Viscosity Solutions of Hamilton-Jacobi Equations

Publicações Matemáticas

Viscosity Solutions of Hamilton-Jacobi Equations

Diogo Gomes
IST - Lisboa

impa

27º Colóquio Brasileiro de Matemática

Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ
Impresso no Brasil / Printed in Brazil
Capa: Noni Geiger / Sérgio R. Vaz

27º Colóquio Brasileiro de Matemática

- A Mathematical Introduction to Population Dynamics - Howard Weiss
- Algebraic Stacks and Moduli of Vector Bundles - Frank Neumann
- An Invitation to Web Geometry - Jorge Vitório Pereira e Luc Pirio
- Bolhas Especulativas em Equilíbrio Geral - Rodrigo Novinski e Mário Rui Páscoa
- C*-algebras and Dynamical Systems - Jean Renault
- Compressive Sensing - Adriana Schulz, Eduardo A. B. da Silva e Luiz Velho
- Differential Equations of Classical Geometry, a Qualitative Theory - Ronaldo Garcia e Jorge Sotomayor
- Dynamics of Partial Actions - Alexander Arbieto e Carlos Morales
- Introduction to Evolution Equations in Geometry - Bianca Santoro
- Introduction to Intersection Homology and Perverse Sheaves - Jean-Paul Brasselet
- Introdução à Análise Harmônica e Aplicações - Adán J. Corcho Fernandez e Marcos Petrúcio de A. Cavalcante
- Introdução aos Métodos de Decomposição de Domínio - Juan Galvis
- Problema de Cauchy para Operadores Diferenciais Parciais - Marcelo Rempel Ebert e José Ruidival dos Santos Filho
- Simulação de Fluidos sem Malha: Uma Introdução ao Método SPH - Afonso Paiva, Fabiano Petronetto, Geovan Tavares e Thomas Lewiner
- Teoria Ergódica para Autômatos Celulares Algébricos - Marcelo Sobottka
- Uma Iniciação aos Sistemas Dinâmicos Estocásticos - Paulo Ruffino
- Uma Introdução à Geometria de Contato e Aplicações à Dinâmica Hamiltoniana - Umberto L. Hryniewicz e Pedro A. S. Salomão
- **Viscosity Solutions of Hamilton-Jacobi Equations - Diogo Gomes**

ISBN: 978-85-244-0306-4

Distribuição: IMPA
Estrada Dona Castorina, 110
22460-320 Rio de Janeiro, RJ
E-mail: ddic@impa.br
http://www.impa.br

Contents

1

Introduction

This book was written for a course in the 27th Brazilian Mathematics Colloquium. This course covered basic notions in viscosity solutions and its applications to deterministic and stochastic optimal control. This books is partially based on a course on Calculus of Variations and Partial Differential Equations that I have taught over the years at the Mathematics Department of Instituto Superior Técnico. I would like to thank my students: Tiago Alcaria, Patrícia Engrácia, Sílvia Guerra, Igor Kravchenko, Anabela Pelicano, Ana Rita Pires, Verónica Quítalo, Lucian Radu, Joana Santos, Ana Santos, and Vitor Saraiva, which took my courses and suggested me several corrections and improvements. Also my post-doc Andrey Byriuk and my colleagues Pedro Girão and Cláudia Nunes Philipart have suggested numerous improvements on the original text. I would like to thank Artur Lopes that challenged me to present the proposal of this course at IMPA.

The structure of this text is the following: we start with a survey of classical mechanics and classical calculus of variations. Then we present the basic tools in classical optimal control. We continue with a discussion of viscosity solutions both for the terminal value problem, discounted cost infinite horizon and stationary problems. We follow with a brief discussion of stochastic optimal control problems and applications to mathematical finance. Zero sum differential games are also discussed as another applications of viscosity solutions. We then present some applications of viscosity solutions to the Aubry Mather theory. We end the book with a presentation of two important results: the characterization of monotone semigroups and the convergence of numerical algorithms.

Many of the results in this book are not done in the largest possible generality. For additional material, the reader should consult the bibliographical references. In each chapter we have a section on bibliographical notes that lists the main references on the material of that chapter.

2

Classical calculus of variations

This chapter is dedicated to the study of classical mechanics and calculus of variations. We start by discussing the minimum action principle, Euler-Lagrange equations and some applications to Classical Mechanics. In section 2.2 we establish further necessary conditions for minimizers. The following section is dedicated to the Hamiltonian formalism. Then, in section 2.4 we consider sufficient conditions. We follow with section Noether's theorem and symmetries. We end the chapter with some bibliographical notes.

2.1 Euler-Lagrange Equations

In classical mechanics, the trajectories $\mathbf{x}(\cdot) : [0,T] \to \mathbb{R}^n$ of a mechanical system are determined by a variational principle, the minimal action principle, of an integral functional. In this section we discuss this approach and discuss several examples.

