UNIVERSIDADE TÉCNICA DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

Weak KAM and Aubry–Mather theories in an optimal switching setting

Diego Marcon Farias

Supervisor: Doctor Diogo Aguiar Gomes
Co–Supervisor: Doctor Alessio Figalli

Thesis approved in public session to obtain the PhD Degree in Mathematics

Jury final classification: Pass with Merit

Jury

Chairperson: Chairman of the IST Scientific Board
Members of the Committee:
Doctor Alessio Figalli
Doctor Carlos Alberto Varelas da Rocha
Doctor Diogo Luís de Castro Vasconcelos de Aguiar Gomes
Doctor Filippo Cagnetti
Doctor António Manuel Atalaia Carvalheiro Serra
Doctor Diogo Martins de Almeida de Araújo Pinheiro

2013

UNIVERSIDADE TÉCNICA DE LISBOA
INSTITUTO SUPERIOR TÉCNICO

INSTITUTO
SUPERIOR
TÉCNICO

Weak KAM and Aubry–Mather theories in an optimal switching setting

Diego Marcon Farias

Supervisor: Doctor Diogo Aguiar Gomes
Co–Supervisor: Doctor Alessio Figalli

Thesis approved in public session to obtain the PhD Degree in Mathematics

Jury final classification: Pass with Merit

Jury

Chairperson: Chairman of the IST Scientific Board
Members of the Committee:
Doctor Alessio Figalli, Professor Catedrático, University of Texas, USA;
Doctor Carlos Alberto Varelas da Rocha, Professor Catedrático do Instituto Superior Técnico, da Universidade Técnica de Lisboa;
Doctor Diogo Luís de Castro Vasconcelos de Aguiar Gomes, Professor Catedrático do Instituto Superior Técnico, da Universidade Técnica de Lisboa;
Doctor Filippo Cagnetti, Professor Auxiliar, University of Sussex, UK;
Doctor António Manuel Atalaia Carvalheiro Serra, Professor Auxiliar do Instituto Superior Técnico, da Universidade Técnica de Lisboa.
Doctor Diogo Martins de Almeida de Araújo Pinheiro, Investigador Auxiliar do Instituto Superior de Economia e Gestão, da Universidade Técnica de Lisboa.

Funding Institutions
Fundação para a Ciência e a Tecnologia, Portugal
2013

Weak KAM and Aubry–Mather theories in an optimal switching setting

Abstract

Dynamical systems defined by Tonelli Lagrangians have been the object of extensive study. In this direction, there is a deep connection between the classical calculus of variations problem and aspects of the weak KAM and Aubry–Mather theories, whence an extremely beautiful theory can be formulated. Such formulation is a combination of results that have been introduced by P.L. Lions, G. Papanicolaou, S.R.S. Varadhan, A. Fathi, J. Mather, R. Mañé, among others.

In this thesis, we extend a number of concepts of this known theory to the case where an optimal switching system is considered. Roughly speaking, an optimal switching problem consists of finding trajectories of a system whose dynamics can be conveniently modified by switching between different settings or "modes", in order to minimize an action functional.

We mainly consider two issues: the analysis of the calculus of variations problem and the study of a generalized weak KAM–type theorem for solutions of a weakly coupled systems of Hamilton–Jacobi equations. Our results include the existence and regularity of action minimizers as well as necessary conditions for minimality, and an extension of Fathi's weak KAM theorem. These can be applied to obtain a third result, namely, the long time behavior of solutions of the time–dependent system.

Keywords: Calculus of variations, dynamical systems, partial differential equations, optimal switching problems, weakly coupled systems, Hamilton–Jacobi equations, weak KAM theory, Aubry–Mather theory, quasivariational inequalities, viscosity solutions.

Teorias KAM fraca e de Aubry–Mather em um sistema de comutação ótimo

Diego Marcon Farias
Doutoramento em Matematica
Orientador: Doutor Diogo Gomes
Co-Orientador: Doutor Alessio Figalli

Resumo

Sistemas dinâmicos definidos a partir de Lagrangianos de Tonelli têm sido intensivamente estudados. Nesta direção, existe uma conexão profunda entre o problema clássico do cálculo de variações e aspectos das teorias KAM fraca e de Aubry–Mather, de onde uma teoria deslumbrante pode ser formulada. Tal formulação é uma combinação de resultados introduzidos por P.L. Lions, G. Papanicolaou, S.R.S. Varadhan, A. Fathi, J. Mather, R. Mañé, entre outros.

Nesta tese, estendemos alguns conceitos desta teoria já conhecida para o caso onde um sistema de comutação ótimo é considerado. Grosseiramente falando, um problema de comutação ótimo consiste em encontrar trajetórias ótimas em um sistema onde a dinâmica pode ser convenientemente modificada por comutação de estados ou modos, de tal maneira que a ação de certo funcional seja minimizada.

Consideramos principalmente duas questões: a análise do problema do cálculo de variações e o estudo de um teorema do tipo KAM fraco para soluções de um sistema fracamente acoplado de equações de Hamilton–Jacobi associado. Nossos resultados incluem a existência e a regularidade de minimizantes para a ação, condições necessárias para minimalidade e uma extensão do Teorema KAM fraco de Fathi. Estes são então aplicados para a obtenção de um terceiro resultado, a saber, o comportamento assimptótico de soluções do sistema com dependência do tempo.

Palavras-chave: Cálculo de variações, sistemas dinâmicos, equações diferenciais parciais, problema de comutação ótimo, sistemas fracamente acoplados, equações de Hamilton–Jacobi, teoria KAM fraca, teoria de Aubry–Mather, desigualdades quasivariacionais, soluções de viscosidade.

Acknowledgements

I thank my supervisors Alessio Figalli and Diogo Gomes for their encouragement, kindness, and precious help in the elaboration of this project. I feel deeply honored to have been working under the guidance of such brilliant mathematicians.

I would like as well to thank Artur Lopes, for making me aware of the UT Austin — Portugal program, and for the confidence he had in me.

To the Portuguese Science Foundation (FCT), which has supported me by a doctoral fellowship through the UT Austin-Portugal partnership, I am greatly thankful.

I thank all my friends from Lisbon and Austin. Especially to Levon Nurbekyan and Vardan Voskanyan, for helpful comments and discussions on topics of this thesis, but also to Emanuel Indrei and Rafayel Teymurazyan, for their help and engagement in different projects.

I also want to thank my family and friends back in Brazil, for their incentive and support. It has been a very difficult task to stay so far away, and for such an enlarged period of time.

Finally, I thank my fiancée Juliane Capaverde, for the inspiration, liveliness, and comfort she provided me with this whole time. She is the one that I cheerfully dedicate this work to.

Diego Marcon Farias

To my future wife,
Juliane.

Contents

Chapter 1

Introduction

1.1 Overview

Dynamical systems defined by Tonelli Lagrangians have been the object of extensive study. Of particular interest is the existing connection between the calculus of variations and both the weak KAM and Aubry–Mather theories.

In the calculus of variations, a Lagrangian $L : TM \to \mathbb{R}$ is a function defined on the vector bundle TM of a smooth manifold M, and its action functional $\mathbb{L}$, defined on the set of absolutely continuous curves $\gamma : [a, b] \to M$, $a \leq b$, is given by

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

In 1940, Leonida Tonelli [26] proved the existence of an absolutely continuous minimizer of the action $\mathbb{L}$, under fixed boundary conditions, that is, he proved the existence of $\gamma \in AC([a, b]; M)$[1] for which $\mathbb{L}(\gamma) \leq \mathbb{L}(\alpha)$, for all curves $\alpha \in AC([a, b]; M)$ with the same endpoints. Tonelli's method is currently known as the direct method in the calculus of variations. Regularity of minimizers and necessary conditions for minimality were already known:

1. Minimizers are not only absolutely continuous, but of class C^2. Furthermore,

[1]We denote by $AC([a, b]; M)$ the set of absolutely continuous curves $\alpha : [a, b] \to M$, see Chapter 2 below.

they satisfy the Euler–Lagrange equation

$$\frac{d}{dt}\left[\frac{\partial L}{\partial v}\big(\gamma(t),\dot\gamma(t)\big)\right]=\frac{\partial L}{\partial x}\big(\gamma(t),\dot\gamma(t)\big),\ \text{in } [a,b];$$

More generally, minimizers have as many derivatives as the Lagrangian.

2. The energy of the system is conserved along a minimizer, that is,

$$E\big(\gamma(t),\dot\gamma(t)\big)=\frac{\partial L}{\partial v}\big(\gamma(t),\dot\gamma(t)\big)\cdot\dot\gamma(t)-L\big(\gamma(t),\dot\gamma(t)\big)$$

is constant in t.

More recent are the weak KAM and the Aubry–Mather theories. The weak KAM theory deals with finding solutions $u : M \to \mathbb{R}$ of the stationary Hamilton–Jacobi equation

$$H\big(x,du(x)\big)=c,\ \forall x\in M,$$

where $H : T^*M \to \mathbb{R}$ is the Hamiltonian associated to the Lagrangian L by the Legendre–Fenchel duality

$$H(x,p):=\sup_{v\in T_xM}\big\{p(v)-L(x,v)\big\},$$

and $c \in \mathbb{R}$ a constant to be specified. In 1997, Albert Fathi [13] proved the following.

Theorem 1 (Weak KAM)**.** *There exists a unique $c_0 \in \mathbb{R}$ such that the Hamilton–Jacobi equation*

$$H\big(x,du_0(x)\big)=c_0,\ \forall x\in M, \tag{1.1}$$

admits a viscosity solution[2] $u_0 : M \to \mathbb{R}$.

Theorem 1 has been proved in a slightly different form by Pierre–Louis Lions, George Papanicolaou, and Srinivasa Varadhan [20], in the case $M = \mathbb{T}^d$, using a different method (namely, they used discounted infinite horizon costs together with a stability result for viscosity solutions of Hamilton–Jacobi equations). In the

[2] See Definition 17, in Chapter 2, for the definition of viscosity solutions of (1.1).

general case, Fathi obtains the solution u_0 by looking at the long–time behavior of the Lax–Oleinik semigroup

$$T_t u(x) := \inf \Big\{ u(\gamma(0)) + \int_0^t \big(L(\gamma,\dot\gamma) + c_0\big)\, ds\ ; \qquad \gamma \in AC([0,t];M), \gamma(t) = x \Big\}, \tag{1.2}$$

where $c_0 \in \mathbb{R}$ is the smallest constant for which (1.1) admits a subsolution. More precisely, he proves that, for a particular function u, the semigroup $T_t u$ converges, as $t \to +\infty$, to a limit u_0 that is a solution of (1.1).

On the other hand, the Aubry–Mather theory attempts to understand properties of solutions and subsolutions through the analysis of invariant objects, like the Aubry set, the Mather set, and Mather measures.

Let us set $h_t : M \times M \to \mathbb{R}$,

$$h_t(x,y) = \inf_{\gamma(0)=x,\ \gamma(t)=y} \int_0^t L\big(\gamma(s),\gamma(s)\big)\, ds.$$

Then, the projected Aubry set $\mathcal{A}$, first introduced by John Mather [24] in 1991, is the set

$$\mathcal{A} = \Big\{ x \in M;\ \liminf_{t\geq 0} \big\{ h_t(x,x) + c_0 t \big\} = 0 \Big\}.$$

The set $\mathcal{A}$ is nonempty, and satisfies many useful properties. We list some of them in the following theorem.

Theorem 2. *The projected Aubry set $\mathcal{A}$ is nonempty, and it satisfies:*

1. (Fathi [15]) *If two solutions of* (1.1) *coincide on the Aubry set $\mathcal{A}$, then they coincide everywhere on M.*

2. (Fathi-Siconolfi [17]) *Every subsolution $u : M \to \mathbb{R}$ of* (1.1) *is differentiable on the projected Aubry set $\mathcal{A}$. Moreover, there exists a C^1 subsolution v of* (1.1) *that coincides with u on $\mathcal{A}$ and it is a strict subsolution on $M\backslash\mathcal{A}$.*

3. (Bernard [3]) *Given a subsolution u of* (1.1)*, there exists a subsolution $v \in C^{1,1}$ that coincides with u on $\mathcal{A}$, and it is strict outside $\mathcal{A}$.*

4. (Mather [23]) *If $x \in \mathcal{A}$ and u is a subsolution of* (1.1)*, then the differential $du(x)$ is independent of u. Moreover, the map*

$$x \in \mathcal{A} \mapsto du(x) \in T_x^* M$$

is Lipschitz.

It is classical that if a viscosity subsolutions of (1.1) is differentiable at a point $x \in M$, then $H(x, du(x)) \le c_0$. Therefore, when we write that v is a strict subsolution on $M \backslash \mathcal{A}$ above, we simply mean the pointwise relation

$$H\big(x, dv(x)\big) < c_0, \ \forall x \in M \backslash \mathcal{A}.$$

To define the Mather set, we first need a couple of definitions (see also Chapter 2, §2.7). A holonomic probability measure is a probability measure $\mu \in \mathcal{P}(TM)$ such that

$$\int_{TM} d\phi(x) \cdot v \ d\mu(x, v) = 0, \ \forall \ \phi \in C^1(M).$$

We write $\mu \in \mathcal{H}$. Following Mather [23] and Mañé [22], we wish to minimize the relaxed action functional

$$\mu \mapsto \mathbb{L}[\mu] := \int_{TM} L(x, v) \ d\mu(x, v)$$

among all holonomic probability measures $\mu \in \mathcal{H}$. It is possible to show that

$$\inf_{\mu \in \mathcal{H}} \mathbb{L}[\mu] = -c_0$$

and that minimizing measures exist, that is, that the infimum above is in fact a minimum. The Mather set $\tilde{\mathcal{M}}$ is then defined as

$$\tilde{\mathcal{M}} := \overline{\bigcup_{\mathbb{L}[\mu] = -c_0} \operatorname{supp} \mu} \subset TM. \tag{1.3}$$

It can be shown that the projected Mather set $\mathcal{M} := \pi(\tilde{\mathcal{M}})$ is contained in the projected Aubry set $\mathcal{A}$ defined above, and an analogous of Theorem 2 holds true on $\mathcal{M}$

By taking advantage of these properties of the Mather set, a general asymptotic result for the Lax–Oleinik semigroup can be proven. For example, Fathi [14] proved the following:

Theorem 3. *Given $u_0 \in C(M)$, the Lax–Oleinik semigroup $T_t u_0$ converges, as $t \to \infty$ to a critical solution $v \in C(M)$ of*

$$H\big(x, du(x)\big) = c_0.$$

This theorem also describes the asymptotic behavior of solutions to the dynamical Hamilton–Jacobi equation. In fact, just observe that given $u_0 \in C(M)$, the Lax–Oleinik semigroup $v(t, x) := T_t u_0(x)$ is a viscosity solution of

$$\partial_t v(t, x) + H\big(x, \partial_x v(t, x)\big) = 0, \text{ on } \mathbb{R}^+ \times M. \tag{1.4}$$

1.2 Main results

In this thesis, we extend the previous results to the case where an optimal switching setting is considered. An optimal switching problem consists of finding trajectories of a system whose dynamics can be conveniently modified by switching between a number of different settings or "modes". Switching from one mode to another is always allowed; however, at every switch, a positive switching cost is incurred. In this way, we are led to minimize a generalized action of the form

$$\mathcal{J}[\gamma, \sigma] := \int_0^t L(\gamma, \dot{\gamma}, \sigma)\ ds + \sum \psi(\sigma^-, \sigma^+), \tag{1.5}$$

where $L : TM \times \mathcal{I} \to \mathbb{R}$ is such that $L(\cdot, \cdot, i) : TM \to \mathbb{R}$ is a Tonelli Lagrangian, $\sigma : [0, t] \to \mathcal{I}$ determines the modes at every instant of $[0, t]$, and $\psi : \mathcal{I} \times \mathcal{I} \to \mathbb{R}$ is a positive switching cost (for more precise definitions, see Chapter 3 below).

The interest in optimal switching problems and its relation with viscosity solutions comes back to 1984, when Italo Capuzzo–Dolcetta and Lawrence Evans [6] extend the notion of viscosity solution to these systems and prove that the value functions are in fact viscosity solutions of a weakly coupled system of Hamilton–Jacobi equations. Apparently, the problem was motivated by a variant of a stochastic problem considered by Lawrence Evans and Avner Friedman [12], where the solution of a Bellman equation is found as the limit of solutions of certain systems of nonlinear equations. Capuzzo–Dolcetta and Evans were motivated by earlier works of S.A. Belbas, and I. Capuzzo Dolcetta, M. Matzeu, J.L. Menaldi (see, for instance [7, 2]).

We consider the following problems:

1. The calculus of variations problem associated to minimizing the action functional $\mathcal{J}$: We study the existence of minimizers of the action as well as generalized necessary conditions for minimality, and a conservation of energy principle.

2. Weak KAM–type theorem for solutions of the system: Is it possible to obtain a version of Fathi's weak KAM theorem for this case?

3. A generalized Aubry–Mather theory associated to the system of Hamilton–Jacobi equations: We are interested in the extension of the concepts of the Aubry set, regular subsolutions, holonomic and minimizing measures, and, of course, its relation with some sort of critical value.

We spend the rest of this section stating our main results. We prove both the existence and the regularity of minimizers of the action of $\mathcal{J}$. Although the optimal switching problem has been studied for about 30 years, and the existence of minimizers for such problems so naturally arises, we were unable to find such result in the literature. In our method, the only extra assumption we require in the m Tonelli Lagrangians is a uniform superlinearity[3]. We also obtain regularity in space for minima. We prove:

Theorem 4. *Assume $\mathcal{I} = \{1, \dots, m\}$ is finite. Then, for every $x, y \in M$, we have existence of minimizers for the action: there exists $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) \in AC \times \mathcal{P}$ with $\gamma_M(0) = x$, $\gamma_M(t) = y$ such that*

$$\mathcal{J}_t[\gamma] = \inf \left\{ \mathcal{J}_t[\alpha] \ ; \ \alpha_M(0) = x, \alpha_M(t) = y \right\}.$$

Moreover,

[3]See assumptions $A1$–$A5$ in Chapter 3

1. *Minimizers are of class $C^2(M)$, meaning $\gamma_M \in C^2(M)$, and solve the Euler–La-grange equation*

$$\frac{d}{ds}\left[\frac{\partial L}{\partial v}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big)\right] = \frac{\partial L}{\partial x}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) \tag{1.6}$$

in $[0,t]\backslash\{t_1, \dots, t_N\}$, where t_k's are the discontinuity points of $\gamma_{\mathcal{I}}$;

2. *Along a minimizing curve, the (generalized) energy functional is conserved:*

$$\begin{aligned} E\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) := \frac{\partial L}{\partial v}\big(\gamma_M(s),\ \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) \cdot \dot\gamma_M(s) \\ - L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) \end{aligned} \tag{1.7}$$

is constant in $[0,t]$.

Next, we describe properties of the associated value functions. These value functions solve, in the viscosity sense, a system of Hamilton–Jacobi equations of the form [6, 19]

$$\max_{x\in M}\left\{H_i\big(x, du_i(x)\big), \max_{j\neq i}\left\{u_i(x) - u_j(x) - \psi(x,i,j)\right\}\right\} = 0,\ \ \forall\ i \in \mathcal{I}. \tag{1.8}$$

One of our main results is a weak KAM–type theorem, analogue to the one in Fathi [13], to the optimal switching setting. We have:

Theorem 5. *There exists a unique constant $c_0 \in \mathbb{R}$ satisfying:*

(i) *The equation*

$$\max_{x\in M}\left\{H_i\big(x, du_i(x)\big) - c_0, \max_{j\neq i}\left\{u_i(x) - u_j(x) - \psi(x,i,j)\right\}\right\} = 0,\ \ \forall\ i \in \mathcal{I}, \tag{1.9}$$

admits a viscosity solution $u : M \times \mathcal{I} \to \mathbb{R}$;

(ii) *For any $c \in \mathbb{R}$ for which (1.9) admits a subsolution with c_0 replaced by c, we have $c \geq c_0$.*

We prove Theorem 5 by extending Fathi's idea of looking at the long time behavior of a generalized Lax–Oleinik semigroup we define[4]. Additionally, we

[4]We learned that Lax–Oleinik operators for similar problems have been considered before [1].

obtain a regularity result for solutions of the system, in the same spirit of Bernard's Theorem [3]. For this matter, we need a notion of a projected Aubry set in the optimal switching setting, which we formulate essentially in the same way it has already been defined by Gomes-Serra [19].

We define the projected Aubry set similarly to [19], and we prove he following theorem:

Theorem 6. *The Aubry set $\mathcal{A} \subset M \times \mathcal{I}$ satisfies the following properties:*

1. *For any $A = (x, i) \in \mathcal{A}$, there exists a curve $\gamma : \mathbb{R} \to \mathcal{A}$ with $\gamma(0) = A$ such that, for any subsolution $u : M \times \mathcal{I} \to \mathbb{R}$ of (1.9), and all $t_1 < t_2$,*

$$u\big(\gamma(t_2)\big) - u\big(\gamma(t_1)\big) = \int_{t_1}^{t_2} L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\big)\ ds \\ + \sum_k \psi\big(\gamma_M(s_{k+1}), \gamma_{\mathcal{I}}(s_k), \gamma_{\mathcal{I}}(s_{k+1})\big),$$

   where the sum above is taken for all k such that $t_1 < s_k < t_2$.

2. *Suppose u is a subsolution and w is a supersolution of (1.9), and that $u \leq w$ on $\mathcal{A}$. Then $u \leq w$ in $M \times \mathcal{I}$.*

1.3 Structure of the thesis

In Chapter 2, we describe carefully the classical calculus of variations theory in dimension 1, including the existence of minimizers, the Euler–Lagrange equations, and the conservation of energy principle. We also describe the connection that the Legendre–Fenchel transform provides between the value function and the viscosity solutions of Hamilton–Jacobi equations, including Fathi's weak KAM theorem. Then, we define elements of the Aubry–Mather theory and show how to obtain subsolutions that are regular in the Aubry set. No results in this chapter are new.

Next, in Chapter 3, we proceed to the description of the optimal switching problem. We prove the existence and regularity of minimizers for a class of "uniformly" Tonelli Lagrangians. A first property of the cost function is also proven, namely, the local semiconcavity.

In Chapter 4, we define the Lax–Oleinik semigroup associate to the system (1.9), and the generalized Weak KAM Theorem. At the end of the chapter, we define the Aubry set and prove Theorem 6.

In Chapter 5 we obtain asymptotic limits of solutions to the dynamical Hamilton–Jacobi system.

Finally, in Chapter 6, we discuss further topics we will present in a future work.

At the end of this thesis, we also have two appendices. In Appendix A we state and prove Tonelli's Existence Theorem of the calculus of variations. In Appendix B we provide a short introduction on semiconvex and semiconcave functions.

Chapter 2

Classical Theory

The purpose of this chapter is to recall the concepts that, in a way or another, will be studied in the later chapters for switching problems. For the convenience of the reader, we make the presentation as self–contained as possible; in the few places where no proof is given, we provide appropriate references.

To begin with, we define Tonelli Lagrangians and its associated action functional in Section 2.1. Next, in Section 2.2, we state Tonelli's Existence Theorem of action minimizers (a proof is postponed to Appendix A), and we prove that they satisfy the Euler–Lagrange equations, that in turn gives us regularity of minimizers. Then, in Section 2.3, we use the Legendre–Fenchel duality to define the Hamiltonian, and we prove the conservation of energy principle. In Section 2.4 we present Fathi's Weak KAM Theorem, and explain the ideas behind Fathi's proof. The existence of $C^{1,1}$ subsolutions is proved in Section 2.5. Next, we proceed to define the main elements of the Aubry–Mather theory in Sections 2.6 and 2.7. Finally, in Section 2.8, we show that weak KAM solutions can be obtained as the asymptotic limit of Lax–Oleinik solutions of time–dependent Hamilton–Jacobi equations.

2.1 Definitions and examples

Throughout this thesis, M denotes a complete differentiable manifold and TM its tangent bundle. We call $L : TM \to \mathbb{R}$ a *Lagrangian* and

$$\mathcal{J}[\gamma] = \mathcal{J}_{a,b}[\gamma] := \int_a^b L\big(\gamma(s), \dot{\gamma}(s)\big)\ ds$$

its *action* over an absolutely continuous curve $\gamma : [a, b] \to M$. We assume L is bounded below and continuous, so that the action is well-defined with values in $\mathbb{R} \cup \{+\infty\}$.

By *action minimizer* we understand a curve $\gamma : [a, b] \to M$ satisfying $\mathcal{J}_{a,b}[\gamma] \leq \mathcal{J}_{a,b}[\alpha]$, for any absolutely continuous curve $\alpha : [a, b] \to M$ with the same end-points. As we have already mentioned in the introduction, Tonelli's theory [26] proves to be very efficient in finding an action minimizer, whence it is fair that the following definition carries his name.

Definition 7. (Tonelli Lagrangian) We say that $L : TM \to \mathbb{R}$ is a *Tonelli Lagrangian* if it satisfies the following four conditions:

1. L is of class $C^2(TM)$;
2. The second order derivative $\frac{\partial^2 L}{\partial v^2} L(x, v)$ is positive definite, for every $(x, v) \in TM$;
3. There exist g, a complete Riemannian metric, and a constant $C \geq 0$ such that
$$\|v\|_x - C \leq L(x, v),$$
for any $v \in T_x M$, where $\|v\|_x := \sqrt{g_x(v, v)}$, for $(x, v) \in TM$;
4. *Superlinear in the fibers above compact subsets*: For every compact $K \subset M$, and for every constant $A \geq 0$, there exists a constant $C \in \mathbb{R}$, depending on K and on A, such that
$$L(x, v) \geq A \|v\|_x + C, \text{ for all } x \in K,\ v \in T_x M$$

Next, we provide a few examples of Tonelli Lagrangians

Example 8. Given a complete smooth Riemannian metric g on M, we define the *quadratic* (or *Riemannian*) Lagrangian as

$$L(x, v) = \frac{1}{2} \|v\|_x^2 .$$

Example 9. If a metric is given on M, as in Example 8, and a potential $U \in C^2$, $U \geq -C$, is added to the quadratic Lagrangian, we again obtain a Tonelli Lagrangian, that is known as the *mechanical* Lagrangian

$$L(x, v) = \frac{1}{2} \|v\|_x^2 + U(x).$$

Example 10. Another example was introduced by Ricardo Mañé [21]: Given a C^2 vector field X, set

$$L(x, v) = \|v - \mathbf{X}(x)\|_x^2 .$$

2.2 Existence and regularity of minimizers

In this section we prove the existence of minimizers for the action of a Tonelli Lagrangian.

Theorem 11 (Existence of action minimizers)**.** *Assume $L : TM \to \mathbb{R}$ is a Tonelli Lagrangian. Then, there exists a minimizer of the action $\mathcal{J}_t$ under fixed endpoints condition. More precisely, given $a < b$ and $x, y \in M$, there exists $\gamma \in AC([a, b]; M)$, with $\gamma(a) = x$, $\gamma(b) = y$ satisfying*

$$\mathcal{J}_t[\gamma] \leq \mathcal{J}_t[\alpha],$$

for all $\alpha \in AC([a, b]; M)$ with $\alpha(a) = x$, $\alpha(b) = y$.

Proof. We provide a proof of this theorem in Appendix A. The proof follows the direct method in the calculus of variations. See also [4, 15]. □

Remark 12. We observe that L being a Tonelli Lagrangian is more than necessary in order to prove the existence of minimizers, like in Theorem 11. Indeed, L does not even need to be C^1. For instance, if we assume the following:

(*i*) L and $\partial L/\partial v$ continuous;

(*ii*) $L(x,\cdot)$ convex, for each fixed x;

(*iii*) L has superlinear growth,

then a minimizer exists. For a proof of this result, see [4, Theorem 3.7].

Next, we prove that minimizers of the action satisfy the Euler–Lagrange equations. The proof is standard and can be found in different textbooks (see, for instance, [4, 15]).

