fenchel's formula 1.3.1

$$H(x_0, \frac{\partial \phi}{\partial x}(t_0, x_0)) = \sup_{v \in T_{x_0}M} \frac{\partial \phi}{\partial x}(t_0, x_0)(v) - L(x_0, v),$$

therefore

$$\frac{\partial \phi}{\partial t}(t_0, x_0) + H(x_0, \frac{\partial \phi}{\partial x}(t_0, x_0)) \leq 0.$$

To prove that U is a supersolution, we consider $\psi \leq U$, with ψ of class C^1. Suppose $U(t_0, x_0) = \psi(t_0, x_0)$, with $t_0 > 0$.

We pick $\gamma : [0, t_0] \to M$ such that $\gamma(t_0) = x_0$ and

$$U(t_0, x_0) = T^-_{t_0} u(x_0) = u(\gamma(0)) + \int_0^{t_0} L(\gamma(s), \dot{\gamma}(s))\, ds.$$

Since $U(0, \gamma(0)) = u(\gamma(0))$, this can be rewritten as

$$U(t_0, x_0) - U(0, \gamma(0)) = \int_0^{t_0} L(\gamma(s), \dot{\gamma}(s))\, ds. \qquad (***)$$

Applying (*) above twice, we obtain

$$U(t_0, x_0) - U(t, \gamma(t)) \leq \int_0^{t_0} L(\gamma(s), \dot{\gamma}(s))\, ds$$
$$U(t, \gamma(t)) - U(0, \gamma(0)) \leq \int_0^{t_0} L(\gamma(s), \dot{\gamma}(s))\, ds.$$

Adding this two inequalities we get in fact by (**) an equality, hence we must have

$$\forall t \in [0, t_0], U(t_0, \gamma(t_0)) - U(t, \gamma(t)) = \int_t^{t_0} L(\gamma(s), \dot{\gamma}(s))\, ds.$$

Since $\psi \leq U$, with equality at (t_0, x_0), we obtain

$$\psi(t_0, \gamma(t_0)) - \psi(t, \gamma(t)) \geq \int_t^{t_0} L(\gamma(s), \dot{\gamma}(s))\, ds.$$

Dividing by $t_0 - t > 0$, and letting $t \to t_0$, we get

$$\frac{\partial \psi}{\partial t}(t_0, x_0) + \frac{\partial \psi}{\partial x}(t_0, x_0)(\dot{\gamma}(t_0)) \geq L(x_0, \dot{\gamma}(t_0)).$$

By Fenchel's formula 1.3.1

$$H(x_0, \frac{\partial \psi}{\partial x}(t_0, x_0)) \geq \frac{\partial \psi}{\partial x}(t_0, x_0)(\dot\gamma(t_0)) - L(x_0, \dot\gamma(t_0)),$$

Therefore

$$\frac{\partial \psi}{\partial t}(t_0, x_0) + H(x_0, \frac{\partial \psi}{\partial x}(t_0, x_0)) \geq 0. \qquad \square$$

7.3 Lower and upper differentials

We need to introduce the notion of lower and upper differentials.

Definition 7.3.1. If $u : M \to \mathbb{R}$ is a map defined on the manifold M, we say that the linear form $p \in T^*_{x_0}M$ is a lower (resp. upper) differential of u at $x_0 \in M$, if we can find a neighborhood V of x_0 and a function $\phi : V \to \mathbb{R}$, differentiable at x_0, with $\phi(x_0) = u(x_0)$ and $d_{x_0}\phi = p$, and such that $\phi(x) \leq u(x)$ (resp. $\phi(x) \geq u(x)$), for every $x \in V$.

We denote by $D^-u(x_0)$ (resp. $D^+u(x_0)$) the set of lower (resp. upper) differentials of u at x_0.

Exercise 7.3.2. *Consider the function $u : \mathbb{R} \to \mathbb{R}, x \mapsto |x|$, for each $x \in \mathbb{R}$, find $D^-u(x)$, and $D^+u(x)$. Same question with $u(x) = -|x|$.*

Definition 7.3.1 is not the one usually given for M an open set of an Euclidean space, see [Bar94], [BCD97] or [Cla90]. It is nevertheless equivalent to the usual definition as we now show.

Proposition 7.3.3. *Let $u : U \to \mathbb{R}$ be a function defined on the open subset U of $\mathbb{R}^n$, then the linear form p is in $D^-u(x_0)$ if and only if*

$$\liminf_{x \to x_0} \frac{u(x) - u(x_0) - p(x - x_0)}{\|x - x_0\|} \geq 0.$$

In the same way $p \in D^+u(x_0)$ if and only if

$$\limsup_{x \to x_0} \frac{u(x) - u(x_0) - p(x - x_0)}{\|x - x_0\|} \leq 0.$$

Proof. Suppose $p \in D^-u(x_0)$, we can find a neighborhood V of x_0 and a function $\phi : V \to \mathbb{R}$, differentiable at x_0, with $\phi(x_0) = u(x_0)$ and $d_{x_0}\phi = p$, and such that $\phi(x) \leq u(x)$, for every $x \in V$. Therefore, for $x \in V$, we can write

$$\frac{\phi(x) - \phi(x_0) - p(x - x_0)}{\|x - x_0\|} \leq \frac{u(x) - u(x_0) - p(x - x_0)}{\|x - x_0\|}.$$

Since $p = d_{x_0}\phi$ the left hand side tends to 0, therefore

$$\liminf_{x \to x_0} \frac{u(x) - u(x_0) - p(x - x_0)}{\|x - x_0\|} \geq 0.$$

Suppose conversely, that $p \in \mathbb{R}^{n*}$ satisfies

$$\liminf_{x \to x_0} \frac{(u(x) - u(x_0) - p(x - x_0))}{\|x - x_0\|} \geq 0.$$

We pick $r > 0$ such that the ball $\mathring{B}(x_0, r) \subset U$, and for $h \in \mathbb{R}^n$ such that $0 < \|h\| < r$, we set

$$\epsilon(h) = \min(0, \frac{u(x_0 + h) - u(x_0) - p(h)}{\|h\|}).$$

It is easy to see that $\lim_{h \to 0} \epsilon(h) = 0$. We can therefore set $\epsilon(0) = 0$. The function $\phi : \mathring{B}(x_0, r) \to \mathbb{R}$, defined by $\phi(x) = u(x_0) + p(x - x_0) + \|x - x_0\|\epsilon(x - x_0)$, is differentiable at x_0, with derivative p, it is equal to u at x_0 and satisfies $\phi(x) \leq u(x)$, for every $x \in \mathring{B}(x_0, r)$. □

Proposition 7.3.4. *Let $u : M \to \mathbb{R}$ be a function defined on the manifold M.*

(i) *For each x in M, we have $D^+u(x) = -D^-(-u)(x) = \{-p \mid p \in D^-(-u)(x)\}$ and $D^-u(x) = -D^+(-u)(x)$.*

(ii) *For each x in M, both sets $D^+u(x), D^-u(x)$ are closed convex subsets of T_x^*M.*

(iii) *If u is differentiable at x, then $D^+u(x) = D^-u(x) = \{d_xu\}$.*

(iv) *If both sets $D^+u(x), D^-u(x)$ are non-empty then u is differentiable at x.*

(v) *if* $v : M \to \mathbb{R}$ *is a function with* $v \leq u$ *and* $v(x) = u(x)$*, then* $D^-v(x) \subset D^-u(x)$ *and* $D^+v(x) \supset D^+u(x)$*.*

(vi) *If* U *is an open convex subset of an Euclidean space and* $u : U \to \mathbb{R}$ *is convex then* $D^-u(x)$ *is the set of supporting linear forms of* u *at* $x \in U$*. In particular* $D^+u(x) \neq \emptyset$ *if and only if* u *is differentiable at* x*.*

(vii) *Suppose* M *has a distance* d *obtained from the Riemannian metric* g*. If* $u : M \to \mathbb{R}$ *is Lipschitz for* d *with Lipschitz constant* $\mathrm{Lip}(u)$*, then for any* $p \in D^{\pm}u(x)$ *we have* $\|p\|_x \leq \mathrm{Lip}(u)$*.*

In particular, if M *is compact then the sets* $D^{\pm}u = \{(x,p) \mid p \in D^{\pm}u(x), x \in M\}$ *are compact.*

Proof. Part (i) and and the convexity claim in part (ii) are obvious from the definition 7.3.1.

To prove the fact that $D^+u(x_0)$ is closed for a given for $x_0 \in M$, we can assume that M is an open subset of $\mathbb{R}^k$. We will apply proposition 7.3.3. If $p_n \in D^+u(x_0)$ converges to $p \in \mathbb{R}^{k*}$, we can write

$$\frac{u(x) - u(x_0) - p(x - x_0)}{\|x - x_0\|} \leq \frac{u(x) - u(x_0) - p_n(x - x_0)}{\|x - x_0\|} + \|p_n - p\|.$$

Fixing n, and letting $x \to x_0$, we obtain

$$\limsup_{x \to x_0} \frac{u(x) - u(x_0) - p(x - x_0)}{\|x - x_0\|} \leq \|p_n - p\|.$$

If we let $n \to \infty$, we see that $p \in D^+u(x_0)$.

We now prove (iii) and (iv) together. If u is differentiable at $x_0 \in M$ then obviously $d_{x_0}u \in D^+u(x_0) \cap D^-u(x_0)$. Suppose now that both $D^+u(x_0)$ and $D^-u(x_0)$ are both not empty, pick $p_+ \in D^+u(x_0)$ and $p_- \in D^-u(x_0)$. For h small, we have

$$p_-(h) + \|h\|\epsilon_-(h) \leq u(x_0 + h) - u(x_0) \leq p_+(h) + \|h\|\epsilon_+(h), \quad (*)$$

where both $\epsilon_-(h)$ and $\epsilon_+(h)$ tend to 0, a $h \to 0$. If $v \in \mathbb{R}^n$, for $t > 0$ small enough, we can replace h by tv in the inequalities (*) above. Forgetting the middle term and dividing by t, we obtain

$$p_-(v) + \|v\|\epsilon_-(tv) \leq p_+(v) + \|v\|\epsilon_+(tv),$$

letting t tend to 0, we see that $p_-(v)+ \leq p_+(v)$, for every $v \in \mathbb{R}^n$. Replacing v by $-v$ gives the reverse inequality $p_+(v)+ \leq p_-(v)$, therefore $p_- = p_+$. This implies that both $D^+u(x_0)$ and $D^-u(x_0)$ are reduced to the same singleton $\{p\}$. The inequality (*) above now gives

$$p(h) + \|h\|\epsilon_-(h) \leq u(x_0 + h) - u(x_0) \leq p(h) + \|h\|\epsilon_+(h),$$

this clearly implies that p is the derivative of u at x_0.

Part (v) follows routinely from the definition.

To prove (vi), we remark that by convexity $u(x_0 + th) \leq (1 - t)u(x_0) + tu(x_0 + h)$, therefore

$$u(x_0 + h) - u(x_0) \geq \frac{u(x_0 + th) - u(x_0)}{t}.$$

If p is a linear form we obtain

$$\frac{u(x_0 + h) - u(x_0) - p(h)}{\|h\|} \geq \frac{u(x_0 + th) - u(x_0) - p(h)}{\|th\|}.$$

If $p \in D^-u(x_0)$, then the $\liminf$ as $t \to 0$ of the right hand side is ≥ 0, therefore $u(x_0 + h) - u(x_0) - p(h) \geq 0$, which shows that p is a supporting linear form. Conversely, a supporting linear form is clearly a lower differential.

It remains to prove (vii). Suppose, for example that $\phi : V \to \mathbb{R}$ is defined on some neighborhood V of a given $x_0 \in M$, that it is differentiable at x_0, and that $\phi \geq u$ on V, with equality at x_0. If $v \in T_{x_0}M$ is given, we pick a C^1 path $\gamma : [0, \delta] \to V$, with $\delta > 0, \gamma(0) = x_0$, and $\dot\gamma(0) = v$.We have

$$\begin{aligned}\forall t \in [0, \delta], |u(\gamma(t)) - u(x_0)| &\leq \mathrm{Lip}(u) d(\gamma(t), x_0)\\ &\mathrm{Lip}(u) \int_0^t \|\dot\gamma(s)\|\, ds.\end{aligned}$$

Therefore $u(\gamma(t)) - u(x_0) \geq -\mathrm{Lip}(u) \int_0^t \|\dot\gamma(s)\|\, ds$. Since $\phi \geq u$ on V, with equality at x_0, it follows that

$$\phi(\gamma(t)) - \phi(x_0) \geq -\mathrm{Lip}(u) \int_0^t \|\dot\gamma(s)\|\, ds.$$

Dividing by $t > 0$, and letting $t \to 0$, we get

$$d_{x_0}\phi(v) \geq -\operatorname{Lip}(u)\|v\|.$$

Since $v \in T_{x_0}M$ is arbitrary, we can change v into $-v$ in the inequality above to conclude that we also have

$$d_{x_0}\phi(v) \leq \operatorname{Lip}(u)\|v\|.$$

It then follows that $\|d_{x_0}\phi\| \leq \operatorname{Lip}(u)$. □

Lemma 7.3.5. *If $u : M \to \mathbb{R}$ is continuous and $p \in D^+u(x_0)$ (resp. $p \in D^-u(x_0)$), there exists a C^1 function $\phi : M \to \mathbb{R}$, such that $\phi(x_0) = u(x_0), d_{x_0}phi = p$, and $\phi(x) > u(x)$ (resp. $\phi(x) < u(x)$) for $x \neq x_0$.*

Moreover, if W is any neighborhood of x_0 and $C > 0$, we can choose ϕ such that $\phi(x) \geq u(x) + C$, for $x \notin W$ (resp. $\phi(x) \leq u(x) - C$).

Proof. Assume first $M = \mathbb{R}^k$. To simplify notations, we can assume $x_0 = 0$. Moreover, subtracting from u the affine function $x \mapsto u(0) + p(x)$. We can assume $u(0) = 0$ and $p = 0$. The fact that $0 \in D^+u(0)$ gives

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

If we take the non-negative part $u^+(x) = \max(u(x), 0)$ of u, this gives

$$\lim_{x \to 0} \frac{u^+(x)}{\|x\|} = 0. \qquad (\spadesuit)$$

If we set

$$c_n = \sup\{u^+(x) \mid 2^{-(n+1)} \leq \|x\| \leq 2^{-n}\}$$

then c_n is finite and ≥ 0, because $u^+ \geq 0$ is continuous. Moreover using that $2^n u_+(x) \leq u^+(x)/\|x\|$, for $\|x\| \leq 2^{-n}$, and the limit in $(\spadesuit)$ above, we obtain

$$\lim_{n \to \infty} [\sup_{m \geq n} 2^m c_m] = 0. \qquad (\heartsuit)$$

We now consider $\theta : \mathbb{R}^k \to \mathbb{R}$ a C^∞ bump function with $\theta = 1$ on the set $\{x \in \mathbb{R}^k \mid 1/2 \leq \|x\| \leq 1\}$, and whose support is contained in $\{x \in \mathbb{R}^k \mid 1/4 \leq \|x\| \leq 2\}$. We define the function $\psi : \mathbb{R}^k \to \mathbb{R}$ by

$$\psi(x) = \sum_{n\in\mathbb{Z}} (c_n + 2^{-2n})\theta(2^n x).$$

This function is well defined at 0 because every term is then 0. For $x \neq 0$, we have $\theta(2^n x) \neq 0$ only if $1/4 < \|2^n x\| < 2$. Taking the logarithm in base 2, this can happen only if $-2 - \log_2\|x\| < n < 1 - \log_2\|x\|$, therefore this can happen for at most 3 consecutive integers n, hence the sum is also well defined for $x \neq 0$. Moreover, if $x \neq 0$, the set $V_x = \{y \neq 0 \mid -1 - \log_2\|x\| < -\log_2\|y\| < 1 - \log_2\|x\|\}$ is a neighborhood of x and

$$\forall y \in V_y, \psi(y) = \sum_{-3-\log_2\|x\|<n<2-\log_2\|x\|} (c_n + 2^{-2n})\theta(2^n y). \qquad (*)$$

This sum is finite with at most 5 terms, therefore θ is C^∞ on $\mathbb{R}^k \setminus \{0\}$.