Consider a mechanical system on $\mathbb{R}^n$ with kinetic energy $K(x,v)$ and potential energy $U(x,v)$. We define the Lagrangian, $L(x,v) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. to be difference between the kinetic energy K and potential energy U of the system, that is, $L = K - U$. The variational formulation of classical mechanics asserts that trajectories of this mechanical system minimize (or are at least critical points) of the action functional

$$S[\mathbf{x}] = \int_0^T L(\mathbf{x}(t), \dot{\mathbf{x}}(t))dt,$$

under fixed boundary conditions. More precisely, a C^1 trajectory $\mathbf{x}(\cdot) : [0,T] \to \mathbb{R}^n$ is a minimizer of the action functional under fixed boundary conditions if for any other C^1 trajectory $\mathbf{y}(\cdot) : [0,T] \to \mathbb{R}^n$ such that $\mathbf{x}(0) = \mathbf{y}(0)$ and $\mathbf{x}(T) = \mathbf{y}(T)$ we have

$$S[\mathbf{x}] \leq S[\mathbf{y}].$$

In particular, this implies that for any C^1 function $\varphi : [0,T] \to \mathbb{R}^n$ with compact support in $(0,T)$, and any $\epsilon \in \mathbb{R}$ we have

$$i(\epsilon) = S[\mathbf{x} + \epsilon\varphi] \geq S[\mathbf{x}] = i(0).$$

Thus $i(\epsilon)$ has a minimum at $\epsilon = 0$, and so, if $i(\cdot)$ is differentiable, $i'(0) = 0$. A trajectory $\mathbf{x}$ is a *critical point* of S, if for all C^1 function $\varphi : [0,T] \to \mathbb{R}^n$ with compact support in $(0,T)$ we have

$$i'(0) = \frac{d}{d\epsilon} S[\mathbf{x} + \epsilon\varphi]\bigg|_{\epsilon=0} = 0.$$

The critical points of the action which are of class C^2 are solutions to an ordinary differential equation, the *Euler-Lagrange equation*, that we derive in what follows. For minimizers of the action functional, further necessary conditions can be derived as will be discussed in section 2.2.

Theorem 1 (Euler-Lagrange equation)**.** *Let $L(x, v) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a C^2 function. Suppose that $\mathbf{x}(\cdot) : [0, T] \to \mathbb{R}^n$ is a C^2 critical point of the action $S[\mathbf{x}]$ under fixed boundary conditions $\mathbf{x}(0)$ and $\mathbf{x}(T)$. Then*

$$\frac{d}{dt} D_v L(\mathbf{x}, \dot{\mathbf{x}}) - D_x L(\mathbf{x}, \dot{\mathbf{x}}) = 0. \tag{2.1}$$

Proof. Let $\mathbf{x}$ be as in the statement. Then for any $\varphi : [0, T] \to \mathbb{R}^n$ with compact support on $(0, T)$, the function

$$i(\epsilon) = S[\mathbf{x} + \epsilon\varphi]$$

has a minimum at $\epsilon = 0$. Thus

$$i'(0) = 0,$$

that is,

$$\int_0^T D_x L(\mathbf{x}, \dot{\mathbf{x}})\varphi + D_v L(\mathbf{x}, \dot{\mathbf{x}})\dot{\varphi} = 0.$$

Integrating by parts, we conclude that

$$\int_0^T \left[\frac{d}{dt} D_v L(\mathbf{x}, \dot{\mathbf{x}}) - D_x L(\mathbf{x}, \dot{\mathbf{x}})\right] \varphi = 0,$$

for all $\varphi : [0, T] \to \mathbb{R}^n$ with compact support in $(0, T)$. This implies (2.1) and ends the proof of the theorem. □

Example 1. In classical mechanics, the kinetic energy T of a particle with mass m following the trajectory $\mathbf{x}(t)$ is:

$$K = m\frac{|\dot{\mathbf{x}}|^2}{2}.$$

The potential energy $U(x)$ depends only on the position x and we assume that it is a smooth function. The corresponding Lagrangian is then

$$L = K - U.$$

So the Euler-Lagrange equation is

$$m\ddot{\mathbf{x}} = -U'(\mathbf{x}),$$

which is the Newton's law. ◀

Exercise 1. *Let $P \in \mathbb{R}^n$, and consider the Lagrangian $L(x,v) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ defined by $L(x,v) = g(x)|v|^2 + P \cdot v - U(x)$, where g and U are C^2 functions. Determine the Euler-Lagrange equation and show that it does not depend on P.*

To understand the behavior of the Euler-Lagrange equation it is sometimes useful to change coordinates. The following proposition shows how this is achieved:

Proposition 2. *Let $\mathbf{x}(\cdot)$ be a critical point of the action*

$$\int_0^T L(\mathbf{x}, \dot{\mathbf{x}})dt.$$

Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be a C^2 diffeomorphism and $\hat{L}$ given by

$$\hat{L}(y,w) = L(g(y), Dg(y)w).$$

Then $\mathbf{y} = g^{-1} \circ \mathbf{x}$ is a critical point of

$$\int_0^T \hat{L}(\mathbf{y}, \dot{\mathbf{y}})dt.$$

Proof. This is a simple computation and is left as an exercise to the reader. □

Before proceeding, we discuss some applications to classical mechanics. As mentioned before, the trajectories of a mechanical system with kinetic energy K and potential energy U are critical points of the action corresponding to the Lagrangian $L = K - U$. In the following examples we use this variational principle to study the motion of a particle in a central field and the planar two body problem.