Theorem 13 (Euler–Lagrange equations and regularity)**.** *If $L : TM \to \mathbb{R}$ is a Tonelli Lagrangian, and $\gamma : [a,b] \to M$ is an action minimizer, then γ is C^2 and it satisfies (in coordinates) the* Euler–Lagrange equation

$$\frac{d}{dt}\left[\frac{\partial L}{\partial v}\big(\gamma(t),\dot\gamma(t)\big)\right] = \frac{\partial L}{\partial x}\big(\gamma(t),\dot\gamma(t)\big), \tag{2.1}$$

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

Proof. Assume γ is minimizing. Then, it is Lipschitz (see [8]). Let $h : [a,b] \to M$ be C^1, with $h(a) = 0$, $h(b) = 0$. Then, since γ is minimizing

$$0 \le \frac{\mathcal{J}[\gamma+\varepsilon h] - \mathcal{J}[\gamma]}{\varepsilon} = \int_a^b \frac{L(\gamma+\varepsilon h, \dot\gamma + \varepsilon \dot h) - L(\gamma,\dot\gamma)}{\varepsilon} =: \int_a^b f(\varepsilon, s) \tag{2.2}$$

Since

$$f(\varepsilon,s) \to \frac{\partial L}{\partial x}(\gamma,\dot\gamma)\cdot h + \frac{\partial L}{\partial v}(\gamma,\dot\gamma)\cdot \dot h, \text{ a.e. as } \varepsilon \to 0$$

and

$$|f(\varepsilon,s)| \le K|h| + |\dot h| \le C,$$

the dominated convergence theorem implies

$$0 \le \int_a^b \frac{\partial L}{\partial x}(\gamma,\dot\gamma)\cdot h + \frac{\partial L}{\partial v}(\gamma,\dot\gamma)\cdot \dot h,$$

for all h. In particular,

$$\begin{aligned} 0 &= \int_a^b \frac{\partial L}{\partial x}(\gamma,\dot\gamma)\cdot h + \frac{\partial L}{\partial v}(\gamma,\dot\gamma)\cdot \dot h \\ &= \int_a^b \left(\frac{\partial L}{\partial x}(\gamma,\dot\gamma) - \frac{d}{ds}\frac{\partial L}{\partial v}(\gamma,\dot\gamma)\right) h \end{aligned} \tag{2.3}$$

Since h is arbitrary, the result follows. □

2.3 Hamiltonian and Conservation of energy

Associated to the Lagrangian $L : TM \to \mathbb{R}$ through the Legendre–Fenchel transform is the *Hamiltonian* $H : T^*M \to \mathbb{R}$,

$$H(x,p) := \sup_{v\in T_xM} \big\{p(v) - L(x,v)\big\}. \tag{2.4}$$

Definition 14. We say that $H : T^*M \to \mathbb{R}$ is a *Tonelli Hamiltonian* if the following conditions are satisfied:

1. H is $C^2(TM)$;

2. The second derivative along the fibers is positive definite: For every $(x,p) \in T^*M$,
$$\frac{\partial H}{\partial p}(x,p) > 0;$$

3. There exist g, a complete Riemannian metric, and a constant $C \geq 0$ such that
$$\|p\|_x^* - C \leq H(x,p),$$
for any $p \in T_x^*M$, where $\|p\|_x^*$ denotes a norm on T^*M;

4. Superlinear in the fibers above compact subsets: For every compact $K \subset M$, and for every constant $A \geq 0$, there exists a constant $C \in \mathbb{R}$, depending on K and on A, such that
$$H(x,p) \geq A\,\|p\|_x^* + C, \text{ for all } x \in K,\ p \in T_x^*M.$$

It is standard to verify that H, given by (2.4) is a Tonelli Hamiltonian (see, for instance [8]). Conversely, if one is given a Tonelli Hamiltonian, in the sense of Definition 14 above, we can recover the Lagrangian $L : TM \to \mathbb{R}$ by the reverse Legendre–Fenchel transform

$$L(x,v) := \sup_{p \in T_x^* M} \big\{ p(v) - H(x,p) \big\},$$

and prove that it is a Tonelli Lagrangian.

Definition 15. The *energy* functional on TM is defined as $E : TM \to \mathbb{R}$,

$$E(x,v) = H\left(x, \frac{\partial L}{\partial x}(x,v)\right) = \frac{\partial L}{\partial v}(x,v) \cdot v - L(x,v). \tag{2.5}$$

The conservation of energy principle states that the energy of the system is conserved along trajectories, and it is the content of the following theorem:

Theorem 16. *The energy is conserved along trajectories of the system, that is,*

$$E\big(\gamma(t), \dot\gamma(t)\big) = \frac{\partial L}{\partial v}\big(\gamma(t), \dot\gamma(t)\big) \cdot \dot\gamma(t) - L\big(\gamma(t), \dot\gamma(t)\big) \tag{2.6}$$

is constant in $[a,b]$*, whenever* $\gamma \in AC([a,b]; M)$ *is an action minimizer.*

Proof. We only sketch the proof. Define the adjoint variables $s \mapsto \big(x(s), p(s)\big) \in T^*M$ by

$$\big(x(s), p(s)\big) = \left(\gamma(s), \frac{\partial L}{\partial v}\big(\gamma(s), \dot\gamma(s)\big)\right). \tag{2.7}$$

Observe that the Euler–Lagrange equation implies $p \in C^1$. By the analysis of the Legendre–Fenchel duality, it is not hard to see that (x,p) is a solution the *Hamiltonian system*:

$$\begin{cases} \dot x(s) = \dfrac{\partial H}{\partial p}\big(x(s), p(s)\big) \\ \dot p(s) = -\dfrac{\partial H}{\partial x}\big(x(s), p(s)\big), \end{cases} \tag{2.8}$$

when $s \mapsto \big(\gamma(s), \dot\gamma(s)\big)$ is a minimizing trajectory. Then

$$\begin{aligned} \frac{d}{ds}E\big(\gamma(s),\dot\gamma(s)\big) &= \frac{d}{ds}H\big(x(s),p(s)\big) \\ &= \frac{\partial H}{\partial x}\big(x(s),p(s)\big)\cdot \dot x(s) + \frac{\partial H}{\partial p}\big(x(s),p(s)\big)\cdot \dot p(s) = 0, \end{aligned} \tag{2.9}$$

and the result follows. □

2.4 Fathi's weak KAM theorem

As mentioned in the introduction, the weak KAM theory is concerned with the existence of solutions to the stationary Hamilton–Jacobi equation

$$H\big(x, du(x)\big) = c, \tag{2.10}$$

for some $c \in \mathbb{R}$. To begin with, we must clarify what notion of solutions we consider. We start with the definition of a viscosity solution:

Definition 17. 1. We say that the continuous function $u : M \to \mathbb{R}$ is a *viscosity subsolution* of (2.10) if for all $x \in M$, and for any C^1 function ϕ such that $u - \phi$ has a maximum at the point x, we have

$$H\big(x, d\phi(x)\big) \leq c. \tag{2.11}$$

2. We say that the continuous function $u : M \to \mathbb{R}$ is a *viscosity supersolution* of (2.10) if for all $x \in M$, and for any C^1 function ϕ such that $u - \phi$ has a minimum at the point x, we have

$$H\big(x, d\phi(x)\big) \geq c. \tag{2.12}$$

3. Finally, we say $u : M \to \mathbb{R}$ is a *viscosity solution* of (2.10) if it is both a viscosity subsolution and a viscosity supersolution.

In this context, the notion of a viscosity solution has different equivalent definitions. In order to present these, we recall the following definition:

Definition 18. For a given function $u : M \to \mathbb{R}$, the *Lax–Oleinik semigroup* $T_t u : M \to \mathbb{R}$ is defined by

$$T_t u(x) := \inf \left\{ u(\gamma(0)) + \int_0^t L\big(\gamma(s), \dot{\gamma}(s)\big)\ ds \right\}, \tag{2.13}$$

where the infimum is taken over all absolutely continuous curves $\gamma \in AC([0,t]; M)$ with $\gamma(t) = x$.

Our main interest in this semigroup is the variational approach that it provides for solutions of (2.10). For the definition of a semiconcave function, see Appendix B.

Proposition 19. *Let $H : T^*M \to \mathbb{R}$ be a Tonelli Hamiltonian. Given a semiconcave function $u : M \to \mathbb{R}$, the following conditions are equivalent:*

(i) u is a viscosity solution of (2.10);

(ii) $H\big(x, du(x)\big) = c$ a.e. in M;

(iii) $u = T_t u + ct$, for every $t \geq 0$.

Proof. See Fathi's book [15]. □

In fact, we have proven:

Proposition 20. *Given a semiconcave function $u : M \to \mathbb{R}$, the following conditions are equivalent:*

(i) u is a viscosity subsolution of (2.10);

(ii) $H\big(x, du(x)\big) \leq c$ a.e. in M;

(iii) $u \leq T_t u + ct$, for every $t \geq 0$.

A similar statement holds for supersolutions.

The previous propositions provide useful ways of searching for solutions, subsolutions, or supersolutions. However, when in search of a solution, in general it is not possible to find a fixed point for the Lax–Oleinik semigroup. In fact, we must be careful with the choice of the constant we consider. The following theorem holds (Fathi [13]; see also Lions, Papanicolaou, Varadhan [20]):

Theorem 21 (Weak KAM)**.** *There exists a unique constant $c_0 \in \mathbb{R}$ for which*

$$H\big(x, du(x)\big) = c_0 \tag{2.14}$$

admits a viscosity solution $u : M \to \mathbb{R}$.

Proof. We only sketch the proof of this theorem, by Fathi's method. Define $c_0 \in \mathbb{R}$ to be the smallest constant $c \in \mathbb{R}$ for which $H\big(x, du_c(x)\big) = c$ admits a subsolution $u_c : M \to \mathbb{R}$. It is simple to verify that this infimum is attained for some function $u : M \to \mathbb{R}$. Since u is a subsolution, Proposition 20 implies $u \leq T_t u + c_0 t$, for all $t \geq 0$. For simplicity, we consider a "normalized" version of the Lax–Oleinik semigroup

$$T_t u(x) = \inf \left\{ u(\gamma(0)) + \int_0^t \Big(L\big(\gamma(s), \dot{\gamma}(s)\big) + c_0 \Big)\, ds \right\}, \tag{2.15}$$

so that

$$u \leq T_t u,$$

for all $t \geq 0$. Now, by using the compactness of M, one proves that this value c_0 is the only one for which we can find a viscosity solution of $H\big(x, du_c(x)\big) = c_0$. Then, by the choice of c_0, we prove that (see our analogue for switching problems: Proposition 77, in Chapter 3)

$$\sup_{t \geq 0} \|T_t u\|_{L^\infty(M)} < +\infty. \tag{2.16}$$

Since the Lax–Oleinik semigroup is monotone, and u is a subsolution, we have

$$T_s u \leq T_t u, \ \text{ for every } 0 \leq s \leq t. \tag{2.17}$$

Thus, (2.16) and (2.17) imply that the pointwise limit

$$u^\infty(x) := \lim_{t \to \infty} T_t u(x) \tag{2.18}$$

is well defined. Moreover, it is easily verified that u^∞ is a fixed point of the

Lax–Oleinik semigroup:

$$T_t u^\infty(x) = T_t(\lim_{s\to\infty} T_s u(x)) = \lim_{s\to\infty} T_{s+t} u(x) = u^\infty(x).$$

□

2.5 Existence of $C^{1,1}$ critical subsolutions

In this section, we present Bernard's proof [3] that the set of $C^{1,1}$ viscosity subsolutions of (2.10) is a dense subset of the set of viscosity subsolutions. More precisely, we prove:

Theorem 22 (Bernard)**.** *Let $H : T^*M \to \mathbb{R}$ be a Tonelli Hamiltonian, as in Definition 14. If the Hamilton–Jacobi equation (2.10) admits a subsolution, then it admits a $C^{1,1}$ subsolution. Furthermore, the set of $C^{1,1}$ viscosity subsolutions of (2.10) is a dense subset of the set of viscosity subsolutions, with the uniform topology.*

The idea of Bernard's proof goes as follows: as before, we define

$$T_t u(x) = \inf_{y\in M} \big\{u(y) + h_t(y,x)\big\}, \tag{2.19}$$

and we consider as well the "backwards" Lax–Oleinik semigroup

$$\breve{T}_t u(x) = \sup_{y\in M} \big\{u(y) - h_t(x,y)\big\}. \tag{2.20}$$

In this section, we consider the normalized function

$$h_t(y,x) = \inf_\gamma \int_0^t \big[L(\gamma,\dot\gamma) + c_0\big].$$

We know that $T_t u$ is semiconcave. Analogously, we can show that $\breve{T}_t u$ is semiconvex. Bernard shows that, if $s > 0$ is sufficiently small, then $T_s(\breve{T}_t u)$ is 'still' semiconvex. Since it is always a semiconcave function, it has to be $C^{1,1}$.

We start with a lemma:

Lemma 23. *If $\mathcal{F}$ is is such that*

$$\mathcal{F} \subseteq \{f \in C^2 \ ; \ \|D^2 f\|_\infty \leq C\}.$$

Then, so is $T_s(\mathcal{F})$, for all sufficiently small $s > 0$. Moreover, for $f \in \mathcal{F}$, and $\gamma : [0,s] \to M$, $\gamma(t) := \pi \circ \phi_t^H(x, d_x f)$, we have

$$T_s f(x) = f(\gamma(0)) + \int_0^s \big(L(\gamma(t), \dot\gamma(t)) + c_0\big)\, dt. \tag{2.21}$$

Proof. This proof makes use the Hamiltonian flow ϕ_t^H. Since we never defined it, we only sketch the proof. Fix $f \in \mathcal{F}$. Since $f \in C^2$, $\big\{\big(x, df(x)\big)\big\}$ is a C^1 graph, and, for sufficiently small $s > 0$, we have that

$$\phi_s^H\big(x, df(x)\big) = \big(y, d(T_s f)(y)\big)$$

is still a C^1 graph, where $y = \gamma(s)$ and $\gamma : [0,s] \to M$ minimizing with $\gamma(0) = x$. Then $T_s f \in C^2$, for all sufficiently small $s > 0$. That the Hessians are uniformly bounded follows from the uniform bound of $f \in \mathcal{F}$. ☐

Proof of Theorem 22. Denote $v(x) = \breve{T}_t u(x)$. By semiconvexity, there exist $C \in \mathbb{R}$ and

$$\mathcal{F} \subseteq \Big\{f \in C^2(M, \mathbb{R}) \ ; \ \|D^2 f\|_\infty \leq C\Big\}$$

such that

$$v(x) = \sup_{f \in \mathcal{F}} f(x), \ \forall\, x \in M. \tag{2.22}$$

Also, for every $p \in \partial^- v(x)$, there exists $f \in F$ such that $v(x) = f(x)$ and $p = df(x)$. Since T_t preserves order, for every $f \in F$, we have $T_t v \geq T_t f$; then

$$T_t v(x) \geq \sup_{f \in F} T_t f(x).$$

Now, for fixed $x \in M$, consider $\gamma : [0,s] \to M$ minimizing curve, with $\gamma(s) = x$:

$$T_s v(x) = v\big(\gamma(0)\big) + \int_0^s \big[L\big(\gamma(t), \dot\gamma(t)\big) + c_0\big]\, dt. \tag{2.23}$$

By minimality, zero is in the subdifferential at $\gamma(0)$ of the function

$$z \mapsto v(z) + h_t(z, x).$$

If we set

$$p := \frac{\partial L}{\partial v}\big(\gamma(0), \dot\gamma(0)\big),$$

we know $-p \in \partial_1^+ h_t\big(\gamma(0), x\big)$; thus,

$$p \in \partial^- v\big(\gamma(0)\big).$$

Then, choose $f \in \mathcal{F}$ such that $v\big(\gamma(0)\big) = f\big(\gamma(0)\big)$ and $p = df\big(\gamma(0)\big)$. We must have

$$T_s f(x) = f\big(\gamma(0)\big) + \int_0^s \big[L\big(\gamma(t), \dot\gamma(t)\big) + c_0\big]\, dt. \tag{2.24}$$

This, together with (2.23), shows that $T_s v$ is semiconvex. Therefore, u is $C^{1,1}$. $\square$

2.6 Aubry set

As mentioned in the introduction, the behavior of subsolutions at each point of its domain $x \in M$ depends heavily on whether x is an element of a certain set, the Aubry set, that we define in this section and study its main properties. For more on the Aubry set, we refer the reader to the lecture notes [25, 18], the book [15], and the paper [3].

Recall that we consider the 'normalized' cost function

$$h_t(x, y) = \inf\left\{ \int_0^t \Big(L\big(\gamma(s), \dot\gamma(s)\big) + c_0\Big)\, ds \ ;\ \gamma \in AC([0,t]; M) \right.$$
$$\left. \text{with } \gamma(0) = x, \gamma(t) = y \right\}.$$

Definition 24. The *(projected) Aubry set* $\mathcal{A} \subset M$ is defined as the set of points $x \in M$ for which

$$\limsup_{t \to +\infty} h_t(x, x) = 0.$$

Our first claim is to show that every subsolution u is differentiable in the projected Aubry set; moreover its differential $du(x)$ is independent of u.

Proposition 25 (Existence of critical curves)**.** *Given $y \in \mathcal{A}$, there exists a curve $\gamma_y : \mathbb{R} \to \mathcal{A}$ such that for any critical subsolution $u : M \to \mathbb{R}$,*

$$u(\gamma(b)) - u(\gamma(a)) = \int_a^b \left[L\big(\gamma(s), \dot{\gamma}(s)\big) + c_0\right] \, ds, \forall \; a \leq b. \tag{2.25}$$

Proof. We postpone the proof of this proposition until Chapter 4.3, where we prove a more general statement. The idea is to consider curves $\gamma^k : [0, t_k] \to M$, with $\gamma^k(0) = y, \gamma^k(t) = x$ for which

$$\int_0^{t_k} L(\gamma^k, \dot{\gamma}^k) \to 0.$$

Such curves exist by the definition of the projected Aubry set. Using the assumptions on L to show that γ^k must converge to a curve γ_y that satisfies the desired properties. □

Corollary 26. *If $x \in \mathcal{A}$, then every critical subsolution $u : M \to \mathbb{R}$ is differentiable at x,*

$$du(x) = \frac{\partial L}{\partial v}\big(x, \dot{\gamma}(0)\big), \tag{2.26}$$

and

$$H\big(x, du(x)\big) = c_0. \tag{2.27}$$

Proof. By working locally on charts, we may assume $M \subseteq \mathbb{R}^d$. Let $\gamma_x : [-\delta, \delta] \to \mathbb{R}^d$ be the critical curve given by Proposition 25, and set, for y in a neighborhood of x, $\alpha_y : [0, \delta] \to \mathbb{R}^d$,

$$\alpha_y(s) := \gamma_x(s - \delta) + \frac{s}{\delta}(y - x).$$

Since $\gamma_x \in C^2$, so is α_y. Observe that the dependence on y is smooth. Define, in a neighborhood of x,

$$\varphi(y) = u(x) + \int_0^\delta L(\alpha_y, \dot{\alpha}_y) \, ds - \int_{-\delta}^0 L(\gamma_x, \dot{\gamma}_x).$$

Then, $\varphi \in C^2$ and touches u from above at x. Analogously, by considering

$$\beta_y(s) := \frac{\delta - s}{s}(y - x) + \gamma_x(s),$$

we obtain a C^2 function touching u from below at the point x. This proves u is differentiable in x. Being u differentiable, Proposition 25 implies

$$du(x) \cdot \dot\gamma(0) = L\big(x, \dot\gamma(0)\big) + c_0 \geq L\big(x, \dot\gamma(0)\big) + H\big(x, du(x)\big), \tag{2.28}$$

and the Legendre–Fenchel duality provides the desired result. □

Since a critical curve is uniquely determined, then for every $x \in \mathcal{A}$, and every subsolution u, the differential $du(x)$ is uniquely determined, namely

$$du(x) = \frac{\partial L}{\partial v}\big(x, \dot\gamma(0)\big).$$

Definition 27. The *Aubry set* $\tilde{\mathcal{A}} \subset T^*M$ is defined as

$$\tilde{\mathcal{A}} = \Big\{\big(x, du(x)\big) \in T^*M \ ; \ x \in \mathcal{A} \text{ and } u \text{ any critical subsolution}\Big\}. \tag{2.29}$$

Theorem 28. *[Mather's Graph Theorem] The restricted projection $\pi : \tilde{\mathcal{A}} \to M$ is injective, $\pi(\tilde{\mathcal{A}}) = \mathcal{A}$, and the inverse*

$$\pi^{-1} : \mathcal{A} \to \tilde{\mathcal{A}}$$

is a Lipschitz graph.

Proof. By (2.26), if $x \in \mathcal{A}$ then $du(x)$ is independent of the subsolution u, and so $x \mapsto du(x)$ is well defined as a map. Moreover, at $x \in \mathcal{A}$, there exist a C^2 function touching u from above and a C^2 function touching u from below. Then, u is $C^{1,1}$ on $\mathcal{A}$; therefore, $x \mapsto du(x)$ is Lipschitz. □

We explain next the behavior of subsolutions in the Aubry set. Bernard's Theorem given in Section 2.5 provides an elegant proof of the following:

Theorem 29 (Bernard)**.** *Let $H : T^*M \to \mathbb{R}$ be a Tonelli Hamiltonian. Then,* *there exists a $C^{1,1}$ critical subsolution u strict outside $\mathcal{A}$, that is,*

$$H\big(x, du(x)\big) < c_0, \ \forall x \notin \mathcal{A}.$$

Proof. Endow the set of critical subsolutions with the C^1 topology. This space is separable, because $C^1(M)$ is, with the same topology. Then, we consider a dense subset u_n of C^1 subsolutions.

Claim: u defined by

$$u(x) = \sum_{n=1}^{\infty} \frac{u_n(x)}{2^n}$$

is a C^1 subsolution that is strict outside $\mathcal{A}$.

Indeed, for each $x \notin \mathcal{A}$, since there exists a critical subsolution which is strict in x, we have $H\big(x, du_n(x)\big) < c_0$, for some n (by density). Then, the convexity of H implies

$$H\big(x, du(x)\big) \leq \sum_n \frac{H\big(x, du_n(x)\big)}{2^n} < c_0.$$

The existence of a $C^{1,1}$ subsolution follows from the density result given by Theorem 22. □

2.7 Mather set

In this section we describe the construction of the Mather set, following Mather [23] and Mañè [22]. To each trajectory $\gamma : [0, +\infty] \to M$, we can associate a family of probability measures $\{\mu_\gamma^t\}_{t \geq 0}$ on TM by

$$\int_{TM} F \, d\mu_\gamma^t := \frac{1}{t} \int_0^t F\big(\gamma(s), \dot{\gamma}(s)\big) \, ds, \tag{2.30}$$

for all $F : TM \to \mathbb{R}$ continuous. Since γ is Lipschitz, μ_γ^t is weakly-* compact; then, there exists a probability measure μ_γ on TM such that

$$\mu_\gamma^{t_n} \rightharpoonup \mu_\gamma.$$

If we consider a function $\varphi \in C^1(M)$, then μ_γ satisfies

$$\begin{aligned}\int d\varphi(x)\cdot v\; d\mu_\gamma &= \lim_{t\to+\infty}\int d\varphi(x)\cdot v\; d\mu_\gamma^t \\ &= \lim_{t\to+\infty}\frac{1}{t}\int_0^t d\varphi\big(\gamma(s)\big)\cdot\dot\gamma(s)\; ds \\ &= \lim_{t\to+\infty}\frac{\varphi\big(\gamma(t)\big)-\varphi\big(\gamma(0)\big)}{t} = 0.\end{aligned} \tag{2.31}$$

Following Mañé [22], we define

Definition 30. We call $\mu \in \mathcal{P}(TM)$ an *holonomic probability measure* if, for any smooth function $\varphi \in C^1(M)$,

$$\int_{TM} d\varphi(x)\cdot v\; d\mu(x,v) = 0. \tag{2.32}$$

We denote by $\mathcal{H}$ the set of holonomic probability measures.

Mather problem. Minimize

$$\mathbb{L}(\mu) := \int_{TM}\big(L(x,v)+c_0\big)\; d\mu(x,v),\;\text{ for } \mu\in\mathcal{H}.$$

It can be shown [22]

$$\mathcal{H} = \overline{\Big\{\mu_\gamma\ ;\ \gamma\in C^1_{per}([0,t];M), t>0\Big\}}.$$

By such characterization, and the construction of μ_γ at the beginning of this section, we can see that

$$\inf_{\mu\in\mathcal{H}}\int_{TM} L(x,v)\; d\mu = -c_0.$$

Proposition 31. *There exists a minimizing measure $\mu\in\mathcal{H}$, that is, μ satisfies*

$$\int_{TM} L(x,v)\; d\mu = -c_0.$$

Proof. By superlinearity,

$$\tilde{\mathcal{H}} := \mathcal{H} \cap \{\mu\ ;\ \mathbb{L}(\mu) \leq -c_0 + 1\}$$

is compact and convex on $\mathcal{P}(TM)$. Therefore, $\mathbb{L}$ attains a minimum in $\mathcal{H}$. $\square$

Definition 32. The *Mather set* $\tilde{\mathcal{M}}$ is defined as

$$\tilde{\mathcal{M}} := \overline{\bigcup_{\mathbb{L}(\mu)=-c_0} \operatorname{supp} \mu} \subset TM. \tag{2.33}$$

The *projected Mather set* is defined as $\mathcal{M} := \pi(\tilde{\mathcal{M}}) \subset M$.

The existence of smooth subsolutions allows us to easily prove the next proposition.

Proposition 33. *The projected Mather set is contained in the projected Aubry set of Section 2.5.*

Proof. Let μ be any minimizing measure, and u any C^1 subsolution. Then, since $\mu \in \mathcal{H}$, the Legendre–Fenchel inequality shows

$$\begin{aligned} 0 = \int_{TM} du(x) \cdot v \; d\mu &\leq \int_{TM} \left[L(x,v) + H\big(x, du(x)\big)\right] \; d\mu \\ &\leq \int_{TM} \left[L(x,v) + c_0\right] \; d\mu = 0. \end{aligned} \tag{2.34}$$

Therefore,

$$H\big(x, du(x)\big) = c_0, \text{ for } x \in \pi(\operatorname{supp} \mu)$$

and $x \in \mathcal{A}$. $\square$

In [23], Mather proved the following theorem, that we obtain as a consequence of the previous proposition.

Corollary 34 (Mather Graph Theorem)**.** *The restricted projection $\pi : \tilde{\mathcal{M}} \to M$ is injective, $\pi(\tilde{\mathcal{M}}) = \mathcal{M}$, and the inverse*

$$\pi^{-1} : \mathcal{M} \to \tilde{\mathcal{M}}$$

is Lipschitz.

Proof. Follows from Proposition 33 combined with Theorem 28. □

2.8 Time–dependent Hamilton–Jacobi equations

Let us now consider the time–dependent Hamilton–Jacobi equation

$$\begin{cases} \partial_t v(t,x) + H(x, \partial_x v(t,x)) = 0, & \text{in } M \times (0,T]; \\ v(0,\cdot) = v_0, & \text{on } M \times \{0\}, \end{cases} \tag{2.35}$$

for a given continuous function $v_0 : M \to \mathbb{R}$. By using the Legendre–Fenchel duality, as in the stationary case, we are led to consider the following candidate for a solution of (2.35):

$$v(t,x) := \inf \left\{ v_0(\gamma(0)) + \int_0^t L\big(\gamma(s), \dot\gamma(s)\big)\, ds \right\} = T_t v_0(x), \tag{2.36}$$

where the infimum is taken over all absolutely continuous curves $\gamma \in AC([0,t]; M)$ with $\gamma(t) = x$. In fact, we have the following:

Theorem 35. *If $v_0 : M \to \mathbb{R}$ is continuous, then $v : M \to \mathbb{R}$ defined by $v(t,x) := T_t v_0(x)$ is a viscosity solution of (2.35).*

Proof. The proof is standard and can be found, for instance, in [11]. See also, Theorem 88, in Chapter 5. □

We observe that when v_0 is a critical subsolution

$$H\big(x, dv_0(x)\big) \le c_0 \text{ on } M,$$

then the weak KAM theorem, Theorem 1, states that $T_t v_0$ converges as $t \to +\infty$ to a critical solution of

$$H\big(x, du(x)\big) = c_0 \text{ on } M.$$

This means that, in the case v_0 is a critical subsolution, the asymptotic behavior of the solution $v(t,x) = T_t v_0$, as $t \to +\infty$ is a weak KAM solution.