We now check that ψ is continuous at 0. Using equation (*), and the limit ($\heartsuit$) we see that

$$\begin{aligned} 0 \leq \psi(x) &\leq \sum_{-3-\log_2\|x\|<n<2-\log_2\|x\|} (c_n + 2^{-2n}) \\ &\leq 5 \sup_{n>-3-\log_2\|x\|} (c_n + 2^{-2n}) \to 0 \text{ as } x \to 0. \end{aligned}$$

To show that ψ is C^1 on the whole of $\mathbb{R}^k$ with derivative 0 at 0, it suffices to show that $d_x\psi$ tends to 0 as $\|x\| \to 0$. Differentiating equation (*) we see that

$$d_x\psi = \sum_{-3-\log_2\|x\|<n<2-\log_2\|x\|} (c_n + 2^{-2n})2^n d_{2^n x}\theta.$$

Since θ has compact support $K = \sup_{x\in\mathbb{R}^n}\|d_x\theta\|$ is finite. The equality above and the limit in ($\heartsuit$) give

$$\|d_x\psi\| \leq 5K \sup\{2^n c_n + 2^{-n} \mid n \geq -2 - \log_2\|x\|\},$$

but the right hand side goes to 0 when $\|x\| \to 0$.

We now show $\psi(x) > u(x)$, for $x \neq 0$. There is an integer n_0 such that $\|x\| \in [2^{-n_0+1}, 2^{-n_0}]$, hence $\theta(2^{n_0}x) = 1$ and $\psi(x) \geq \theta(2^{n_0}x)(c_{n_0} + 2^{-2n_0}) \geq c_{n_0} + 2^{-2n_0}$, since $c_{n_0} = \sup\{u^+(y) \mid \|y\| \in [2^{(-n_0+1)}, 2^{-n_0}]\}$, we obtain $c_{n_0} \geq u^+(x)$ and therefore $\psi(x) > u^+(x) \geq u(x)$.

It remains to show that we can get rid of the assumption $M = \mathbb{R}^k$, and to show how to obtain the desired inequality on the complement of W. We pick a small open neighborhood $U \subset W$ of x_0 which is diffeomorphic to an Euclidean space. By what we have done, we can find a C^1 function $\psi : U \to \mathbb{R}$ with $\psi(x_0) = u(x_0), d_{x_0}\psi = p$, and $\psi(x) > u(x)$, for $x \in U \setminus \{x_0\}$. We then take a C^∞ bump function $\varphi : M \to [0,1]$ which is equal to 1 on a neighborhood of x_0 and has compact support contained in $U \subset W$. We can find a C^∞ function $\tilde{\psi} : M \to \mathbb{R}$ such that $\tilde{\psi} \geq u + C$. It is easy to check that the function $\phi : M \to \mathbb{R}$ defined by $\phi(x) = (1 - \varphi(x))\tilde{\psi}(x) + \varphi(x)\psi(x)$ has the required property. $\square$

The following simple lemma is very useful.

Lemma 7.3.6. *Suppose $\psi : M \to \mathbb{R}$ is C^r, with $r \geq 0$. If $x_0 \in M, C \geq 0$, and W is a neighborhood of x_0, there exist two C^r functions $\psi_+, \psi_- : M \to \mathbb{R}$, such that $\psi_+(x_0) = \psi_-(x_0) = \psi(x_0)$, and $\psi_+(x) > \psi(x) > \psi_-(x)$, for $x \neq x_0$. Moreover $\psi_+(x) - C > \psi(x) > \psi_-(x) + C$, for $x \notin W$. If $r \geq 1$, then necessarily $d_{x_0}\psi_+ = d_{x_0}\psi_- = d_{x_0}\tilde{\psi}$*

Proof. The last fact is clear since $\psi_+ - \psi$ (resp. $\psi_- - \psi$) achieves a minimum (resp. maximum) at x_0.

Using the same arguments as in the end of the proof in the previous lemma to obtain the general case, it suffices to assume $C = 0$ and $M = \mathbb{R}^n$. In that case, we can take $\psi_\pm(x) = \psi(x) \pm \|x - x_0\|^2$. $\square$

Proposition 7.3.7. *Suppose M is a compact manifold. Let $L : TM \to \mathbb{R}$ be a C^2 Tonelli Lagrangian. Consider the associated Lax-Oleinik semi-groups T_t^-, T_t^+. Suppose that $t > 0$, and that $\gamma : [0,t] \to M$ is a C^1 curve with $\gamma(t) = x$ (resp. $\gamma(0) = x$), and such that $T_t^- u(x) = u(\gamma(0)) + \int_0^t L(\gamma(s), \dot{\gamma}(s))\, ds$ (resp. $T_t^+ u(x) = u(\gamma(t)) - \int_0^t L(\gamma(s), \dot{\gamma}(s))\, ds$), then $\partial L/\partial v(\gamma(t), \dot{\gamma}(t)) \in$*

$D^+[T_t^- u](x)$, *and* $\partial L/\partial v(\gamma(0), \dot\gamma(0)) \in D^- u(\gamma(0))$ *(resp.* $\partial L/\partial v(\gamma(0), \dot\gamma(0)) \in D^-[T_t^- u](x)$ *and* $\partial L/\partial v(\gamma(t), \dot\gamma(t)) \in D^+[u](\gamma(t))$*).*

Proof. **A Faire!!!!!** □

7.4 Criteria for viscosity solutions

We fix in this section a continuous function $H : T^*M \to \mathbb{R}$.

Theorem 7.4.1. *Let* $u : M \to \mathbb{R}$ *be a continuous function.*

(i) u *is a viscosity subsolution of* $H(x, d_x u) = 0$ *if and only if for each* $x \in M$ *and each* $p \in D^+u(x)$ *we have* $H(x, p) \leq 0$.

(ii) u *is a viscosity supersolution of* $H(x, d_x u) = 0$ *if and only if for each* $x \in M$ *and each* $p \in D^-u(x)$ *we have* $H(x, p) \geq 0$.

Proof. Suppose that u is a viscosity subsolution. If $p \in D^+u(x)$, since u is continuous, it follows from 7.3.5 that there exists a C^1 function $\phi : M \to \mathbb{R}$, with $\phi \geq u$ on M, $u(x_0) = \phi(0)$ and $d_x\phi = p$. By the viscosity subsolution condition $H(x, p) = H(x, d_x\phi) \leq 0$, .

Suppose conversely that for each $x \in M$ and each $p \in D^+u(x_0)$ we have $H(x, p) \leq 0$. If $\phi : M \to \mathbb{R}$ is C^1 with $u \leq \phi$, then at each point x where $u(x) = \phi(x)$, we have $d_x\phi \in D^+u(x)$ and therefore $H(x, d_x\phi) \leq 0$. □

Since $D^\pm u(x)$ depends only on the values of u in a neighborhood of x, the following corollary is now obvious. It shows the local nature of the viscosity conditions.

Corollary 7.4.2. *Let* $u : M \to \mathbb{R}$ *be a continuous function.*

If u *is a viscosity subsolution (resp. supersolution, solution) of* $H(x, d_x u) = 0$ *on* M*, then any restriction* $u_{|U}$ *to an open subset* $U \subset M$ *is itself a viscosity subsolution (resp. supersolution, solution) of* $H(x, d_x u) = 0$ *on* U.

Conversely, if there exists an open cover $(U_i)_{i \in I}$ *of* M *such that every restriction* $u_{|U_i}$ *is a viscosity subsolution (resp. supersolution, solution) of* $H(x, d_x u) = 0$ *on* U_i*, then* u *itself is a viscosity subsolution (resp. supersolution, solution) of* $H(x, d_x u) = 0$ *on* M.

Here is another straightforward consequence of theorem 7.4.1.

Corollary 7.4.3. *Let $u : M \to \mathbb{R}$ be a locally Lipschitz function. If u is a viscosity subsolution (resp. supersolution solution) of $H(x, d_x u) = 0$, then $H(x, d_x u) \leq 0$ (resp. $H(x, d_x u) \geq 0, H(x, d_x u) = 0$) for almost every $x \in M$.*

In particular, a locally Lipschitz viscosity solution is always a very weak solution.

We end this section with one more characterization of viscosity solutions.

Proposition 7.4.4 (Criterion for viscosity solution)**.** *Suppose that $u : M \to \mathbb{R}$ is continuous. To check that u is a viscosity subsolution (resp. supersolution) of $H(x, d_x u) = 0$, it suffices to show that for each C^∞ function $\phi : M \to \mathbb{R}$ such that $u - \phi$ has a unique strict global maximum (resp. minimum), attained at x_0, we have $H(x_0, d_{x_0}\phi) \leq 0$ (resp. $H(x_0, d_{x_0}\phi) \geq 0$).*

Proof. We treat the subsolution case. We first show that if $\phi : M \to \mathbb{R}$ is a C^∞ function such that $u - \phi$ achieves a (not necessarily strict) maximum at x_0, then we have $H(x_0, d_{x_0}\phi) \leq 0$. In fact applying 7.3.6, we can find a C^∞ function $\phi_+ : M \to \mathbb{R}$ such that $\phi_+(x_0) = \phi(x_0), d_{x_0}\phi_+ = d_{x_0}\phi, \phi_+(x) > \phi(x)$, for $x \neq x_0$. The function $u - \phi_+$ has a unique strict global maximum achieved at x_0, therefore $H(x_0, d_{x_0}\phi_+) \leq 0$. Since $d_{x_0}\phi_+ = d_{x_0}\phi$, this finishes our claim.

Suppose now that $\psi : M \to \mathbb{R}$ is C^1 and that $u - \psi$ has a global maximum at x_0, we must show that $H(x_0, d_{x_0}\psi) \leq 0$. We fix a *relatively compact* open neighborhood W of x_0, by 7.3.6, applied to the continuous function ψ, there exists a C^1 function $\psi_+ : M \to \mathbb{R}$ such that $\psi_+(x_0) = \psi(x_0), d_{x_0}\psi_+ = d_{x_0}\psi, \psi_+(x) > \psi(x)$, for $x \neq x_0$, and even $\psi_+(x) > \psi(x) + 3$, for $x \notin W$. It is easy to see that $u - \psi_+$ has a strict global maximum at x_0, and that $u(x) - \psi_+(x) < u(x_0) - \psi_+(x_0) - 3$, for $x \notin W$. By smooth approximations, we can find a sequence of C^∞ functions $\phi_n : M \to \mathbb{R}$ such that ϕ_n converges to ψ_+ in the C^1 topology uniformly on compact subsets, and $\sup_{x \in M} |\phi_n(x) - \psi_+(x)| < 1$. This last condition together with $u(x) - \psi_+(x) < u(x_0) - \psi_+(x_0) - 3$, for $x \notin W$, gives $u(x) - \phi_n(x) < u(x_0) - \phi_n(x_0) - 1$, for $x \notin W$. This implies that the maximum of $u - \phi_n$ on the compact set $\bar{W}$ is a global maximum of $u - \phi_n$. Choose $y_n \in \bar{W}$ where $u - \phi_n$ attains its global maximum. Since ϕ_n is C^∞, from the beginning of the proof we must have $H(y_n, d_{y_n}\phi_n) \leq 0$. Extracting a subsequence, if necessary, we can assume that y_n converges to $y_\infty \in \bar{W}$. Since ϕ_n converges to ψ_+ uniformly on the compact set $\bar{W}$, necessarily $u - \psi_+$ achieves its maximum on $\bar{W}$ at y_∞. This implies that $y_\infty = x_0$, because the strict global maximum of $u - \tilde{\psi}$ is precisely attained at $x_0 \in W$. The convergence of ϕ_n to ψ_+ is in the C^1 topology, therefore $(y_n, d_{y_n}\phi_n) \to (x_0, d_{x_0}\psi_+)$, and hence $H(y_n, d_{y_n}\phi_n) \to H(x_0, d_{x_0}\psi_+)$, by continuity of H. But $H(y_n, d_{y_n}\phi_n) \leq 0$ and $d_{x_0}\psi = d_{x_0}\psi_+$, hence $H(x_0, d_{x_0}\psi) \leq 0$. $\square$

7.5 Coercive Hamiltonians

Definition 7.5.1 (Coercive)**.** A continuous function $H : T^*M \to \mathbb{R}$ is said to be coercive above every compact subset, if for each compact subset $K \subset M$ and each $c \in \mathbb{R}$ the set $\{(x, p) \in T^*M \mid x \in K, H(x, p) \leq c\}$ is compact.

Choosing any Riemannian metric on M, it is not difficult to see that H is coercive, if and only if for each compact subset $K \subset M$, we have $\lim_{\|p\| \to \infty} H(x, p) = +\infty$ the limit being uniform in $x \in K$.

Theorem 7.5.2. *Suppose that $H : T^*M \to \mathbb{R}$ is coercive above every compact subset, and $c \in \mathbb{R}$ then a viscosity subsolution of $H(x, d_x u) = c$ is necessarily locally Lipschitz, and therefore satisfies $H(x, d_x u) \leq c$ almost everywhere.*

Proof. Since this is a local result we can assume $M = \mathbb{R}^k$, and prove only that u is Lipschitz on a neighborhood of the origin 0. We will consider the usual distance d given by $d(x, y) = \|y - x\|$, where we have chosen the usual Euclidean norm on $\mathbb{R}^k$. We set

$$\ell_0 = \sup\{\|p\| \mid p \in \mathbb{R}^{k*}, \exists x \in \mathbb{R}^k, \|x\| \leq 3, H(x, p) \leq c\}.$$

Suppose $u : \mathbb{R}^k \to \mathbb{R}$ is a subsolution of $H(x, d_x u) = c$. Choose $\ell \geq \ell_0 + 1$ such that

$$2\ell > \sup\{|u(y) - u(x)| \mid x, y \in \mathbb{R}^k, \|x\| \leq 3, \|y\| \leq 3\}.$$

Fix x, with $\|x\| \leq 1$, and define $\phi : \mathbb{R}^k \to \mathbb{R}$ by $\phi(y) = \ell\|y - x\|$. Pick $y_0 \in \bar{B}(x, 2)$ where the function $y \mapsto u(y) - \phi(y)$ attains its maximum for $y \in \bar{B}(x,2)$. We first observe that y_0 is not on the boundary of $\bar{B}(x,2)$. In fact, if $\|y-x\| = 2$, we have $u(y)-\phi(y) = u(y)-2\ell < u(x) = u(x)-\phi(x)$. In particular y_0 is a local maximum of $u-\phi$. If y_0 is not equal to x, then $d_{y_0}\phi$ exists, with $d_{y_0}\phi(v) = \ell\langle y_0 - x, v\rangle/\|y_0 - x\|$, and we obtain $\|d_{y_0}\phi\| = \ell$. On the other hand, since $u(y) \geq u(y_0) - \phi(y_0) + \phi(y)$, for y in a neighborhood of y_0, we get $d_{y_0}\phi \in D^+u(y_0)$, and therefore have $H(y_0, d_{y_0}\phi) \leq c$. By the choice of ℓ_0, this gives $\|d_{y_0}\phi\| \leq \ell_0 < \ell_0 + 1 \leq \ell$. This contradiction shows that $y_0 = x$, hence $u(y) - \ell\|y-x\| \leq u(x)$, for every x of norm ≤ 1, and every $y \in \bar{B}(x,2)$. This implies that u has Lipschitz constant $\leq \ell$ on the unit ball of $\mathbb{R}^k$. □

7.6 Viscosity and weak KAM

In this section we finish showing that weak KAM solutions and viscosity solutions are the same.