Example 2 (Central field motion)**.** Consider the Lagrangian of a particle in the plane subjected to a radial potential field.

$$L(\mathbf{x}, \mathbf{y}, \dot{\mathbf{x}}, \dot{\mathbf{y}}) = \frac{\dot{\mathbf{x}}^2 + \dot{\mathbf{y}}^2}{2} - U(\sqrt{\mathbf{x}^2 + \mathbf{y}^2}).$$

Using polar coordinates, (r, θ). That is $(x, y) = (r\cos\theta, r\sin\theta) = g(r, \theta)$, We can change coordinates (see proposition 2) and obtain

$$\hat{L}(\mathbf{r}, \theta, \dot{\mathbf{r}}, \dot{\theta}) = \frac{\mathbf{r}^2\dot{\theta}^2 + \dot{\mathbf{r}}^2}{2} - U(\mathbf{r}).$$

In these new coordinates the Euler-Lagrange equations can be written as

$$\frac{d}{dt}\mathbf{r}^2\dot{\theta} = 0 \qquad \frac{d}{dt}\dot{\mathbf{r}} = -U'(\mathbf{r}) + \mathbf{r}\dot{\theta}^2.$$

The first equation implies that $\mathbf{r}^2\dot{\theta} \equiv \eta$ is conserved. Therefore $\mathbf{r}\dot{\theta}^2 = \frac{\eta^2}{\mathbf{r}^3}$. Multiplying the second equation by $\dot{\mathbf{r}}$ we get

$$\frac{d}{dt}\left[\frac{\dot{\mathbf{r}}^2}{2} + U(\mathbf{r}) + \frac{\eta^2}{2\mathbf{r}^2}\right] = 0.$$

Consequently

$$E_\eta = \frac{\dot{\mathbf{r}}^2}{2} + U(\mathbf{r}) + \frac{\eta^2}{2\mathbf{r}^2}$$

is a conserved quantity. Thus, one can solve for $\dot{\mathbf{r}}$ as a function of $\mathbf{r}$ (given the values of the conserved quantities E_η and η) and so obtain a first-order differential equation for the trajectories. ◀

Example 3 (Planar two-body problem)**.** Consider now the problem of two point bodies in the plane, describing trajectories $(\mathbf{x}_1, \mathbf{y}_1)$ and

$(\mathbf{x}_2, \mathbf{y}_2)$, whose interaction potential energy U depends only on its distance $\sqrt{(\mathbf{x}_1 - \mathbf{x}_2)^2 + (\mathbf{y}_1 - \mathbf{y}_2)^2}$. We will show how to reduce this problem to the one of a single body under a radial field.

The Lagrangian of this system is

$$L = m_1 \frac{\dot{\mathbf{x}}_1^2 + \dot{\mathbf{y}}_1^2}{2} + m_2 \frac{\dot{\mathbf{x}}_2^2 + \dot{\mathbf{y}}_2^2}{2} - U(\sqrt{(\mathbf{x}_1 - \mathbf{x}_2)^2 + (\mathbf{y}_1 - \mathbf{y}_2)^2}).$$

We will choose new coordinates (X, Y, x, y), where (X, Y) is the center of mass

$$X = \frac{m_1 x_1 + m_2 x_2}{m_1 + m_2} \qquad Y = \frac{m_1 y_1 + m_2 y_2}{m_1 + m_2}$$

and (x, y) the relative position of the two bodies

$$x = x_1 - x_2 \qquad y = y_1 - y_2.$$

In these new coordinates we can write the Lagrangian, using proposition 2,

$$\hat{L} = \hat{L}_1(\dot{\mathbf{X}}, \dot{\mathbf{Y}}) + \hat{L}_2(\mathbf{x}, \mathbf{y}, \dot{\mathbf{x}}, \dot{\mathbf{y}}).$$

Therefore, the equations for the variables $\mathbf{X}$ and $\mathbf{Y}$ decouple from the ones for $\mathbf{x}, \mathbf{y}$. Elementary computations show that

$$\frac{d^2}{dt^2}\mathbf{X} = \frac{d^2}{dt^2}\mathbf{Y} = 0.$$

Thus $\mathbf{X}(t) = X_0 + V_X t$ and $\mathbf{Y}(t) = Y_0 + V_Y t$, for suitable constants V_X and V_Y.