A natural question that arises is the following: Is the asymptotic behavior of solutions of (2.35) the same, for any continuous initial condition? In other words, given any $v_0 : M \to \mathbb{R}$ continuous (not necessarily a subsolution), can we say that $v(t,x) = T_t u(x)$ converges, as $t \to +\infty$, to a critical solution? The following theorem gives a positive answer to this question:

Theorem 36. *Given $u_0 \in C(M)$, the Lax–Oleinik semigroup $T_t u_0$ converges, as $t \to \infty$ to a critical solution $v \in C(M)$ of*

$$H\big(x, du(x)\big) = c_0.$$

Moreover, v can be written as:

$$v(x) = \inf_{y \in \mathcal{A}} \big\{ h(y,x) + \inf_{z \in M} \{ u_0(z) + h(z,y) \} \big\},$$

where $h : M \times M \to \mathbb{R}$ denotes the Peierls barrier

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

We present two proofs of this theorem. The first one is the original one of Fathi [14], exploiting the invariance of Mather measures through the flow. The second one is somehow more elementary and uses critical curves to avoid the use of the flow.

First proof. The set $\{T_t u\}_{t \geq t_0}$, for a fixed $t_0 > 0$, is uniformly Lipschitz. Moreover, take $\tilde{u}$ a critical solution given by Theorem 21, then

$$\|T_t u - \tilde{u}\|_{L^\infty} \leq \|u - \tilde{u}\|_{L^\infty},$$

and so the family is also uniformly bounded. It follows that there exists a sequence $t_n \to +\infty$ such that

$$T_{t_n} u \to u_\infty,$$

by the Arzelà–Ascoli Theorem. Observe that for any C^1 function $w : M \to \mathbb{R}$, by the Legendre–Fenchel inequality, we have

$$w\big(\gamma(t)\big) - w\big(\gamma(0)\big) \le \int_0^t \big(L(\gamma,\dot\gamma) + \sup_x H\big(x, dw(x)\big)\big)\ ds, \tag{2.37}$$

for any curve $\gamma \in AC([0,t];M)$. By approximation, (2.37) holds true for any Lipschitz function w; then, since $T_{t_n}u$ is Lipschitz,

$$T_{t_n}u\big(\gamma(t)\big) - T_{t_n}u\big(\gamma(0)\big) \le \int_0^{t_n} \Big(L(\gamma,\dot\gamma) + \sup_x H\big(x, dT_{t_n}u(x)\big)\Big)\ ds, \tag{2.38}$$

for any $\gamma \in AC([0,t];M)$.
Claim. For every $x \in M$,

$$H\big(x, dT_{t_n}u(x)\big) \to c_0 \text{ as } t_n \to \infty. \tag{2.39}$$

In fact, if consider $\gamma_n : [0,t_n] \to M$ extremal with $\gamma_n(t_n) = x$, then we have (compare to Proposition 94 of Chapter 4)

$$dT_{t_n}u(x) = \frac{\partial L}{\partial v}\big(x, \dot\gamma_n(t_n)\big), \tag{2.40}$$

and the claim is the content of Lemma 37 below.

Now, by letting $t_n \to +\infty$ in (2.38), we obtain

$$u_\infty\big(\gamma(t)\big) - u_\infty\big(\gamma(0)\big) \le \int_0^t \big(L(\gamma,\dot\gamma) + c_0\big)\ ds, \tag{2.41}$$

and we conclude u_∞ is a critical subsolution. To prove that it is actually a critical solution, consider a subsequence of t_n such that $s_n = t_{n+1} - t_n \to \infty$. We have

$$\begin{aligned} \|T_{s_n}u_\infty - u_\infty\| &\le \|T_{s_n}u_\infty - T_{s_n}T_{t_n}u\| + \big\|T_{t_{n+1}}u - u_\infty\big\| \\ &\le \|u_\infty - T_{t_n}u\| + \big\|T_{t_{n+1}}u - u_\infty\big\| \to 0. \end{aligned} \tag{2.42}$$

This shows that u_∞ is a fixed point of the semigroup. □

Lemma 37 (Carneiro's Theorem)**.** *For any $\gamma_n : [0,t_n] \to M$ minimizer, with*

$t_n \to \infty$, *there exists a subsequence* $s_n \in [0, t_n]$, *such that*

$$H\left(\gamma_n(s_n), \frac{\partial L}{\partial v}\big(\gamma_n(s_n), \dot\gamma_n(s_n)\big)\right) \to c_0. \tag{2.43}$$

Proof. We use Riesz Representation Theorem to define a family of probability measures $\mu_n \in \mathcal{P}(TM)$ with support on the speed curves $(\gamma_n(\cdot), \dot\gamma_n(\cdot)) \subset TM$: For any $\theta \in C(TM)$, set

$$\int_{TM} \theta \; d\mu_n := \frac{1}{t_n} \int_0^{t_n} \theta(\gamma_n(s), \dot\gamma_n(s)) \; ds. \tag{2.44}$$

By compactness, the support of all measures μ_n lie in a fixed compact set $K \subset TM$ of the tangent bundle. Then, we can extract a subsequence such that $\mu_n \rightharpoonup \mu$. Such a limit is invariant under the Euler–Lagrange flow as well, and satisfies

$$\int_{TM} L \; d\mu = \lim_{t_n \to \infty} \frac{1}{t_n} \int_0^{t_n} L(\gamma_n(s), \dot\gamma_n(s)) \; ds = -c_0. \tag{2.45}$$

This implies that $\operatorname{supp}\mu \subset \tilde{\mathcal{M}}$. Take a sequence $(\gamma_n(s_n), \dot\gamma_n(s_n)) \in \operatorname{supp}\mu_n$ converging to a point of $\operatorname{supp}\mu$. Since

$$H\left(x, \frac{\partial L}{\partial v}(x, v)\right) = c_0 \tag{2.46}$$

for points $(x, v) \in \tilde{\mathcal{M}}$, the result follows from continuity. □

2.8.1 Convergence of the Lax–Oleinik semigroup; A second proof

The second proof we present is due to Davini and Siconolfi [10]. The idea is to show that the limit $T_t u_0$ exists, as $t \to +\infty$, by looking at the 'semilimits':

$$\underline{u}(x) = \sup\left\{ \limsup_{n\to\infty} (T_{t_n} u_0)(x_n) \right\}$$

and

$$\overline{u}(x) = \inf\left\{\liminf_{n\to\infty}(T_{t_n}u_0)(x_n)\right\},$$

where the supremum and the infimum above are taken over all sequences $x_n \to x$ and $t_n \to +\infty$. By setting

$$\omega(u_0) = \left\{\psi\ ;\ \psi = \lim_n T_{t_n}u_0 \text{ for some } t_n \to +\infty\right\},$$

we have

$$\underline{u}(x) = \sup\left\{\psi(x)\ \middle|\ \psi \in \omega(u_0)\right\} \tag{2.47}$$

and

$$\overline{u}(x) = \inf\left\{\psi(x)\ \middle|\ \psi \in \omega(u_0)\right\}. \tag{2.48}$$

Note that these are well defined by the Arzelà–Ascoli Theorem, since the family $\{T_t u_0\}$ is uniformly bounded and uniformly Lipschitz in t. To prove Theorem 36, we claim, following Davini-Siconolfi [10], that

$$\underline{u} = \overline{u} = v$$

on M, where $v : M \to \mathbb{R}$ is the solution of (2.10) given by the following representation formula

$$v(x) := \inf_{y\in\mathcal{A}}\left\{h(y,x) + \inf_{z\in M}\left\{u_0(z) + h(z,y)\right\}\right\}. \tag{2.49}$$

We start with a lemma

Lemma 38. *Let $\underline{u}$ and $\overline{u}$ be as defined above. Then $\underline{u}$ is a critical subsolution and $\overline{u}$ is a critical supersolution.*

Proof. We prove that $\underline{u}$ is a subsolution. The other is analogous. Fix $t > 0$. By the definition of $\underline{u}$, we know $\underline{u} \geq \varphi$, for any $\varphi \in \omega(u_0)$, so that

$$T_t\underline{u} \geq T_t\varphi,\ \forall\ \varphi \in \omega(u_0).$$

Then,

$$T_t\underline{u} \geq \sup\left\{T_t\varphi\ ;\ \varphi \in \omega(u_0)\right\}.$$

The proof is finished once we prove

$$\big\{T_t\varphi \ ; \ \varphi \in \omega(u_0)\big\} = \big\{\varphi \ ; \ \varphi \in \omega(u_0)\big\}.$$

If $\varphi \in \omega(u_0)$, we know

$$\varphi = \lim_n T_{t_n} u_0$$

for some $t_n \to +\infty$. Up to a subsequence, we also know

$$T_{t_n - t} u_0 \to \tilde{\varphi};$$

then, $T_t \tilde{\varphi} = \lim T_{t_n} u_0 = \varphi$. □

Proposition 39. *Set*

$$v_0(y) := \inf_{z \in M} \Big\{u_0(z) + h(z, y)\Big\},$$

for $y \in M$, so that

$$v(x) = \inf_{y \in \mathcal{A}} \big\{v_0(y) + h(y, x)\big\}. \tag{2.50}$$

Then, the following hold true:

1. *v_0 is the maximal subsolution with $v_0 \leq u_0$ on M;*

2. *v is a solution and it equals v_0 on $\mathcal{A}$;*

3. *If $u_0(y) - u_0(x) \leq h(y, x)$, for all $x, y \in M$, then*

$$v(x) := \inf_{y \in \mathcal{A}} \big\{u_0(y) + h(y, x)\big\}. \tag{2.51}$$

   in M, and $v_0 = u_0$ on $\mathcal{A}$.

Proof. See [10, Theorem 3.1]. □

Proposition 40. *Let u be either a subsolution or a supersolution, and let v be the function defined by (2.49). Then, as $t \to +\infty$,*

$$T_t u \to v.$$

We have proved this before when u is a subsolution (Theorem 21). However, we present this short proof, by Davini-Siconolfi.

Proof. Assume first that u_0 is a subsolution; then, $u_0 \leq T_t u_0$, for all $t \geq 0$. By Proposition 39, v is the maximal subsolution with $v = u_0$ on $\mathcal{A}$, and so we have $u_0 \leq v$ on M. Thus, by the monotonicity of the Lax–Oleinik semigroup, and by observing $v = T_t v$, for any $t \geq 0$, we obtain

$$u_0 \leq T_t u_0 \leq v \text{ on } M.$$

Now, $v = u_0$ on $\mathcal{A}$ implies

$$T_t u_0 = v \text{ on } \mathcal{A}, \ \forall\, t \geq 0.$$

This means

$$\underline{u} = \overline{u} = v \text{ on } \mathcal{A}.$$

Hence, the comparison principle implies the same equality on M.

Next, assume u_0 is a supersolution, so that $u_0 \geq T_t u_0$, and consider

$$v_0(y) := \inf_{z \in M} \big\{u_0(z) + h(z, y)\big\},$$

the maximal subsolution with $v_0 \leq u_0$ on M. Then, by monotonicity,

$$v_0 \leq T_t v_0 \leq T_t u_0 \leq u_0,$$

and the maximality of v_0 implies $v_0 = T_t v_0$, whence v_0 is a solution. This in turn implies $v = v_0$ on M. Thus, again by monotonicity, we obtain

$$v \leq T_t u_0 \leq u_0 \text{ on } M, \text{ for all } \ t \geq 0,$$

which implies $v \leq \overline{u} \leq \underline{u} \leq u_0$ on M. Since $\underline{u}$ is a subsolution, we must have $\underline{u} \leq v$ and the proof is finished. ∎

Proposition 41. *Let $\underline{u}$, $\overline{u}$, and v be given as before. Then,*

$$v \leq \overline{u} \leq \underline{u} \text{ on } M. \tag{2.52}$$

Proof. As in the second part of the previous proof, v_0 satisfies $v_0 \leq u_0$ on M. Then $T_t v_0 \leq T_t u_0$ on M. Since v_0 is a subsolution, we have $T_t v_0 \to v$; therefore, both $v \leq \overline{u}$ and $v \leq \underline{u}$ hold. That $\overline{u} \leq \underline{u}$ holds is immediate. □

Remark 42. By Proposition 41, in order to obtain the desired convergence result for the Lax–Oleinik semigroup, all we need to prove is that $v = \underline{u}$ on $\mathcal{A}$. This is what we prove next, with the help of critical curves constructed in Section 2.6.

Proposition 43. *Let $\underline{u}$, $\overline{u}$, and v be given as before. Then, $\underline{u} \leq v$ on $\mathcal{A}$. Moreover, this implies*

$$v = \overline{u} = \underline{u} \text{ on } M.$$

Proof. Let $\psi \in \omega(u_0)$, and we claim $\psi \leq v$ on $\mathcal{A}$. Since ψ be in the ω-limit set of u_0, $\psi = \lim_n T_{\sigma_n} u_0$ for some divergent sequence $\sigma_n \to +\infty$. It is then not difficult to obtain a sequence $s_n \to +\infty$ such that

$$\psi = \lim_{n\to\infty} T_{s_n}\psi.$$

Let γ be a critical curve, and $x \in \omega(\gamma)$, so that $x = \lim_n \gamma(t_n)$, for some divergence sequence $t_n \to +\infty$. Up to extracting a subsequence, we can assume $\tau_n = t_n - s_n \to +\infty$. Since γ is a critical curve and v is a subsolution (it is in fact a solution), we have (compare to Lemma 94)

$$T_{s_n}\psi\big(\gamma(t_n)\big) - \psi\big(\gamma(t+\tau_n)\big) \leq v\big(\gamma(t_n)\big) - v\big(\gamma(t+\tau_n)\big) + |t|\rho(t/s_n). \tag{2.53}$$

If we set

$$\eta = \lim_n \gamma(\cdot + \tau_n),$$

then, η is also a critical curve and, by letting $n \to +\infty$ in (2.53), we obtain

$$\psi(x) - \psi(\eta(t)) \leq v(x) - v(\eta(t)).$$

It only remains to show that

$$\liminf_t \big\{\psi\big(\eta(t)\big) - v\big(\eta(t)\big)\big\} \leq 0.$$

To this order, observe

$$v(\eta(t)) - v(\eta(0)) = \int_0^t L(\gamma, \dot{\gamma}) \geq T_t u_0(\eta(t)) - u_0(\eta(0)).$$

Thus, since $\eta(\mathbb{R}) \subset \mathcal{A}$ and $v = u_0$ on $\mathcal{A}$,

$$\begin{aligned} \psi\big(\eta(t)\big) - v\big(\eta(t)\big) &\leq \psi\big(\eta(t)\big) - T_t u_0(\eta(t)) + u_0(\eta(0)) - v\big(\eta(0)\big) \\ &\leq \max_{y \in \mathcal{A}} \big|\psi - T_t u_0\big|. \end{aligned} \tag{2.54}$$

Since along the sequence $\sigma_n \to +\infty$, we have $T_{\sigma_n} u_0 \to \psi$, our claim is proved. $\square$

Chapter 3

Optimal switching problem

In this chapter we study the optimal switching problem. In Section 3.1 we state the problem and relatively mild assumptions on both L and ψ we need in what follows. Next, in Section 3.2 we prove the main result of this chapter, namely the existence of minimizers for the action $\mathcal{J}_t$, given by (3.2) below. Although we believe this is a very natural problem, we were unable to find it in the literature. In Section 3.3 we obtain necessary conditions for $(\gamma_M, \gamma_{\mathcal{I}})$ to be a minimizer and, as a corollary, obtain regularity of γ_M, from the usual regularity theory for Tonelli Lagrangians. In section 3.4 we prove that a conservation of energy principle holds, which is what we naturally expect for such problems. Finally, in Section 3.5 we prove that the cost function h_t (see definitions below) is locally semiconcave, as the usual cost function in the 'classical' theory of the calculus of variations is. In the literature, Lipschitz regularity when $M = \mathbb{R}^n$ has been proved before (see [19, Proposition 2.4]).

3.1 Setting of the problem

We consider the *optimal switching problem.* Intuitively, it consists of minimizing an action functional in a system that allows switching between different "modes". In other words, we have the option of switching between given different settings of the system for a given price, whenever it is convenient in order to minimize the cost. We proceed for a more detailed description of the problem.

We call, as in the classical case,

$$\begin{aligned} L : TM \times \mathcal{I} &\to \mathbb{R} \\ (x, v, i) &\mapsto L(x, v, i) \end{aligned} \tag{3.1}$$

a *Lagrangian*, where $\mathcal{I}$ is a given family of indices that prescribes the different settings of our system – each $i \in \mathcal{I}$ is called a *mode*. We define $\mathcal{P}([a,b];\mathcal{I})$ as the set of piecewise constant functions

$$\sigma = \sum_{i=0}^{N} \sigma_i \chi_{[t_i, t_{i+1})}, \quad a = t_0 \leq t_1 \leq \cdots \leq t_{N+1} = b,$$

taking values in the index set $\mathcal{I}$. Then, in analogy with the classical case, we also define the *action* functional of L by

$$\mathcal{J}_t[\gamma] = \int_0^t L(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)) \ ds + \sum_{k=0}^{N-1} \psi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})\big), \tag{3.2}$$

where $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$, with $\gamma_M \in AC([0,t];M)$, $\gamma_{\mathcal{I}} \in \mathcal{P}([0,t];\mathcal{I})$, and $\gamma_{\mathcal{I}}(t_k) \in \mathcal{I}$ is the value assumed by the piecewise constant curve $\gamma_{\mathcal{I}}$ in the largest subinterval $[t_i, t_{i+1}) \subset [0,t]$ where it is constant. The function $\psi : M \times \mathcal{I} \times \mathcal{I} \to \mathbb{R}$ is assumed to be non-negative, and it is called the *switching cost*. The *cost function* is set as

$$h_t(x,i,y,j) := \inf \Big\{ \mathcal{J}_t[\gamma] \ ; \ \gamma \in AC \times \mathcal{P}, \ \gamma(0) = (x,i), \ \gamma(t) = (y,j) \Big\}. \tag{3.3}$$

As it might be already clear, the idea is that, when trying to minimize the action functional, we might decide to switch from the Lagrangian i to the Lagrangian j when we arrive at x, because it might decrease the action, even though we pay a certain fee $\psi(x,i,j)$ to do so.

We make the following assumptions on L and ψ:

A1. The Lagrangian $L : TM \times \mathcal{I} \to \mathbb{R}$ is continuous, with $\mathcal{I} = \{1, \ldots, m\}$ a finite set endowed the discrete topology. Moreover, for every $i \in \mathcal{I}$, $L(\cdot, \cdot, i)$ is a Tonelli Lagrangian, as in Definition 7;

A2. $L(\cdot, \cdot, i)$ is superlinear above compact sets, uniformly in $i \in \mathcal{I}$, meaning that, given $K \subset M$ compact, for every constant $A \geq 0$, there exists a constant $C \in \mathbb{R}$, depending on K and A, such that

$$L(x, v, i) \geq A \|v\|_x + C, \text{ for all } x \in K,\ v \in T_x M;$$

A3. The switching cost function $\psi : M \times \mathcal{I} \times \mathcal{I} \to \mathbb{R}$ is continuous and satisfies a triangle inequality: For all distinct $i, j, k \in \mathcal{I}$, and all $x \in M$, we have

$$\psi(x, i, j) < \psi(x, i, k) + \psi(x, k, j);$$

A4. We also assume that and $\psi(\cdot, i, i) \equiv 0$ and a bound from below on the switching cost:

$$\min_{x \in M, i \neq j} \psi(x, i, j) > 0;$$

A5. For all $i, j \in \mathcal{I}$, $\psi(\cdot, i, j) \in C^2(M)$.

Remark 44. Condition $A3$ is natural in the sense that it does not allow us to switch from Lagrangian i to j and then to k in a short period of time, just because it would be cheaper than going straight from i to k.

Remark 45. Condition $A4$ is automatically satisfied, for instance, when M is compact and the set of modes $\mathcal{I}$ is finite.

Notation We denote the elements of $M \times \mathcal{I}$ by $A = (x, i)$, $B = (y, j)$, and $C = (z, k)$ and $\gamma(t) = \big(\gamma_M(t), \gamma_{\mathcal{I}}(t)\big)$. This makes the presentation more elegant and clear.

The purpose of this chapter is to prove the following two theorems:

Theorem 46. *Let M be a complete differentiable manifold, and $\mathcal{I} = \{1, \ldots, m\}$ a finite set. Also, let $L : TM \times \mathcal{I} \to \mathbb{R}$ be a Lagrangian satisfying* A1 *and* A2*, and $\psi : M \times \mathcal{I} \times \mathcal{I} \to \mathbb{R}$ a switching cost satisfying* A3 *and* A4*. Then, for every $A, B \in M \times \mathcal{I}$, there exists a minimizer for the action: there exists $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) \in AC \times \mathcal{P}$ with $\gamma(0) = A$, $\gamma(t) = B$ such that*

$$\mathcal{J}_t[\gamma] = \inf \Big\{ \mathcal{J}_t[\alpha] \ ; \ \alpha \in AC \times \mathcal{P}, \alpha(0) = A, \alpha(t) = B \Big\}.$$

Moreover,

1. *Minimizers are of class $C^2([0,t];M)$, meaning $\gamma_M \in C^2([0,t];M)$, and solve the Euler–Lagrange equation*

$$\frac{d}{ds}\left[\frac{\partial L}{\partial v}\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(s)\big)\right] = \frac{\partial L}{\partial x}\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(s)\big) \tag{3.4}$$

   in $[0,t]\backslash\{t_1,\dots,t_N\}$, where t_k's are the points where $\gamma_{\mathcal{I}}$ has a jump;

2. *Along a minimizing curve, the (generalized) energy functional is conserved:*

$$E\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(s)\big) := \frac{\partial L}{\partial v}\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(s)\big)\cdot\dot\gamma_M(s) - L\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(s)\big) \tag{3.5}$$

   is constant in $[0,t]$.

In the following sections, we separate the proof of Theorem 46 into various propositions, and in Section 3.5 we obtain the following regularity result for the cost function:

Theorem 47. *Suppose $L : TM \to \mathbb{R}$ is a Tonelli Lagrangian satisfying A1 and A2, and $\psi : M \times \mathcal{I} \times \mathcal{I} \to \mathbb{R}$ a switching cost satisfying A3 and A4. Then,*

1. *For any $i,j \in \mathcal{I}$, the restricted cost function $h_t(\cdot,i,\cdot,j) : M \times M \to \mathbb{R}$ given by (3.3) is locally semiconcave on $M \times M$.*

2. *For any $x,y \in M$ and any $i,j,k \in \mathcal{I}$,*

$$h_t(x,i,y,j) - h_t(x,i,y,k) \le \psi(x,j,k).$$

3.2 Existence of action minimizers

In this section we prove the existence of minimizers for the action $\mathcal{J}_t$, under fixed boundary conditions. As already mentioned before, we were unable to find this result in the literature. Our proof follows the direct method in the calculus of variations. We refer the reader to [8] for an introduction to the calculus of variations methods, or to Appendix A for a concise presentation of the results we need.

Recall we denote elements of $M \times \mathcal{I}$ by $A = (x, i)$, $B = (y, j)$, and $C = (z, k)$ and $\gamma(t) = (\gamma_M(t), \gamma_{\mathcal{I}}(t))$. Now we are ready to prove the existence of minimizers:

Proposition 48. *Assume A1–A4 hold, where $\mathcal{J}_t$ is defined by (3.2) and h_t by (3.3). Then, for every $A, B \in M \times \mathcal{I}$, there exists a minimizer for the action of $\mathcal{J}_t$: there exists $\gamma \in AC \times \mathcal{P}$ with $\gamma(0) = A$ and $\gamma(t) = B$ such that*

$$h_t(A, B) = \mathcal{J}_t[\gamma].$$

Proof. Assume, without loss of generality, that $L \geq 0$. Let $\gamma^k = (\gamma^k_M,\ \gamma^k_{\mathcal{I}})$ be a minimizing sequence: $\gamma^k(0) = (x, i)$, $\gamma^k(t) = (y, j)$ and

$$\mathcal{J}_t[\gamma^k] \to h_t(A, B), \tag{3.6}$$

and consider partitions

$$P^k = \{0 < t^k_1 < \cdots < t^k_{N^k} < t\}$$

of the interval $[0, t]$ that indicate the maximal sets where the $\gamma^k_{\mathcal{I}}$'s are constant.

Step 1. Our assumptions ensure that $N^k \not\to \infty$ as $k \to \infty$, that is, we do not reach a minimizer by increasing the number of switches. Indeed, assume, by contradiction, that this is not the case. By $A4$, we have

$$\delta := \min_{x \in M, i \neq j \in \mathcal{I}} \psi(x, i, j) > 0. \tag{3.7}$$

Then, for any $k \in \mathbb{N}$,

$$N^k \delta \leq \sum_{i=0}^{N^k-1} \int_{t^k_i}^{t^k_{i+1}} L(\gamma^k_M(s), \dot{\gamma}^k_M(s), \gamma^k_{\mathcal{I}}(t^k_i))\, ds + \sum_{i=0}^{N^k-1} \psi(\gamma^k_{\mathcal{I}}(t^k_i), \gamma^k_{\mathcal{I}}(t^k_{i+1})) = \mathcal{J}_t[\gamma^k],$$

and $\mathcal{J}_t[\gamma^k]$ would have to be unbounded, a contradiction (h_t is finite; for instance, consider a geodesic from x to y and one switch from i to j).

Step 2. Convergence to a minimizer candidate. Step 1 implies we may assume

$N^k \equiv N$, without any loss of generality. We can then write

$$P^k := \{0 \leq t_1^k \leq \cdots \leq t_N^k \leq t\},$$

where now the intervals might not be the maximal ones where $\gamma_{\mathcal{I}}^k$'s are constant. Since, for every i, $(t_i^k)_k$ is a bounded sequence of real numbers, it converges, up to a subsequence, to some $t_i \in [0,t]$. We have

$$0 \leq t_1 \leq \cdots \leq t_N \leq t.$$

From assumption $A1$ and the third condition in the definition of a Tonelli Lagrangian, we obtain

$$\int_0^t \left\|\dot{\gamma}_M^k(s)\right\|_{\gamma_M^k(s)} \, ds + C \leq \int_0^t L(\gamma_M^k(s), \dot{\gamma}_M^k(s), \gamma_{\mathcal{I}}^k(s)) \, ds$$

for some constant $C \in \mathbb{R}$. This implies that for all $s \in [0,t]$, $\gamma_M^k(s)$ lives in a compact set. Then, $A1$ and the results of Appendix A) imply that there exists $\gamma_M : [0,t] \to M$ absolutely continuous such that, up to a subsequence,

$$\begin{aligned} \gamma_M^k &\to \gamma_M \text{ uniformly in } [0,t], \\ \dot{\gamma}_M^k &\rightharpoonup \dot{\gamma}_M \text{ weakly in } L^1. \end{aligned} \tag{3.8}$$

With the possible exception of the initial and final points, we exclude the cases where $t_i = t_{i+1}$, obtaining a limiting partition

$$\{0 = t_{i_0} \leq t_{i_1} < t_{i_2} < \cdots < t_{i_\ell} \leq t_{i_{\ell+1}} = t\},$$

where we define $\gamma_{\mathcal{I}}$: for $\varepsilon > 0$ small, we have, once k is sufficiently large, that $\gamma_{\mathcal{I}}^k \equiv \gamma_{\mathcal{I}}^k(t_{i_j^*})$ in each of the subintervals $[0, t_{i_1} - \varepsilon), (t_{i_1} + \varepsilon, t_{i_2} - \varepsilon), \ \ldots, (t_{i_{\ell-1}} + \varepsilon, t_{i_\ell} - \varepsilon), (t_{i_\ell} + \varepsilon, t]$, and so we have the pointwise convergence

$$\gamma_{\mathcal{I}}^k \to \gamma_{\mathcal{I}} := \sum_{j=0}^{\ell-1} \gamma_{\mathcal{I}}(t_{i_{j+1}^*}) \chi_{[t_{i_j^*}, t_{i_{j+1}^*})}, \tag{3.9}$$

except at most in the points $t_{i_j^*}$. We have $\gamma_I^k(0) \to i$, $\gamma_I^k(t) \to j$, because these are constant. Assumption $A3$ suggests we should choose

$$\gamma_{\mathcal{I}}(t_{i_j^*}) := \lim_{s \to t_{i_j^*}^+} \gamma_{\mathcal{I}}(s),$$

so that we switch less times.