Theorem 7.6.1. *Let $L : TM \to \mathbb{R}$ be a Tonelli Lagrangian on the compact manifold M. Denote by $H : T^*M \to \mathbb{R}$ its associated Hamiltonian. A continuous function $u : U \to \mathbb{R}$ is a viscosity subsolution of $H(x, d_x u) = c$ on the open subset U if and only if $u \prec L + c$.*

Proof. By proposition 7.2.7, it remains to prove that a viscosity subsolution of $H(x, d_x u) = c$ is dominated by $L + c$ on U. Since H is superlinear, we can apply theorem 7.5.2 to conclude that u is locally Lipschitz, and hence, by corollary 7.4.3, we obtain $H(x, d_x u) \leq c$ almost everywhere on U. From lemma 4.2.3, we infer $u \prec L + c$ on U. □

Theorem 7.6.2. *Let $L : TM \to \mathbb{R}$ be a Tonelli Lagrangian on the compact manifold M. Denote by $H : T^*M \to \mathbb{R}$ its associated Hamiltonian. A continuous function $u : M \to \mathbb{R}$ is a viscosity solution of $H(x, d_x u) = c$ if and only if it is Lipschitz and satisfies $u = T_t^- u + ct$, for each $t \geq 0$. (In particular, we must have c=c[0].)*

Proof. If u satisfies $u = T_t^- u + ct$, for each $t \geq 0$, then necessarily $c = c[0]$, and therefore, by proposition 7.2.7, the function u is a viscosity solution of $H(x, d_x u) = c = c[0]$.

Suppose now that u is a viscosity solution. From proposition 7.6.1, we know that $u \prec L + c$, and that u is Lipschitz. We then define $\tilde{u}(t,x) = T_t^- u(x)$. We must show that $\tilde{u}(t,x) = u(x) - ct$. Since we know that $\tilde{u}$ is locally Lipschitz on $]0,+\infty[\times M$, it suffices to show that $\partial_t \tilde{u}(t,x) = -c$ at each (t,x) where $\tilde{u}$ admits a derivative. In fact, since $\tilde{u}$ is locally Lipschitz, it is differentiable almost everywhere, therefore for almost every x the derivative $d_{(t,x)}\tilde{u}$ exists for almost every t. If we fix such an x, it follows that $\partial_t \tilde{u}(t,x) = -c$, for almost every t. But, since the $t \mapsto \tilde{u}(t,x)$ is locally Lipschitz, it is the integral of its derivative, therefore $\tilde{u}(t,x) - \tilde{u}(0,x) = -ct$. This is valid for almost every $x \in M$, and by continuity for every $x \in M$.

It remains to show that at a point (t,x) where $\tilde{u}$ is differentiable, we have $\partial_t \tilde{u}(t,x) = -c$. From proposition 7.2.8, we know that $\tilde{u}$ is a viscosity solution of $\partial_t \tilde{u} + H(x, \partial_x \tilde{u}) = 0$. Hence we have to show that $H(x, \partial_x \tilde{u}(t,x)) = c$. In fact, we already have that $H(x, \partial_x \tilde{u}(t,x)) \leq c$, because $\tilde{u}(t,\cdot) = T_t^- u$ is dominated by $L + c$, like u. It remains to prove that $H(x, \partial_x \tilde{u}(t,x)) \geq c$. To do this, we identify the derivative of $\partial_x \tilde{u}(t,x)$. We choose $\gamma : [0,t] \to M$ with $\gamma(t) = x$ and $T_t^- u(x) = u(\gamma(0)) + \int_0^t L(\gamma(s), \dot{\gamma}(s))\, ds$. The curve γ is a minimizer of the action. In particular, the curve $s \mapsto (\gamma(s), \dot{\gamma}(s))$ is a solution of the Euler-Lagrange equation, it follows that the energy $H(\gamma(s), \frac{\partial L}{\partial v}(\gamma(s)), \dot{\gamma}(s)))$ is constant on $[0,t]$. By proposition 7.3.7, we have $\partial L/\partial v(\gamma(t)), \dot{\gamma}(t)) \in D^+(T_t^+ u)(x)$, therefore $\partial_x \tilde{u}(t,x) = \frac{\partial L}{\partial v}(\gamma(t)), \dot{\gamma}(t))$. Using the fact that $H(\gamma(s), \frac{\partial L}{\partial v}(\gamma(s)), \dot{\gamma}(s)))$ is constant, we are reduced to see that $H(\gamma(0), \frac{\partial L}{\partial v}(\gamma(0)), \dot{\gamma}(0))) \geq c$. But the same proposition 7.3.7 yields also $\partial L/\partial v(\gamma(0), \dot{\gamma}(0)) \in D^- u(\gamma(0))$. We can therefore conclude using theorem 7.4.1, since u is a viscosity solution of $H(x, d_x u) = c$. □

Chapter 8

More on Viscosity Solutions

We further develop the theory of viscosity solutions. Although many things are standard, whatever is not comes from joint work with Antonio Siconolfi, see [FS04] and [FS05].

8.1 Stability

Theorem 8.1.1 (Stability). *Suppose that the sequence of continuous functions $H_n : T^*M \to \mathbb{R}$ converges uniformly on compact subsets to $H : T^*M \to \mathbb{R}$. Suppose also that $u_n : M \to \mathbb{R}$ is a sequence of continuous functions converging uniformly on compact subsets to $u : M \to \mathbb{R}$. If, for each n, the function u_n is a viscosity subsolution (resp. supersolution, solution) of $H_n(x, d_x u_n) = 0$, then u is a viscosity subsolution (resp. supersolution, solution) of $H(x, d_x u) = 0$.*

Proof. We show the subsolution case. We use the criterion 7.4.4. Suppose that $\phi : M \to \mathbb{R}$ is a C^∞ function such that $u - \phi$ has a unique strict global maximum, achieved at x_0, we have to show $H(x_0, d_{x_0}\phi) \leq 0$. We pick a relatively compact open neighborhood W of x_0. For each n, choose $y_n \in \bar{W}$ where $u_n - \phi$ attains its maximum on the compact subset $\bar{W}$. Extracting a subsequence, if necessary, we can assume that y_n converges to $y_\infty \in \bar{W}$. Since u_n converges to u uniformly on the compact set $\bar{W}$, necessarily

$u - \phi$ achieves its maximum on $\bar{W}$ at y_∞. But $u - \phi$ has a strict global maximum at $x_0 \in W$ therefore $y_\infty = x_0$. By continuity of the derivative of ϕ, we obtain $(y_n, d_{y_n}\phi) \to (x_0, d_{x_0}\phi)$. Since W is an open neighborhood of x_0, dropping the first terms if necessary, we can assume $y_n \in W$, this implies that y_n is a local maximum of $u_n - \phi$, therefore $d_{y_n}\phi \in D^+u_n(y)$. Since u_n is a viscosity subsolution of $H_n(x, d_x u_n) = 0$, we get $H_n(y_n, d_{y_n}\phi) \leq 0$. The uniform convergence of H_n on compact subsets now implies $H(x_0, d_{x_0}\phi) = \lim_{n\to\infty} H_n(y_n, d_{y_n}\phi) \leq 0$. □

8.2 Construction of viscosity solutions

Proposition 8.2.1. *Let $H : T^*M \to \mathbb{R}$ be a continuous function. Suppose $(u_i)_{i\in I}$ is a family of continuous functions $u_i : M \to \mathbb{R}$ such that each u_i is a subsolution (resp. supersolution) of $H(x, d_x u) = 0$. If $\sup_{i\in I} u_i$ (resp. $\inf_{i\in u} u_i$) is finite and continuous everywhere, then it is also a subsolution (resp. supersolution) of $H(x, d_x u) = 0$.*

Proof. Set $u = \sup_{i\in I} u_i$. Suppose $\phi : M \to \mathbb{R}$ is C^1, with $\phi(x_0) = u(x_0)$ and $\phi(x) > u(x)$, for every $x \in M \setminus \{x_0\}$. We have to show $H(x_0, d_{x_0}\phi) \leq 0$. Fix some distance d on M. By continuity of the derivative of ϕ, it suffices to show that for each $\epsilon > 0$ small enough there exists $x \in \mathring{B}(x_0, \epsilon)$, with $H(x, d_x\phi) \leq 0$.

For $\epsilon > 0$ small enough, the closed ball $\bar{B}(x_0, \epsilon)$ is compact. Fix such an $\epsilon > 0$. There is a $\delta > 0$ such that $\phi(y) - \delta \geq u(y) = \sup_{i\in I} u_i(y)$, for each $y \in \partial B(x_0, \epsilon)$.

Since $\phi(x_0) = u(x_0)$, we can find $i_\epsilon \in I$ such that $\phi(x_0) - \delta < u_{i_\epsilon}(x_0)$. It follows that the maximum of the continuous function $u_{i_\epsilon} - \phi$ on the compact set $\bar{B}(x_0, \epsilon)$ is not attained on the boundary, therefore $u_{i_\epsilon} - \phi$ has a local maximum at some $x_\epsilon \in \mathring{B}(x_0, \epsilon)$. Since the function u_{i_ϵ} is a viscosity subsolution of $H(x, d_x u) = 0$, we have $H(x_\epsilon, d_{x_\epsilon}\phi) \leq 0$. □

Theorem 8.2.2 (Perron Method)**.** *Suppose the Hamiltonian $H : TM \to \mathbb{R}$ is coercive above every compact subset. If M is connected and there exists a viscosity subsolution $u : M \to \mathbb{R}$ of $H(x, d_x u) = 0$, then for every $x_0 \in M$, the function $S_{x_0} : M \to \mathbb{R}$ defined by $S_{x_0}(x) = \sup_v v(x)$, where the supremum is taken over* *all viscosity subsolutions v satisfying $v(x_0) = 0$, has indeed finite values and is a viscosity subsolution of $H(x, d_x u) = 0$ on M.*

Moreover, it is a viscosity solution of $H(x, d_x u) = 0$ on $M \setminus \{x_0\}$.

Proof. Call $\mathcal{SS}_{x_0}$ the family of viscosity subsolutions $v : M \to \mathbb{R}$ of $H(x, d_x v) = 0$ satisfying $v(x_0) = 0$.

Since H is coercive above every compact subset of M, by theorem 7.5.2, we know that each element of this family is locally Lipschitz. Moreover, since for each compact set K, the set $\{(x,p) \mid x \in K, H(x,p) \leq 0\}$ is compact, it follows that the family of restrictions $v_{|K}, v \in \mathcal{SS}_{x_0}$ is equi-Lipschitzian. We now show, that S_{x_0} is finite everywhere. Since M is connected, given $x \in M$, there exists a compact connected set K_{x,x_0} containing both x and x_0. By the equicontinuity of the family of restrictions $\{v_{|K_{x,x_0}} \mid v \in \mathcal{SS}_{x_0}\}$, we can find $\delta > 0$, such that for each $y, z \in K_{x,x_0}$ with $d(y,z) \leq \delta$, we have $|v(y) - v(z)| \leq 1$, for each $v \in \mathcal{SS}_{x_0}$.

By the set K_{x,x_0} is connected , we can find a sequence $x_0, x_1, \cdots, x_n = x$ in K_{x,x_0} with $d(x_i, x_{i+1}) \leq \delta$. It follows that $|v(x)| = |v(x) - v(x_0)| \leq \sum_{i=0}^{n-1} |v(x_{i+1}) - v(x_i)| \leq n$, for each $v \in \mathcal{SS}_{x_0}$. Therefore $\sup_{v \in \mathcal{SS}_{x_0}} v(x)$ is finite everywhere. Moreover, as a finite-valued supremum of a family of locally equicontinuous functions, it is continuous.

By the previous proposition 8.2.1, the function S_{x_0} is a viscosity subsolution on M itself. It remains to show that it is a viscosity solution of $H(x, d_x u)$ on $M \setminus \{x_0\}$.

Suppose $\psi : M \to \mathbb{R}$ is C^1 with $\psi(x_1) = S_{x_0}(x_1)$, where $x_1 \neq x_0$, and $\psi(x) < S_{x_0}(x)$ for every $x \neq x_1$. We want to show that necessarily $H(x_1, d_{x_1}\psi) \geq 0$. If this were false, by continuity of the derivative of ψ, endowing M with a distance defining its topology, we could find $\epsilon > 0$ such that $H(y, d_y \psi) < 0$, for each $y \in \bar{B}(x_1, \epsilon)$. Taking $\epsilon > 0$ small enough, we assume that $\bar{B}(x_1, \epsilon)$ is compact and $x_0 \notin \bar{B}(x_1, \epsilon)$. Since $\psi < S_{x_0}$ on the boundary $\partial B(x_1, \epsilon)$ of $\bar{B}(x_1, \epsilon)$, we can pick $\delta > 0$, such that $\psi(y) + \delta \leq S_{x_0}(y)$, for every $y \in \partial B(x_1, \epsilon)$. We define $\tilde{S}_{x_0}$ on $\bar{B}(x_1, \epsilon)$ by $\tilde{S}_{x_0}(x) = \max(\psi(x) + \delta/2, S_{x_0}(x))$. The function $\tilde{S}_{x_0}$ is a viscosity subsolution of $H(x, d, u)$ on $\mathring{B}(x_1, \epsilon)$ as the maximum of the two viscosity subsolutions $\psi + \delta/2$ and S_{x_0}. Moreover, this function $\tilde{S}_{x_0}$ coincides with S_{x_0} outside $K = \{x \in \mathring{B}(x_1, \epsilon) \mid \psi(x)+\delta/2 \geq S_{x_0}(x))\}$ which is a compact subset of $\mathring{B}(x_1, \epsilon)$, therefore we can extend it to M itself by $\tilde{S}_{x_0} = S_{x_0}$ on $M \setminus K$. It is a viscosity subsolution of $H(x, d_x u)$ on M itself, since its restrictions to both open subsets $M \setminus K$ and $\mathring{B}(x_1, \epsilon)$ are viscosity subsolutions and $M = \mathring{B}(x_1, \epsilon) \cup (M \setminus K)$.

But $\tilde{S}_{x_0}(x_0) = S_{x_0}(x_0) = 0$ because $x_0 \notin \bar{B}(x_1, \epsilon)$. Moreover $\tilde{S}_{x_0}(x_1) = \max(\psi(x_1)+\delta/2, S_{x_0}(x_1)) = \max(S_{x_0}(x_1)+\delta/2, S_{x_0}(x_1)) = S_{x_0}(x_1)+\delta/2 > S_{x_0}(x_1)$. This contradicts the definition of S_{x_0}. $\square$

The next argument is inspired by the construction of Busemann functions in Riemannain Geometry, see [BGS85].

Corollary 8.2.3. *Suppose that $H : T^*M \to \mathbb{R}$ is a continuous Hamiltonian coercive above every compact subset of the connected* non-compact *manifold M. If there exists a viscosity subsolution of $H(x, d_x u) = 0$ on M, then there exists a viscosity* solution *on M.*

Proof. Fix $\hat{x} \in M$, and pick a sequence $x_n \to \infty$ (this means such that each compact subset of M contains only a finite number of points in the sequence).