Since

$$L_2 = \frac{m_1 m_2}{m_1 + m_2} \frac{\dot{\mathbf{x}}^2 + \dot{\mathbf{y}}^2}{2} - U(\sqrt{\mathbf{x}^2 + \mathbf{y}^2}),$$

the problem now is reduced to the previous example. ◀

Exercise 2 (Two body problem)**.** *Consider a system of two point bodies in* $\mathbb{R}^3$ *with masses* m_1 *and* m_2*. Assume further that the interaction depends only on the distance between the bodies. Show that by choosing appropriate coordinates, the motion can be reduced to the one of a single point particle with mass* $M = \frac{m_1 m_2}{m_1+m_2}$ *under a radial potential. Show that the orbit of a particle under a radial field lies in a fixed plane for all times, by proving that* $\mathbf{r} \times \dot{\mathbf{r}}$ *is conserved.*

Exercise 3. *Let* $\mathbf{x}(\cdot) : [0,T] \to \mathbb{R}^n$ *be a solution to the Euler-Lagrange equation associated to a* C^2 *Lagrangian* $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$*. Show that*

$$E(t) = L(\mathbf{x}, \dot{\mathbf{x}}) - \dot{\mathbf{x}} D_v L(\mathbf{x}, \dot{\mathbf{x}})$$

is constant in time. For mechanical systems this is simply the conservation of energy. Occasionally, the identity $\frac{d}{dt}E(t) = 0$ *is also called the Beltrami identity.*

Exercise 4. *Consider a system of* n *point bodies of mass* m_i*, and positions* $\mathbf{r}_i \in \mathbb{R}^3$*,* $1 \leq i \leq n$*. Suppose the kinetic energy is* $T = \sum_i \frac{m_i}{2}|\dot{\mathbf{r}}|^2$ *and the potential energy is* $U = -\sum_{i,j\neq i} \frac{m_i m_j}{2|\mathbf{r}_i - \mathbf{r}_j|}$*. Let* $I = \sum_i m_i |\mathbf{r}_i|^2$*. Show that*

$$\frac{d^2}{dt^2} I = 4T + 2U,$$

which is strictly positive if the energy $T + U$ *is positive. What implications does this identity have for the stability of planetary systems?*

Exercise 5 (Jacobi metric)**.** *Let* $L(x, v) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *be a* C^2 *Lagrangian. Let* $\mathbf{x}(\cdot) : [0,T] \to \mathbb{R}^n$ *be a solution to the corresponding Euler-Lagrange*

$$\frac{d}{dt} D_v L - D_x L = 0, \tag{2.2}$$

for the Lagrangian

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

Let $E(t) = \frac{|\dot{\mathbf{x}}(t)|^2}{2} + V(\mathbf{x}(t))$.

1. *Show that* $\dot{E} = 0$.
2. *Let* $E_0 = E(0)$. *Show that* $\mathbf{x}$ *is a solution to the Euler-Lagrange equation*
$$\frac{d}{dt} D_v L_J - D_x L_J = 0 \tag{2.3}$$
associated to $L_J = \sqrt{E_0 - V(\mathbf{x})}|\dot{\mathbf{x}}|$.
3. *Show that any reparametrization of* $\mathbf{x}$ *is also a solution to* (2.3) *and observe that the functional*
$$\int_0^T \sqrt{E_0 - V(\mathbf{x})}|\dot{\mathbf{x}}|$$
represents the distance between $\mathbf{x}(0)$ *and* $\mathbf{x}(T)$ *using the Jacobi metric* $g = \sqrt{E_0 - V(x)}$.
4. *Show that the solutions to the Euler-Lagrange* (2.3) *when reparametrized in suitable way satisfy* (2.2)

Exercise 6 (Braquistochrone problem)**.** *Let* (x_1, y_1) *be a point in a (vertical) plane. Show that the curve* $\mathbf{y} = \mathbf{u}(x)$ *that connects* $(0,0)$ *to* (x_1, y_1) *in such a way that a particle moving under the influence of the gravity* g *reaches* (x_1, y_1) *in the minimum amount of time minimizes*
$$\int_0^{x_1} \sqrt{\frac{1+\dot{\mathbf{u}}^2}{-2g\mathbf{u}}}dx.$$

Hint: *use the fact that the sum of kinetic and potential energy is constant.*

Determine the Euler-Lagrange equation and study its solutions, using exercise 3.

Exercise 7. *Consider a second-order variational problem:*

$$\min_{\mathbf{x}} \int_0^T L(\mathbf{x}, \dot{\mathbf{x}}, \ddot{\mathbf{x}})$$

where the minimum is taken over all trajectories $\mathbf{x}(\cdot) : [0, T] \to \mathbb{R}^n$ *with fixed boundary data* $\mathbf{x}(0), \mathbf{x}(T), \dot{\mathbf{x}}(0), \dot{\mathbf{x}}(T)$*. Determine the Euler-Lagrange equation.*

2.2 Further necessary conditions

A classical strategy in the study of variational problems consists in establishing necessary conditions for minimizers. If there exists a minimizer and if the necessary conditions have a unique solution, then this solution has to be the unique minimizer and thus the problem is solved. In addition to Euler-Lagrange equations, several additional necessary can be derived. In this section we discuss boundary conditions which arise, for instance when the end-points are not fixed, and second-order conditions.

2.2.1 Boundary conditions

In the case in which no boundary conditions are imposed a-priori, it is possible to prove that the minimizers satisfy certain boundary conditions automatically. These boundary conditions called natural boundary conditions.