Step 3. Lower semicontinuity. Since $L \geq 0$,

$$\begin{aligned}\int_0^t L(\gamma_M^k(s), \dot\gamma_M^k(s), \gamma_{\mathcal{I}}^k(s))\, ds \geq & \int_0^{t_{i_1}-\varepsilon} L(\gamma_M^k(s), \dot\gamma_M^k(s), \gamma_{\mathcal{I}}(t_{i_0^*}))\, ds \\ & + \int_{t_{i_1}+\varepsilon}^{t_{i_2}-\varepsilon} L(\gamma_M^k(s), \dot\gamma_M^k(s), \gamma_{\mathcal{I}}(t_{i_1^*}))\, ds \\ & + \cdots + \int_{t_{i_\ell}+\varepsilon}^{t} L(\gamma_M^k(s), \dot\gamma_M^k(s), \gamma_{\mathcal{I}}(t_{i_\ell^*}))\, ds \end{aligned} \tag{3.10}$$

Notice that it might happen that some of the intervals $[t_i^k, t_{i+1}^k]$ collapse when $k \to \infty$, and we are completely ignoring some of the switches. On the other hand, the triangle inequality $A3$ and the continuity of ψ in the state-variable give

$$\begin{aligned}\sum_{i=0}^{N-1} \psi\big(\gamma_M(t_{i+1}^k), \gamma_{\mathcal{I}}(t_i^k), \gamma_{\mathcal{I}}(t_{i+1}^k)\big) + \varepsilon \geq & \ \psi\big(x, i, \gamma_{\mathcal{I}}(0)^+\big) \\ & + \sum_{j=0}^{\ell-1} \psi\big(\gamma_M(t_{i_{j+1}^*}), \gamma_{\mathcal{I}}(t_{i_j^*}), \gamma_{\mathcal{I}}(t_{i_{j+1}^*})\big) \\ & + \psi\big(y, \gamma_{\mathcal{I}}(t)^-, j\big) \end{aligned} \tag{3.11}$$

for all k sufficiently big. So, by adding (3.10) and (3.11), taking the $\liminf$ on both sides, and using the lower semicontinuity with respect to the convergence in

(3.8) for each the integrals on the right hand side of (3.10), we obtain

$$
\begin{aligned}
h_t(A,B) &= \liminf_k \left(\int_0^t L(\gamma_M^k(s), \dot\gamma_M^k(s), \gamma_{\mathcal{I}}^k(s))\, ds + \sum_{i=0}^{N-1} \psi(\gamma_{\mathcal{I}}(t_i^k), \gamma_{\mathcal{I}}(t_{i+1}^k)) \right) \\
&\geq \int_0^{t_{i_1^*}-\varepsilon} L(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(t_{i_0^*}))\, ds + \int_{t_{i_1^*}+\varepsilon}^{t_{i_2^*}-\varepsilon} L(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(t_{i_1^*}))\, ds \\
&\quad + \cdots + \int_{t_{i_\ell^*}+\varepsilon}^{t} L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(t_{i_\ell^*})\big)\, ds + \psi\big(x, i, \gamma_{\mathcal{I}}(0)^+\big) \\
&\quad + \sum_{j=0}^{\ell-1} \psi\big(\gamma_{\mathcal{I}}(t_{i_j^*}), \gamma_{\mathcal{I}}(t_{i_{j+1}^*})\big) + \psi\big(y, \gamma_{\mathcal{I}}(t)^-, j\big),
\end{aligned}
\tag{3.12}
$$

for all $\varepsilon > 0$. Hence, let $\varepsilon \to 0$ to conclude

$$
\begin{aligned}
h_t(A,B) \geq &\int_0^t L(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s))\, ds + \psi\big(x, i, \gamma_{\mathcal{I}}(0)^+\big) \\
&+ \sum_{j=0}^{\ell-1} \psi\big(\gamma_M(t_{i_{j+1}^*}), \gamma_{\mathcal{I}}(t_{i_j^*}), \gamma_{\mathcal{I}}(t_{i_{j+1}^*})\big) + \psi\big(y, \gamma_{\mathcal{I}}(t)^-, j\big) = \mathcal{J}_t[\gamma];
\end{aligned}
\tag{3.13}
$$

therefore, $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ is a minimizer of the action of $\mathcal{J}_t$. □

3.3 Euler–Lagrange equations

In this section we obtain necessary conditions for minimality. First, we prove that the Euler–Lagrange equations are satisfied, except possibly where $\gamma_{\mathcal{I}}$ has a jump, and conditions for the Lagrangian L for every time where a jump exists. We also obtain regularity results for the minimizers of the optimal switching problem. Of course, $\gamma_{\mathcal{I}} \in \mathcal{P}$ is piecewise constant and so we focus on the regularity of γ_M.

Assume $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ is a minimizer of (3.3). Two variations are quite natural. We can fix γ_M and see what happens when $\gamma_{\mathcal{I}}$ changes, or the other way around. On the other hand, we can fix $\gamma_{\mathcal{I}}$ and see what happens when we change γ_M smoothly.

Proposition 49 (Euler-Lagrange equations). *Assume conditions A1 through A4* *hold. If $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ is a minimizer for the action $\mathcal{J}_t$ given by (3.2), then, in coordinates, the Euler–Lagrange equations are satisfied:*

$$\frac{d}{ds}\left[\frac{\partial L}{\partial v}\big(\gamma_M(s), \dot{\gamma_M}(s), \gamma_{\mathcal{I}}(s)\big)\right] = \frac{\partial L}{\partial x}\big(\gamma_M(s), \dot{\gamma_M}(s), \gamma_{\mathcal{I}}(s)\big) \tag{3.14}$$

in $U \equiv [0,t]\backslash\{t_1, \ldots, t_N\}$, where the t_i's are the switching times.

Proof. Observe that $\gamma_M|_{[t_i,t_{i+1}]}$ must be a minimizer of

$$\int_{t_i}^{t_{i+1}} L(\alpha(s), \dot{\alpha}(s), \gamma_{\mathcal{I}}(t_i))\, ds,$$

among all curves $\alpha \in AC([t_i, t_{i+1}], M)$, with $\alpha(t_i) = \gamma_M(t_i)$, $\alpha(t_{i+1}) = \gamma_M(t_{i+1})$. Then, (3.14) follows from the classical Euler–Lagrange equations applied to the Lagrangian $L(\cdot, \cdot, \gamma_{\mathcal{I}}(t_i))$ (see Theorem 13, in Chapter 2). □

Remark 50. It is important to consider the subset U in the last proof since $L(\ \cdot\ ,\ \cdot\ , \gamma_{\mathcal{I}}) \equiv L(\ \cdot\ ,\ \cdot\ , \gamma_{\mathcal{I}}(t_i))$ is constant in each of the subintervals $(t_i, t_{i+1}) \subset [0,t]$.

As it is customary in the calculus of variations, the Euler–Lagrange equations provide regularity properties of minimizers. For Tonelli Lagrangian actions, $L \in C^r$ implies $\gamma \in C^r$. For a proof of such a statement, see Theorem 13, in Chapter 2 (see also [15]).

Corollary 51. *Assume conditions A1 through A4 hold. If $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ is a minimizer for the action $\mathcal{J}_t$ given by (3.2). Then, γ_M is a curve of class C^2 in $U = [0,1]\backslash\{t_1, \ldots, t_N\}$, Moreover, if we assume the Lagrangian $L : TM \times \mathcal{I} \to \mathbb{R}$ is of class C^r, so is γ_M in U.*

Now, with the help of the Euler–Lagrange equations, we obtain the "continuity" of the first derivative of L with respect to v.

Proposition 52 (Necessary conditions for minimality II)**.** *Assume A1–A5. If $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) \in AC \times \mathcal{P}$ is a minimizer for the action $\mathcal{J}_t$, given by (3.2), then*

$$\begin{aligned}\frac{\partial L}{\partial v}\big(\gamma_M(t_i), \dot{\gamma}_M(t_i)^-, \gamma_{\mathcal{I}}(t_{i-1})\big) =& \frac{\partial L}{\partial v}\big(\gamma(t_i), \dot{\gamma}(t_i)^+, \gamma_{\mathcal{I}}(t_i)\big)\\ &+ \partial_x\psi\big(\gamma_M(t_i), \gamma_{\mathcal{I}}(t_{i-1}), \gamma_{\mathcal{I}}(t_i)\big),\end{aligned} \tag{3.15}$$

for $i = 1, \dots, N$.

Proof. Let $\bar{t} = \min\{t_i - t_{i-1}, t_{i+1} - t_i\}$ and consider $\omega \in C^\infty([0,t]; M)$ such that $\omega \equiv 1$ in $[t_i - \bar{t}/2, t_i + \bar{t}/2]$ and $\omega \equiv 0$ outside $[t_i - \bar{t}, t_i + \bar{t}]$. Then

$$\begin{aligned}
0 &= \frac{d}{ds}\mathcal{J}_t[\gamma_M + s\omega, \gamma_{\mathcal{I}}] \\
&= \int_0^t \left(\frac{\partial L}{\partial x}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}})\omega + \frac{\partial L}{\partial v}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}})\dot{\omega}\right) ds \\
&\qquad\qquad + \sum_{k=0}^{N} \partial_x \psi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})\big)\omega(t_{k+1}) \\
&= \int_{t_{i-1}}^{t_i} \left(\frac{\partial L}{\partial x}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}})\omega + \frac{\partial L}{\partial v}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}})\dot{\omega}\right) ds \\
&\quad + \int_{t_i}^{t_{i+1}} \left(\frac{\partial L}{\partial x}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}})\omega + \frac{\partial L}{\partial v}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}})\dot{\omega}\right) ds \\
&\qquad\qquad + \sum_{k=0}^{N} \partial_x \psi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})\big)\omega(t_{k+1}) \\
&= \int_{t_{i-1}}^{t_{i+1}} \left(\frac{\partial L}{\partial x}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) - \frac{d}{ds}\frac{\partial L}{\partial v}(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}})\right)\omega\, ds \\
&\quad + \frac{\partial L}{\partial v}\big(\gamma_M(t_i), \dot{\gamma}_M(t_i)^-, \gamma_{\mathcal{I}}(t_{i-1})\big) - \frac{\partial L}{\partial v}\big(\gamma_M(t_i), \dot{\gamma}_M(t_i)^+, \gamma_{\mathcal{I}}(t_i)\big) \\
&\qquad\qquad + \partial_x \psi\big(\gamma_M(t_i), \gamma_{\mathcal{I}}(t_{i-1}), \gamma_{\mathcal{I}}(t_i)\big) \\
&= \frac{\partial L}{\partial v}\big(\gamma_M(t_i), \dot{\gamma}_M(t_i)^-, \gamma_{\mathcal{I}}(t_{i-1})\big) - \frac{\partial L}{\partial v}\big(\gamma_M(t_i), \dot{\gamma}_M(t_i)^+, \gamma_{\mathcal{I}}(t_i)\big) \\
&\qquad\qquad + \partial_x \psi\big(\gamma_M(t_i), \gamma_{\mathcal{I}}(t_{i-1}), \gamma_{\mathcal{I}}(t_i)\big),
\end{aligned} \tag{3.16}$$

where in the fourth equality we have used the fact that, by construction, $\omega(t_i) = 1$ and $\omega(t_{i-1}) = \omega(t_{i+1}) = 0$, and we have used, in the last equality, the Euler–Lagrange equation of Theorem 49. □

Remark 53. Proposition 52 shows, in particular, that when the switching cost ψ is independent of the state variable, that is, when $\psi(x, i, j) \equiv \psi(i, j)$, for any $i, j \in \mathcal{I}$, then the function

$$t \mapsto \frac{\partial L}{\partial v}\big(\gamma_M(t), \dot{\gamma}_M(t), \gamma_{\mathcal{I}}(t)\big)$$

is continuous in $[0,t]$, whenever γ is an action minimizer.

3.4 Conservation of Energy

In this section we define the energy functional of our system and prove a conservation of energy principle that generalizes the classical case of a single Lagrangian.

Definition 54. The energy functional $E : TM \times \mathcal{I} \to \mathbb{R}$ is defined as

$$E(x,v,i) := \frac{\partial L}{\partial v}(x,v,i) \cdot v - L(x,v,i). \tag{3.17}$$

Proposition 55 (Conservation of Energy)**.** *Along a minimizer of the action $\mathcal{J}_t$, given by (3.2), the energy of the system is conserved, that is,*

$$E(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(s)) := \frac{\partial L}{\partial v}\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(s)\big)\cdot\dot\gamma_M(s) - L\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(s)\big)$$

is constant in $[0,t]$, when $\gamma = (\gamma_M,\gamma_{\mathcal{I}})$ is an action minimizer.

Proof. It suffices to check it at the switching times. Let $\omega : [0,t] \to [0,t]$ be smooth and compactly supported in $[t_{i-1},t_{i+1}]$ with $\omega(t_{i-1}) = \omega(t_{i+1}) = 0$ and $\omega \equiv 1$ in $[t_i - \delta, t_i + \delta]$ for some small $\delta > 0$. Define $\gamma_M^\varepsilon : [0,t] \to M$ as $\gamma_M^\varepsilon(s) = \gamma_M\big(s - \varepsilon w(s)\big)$, so that the function

$$\begin{aligned} f(\varepsilon) := &\int_0^{t_{i-1}} L\big(\gamma_M^\varepsilon(s),\dot\gamma_M^\varepsilon(s),\gamma_{\mathcal{I}}(s)\big) + \int_{t_{i-1}}^{t_i+\varepsilon} L\big(\gamma_M^\varepsilon(s),\dot\gamma_M^\varepsilon(s),\gamma_{\mathcal{I}}(t_{i-1})\big)\, ds \\ &+ \int_{t_i+\varepsilon}^{t_{i+1}} L\big(\gamma_M^\varepsilon(s),\dot\gamma_M^\varepsilon(s),\gamma_{\mathcal{I}}(t_i)\big)\, ds + \int_{t_{i+1}}^{t} L\big(\gamma_M^\varepsilon(s),\dot\gamma_M^\varepsilon(s),\gamma_{\mathcal{I}}(s)\big) \\ &+ \sum \psi\big(\gamma_M(t_{i+1}),\gamma_{\mathcal{I}}(t_i),\gamma_{\mathcal{I}}(t_{i+1})\big) \end{aligned} \tag{3.18}$$

has a minimum at $\varepsilon = 0$. By differentiating with respect to ε and setting $\varepsilon = 0$, we get

$$\begin{aligned}
0 =& L\big(\gamma(t_i), \dot\gamma(t_i)^-, \gamma_{\mathcal I}(t_{i-1})\big) \\
&+ \int_{t_{i-1}}^{t_i} \left(\frac{\partial L}{\partial x}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal I}(t_{i-1})\big) \cdot \dot\gamma_M \omega \right. \\
&+ \left. \frac{\partial L}{\partial v}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal I}(t_{i-1})\big) \cdot (\ddot\gamma_M \omega + \dot\gamma_M \dot\omega) \right) \, ds - L\big(\gamma_M(t_i), \dot\gamma_M(t_i)^+, \gamma_{\mathcal I}(t_i)\big) \\
&- \int_{t_{i-1}}^{t_i} \left(\frac{\partial L}{\partial x}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal I}(t_i)\big) \cdot \dot\gamma_M \omega \right. \\
&+ \left. \frac{\partial L}{\partial v}\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal I}(t_i)\big) \cdot (\ddot\gamma_M \omega + \dot\gamma_M \dot\omega) \right) \, ds.
\end{aligned} \tag{3.19}$$

By integration by parts and the Euler–Lagrange equations, we conclude

$$\begin{aligned}
& L\big(\gamma_M(t_i), \dot\gamma_M(t_i)^-, \gamma_{\mathcal I}(t_{i-1})\big) - \frac{\partial L}{\partial v}\big(\gamma_M(t_i), \dot\gamma_M(t_i)^-, \gamma_{\mathcal I}(t_{i-1})\big) \cdot \dot\gamma(t_i)^- \\
&\quad = L\big(\gamma_M(t_i), \dot\gamma_M(t_i)^+, \gamma_{\mathcal I}(t_i)\big) - \frac{\partial L}{\partial v}\big(\gamma_M(t_i), \dot\gamma_M(t_i)^+, \gamma_{\mathcal I}(t_i)\big) \cdot \dot\gamma^+(t_i),
\end{aligned} \tag{3.20}$$

as desired. □

3.5 Semiconcavity of the cost

In order to prove the local semiconcavity of our cost, we use conservation of energy to obtain a bound on the speed curves of minimizers of the action, given by the next lemma. It basically states that the speed curve $(\gamma_M, \dot\gamma_M, \gamma_{\mathcal I})$ of a minimizer γ with initial points on a given compact $K \subset M \times \mathcal I$ is contained in a fixed compact of $TM \times \mathcal I$.

Lemma 56. *Let $K \subset M$ be a compact subset of M and assume $\gamma = (\gamma_M, \gamma_{\mathcal I})$ is a minimizer for $\mathcal J_t$, given by (3.2), with $\gamma_M(0) = x \in K$, $\gamma_M(t) = y \in K$. Then, there exist a compact set $\tilde K \subset TM$ and a constant $C > 0$ such that $\gamma_M(s) \in \tilde K$ and $\|\dot\gamma_M(s)\|_{\gamma_M(s)} \le C$, for every $s \in [0, t]$.*

Proof. Assume, without loss of generality, $K = \overline{B_R(x_0)}$ (observe this ball is compact, since our metric g is complete). Clearly, $d(x, y) \le 2R$ and then the constant speed geodesic $\alpha : [0,t] \to M$ connecting x to y whose length is $d(x,y)$ is completely contained in the ball $\overline{B_{3R}(x_0)}$ and

$$\|\dot{\alpha}(s)\|_{\alpha(s)} = \frac{d(x,y)}{t} \leq \frac{2R}{t}.$$

Note that the Lagrangian $L(\cdot,\cdot,i)$ is bounded on the compact set

$$\mathcal{L} := \left\{ (z,v) \ ; \ d(z,x_0) \leq 3R, \|v\|_z \leq \frac{2R}{t} \right\} \subset TM,$$

for all $i \in \mathcal{I}$. Say $L(z,v,i) \leq B$, for every $(z,v,i) \in \mathcal{L} \times \mathcal{I}$. Then, in particular,

$$\int_0^t L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\big) \ ds \leq \int_0^t L\big(\alpha(s), \dot{\alpha}(s), \gamma_{\mathcal{I}}(s)\big) \ ds \leq Bt. \tag{3.21}$$

Now, by the third condition in the Definition 7 of a Tonelli Lagrangian, there exists $C \in \mathbb{R}$ such that

$$Ct + \int_0^t \|\dot{\gamma}_M(s)\|_{\gamma_M(s)} \ ds \leq \int_0^t L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\big) \ ds. \tag{3.22}$$

By putting together (3.21) and (3.22), we get

$$\frac{1}{t} \int_0^t \|\dot{\gamma}_M(s)\|_{\gamma_M(s)} \ ds \leq B - C =: C_2. \tag{3.23}$$

Then, for some $\bar{s} \in (0,t)$, $\|\dot{\gamma}_M(\bar{s})\|_{\gamma_M(\bar{s})} \leq C_2$, and for every $s \in [0,t]$,

$$\gamma_M(s) \in B_{tC_2}(\gamma_M(0)) \subset B_{tC_2+3R}(x_0) \equiv \bar{K}.$$

Now, the energy E given by (3.17) is bounded on compact sets, so if we set θ the maximum of E on the compact set

$$\big\{(x,v,i) \in TM \times \mathcal{I} \ ; \ x \in \bar{K} \ , \ \|v\|_x \leq C_2\big\},$$

we get $E\big(\gamma_M(\bar{s}), \dot{\gamma}_M(\bar{s}), \gamma_{\mathcal{I}}(\bar{s})\big) \leq \theta$. Then, by the conservation of energy (see Proposition 55), $E\big(\gamma_M(t), \dot{\gamma}_M(t), \gamma_{\mathcal{I}}(t)\big) \leq \theta$, for every $s \in [0,t]$ and the speed curve of γ_M is contained in the compact

$$\big\{(x,v) \in TM \ ; \ x \in \bar{K}, \ E(x,v,i) \le \theta, \ \forall i\big\}.$$

Thus, there exists $\tilde{K}$ compact, and a constant C such that $\gamma_M([0,t]) \subset \tilde{K}$ and $\|\dot{\gamma}_M(s)\| \le C$, for all $s \in [0,t]$. $\square$

The local semiconcavity of c now follows as in the single Lagrangian case (cf. [16]).

Proposition 57. *For any $i, j \in \mathcal{I}$, the cost function $\bar{h}_t = h_t(\cdot, i, \cdot, j) : M \times M \to \mathbb{R}$, where h_t is given by (3.3), is locally semiconcave.*

Proof. To simplify the notation, we consider the case $M = \mathbb{R}^n$. The general case follows by the same reasoning when written in charts. For $x_1, x_2 \in B_1(0)$, let $\gamma \in AC \times \mathcal{P}$ be such that $\gamma_M(0) = x_1$, $\gamma_M(t) = x_2$ and

$$\begin{aligned}\bar{h}_t(x_1,x_2) = &\int_0^t L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\big)\ ds + \psi\big(x_1, i, \gamma_{\mathcal{I}}(0)^+\big) \\ &+ \sum_{i=0}^{N-1} \psi\big(\gamma_M(t_{i+1}), \gamma_{\mathcal{I}}(t_i), \gamma_{\mathcal{I}}(t_{i+1})\big) + \psi\big(x_2, \gamma_{\mathcal{I}}(t)^-, j\big)\end{aligned} \tag{3.24}$$

that is, γ is an action minimizer from (x_1, i) to (x_2, j). We show that $\bar{h}_t$ is semiconcave in $B_1(x_1) \times B_1(x_2)$. By Lemma 56, there exist $K \subset M$ compact and C constant such that

$$\gamma_M([0,t]) \subset K \text{ and } \|\dot{\gamma}_M(s)\| \le C \text{ for all } s \in [0,t].$$

We then choose $\varepsilon > 0$ such that $C\varepsilon < 1$. This implies

$$\gamma_M([0,\varepsilon]) \subset B_2(x_1) \text{ and } \gamma_M([t-\varepsilon,t]) \subset B_2(x_2).$$

For $y_1, y_2 \in B_2$, we define

$$\tilde{\gamma}_M(s) = \begin{cases} \dfrac{\varepsilon - s}{\varepsilon}\ y_1 + \gamma_M(s)\ , & \text{for } s \in [0, \varepsilon] \\ \gamma(s)\ , & \text{for } s \in [\varepsilon, t - \varepsilon] \\ \dfrac{s - t + \varepsilon}{\varepsilon}\ y_2 + \gamma_M(s)\ , & \text{for } s \in [t - \varepsilon, t]. \end{cases} \tag{3.25}$$

Observe that $\tilde{\gamma}_M$ and $\dot{\tilde{\gamma}}_M$ are bounded as well. Since $\tilde{\gamma}_M(0) = x_1 + y_1$ and $\tilde{\gamma}_M(t) = x_2 + y_2$, and we can take $\varepsilon > 0$ sufficiently small so that $\varepsilon < \min\{t_1, t - t_N\}$,

$$\begin{aligned} \bar{h}_t(x_1 + y_1,\, x_2 + y_2) - h_t(x_1, x_2) \leq &\int_0^t L(\tilde{\gamma}_M, \dot{\tilde{\gamma}}_M, \gamma_{\mathcal{I}})\ ds - \int_0^t L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}})\ ds \\ &+ \psi\big(x_1, i, \gamma_{\mathcal{I}}(0)^+\big) + \sum_{i=0}^{N-1} \psi\big(\tilde{\gamma}_M(t_{i+1}), \gamma_{\mathcal{I}}(t_i), \gamma_{\mathcal{I}}(t_{i+1})\big) \\ &+ \psi\big(x_2, \gamma_{\mathcal{I}}(t)^-, j\big) - \psi\big(x_1, i, \gamma_{\mathcal{I}}(0)^+\big) \\ &- \sum_{i=0}^{N-1} \psi\big(\gamma_M(t_{i+1}), \gamma_{\mathcal{I}}(t_i), \gamma_{\mathcal{I}}(t_{i+1})\big) + \psi\big(x_2, \gamma_{\mathcal{I}}(t)^-, j\big) \\ = \int_0^\varepsilon L\Big(\frac{\varepsilon - s}{\varepsilon} y_1 + \gamma_M(s), -\frac{1}{\varepsilon} y_1 + \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\Big)\ ds \hspace{-6em} & \\ &+ \int_\varepsilon^{t-\varepsilon} L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}})\ ds \\ &+ \int_{t-\varepsilon}^t L\Big(\frac{s - t + \varepsilon}{\varepsilon} y_2 + \gamma_M(s), \frac{1}{\varepsilon} y_2 + \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\Big)\ ds \\ &- \int_0^t L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}})\ ds \\ &+ \psi\big(x_1 + y_1, i, \gamma_{\mathcal{I}}(0)^+\big) - \psi\big(x_1, i, \gamma_{\mathcal{I}}(0)^+\big) \\ &+ \psi\big(x_2 + y_2, \gamma_{\mathcal{I}}(t)^-, j\big) - \psi\big(x_2, \gamma_{\mathcal{I}}(t)^-, j\big). \end{aligned} \tag{3.26}$$

Thus,

$$
\begin{aligned}
&\bar{h}_t(x_1+y_1,x_2+y_2)-h_t(x_1,x_2)\\
&\quad\le \int_0^{\varepsilon}\Big[L\Big(\frac{\varepsilon-s}{\varepsilon}y_1+\gamma_M(s),-\frac{1}{\varepsilon}y_1+\dot\gamma_M(s),\gamma_{\mathcal{I}}(s)\Big)-L(\gamma_M,\dot\gamma_M,\gamma_{\mathcal{I}})\Big]\,ds\\
&\qquad+\int_{t-\varepsilon}^{t}\Big[L\Big(\frac{s-1+\varepsilon}{\varepsilon}y_2+\gamma_M(s),\frac{1}{\varepsilon}y_2+\dot\gamma_M(s),\gamma_{\mathcal{I}}(s)\Big)-L(\gamma_M,\dot\gamma_M,\gamma_{\mathcal{I}})\Big]\,ds\\
&\qquad+\psi\big(x_1+y_1,i,\gamma_{\mathcal{I}}(0)^+\big)-\psi\big(x_1,i,\gamma_{\mathcal{I}}(0)^+\big)+\psi\big(x_2+y_2,\gamma_{\mathcal{I}}(t)^-,j\big)\\
&\qquad-\psi\big(x_2,\gamma_{\mathcal{I}}(t)^-,j\big).
\end{aligned}
\tag{3.27}
$$

We know L and ψ are both C^2, so they are locally semiconcave. Thus, by choosing a common modulus of continuity for $L(\,\cdot\,,\,\cdot,i)$ and $\psi(\cdot,i,j)$, $i,j\in\mathcal{I}$, in the compact $\overline{B_4(x_1)\cup B_4(x_2)}\times\overline{B_C(x_1)\cup B_C(x_2)}$, we get

$$
\begin{aligned}
\bar{h}_t(x_1+y_1,x_2+y_2)-\bar{h}_t(x_1,x_2)&\le F_t(y_1,y_2)+\|v\|\,w(\|v\|)+\|z\|\,w(\|z\|)\\
&\le F_t(y_1,y_2)+\|(v,z)\|\,w(\|(v,z)\|),
\end{aligned}
\tag{3.28}
$$

where

$$
\begin{aligned}
F_t(y_1,y_2):=&\int_0^{\varepsilon}\Big(\frac{\varepsilon-s}{\varepsilon}\frac{\partial L}{\partial x}\big(\gamma_M,\dot\gamma_M,\gamma_{\mathcal{I}}\big)\cdot y_1-\frac{1}{\varepsilon}\frac{\partial L}{\partial v}\big(\gamma_M,\dot\gamma_M,\gamma_{\mathcal{I}}\big)\cdot y_1\Big)\,ds\\
&+\int_{t-\varepsilon}^{t}\Big(\frac{s-1+\varepsilon}{\varepsilon}\,\frac{\partial L}{\partial x}\big(\gamma_M,\dot\gamma_M,\gamma_{\mathcal{I}}\big)\cdot y_2+\frac{1}{\varepsilon}\frac{\partial L}{\partial v}\big(\gamma_M,\dot\gamma_M,\gamma_{\mathcal{I}}\big)\cdot y_2\Big)\,ds\\
&+\partial_x\psi\big(x_1,i,\gamma_{\mathcal{I}}(0)^+\big)\cdot y_1+\partial_x\psi\big(x_2,\gamma_{\mathcal{I}}(t)^-,j\big)\cdot y_2
\end{aligned}
\tag{3.29}
$$

is linear, and

$$
v:=\Big(\frac{\varepsilon-s}{\varepsilon}\,y_1,-\frac{1}{\varepsilon}\,y_1\Big),\qquad z:=\Big(\frac{s-t+\varepsilon}{\varepsilon}\,y_2,\frac{1}{\varepsilon}\,y_2\Big).
$$

It follows that h_t is locally semiconcave. □

Remark 58. Proposition 57 implies in particular that $\bar{h}_t$ is differentiable almost everywhere.