By arguments analogous to the ones used in the previous proof, the sequence S_{x_n} is locally equicontinuous and moreover, for each $x \in M$, the sequence $S_{x_n}(x) - S_{x_n}(\hat{x})$ is bounded. Therefore, by Ascoli's theorem, extracting a subsequence if necessary, we can assume that $S_{x_n} - S_{x_n}(\hat{x})$ converges uniformly to a continuous function $u : M \to \mathbb{R}$. It now suffices to show that the restriction of u to an arbitrary open relatively compact subset V of M is a viscosity solution of $H(x, d_x u) = 0$ on V. Since $\{n \mid x_n \in \bar{V}\}$ is finite, for n large enough, the restriction of $S_{x_n} - S_{x_n}(\hat{x})$ to V is a viscosity solution; therefore by the stability theorem 8.1.1, the restriction of the limit u to V is also a viscosity solution. $\square$

The situation is different for compact manifolds as can be seen from the following theorem:

Theorem 8.2.4. *Suppose $H : T^*M \to \mathbb{R}$ is a coercive Hamiltonian on the compact manifold M. If there exists a viscosity subsolution of $H(x, d_x u) = c_1$ and a viscosity supersolution of $H(x, d_x u) = c_2$, then necessarily $c_2 \leq c_1$.*

In particular, there exists a most one c for which the Hamilton-Jacobi equation $H(x, d_x u) = c$ has a global viscosity solution $u : M \to \mathbb{R}$. This only possible value is the smallest c for which $H(x, d_x u) = c$ admits a global viscosity subsolution $u : M \to \mathbb{R}$.

In order to prove this theorem, we will need a lemma (the last part of the lemma will be used later).

Lemma 8.2.5. *Suppose M compact, and $u : M \to \mathbb{R}$ is a Lipschitz viscosity subsolution (resp. supersolution) of $H(x, d_x u) = c$. For every $\epsilon > 0$, there exists a locally semi-convex (resp. semi-concave) function $u_\epsilon : M \to \mathbb{R}$ such that $\|u_\epsilon - u\|_\infty < \epsilon$, and u_ϵ is a viscosity subsolution (resp. supersolution) of $H(x, d_x u) = c + \epsilon$ (resp. $H(x, d_x u) = c - \epsilon$).*

Moreover, if $K \subset U$ are respectively a compact and an open subsets of M such that the restriction $u_{|U}$ is a viscosity subsolution (resp. supersolution) of $H(x, d_x u) = c'$ on U, for some $c' < c$ (resp. $c' > c$), we can also impose that restriction $u_{\epsilon|U}$ is a viscosity subsolution (resp. supersolution) of $H(x, d_x u) = c' + \epsilon$ (resp. $H(x, d_x u) = c' - \epsilon$) on a neighborhood of K.

Proof. Suppose $\epsilon > 0$ is given. Fix a C^∞ Riemannian metric g on M. We call T_t^- and T_t^+ the two Lax-Oleinik semi-groups associated to the Lagrangian $L(x, v) = \frac{1}{2} g_x(v, v) = \frac{1}{2}\|v\|_x^2$. If $\epsilon > 0$ and $u : M \to \mathbb{R}$ is a Lipschitz viscosity subsolution of $H(x, d_x u) = c$, we consider the locally semi-convex function $T_t^+ u$. By the analogous of part (7) of corollary **??**, the map $t \mapsto T_t^+ u$ is continuous as a map with values in $\mathcal{C}^0(M, \mathbb{R})$ endowed with the sup norm, therefore $\|T_t^+ u - u\|_\infty < \epsilon$, for each $t > 0$ small enough. Assume that $t > 0$. Since $T_t^+ u$ is locally semi-convex, at each point x, the set $D^- T_t^+ u(x)$ is not empty, therefore the points x where $D^+ T_t^+ u(x) \neq \emptyset$ are the points where to $T_t^+ u$ is differentiable. Hence to check that it is a subsolution of $H(x, du) = c + \epsilon$, it suffices to show that if $d_x T_t^+ u$ exists then $H(x, d_x T_t^+ u) \leq c + \epsilon$.

Suppose that $d_x T_t^+ u$ exists. Choose a geodesic $\gamma : [0, t] \to M$ with $\gamma(0) = x$, and $T_t^+ u(x) = \int_0^t \frac{1}{2}\|\dot\gamma(t)\|_{\gamma(t)}^2 dt - u(\gamma(t))$. We have $d_x T_t^+ u(x) = \dot\gamma(0)^\sharp$, where for $v \in T_y M$, the linear form $v^\sharp \in T_y^* M$ is given by $v^\sharp(w) = g_y(v, w)$, for every $w \in T_y M$. Moreover, we also have $\dot\gamma(t)^\sharp \in D^+ u(\gamma(t))$. Therefore $d_x T_t^+ u \in \varphi_{-t}^{g*}(\mathrm{Graph}(D^+ u))$, where $\varphi_t^{g*} : T^*M \to T^*M$ is the Hamiltonian form of the geodesic flow $\varphi_t^g : TM \to TM$ of g, and $\mathrm{Graph}(D^+u) = \{(x,p) \mid x \in M, p \in D^+u(x)\}$.

Since u is a subsolution of $H(x, d_xu) = c$, we have $H(x,p) \leq c$, for each $p \in D^+u(x)$. The compactness of $P = \{p \mid p \in T_x^*M, \|p\|_x \leq \mathrm{Lip}(u)\}$, which contains $\mathrm{Graph}(D^+u)$, and the continuity of both the flow φ_t^{g*} and the Hamiltonian H imply that there exists $t_0 > 0$ such that $\varphi_{-t}^{g*} D^+u \subset H^{-1}(]-\infty, c+\epsilon])$, for every $t \in [0, t_0]$.

To prove the last part, we choose an open subset V with $K \subset V \subset \bar{V} \subset U$. We observe that $\mathrm{Graph}(D^+u)$ is contained in the compact set $\tilde{P} = \{(x,p) \in P \mid x \notin U, H(x,p) \leq c\} \cup \{(x,p) \in P \mid x \in \bar{U}, H(x,p) \leq c'\}$. Again by compactness, we can find a t_0' such that the intersection of $\varphi_{-t}^{g*} D^+u$ with $T^*V = \{(x,p) \mid x \in V\}$ is contained in $H^{-1}(]-\infty, c'+\epsilon])$, for every $t \in [0, t_0']$. □

Proof of theorem 8.2.4. Suppose $c_1 < c_2$, and choose $\epsilon > 0$, with $c_1 + \epsilon < c_2$, by the previous lemma 8.2.5, we can find a locally semi-convex function $u_1 : M \to \mathbb{R}$ which is a viscosity subsolution of $H(x, d_xu) = c_1 + \epsilon$.

We now show that for every $x \in M$, there exists $p \in D^-u_1(x)$ with $H(x,p) \leq c_1 + \epsilon$. Since a locally semi-convex function is Lipschitz, by Rademacher's theorem, if $x \in M$, we can find a sequence of points $x_n \in M$ converging to x such that the derivative $d_{x_n}u_1$ exists. We have $H(x_n, d_{x_n}u_1) \leq c_1 + \epsilon$. Since u_1 is Lipschitz, the points $(x_n, d_{x_n}u_1)$ are contained in a compact subset of T^*M. Extracting a sequence if necessary, we can assume that $(x_n, d_{x_n}u_1) \to (x,p)$, of course $H(x,p) \leq c_1+\epsilon$, and $p \in D^-u_1(x)$, because u_1 is locally semi-convex.

We fix $u_2 : M \to \mathbb{R}$ a viscosity supersolution of $H(x, d_xu) = c_2$. Call x_0 a point where the continuous function $u_2 - u_1$ on the compact manifold M achieves its minimum. We have $u_2 \geq u_2(x_0) - u_1(x_0) + u_1$ with equality at x_0, therefore $D^-u_1(x_0) \subset D^-u_2(x_0)$, this is impossible since $D^-u_2(x_0) \subset H^{-1}([c_2, +\infty[), D_-u_1(x_0) \cap H^{-1}(]-\infty, c_1+\epsilon]) \neq \phi$ and $c_1 + \epsilon < c_2$. □

Exercise 8.2.6. *Prove a non-compact version of lemma 8.2.5:*

Suppose $H : T^*M \to \mathbb{R}$ is a continuous Hamiltonian on the (not necessarily compact) manifold M. If $u : M \to \mathbb{R}$ is a locally Lipschitz viscosity subsolution (resp. supersolution) of $H(x, d_xu) =$

c, show that for every $\epsilon > 0$, and every relatively compact open subset U, there exists a locally semi-convex (resp. semi-concave) function $u_\epsilon : U \to \mathbb{R}$ such that $\|u_\epsilon - u_{|U}\|_\infty < \epsilon$, and u_ϵ is a viscosity subsolution (resp. supersolution) of $H(x, d_x u) = c + \epsilon$ (resp. $H(x, d_x u) = c - \epsilon$) on U.

Definition 8.2.7 (strict subsolution)**.** We say that a viscosity subsolution $u : M \to \mathbb{R}$ of $H(x, d_x u) = c$ is strict at $x_0 \in M$ if there exists an open neighborhood V_{x_0} of x_0, and $c_{x_0} < c$ such that $u|V_{x_0}$ is a viscosity subsolution of $H(x, d_x u) = c_{x_0}$ on V_{x_0}.

Here is a way to construct viscosity subsolutions which are strict at some point.

Proposition 8.2.8. *Suppose that $u : M \to R$ is a viscosity subsolution of $H(y, d_y u) = c$ on M, that is also a viscosity solution on $M \setminus \{x\}$. If u is not a viscosity solution of $H(y, d_y u) = c$ on M itself then there exists a viscosity subsolution of $H(y, d_y u) = c$ on M which is strict at x.*

Proof. If u is not a viscosity solution, since it is a subsolution on M, it is the supersolution condition that is violated. Moreover, since u is a supersolution on $M \setminus \{x\}$, the only possibility is that there exists $\psi : M \to R$ of class C^1 such that $\psi(x) = u(x), \psi(y) < u(y)$, for $y \neq x$, and $H(x, d_x\psi) < c$. By continuity of the derivative of ψ, we can find a compact ball $\bar{B}(x, r)$, with $r > 0$, and a $c_x < c$ such that $H(y, d_y\psi) < c_x$, for every $y \in \bar{B}(0, r)$. In particular, the C^1 function ψ is a subsolution of $H(z, d_z v) = c_x$ on $\mathring{B}(x, r)$, and therefore also of $H(z, d_z v) = c$ on the same set since $c_x < c$.

We choose $\delta > 0$ such that for every $y \in \partial B(x, r)$ we have $u(y) > \psi(y) + \delta$. This is possible since $\partial B(x, r)$ is a compact subset of $M \setminus \{x\}$ where we have the strict inequality $\psi < u$.

If we define $\tilde{u} : M \to R$ by $\tilde{u}(y) = u(y)$ if $y \notin \bar{B}(x, r)$ and $\tilde{u}(y) = \max(u(y), \psi(y) + \delta)$, we obtain the desired viscosity subsolution of $H(y, d_y u) \leq c$ which is strict at x. In fact, by the choice of $\delta > 0$, the subset $K = \{y \in \bar{B}(x, r) \mid \psi(y) + \delta \leq u(y)\}$ is compact and contained in the open ball $\mathring{B}(x, r)$. Therefore M is covered by the two open subsets $M \setminus K$ and $\mathring{B}(x, r)$. On the first open subset $\tilde{u}$ is equal to u, it is therefore a subsolution of $H(y, d_y u) = c$ on that subset. On the second open subset $\mathring{B}(x, r)$, the function $\tilde{u}$ is the maximum of u and $\psi + \delta$ which are both subsolutions of $H(y, d_y u) = c$ on $\mathring{B}(x, r)$, by proposition 8.2.1, it is therefore a subsolution of $H(y, d_y u) = c$ on that second open subset. Since $u(x) = \psi(x)$; we have $\tilde{u}(x) = \psi(x) + \delta > u(x)$, therefore by continuity $\tilde{u} = \psi + \delta$ on a neighborhood $N \subset \mathring{B}(x, r)$ of x. On that neighborhood $H(y, d_y\psi) < c_x$, hence $\tilde{u}$ is strict at x. □

8.3 Quasi-convexity and viscosity subsolutions

In this section we will be mainly interested in Hamiltonians $H : T^*M \to \mathbb{R}$ quasi-convex in the fibers, i.e. for each $x \in M$, the function $p \mapsto H(x, p)$ is quasi-convex on the vector space $T^*_x M$, see definition 1.5.1

Our first goal in this section is to prove the following theorem:

Theorem 8.3.1. *Suppose $H : T^*M \to \mathbb{R}$ is quasi-convex in the fibers. If $u : M \to \mathbb{R}$ is locally Lipschitz and $H(x, d_x u) \leq c$ almost everywhere, for some fixed $c \in \mathbb{R}$, then u is a viscosity subsolution of $H(x, d_x u) = c$.*

Before giving the proof of the theorem we need some preliminary material.

Let us first recall from definition 4.2.4 that the Hamiltonian constant $\mathbb{H}_U(u)$ of a locally Lipschitz function $u : U \to R$, where U is an open subset of M is the essential supremum on U of $H(x, d_x u)$.

We will use some classical facts about convolution. Let $(\rho_\delta)_{\delta>0}$ be a family of functions $\rho_\delta : \mathbb{R}^k \to [0, \infty[$ of class C^∞, with $\rho_\delta(x) = 0$, if $\|x\| \geq \delta$, and $\int_{\mathbb{R}^k} \rho_\delta(x)\, dx = 1$. Suppose that V, U are open subsets of $\mathbb{R}^k$, with $\bar{V}$ compact and contained in U. Call $2\delta_0$ the Euclidean distance of the compact set $\bar{V}$ to the boundary of U, we have $\delta_0 > 0$, therefore the closed δ_0-neighborhood

$$\bar{N}_{\delta_0}(\bar{V}) = \{y \in \mathbb{R}^k \mid \exists x \in \bar{V}, \|y - x\| \leq \delta_0\}$$

of $\bar{V}$ is compact and contained in U.

If $u : U \to \mathbb{R}$ is a continuous function, then for $\delta < \delta_0$, the convolution

$$u_\delta(x) = \rho_\delta * u(x) = \int_{\mathbb{R}^k} \rho_\delta(y) u(x-y)\, dy.$$

makes sense and is of class C^∞ on a neighborhood of $\bar{V}$. Moreover, the family u_δ converges uniformly on $\bar{V}$ to u, as $t \to 0$.

Lemma 8.3.2. *Under the hypothesis above, suppose that $u : U \to \mathbb{R}$ is a locally Lipschitz function. Given any Hamiltonian $H : T^*U \to \mathbb{R}$ quasi-convex in the fibers and any $\epsilon > 0$, for every $\delta > 0$ small enough, we have* $\sup_{x\in V}|u_\delta(x) - u(x)| \le \epsilon$ *and* $\mathbb{H}_V(u_\delta) \le \mathbb{H}_U(u) + \epsilon$.