Example 4. Consider the problem of minimizing the integral

$$\int_0^T L(\mathbf{x}, \dot{\mathbf{x}}) dt, \tag{2.4}$$

over all C^2 curves $\mathbf{x}(\cdot) : [0,T] \to \mathbb{R}^n$. Note that the boundary values for the trajectory $\mathbf{x}(\cdot)$ at $t = 0, T$ are not prescribed a-priori.

Let $\mathbf{x}$ be a minimizer of (2.4) (with free endpoints). Then for all $\varphi : [0,T] \to \mathbb{R}^n$, not necessarily compactly supported,

$$\int_0^T D_x L(\mathbf{x}, \dot{\mathbf{x}})\varphi + D_v L(\mathbf{x}, \dot{\mathbf{x}})\dot{\varphi} dt = 0.$$

Integrating by parts and using the fact that $\mathbf{x}$ is a solution to the Euler-Lagrange equation, we conclude that

$$D_v L(\mathbf{x}(0), \dot{\mathbf{x}}(0)) = D_v L(\mathbf{x}(T), \dot{\mathbf{x}}(T)) = 0.$$

◀

Exercise 8. *Consider the problem of minimizing the integral*

$$\int_0^T L(\mathbf{x}, \dot{\mathbf{x}})dt,$$

over all C^2 curves $\mathbf{x}(\cdot) : [0,T] \to \mathbb{R}^n$ such that $\mathbf{x}(0) = \mathbf{x}(T)$. Deduce that

$$D_v L(\mathbf{x}(0), \dot{\mathbf{x}}(0)) = D_v L(\mathbf{x}(T), \dot{\mathbf{x}}(T)).$$

Exercise 9. *Consider the problem of minimizing*

$$\int_0^T L(\mathbf{x}, \dot{\mathbf{x}})dt + \psi(\mathbf{x}(T)),$$

with $\mathbf{x}(0)$ fixed and $\mathbf{x}(T)$ free. Derive a boundary condition at $t = T$ for the minimizers.

Exercise 10 (Free boundary)**.**

Consider the problem of minimizing

$$\int_0^T L(\mathbf{x}, \dot{\mathbf{x}}),$$

over all terminal times T and all C^2 curves $\mathbf{x} : [0,T] \to \mathbb{R}^n$. Show that $\mathbf{x}$ is a solution to the Euler-Lagrange equation and that

$$L(\mathbf{x}(T), \dot{\mathbf{x}}(T)) = 0,$$
$$D_x L(\mathbf{x}(T), \dot{\mathbf{x}}(T))\dot{\mathbf{x}}(T) + D_v L(\mathbf{x}(T), \dot{\mathbf{x}}(T))\ddot{\mathbf{x}}(T) \geq 0,$$
$$D_v L(\mathbf{x}(T), \dot{\mathbf{x}}(T)) = 0.$$

Let $q \in \mathbb{R}$ and $L : \mathbb{R}^2 \to \mathbb{R}$ given by

$$L(x,v) = \frac{(v-q)^2}{2} + \frac{x^2}{2} - 1$$

If possible, determine T and $\mathbf{x}(\cdot)$ that are (local) minimizers of

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

with $\mathbf{x}(0) = 0$.

2.2.2 Second-order conditions

If $f : \mathbb{R} \to \mathbb{R}$ is a C^2 function which has a minimum at a point x_0 then $f'(x_0) = 0$ and $f''(x_0) \geq 0$. For the minimal action problem, the analog of the vanishing of the first derivative is the Euler-Lagrange equation. We will now consider the analog to the second derivative being non-negative.

The next theorem concerns second-order conditions for minimizers:

Theorem 3 (Jacobi's test)**.** *Let $L(x,v) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a C^2 Lagrangian. Suppose $\mathbf{x}(\cdot) : [0,T] \to \mathbb{R}^n$ is a C^1 minimizer of the action under fixed boundary conditions. Then, for each $\eta : (0,T) \to \mathbb{R}^n$, with compact support in $(0,T)$, we have*

$$\int_0^T \frac{1}{2}\eta^T D^2_{xx} L(\mathbf{x}, \dot{\mathbf{x}})\eta + \eta^T D^2_{xv} L(\mathbf{x}, \dot{\mathbf{x}})\dot{\eta} + \frac{1}{2}\dot{\eta}^T D^2_{vv} L(\mathbf{x}, \dot{\mathbf{x}})\dot{\eta} \geq 0. \quad (2.5)$$

Proof. If $\mathbf{x}$ is a minimizer, the function $\epsilon \mapsto I[\mathbf{x}+\epsilon\eta]$ has a minimum at $\epsilon = 0$. By computing $\frac{d^2}{d\epsilon^2} I[\mathbf{x}+\epsilon\eta]$ at $\epsilon = 0$ we obtain (2.5). □

A corollary of the previous theorem is Lagrange's test that we state next:

Corollary 4 (Lagrange's test)**.** *Let $L(x,v) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a C^2 Lagrangian. Suppose $\mathbf{x}(\cdot) : [0,T] \to \mathbb{R}^n$ is a C^1 minimizer of the action under fixed boundary conditions. Then*

$$D^2_{vv} L(\mathbf{x}, \dot{\mathbf{x}}) \geq 0.$$

Proof. Use Theorem 3 with $\eta = \epsilon\xi(t)\sin\frac{t}{\epsilon}$, for $\xi : (0,T) \to \mathbb{R}^n$, with compact support in $(0,T)$, and let $\epsilon \to 0$. □