Next, we show how Proposition 57 yields an important characterization of the superdifferential of $\bar{h}_t$.

Corollary 59. *For any action minimizer $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ satisfying $\gamma_M(0) = x$ and $\gamma_M(t) = y$, the linear functional F_t given by (3.29) is a superdifferential of $\bar{h}_t$ at the point (x, y), and it can be written as*

$$\begin{aligned}F_t(y_1, y_2) =& \frac{\partial L}{\partial v}\big(\gamma_M(t), \dot{\gamma}_M(t), \gamma_{\mathcal{I}}(t)^-\big) \cdot y_2 - \frac{\partial L}{\partial v}\big(\gamma_M(0), \dot{\gamma}_M(0), \gamma_{\mathcal{I}}(0)^+\big) \cdot y_1 \\ &+ \partial_x\psi\big(\gamma_M(0), i, \gamma_{\mathcal{I}}(0)^+\big) \cdot y_1 + \partial_x\psi\big(\gamma_M(t), \gamma_{\mathcal{I}}(t)^-, j\big) \cdot y_2.\end{aligned} \tag{3.30}$$

In particular, if $\bar{h}_t$ is differentiable at (x, y), then

$$\begin{aligned}d_{(x,y)}\bar{h}_t \cdot (y_1, y_2) =& \frac{\partial L}{\partial v}\big(\gamma_M(t), \dot{\gamma}_M(t), \gamma_{\mathcal{I}}(t)^-\big) \cdot y_2 - \frac{\partial L}{\partial v}\big(\gamma_M(0), \dot{\gamma}_M(0), \gamma_{\mathcal{I}}(0)^+\big) \cdot y_1 \\ &+ \partial_x\psi\big(\gamma_M(0), i, \gamma_{\mathcal{I}}(0)^+\big) \cdot y_1 + \partial_x\psi\big(\gamma_M(t), \gamma_{\mathcal{I}}(t)^-, j\big) \cdot y_2.\end{aligned} \tag{3.31}$$

Proof. Clearly, F_t given by (3.29) is a superdifferential for $\bar{h}_t$. We now prove the representation formula (3.30). By Theorem 49, the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial L}{\partial v}\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\big) = \frac{\partial L}{\partial x}\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\big) \tag{3.32}$$

is satisfied in the set $V \equiv [0, t_1) \cup (t_1, t_2) \cup \cdots \cup (t_{N-1}, t_N) \cup (t_N, 1]$. Substituting (3.32) in (3.29) and integrating by parts (recall $\varepsilon < \min\{t_1, t - t_N\}$), we get

$$\begin{aligned}F_t(y_1, y_2) \;=\;& -\frac{\partial L}{\partial v}\big(\gamma_M(0), \dot{\gamma}_M(0)^+, \gamma_{\mathcal{I}}(0)^+\big) \cdot y_1 \\ &+\frac{\partial L}{\partial v}\big(\gamma_M(t), \dot{\gamma}_M(t)^-, \gamma_{\mathcal{I}}(t)^-\big) \cdot y_2 \\ &+\partial_x\psi\big(\gamma_M(0), i, \gamma_{\mathcal{I}}(0)^+\big) \cdot y_1 + \partial_x\psi\big(\gamma_M(t), \gamma_{\mathcal{I}}(t)^-, j\big) \cdot y_2\end{aligned}$$

Then, apply Theorem 52 to conclude the proof of the corollary. □

Remark 60. Observe that the terms $\partial_x\psi$ disappear in case we have no switches at the endpoints.

In order to finish the proof of Theorem 47, we only need to prove the following proposition:

Proposition 61. *Suppose $L : TM \to \mathbb{R}$ is a Tonelli Lagrangian satisfying A1 and*

A2, and $\psi : M \times \mathcal{I} \times \mathcal{I} \to \mathbb{R}$ a switching cost satisfying A3 and A4. Then, for any $x, y \in M$ and any $i, j, k \in \mathcal{I}$,

$$h_t\big((x,i),(y,j)\big) - h_t\big((x,i),(y,k)\big) \leq \psi(x,k,j).$$

Proof. Let $\gamma \in AC([0,t];M) \times \mathcal{P}([0,t];\mathcal{I})$ be such that $\gamma(0) = (x,i)$, $\gamma(t) = (y,k)$, and

$$h_t\big((x,i),(y,k)\big) = \mathcal{J}_t[\gamma].$$

Define $\tilde{\gamma} : [0, t+\delta] \to M \times \mathcal{I}$ by

$$\tilde{\gamma}(s) = \begin{cases} \gamma(s), & \text{in } [0,t); \\ (x,j), & \text{in } [t, t+\delta]. \end{cases} \tag{3.33}$$

Then

$$h_{t+\delta}\big((x,i),(y,j)\big) \leq \mathcal{J}_{t+\delta}[\tilde{\gamma}] = h_t\big((x,i),(y,k)\big) + \delta L(x,0,j) + \psi(x,k,j).$$

Let $\delta \to 0$ to conclude the proof. $\square$

Chapter 4

Weakly coupled systems of Hamilton–Jacobi equations

The purpose of this chapter is to analyze aspects of the weak KAM theory in the optimal switching setting. Let $L : TM \to \mathbb{R}$ be a Tonelli Lagrangian, as in Definition 7, for which conditions $A1$–$A5$ are valid. By the Legendre–Fenchel duality, we associate the *Hamiltonian* $H : T^*M \times \mathcal{I} \to \mathbb{R}$ by

$$H(x, p, i) := \sup_{v \in T_xM} \big\{ p(v) - L(x, v, i) \big\}. \tag{4.1}$$

As in the case of a single Lagrangian, it follows that each $H(\cdot, \cdot, i)$ is a Tonelli Hamiltonian, in the sense of Definition 14. Furthermore, we have

A2'. $H(\cdot, \cdot, i)$ is superlinear above compact sets, uniformly in $i \in \mathcal{I}$, meaning that, given $K \subset M$ compact, for every constant $A \geq 0$, there exists a constant $C \in \mathbb{R}$, depending on K and A, such that

$$H(x, p, i) \geq A \, \|p\|_x^* + C, \text{ for all } x \in K, \ p \in T_x^*M;$$

As we will see in Section 4.1 below, the optimal switching problem is closely related to the following system of Hamilton–Jacobi equations:

$$\max \Big\{ H\big(x, du(x, i), i\big) - c, \ \max_{j \neq i} \big\{ u(x, i) - u(x, j) - \psi(x, i, j) \big\} \Big\} = 0, \tag{4.2}$$

for every $i \in \mathcal{I}$, and for some $c \in \mathbb{R}$. If we define $\Psi u : M \times \mathcal{I} \to \mathbb{R}$ by

$$\Psi u(x,i) = \min_{j \neq i} \big\{ u(x,j) + \psi(x,i,j) \big\}, \tag{4.3}$$

we can write, more concisely,

$$\max \Big\{ H\big(x, du(x,i), i\big) - c,\ u(x,i) - \Psi u(x,i) \Big\} = 0, \tag{4.4}$$

for every $i \in \mathcal{I}$, and for some $c \in \mathbb{R}$. We prove the following theorem:

Theorem 62 (Weak KAM)**.** *There exists a unique $c_0 \in \mathbb{R}$ for which the weakly coupled system of Hamilton–Jacobi equations*

$$\max \Big\{ H\big(x, du(x,i), i\big) - c_0,\ u(x,i) - \Psi u(x,i) \Big\} = 0 \tag{4.5}$$

admits a viscosity solution[1] $u : M \times \mathcal{I} \to \mathbb{R}$. *We call such a number c_0 a (generalized) Mañé critical value.*

Our proof of this theorem utilizes Fathi's idea [13] of understanding the long–time behavior of the Lax–Oleinik semigroup associated to a viscosity subsolution of (4.5).

As we have mentioned before, we are not the first ones to consider such systems; in fact, in a similar fashion, it is present in the early work of Capuzzo Dolcetta-Evans [6] (see also Gomes-Serra [19]).

Here is how this chapter is organized: In Section 4.1 we define the Lax–Oleinik semigroup associated to the optimal switching problem of last chapter, and prove some of its properties. We also define viscosity solutions of the system (4.4) and explain how it relates to the Lax–Oleinik semigroup. Next, in Section 4.2, we prove a generalized version of Fathi's weak KAM theorem, for the system studied in Section 4.1. Finally, in Section 4.3, we define the Aubry set following [19] and prove the results we need for the long time behavior of Lax–Oleinik of net chapter.

[1] We define viscosity solutions of (4.5) in Definition 66 below.

4.1 Lax–Oleinik semigroup and viscosity solutions

In this section we define a Lax–Oleinik semigroup (see Definition 63 below) and we show how it provides a variational formulation for viscosity solutions of the weakly coupled system (4.4). First, we give a motivation for the definition of our semigroup.

Motivation. Assume u is a smooth solution of (4.4), and let $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) : [0,t] \to M \times \mathcal{I}$ be such that $\gamma(t) = (x,i) = A$, and $\dot{\gamma}_M(t) = v$, for some $v \in T_xM$. Thus, from (4.4), we can write

$$H\Big(x, du\big(\gamma(t)\big), i\Big) \leq c,$$

and by the definition of the Hamiltonian $H(\cdot,\cdot,i)$, we have

$$\frac{d}{dt}u\big(\gamma(t)\big) = du\big(\gamma(t)\big) \cdot \dot{\gamma}_M(t) \leq L\big(\gamma_M(t), \dot{\gamma}_M(t), \gamma_{\mathcal{I}}(t)\big) + c.$$

Thus, by integrating the last inequality from t_N to t, we get

$$u\big(\gamma_M(t), i\big) \leq u\big(\gamma_M(t_N), i\big) + \int_{t_N}^{t} \big[L\big(\gamma_M(s), \dot{\gamma}_M(s), i\big) + c\big]\, ds.$$

Now, (4.4) also gives us

$$u\big(\gamma_M(t_N), i\big) \leq u\big(\gamma_M(t_N), \gamma_{\mathcal{I}}(t_{N-1})\big) + \psi\big(\gamma_M(t_N), \gamma_{\mathcal{I}}(t_{N-1}), \gamma_{\mathcal{I}}(t_N)\big)$$

so that

$$\begin{aligned} u\big(\gamma_M(t), i\big) \leq u\big(\gamma_M(t_N), \gamma_{\mathcal{I}}(t_{N-1})\big) &+ \int_{t_N}^{t} \big[L\big(\gamma_M(s), \dot{\gamma}_M(s), i\big) + c\big]\, ds \\ &+ \psi\big(\gamma_M(t_N), \gamma_{\mathcal{I}}(t_{N-1}), \gamma_{\mathcal{I}}(t_N)\big). \end{aligned} \tag{4.6}$$

By iterating this argument, we get that

$$
\begin{aligned}
u(x,i) \leq u\big(\gamma_M(0),\gamma_{\mathcal{I}}(0)\big) &+ \sum_{k=0}^{N}\int_{t_k}^{t_{k+1}} \big[L\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(t_k)\big)+c\big]\ ds \\
&+ \sum_{k=0}^{N-1}\psi\big(\gamma_M(t_{k+1}),\gamma_{\mathcal{I}}(t_k),\gamma_{\mathcal{I}}(t_{k+1})\big).
\end{aligned}
\tag{4.7}
$$

Furthermore,

$$
\begin{aligned}
u(x,i) \leq \inf\Bigg\{u\big(\gamma_M(0),\gamma_{\mathcal{I}}(0)\big) &+ \sum_{k=0}^{N}\int_{t_k}^{t_{k+1}} \big[L\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(t_k)\big)+c\big]\ ds \\
&+ \sum_{k=0}^{N-1}\psi\big(\gamma_M(t_{k+1}),\gamma_{\mathcal{I}}(t_k),\gamma_{\mathcal{I}}(t_{k+1})\big)\Bigg\},
\end{aligned}
\tag{4.8}
$$

where the infimum is taken over all absolutely continuous curves $\gamma_M \in AC([0,t];M)$ with $\gamma_M(t)=x$ and all piecewise constant functions

$$
\gamma_{\mathcal{I}} = \sum_{k=0}^{N-1}\gamma_{\mathcal{I}}(t_k)\chi_{[t_k,t_{k+1})} \in \mathcal{P}
$$

with $\gamma_{\mathcal{I}}(t_N)=i$. Note that up to now we have only utilized that the maximum in (4.4) is less than or equal to zero. Next, we show that the right hand side of (4.8) actually provides the variational formulation of a viscosity solution of (4.4), as we wanted.

Definition 63. [Lax–Oleinik semigroup] For a given function $u : M\times\mathcal{I}\to\mathbb{R}$, where $\mathcal{I}=\{1,\dots,m\}$, we define the *Lax–Oleinik semigroup* associated to the Lagrangian L by $T_t u : M\times\mathcal{I}\to\mathbb{R}$,

$$
\begin{aligned}
T_t u(A) := \inf\Bigg\{u\big(\gamma(0)\big) &+ \sum_{k=0}^{N}\int_{t_k}^{t_{k+1}} L\big(\gamma_M(s),\dot\gamma_M(s),\gamma_{\mathcal{I}}(t_k)\big)\ ds \\
&+ \sum_{k=0}^{N-1}\psi\big(\gamma_M(t_{k+1}),\gamma_{\mathcal{I}}(t_k),\gamma_{\mathcal{I}}(t_{k+1})\big)\Bigg\},
\end{aligned}
\tag{4.9}
$$

where the infimum is taken over all absolutely continuous curves $\gamma_M \in AC([0,t];M)$ and all piecewise constant functions

$$\gamma_{\mathcal{I}} = \sum_{k=0}^{N-1} \gamma_{\mathcal{I}}(t_k)\chi_{[t_k,t_{k+1})} \in \mathcal{P}$$

with $\gamma(t) = \big(\gamma_M(t), \gamma(t)\big) = (x,i) = A$, as usual.

Remark 64. If we define $h_t : M \times \mathcal{I} \times M \times \mathcal{I} \to \mathbb{R}$ to be

$$h_t(B,A) := \inf \int_0^t L\big(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}}\big) + \sum_{k=0}^{N-1} \psi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})\big), \tag{4.10}$$

where $\gamma = (\gamma_M, \gamma_{\mathcal{I}}) \in AC \times \mathcal{P}$ satisfies $\gamma(0) = B$, $\gamma(t) = A$, then we can more concisely write

$$T_t u(A) = \inf_{B \in M \times \mathcal{I}} \big\{u(B) + h_t(B,A)\big\}.$$

Proposition 65. *For $u \in C(M \times \mathcal{I})$, $t > 0$, and $i \in \mathcal{I}$, the map $x \mapsto T_t u(x,i)$ is semiconcave. In particular, when written in charts, we have*

$$T_t u(x+h,i) - T_t u(A) \le \frac{\partial L}{\partial v}\big(\gamma_M(t), \dot\gamma_M(t), \gamma_{\mathcal{I}}(t)^-\big) \cdot h + K\,\|h\|^2 \tag{4.11}$$

and consequently

$$\frac{\partial L}{\partial v}\big(\gamma_M(t), \dot\gamma_M(t), \gamma_{\mathcal{I}}(t)^-\big) \in \partial^+\big(T_t u\big)(x,i).$$

Proof. By compactness, there exists $B \in M \times \mathcal{I}$ such that

$$T_t u(A) = u(B) + h_t(B,A).$$

Thus,

$$T_t u(x+h,i) - T_t u(x,i) \le h_t\big(B,(x+h,i)\big) - h_t(B,A).$$

Hence, by reasoning as in Proposition 57, we have

$$T_t u(x+h,i) - T_t u(A) \le \frac{\partial L}{\partial v}\big(\gamma_M(t), \dot\gamma_M(t), \gamma_{\mathcal{I}}(t)^-\big) \cdot h + K\,\|h\|^2 \tag{4.12}$$

as desired. $\square$

Let us now define what we mean by a viscosity solution of the system (4.4). The following extends naturally the concept of a viscosity solution as introduced by Crandall-Lions [9] (see also [6, 19]).

Definition 66. 1. We say that the continuous function $u : M \times \mathcal{I} \to \mathbb{R}$ is a *viscosity subsolution* to the system (4.4) if, for any $x \in M$ and $i \in \mathcal{I}$, whenever the function $\phi : M \to \mathbb{R}$ is such that $u(\cdot, i) - \phi$ admits a maximum at x, we have

$$\max \Big\{ H\big(x, d\phi(x), i\big) - c, u(x,i) - \Psi u(x,i) \Big\} \leq 0. \tag{4.13}$$

2. We say that $u : M \times \mathcal{I} \to \mathbb{R}$ is a *viscosity supersolution* to the system (4.4) if, for any $x \in M$ and $i \in \mathcal{I}$, whenever the function $\phi : M \to \mathbb{R}$ is such that $u(\cdot, i) - \phi$ admits a minimum at x, we have

$$\max \Big\{ H\big(x, d\phi(x), i\big) - c, u(x,i) - \Psi u(x,i) \Big\} \geq 0. \tag{4.14}$$

3. The function u is said to be a *viscosity solution* if it is both viscosity subsolution and a viscosity supersolution.

Remark 67. For $u : M \times \mathcal{I} \to \mathbb{R}$ to be a viscosity subsolution to the system (4.4), it is necessary and sufficient that the following two conditions are satisfied:

(i) the function $u(\cdot, i)$ satisfies

$$H\big(x, du(x,i), i\big) \leq c, \tag{4.15}$$

in the viscosity sense, and;

(ii) the function $u(\cdot, i)$ satisfies

$$u(x,i) \leq \Psi u(x,i) = \min_{j \neq i} \big\{ u(x,j) + \psi(x,i,j) \big\}.$$

Analogously, for $u : M \times \mathcal{I} \to \mathbb{R}$ to be a viscosity supersolution to the system (4.4), it is necessary and sufficient that *at least one* of following two conditions are satisfied:

(i) the function $u(\cdot, i)$ satisfies

$$H\big(x, du(x,i), i\big) \geq c, \tag{4.16}$$

in the viscosity sense, or;

(ii) the function $u(\cdot, i)$ satisfies

$$u(x,i) \geq \Psi u(x,i).$$

Before providing an important connection between our Lax–Oleinik semigroup and the notion solutions of (4.4) defined above, we prove a few lemmas.

Lemma 68. *Let $v \in C(M \times \mathcal{I})$ and $A = (x,i) \in M \times \mathcal{I}$. Then there exists $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ minimizer for the Lax–Oleinik operator:*

$$T_t v(A) = v\big(\gamma(0)\big) + \mathcal{J}_t[\gamma]. \tag{4.17}$$

Moreover, if v is a subsolution of (4.4), then we can choose γ to have no switch at the initial time.

Proof. Let $A = (x,i) \in M \times \mathcal{I}$. Recall

$$T_s v(A) = \inf_{B \in M \times \mathcal{I}} \big\{v(B) + h_t(B,A)\big\}.$$

By compactness, and by the existence theorem of Section 2, there exists $B = (y,j) \in M \times \mathcal{I}$ and $\gamma : [0,s] \to M \times \mathcal{I}$, with $\gamma_M(0) = y$, $\gamma_M(s) = x$, such that

$$\begin{aligned} T_s v(A) &= v(B) + h_s(B,A) \\ &= v(\gamma(0)) + \int_0^s L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) + \psi\big(\gamma_M(0), \gamma_{\mathcal{I}}(0), \gamma_{\mathcal{I}}(0)^+\big) \\ &\quad + \sum_{\ell=1}^{N-1} \psi\big(\gamma_M(t_{\ell+1}), \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big) + \psi(x, \gamma_{\mathcal{I}}(t)^-, \gamma_{\mathcal{I}}(t)). \end{aligned} \tag{4.18}$$

We now claim that, when v is a subsolution, γ might be chosen so that $\gamma_{\mathcal{I}}(0) = \gamma_{\mathcal{I}}(0)^+ = j$. Indeed, if we set $\tilde{B} = \big(\gamma_M(0), \gamma_{\mathcal{I}}(0)^+\big)$, we have

$$\begin{aligned} T_s v(A) &\leq v(\tilde{B}) + h_t(\tilde{B}, A) \\ &= v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0)^+\big) + \int_0^s L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) \\ &\quad + \sum_{\ell=1}^{N-1} \psi\big(\gamma_M(t_{\ell+1}), \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big) + \psi(x, \gamma_{\mathcal{I}}(t)^-, \gamma_{\mathcal{I}}(t)), \end{aligned} \tag{4.19}$$

so that (4.18) and (4.19) give

$$v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0)\big) + \psi\big(\gamma_M(0), \gamma_{\mathcal{I}}(0), \gamma_{\mathcal{I}}(0)^+\big) \leq v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0)^+\big).$$

Since v is a subsolution, the other inequality is also true, and

$$v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0)\big) + \psi\big(\gamma_M(0), \gamma_{\mathcal{I}}(0), \gamma_{\mathcal{I}}(0)^+\big) = v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0)^+\big),$$

that is, we may choose not to switch at time $t = 0$. $\square$

Lemma 69. *Let v be a subsolution of (4.4) and let $(x, i) \in M \times \mathcal{I}$. Then, for sufficiently small $s > 0$, we might have at most one switch.*

Proof. Since v is Lipschitz, there exists $K_1 > 0$ such that

$$|v(z, i) - v(y, i)| \leq K_1 d(z, y),$$

for any $z, y \in M$, and for all $i \in \mathcal{I}$. By superlinearity, for any $K_2 > 0$, there exists a constant $C \geq 0$ such that

$$(K_1 + K_2) \left\| v \right\|_x - C \leq L(x, v, i).$$

Thus, using that v is a subsolution, we have

$$
\begin{aligned}
T_s v(A) \geq\ & v\big(x, \gamma_{\mathcal{I}}(0)\big) + (K_1 + K_2) \int_0^s \|\dot\gamma_M(\tau)\|_{\gamma_M(\tau)}\, d\tau - Cs + v\big(\gamma_M(0), \gamma_{\mathcal{I}}(0)\big) \\
& - v\big(x, \gamma_{\mathcal{I}}(0)\big) + \sum_{\ell=1}^{N-1} \psi\big(\gamma_M(t_{\ell+1}), \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big) + \psi(x, \gamma_{\mathcal{I}}(t)^-, i) \\
\geq\ & v\big(x, i\big) - K_1 d\big(\gamma_M(0), x\big) + (K_1 + K_2) \int_0^s \|\dot\gamma_M(\tau)\|_{\gamma_M(\tau)}\, d\tau - Cs \\
& + \sum_{\ell=1}^{N-1} \psi\big(\gamma_M(t_{\ell+1}), \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big) + \psi(x, \gamma_{\mathcal{I}}(t)^-, i) - \psi(x, \gamma_{\mathcal{I}}(0), i) \\
\geq\ & v\big(x, i\big) + K_2 \int_0^s \|\dot\gamma_M(\tau)\|_{\gamma_M(\tau)}\, d\tau - Cs + \sum_{\ell=1}^{N-1} \psi\big(x, \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big) \\
& + \psi(x, \gamma_{\mathcal{I}}(t)^-, i) - \psi(x, \gamma_{\mathcal{I}}(0), i) + \sum_{\ell=1}^{N-1} \Big[\psi\big(\gamma_M(t_{\ell+1}), \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big) \\
& \qquad - \psi\big(x, \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big)\Big] \\
\geq\ & v\big(x, i\big) + K_2 \int_0^s \|\dot\gamma_M(\tau)\|_{\gamma_M(\tau)}\, d\tau - Cs + \sum_{\ell=1}^{N-1} \psi\big(x, \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big) \\
& + \psi(x, \gamma_{\mathcal{I}}(t)^-, i) - \psi(x, \gamma_{\mathcal{I}}(0), i) + K_3 \sum_{\ell=1}^{N-1} d\big(\gamma_{\mathcal{I}}(t_{\ell+1}), x\big).
\end{aligned}
\tag{4.20}
$$

Choose $K_2 \geq K_3(\tilde N - 1)$ (recall the number of switches of minimizing curves is uniformly bounded) so that

$$
\begin{aligned}
T_s v(A) \geq v\big(A\big) - Cs + \sum_{\ell=1}^{N-1} \psi\big(x, \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big) \\
+ \psi(x, \gamma_{\mathcal{I}}(t)^-, i) - \psi(x, \gamma_{\mathcal{I}}(0), i).
\end{aligned}
\tag{4.21}
$$

Then, if we would have two or more switches,

$$
\sum_{\ell=1}^{N-1} \psi\big(x, \gamma_{\mathcal{I}}(t_\ell), \gamma_{\mathcal{I}}(t_{\ell+1})\big) + \psi(x, \gamma_{\mathcal{I}}(t)^-, i) - \psi(x, \gamma_{\mathcal{I}}(0), i) \geq \delta_0 > 0; \tag{4.22}
$$

therefore

$$T_s v(A) \geq v\big(A\big) + \delta_0 - Cs, \tag{4.23}$$

a contradiction when $s > 0$ is sufficiently small. $\square$

Lemma 70. *Let v be a subsolution and let $(x,i) \in M \times \mathcal{I}$. Assume*

$$v(x,i) < \min_{j \neq i} \big\{ v(x,j) + \psi(x,i,j) \big\}. \tag{4.24}$$

Then, there exists $s_0 = s_0(x)$ such that

$$T_s v(x,i) = \inf_{\gamma_M(s)=x} \left\{ v(\gamma_M(0),i) + \int_0^s L(\gamma_M, \dot{\gamma}_M, i) \right\}, \ \forall s \in [0, s_0]. \tag{4.25}$$

In other words, it is not worth switching between modes.

Proof. Observe that (4.24) holds true in a neighborhood of x. We claim that $s > 0$ can be taken small so that any minimizing curve γ for $T_t(x,i)$ is completely contained in such neighborhood. In fact, $T_t u(y,i) \leq C$ for $(t,y) \in [0,t_0] \times M$. Then, for any $\varepsilon > 0$, there exists $C_\varepsilon \geq 0$ such that

$$-C + \frac{1}{\varepsilon} \int_0^s \|\dot{\gamma}_M\|_{\gamma_M} - C_\varepsilon s \leq u\big(\gamma(0)\big) + \int_0^s L\big(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}\big) \leq C, \tag{4.26}$$

which in turn implies

$$d\big(\gamma_M(0), x\big) \leq \int_0^s \|\dot{\gamma}_M\|_{\gamma_M} \leq C\varepsilon + C_\varepsilon \varepsilon s. \tag{4.27}$$

Hence, given any $\delta > 0$, we first choose $\varepsilon < \delta/2C$, then $s > 0$ sufficiently small so that $C_\varepsilon \varepsilon s \leq \delta/2$, and we have $d\big(\gamma_M(0), x\big) \leq \delta$.