Proof. Because u is locally Lipschitz the derivative $d_z u$ exists for almost every $z \in U$. We first show that, for $\delta < \delta_0$, we must have

$$\forall x \in V, d_x u_\delta = \int_{\mathbb{R}^k} \rho_\delta(y) d_{x-y} u\, dy. \tag{*}$$

In fact, since u_δ is C^∞, it suffices to check that

$$\lim_{t\to 0} \frac{u_\delta(x+th) - u_\delta(x)}{t} = \int_{\mathbb{R}^k} \rho_\delta(y) d_{x-y} u(h)\, dy, \tag{**}$$

for $x \in V, \delta < \delta_0$, and $h \in \mathbb{R}^k$. Writing

$$\frac{u_\delta(x+th) - u_\delta(x)}{t} = \int_{\mathbb{R}^k} \rho_\delta(y) \frac{u(x+th-y) - u(x-y)}{t}\, dy,$$

We see that we can obtain (**) from Lebesgue's dominated convergence theorem, since ρ_δ has a compact support contained in $\{y \in \mathbb{R}^k \mid \|y\| < \delta\}$, and for $y \in \mathbb{R}^k, t \in \mathbb{R}$ such that $\|y\| < \delta, \|th\| < \delta_0 - \delta$, the two points $x + th - y, x - y$ are contained in the compact set $\bar{N}_{\delta_0}(\bar{V})$ on which u is Lipschitz. Equation (*) yields

$$H(x, d_x u_\delta) = H(x, \int_{\mathbb{R}^k} \rho_\delta(y) d_{x-y} u\, dy). \tag{***}$$

Since $\bar{N}_{\delta_0}(\bar{V})$ is compact and contained in U, and u is locally Lipschitz, we can find $K < \infty$ such that $\|d_z u\| \le K$, for each $z \in \bar{N}_{\delta_0}(\bar{V})$ for which $d_z u$ exists. Since H is continuous, by a compactness argument, we can find $\delta_\epsilon \in]0, \delta_0[$, such that for $z, z' \in \bar{N}_{\delta_0}(\bar{V})$, with $\|z - z'\| \leq \delta_\epsilon$, and $\|p\| \leq K$, we have $|H(z', p) - H(z, p)| \leq \epsilon$. If $\delta \leq \delta_\epsilon$, since $\rho_\delta(y) = 0$, if $\|y\| \geq \delta$, we deduce that for all x in V and almost every y with $\|y\| \leq \delta$, we have

$$H(x, d_{x-y}u) \leq H(x - y, d_{x-y}u) + \epsilon \leq \mathbb{H}_U(u) + \epsilon.$$

Since H is quasi-convex in the fibers, and $\rho_\delta \, dy$ is a probability measure whose support is contained in $\{y \mid \|y\| \geq \delta\}$, we can now apply proposition 1.5.6 to obtain

$$\forall \delta \leq \delta_\epsilon, H(x, \int_{\mathbb{R}^k} \rho_\delta(y) d_{x-y}u \, dy) \mathbb{H}_U(u) + \epsilon.$$

It from inequality (***) above that $H(x, d_x u_\delta) \leq \mathbb{H}_U(u) + \epsilon$, for $\delta \leq \delta_\epsilon$ and $x \in V$. This gives $\mathbb{H}_V(u_\delta) \leq \mathbb{H}_U(u) + \epsilon$, for $\delta \leq \delta_\epsilon$. The inequality $\sup_{x \in V} |u_\delta(x) - u(x)| < \epsilon$ also holds for every δ small enough, since u_δ converges uniformly on $\bar{V}$ to u, as $t \to 0$. □

Proof of theorem 8.3.1. We have to prove that for each $x_0 \in M$, there exists an open neighborhood V of x_0 such that $u_{|V}$ is a viscosity subsolution of $H(x, d_x u)$ on V. In fact, if we take V any open neighborhood such that $\bar{V}$ is contained in a domain of a coordinate chart, we can apply lemma 8.3.2 to obtain a sequence $u_n : V \to \mathbb{R}, n \geq 1$, of C^∞ functions such that u_n converges uniformly to $u_{|V}$ on V and $H(x, d_x u_n) \leq c + 1/n$. If we define $H_n(x, p) = H(x, p) - c - 1/n$, we see that u_n is a smooth classical, and hence viscosity, subsolution of $H_n(x, d_x w) = 0$ on V. Since H_n converges uniformly to $H - c$, the stability theorem 8.1.1 implies that $u_{|V}$ is a viscosity subsolution of $H(x, d_x u) - c = 0$ on V. □

Corollary 8.3.3. *Suppose that the Hamiltonian $H : T^*M \to \mathbb{R}$ is continuous and quasi-convex in the fibers. For every $c \in \mathbb{R}$, the set of Lipschitz functions $u : M \to \mathbb{R}$ which are viscosity subsolutions of $H(x, d_x u) = c$ is convex.*

Proof. If $u_1, \dots, u_n$ are such viscosity subsolutions. By 7.4.3, we know that at every x where $d_x u_j$ exists we must have $H(x, d_x u_j) \leq c$. If we call A the set of points x where $d_x u_j$ exists for each $j = 1, \dots, n$, then A has full Lebesgue measure in M. If $a_1, \dots, a_n \geq 0$, and $a_1 + \cdots + a_n = 1$, then $u = a_1 u_1 + \cdots + a_n u_n$ is differentiable at each point of $x \in A$ with $d_x u = a_1 d_x u_1 + \cdots + a_n d_x u_n$. Therefore by the quasi-convexity of $H(x, p)$ in the variable p, for every $x \in A$, we obtain $H(x, d_x u) = H(x, a_1 d_x u_1 + \cdots + a_n d_x u_n) \leq \max_{i=1}^n H(x, d_x u_i) \leq c$. Since A is of full measure, by theorem 8.3.1, we conclude that u is also a viscosity subsolution of $H(x, d_x u) = c$. $\square$

The next corollary shows that the viscosity subsolutions are the same as the very weak subsolutions, at least in the geometric cases we have in mind. This corollary is clearly a consequence of theorems 7.5.2 and 8.3.1.

Corollary 8.3.4. *Suppose that the Hamiltonian $H : T^*M \to \mathbb{R}$ is continuous, coercive, and quasi-convex in the fibers. A continuous function $u : M \to \mathbb{R}$ is a viscosity subsolution of $H(x, d_x u) = c$, for some $c \in \mathbb{R}$ if and only if u is locally Lipschitz and $H(x, d_x u) \leq c$, for almost every $x \in M$.*

We now give a global version of lemma 8.3.2.

Theorem 8.3.5. *Suppose that $H : T^*M \to \mathbb{R}$ is a Hamiltonian, which is quasi-convex in the fibers. Let $u : M \to \mathbb{R}$ be a locally Lipschitz viscosity subsolution of $H(x, d_x u) = c$ on M. For every couple of continuous functions $\delta, \epsilon : M \to]0, +\infty[$, we can find a C^∞ function $v : M \to \mathbb{R}$ such that $|u(x) - v(x)| \leq \delta(x)$ and $H(x, d_x v) \leq c + \epsilon(x)$, for each $x \in M$.*

Proof. We endow M with an auxiliary Riemannian metric. We pick up a locally finite countable open cover $(V_i)_{i \in \mathbb{N}}$ of M such that each closure $\bar{V}_i$ is compact and contained in the domain U_i of a chart which has a compact closure $\bar{U}_i$ in M. The local finiteness of the cover $(V_i)_{i \in \mathbb{N}}$ and the compactness of $\bar{V}_i$ imply that the set $J(i) = \{j \in \mathbb{N} \mid V_i \cap V_j \neq \emptyset\}$ is finite. Therefore, denoting by $\#A$ for the number of elements in a set A, we obtain

$$
\begin{aligned}
j(i) &= \# J(i) = \#\{j \in \mathbb{N} \mid V_i \cap V_j \neq \emptyset\} < +\infty, \\
\tilde{j}(i) &= \max_{\ell \in J(i)} j(\ell) < +\infty.
\end{aligned}
$$

We define $R_i = \sup_{x \in \bar{U}_i} \|d_x u\|_x < +\infty$, where the sup is in fact taken over the subset of full measure of $x \in U_i$ where the locally

Lipschitz function u has a derivative. It is finite because $\bar{U}_i$ is compact. Since $J(i)$ is finite, the following quantity $\tilde{R}_i$ is also finite

$$\tilde{R}_i = \max_{\ell \in J(i)} R_\ell < +\infty.$$

We now choose $(\theta_i)_{i\in\mathbb{N}}$ a C^∞ partition of unity subordinated to the open cover $(V_i)_{i\in\mathbb{N}}$. We also define

$$K_i = \sup_{x\in M} \|d_x\theta_i\|_x < +\infty,$$

which is finite since θ_i is C^∞ with support in V_i which is relatively compact.

Again by compactness, continuity, and finiteness routine arguments the following numbers are > 0

$$\delta_i = \inf_{x\in\bar{V}_i} \delta(x) > 0, \tilde{\delta}_i = \min_{\ell\in J(i)} \delta_\ell > 0$$
$$\epsilon_i = \inf_{x\in\bar{V}_i} \epsilon(x) > 0, \tilde{\epsilon}_i = \min_{\ell\in J(i)} \epsilon_\ell > 0.$$

Since $\bar{V}_i$ is compact, the subset $\{(x,p) \in T^*M \mid x \in \bar{V}_i, \|p\|_x \leq \tilde{R}_i + 1\}$ is also compact, therefore by continuity of H, we can find $\eta_i > 0$ such that

$$\forall x \in \bar{V}_i, \forall p, p' \in T^*_x M, \|p\|_x \leq \tilde{R}_i + 1, \|p'\|_x \leq \eta_i, H(x,p) \leq c + \frac{\epsilon_i}{2}$$
$$\Rightarrow H(x, p+p') \leq c + \epsilon_i.$$

We can now choose $\tilde{\eta}_i > 0$ such that $\tilde{j}(i)K_i\tilde{\eta}_i < \min_{\ell\in J(i)} \eta_\ell$. Noting that $H(x,p)$ and $\|p\|_x$ are both quasi-convex in p, and that $\bar{V}_i$ is compact and contained in the domain U_i of a chart, by lemma 8.3.2, for each $i \in \mathbb{N}$, we can find a C^∞ function $u_i : V_i \to \mathbb{R}$ such that

$$\forall x \in V_i, |u(x) - u_i(x)| \leq \min(\tilde{\delta}_i, \tilde{\eta}_i),$$
$$H(x, d_x u_i) \leq \sup_{z\in V_i} H(z, d_z u) + \frac{\tilde{\epsilon}_i}{2} \leq c + \frac{\tilde{\epsilon}_i}{2}$$
$$\|d_x u_i\|_x \leq \sup_{z\in V_i} \|d_z u\|_z + 1 = R_i + 1,$$

where the sup in the last two lines is taken over the set of points $z \in V_i$ where $d_z u$ exists.

We now define $v = \sum_{i\in\mathbb{N}} \theta_i u_i$, it is obvious that v is C^∞. We fix $x \in M$, and choose $i_0 \in \mathbb{N}$ such that $x \in V_{i_0}$. If $\theta_i(x) \neq 0$ then necessarily $V_i \cap V_{i_0} \neq \emptyset$ and therefore $i \in J(i_0)$. Hence $\sum_{i\in J(i_0)} \theta_i(x) = 1$, and $v(x) = \sum_{i\in J(i_0)} \theta_i(x)u_i(x)$.We can now write

$$\begin{aligned}|u(x) - v(x)| &\leq \sum_{i\in J(i_0)} \theta_i(x)|u(x) - u_i(x)| \leq \sum_{i\in J(i_0)} \theta_i(x)\tilde{\delta}_i \\ &\leq \sum_{i\in J(i_0)} \theta_i(x)\delta_{i_0} = \delta_{i_0} \leq \delta(x).\end{aligned}$$

We now estimate $H(x, d_x u)$. First we observe that $\sum_{i\in J(i_0)} \theta_i(y) = 1$, and $v(x) = \sum_{i\in J(i_0)} \theta_i(y)u_i(y)$, for every $y \in V_{i_0}$. Since V_{i_0} is a neighborhood of x, we can differentiate to obtain $\sum_{i\in J(i_0)} d_x\theta_i = 0$, and

$$d_x v = \underbrace{\sum_{i\in J(i_0)} \theta_i(x)d_x u_i}_{p(x)} + \underbrace{\sum_{i\in J(i_0)} u_i(x)d_x\theta_i}_{p'(x)}\,.$$

Using the quasi-convexity of H in p, we get

$$H(x, p(x)) \leq \max_{i\in J(i_0)} H(x, d_x u_i) \leq \max_{i\in J(i_0)} c + \frac{\tilde{\epsilon}_i}{2} \leq c + \frac{\epsilon_{i_0}}{2}, \qquad (*)$$

where for the last inequality we have used that $i \in J(i_0)$ means $V_i \cap V_{i_0} \neq \emptyset$, and therefore $i_0 \in J(i)$, which implies $\tilde{\epsilon}_i \leq \epsilon_{i_0}$, by the definition of $\tilde{\epsilon}_i$.

In the same way, we have

$$\|p(x)\|_x \leq \max_{i\in J(i_0)})\|d_x u_i\|_x \leq \max_{i\in J(i_0)} R_i + 1 \leq \tilde{R}_{i_0} + 1. \qquad (**)$$

We now estimate $\|p'(x)\|_x$. Using $\sum_{i\in J(i_0)} d_x\theta_i = 0$, we get

$$p'(x) = \sum_{i\in J(i_0)} u_i(x)d_x\theta_i = \sum_{i\in J(i_0)} (u_i(x) - u(x))d_x\theta_i.$$

Therefore

$$\begin{aligned}\|p'(x)\|_x &= \|\sum_{i\in J(i_0)} (u_i(x) - u(x))d_x\theta_i\|_x = \sum_{i\in J(i_0)} |u_i(x) - u(x)|\|d_x\theta_i\|_x \\ &\leq \sum_{i\in J(i_0)} \tilde{\eta}_i K_i. \qquad (***)\end{aligned}$$

From the definition of $\tilde{\eta}_i$, we get $K_i\tilde{\eta}_i \leq \frac{\eta_{i_0}}{j(i_0)}$, for all $i \in J(i_0)$. Hence $\|p'(x)\|_x \leq \sum_{i\in J(i_0)} \frac{\eta_{i_0}}{j(i_0)} = \eta_{i_0}$. The definition of η_{i_0}, together with the inequalities (*), (**) and (***), above implies $H(x, d_x v) = H(x, p(x) + p'(x)) \leq c + \epsilon_{i_0} \leq c + \epsilon(x)$. □

Theorem 8.3.6. *Suppose $H : T^*M \to \mathbb{R}$ is a Hamiltonian quasi-convex in the fibers. Let $u : M \to \mathbb{R}$ be a locally Lipschitz viscosity subsolution of $H(x, d_x u) = c$ which is strict at every point of an open subset $U \subset M$. For every continuous function $\epsilon : U \to]0, +\infty[$, we can find a viscosity subsolution $u_\epsilon : M \to \mathbb{R}$ of $H(x, d_x u) = c$ such that $u = u_\epsilon$ on $M \setminus U$, $|u(x) - u_\epsilon(x)| \leq \epsilon(x)$, for every $x \in M$, and the restriction $u_{\epsilon|U}$ is a C^∞ with $H(x, d_x u) < c$ for each $x \in U$.*

Proof. We define $\tilde{\epsilon} : M \to \mathbb{R}$ by $\tilde{\epsilon}(x) = \min(\epsilon(x), d(x, M \setminus U)^2)$, for $x \in U$, and $\tilde{\epsilon}(x) = 0$, for $x \notin U$. It is clear that $\tilde{\epsilon}$ is continuous on M and $\tilde{\epsilon} > 0$ on U.

For each $x \in U$, we can find $c_x < c$, and $V_x \subset V$ an open neighborhood of x such that $H(y, d_y u) \leq c_x$, for almost every $y \in V_x$. The family $(V_x)_{x\in U}$ is an open cover of U, therefore we can find a locally finite partition of unity $(\varphi_x)_{x\in U}$ on U submitted to the open cover $(V_x)_{x\in U}$. We define $\delta : U \to]0, +\infty[$ by $\delta(g) = \sum_{x\in U} \varphi_x(y)(c - c_x)$, for $y \in U$. It is not difficult to check that $H(y, d_y u) \leq c - \delta(y)$ for almost every $y \in U$.