Exercise 11. *Let $L : \mathbb{R}^{2n} \to \mathbb{R}$ be a continuous Lagrangian and let $\mathbf{x}(\cdot)$ be a continuous piecewise C^1 trajectory. Show that for each $\delta > 0$ there exists a trajectory $\mathbf{y}_\delta(\cdot)$ of class C^1 such that*

$$\left| \int_0^T L(\mathbf{x}, \dot{\mathbf{x}}) - \int_0^T L(\mathbf{y}_\delta, \dot{\mathbf{y}}_\delta) \right| < \delta.$$

As a corollary, show that the value of the infimum of the action over piecewise C^1 trajectories is the same as the infimum over trajectories globally C^1. Note, however, that the minimizer may not be C^1.

Exercise 12 (Weierstrass test)**.** *Let $\mathbf{x}(\cdot)$ be a C^1 minimum of the action corresponding to a Lagrangian L. Let $v, w \in \mathbb{R}^n$ and $0 \leq \lambda \leq 1$ be such that $\lambda v + (1-\lambda) w = 0$. Show that*

$$\lambda L(\mathbf{x}, \dot{\mathbf{x}} + v) + (1-\lambda) L(\mathbf{x}, \dot{\mathbf{x}} + w) \geq L(\mathbf{x}, \dot{\mathbf{x}}).$$

Hint: *To prove the inequality at a point t_0, choose η such that*

$$\dot{\eta}(t) = \begin{cases} v & \text{if} \quad t_0 \leq t \leq t + \lambda\epsilon \\ w & \text{if} \quad t + \lambda\epsilon < t \leq t_0 + \epsilon \\ 0 & \text{otherwise} \end{cases}$$

and consider $I[\mathbf{x} + \eta]$, as $\epsilon \to 0$.

2.3 Hamiltonian dynamics

In this section we introduce the Hamiltonian formalism of Classical Mechanics. We start by discussing the main properties of the Legendre transform. Then we derive Hamilton's equations. Afterward we discuss briefly the classical theory of canonical transformations. The section ends with a discussion of additional variational principles.

2.3.1 Legendre transform

Before we proceed, we need to discuss the Legendre transform of convex functions. The Legendre transform is used to define the Hamiltonian of a mechanical system and it plays an essential role in many problems in calculus of variations. Additionally, it illustrates many of the tools associated with convexity.

Let $L(v) : \mathbb{R}^n \to \mathbb{R}$ be a convex function, satisfying the following superlinear growth condition:

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

The *Legendre transform* L^* of L is

$$L^*(p) = \sup_v \left[-v \cdot p - L(v)\right].$$

This is the usual definition of Legendre transform in optimal control, see [FS06] or [BCD97]. However, it differs by a sign from the Legendre transform usually used in classical mechanics:

$$L^\sharp(p) = \sup_v \left[v \cdot p - L(v)\right],$$

as it is defined, for instance, in [AKN97] or [Eva98b]. They are related by the elementary identity

$$L^*(p) = L^\sharp(-p).$$

We will frequently denote $L^*(p)$ by $H(p)$. The Legendre transform of H is denoted by H^* and is

$$H^*(v) = \sup_v -p \cdot v - H(p).$$

In classical mechanics, the Lagrangian L can depend also on a position coordinate $x \in \mathbb{R}^n$, $L(x, v)$, but for purposes of the Legendre transform x is taken as a parameter. In this case we write also $H(p, x) = L^*(p, x)$.

Proposition 5. *Let $L(x, v)$ be a C^2 function, which for each x fixed is strictly convex and superlinear in v. Let $H = L^*$. Then*

1. *$H(p, x)$ is convex in p;*
2. *$H^* = L$;*
3. *for each x*
$$\lim_{|p|\to\infty} \frac{H(p, x)}{|p|} = \infty;$$

4. *let* v^* *be defined by* $p = -D_vL(x, v^*)$, *then*

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

5. *in a similar way, let* p^* *be given by* $v = -D_pH(p^*, x)$, *then*

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

6. *if* $p = -D_vL(x, v)$ *or* $v = -D_pH(p, x)$, *then*

$$D_xL(x, v) = -D_xH(p, x).$$

Proof. The first statement follows from the fact that the supremum of convex functions is a convex function. To prove the second point, observe that

$$\begin{aligned} H^*(x, w) &= \sup_p \left[-w \cdot p - H(p, x)\right] \\ &= \sup_p \inf_v \left[(v - w) \cdot p + L(x, v)\right]. \end{aligned}$$

For $v = w$ we conclude that

$$H^*(x, w) \leq L(x, w).$$

The opposite inequality is obtained by observing that, since L is convex in v, for each $w \in \mathbb{R}^n$ there exists $s \in \mathbb{R}^n$ such that

$$L(x, v) \geq L(x, w) + s \cdot (v - w)$$

and, therefore,

$$H^*(x, w) \geq \sup_p \inf_v \left[(p + s) \cdot (v - w) + L(x, w)\right] \geq L(x, w),$$

by letting $p = -s$.