Now, suppose $\gamma_{\mathcal{I}}(0) = j \neq i$, that is, we have one switch, for $s > 0$ small. Then, since

$$v\big(\gamma_M(0), j\big) + \psi\big(\gamma_M(0), j, i\big) \geq v\big(\gamma_M(0), i\big) + \delta_0, \tag{4.28}$$

for some $\delta_0 > 0$, the same computation from the previous lemma shows that

$$T_s v(A) \geq v\big(A\big) + \delta_0 - Cs, \tag{4.29}$$

where now, instead of (4.20), we use (4.28). This is again a contradiction and we must have no switches for all times $0 < s < s_0(x)$, if $s_0(x)$ is sufficiently small. $\square$

Next, we provide the aforementioned connection between our Lax–Oleinik semigroup and the viscosity solutions of (4.4).

Proposition 71. *A Lipschitz function $u : M \times \mathcal{I} \to \mathbb{R}$ is a viscosity subsolution of (4.4) if, and only if, $u \leq T_t u + ct$, for all $t \geq 0$.*

Proof. To simplify the notation we suppose without loss of generality that $c = 0$. If u is a subsolution, then, for any curve $\gamma : [0, t] \to M \times \mathcal{I}$ with $\gamma(t) = A$, we write

$$\begin{aligned}
u(x,i) - u\big(\gamma(0)\big) &= \sum_{k=0}^{N} \Big[u\big(\gamma(t_{k+1})\big) - u\big(\gamma(t_k)\big)\Big] \\
&= \sum_{k=0}^{N} \Big[u\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k)\big) - u\big(\gamma_M(t_k), \gamma_{\mathcal{I}}(t_k)\big)\Big] \\
&\quad + \sum_{k=0}^{N} \Big[u\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_{k+1})\big) - u\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k)\big)\Big] \\
&\leq \sum_{k=0}^{N} \int_{t_k}^{t_{k+1}} L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(t_k)\big)\, ds \\
&\quad + \sum_{k=0}^{N-1} \psi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})\big),
\end{aligned} \tag{4.30}$$

which implies $u \leq T_t u$.

Now, given (x_0, i_0), we take γ satisfying $\gamma(t_0) = (x_0, i_0)$, $\gamma_M(t) = x$, $\gamma_{\mathcal{I}} \equiv i_0$, and $\dot{\gamma}_M(t_0) = v$, so that

$$u(x_0, i_0) \leq u\big(\gamma_M(t), i_0\big) + \int_t^{t_0} L\big(\gamma_M, \dot{\gamma}_M, i_0\big)\, ds$$

and any test function ϕ such that $u(\cdot, i_0) - \phi$ has a maximum at x_0 satisfies

$$\frac{\phi(x_0) - \phi(\gamma_M(t))}{t_0 - t} \leq \frac{1}{t_0 - t} \int_t^{t_0} L\big(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}\big)\ ds;$$

hence

$$d\phi(x_0) \cdot v \leq L(x_0, v, i_0).$$

Since v is arbitrary, we obtain

$$H\big(x_0, d\phi(x_0), i_0\big) \leq 0.$$

To complete the proof, we observe that the Lax–Oleinik semigroup is a solution to a time–dependent weakly coupled system of Hamilton–Jacobi equations, see Proposition 88. It follows that the Lax–Oleinik semigroup satisfies

$$T_t u(x, i) \leq T_t u(x, j) + \psi(x, i, j).$$

Let $t \to 0$ to obtain

$$u(x, i) \leq u(x, j) + \psi(x, i, j).$$

□

Proposition 72. *A semiconcave function $u : M \times \mathcal{I} \to \mathbb{R}$ is a solution of* (4.4) *if, and only if, $u = T_t u$, $\forall t$.*

Proof. Again, to simplify the notation we suppose without loss of generality that $c = 0$. The 'if' part is the same as in Proposition 88, of Chapter 5. Assume now that $u : M \times \mathcal{I} \to \mathbb{R}$ is a solution of (4.4) and fix $t > 0$. This means that, for any $i \in \mathcal{I}$,

$$\max\Big\{H\big(x, du(x), i\big), u(x, i) - \Psi u(x, i)\Big\} = 0, \tag{4.31}$$

in the viscosity sense of Definition 66. Fix $x \in M$ and $i \in \mathcal{I}$. Let us consider two cases.

Case 1. $u(x, i) < \Psi u(x, i)$. In this case, such inequality is true in a neighborhood $\tilde{U}$ of x. We first claim that $u(\cdot, i) = T_s u(\cdot, i)$ in a neighborhood of x, for all sufficiently small times $0 < s < s(x)$. Indeed, the initial point of any minimizer must be contained in $\tilde{U}$, if $s > 0$ is chosen sufficiently small, and $T_s u(y, j)$ cannot have any switches, for any $y \in U \subset \tilde{U}$ (See Lemma 70). In this way, since we have

$$H\big(y, du(y), i\big) = 0 \text{ on } U$$

and u is semiconcave, the classical theory implies $u(y, i) = T_s u(y, i)$, for all $y \in U$. Now, set

$$A := \{s; u(y, i) = T_s u(y, i) \text{ in } U\}.$$

By continuity, this set is closed in $[0, +\infty)$. But the reasoning above (combined with the fact that $u \leq T_t u$) shows it is also open. Hence, $A = [0, +\infty)$.

Case 2. $u(x, i) = \Psi u(x, i)$. Then, we have

$$u(x, i) = u(x, j) + \psi(x, j, i), \tag{4.32}$$

for some $j \in \mathcal{I} \backslash \{i\}$. It is not hard to see that j is such that

$$u(x, j) < \Psi u(x, j);$$

thus, by the same reasoning of case 1 (and by using the equation for the index j), we must have $u(x, j) = T_t u(x, j)$. This, together with (4.32) and Proposition 88, implies

$$u(x, i) = u(x, j) + \psi(x, j, i) = T_t u(x, j) + \psi(x, j, i) \geq T_t u(x, i).$$

The other inequality follows from the previous proposition and the fact that u is a subsolution. $\square$

Until the end of this section, we list some of the various useful properties that the Lax–Oleinik semigroup, as we defined above, satisfy.

Proposition 73. *For every $i \in \mathcal{I}$, the map $u \mapsto T_t u(\cdot, i)$ is a weak contraction in the $L^\infty(M)$–norm, that is, for every $g, h : M \times \mathcal{I} \to \mathbb{R}$ continuous,*

$$\|T_t g(\cdot, i) - T_t h(\cdot, i)\|_{L^\infty(M)} \leq \|g(\cdot, i) - h(\cdot, i)\|_{L^\infty(M)}.$$

Proof. The definition of the semigroup readily implies $T_t(g+c)(\cdot,i) = T_t g(\cdot,i)+c$, for $c \in \mathbb{R}$. Also, if $g(\cdot,i) \le h(\cdot,i)$, for every $i \in \mathcal{I}$, then $T_t g(\cdot,i) \le T_t h(\cdot,i)$. So, from

$$h(\cdot,i) - \|g(\cdot,i) - h(\cdot,i)\|_{L^\infty(M)} \le g(\cdot,i) \le h(\cdot,i) + \|g(\cdot,i) - h(\cdot,i)\|_{L^\infty(M)}$$

we conclude

$$T_t h(\cdot,i) - \|g(\cdot,i) - h(\cdot,i)\|_{L^\infty(M)} \le T_t g(\cdot,i) \le T_t h(\cdot,i) + \|g(\cdot,i) - h(\cdot,i)\|_{L^\infty(M)},$$

as wanted. □

Proposition 74. *The Lax–Oleinik semigroup defined above is in fact a semigroup, that is,*

$$T_t(T_s u) = T_{t+s}(u), \ \forall\ t,s \ge 0. \tag{4.33}$$

Proof. First, we claim that h_t, as in Remark 64, satisfies

$$h_{t+s}(A,B) = \inf_{C\in M\times\mathcal{I}} \big\{h_t(A,C) + h_s(C,B)\big\}. \tag{4.34}$$

Indeed, for a given $C \in M \times \mathcal{I}$, consider

$$\gamma^1 : [0,t] \to M \times \mathcal{I},$$
$$\gamma^2 : [0,s] \to M \times \mathcal{I},$$

with $\gamma^1(0) = A$, $\gamma^1(t) = C$, and $\gamma^2(0) = C$, $\gamma^2(t) = B$ so that, if we define $\gamma^3 : [0,t+s] \to M \times \mathcal{I}$ by the formula

$$\gamma^3(\tau) := \begin{cases} \gamma^1(\tau), & \text{for } \tau \in [0,t]; \\ \gamma^2(\tau - t), & \text{for } \tau \in [t,t+s], \end{cases} \tag{4.35}$$

and denoting with t^3_k the new partition that arises, we get

$$\begin{aligned}h_{t+s}(A,B) &\leq \int_0^{t+s} L(\gamma^3_M,\dot\gamma^3_M,\gamma^3_{\mathcal{I}}) + \sum_{i=0}^{N_1+N_2} \psi(\gamma^3_M(t^3_{i+1}),\gamma^3_{\mathcal{I}}(t^3_i),\gamma^3_{\mathcal{I}}(t^3_{i+1}))\\ &= \int_0^{t} L(\gamma^1_M,\dot\gamma^1_M,\gamma^1_{\mathcal{I}}) + \sum_{i=0}^{N_1-1} \psi\big(\gamma_M(t^1_{i+1}),\gamma^1_{\mathcal{I}}(t^1_i),\gamma^1_{\mathcal{I}}(t^1_{i+1})\big) \qquad (4.36)\\ &\quad + \int_0^{s} L(\gamma^2_M,\dot\gamma^2_M,\gamma^2_{\mathcal{I}}) + \sum_{i=0}^{N_2-1} \psi\big(\gamma_M(t^2_{i+1}),\gamma^2_{\mathcal{I}}(t^1_i),\gamma^2_{\mathcal{I}}(t^2_{i+1})\big).\end{aligned}$$

(note $\psi\big(\sigma^3(t_{N_1}),\sigma^3(t)\big)=0$) Then, by taking the infimum over γ^1,γ^2 we obtain

$$h_{t+s}(A,B) \leq h_t(A,C) + h_s(C,B),$$

and since this is true for all $C \in M\times\mathcal{I}$, we infer

$$h_{t+s}(A,B) \leq \inf_{C\in M\times\mathcal{I}} \big\{h_t(A,C)+h_s(C,B)\big\}.$$

For the reverse inequality, we fix $\varepsilon>0$, and choose $\gamma:[0,t+s]\to M\times\mathcal{I}$ such that $\gamma(0)=A$, $\gamma(t+s)=B$ for which

$$h_{t+s}(A,B)+\varepsilon \geq \int_0^{t+s} L(\gamma_M,\dot\gamma_M,\gamma_{\mathcal{I}}) + \sum_{i=0}^{N-1} \psi\big(\gamma_M(t_{i+1}),\gamma_{\mathcal{I}}(t_i),\gamma_{\mathcal{I}}(t_{i+1})\big). \qquad (4.37)$$

Let $\gamma(t)=(z,k)$ be the intermediate point of γ at time t. Now, we define $\gamma^1 : [0,t]\to M\times\mathcal{I}, \gamma^2:[0,s]\to M\times\mathcal{I}$, respectively, by

$$\gamma^1 = \gamma|_{[0,t]}$$

and

$$\gamma^2(\tau) = \gamma|_{[t,t+s]}(\tau+t)$$

Observe that

$$h_t(A,\gamma(t)) \leq \int_0^{t} L(\gamma^1_M,\dot\gamma^1_M,\gamma^1_{\mathcal{I}}) + \sum_{i=0}^{N_1-1} \psi\big(\gamma_M(t^1_{i+1}),\gamma^1_{\mathcal{I}}(t^1_i),\gamma^1_{\mathcal{I}}(t^1_{i+1})\big), \qquad (4.38)$$

and

$$h_s(\gamma(t), B) \leq \int_t^{t+s} L(\gamma_M^2, \dot{\gamma}_M^2, \gamma_{\mathcal{I}}^2) + \sum_{i=0}^{N_2-1} \psi\big(\gamma_M(t_{i+1}^2), \gamma_{\mathcal{I}}^2(t_i^1), \gamma_{\mathcal{I}}^2(t_{i+1}^2)\big), \tag{4.39}$$

Hence, by adding the last two inequalities, and combining it with (4.37), we conclude

$$\begin{aligned} h_{t+s}(A, B) + \varepsilon &\geq h_t(A, \gamma(t)) + h_s(\gamma(t), B) \\ &\geq \inf_{C \in M \times \mathcal{I}} \big\{ h_t(A, C) + h_s(C, B) \big\}, \end{aligned} \tag{4.40}$$

and since ε is arbitrary, the claim is proved.

We now turn to the proof of (4.33), which becomes simple with the help of (4.34). In fact, for $A \in M \times \mathcal{I}$, we have

$$\begin{aligned} T_{t+s}u(A) &= \inf_{B \in M \times \mathcal{I}} \big\{ u(B) + h_{t+s}(B, A) \big\} \\ &= \inf_{B \in M \times \mathcal{I}} \big\{ u(B) + \inf_{C \in M \times \mathcal{I}} \big[h_t(B, C) + h_s(C, A) \big] \big\} \\ &= \inf_{C \in M \times \mathcal{I}} \big\{ h_s(C, A) + \inf_{B \in M \times \mathcal{I}} \big[u(B) + h_t(B, C) \big] \big\} \\ &= \inf_{C \in M \times \mathcal{I}} \big\{ T_t(C) + h_s(C, A) \big\} \\ &= T_s(T_t u)(A), \end{aligned} \tag{4.41}$$

as desired. □

4.2 Weak KAM theorem in the optimal switching setting

The aim of this section is to prove Theorem 62. As in the Weak KAM Theorem for the action of a single Lagrangian on a compact manifold, we proceed to show that the associated Lax–Oleinik semigroup (see Definition 63 above) converges to a viscosity solution of the system of Hamilton–Jacobi equations (4.5)

We start by showing that the critical value c_0 is well defined as the infimum of $c \in \mathbb{R}$ for which (4.4) admits a viscosity subsolution:

Proposition 75. *There exists a unique $c_0 \in \mathbb{R}$ satisfying:*

(i) The equation (4.4) admits a subsolution, with c replaced by c_0;

(ii) For any $c \in \mathbb{R}$ such that (4.4) admits a subsolution, $c \geq c_0$.

Proof. Define $c_0 \in \mathbb{R}$ by

$$c_0 := \inf \left\{ c \in \mathbb{R}; (4.4) \text{ admits a subsolution} \right\}.$$

Notice that the infimum is over a nonempty set, because for big values of c, $u \equiv 0$ is a subsolution. Moreover, we clearly have $c_0 \geq \max_i c_0(L_i)$, where $c_0(L_i)$ is the Mañé critical value of $L(\cdot, \cdot, i)$; thus, $c_0 \in \mathbb{R}$.

In order to show that this infimum is in fact a minimum, let $c^j \in \mathbb{R}$ be a minimizing sequence for c_0, so that there exists a subsolution u^j for (4.4) with c replaced by c^j, and satisfying $c^j \to c_0$. Since $u^j(\cdot, i)$ is uniformly Lipschitz, we know that there exists u such that $u^j(\cdot, i) \to u(\cdot, i)$, up to a subsequence. It is then clear that this limit u is a subsolution of (4.4) with constant c_0. □

From this point on, so that the notation is simplified, we consider a 'normalized' definition for the Lax–Oleinik semigroup, by adding the constant c_0, from last proposition, to the Lagrangians $L(\cdot, \cdot, i)$:

$$\begin{aligned} T_t u(A) := \inf \Bigg\{ u(\gamma(0)) + \sum_{k=0}^{N} \int_{t_k}^{t_{k+1}} \left[L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(t_k)\big) + c_0 \right] ds \\ + \sum_{k=0}^{N-1} \psi\big(\gamma_M(t_k), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})\big) \Bigg\}. \end{aligned} \tag{4.42}$$

Lemma 76. *Assume $u : M \times \mathcal{I} \to \mathbb{R}$ is a subsolution of (4.4). Then, for any $t \in [0, \infty)$, there exists $A_t = (x_t, i_t) \in M \times \mathcal{I}$ such that $u(A_t) = T_t u(A_t)$.*

Proof. Assume, by contradiction, that this is not the case. Then there exists $t_0 > 0$ such that, for every $A \in M \times \mathcal{I}$,

$$u(A) < T_{t_0} u(A).$$

Since M is compact, we can find $\varepsilon > 0$ such that, for all $A \in M \times \mathcal{I}$,

$$u(A) < T_{t_0}u(A) - \varepsilon =: \tilde{T}_{t_0}u(A), \tag{4.43}$$

so that

$$\begin{aligned}\tilde{T}_t u(A) := \inf \Bigg\{ u(\gamma(0)) + \sum_{k=0}^{N} \int_{t_k}^{t_{k+1}} \Big(L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(t_k)\big) + \tilde{c} \Big)\, ds \\ + \sum_{k=0}^{N-1} \psi\big(\gamma_M(t_k), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})\big) \Bigg\},\end{aligned}$$

with $\tilde{c} = c_0 - \varepsilon/t_0$. Since u is a subsolution, the monotonicity of the semigroup $\tilde{T}_t$ implies that for any $\tau > 0$, and any $n \in \mathbb{N}$,

$$\tilde{T}_\tau u \le \tilde{T}_{nt_0+\tau} u.$$

This in turn implies

$$\tilde{u} := \sup_{\tau \ge 0} \tilde{T}_\tau u = \sup_{0 \le \tau \le t_0} \tilde{T}_{nt_0+\tau} u.$$

Observe now that, for any $h > 0$,

$$\begin{aligned}\tilde{T}_h \tilde{u} &= \sup_{\tau \ge 0} \tilde{T}_{\tau+h} u \\ &= \sup_{\tau \ge h} \tilde{T}_\tau u \ge \tilde{u},\end{aligned} \tag{4.44}$$

so that $\tilde{u}$ is a subsolution to

$$\max \Big\{ H(x, d_x u_i, i) - \tilde{c},\ \max_{j \ne i} \ \big\{ u(x,i) - u(x,j) - \psi(x,i,j) \big\} \Big\} = 0,\ \forall\ i. \tag{4.45}$$

But then, by the definition of c_0, we have

$$c_0 \le \tilde{c} = c_0 - \varepsilon/t_0,$$

which is clearly a contradiction. $\square$

Proposition 77. *For every $i \in \mathcal{I}$, $T_t u(\cdot, i)$ is uniformly bounded in t, meaning*

$$\sup_{t \geq 0} \|T_t u(\cdot, i)\|_{L^\infty(M)} < +\infty.$$

Proof. For $t \geq 0$, there exists $A_t = (x_t, i_t) \in M \times \mathcal{I}$ such that $(T_t u - u)(A_t) = 0$. Since $T_t u(\cdot, i_t) - u(\cdot, i_t)$ is uniformly Lipschitz and vanishes at one point, we have

$$\begin{aligned} |T_t u(x, i_t) - u(x, i_t)| &= |(T_t u - u)(x, i_t) - (T_t u - u)(A_t)| \\ &\leq 2C d(x, x_t^i) \\ &\leq 2C \operatorname{diam}(M) \equiv C_M, \end{aligned} \tag{4.46}$$

with C independent of t. Then, *for any* $i \in \mathcal{I}$, we know

$$T_t u(x, i_t) - \psi(x, i_t, i) \leq T_t u(x, i) \leq T_t u(x, i_t) + \psi(x, i, i_t).$$

Here, we again use that $T_t u$ is a solution to (5.1), as stated in Theorem 88. Thus,

$$|T_t u(\cdot, i)| \leq \|u(\cdot, i_t)\|_{L^\infty(M)} + C_M + \sup_{x \in M, i \neq j} \psi(x, i, j),$$

whence,

$$\sup_{t \geq 0} \|T_t u(\cdot, i)\|_{L^\infty(M)} \leq \|u(\cdot, i_t)\|_{L^\infty(M)} + C.$$

□

Proposition 78. *If $u : M \times \mathcal{I} \to \mathbb{R}$ is a subsolution of (4.4), then $T_t u$ converges to a fixed point of the Lax–Oleinik semigroup T_t.*

Proof. By assumption, u is a subsolution and so $u \leq T_t u$. By the monotonicity of the Lax–Oleinik semigroup, we have

$$T_s u \leq T_{s+t} u.$$

Since it is also uniformly bounded in t, the pointwise limit

$$u^\infty(x, i) := \lim_{t \to +\infty} T_t u(x, i)$$

is well defined, everywhere in M. Since T_t is continuous in $t \geq 0$, we conclude

$$T_s u^\infty(x,i) = \lim_{t\to+\infty} T_{s+t}u(x,i) = u^\infty(x,i).$$

$\square$

As a corollary, taking Proposition 72 into account, we obtain Theorem 62, which was the goal of this section.

4.3 The Aubry set

In this section, we start our analysis of the Aubry–Mather theory in our setting. We define the Aubry set $\mathcal{A}$ in a fashion similar to [19]. We prove the following theorem:

Definition 79. The *projected Aubry set* $\mathcal{A}$ is defined as the set of $B \in M \times \mathcal{I}$ for which

$$\liminf_{t\to+\infty} h_t(B,B) = 0. \tag{4.47}$$

The set $\mathcal{A}$ is defined in Gomes-Serra [19] as the set of $B \in M \times \mathcal{I}$ for which (ii) of Proposition 80 below is satisfied. We prove that this definition is in fact equivalent to the one we propose. The proof is standard and follows the same steps of the proof of an equivalent statement in the classical theory. We recall that, by Proposition 71, any critical subsolution satisfies

$$u\big(\gamma(t)\big) - u\big(\gamma(0)\big) \leq \mathcal{J}_t[\gamma], \ \forall \gamma \in AC([0,t];M) \times \mathcal{P}([0,t];\mathcal{I}).$$

In particular, the action on loops is always nonnegative.

Proposition 80. *The following are equivalent:*

(i) $B \in \mathcal{A}$;

(ii) $\inf\left\{\mathcal{J}_t[\gamma] \;\middle|\; t \geq \delta, \gamma(0) = \gamma(t) = B\right\} = 0$, *for some* $\delta > 0$;

(iii) $\inf\left\{\mathcal{J}_t[\gamma] \;\middle|\; t \geq \delta, \gamma(0) = \gamma(t) = B\right\} = 0$, *for every* $\delta > 0$.

Proof. $(i) \implies (ii)$. Let $\varepsilon > 0$. If $B \in \mathcal{A}$, then there exist a sequence $t_k \to +\infty$ for which $h_{t_k}(B, B) \to 0$. Let $\gamma^k : [0, t_k] \to M \times \mathcal{I}$ with $\gamma(0) = \gamma(t) = B$ be such that

$$h_{t_k}(B, B) + \frac{\varepsilon}{2} \geq \mathcal{J}_{t_k}[\gamma^k].$$

Without loss of generality, we may assume $t_k \geq 1$, for all k. Since $h_{t_k}(B, B) \to 0$, if k is large enough

$$h_{t_k}(B, B) \leq \frac{\varepsilon}{2}$$

and

$$\varepsilon \geq \mathcal{J}_{t_k}[\gamma^k]$$

for some $t_k \geq 1$. Then (ii) holds with $\delta = 1$.

$(ii) \implies (iii)$. Since (ii) holds, there exists $\delta_0 > 0$ such that (4.47) holds true. Now fix any $\delta > 0$. If $\delta < \delta_0$,

$$\inf \left\{ J_t[\gamma] \;\middle|\; t \geq \delta \right\} \leq \inf \left\{ J_t[\gamma] \;\middle|\; t \geq \delta_0 \right\} = 0, \tag{4.48}$$

and we are done, because the action on loops is nonnegative. Otherwise, $\delta > \delta_0$ and we let $m \in \mathbb{N}$ be such that $\delta < m\delta_0$. In this way, for any $\gamma : [0, t] \to M \times \mathcal{I}$ with with period $t \geq \delta_0$, we can define a curve $\bar{\gamma} : [0, mt] \to M \times \mathcal{I}$ by concatenation of γ with itself, m times, and we obtain

$$J_{mt}[\bar{\gamma}] = mJ_t[\gamma].$$

We do not create any new switches, for γ is a loop. Since $mt \geq \delta_0$, both infima must coincide.

$(iii) \implies (i)$. Take $\delta = k$ and define a sequence. $\square$

Next, we extend the definition of critical curves to our setting and prove its existence for points of the projected Aubry set $\mathcal{A}$ (compare proof with Theorem 25).

Definition 81 (Critical curve)**.** We say $\gamma : \mathbb{R} \to M \times \mathcal{I}$ is a critical curve if, for any subsolution $u : M \times \mathcal{I} \to \mathbb{R}$ of (4.4), and all $t_1 < t_2$,

$$u\big(\gamma(t_2)\big)-u\big(\gamma(t_1)\big) = \int_{t_1}^{t_2} L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(s)\big)\, ds+\sum_k \psi\big(\gamma_M(s_{k+1}), \gamma_{\mathcal{I}}(s_k), \gamma_{\mathcal{I}}(s_{k+1})\big),$$

where the sum above is taken for all k such that $t_1 < s_{k+1} < t_2$.