We can apply theorem 8.3.5 to the Hamiltonian $\tilde{H} : T^*U \to \mathbb{R}$ defined by $\tilde{H}(y, p) = H(y, p) + \delta(y)$ and $u|U$ which satisfies $\tilde{H}(y, d_y u) \leq c$ for almost every $y \in U$, we can therefore find a C^∞ function $u_\epsilon : U \to \mathbb{R}$, with $|u_\epsilon(y) - u(y)| \leq \tilde{\epsilon}(y)$, and $\tilde{H}(y, d_y u_\epsilon) \leq c + \delta(y)/2$, for each $y \in U$. Therefore, we obtain $|u_\epsilon(y) - u(y)| \leq \epsilon(y)$, and $H(y, d_y u_\epsilon) \leq c - \delta(y)/2 < c$, for each $y \in U$. Moreover, since $\tilde{\epsilon}(y) \leq d(y, M \setminus U)^2$, it is clear that we can extend continuously u_ϵ by u on $M \setminus U$. This extension satisfies $|u_\epsilon(x) - u(x)| \leq d(x, M \setminus U)^2$, for every $x \in M$. We must verify that u_ϵ is a viscosity subsolution of $H(x, d_x u_\epsilon) = c$. This is clear on U, since u_ϵ is C^∞ on U, and $H(y, d_y u_\epsilon) < c$, for $y \in U$. It remains to check that if $\phi : M \to \mathbb{R}$ is such that $\phi \geq u_\epsilon$ with equality at $x_0 \notin U$ then $H(x_0, d_{x_0}\phi) \leq c$. For this, we note that $u_\epsilon(x_0) = u(x_0)$, and $u(x) - u_\epsilon(x) \leq d(x, M \setminus U)^2 \leq d(x, x_0)^2$. Hence $u(x) \leq \phi(x) + d(x, x_0)^2$, with equality at x_0. The function

$x \to \phi(x) + d(x, x_0)^2$ has a derivative at x_0 equal to $d_{x_0}\phi$, therefore $H(x_0, d_{x_0}\phi) \le c$, since u is a viscosity solution of $H(x, d_x u) \le c$. $\square$

8.4 The viscosity semi-distance

We will suppose that $H : T^*M \to \mathbb{R}$ is a continuous Hamiltonian coercive above every compact subset of the connected manifold M.

We define c[0] as the infimum of all $c \in \mathbb{R}$, such that $H(x, d_x u) = c$ admits a global subsolution $u : M \to \mathbb{R}$. This definition is coherent with the one we gave in earlier chapters for particular Hamiltonians.

As before we denote by $\mathcal{SS}^c$ the set of viscosity subsolutions of $H(x, d_x u) = c$, and by $\mathcal{SS}^c_{\hat{x}} \subset \mathcal{SS}^c$ the subset of subsolutions vanishing at a given $\hat{x} \in M$. Of course, since we can always add a constant to a viscosity subsolution and still obtain a subsolution, we have $\mathcal{SS}^c_{\hat{x}} \neq \emptyset$ if and only if $\mathcal{SS}^c \neq \emptyset$, and in that case $\mathcal{SS}^c = \mathbb{R} + \mathcal{SS}^c_{\hat{x}}$.

Proposition 8.4.1. *Under the above hypothesis, the constant $c[0]$ is finite and there exists a global $u : M \to \mathbb{R}$ viscosity subsolution of $H(x, d_x u) = c[0]$.*

Proof. Fix a point $\hat{x} \in M$. Subtracting $u(\hat{x})$ if necessary, we will assume that all the viscosity subsolutions of $H(x, d_u) = c$ we consider vanish at $\hat{x}$. Since H is coercive above every compact subset of M, for each c the family of functions in $\mathcal{SS}^c_{\hat{x}}$ is locally equi-Lipschitzian, therefore

$$\forall x \in M, \sup_{v \in \mathcal{SS}^c_{\hat{x}}} |v(x)| < +\infty,$$

since M is connected, and every $v \in \mathcal{SS}^c_{\hat{x}}$ vanish at $\hat{x}$. We pick a sequence $c_n \searrow c[0]$, and a sequence $u_n \in \mathcal{SS}^{c_n}_{\hat{x}}$. Since, by Ascoli's theorem, the family $\mathcal{SS}^c_{\hat{x}}$ is relatively compact in the topology of uniform convergence on each compact subset, extracting a sequence if necessary, we can assume that u_n converges uniformly to u on each compact subset of M. By the stability theorem 8.1.1, the function u is a viscosity subsolution of $H(x, d_x u) = c_n$, for each

n, it is therefore locally Lipschitz. Pick a point where $d_{x_0}u$ exists, we have $H(x_0, d_{x_0}u) \leq c_n$, for each n, therefore $c[0] \geq H(x_0, d_{x_0}u)$ has to be finite. Again by the stability theorem 8.1.1, the function u is a viscosity subsolution of $H(x, d_x u) = c[0]$. □

For $c \geq c[0]$, we define

$$S^c(x,y) = \sup_{u \in \mathcal{SS}^c} u(y) - u(x) = \sup_{u \in \mathcal{SS}^c_x} u(y).$$

It follows from the 8.2.2, that for each $x \in M$ the function $S^c(x,.)$ is a viscosity subsolution of $H(y, d_y u) = c$ on M itself, and a viscosity solution on $M \setminus \{x\}$.

Theorem 8.4.2. *For each $c \geq c[0]$, the function S^c is a semi-distance, i.e. it satisfies*

(i) *for each $x \in M, S^c(x,x) = 0$,*

(ii) *for each $x, y, z \in M$, $S^c(x,z) \leq S^c(x,y) + S^c(y,z)$*

Moreover, for $c > c[0]$, the symmetric semi-distance, $\hat{S}^c(x,y) = S^c(x,y) + S^c(y,x)$ is a distance which is locally Lipschitz-equivalent to any distance coming from a Riemannian metric.

Proof. The fact that S^c is a semi-distance follows easily from the definition

$$S^c(x,y) = \sup_{u \in \mathcal{SS}^c} u(y) - u(x).$$

Fix a Riemannian metric on the connected manifold M whose associated norm is denoted by $\|\cdot\|$, and associated distance is d. Given a compact subset $K \subset M$, the constant $\sup\{\|p\| \mid x \in K, p \in T_x M, H(x,p) \leq c\}$, is finite since H is coercive above compact subsets of M. It follows from this that for each compact subset $K \subset M$, there exists a constant $L_K < \infty$ such that.

$$\forall x, y \in K, S^c(x,y) \leq L_K d(x,y).$$

It remains to show a reverse inequality for $c > c[0]$. Fix such a c, and a compact set $K \subset M$. Choose $\delta > 0$, such that $\bar{N}_\delta(K) = \{x \in M \mid d(x,K) \leq \delta\}$ is also compact. By the compactness of the set

$$\{(x,p) \mid x \in \bar{N}_\delta(K), H(x,p) \leq c[0]\},$$

and the continuity of H, we can find $\epsilon > 0$ such that

$$\forall x \in \bar{N}_\delta(K), \forall p, p' \in T_xM, H(x,p) \leq c[0] \text{ and } \|p'\| \leq \epsilon \\ \Rightarrow H(x, p+p') \leq c. \tag{*}$$

We can find $\delta_1 > 0$, such that the radius of injectivity of the exponential map, associated to the Riemannian metric, is at least δ_1 at every point x in the compact subset $\bar{N}_\delta(K)$. In particular, the distance function $x \mapsto d(x, x_0)$ is C^∞ on $\mathring{B}(x_0, \delta_1) \setminus \{x_0\}$, for every $x_0 \in \bar{N}_\delta(K)$. The derivative of $x \mapsto d(x, x_0)$ at each point where it exists has norm 1, since this map has (local) Lipschitz constant equal to 1. We can assume $\delta_1 < \delta$. We now pick $\phi : \mathbb{R} \to \mathbb{R}$ a C^∞ function, with support in $]1/2, 2[$, and such that $\phi(1) = 1$. If $x_0 \in K$ and $0 < d(y, x_0) \leq \delta_1/2$, the function

$$\phi_y(x) = \phi(\frac{d(x, x_0)}{d(y, x_0)})$$

is C^∞. In fact, if $d(x, x_0) \geq \delta_1$, then ϕ_y is zero in a neighborhood of x, since $d(x, x_0)/d(y, x_0) \geq \delta_1/(\delta_1/2) = 2$; if $0 < d(x, x_0) < \delta_1 < \delta$, then it is C^∞ on a neighborhood of x; finally $\phi_y(x) = 0$ for x such that $d(x, x_0) \leq d(y, x_0)/2$. In particular, we obtained that $d_x\phi_y = 0$, unless $0 < d(x, x_0) < \delta$, but at each such x, the derivative of $z \mapsto d(z, x_0)$ exists and has norm 1. It is then not difficult to see that $\sup_{x \in M} \|d_x\phi_y\| \leq A/d(y, x_0)$, where $A = \sup_{t \in \mathbb{R}} |\phi'(t)|$.

Therefore if we set $\lambda = \epsilon d(y, x_0)/A$, we see that $\|\lambda d_x\phi_y\| \leq \epsilon$, for $x \in M$. Since ϕ is 0 outside the ball $B(x_0, \delta_1) \subset N_{\delta_1}(K)$, it follows from the property (*) characterizing ϵ that we have

$$\forall (x, p) \in T^*M, H(x, p) \leq c[0] \Rightarrow H(x, p + \lambda d_x\phi_y) \leq c.$$

Since $S^{c[0]}(x_0, \cdot)$ is a viscosity subsolution of $H(x, d_xu) = c[0]$, and ϕ_y is C^∞, we conclude that the function $u(.) = S^{c[0]}(x_0, .) + \lambda\phi_y(.)$ is a viscosity subsolution of $H(x, d_xu) = c$. But the value of u at x_0 is 0, and its value at y is $S^{c[0]}(x_0, y) + \lambda\phi_y(y) = S^{c[0]}(x_0, y) + \epsilon d(y, x_0)/A$, since $\phi_y(y) = \phi(1) = 1$. Therefore $S^c(x_0, y) \geq S^{c[0]}(x_0, y) + \epsilon d(y, x_0)/A$. Hence we obtained

$$\forall x, y \in K, d(x, y) \leq \delta_1/2 \Rightarrow S^c(x, y) \geq S^{c[0]}(x, y) + \epsilon A^{-1} d(x, y).$$

Adding up and using $S^{c[0]}(x,y)+S^{c[0]}(y,x)\geq S^{c[0]}(x,x)=0$, we get

$$\forall x,y\in K, d(x,y)\leq \delta_1/2 \Rightarrow S^c(x,y)+S^c(y,x)\geq \frac{2\epsilon}{A}d(x,y). \quad \square$$

8.5 The projected Aubry set

Theorem 8.5.1. *Assume that $H: T^*M\to\mathbb{R}$ is a Hamiltonian coercive above every compact subset of the connected manifold M. For each $c\geq c[0]$, and each $x\in M$, the following two conditions are equivalent:*

(i) *The function $S^c(x,\cdot)$ is a viscosity solution of $H(z,d_zu)=c$.*

(ii) *There is no viscosity subsolution of $H(z,d_zu)=c$ which is strict at x.*

Proof. The implication (ii)⇒(i) follows from proposition 8.2.8.

To prove (i)⇒(ii), fix $x\in M$ such that S_x is a viscosity solution on the whole of M, and suppose that $u: M\to\mathbb{R}$ is a viscosity subsolution of $H(y,d_yu)=c$ which is strict at x. Therefore we can find an open neighborhood V_x of x, and a $c_x<c$ such that $u_{|V_x}$ is a viscosity subsolution of $H(y,d_yu)=c_x$ on V_x. We can assume without loss of generality that V_x is an open subset of $\mathbb{R}^n$ and $u(x)=0$. We have

$$u(y)\leq S(x,y)$$

and $u(x)=S(x,x)=0$. On $V_x\subset\mathbb{R}^n$, we can define $u_1(y)=u(y)-\frac{1}{2}\|x-y\|^2$. Define $\epsilon(\delta)>0$ by

$$\epsilon(\delta)=\max_{\|x-y\|\leq\delta}\{H(y,p+p')-c_x \mid H(y,p)\leq c_x\|p'\|\leq\delta\}.$$

Since H is continuous and coercive above compact subsets we have $\epsilon(\delta)\to 0$, when $\delta\to 0$. Since the derivative at y_0 of $y\to\frac{1}{2}\|y-x\|^2$ is $\langle y_0-x,\cdot\rangle$, we see that $u_{1|\mathring{B}(x,\delta)}$ is a viscosity sub solution of $H(y,d_yu_1)=c_x+\epsilon(\delta)$. We fix $\delta>0$ such that $c_x+2\epsilon(\delta)<c$. By 8.2.5, we can find a real-valued function u_2 defined on a

neighborhood of $\bar{B}(x, \delta/2)$ and semi-convex such that u_2 is as close as we want to u_1 on $\bar{B}(x\delta/2)$, and u_2 is a viscosity subsolution of $H(x, d_x u_2) = c_x + 2\epsilon(\delta)$. In a neighborhood of $\bar{B}(0, \delta/2)$. We have $u_1(x) = S(x, x) = 0$ and $u_1(y) \leq u(y) - \frac{1}{2}\|x-y\|^2 \leq u(y) \leq S(x, y)$ hence $S(x, y) - u_1(y) \geq \frac{1}{2}\delta$ on the boundary $\partial B(x, \delta/2)$. We can therefore choose u_2 close enough to u_1 so that $S(x, \cdot) - u_2(\cdot)$ attains its minimum on $\bar{B}(x, \delta/2)$ at a point $y_0 \in \mathring{B}(x, \delta/2)$. Therefore $S(x, y) \geq S(x, y_0) - u_2(y_0) + u_2(y)$ in a neighborhood of y_0, and therefore $D^- u_2(y_0) \subset D^- S_x(y_0)$. Since u_2 is semi-convex and is a viscosity subsolution of $H(y, dyu_2) = c_x + 2\epsilon(\delta)$ on a neighborhood of $\bar{B}(x, \delta/2)$, by an argument analogous to the proof of theorem 8.2.4, we can find $p_0 \in D^- u_2(x_0)$ with $H(y_0, p_0) \leq c_x + 2\epsilon(\delta)$. Since $D^- u_2(x_0) \subset D^- S_x(y_0)$, and S_x is a viscosity solution of $H(y, d_y S_x) = c$ on M, we must have $H(y_0, p_0) \geq c$. This is a contradiction since $c > c_x + 2\epsilon(\delta)$. □

Definition 8.5.2 (Projected Aubry set)**.** If $H : T^*M \to \mathbb{R}$ is a continuous Hamiltonian, coercive above every compact subset of the connected manifold M. We define the projected Aubry set as the set of $x \in M$ such that that $S^{c[0]}(x, \cdot)$ is a viscosity solution of $H(z, d_z u) = c[0]$.