To prove the third point observe that

$$\frac{H(p,x)}{|p|} \geq \lambda - \frac{L(x, -\lambda\frac{p}{|p|})}{|p|},$$

by choosing $v = -\lambda\frac{p}{|p|}$. Thus, we conclude

$$\liminf_{|p|\to\infty} \frac{H(p,x)}{|p|} \geq \lambda.$$

Since λ is arbitrary, we have

$$\liminf_{|p|\to\infty} \frac{H(p,x)}{|p|} = \infty.$$

To establish the fourth point, note that for fixed p the function

$$v \mapsto v \cdot p + L(x,v)$$

is differentiable and strictly convex. Consequently, its minimum, which exists by coercivity and is unique by the strict convexity, is achieved for

$$-p - D_v L(x,v) = 0.$$

Note also that v as function of p is a differentiable function by the inverse function theorem.

The proof of the fifth point is similar.

Finally, to prove the last item, observe that for

$$p(x,v) = -D_v L(x,v),$$

we have

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

Differentiating this last equation with respect to x and using

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

we obtain

$$D_x H = -D_x L.$$

□

Exercise 13. *Compute the Legendre transform of the following functions:*

1.
$$L(x,v) = \frac{1}{2}a_{ij}(x)v_i v_j + h_i(x)v_i - U(x),$$
where a_{ij} is a positive definite matrix and $h(x)$ an arbitrary vector field.

2.
$$L(x,v) = \sqrt{a_{ij}(x)v_i v_j},$$
where a_{ij} is a positive definite matrix.

3.
$$L(x,v) = \frac{1}{2}|v|^\lambda - U(x),$$
with $\lambda > 1$.

2.3.2 Hamiltonian formalism

To motivate the Hamiltonian formalism, we consider the following alternative problem. Rather than looking for curves $\mathbf{x}(\cdot) : [0,T] \to \mathbb{R}^n$, which minimize the action

$$\int_0^T L(\mathbf{x}, \dot{\mathbf{x}})dt$$

we can consider extended curves $(\mathbf{x}(\cdot), \mathbf{v}(\cdot)) : [0,T] \to \mathbb{R}^{2n}$ which minimize the action

$$\int_0^T L(\mathbf{x}, \mathbf{v})dt \tag{2.6}$$

and that satisfy the additional constraint $\dot{\mathbf{x}} = \mathbf{v}$. Obviously, this problem is equivalent to the original one, however it motivates the introduction of a Lagrange multiplier $\mathbf{p}$ in order to enforce the constraint. Therefore, we will look for critical points of

$$\int_0^T L(\mathbf{x}, \mathbf{v}) + \mathbf{p} \cdot (\mathbf{v} - \dot{\mathbf{x}}) dt. \tag{2.7}$$

Proposition 6. *Let $(\mathbf{x}, \mathbf{v})$ be a critical point of (2.6) under fixed boundary conditions and under the constraint $\dot{\mathbf{x}} = \mathbf{v}$ (the choice of $\mathbf{p}$ is irrelevant since the corresponding term always vanishes). Let*

$$\mathbf{p} = -D_v L(\mathbf{x}, \mathbf{v}).$$

Then the curve $(\mathbf{x}, \mathbf{v}, \mathbf{p})$ is a critical point of (2.7) under fixed boundary conditions. Additionally, any critical point $(\mathbf{x}, \mathbf{v}, \mathbf{p})$ of (2.7) satisfies

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{v} \\ \mathbf{p} = -D_v L(\mathbf{x}, \mathbf{v}) \\ \dot{\mathbf{p}} = D_x L(\mathbf{x}, \mathbf{v}), \end{cases}$$

and so $\mathbf{x}$ is a critical point of (2.6). Furthermore, the Euler-Lagrange equation can be rewritten as

$$\dot{\mathbf{p}} = D_x H(\mathbf{p}, \mathbf{x}) \qquad \dot{\mathbf{x}} = -D_p H(\mathbf{p}, \mathbf{x}).$$

Proof. Let ϕ, ψ and η be $C^2([0, T], \mathbb{R}^n)$ with compact support in $(0, T)$. Then, at $\epsilon = 0$

$$\begin{aligned} \frac{d}{d\epsilon} \int_0^T L(\mathbf{x} + \epsilon\phi, \mathbf{v} + \epsilon\psi) + (\mathbf{p} + \epsilon\eta) \cdot (\mathbf{v} - \dot{\mathbf{x}}) + \epsilon(\mathbf{p} + \epsilon\eta) \cdot (\psi - \dot{\phi}) \\ = \int_0^T D_x L(\mathbf{x}, \dot{\mathbf{x}})\phi + D_v L\psi + \mathbf{p} \cdot (\psi - \dot{\phi}) + \eta \cdot (\mathbf{v} - \dot{\mathbf{x}}) \\ = \int_0^T \left[D_x L(\mathbf{x}, \dot{\mathbf{x}}) + \dot{\mathbf{p}}\right] \phi = 0. \end{aligned}$$