Proposition 82 (Existence of critical curves)**.** *Given $B = (y, j) \in \mathcal{A}$, there exists a critical curve $\gamma : \mathbb{R} \to M \times \mathcal{I}$ with $\gamma(0) = B$.*

Proof. Let $\eta^k : [0, t_k] \to M \times \mathcal{I}$, with $\eta^k(0) = B = \eta^k(t_k)$ and $t_k \geq k$, be such that

$$\mathcal{J}_{t_k}[\eta^k] = \int_0^{t_k} L\big(\eta^k_M(s), \dot\eta^k_M(s), \eta_{\mathcal{I}}(s)\big)\; ds + \sum_{j=1}^{N^k-1} \psi\big(\eta^k_{\mathcal{I}}(s_j)\eta^k_{\mathcal{I}}(s_{j+1})\big) \to 0.$$

If we set $\gamma^k : [-t_k/2, t_k/2] \to M \times \mathcal{I}$ as $\gamma^k(s) := \eta^k(s + t_k/2)$, we have

$$\int_{-t_k/2}^{t_k/2} L\big(\gamma^k_M(s), \dot\gamma^k_M(s), \gamma^k_{\mathcal{I}}(s)\big)\; ds + \sum_{j=1}^{N^k-1} \psi\big(\gamma^k_M(s_j), \gamma^k_{\mathcal{I}}(s_j)\gamma^k_{\mathcal{I}}(s_{j+1})\big) \to 0.$$

We observe that the number of switches grows at most linearly, otherwise $\mathcal{J}_{t_k}[\eta^k]$ would diverge. Since the action of a loop is nonnegative, it is easy to see that γ^k can be assumed to be an action minimizer. Then

$$0 \leq \frac{1}{t_k}\int_{-t_k/2}^{t_k/2} L\big(\gamma^k_M(s), \dot\gamma^k_M(s), \gamma^k_{\mathcal{I}}(s)\big)\; ds + \frac{1}{t_k}\sum_{j=1}^{N-1} \psi\big(\gamma^k_M(s_j), \gamma^k_{\mathcal{I}}(s_j)\gamma^k_{\mathcal{I}}(s_{j+1})\big) \to 0$$

and there exists $t^k_0 \in [-t_k/2, t_k/2]$ for which $L\big(\gamma^k_M(t^k_0), \dot\gamma^k_M(t^k_0), \gamma^k_{\mathcal{I}}(t^k_0)\big) \leq C$; then, by assumption 3 in the Definition 7 of a Tonelli Lagrangian,

$$\big\|\dot\gamma^k_M(t^k_0)\big\|_{\gamma^k_M(t^k_0)} \leq L\big(\gamma^k_M(t^k_0), \dot\gamma^k_M(t^k_0), \gamma^k_{\mathcal{I}}(t^k_0)\big) + C \leq C.$$

Now, conservation of energy implies that the speed curve $(\gamma^k_M, \dot\gamma^k_M)$ is contained in a compact subset of TM. Thus, by a diagonal argument, we can construct a curve

$\gamma_M : \mathbb{R} \to M$ such that, for every fixed interval $[-T, T]$

$$\begin{aligned} \gamma_M^k &\to \gamma_M, \ \text{uniformly in } [-T,T], \\ \dot\gamma_M^k &\rightharpoonup \dot\gamma_M, \ \text{weakly in } L^1\big([-T,T]\big). \end{aligned} \tag{4.49}$$

By following the proof of Theorem 48, we can also defined a limit $\gamma_{\mathcal{I}}$ of $\gamma_{\mathcal{I}}^k$, for which lower semicontinuity of the action functional holds in $[-T, T]$ (see (3.9)). Finally, we prove γ is a critical curve. Indeed, given u subsolution of (4.4), and $a \le b$, we have

$$0 \le u(\gamma^k(a)) - u(\gamma^k(b)) + \int_a^b L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k)\, ds + \sum_{[a,b]} \psi\big(\gamma_M^k, \gamma_{\mathcal{I}}^{k-}, \gamma_{\mathcal{I}}^{k+}\big) =: I_1^k \tag{4.50}$$

and, since γ^k is a loop,

$$\begin{aligned} 0 \le u(\gamma^k(b)) - u(\gamma^k(a)) + \int_b^{t_k/2} L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k) + \int_{-t_k/2}^{a} L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k) \\ + \sum_{[-t_k/2, t_k/2]\setminus[a,b]} \psi(\gamma_M^k, \gamma_{\mathcal{I}}^{k-}, \gamma_{\mathcal{I}}^{k+}) =: I_2^k. \end{aligned} \tag{4.51}$$

By adding these, we obtain

$$0 \le \int_{-t_k/2}^{t_k/2} L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k) + \sum_{[-t_k/2, t_k/2]} \psi(\gamma_M^k, \gamma_{\mathcal{I}}^{k-}, \gamma_{\mathcal{I}}^{k+}) \to 0.$$

In this way, $I_1^k \ge 0$ and $I_2^k \ge 0$ are such that $I_1^k + I_2^k \to 0$. In particular, each of them converges to zero and we have

$$\lim_k \left\{ u(\gamma^k(a)) - u(\gamma^k(b)) + \int_a^b L(\gamma_M^k, \dot\gamma_M^k, \gamma_{\mathcal{I}}^k)\, ds + \sum_{[a,b]} \psi\big(\gamma_M^k, \gamma_{\mathcal{I}}^{k-}, \gamma_{\mathcal{I}}^{k+}\big) \right\} = 0.$$

Hence, lower semicontinuity implies

$$u(\gamma(b)) - u(\gamma(a)) = \liminf_k \left\{ \int_a^b L(\gamma_M^k, \dot{\gamma}_M^k, \gamma_{\mathcal{I}}^k)\, ds + \sum_{[a,b]} \psi\big(\gamma_M^k, \gamma_{\mathcal{I}}^{k-}, \gamma_{\mathcal{I}}^{k+}\big) \right\}$$
$$\geq \int_a^b L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) + \sum_{[a,b]} \psi\big(\gamma_M^k, \gamma_{\mathcal{I}}^{-}, \gamma_{\mathcal{I}}^{+}\big).$$

The other inequality is clear, by Proposition 71. □

Proposition 83. *Every critical curve is contained in the projected Aubry set $\mathcal{A}$.*

Proof. Gomes and Serra proved that for every $(x, i) \in \mathcal{A}$, and every subsolution u of (4.4), $u(\cdot, i)$ cannot be a strict subsolution at x (see [19, remarks after Theorem 3.6]), that is, at least one of the following is satisfied:

(i) $H\big(x, du(x, i), i\big) = 0$, in the viscosity sense or;

(ii) for some $j \neq i$, $v(x, i) - v(x, j) = \psi(x, i, j)$.

The proof then follows, since the next proposition ensures there exist subsolutions that are strict outside the Aubry set $\mathcal{A}$. □

Gomes and Serra [19] proved the following:

Proposition 84. *If the open subset $U \subset M \times \mathcal{I}$ is such that $\overline{U} \subseteq M \times \mathcal{I} \backslash \mathcal{A}$, then there exists a subsolution v of (4.4) which is strict in U, meaning*

- *for all $(y, j) \notin \mathcal{A}$, and $p \in \partial^+ v_j(y)$, we have $H(y, p, j) < 0$ and*
- *for all $i \neq j$, $v_j(y) - v_i(y) < \psi(i, j)$.*

Furthermore, v is smooth in $(M \times \mathcal{I}) \backslash \mathcal{A}$.

Proof. See [19, Lemma 4.2]. □

We now use this result to prove a comparison principle:

Corollary 85 (Comparison Principle)**.** *Suppose u is a subsolution, w is a supersolution, and that $u \leq w$ on $\mathcal{A}$. Then $u \leq w$ in $M \times \mathcal{I}$.*

Proof. Suppose to find a contradiction that

$$\min_{M\times\mathcal{I}}(w-u)<0.$$

Then any minimizer must be outside $\mathcal{A}$. Consider v_n subsolutions that approximate u and are smooth outside $\mathcal{A}$, and (x_n, i_n) such that

$$w(x_n,i_n)-v_n(x_n,i_n)=\min_{M\times\mathcal{I}}(w-v_n). \tag{4.52}$$

For $n >> 1$, this minimum is still negative, and we must have $(x_n, i_n) \notin \mathcal{A}$. But (4.52) implies $d_{x_n}v_n \in \partial^- w(x_n, i_n)$. Since w is a supersolution, we must have either

- $H(x_n, d_{x_n}v_n, i_n) \geq 0$ or;
- for some $j \neq i_n$, $v_n(x_n, i_n) - v_n(x_n, j) = \psi(x_n, i_n, j)$.

Since both contradict the previous proposition, our result is proved. □

Next, we study properties of a special solution of (4.4), motivated by Davini-Siconolfi [10]. These properties will be important, for instance, in the next chapter when we describe the large time behavior of the Lax–Oleinik semigroup.

Given $u_0 \in C(M\times\mathcal{I})$, we consider

$$v(A):=\inf_{B\in\mathcal{A}}\Big\{h(B,A)+\inf_{C\in M\times\mathcal{I}}\big\{u_0(C)+h(C,B)\big\}\Big\}, \tag{4.53}$$

where $h : M\times\mathcal{I}\times M\times\mathcal{I}\to\mathbb{R}$ denotes the (generalized) Peierls barrier

$$h(A,B)=\liminf_{t\to+\infty}h_t(A,B).$$

We prove v is a critical solution defined on $M\times\mathcal{I}$. Indeed, we have (compare to [10, Theorem 3.1])

Theorem 86. *Set*

$$v_0(B):=\inf_{C\in M\times\mathcal{I}}\big\{u_0(C)+h(C,B)\big\},$$

for $B \in M \times \mathcal{I}$, so that

$$v(A) = \inf_{B \in \mathcal{A}} \big\{ v_0(B) + h(B, A) \big\}. \tag{4.54}$$

Then, the following hold true:

1. *v_0 is the maximal subsolution with $v_0 \leq u_0$ on $M \times \mathcal{I}$;*

2. *v is a solution and it equals v_0 on $\mathcal{A}$;*

3. *If $u_0(B) - u_0(A) \leq h(B, A)$, for all $A, B \in M \times \mathcal{I}$, then*

$$v(A) := \inf_{B \in \mathcal{A}} \big\{ u_0(B) + h(B, A) \big\}. \tag{4.55}$$

   in $M \times \mathcal{I}$, and $v_0 = u_0$ on $\mathcal{A}$.

Proof. 1. By setting $C = B$, we get that $v_0 \leq u_0$ on $M \times \mathcal{I}$. Also, by considering $C \in M \times \mathcal{I}$ such that

$$v_0(B) = u_0(C) + h(C, B),$$

we obtain

$$v_0(A) - v_0(B) \leq h(C, A) - h(C, B) \leq h(B, A),$$

and we are done.

2. This is what says [19, Proposition 4.3].

3. If $u_0(B) - u_0(A) \leq h(B, A)$, for all $A, B \in M \times \mathcal{I}$, then u_0 is a subsolution and we must have $u_0 = v_0$, which implies (4.55). □

Chapter 5

Time–dependent weakly coupled Hamilton–Jacobi system

In this chapter we study the asymptotic behavior of solutions to the time–dependent weakly coupled Hamilton–Jacobi system associated to the optimal switching problem.

In Section 5.1 we prove that the Lax–Oleinik semigroup of Chapter 4 is a viscosity solution to the time–dependent equations. In Section 5.2 we prove the main result of this chapter: we prove that the Lax–Oleinik solution to the time–dependent system goes asymptotically to a weak KAM solution, as $t \to +\infty$.

5.1 Viscosity solutions

We now study the following time–dependent system of weakly coupled Hamilton–Jacobi equations: $\forall i \in \mathcal{I}$,

$$\max\left\{\partial_t u(t,x,i)+H\big(x,\partial_x u(t,x,i),i\big),\ \max_{j\neq i}\left\{u(t,x,i)-u(t,x,j)-\psi(x,i,j)\right\}\right\}=0, \tag{5.1}$$

on $[0,+\infty)\times M$. The definition of viscosity solutions to this system is completely analogous to the stationary case studied in Chapter 4.

Definition 87. 1. We say that the continuous function $u : [0,+\infty)\times M\times\mathcal{I} \to \mathcal{I}$ is a *viscosity subsolution* to the system (5.1) if, for each $i, j \in \mathcal{I}$,

(i) for any $i \in \mathcal{I}$, the function $u(\cdot,\cdot,i)$ satisfies

$$\partial_t u(t,x,i) + H\big(x, \partial_x u(t,x,i), i\big) \leq 0,$$

in the viscosity sense, and;

(ii) for any $i, j \in \mathcal{I}$, the functions $u(\cdot,\cdot,i)$ and $u(\cdot,\cdot,j)$ satisfy

$$u(t,x,i) \leq u(t,x,j) + \psi(x,i,j).$$

2. We say that $u : [0,+\infty) \times M \times \mathcal{I} \to \mathbb{R}$ is a *viscosity supersolution* to the system (5.1) if, for every $i \in \mathcal{I}$, at least one of the following is satisfied:

(i) the function $u(\cdot,\cdot,i)$ satisfies

$$\partial_t u(t,x,i) + H\big(x, \partial_x u(t,x,i), i\big) \geq 0,$$

in the viscosity sense, or;

(ii) there exists $j \in \mathcal{I}$, $j \neq i$, such that the functions $u(\cdot,\cdot,i)$ and $u(\cdot,\cdot,j)$ satisfy

$$u(t,x,i) \geq u(t,x,j) + \psi(x,i,j).$$

3. The function u is said to be a *viscosity solution* if it is both viscosity subsolution and a a viscosity supersolution.

In the following proposition we prove that indeed we obtain a solution of (5.1) by the Lax–Oleinik formula.

Proposition 88. *Let $u_0 \in C(M \times \mathcal{I})$. Then, $u : [0,+\infty) \times M \times \mathcal{I} \to \mathbb{R}$ defined by*

$$u(t,A) = T_t u(A), \ \ t \geq 0, A \in M \times \mathcal{I},$$

is a viscosity solution of (5.1).

Proof. Fix $(t_0, x_0, i) \in \mathbb{R} \times M \times \mathcal{I}$. To prove that it is a subsolution, we first verify the inequality of item (i), in the viscosity sense. Choose $\gamma_{\mathcal{I}} : [0, t_0] \to \{1, \ldots, m\}$ to be arbitrary in $[0,t)$, with $\gamma_{\mathcal{I}} \equiv i$ in the interval ending t, where it is constant, and $\gamma_{\mathcal{I}} \equiv i$ in $[t, t_0]$ as well. Choose any curve $\gamma_M \in AC([0, t_0]; M)$ satisfying $\gamma_M(t) = x$, $\gamma_M(t_0) = x_0$, and $\dot{\gamma}_M(t_0) = v \in T_{x_0}M$. Then,

$$\begin{aligned} u(t_0, x_0, i) \leq u\big(0, \gamma_M(0), \gamma_{\mathcal{I}}(0)\big) + \int_0^t L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)\big) + \sum \psi\big(\gamma_M, \gamma_{\mathcal{I}}^-, \gamma_{\mathcal{I}}^+\big) \\ + \int_t^{t_0} L\big(\gamma_M(s), \dot{\gamma}_M(s), i\big)\, ds. \end{aligned} \tag{5.2}$$

Since $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ is arbitrary in $[0, t]$, we obtain

$$u(t_0, x_0, i) \leq u\big(t, \gamma_M(t), i\big) + \int_t^{t_0} L\big(\gamma(s), \dot{\gamma}(s), i\big)\, ds. \tag{5.3}$$

So, if ϕ is a test function such that $u(\cdot, \cdot, i) - \phi$ has a maximum at (t_0, x_0), we have

$$\phi(t_0, x_0) - \phi(t, \gamma_M(t)) \leq u(t_0, x_0, i) - u\big(t, \gamma_M(t), i\big) \leq \int_t^{t_0} L\big(\gamma_M(s), \dot{\gamma}_M(s), i\big)\, ds,$$

so that

$$\frac{\phi(t_0, x_0) - \phi(t, \gamma_M(t))}{t_0 - t} \leq \fint_t^{t_0} L\big(\gamma_M(s), \dot{\gamma}_M(s), i\big)\, ds.$$

Let $t \to t_0$ to get

$$\partial_t \phi(t_0, x_0) + \partial_x \phi(t_0, x_0) \cdot v \leq L(t_0, v, i).$$

Since $v \in T_{x_0}M$ is arbitrary, we have proven

$$\partial_t \phi(t_0, x_0) + H\big(x_0, \partial_x \phi(t_0, x_0), i\big) \leq 0.$$

We now verify our second condition in order to be a viscosity subsolution. Fix $i, j \in \mathcal{I}$, and we want to show that for all (t, x),

$$u(t, x, i) \leq u(t, x, j) + \psi(x, i, j).$$

For given $\varepsilon > 0$, let $\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ be such that $\gamma(t) = (x, j)$ and

$$\begin{aligned} u(t, x, j) + \varepsilon \geq u(0, \gamma(0)) + \sum_{k=0}^{N} \int_{t_k}^{t_{k+1}} L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(t_k)\big)\ ds \\ + \sum_{k=0}^{N-1} \psi(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})). \end{aligned} \tag{5.4}$$

Thus, if we stay at the point x for a time $\delta > 0$ longer, by using the Lagrangian i, we obtain

$$\begin{aligned} u(t+\delta, x, i) &\leq u(0, \gamma(0)) + \sum_{k=0}^{N} \int_{t_k}^{t_{k+1}} L\big(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(t_k)\big)\ ds \\ &\quad + \delta L(x, 0, i) + \sum_{k=0}^{N-1} \psi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})\big) + \psi\big(x, i, j\big) \\ &\leq u(t, x, j) + \varepsilon + \delta L(x, 0, i) + \psi(x, i, j) \end{aligned} \tag{5.5}$$

But then, by letting $\delta \to 0$ and using the continuity of $u(\cdot, \cdot, i)$, we obtain

$$u(t, x, i) \leq u(t, x, j) + \psi(x, i, j) + \varepsilon;$$

therefore, since this is true for an arbitrary $\varepsilon > 0$, u is a viscosity subsolution of (5.1).

In order to prove that it is also a supersolution, we first observe that if $v : \mathbb{R} \times M \times \mathcal{I} \to \mathbb{R}$, with the same initial condition as u (meaning $v(0, x, i) = u(0, x, i)$, $\forall i$) is a subsolution, then necessarily $v(\cdot, \cdot, i) \leq u(\cdot, \cdot, i)$, for all i. Indeed, for given

$\gamma = (\gamma_M, \gamma_{\mathcal{I}})$ with $\gamma(t) = (x, i)$, we compute

$$
\begin{aligned}
v(t,x,i) &- v\big(0, \gamma_M(0), \gamma_{\mathcal{I}}(0)\big) \\
&= \sum_{k=0}^{N} \Big[v\big(t_{k+1}, \gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_{k+1})\big) - v\big(t_k, \gamma_M(t_k), \gamma_{\mathcal{I}}(t_k)\big) \Big] \\
&= \sum_{k=0}^{N} \Big[v\big(t_{k+1}, \gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k)\big) - v\big(t_k, \gamma_M(t_k), \gamma_{\mathcal{I}}(t_k)\big) \Big] \\
&\quad + \sum_{k=0}^{N} \Big[v\big(t_{k+1}, \gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_{k+1})\big) - v\big(t_{k+1}, \gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k)\big) \Big] \\
&\le \sum_{k=0}^{N} \int_{t_k}^{t_{k+1}} L\big(\gamma_M(s), \dot\gamma_M(s), \gamma_{\mathcal{I}}(t_k)\big)\, ds \\
&\quad + \sum_{k=0}^{N-1} \psi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})\big).
\end{aligned}
\tag{5.6}
$$

(note we use $\gamma_{\mathcal{I}}(t_N) = \gamma_{\mathcal{I}}(t) = i$) Since the initial conditions are the same, and the previous inequality holds for any γ, we conclude $v(t,x,i) \le u(t,x,i)$. Now we proceed to prove that u is in fact a supersolution. For fixed x_0, and t_0, assume, in order to find a contradiction, that both conditions in the definition of supersolution fail to be true, that is, that there exist $i \in \mathcal{I}$ and ϕ such that $u(\cdot,\cdot,i) - \phi$ has a minimum at (t_0, x_0), with

$$\partial_t \phi(t_0, x_0) + H\big(x_0, \partial_x \phi(t_0, x_0), i\big) < 0,$$

and that

$$u(t_0, x_0, i) < u(t_0, x_0, j) + \psi(x_0, i, j), \ \forall j \in \mathcal{I}.$$

By continuity, these inequalities hold true at least in a small neighborhood $B_\delta(t_0, x_0)$ of (t_0, x_0). Then, if we set $\varepsilon > 0$ any positive number satisfying

$$\varepsilon < \max_{\partial B_\delta(t_0,x_0)} \{u(t,x,i) - \phi(t,x)\},$$

$$\varepsilon < \max_{B_\delta(t_0,x_0)} \{u(t,x,j) + \psi(x,i,j) - u(t,x,i)\},$$

and consider the function v defined as

$$v(t,x) := \begin{cases} \max\left\{\phi(t,x)+\varepsilon,\ u(t,x,i)\right\}, & \text{for } (t,x) \in B_\delta(t_0,x_0); \\ u(t,x,i), & \text{elsewhere,} \end{cases} \tag{5.7}$$

one verifies that $\tilde{u}$ defined as

$$\tilde{u}(t,x,j) := \begin{cases} v(t,x,i), & \text{for } j = i; \\ u(t,x,j), & \text{for } j \neq i, \end{cases} \tag{5.8}$$

is also a subsolution. But $v(t_0,x_0,i) > u(t_0,x_0,i)$, a contradiction. □

5.2 Large time behavior of the generalized Lax–Oleinik semigroup

By combining our Theorem 62 of Chapter 4 with Proposition 88 above, we see that, when $u_0 : M \times \mathcal{I} \to \mathbb{R}$ is a critical subsolution, then the Lax–Oleinik solution $u = T_t u_0$ converges, as $t \to +\infty$, to a solution of the critical Hamilton–Jacobi system (4.4).

Naturally, we are interested in the following question: Does the Lax–Oleinik semigroup converge for any given "initial" function $u_0 : M \times \mathcal{I} \to \mathbb{R}$? The main theorem of this section answers affirmatively this question:

Theorem 89. *Let $u_0 : M \times \mathcal{I} \to \mathbb{R}$ be a continuous function. Then, the Lax–Oleinik semigroup $T_t u_0 : M \times \mathcal{I} \to \mathbb{R}$ converges, as $t \to +\infty$, to a critical solution v of (4.4), given by (4.53).*

We observe that the long time behavior for different but related systems has been studied recently by Filippo Cagnetti, Diogo Gomes, Hiroyoshi Mitake, and Hung V. Tran [5].

Our method for proving Theorem 89 follows Davini-Siconolfi strategy [10] in the classical case that we presented in Subsection 2.8.1. More precisely, we set

$$\omega(u_0) := \left\{\psi : M \times \mathcal{I} \to \mathbb{R}\ ;\ \psi = T_{t_n} u,\ \text{for some } t_n \to +\infty\right\} \tag{5.9}$$

and define the 'semilimits':

$$\underline{u}(A) = \sup \left\{ \psi(A) \;\middle|\; \psi \in \omega(u_0) \right\} \tag{5.10}$$

and

$$\overline{u}(A) = \inf \left\{ \psi(A) \;\middle|\; \psi \in \omega(u_0) \right\}; \tag{5.11}$$

these are well defined since the family $\{T_t u\}_{t>0}$ is uniformly bounded and uniformly Lipschitz.

Proposition 90. *As defined above, $\underline{u}$ is a subsolution of (4.4) and $\overline{u}$ is a supersolution of (4.4).*

Proof. By using Propositions 71 and 72, the proof follows exactly as in the proof of Proposition 38. □

Next, we give a second proof of our weak KAM theorem for the weakly coupled equations, Theorem 62. The proof is an easy variation of [10, Theorem 3.4].

Proposition 91. *Let u be either a subsolution or a supersolution, and let v be the function defined by (4.53). Then*

$$T_t u \to v.$$

Proof. Assume u_0 is a subsolution, so that $u_0 \leq T_t u_0$. Since v is the maximal subsolution with $v = u_0$ on $\mathcal{A}$, we have $u_0 \leq v$ on $M \times \mathcal{I}$. Then, by noticing $v = T_t v$, for any $t > 0$, we have

$$u_0 \leq T_t u_0 \leq v \text{ on } M \times \mathcal{I}.$$

Now, $v = u_0$ on $\mathcal{A}$ implies

$$T_t u_0 = v \text{ on } \mathcal{A}, \ \forall\, t > 0.$$

This means

$$\underline{u} = \overline{u} = v \text{ on } \mathcal{A},$$

and the comparison principle implies the same equality on the whole $M \times \mathcal{I}$.

Next, assume u_0 is a supersolution, so that $u_0 \geq T_t u_0$. Consider

$$v_0(B) := \inf_{C \in M \times \mathcal{I}} \big\{ u_0(C) + h(C, B) \big\},$$

the maximal subsolution with $v_0 \leq u_0$ on $M \times \mathcal{I}$. Since v_0 is a subsolution, we have $v_0 \leq T_t v_0$. Then

$$v_0 \leq T_t v_0 \leq T_t u_0 \leq u_0 \text{ on } M \times \mathcal{I},$$

and the maximality of v_0 implies that these are all equalities. In particular, $v_0 = T_t v_0$, that is, v_0 is a solution, which implies $v = v_0$ on $M \times \mathcal{I}$. By monotonicity, we get

$$v \leq T_t u_0 \leq u_0 \text{ on } M \times \mathcal{I}, \ \forall\, t > 0,$$

which implies $v \leq \overline{u} \leq \underline{u} \leq u_0$ on $M \times \mathcal{I}$. Since $\underline{u}$ is a subsolution, we must have $\underline{u} \leq v$ and the proof is finished. □

Proposition 92. *Let $\underline{u}, \overline{u}$ be given by* (5.10) *and* (5.11)*, respectively, and v be the solution of* (4.4) *given by* (4.53)*. We have*

$$v \leq \overline{u} \leq \underline{u} \text{ on } M \times \mathcal{I}. \tag{5.12}$$

Proof. As in the second part of the previous proof, v_0 satisfies $v_0 \leq u_0$ on $M \times \mathcal{I}$. Then $T_t v_0 \leq T_t u_0$ on $M \times \mathcal{I}$. Since v_0 is a subsolution, we have $T_t v_0 \to v$ and both $v \leq \overline{u}$ and $v \leq \underline{u}$ are valid. The other inequality is trivial. □

Remark 93. In order to obtain the convergence result, all we need is $v = \underline{u}$ on $\mathcal{A}$. This is what we prove in the next section.

Lemma 94. *There exists a modulus of continuity ρ for which, if γ is a critical curve and λ is a constant sufficiently close to 1, then*

$$\int_{t_1}^{t_2} L\big(\gamma_M(\lambda s), \lambda \dot{\gamma}_M(\lambda s), \gamma_{\mathcal{I}}(\lambda s)\big)\, ds + \sum \psi \leq \frac{h\big(\gamma(\lambda t_1), \gamma(\lambda t_2)\big)}{\lambda} + |\lambda - 1| \rho(|\lambda - 1|)(t_2 - t_1). \tag{5.13}$$

Proof. The speed curve of a critical curve is bounded, say $(\gamma, \dot{\gamma}) \subset K$, $K \subset TM$ compact.

By the mean value theorem, we have

$$
\begin{aligned}
L\big(\gamma_M(\lambda s), \lambda\dot\gamma_M(\lambda s), \gamma_{\mathcal{I}}(\lambda s)\big) &- L\big(\gamma_M(\lambda s), \dot\gamma_M(\lambda s), \gamma_{\mathcal{I}}(\lambda s)\big) \\
&= (\lambda - 1)\frac{\partial L}{\partial v}\big(\gamma_M(\lambda s), \mu\dot\gamma_M(\lambda s), \gamma_{\mathcal{I}}(\lambda s)\big)\cdot\dot\gamma_M(\lambda s).
\end{aligned}
\tag{5.14}
$$

Then, if we take ρ a modulus of continuity for $\partial L/\partial v$ in $K := \{(x, \lambda v); (x, v) \in K, |\lambda - 1| \le \delta\}$, we have

$$
\begin{aligned}
L\big(\gamma_M(\lambda s), \lambda\dot\gamma_M(\lambda s), \gamma_{\mathcal{I}}(\lambda s)\big) \le L\big(\gamma_M(\lambda s), \dot\gamma_M(\lambda s), \gamma_{\mathcal{I}}(\lambda s)\big) \\
+ |\lambda - 1|\rho(|\lambda - 1|).
\end{aligned}
\tag{5.15}
$$

Thus,

$$
\begin{aligned}
\int_{t_1}^{t_2} L\big(\gamma_M(\lambda s), \lambda\dot\gamma_M(\lambda s), \gamma_{\mathcal{I}}(\lambda s)\big) + \sum\psi \le \int_{t_1}^{t_2} & L\big(\gamma_M(\lambda s), \dot\gamma_M(\lambda s), \gamma_{\mathcal{I}}(\lambda s)\big) \\
& + \sum\psi + C|\lambda - 1|\rho(C|\lambda - 1|)(t_2 - t_1) \\
& = \frac{1}{\lambda} h\big(\gamma(\lambda t_1), \gamma(\lambda t_2)\big) \\
& + C|\lambda - 1|\rho(C|\lambda - 1|)(t_2 - t_1),
\end{aligned}
\tag{5.16}
$$

as desired. □

Proposition 95. *Let $\underline{u}$ be given by (5.10), and v be the solution of (4.4) given by (4.53). Then, $\underline{u} \le v$ on $\mathcal{A}$.*

Proof. Let $\phi : M \times \mathcal{I} \to \mathbb{R}$ be in the ω-limit set of u_0, so that ϕ is the limit of $T_t u_0$ for a particular divergent (to $+\infty$) sequence of t's. It is not difficult to obtain from there a sequence $s_n \to +\infty$ such that

$$
\phi = \lim_{n\to+\infty} T_{s_n}\phi.
$$

Observe $T_{s_n}\phi$ is the classical Lax–Oleinik semigroup of the function ϕ. Let $A = (x, i) \in \omega(\gamma)$, so that $A = \lim_{t_n\to+\infty}\gamma(t_n)$, for some sequence $t_n \to +\infty$. Up to extracting subsequences, we can assume $\tau_n = t_n - s_n \to +\infty$. Since γ is a critical curve and v is a subsolution (it is in fact a solution), we have, by Lemma 94,

$$\begin{aligned} T_{s_n}\phi(\gamma(t_n)) - \phi(\gamma(t+\tau_n)) &\le \int_{t+\tau_n}^{t_n} L(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}}) + \sum \psi(\gamma_{\mathcal{I}}^-, \gamma_{\mathcal{I}}^+) + |t|\rho(t/s_n) \\ &= v(\gamma(t_n)) - v(\gamma(t+\tau_n)) + |t|\rho(t/s_n). \end{aligned}$$

If we set

$$\eta = \lim_n \gamma(\cdot + \tau_n),$$

then, η is also a critical curve and, by letting $n \to +\infty$, we obtain

$$\phi(A) - \phi(\eta(t)) \le v(A) - v(\eta(t)).$$

It only remains to show that

$$\liminf_t \left\{\phi\big(\eta(t)\big) - v\big(\eta(t)\big)\right\} \le 0.$$

To this order, observe

$$v(\eta(t)) - v(\eta(0)) = \int_0^t L(\gamma_M, \dot\gamma_M, \gamma_{\mathcal{I}}) + \sum \psi(\gamma_{\mathcal{I}}^-, \gamma_{\mathcal{I}}^+) \ge T_t u_0(\eta(t)) - u_0(\eta(0)).$$

Thus, since $\eta(\mathbb{R}) \subset \mathcal{A}$ and $v = u_0$ on $\mathcal{A}$,

$$\begin{aligned} \phi\big(\eta(t)\big) - v\big(\eta(t)\big) &\le \phi\big(\eta(t)\big) - T_t u_0(\eta(t)) + u_0(\eta(0)) - v\big(\eta(0)\big) \\ &\le \max_{B\in\mathcal{A}} \big|\phi - T_t u_0\big|. \end{aligned} \tag{5.17}$$

Since along a particular subsequence $\sigma_n \to +\infty$, we know $T_{\sigma_n} u_0 \to \phi$, our claim is proved. □

Chapter 6

Further developments

In this chapter we discuss further aspects of our optimal switching problem that can be developed in a future work.