Proposition 8.5.3. *Assume that $H : T^*M \to \mathbb{R}$ s a continuous Hamiltonian, convex is the fibers, and coercive above every compact subset of the connected manifold M. There exists a viscosity subsolution $v : M \to \mathbb{R}$ of $H(x, d_x v) = c[0]$, which is strict at every $x \in M \setminus \mathcal{A}$.*

Proof. We fix some base point $\hat{x} \in M$. For each $x \notin \mathcal{A}$, we can find $u_x : M \to \mathbb{R}$, an open subset V_x containing x, and $c_x < c[0]$, such that u_x is a viscosity subsolution of $H(y, d_y u_x) = c[0]$ on M, and $u_x|V_x$ is a viscosity subsolution of $H(y, d_y u_x) \leq c_x$, on V_x. Subtracting $u_x(\hat{x})$ if necessary, we will assume that $u_x(\hat{x}) = 0$. Since $U = M \setminus \mathcal{A}$ is covered by the family of open sets $V_x, x \notin \mathcal{A}$, we can extract a countable subfamily $(V_{x_i})_{i \in \mathbb{N}}$ covering U. Since H is coercive above every compact set the sequence $(u_{x_i})_{i \in \mathbb{N}}$ is locally equi-Lipschitzian. Therefore, since M is connected, and all the u_{x_i} vanish at $\hat{x}$, the sequence $(u_{x_i})_{i \in \mathbb{N}}$ is uniformly bounded on every compact subset of M. It follows that the sum $V = \sum_{i \in \mathbb{N}} \frac{1}{2^{i+1}} u_{x_i}$ is uniformly convergent on each compact subset. If we set $u_n =$

$(1-2^{-(n+1)})^{-1}\sum_{0\leq i\leq n}\frac{1}{2^{i+1}}u_{x_i}$, then u_n is a viscosity subsolution of $H(x,d_xu_n)=c[0]$ as a convex combination of viscosity subsolutions, see proposition 8.3.3. Since u_n converges uniformly on compact subsets to u, the stability theorem 8.1.1 implies that v is also a viscosity subsolution of $H(x,d_xv)=c[0]$.

On the set $V_{x_{n_0}}$, we have $H(x,d_xu_{x_{n_0}})\leq c_{x_{n_0}}$, for almost every $x\in V_{x_{n_0}}$. Therefore, if we fix $n\geq n_0$, we see that for almost every $x\in V_{x_{n_0}}$ we have

$$\begin{aligned}H(x,d_xu_n)&\leq(1-2^{-(n+1)})^{-1}\sum_{i=0}^{n}\frac{1}{2^{i+1}}H(x,d_xu_{x_i})\\&\leq(1-2^{-(n+1)})^{-1}\left[\sum_{i=0}^{n}\frac{1}{2^{i+1}}c[0]+\frac{(c_{x_{n_0}}-c[0])}{2^{n_0+1}}\right].\end{aligned}$$

Therefore $u_n|V_{x_{n_0}}$ is a viscosity subsolution of $H(x,d_xu_n)\leq c[0]+(c_{x_{n_0}}-c[0])/2^{n_0+1}$.

By the stability theorem, this is also true for $v|V_{x_{n_0}}$. Since $c_{x_{n_0}}-c[0]<0$, we conclude that $u|V_{x_{n_0}}$ is a strict subsolution of $H(x,d_xv)=c[0]$, for each $x\in V_{x_{n_0}}$, and therefore at each $x\in U\subset\cup_{n\in\mathbb{N}}V_{x_n}$. □

Theorem 8.5.4. *Assume that $H:T^*M\to\mathbb{R}$ is a Hamiltonian convex in the fibers and coercive, where M is a compact connected manifold. Its projected Aubry set $\mathcal{A}$ is not empty.*

If two viscosity solutions of $H(x,d_xu)=c[0]$ coincide on $\mathcal{A}$, they coincide on M.

Theorem 8.5.5. *Suppose $u_1,u_2:M\to\mathbb{R}$ are respectively a viscosity subsolution and a viscosity supersolution of $H(x,d_xu)=c[0]$. If $u_1\leq u_2$ on the projected Aubry set $\mathcal{A}$, then $u_1\leq u_2$ everywhere on M.*

Proposition 8.5.6. *Assume that $H:T^*M\to\mathbb{R}$ is a Hamiltonian convex in the fibers and coercive, where M is a compact connected manifold. If M is compact and connected, for each viscosity subsolution $u:M\to\mathbb{R}$ of $H(x,d_xu)=c[0]$, and each $\epsilon>0$, we can find a viscosity subsolution $u_\epsilon:M\to\mathbb{R}$of $H(x,d_xu_\epsilon)=c[0]$ such that $\|u-u_\epsilon\|_\infty<\epsilon$, and u_ϵ is C^∞ on $M\setminus\mathcal{A}$, with $H(x,d_xu_\epsilon)<c[0]$, for each $x\in M\setminus\mathcal{A}$.*

Proof. Call v the strict subsolution given by the previous proposition 8.5.3. By a similar argument to the one used in the proof of that proposition $v_\delta = (1-\delta)u + \delta v$ is a viscosity subsolution of $H(x, d_x v_\delta) = c[0]$ which is strict a each point of $M \setminus \mathcal{A}$, and $v_\delta \to u$ uniformly as $\delta \to 0$. It then suffices to choose δ small enough and to apply 8.3.6 to v_δ to obtain the function u_ϵ. □

Proof of theorem 8.5.5. Assume $m = \inf(u_2 - u_1) < 0$. Choose $\epsilon > 0$ such that $m + 2\epsilon < 0$. If we apply proposition 8.5.6, we obtain $\tilde{u}_1 : M \to \mathbb{R}$, with $\|\tilde{u}_1 - u_1\|_\infty < \epsilon$, and $\tilde{u}_1$ of class C^∞on $M \setminus \mathcal{A}$, with $H(x, d_x\tilde{u}_1) < c[0]$, for every $x \notin \mathcal{A}$. We have $u_2(x) - \tilde{u}_1(x) \geq u_2(x) - u_1(x) + \tilde{u}_1(x) - u_1(x) \geq u_2(x) - u(x) - \epsilon$, therefore $u_2(x) - \tilde{u}_1(x) \geq -\epsilon$, for $x \in \mathcal{A}$. Moreover, $\inf(u_2 - \tilde{u}_1) \leq \inf(u_2 - u_1) + \|u_1 - \tilde{u}_1\|_\infty \leq m + \epsilon$. Since $m + \epsilon < -\epsilon$, on the compact space M, the infimum of $(u_2 - \tilde{u}_1)$ is attained at a point $x_0 \notin \mathcal{A}$. Since $u_2(x) \geq [u_2(x_0) - \tilde{u}_1(x_0)] + \tilde{u}_1(x)$, with equality at x_0, the function $\tilde{u}_1$ is differentiable on $M \setminus \mathcal{A} \ni x_0$, and u_2 is a supersolution of $H(x, d_x u_2) = c[0]$, we must have $H(x, d_x\tilde{u}_1) \geq c[0]$. This is impossible by the choice of $\tilde{u}_1$. □

8.6 The representation formula

We still assume that M is compact, and that $H : T^*M \to \mathbb{R}$ is a coercive Hamiltonian convex in the fibers.

Theorem 8.6.1. *Any viscosity solution $u : M \to \mathbb{R}$ for $H(x, d_x u) = c[0]$ satisfies*

$$\forall x \in M, u(x) = \inf_{x_0 \in \mathcal{A}} u(x_0) + S^{c[0]}(x_0, x)$$

.

This theorem follows easily from the uniqueness theorem 8.5.4 and the following one:

Theorem 8.6.2. *For any function $v : \mathcal{A} \to \mathbb{R}$ bounded below, the function*

$$\tilde{v}(x) = \inf_{x_0 \in \mathcal{A}} v(x_0) + S^{c[0]}(x_0, x)$$

is a viscosity solution of $H(x, d_x v) = c[0]$. Moreover, we have $\tilde{v}_{|\mathcal{A}} = v$, if and only if

$$\forall x, y \in \mathcal{A}, v(y) - v(x) \leq S(x, y).$$

We start with a lemma.

Lemma 8.6.3. *Suppose $H : T^*M \to \mathbb{R}$ is a continuous Hamiltonian convex in the fibers, and coercive above each compact subset of the connected manifold M. Let $u_i : M \to \mathbb{R}, i \in I$ be a family of viscosity subsolutions of $H(x, d_x u) = c$. If $\inf_{i\in I} u_i(x_0)$, is finite for some $x_0 \in M$, then $\inf_{i\in I} u_i$ is finite everywhere. In that case, the function $u = \inf_{i\in I} u_i$ is a viscosity subsolution of $H(x, d_x u) \leq c$.*

Proof. We choose an auxiliary Riemannian metric on M, and use the associated distance.

By the coercivity condition, the family $(u_i)_{i\in I}$ is locally equi-Lipschitzian, therefore for if K compact connected subset of M, there exists a constant $C(K)$ such that

$$\forall x, y \in K, \forall i \in I, |u_i(x) - u_i(y)| \leq C(K).$$

If $x \in M$ is given, we can find a compact connected subset K_x containing x_0 and x, it follows that

$$\inf_{i\in I} u_i(x_0) \leq \inf_{i\in I} u_i(x) + C(K_x)$$

therefore $\inf_{i\in I} u_i$ is finite everywhere. It now suffices to show that for a given $\tilde{x} \in M$, we can find an open neighborhood V of $\tilde{x}$ such that $\inf_{i\in I} u_i|V$ is a viscosity subsolution of $H(x, d_x u) = c$ on V. We choose an open neighborhood V of $\tilde{x}$ such that its closure $\bar{V}$ is compact. Since $\mathcal{C}^0(\bar{V}, \mathbb{R})$ is metric and separable in the topology of uniform convergence, we can find a countable subset $I_0 \subset I$ such that $u_{i|\bar{V}}, i \in I_0$ is dense in $\{u_{i|\bar{V}} \mid i \in I\}$, for the topology of uniform convergence. Therefore $\inf_{i\in I} u_i = \inf_{i\in I'} u_i = \inf_{i\in I_0} u_i$ on $\bar{V}$. Since I_0 is countable, we have reduced to the case $I_0 = \{0, \cdots, N\}$, or $I_0 = \mathbb{N}$.

Let us start with the first case. Since $u_0, \cdots, u_N$, and $u = \inf_{i=0}^N u_i$ are all Lipschitzian on V, we can find $E \subset V$ of full Lebesgue measure such that $d_x u, d_x u_0, \cdots, d_x u_N$ exists, for each $x \in E$. At each such $x \in E$, we necessarily have $d_x u \in \{d_x u_0, \ldots, d_x u_N\}$. In fact, if n is such that $u(x) = u_n(x)$, since $u \leq u_n$ with equality at x and both derivative at x exists, they must be equal. Since each u_i is a viscosity subsolution of $H(x, d_x v) = c$, we obtain

$H(x, d_x u) \leq c$, for every x in the subset E of full measure in V. The convexity of H in the fibers imply that u is a viscosity subsolution of $H(x, d_x u) = c$ in V. It remains to consider the case $I_0 = \mathbb{N}$. Define $u^N(x) = \inf_{0 \leq i \leq N} u_i(x)$, by the previous case, u^N is a viscosity subsolution of $H(x, d_x u^N) = c$ on V.

Now $u^N(x) \to \inf_{i \in I_0} u_i(x)$, for each $x \in \bar{V}$, the convergence is in fact, uniform on $\bar{V}$ since $(u_i)_{i \in I_0}$ is equi-Lipschitzian on the compact set $\bar{V}$. It remains to apply the stability theorem 8.1.1. $\square$

Chapter 9

Mañé's Point of View

Ricardo Mañé's last paper [Mn97] contained a version of the weak KAM theorem. The point of view is probably the closest to the theory of optimal contral. His ideas after his untimely death were carried out much further by G. Contreras, J. Delgado, R. Iturriaga, Gabriel and Miguel Paternain [CDI97, CIPP98].

There is an excellent reference on Mañé's point of view and th subsequent developments [CI99].

9.1 Mañé's potential

As in definition 5.3.1, we set

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

Where the infimum is taken over all continuous piecewise C^1 curves $\gamma : [0, t] \to M$ with $\gamma(0) = x, \gamma(t) = y$.

Definition 9.1.1 (Mañé's potential)**.** Fix $c \in \mathbb{R}$ for each $x, y, \in M$, we set

$$m_c(x, y) = \inf_{t>0} h_t(x, y) + ct$$

Here are some properties of m_c:

Proposition 9.1.2. *For each $c \in \mathbb{R}$, the Mañé potential m_c has values in $\mathbb{R} \cup \{-\infty\}$, and satisfies the following properties:*

(i) *If* $x, u \in M$, *and* $c, c' \in \mathbb{R}$, *with* $c \leq c'$, *we have* $m_c(x,y) \leq m_{c'}(x,y)$.

(ii) *For all* $c \in \mathbb{R}, y, z \in M$, *we have*

$$m_c(x,z) \leq m_c(x,y) + m_c(x,z).$$

(iii) *If* $A = \sup\{L(x,v) \mid (x,v) \in TM, \|v\|_x \leq 1\}$ *we have* $m_c(x,y) \leq (A+c)d(x,y)$.

(iv) *For a given* $c \in \mathbb{R}$ *either* m_c *is equal identically to* $-\infty$ *or* m_c *is finite everywhere.*

(v) *For every* $c \in \mathbb{R}$, *either* $m_c \equiv_\infty$ *or* $m_c(x,x) = 0$ *for every* $x \in M$.

(vi) *If* m_c *is finite then it is Lipschitz.*

(vii) *For* $u : M \to \mathbb{R}$, *we have* $u \prec L + c$ *if and only if*

$$\forall x, y \in M, u(y) - u(x) \leq m_c(x,y).$$

(viii) *If* m_c *is finite, then for each* $x \in M$, *the function* $m_{c,x} : M \to \mathbb{R}, y \mapsto m_c(x,y)$ *(resp.* $-m_c^x : M \to \mathbb{R}, y \mapsto -m_c(y,x)$*) is dominated by* $L + c$.

(ix) *The Mañé critical value* $c[0]$ *is equal to the infimum of the set of* $c \in \mathbb{R}$ *such that* m_c *is finite. Moreover, the critical Mañé potential* $m^0 = m_{c[0]}$ *is finite everywhere.*

Proof. Property (i) is obvious. Property 3 (ii) results from

$$h_t(x,z) \leq h_t(x,y) = h_t(y,z).$$

For property (iii), if we use a geodesic $\gamma_{x,y} : [0, d(x,y)] \to M$ from x to y parametrized by arc-length, we see that $m_c(x,y) \leq h_{d(x,y)}(x,y) + cd(x,y) \leq \mathbb{L}(\gamma_{x,y}) + cd(x,y) \leq (A+c)d(x,y)$. For property (iv), we remark that $m_c(x',y') \leq m_c(x',x) + m_c(x,y) + m_c(y,y') \leq m_c(x,y) + (A+c)[d(x',x) + d(y',y)]$ hence if $m_c(x,y) = -\infty$ for some $(x,y) \in M \times M$ then $m_c(x',y') = -\infty$ for every $(x',y') \in M \times M$.

For property (v), using constant paths, we first remark that $m_c(x,x) \leq (L(x,0)+c)t$, for every $t>0$, therefore $m_c(x,x) \leq 0$. Moreover, by (ii)

$$m_c(x,x) \leq m_c(x,x) + m_c(x,x) \leq \cdots \leq n m_c(x,x).$$

Hence $m_c(x,x) < 0$ implies $m_c(x,x) = -\infty$.