6.1 Existence of $C^{1,1}$ subsolutions

Motivated by Bernard's results, we present the following conjecture.

Conjecture 96 (Existence of $C^{1,1}$ subsolutions)**.** *Given a subsolution u of* (1.9)*, there exists a subsolution v such that, for every mode $i \in \mathcal{I}$, $v(\cdot, i)$ is in $C^{1,1}(M)$, at least when $\mathcal{I} = \{1, 2\}$.*

In this chapter we present a few results that could be proven in case this conjecture is to be proven.

Remark 97. By approximation, we see that there exists a $C^{1,1}$ subsolution that is strict in U.

We improve Proposition 84 by showing that the set of $C^{1,1}$ subsolutions that are strict outside $\mathcal{A}$ is dense in the set of subsolutions.

Proposition 98. *Given a subsolution $u : M \to \mathbb{R}$, there exists a $C^{1,1}$ subsolution that coincides with u on $\mathcal{A}$.*

Proof. Let u be any subsolution. We claim that, for any $B = (y, j) \in \mathcal{A}$, and for any $t \geq 0$,

$$u(B) = T_t u(B) = \check{T}_t u(B).$$

Indeed, take a critical curve η associated to B and consider $\gamma(s) = \gamma(s+t)$, also critical. Then

$$u(B) = u(\gamma(0)) + \int_0^t L(\gamma_M, \dot{\gamma}_M, \gamma_{\mathcal{I}}) + \sum \psi(\gamma_M, \gamma_{\mathcal{I}}^-, \gamma_{\mathcal{I}}^+) \geq T_t u(B).$$

Since, u is a subsolution, that $u(B) \leq T_t u(B)$ is immediate. This proves the first equality; the other one is analogous. So, if we are able to construct a $C^{1,1}$ subsolution by an analogous of Bernard's method, namely $v = \breve{T}_\varepsilon(T_t)u$, for ε sufficiently small, we have $u = v$ on $\mathcal{A}$. $\square$

6.2 Mather set and minimizing measures

Let $\gamma : [0, +\infty] \to M \times \mathcal{I}$ be any trajectory on $M \times \mathcal{I}$. Define a measure $\mu_\gamma^t \in \mathcal{P}(TM \times I)$ by setting, for any $F : TM \times I \to \mathbb{R}$ continuous,

$$\int_{TM \times I} F \, d\mu_\gamma^t = \frac{1}{t} \int_0^t F(\gamma_M(s), \dot{\gamma}_M(s), \gamma_{\mathcal{I}}(s)) \, ds.$$

If the velocities are bounded, then there exists $\mu_\gamma \in P(TM \times \mathcal{I})$ for which

$$\mu_\gamma^t \overset{*}{\rightharpoonup} \mu_\gamma.$$

Also, for any $\theta : M \times \mathcal{I} \to \mathbb{R}$ continuous, define a measure ν_γ^t on $M \times \mathcal{I} \times \mathcal{I}$ by

$$\int_{M \times I \times I} \theta \, d\nu_\gamma^t = \frac{1}{t} \sum_{k=1}^{N_t - 1} \theta\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k), \gamma_{\mathcal{I}}(t_{k+1})\big).$$

We observe that the number of switches grows at most linearly (when γ is minimizing), and then $|\nu_\gamma^t|(M \times \mathcal{I} \times \mathcal{I}) \leq C$. Hence, there exists ν_γ such that

$$\nu_\gamma^t \overset{*}{\rightharpoonup} \nu_\gamma.$$

If, in particular, we choose

$$F(x, v, i) = d_x\varphi(\cdot, i) \cdot v, \text{ and } \theta(x, i, j) = \varphi(x, j) - \varphi(x, i),$$

for any $\varphi \in C^1(M \times \mathcal{I})$, we have

$$\begin{aligned}
\int_{TM\times I} F \, d\mu_\gamma + \int_{M\times I\times I} \theta \, d\nu_\gamma &= \lim_{t\to+\infty} \left[\int_{TM\times I} F \, d\mu_\gamma^t + \int_{M\times I\times I} \theta \, d\nu_\gamma^t \right] \\
&= \lim_t \Bigg[\frac{1}{t} \int_0^t d_x\varphi\big(\gamma_M(s), \gamma_{\mathcal{I}}(s)\big) \cdot \dot\gamma_M(s) \, ds \\
&\qquad + \frac{1}{t} \sum_{k=0}^{N_t-1} \Big(\varphi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k)\big) - \varphi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_{k+1})\big) \Big) \Bigg] \\
&= \lim_t \Bigg[\frac{1}{t} \sum_{k=0}^{N_t} \Big(\varphi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k)\big) - \varphi\big(\gamma_M(t_k), \gamma_{\mathcal{I}}(t_k)\big) \Big) \\
&\qquad + \frac{1}{t} \sum_{k=0}^{N_t-1} \Big(\varphi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_{k+1})\big) - \varphi\big(\gamma_M(t_{k+1}), \gamma_{\mathcal{I}}(t_k)\big) \Big) \Bigg] \\
&= \lim_t \left[\frac{\varphi\big(\gamma_M(t), \gamma_{\mathcal{I}}(t_N)\big) - \varphi\big(\gamma_M(0), \gamma_{\mathcal{I}}(0)\big)}{t} \right] = 0.
\end{aligned} \tag{6.1}$$

As in the classical case, motivated by this, we define:

Definition 99 (Holonomic measures)**.** The pair of measures (μ, ν), with $\mu \in P(TM \times \mathcal{I})$ and ν measure in $M \times \mathcal{I} \times \mathcal{I}$, is said to be *holonomic*, if it satisfies, for every $\varphi \in C^1(M \times \mathcal{I})$,

$$\int_{TM\times I} \big(d_x\varphi(x,v,i)\cdot v\big) \, d\mu(x,v,i) + \int_{M\times I\times I} \big(\varphi(x,j) - \varphi(x,i)\big) \, d\nu(x,i,j) = 0. \tag{6.2}$$

We then write $(\mu, \nu) \in \mathcal{H}$.

We are then interested in minimizing a relaxed version of the optimal switching problem:

$$\int_{TM\times I} \big(L(x,v,i) + c_0\big) \, d\mu(x,v,i) + \int_{M\times I\times I} \psi(x,i,j) \, d\nu(x,i,j), \tag{6.3}$$

among all holonomic pairs $(\mu, \nu) \in \mathcal{H}$. A minimizer $(\mu, \nu) \in \mathcal{H}$ of this problem is called a *minimizing measure*. We write $(\mu, \nu) \in \mathcal{H}_{min}$.

Definition 100. We define the *Mather set* as

$$\tilde{\mathcal{M}} = \overline{\bigcup_{(\mu,\nu)\in\mathcal{H}_{min}} \big(\operatorname{supp}\mu \otimes \operatorname{supp}\nu\big)} \subset TM \times \mathcal{I} \times M \times \mathcal{I} \times \mathcal{I}.$$

Theorem 101. *The Mather set $\tilde{\mathcal{M}}$ is nonempty, and its projection onto M is a compact subset of the projected Aubry set $\mathcal{A}$.*

Proof. First we prove the Mather set is nonempty. Let $\gamma^k : [0, t_k] \to M \times \mathcal{I}$ be a sequence of loops, with $t_k \to +\infty$, such that

$$\mathcal{J}_{t_k}[\gamma^k] \to 0.$$

Such a sequence of minimizing trajectories does exist, because the Aubry set is nonempty. Assume too that γ^k is minimizing, which can be made, since the action of loops is nonnegative. Define a measure μ on $TM \times I \times I$ by setting

$$\int_{TM\times I} F(x,v,i)\ d\mu^k(x,v,i) = \frac{1}{t_k}\int_0^{t_k} F(\gamma_M^k(s), \dot{\gamma}_M^k(s), \gamma_{\mathcal{I}}^k(s))\ ds$$

and define a measure ν on $M \times \mathcal{I} \times \mathcal{I}$ by

$$\int_{M\times I\times I} \varphi(x,i,j)\ d\nu^k(x,i,j) = \frac{1}{t_k}\sum \varphi\big(\gamma_M^k(t_k), \gamma_{\mathcal{I}}^k(t_k), \gamma_{\mathcal{I}}^k(t_{k+1})\big).$$

Since the velocities are bounded, there exists $\mu \in \mathcal{P}(TM \times \mathcal{I})$ such that

$$\mu^k \overset{*}{\rightharpoonup} \mu.$$

Analogously, since φ is continuous, also ν^k are bounded and there exists ν such that

$$\nu^k \overset{*}{\rightharpoonup} \nu.$$

We have that

$$\int_{TM\times I} \big(L(x,v,i) + c_0\big)\ d\mu(x,v,i) + \int_{M\times I\times I} \psi(x,i,j)\ d\nu(x,i,j) = 0,$$

and so $(\mu, \nu) \in \mathcal{H}_{min}$. This shows $\tilde{\mathcal{M}} \neq \emptyset$. Now, given $(\mu, \nu) \in \mathcal{H}_{min}$, if u is a C^1 subsolution, then by the Legendre–Fenchel inequality and the definition of a subsolution, we obtain

$$\begin{aligned}\int_{TM\times I} \big(d_x u(x,i)\cdot v\big)\ d\mu(x,v,i) + \int_{M\times I\times I} \big(u(x,j)-u(x,i)\big)\ d\nu(x,i,j)\\ \leq \int_{TM\times I} \big(L(x,v,i)+c_0\big)\ d\mu(x,v,i) + \int_{M\times I\times I} \psi(x,i,j)\ d\nu(x,i,j) = 0,\end{aligned} \tag{6.4}$$

so that (μ, ν) is minimizing. Moreover, by holonomy, we conclude

$$d_x u(x,i)\cdot v = L(x,v,i)+c_0, \quad \mu\text{–a.e. } (x,v,i).$$

$$u(x,j)-u(x,i) = \psi(x,i,j), \quad \nu\text{–a.e. } (x,i,j).$$

This, in particular, shows that the projected Mather set is contained in the projected Aubry set. Indeed, if it were not, we would be able to construct a strict smooth subsolution at (x, i). $\square$

Appendices

Appendix A

Lowersemicontinuity and compactness

For convenience of the reader, we present in this first appendix, the fundamental results for the direct method of the calculus of variations. The idea is simple and it goes as follows: Suppose we want to minimize the action functional

$$\mathcal{J}[\gamma] = \int_0^t L(\gamma(s), \dot{\gamma}(s))\, ds,$$

over all absolutely continuous curves $\gamma \in AC([0,t]; M)$ with $\gamma(0) = x$, $\gamma(t) = y$.

1. Consider a minimizing sequence: $\gamma^k \in AC([0,t]; M)$ with $\gamma^k(0) = x$, $\gamma^k(t) = y$, and
$$\mathcal{J}[\gamma^k] \to \inf \mathcal{J};$$

2. Assume we obtain a compactness result: Up to a subsequence, there exists $\gamma \in AC([0,t]; M)$ with $\gamma(0) = x$, $\gamma(t) = y$ such that $\gamma^k \to \gamma$;

3. Prove that our functional $\mathcal{J}$ is lower semicontinuous with respect to the convergence in 2:
$$\mathcal{J}[\gamma] \le \liminf_k \mathcal{J}[\gamma^k].$$

If we are able to make this work, then 1, 2, and 3 imply

$$\mathcal{J}[\gamma] = \lim_k \mathcal{J}[\gamma^k] = \inf \mathcal{J}.$$

In what follows, we make this argument rigorous. The material of this appendix is certainly not new, and can be found in many textbooks. See, for instance, [15, Proposition 3.1.4].

Proposition 102 (Compactness)**.** *Suppose $\gamma^k \in AC([0,t];\mathbb{R}^d)$ satisfies*

$$\int_0^t S\big(\|\dot{\gamma}^k\|\big) \le C,$$

for some superlinear function $S : \mathbb{R}^+ \to \mathbb{R}^+$. Assume also that, for some $t_0 \in [0,t]$, $\|\gamma(t_0)\| \le C$. Then, there exists a subsequence γ^{k_j} satisfying:

1. *$\gamma^{k_j} \to \gamma$, for some $\gamma \in AC([0,t];\mathbb{R}^d)$;*

2. *$\dot{\gamma}^{k_j} \rightharpoonup \dot{\gamma}$, weakly in L^1, that is, for any $\phi \in L^\infty([0,t])$, we have*

$$\int_0^t \phi\dot{\gamma}^{k_j} \to \int_0^t \phi\dot{\gamma}.$$

Proof. **Step 1:** $\{\gamma^k\}$ is equicontinuous. Since S is superlinear, for any $k \ge 1$, there exists $C(k) \in \mathbb{R}$ such that

$$S(\tau) \ge k|\tau| - C(k).$$

Then, for any $t_1, t_2 \in [0,t]$, we have

$$\begin{aligned} |\gamma^k(t_2) - \gamma^k(t_1)| &\le \int_{t_1}^{t_2} |\dot{\gamma}^k| \\ &\le \frac{1}{k}\int_{t_1}^{t_2} S(|\dot{\gamma}^k|) + \frac{C(k)}{k}|t_2 - t_1| \\ &\le \frac{C}{k} + \frac{C(k)}{k}|t_2 - t_1| \end{aligned} \tag{A.1}$$

So, given $\varepsilon > 0$, if we first choose k such that $C/k \le \varepsilon/2$, and then choose $\delta > 0$

such that $C(k)\delta/k \leq \varepsilon/2$, we have that

$$|t_2 - t_1| < \delta \implies |\gamma^k(t_2) - \gamma^k(t_1)| < \varepsilon,$$

as desired. In particular, since it is bounded in a point, by Arzelà–Ascoli Theorem, there exists γ continuous for which $\gamma^k \to \gamma$, uniformly.

Step 2: γ is absolutely continuous. Indeed, given a disjoint family of subintervals $[s_i, t_i] \subseteq [0, t]$, we know

$$\begin{aligned}\sum_i |\gamma^k(t_i) - \gamma^k(s_i)| &\leq \sum_i \int_{s_i}^{t_i} \|\dot{\gamma}^k\| \\ &\leq \frac{1}{k}\sum_i \int_{s_i}^{t_i} S(\|\dot{\gamma}^k\|) + \frac{C(k)}{k}\sum_i |t_i - s_i| \\ &\leq \frac{C}{k} + \frac{C(k)}{k}\sum_i |t_i - s_i|\end{aligned} \tag{A.2}$$

Let $k \to +\infty$ to obtain

$$\sum_i |\gamma(t_i) - \gamma(s_i)| \leq \frac{C}{k} + \frac{C(k)}{k}\sum_i |t_i - s_i|$$

Now, given $\varepsilon > 0$, by choosing $\delta > 0$ as in step 1, we prove

$$\sum_i |t_i - s_i| < \delta \implies \sum_i |\gamma(t_i) - \gamma^k(s_i)| < \varepsilon;$$

therefore, γ is absolutely continuous.

Step 3: $\dot{\gamma}^k \rightharpoonup \dot{\gamma}$ weakly in L^1.

It is enough to prove it for characteristic functions of Borel sets. Indeed, these are dense in L^∞. If E is a finite union of disjoint intervals (a_i, b_i)

$$E = \bigcup_{i=1}^{N} (a_i, b_i),$$

then
$$\int_E \dot\gamma^k = \sum_{i=1}^N \left[\gamma^k(b_i) - \gamma^k(a_i)\right] \to \sum_{i=1}^N \left[\gamma(b_i) - \gamma(a_i)\right] = \int_E \dot\gamma.$$

Next, let E be an infinite union of disjoint intervals (a_i, b_i)

$$E = \bigcup_{i=1}^{\infty} (a_i, b_i),$$

and let $\varepsilon > 0$. Since γ is absolutely continuous, there exists $\delta > 0$ for which

$$\sum_{i=i_0}^{\infty} |b_i - a_i| < \delta \implies \sum \|\gamma(b_i) - \gamma(a_i)\| < \varepsilon/2.$$

It is easy to see that such i_0 does exist. Set $E_0 = \bigcup_{i=i_0}^{\infty}(a_i, b_i)$. Then

$$\left\| \int_E (\dot\gamma^k - \dot\gamma) \right\| \leq \left\| \int_{E_0} (\dot\gamma^k - \dot\gamma) \right\| + \left\| \int_{E \setminus E_0} (\dot\gamma^k - \dot\gamma) \right\| \tag{A.3}$$

The first term goes to zero by the previous case, and the second by absolute continuity. Finally, if E is any Borel set, we approximate E by a decreasing sequence of open sets $\mathcal{O} \supseteq E$, and use Lebesgue Convergence Theorem to conclude. □

Proposition 103 (Lower semicontinuity)**.** *Suppose $L : TM \to \mathbb{R}$ is a Lagrangian that is C^1, bounded below, and convex in v. Assume $\gamma^k, \gamma \in AC([0,t]; M)$ satisfy*

$$\gamma^k \to \gamma \text{ uniformly in } [0,t]$$

and

$$\dot\gamma^k \rightharpoonup \dot\gamma \text{ weakly in } L^1.$$

Then

$$\int_0^t L(\gamma, \dot\gamma) \leq \liminf_k \int_0^t L(\gamma^k, \dot\gamma^k).$$

Proof. Assume, without loss of generality, $L \geq 0$. For $\varepsilon > 0$, set

$$F_\varepsilon := \big\{ s \in [0,t] \in M \ ; \ \dot\gamma(s) \text{ exists}, |\dot\gamma(s)| \leq 1/\varepsilon \big\}.$$

It is easy to see that $|[0,t]\backslash F_\varepsilon| \to 0$ as $\varepsilon \to 0$, because $\dot\gamma \in L^1$. By convexity in v, we have

$$L(x,v) \geq L(x,w) + \frac{\partial L}{\partial v}(x,w)\cdot(v-w),$$

and then

$$\int_0^t L(\gamma^k,\dot\gamma^k) \geq \int_{F_\varepsilon} L(\gamma^k,\dot\gamma) + \int_{F_\varepsilon} \frac{\partial L}{\partial v}(\gamma^k,\dot\gamma)\cdot(\dot\gamma^k - \dot\gamma).$$

Now the second term in the right hand side above goes to zero, when $k \to +\infty$, for

$$\frac{\partial L}{\partial v}(\gamma^k,\dot\gamma) \to \frac{\partial L}{\partial v}(\gamma,\dot\gamma) \text{ uniformly}$$

and $\dot\gamma^k \rightharpoonup 0$ weakly in L^1. Since L is locally Lipschitz, we have

$$|L(\gamma^k,\dot\gamma) - L(\gamma,\dot\gamma)| \leq C_\varepsilon|\gamma^k - \gamma|, \text{ on } F_\varepsilon.$$

Hence, by letting $k \to +\infty$, we have

$$\liminf_k \int_0^t L(\gamma^k,\dot\gamma^k) \geq \int_{F_\varepsilon} L(\gamma,\dot\gamma).$$

Since $\varepsilon > 0$ is arbitrary, the proof is finished. □

As a corollary, we obtain Tonelli's Theorem on the existence of minimizers to the action.

Theorem 104 (Tonelli)**.** *Suppose $L : TM \to \mathbb{R}$ is a Lagrangian that is $C^1(TM)$, bounded below, and convex in v. Assume further that*

$$L(x,v) \geq S\big(\,\|v\|\,\big),$$

for some superlinear function S. Then, there exists an absolutely continuous minimizer γ for the action:

$$\mathcal{J}[\gamma] = \inf_\alpha \mathcal{J}[\alpha].$$

Proof. Consider a minimizing sequence: $\gamma^k \in AC([0,t];M)$ with $\gamma^k(0) = x$, $\gamma^k(t) = y$, and

$$\mathcal{J}[\gamma^k] \to \inf \mathcal{J}.$$

Then,

$$\int_0^t S\big(\,\|\dot{\gamma}^k\|\,\big) \le \mathcal{J}[\gamma^k] \le C,$$

so that, by compactness, there exists $\gamma \in AC([0,t];M)$ with $\gamma(0) = x$, $\gamma(t) = y$ such that

$$\gamma^k \to \gamma \text{ uniformly in [0,t]}$$

and

$$\dot{\gamma}^{k_j} \rightharpoonup \dot{\gamma}, \text{ weakly in } L^1.$$

Finally, by lower semicontinuity,

$$\mathcal{J}[\gamma] \le \liminf_k \mathcal{J}[\gamma^k],$$

which in turn implies $\mathcal{J}[\gamma] = \inf_\alpha \mathcal{J}[\alpha]$. □

Appendix B

Semiconvexity and semiconcavity

Here, we recall how a semiconvex function is defined and some of its properties. [3, 15, 16]

Definition 105. Let $O \subseteq \mathbb{R}^n$ be a convex set. We say that a function $f : O \to \mathbb{R}$ is *semiconvex* if there exists $C \in \mathbb{R}$ such that, for each $x \in O$, there exists a linear functional $l_x : \mathbb{R}^n \to \mathbb{R}$ with

$$f(y) - f(x) \geq \langle l_x, y - x \rangle - C|y - x|^2, \tag{B.1}$$

for any $y \in U$.

Proposition 106. *The following are equivalent:*

(i) f is semiconvex;

(ii) There exists $\varphi \in C^2(O)$, with $\|D^2\varphi(x)\| \leq M$, such that $f + \varphi$ is convex;

(iii) There exists $\varphi \in C^\infty(O)$, with $\|D^2\varphi(x)\| \leq M$, such that $f + \varphi$ is convex;

We denote by $\mathcal{C}^2_b(O, \mathbb{R})$ the set of $\varphi \in C^2(O)$ with $\|D^2\varphi(x)\| \leq M$, for some $M > 0$.

Proposition 107. *Let $f : O \to \mathbb{R}$ be a semiconvex function. Then, there exists $\mathcal{F} \subseteq \mathcal{C}^2_b(O, \mathbb{R})$ so that*

$$f = \max_{h \in \mathcal{F}} h. \tag{B.2}$$

Moreover, for every $l \in \partial^- f(x)$*, there exists* $h \in \mathcal{F}$ *such that* $f(x) = h(x)$ *and* $l = df(x)$*.*

Proposition 108. *The function* $f : O \to \mathbb{R}$ *is in* $C^{1,1}(O)$ *if, and only if, it is a locally semiconvex and a locally semiconcave function.*

Proposition 109. *Let* $f : O \to \mathbb{R}$ *be a locally semiconvex function. Then, for any* C^2 *function* $\varphi : U \to O$*, also* $f \circ \varphi : U \to R$ *is locally semiconvex.*

Thanks to Proposition 109, the notion of semiconvexity (and semiconcavity) is successfully adapted to a Riemannian manifold setting. Its properties remain true.

Bibliography

[1] M. Bardi and I. 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.

[2] S.A. Belbas. Optimal switching control of diffusion processes: The associated implicit variational problems. In *Decision and Control including the Symposium on Adaptive Processes, 1981 20th IEEE Conference on*, volume 20, pages 1380–1383, 1981.

[3] P. Bernard. Existence of $C^{1,1}$ critical subsolutions of the Hamilton–Jacobi equation on compact manifolds. *Annales Scientifiques de l'École Normale Supérieure*, 40(3):445–452, 2007.

[4] G. Buttazzo, M. Giaquinta, and S. Hildebrandt. *One-dimensional variational problems*, volume 15 of *Oxford Lecture Series in Mathematics and its Applications.* The Clarendon Press Oxford University Press, New York, 1998.

[5] F. Cagnetti, D. Gomes, H Mitake, and H.V. Tran. A new method for large time behavior of convex Hamilton–Jacobi equations I: degenerate equations and weakly coupled system. *to appear*, 2012.

[6] I. Capuzzo Dolcetta and L. C. Evans. Optimal switching for ordinary differential equations. *SIAM Journal of Control and Optimization*, 22(1):143–161, 1984.

[7] I. Capuzzo Dolcetta, M. Matzeu, and J.-L. Menaldi. On a system of first order quasi-variational inequalities connected with optimal switching problem. *Systems and Control Letters*, 3:113–116, 1983.

[8] F. H. Clarke. *Methods of dynamic and nonsmooth optimization*, volume 57 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1989.

[9] M. G. Crandall and P.-L. Lions. Viscosity solutions of Hamilton-Jacobi equations. *Transactions of the American Mathematical Society*, 277(1):1–42, 1983.

[10] A. Davini and A. Siconolfi. A generalized dynamical approach to the large time behavior of solutions of Hamilton-Jacobi equations. *SIAM Journal of Mathematical Analysis*, 38(2):478–502, 2006.

[11] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[12] L.C. Evans and A. Friedman. Optimal stochastic switching and the Dirichlet problem for the Bellman equation. *Transactions of American Mathematical Society*, 253:365–389, 1979.

[13] A. Fathi. Théorème KAM faible et théorie de Mather sur les systèmes legrangiens. *Comptes Rendus de l'Académie des Sciences*, 324:1043–1046, 1997.

[14] A. Fathi. Sur la convergence du semi–groupe de Lax–Oleinik. *Comptes Rendus de l'Académie des Sciences*, 327:267–270, 1998.

[15] A. Fathi. *Weak KAM Theorem in Lagrangian Dynamics*. Preliminary Version Number 10, June 2008.

[16] A. Fathi and A. Figalli. Optimal transportation on non-compact manifolds. *Israel Journal of Mathematics*, 175(1):1–59, 2010.

[17] A. Fathi and A. Siconolfi. PDE aspects of Aubry–Mather theory for quasiconvex Hamiltonians. *Calculus of Variations*, 22:185–228, 2005.

[18] A. Figalli and L. Rifford. Aubry sets, hamilton-jacobi equations, and mañé conjecture. *Contemporary Mathematics*, to appear.

[19] D. Gomes and A. Serra. Systems of weakly coupled Hamilton-Jacobi equations with implicit obstacles. *Canadian Journal of Mathematics*, 64:1289–1309, 2012.

[20] P.L. Lions, G. Papanicolaou, and S.R.S. Varadhan. Homogenization of Hamilton-Jacobi equations. Unpublished preprint, 1987.

[21] R. Mañé. On the minimizing measures of Lagrangian dynamical systems. *Nonlinearity*, 5(3):623–638, 1992.

[22] R. Mañé. Generic properties and problems of minimizing measures of Lagrangian systems. *Nonlinearity*, 9(2):273–310, 1996.

[23] J. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Mathematische Zeitschrift*, 207(2):169–207, 1991.

[24] J. Mather. Variational construction of connecting orbits. *Université de Grenoble. Annales de l'Institut Fourier*, 43(5):1349–1386, 1993.

[25] L. Rifford. Regularity of weak KAM solutions and Mañé's conjecture. to appear.

[26] L. Tonelli. L'analisi funzionale nel calcolo delle variazioni. *Annali della Scuola Normale Superiore di Pisa*, 9(3):289–302, 1940.