Property (vi) follows from the proof of (iv), since we obtained there

$$m_c(x',y') \leq m_c(x,y) + (A+c)[d(x,x') + d(y,y')]$$

which gives by symmetry

$$|m_c(x',y') - m_c(x,y)| \leq [A+c][d(x,x') + d(y,y')].$$

Property (vii) is obvious since $u \prec L+c$ if and only if

$$\forall t > 0, u(y) - u(x) \leq h_t(x,y) + ct.$$

For (viii), the inequality obtained in (ii)

$$m_c(x,z) \leq \tilde{S}_c(x,y) + m_c(y,z)$$

gives, when m_c is finite

$$m_c(x,z) - m_c(x,y) \leq m_c(y,z).$$

But this can be rewritten as

$$m_{c,x}(z) - m_{c,x}(y) \leq m_c(y,z),$$

therefore $m_{c,x} \prec L+c$ by (vii).

For (ix), if $c \geq c[0]$, there exists $u : M \to \mathbb{R}$ with $u \prec L+c$ therefore by (vii), we have m_c finite.

Conversely if m_c is finite $m_{c,x} \prec L+c$ therefore $c \geq c[0]$. □

Corollary 9.1.3. *For each $c \geq c[0]$, the Mañé potential m_c is equal to the viscosity semi-distance S^c, and therefore*

$$\forall x,y \in M, m_c(x,y) = \sup\{u(y) - u(x) \mid u \prec L+c\}.$$

Proof. The function $m_{c,x}$ (resp. S^c_x) is a viscosity subsolution of $H(x,d_u) = c$ (resp. is dominated by $L+c$), therefore $m_c(x,y) = m_{c,x}(y) - m_{c,x}(x) \leq S^c(x,y)$ (resp. $S^c(x,y) = S^c_x(y) - S^c_x(x) \leq m_c(x,y)$). □

Definition 9.1.4 (Mañé's critical potential)**.** We will call $m^0 = m_{c[0]}$ the Mañé critical potential.

9.2 Semi-static and static curves

Proposition 9.2.1. *Given $c \in \mathbb{R}$, a curve $\gamma : [a,b] \to M$ is an absolute $(L+c)$-minimizer if and only if*

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

Proof. Suppose that $\gamma : [a,b] \to M$ is an absolute $(L+c)$-minimizer, then for any curve $\delta : [0,t] \to M$, with $t > 0, \delta(0) = \gamma(a)$, and $\delta(t) = \gamma(b)$, we have

$$\int_0^t L(\delta(s),\dot\delta(s))\,ds + ct \geq \int_a^b L(\gamma(s),\dot\gamma(s))ds + c(b-a)$$

therefore $h_t(\gamma(a),\gamma(b)) + ct \geq \int_a L(\gamma(s),\dot\gamma(s))\,ds + c(b-a)$.

On the other hand reparametrizing γ linearly by $[0, b-a]$, we see that $h_{b-a}(\gamma(b),\gamma(a)) + c(b-a) \leq \int_a^b L(\gamma(s),\dot\gamma(s))\,ds + c(b-a)$. It follows that

$$\begin{aligned} m_c(\gamma(b),\gamma(a)) &= h_{b-a}(\gamma(b),\gamma(a)) + c(b-a) \\ &= \int_a^b L(\gamma(s),\dot\gamma(s))ds + c(b-a). \end{aligned}$$

Conversely, since $m_c(\gamma(a),\gamma(a)) = 0$, the equality

$$m_c(\gamma(a),\gamma(b)) = \int_a^b L(\gamma(s),\dot\gamma(s)) + ds + c(b-a),$$

can be rewritten as

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

This means that γ is $(m_{c,\gamma(a)}, L, c)$-calibrated. $\square$

Definition 9.2.2 (Semi-static curve)**.** A curve $\gamma : [a,b] \to M$ is called semi-static, if $a < b$ and $m^0(\gamma(a),\gamma(b)) = \int_a^b L(\gamma(s),\dot\gamma(s))ds + c[0](b-a)$. (Recall that $m^0 = m_{c[0]}$ is the Mañé potential).

We therefore have the following proposition;

Proposition 9.2.3. *A curve $\gamma : [a,b] \to M$ semi-static if and only if it is absolutely minimizing, if and only if it is $(u, L, c[0])$-calibrated for some $u : M \to \mathbb{R}$ dominated by $L + c[0]$.*

Mañé has also defined a notion of static curve.

Definition 9.2.4 (Static curve)**.** A curve $\gamma : [a,b] \to M$ is static, if $a < b$ and

$$\int_a^b L(\gamma(s), \dot{\gamma}(s))\, ds + c[0](b-a) = -m^0(\gamma(b), \gamma(a)).$$

Proposition 9.2.5. *A curve is static if and only if it is a part of a projected Aubry curve*

Proof. We have

$$0 = \int_a^b L(\gamma(s), \dot{\gamma}(s))\, ds + c[0](b-a) + m^0(\gamma(b), \gamma(a)) = 0.$$

For every $\epsilon > 0$, we can find a curve $\delta_\epsilon : [b, b_\epsilon] \to M$ with $\delta_\epsilon(b) = \gamma(b), \delta_\epsilon(b_\epsilon) = \gamma(a)$, and

$$\int_b^{b_\epsilon} L(\delta_\epsilon(s), \dot{\delta}_\epsilon(s))ds + c[0](b_\epsilon - b) \leq m^0(\gamma(b), \gamma(a)) + \epsilon.$$

Therefore, if we consider the concatenated closed curve $\gamma * \delta_\epsilon$, we find a curve $\tilde{\delta}_\epsilon$ that is a loop at $\gamma(a)$, is parametrized by an interval of length $\ell_\epsilon \geq b - a > 0$ and satisfies $\mathbb{L}(\tilde{\delta}_\epsilon) + c[0]\ell_\epsilon \leq \epsilon$. Going n times through the loop $\tilde{\delta}_{\epsilon/n}$, we find a loop $\tilde{\delta}_\epsilon$ at $\gamma(a)$, parametrized by an interval of length $n\tilde{\ell}_{\epsilon/n} \geq n(b-a)$, with $\mathbb{L}(\tilde{\delta}_{n,\epsilon}) + c[0]n\tilde{\ell}_{\epsilon/n} \leq \epsilon$. Since $b - a > 0$, we have $n(b-a) \to +\infty$ as $n \to +\infty$. It follows that $h(\gamma(a), \gamma(a)) \leq \epsilon$, for every $\epsilon > 0$, where h is the Peierls barrier. Therefore $\gamma(a) \in \mathcal{A}$. Since the loop $\tilde{\delta}_\epsilon = \gamma * \tilde{\delta}_\epsilon$ goes through every point of $\gamma([a,b])$ a similar argument shows $\gamma([a,b]) \subset \mathcal{A}$.

It remains to show that γ is $(u, L, c[0])$-calibrated for every $u : M \to \mathbb{R}$ which is dominated by $L + c[0]$.

In fact, if we add up the two inequalities

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

$$u(\gamma(a)) - u(\gamma(b)) \leq m^0(\gamma(b), \gamma(a)),$$

we obtain the equality $0 = 0$ therefore both inequalities above must be equalities. $\square$

Bibliography

[Bar94] Guy Barles. *Solutions de viscosité des équations de Hamilton-Jacobi.* Springer-Verlag, Paris, 1994.

[BCD97] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.

[BGH98] Giuseppe Buttazo, Mariano Giaquinta, and Stefan Hildebrandt. *One-dimensional variational problems, An introduction.* Oxford University Press, Oxford, 1998. Oxford Lecture Series in Mathematics and its Applications, Vol. 15.

[BGS85] Werner Ballmann, Mikhael Gromov, and Viktor Schroeder. *Manifolds of nonpositive curvature*, volume 61 of *Progress in Mathematics.* Birkhäuser Boston Inc., Boston, MA, 1985.

[Bos86] Jean-Benoît Bost. Tores invariants des systèmes dynamiques hamiltoniens (d'après Kolmogorov, Arnol'd, Moser, Rüssmann, Zehnder, Herman, Pöschel,...). *Astérisque*, (133-134):113–157, 1986. Seminar Bourbaki, Vol. 1984/85.

[Bou76] Nicolas Bourbaki. *Éléments de mathématique.* Hermann, Paris, 1976. Fonctions d'une variable réelle, Théorie élémentaire, Nouvelle édition.

[Bou81] Nicolas Bourbaki. *Espaces vectoriels topologiques. Chapitres 1 à 5.* Masson, Paris, new edition, 1981. Éléments de mathématique. [Elements of mathematics].

[BS00] Guy Barles and P.E. Souganidis. On the long time behavior of Hamilton-Jacobi equations. *J. Math. Analysis*, 2000. to appear.

[Car95] M. J. Dias Carneiro. On minimizing measures of the action of autonomous Lagrangians. *Nonlinearity*, 8(6):1077–1085, 1995.

[CC95] Luis A. Caffarelli and Xavier Cabré. *Fully nonlinear elliptic equations.* American Mathematical Society, Providence, RI, 1995.

[CDI97] Gonzalo Contreras, Jorge Delgado, and Renato Iturriaga. Lagrangian flows: the dynamics of globally minimizing orbits. II. *Bol. Soc. Brasil. Mat. (N.S.)*, 28(2):155–196, 1997.

[CEL84] M. G. Crandall, L. C. Evans, and P.-L. Lions. Some properties of viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 282(2):487–502, 1984.

[CHL91] Dean A. Carlsson, Alain B. Haurie, and Arie Leizarowitz. *Infinite horizon optimal control: determistic and stochastic systems.* Springer-Verlag, New York Berlin Heidelberg, 1991. 2nd, rev. and enl. ed.

[CI99] Gonzalo Contreras and Renato Iturriaga. *Global minimizers of autonomous Lagrangians.* Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1999.

[CIPP98] G. Contreras, R. Iturriaga, G. P. Paternain, and M. Paternain. Lagrangian graphs, minimizing measures and Mañé's critical values. *Geom. Funct. Anal.*, 8(5):788–809, 1998.

[CL83] Michael G. Crandall and Pierre-Louis Lions. Viscosity solutions of Hamilton-Jacobi equations. *Trans. Amer. Math. Soc.*, 277(1):1–42, 1983.

[Cla90] F. H. Clarke. *Optimization and nonsmooth analysis*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

[DG82] James Dugundji and Andrzej Granas. *Fixed point theory. I.* Państwowe Wydawnictwo Naukowe (PWN), Warsaw, 1982.

[Dug66] James Dugundji. *Topology*. Allyn and Bacon Inc., Boston, Mass., 1966.

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

[Eva98] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[Fat97a] Albert Fathi. Solutions KAM faibles conjuguées et barrières de Peierls. *C. R. Acad. Sci. Paris Sér. I Math.*, 325(6):649–652, 1997.

[Fat97b] Albert Fathi. Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(9):1043–1046, 1997.

[Fat98a] Albert Fathi. Orbites hétéroclines et ensemble de Peierls. *C. R. Acad. Sci. Paris Sér. I Math.*, 326(10):1213–1216, 1998.

[Fat98b] Albert Fathi. Sur la convergence du semi-groupe de Lax-Oleinik. *C. R. Acad. Sci. Paris Sér. I Math.*, 327(3):267–270, 1998.

[Fat03] Albert Fathi. Regularity of C^1 solutions of the Hamilton-Jacobi equation. *Ann. Fac. Sci. Toulouse Math. (6)*, 12(4):479–516, 2003.

[Fle69] Wendell H. Fleming. The Cauchy problem for a nonlinear first order partial differential equation. *J. Differential Equations*, 5:515–530, 1969.

[FM00] Albert Fathi and John N. Mather. Failure of the convergence of the Lax-Oleinik semi-group in the time periodic case. *Bull. Soc. Math. France*, 128, 2000.

[FS04] Albert Fathi and Antonio Siconolfi. Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation. *Invent. Math.*, 155(2):363–388, 2004.

[FS05] Albert Fathi and Antonio Siconolfi. PDE aspects of Aubry-Mather theory for quasiconvex Hamiltonians. *Calc. Var. Partial Differential Equations*, 22(2):185–228, 2005.

[GK90] Kazimierz Goebel and W. A. Kirk. *Topics in metric fixed point theory*. Cambridge University Press, Cambridge, 1990.

[Gro90] M. Gromov. Convex sets and Kähler manifolds. In *Advances in differential geometry and topology*, pages 1–38. World Sci. Publishing, Teaneck, NJ, 1990.

[Her89] Michael-R. Herman. Inégalités "a priori" pour des tores lagrangiens invariants par des difféomorphismes symplectiques. *Inst. Hautes Études Sci. Publ. Math.*, (70):47–101 (1990), 1989.

[Kis92] Christer O. Kiselman. Regularity classes for operations in convexity theory. *Kodai Math. J.*, 15(3):354–374, 1992.

[Kni86] Gerhard Knieper. *Mannigfaltigkeiten ohne konjugierte Punkte*. Universität Bonn Mathematisches Institut, Bonn, 1986. Dissertation, Rheinische Friedrich-Wilhelms-Universität, Bonn, 1985.

[Lio82] Pierre-Louis Lions. *Generalized solutions of Hamilton-Jacobi equations*. Pitman (Advanced Publishing Program), Boston, Mass., 1982.

[LPV87] Pierre-Louis Lions, George Papanicolaou, and S.R.S. Varadhan. Homogenization of Hamilton-Jacobi equation. unpublished preprint, 1987.

[Mat91] John N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.*, 207(2):169–207, 1991.

[Mat93] John N. Mather. Variational construction of connecting orbits. *Ann. Inst. Fourier (Grenoble)*, 43(5):1349–1386, 1993.

[Min61] George J. Minty. On the maximal domain of a "monotone" function. *Michigan Math. J.*, 8:135–137, 1961.

[Min64] George J. Minty. On the monotonicity of the gradient of a convex function. *Pacific J. Math.*, 14:243–247, 1964.

[Mn] Ricardo Mañé. *Global Variational methods in conservative dynamics*. IMPA, Rio de Janeiro. Lecture notes, 18 Colóquio Brasileiro de Matemática.

[Mn97] Ricardo Mañé. Lagrangian flows: the dynamics of globally minimizing orbits. *Bol. Soc. Brasil. Mat. (N.S.)*, 28(2):141–153, 1997.

[NR97a] Gawtum Namah and Jean-Michel Roquejoffre. Comportement asymptotique des solutions d'une classe d'équations paraboliques et de Hamilton-Jacobi. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(12):1367–1370, 1997.

[NR97b] Gawtum Namah and Jean-Michel Roquejoffre. Convergence to periodic fronts in a class of semilinear parabolic equations. *NoDEA Nonlinear Differential Equations Appl.*, 4(4):521–536, 1997.

[NR99] Gawtum Namah and Jean-Michel Roquejoffre. Remarks on the long time behaviour of the solutions of Hamilton-Jacobi equations. *Comm. Partial Differential Equations*, 24(5-6):883–893, 1999.

[Roq98a] Jean-Michel Roquejoffre. Comportement asymptotique des solutions d'équations de Hamilton-Jacobi monodimensionnelles. *C. R. Acad. Sci. Paris Sér. I Math.*, 326(12):185–189, 1998.

[Roq98b] Jean-Michel Roquejoffre. Convergence to steady states of periodic solutions in a class of Hamilton-Jacobi equations. preprint, 1998.

[Rud87] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York, third edition, 1987.

[Rud91] Walter Rudin. *Functional analysis*. McGraw-Hill Inc., New York, second edition, 1991.

[RV73] A. Wayne Roberts and Dale E. Varberg. *Convex functions*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1973. Pure and Applied Mathematics, Vol. 57.

[Smi83] Kennan T. Smith. *Primer of modern analysis*. Springer-Verlag, New York, second edition, 1983.

[WZ77] Richard L. Wheeden and Antoni Zygmund. *Measure and Integral*. Marcel Dekker, Inc., New York and Basel, 1977. Monographs and textbooks in Pure and Applied Mathematics, Vol. 43.

Index