If $p = -D_v L(x, v)$, then v maximizes

$$-p \cdot v - L(x, v).$$

Let

$$H(p, x) = \max_v \left[-p \cdot v - L(x, v)\right].$$

By proposition 5 we have

$$D_x H(p, x) = -D_x L(x, v)$$

whenever

$$p = -D_v L(x, v)$$

. Additionally, we also have

$$v = -D_p H(p, x).$$

Therefore, the Euler-Lagrange equation can be rewritten as

$$\dot{\mathbf{p}} = D_x H(\mathbf{p}, \mathbf{x}) \qquad \dot{\mathbf{x}} = -D_p H(\mathbf{p}, \mathbf{x}).$$

These are the *Hamilton equations.* □

Exercise 14. *Suppose $H(p, x) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a C^1 function. Show that the energy, which coincides with H, is conserved by the Hamiltonian flow since*

$$\frac{d}{dt} H(\mathbf{p}, \mathbf{x}) = 0.$$

2.3.3 Canonical transformations

Before discussing canonical transformations we will need to recall some basic facts about differential forms in $\mathbb{R}^n$. Firstly, recall that given a C^1 function $f : \mathbb{R}^n \to \mathbb{R}$ its differential, denoted by df is a mapping $df : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ that for any point $x \in \mathbb{R}^n$ and each direction $v \in \mathbb{R}^n$ it associates the derivative of f in the direction v:

$$df(x)(v) = \frac{d}{dt} f(x + vt)\bigg|_{t=0}.$$

Note that for each $x \in \mathbb{R}^n$ this mapping is linear in v. For example, for each coordinate $i \in \{1, \ldots, n\}$ we can consider the projection $x \mapsto x_i$, whose differential is dx_i.

A (first order) differential form is any mapping

$$\Lambda : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R},$$

linear on the second coordinate. Clearly we can write

$$\Lambda = \sum_i f_i(x) dx_i,$$

where $f_i(x) = \Lambda(x)(e_i)$.

A important example is the differential df of a C^1 function f which is a differential form that can be written as

$$df = \sum_i \frac{\partial f}{\partial x_i} dx_i.$$

The integral of a differential Λ form along a path $\gamma : [0, T] \to \mathbb{R}^n$ is simply

$$\int_0^T \Lambda(\gamma(t))(\dot{\gamma}(t)) dt = \sum_i \int_0^T f_i(\gamma(t)) \dot{\gamma}_i(t) dt.$$

Exercise 15 (Poincaré-Cartan invariant)**.** *Let, for each fixed $t \in \mathbb{R}$,*

$$\gamma = (\mathbf{x}(s, t), \mathbf{p}(s, t)),$$

be a closed curve in $\mathbb{R}^{2n}$ *for* $s \in [0,1]$*. Suppose that*

$$\frac{d}{dt}\mathbf{x} = -D_p H(\mathbf{p}, \mathbf{x}) \qquad \frac{d}{dt}\mathbf{p} = D_x H(\mathbf{p}, \mathbf{x}).$$

Show that

$$\oint \mathbf{p} d\mathbf{x} \equiv \int_0^1 \mathbf{p} \cdot \frac{\partial \mathbf{x}}{\partial s} ds$$

is independent of t*.*

Exercise 16. *Show that the critical points of*

$$\int_0^T \mathbf{p} d\mathbf{x} + H(\mathbf{p}, \mathbf{x}) dt$$

under fixed boundary conditions satisfy the Hamilton equations.

Let $(\mathbf{x}, \mathbf{p})$ be a solution of the Hamilton equation. By exercise 16, $(\mathbf{x}, \mathbf{p})$ is a critical point of

$$\int \mathbf{p} d\mathbf{x} + H dt.$$

Let $S(x) : \mathbb{R}^n \to \mathbb{R}$ be a C^1 function. Then $(\mathbf{x}, \mathbf{p})$ is also a critical point of

$$\int \mathbf{p} d\mathbf{x} + H dt - dS$$

because the last integral differs from the previous only be the addition of the differential of a function S. Consider now a change of coordinates $P(x,p)$, $X(x,p)$. In general the functional $\int pdx + Hdt - dS$ when rewritten in terms of the new coordinates (P, X) does not have the form $\int PdX + \bar{H}(P,X)dt$, and, therefore, the Hamilton equations in these new coordinates may not have the standard form. A change of coordinates $(x,p) \mapsto (X(x,p), P(x,p))$ is called *canonical* if there exist functions S and $\overline{H}(P,X)$ such that

$$pdx + Hdt - dS = PdX + \bar{H}dt. \tag{2.8}$$

Consider now a trajectory $(\mathbf{x}, \mathbf{p})$ of Hamilton's equations. Suppose the coordinate change $(x, p) \mapsto (X(x, p), P(x, p))$ is canonical. Then the trajectory written in the new coordinates $(\mathbf{X}, \mathbf{P})$ is a critical point of the function

$$\int \mathbf{P} d\mathbf{X} + \bar{H} dt.$$

Therefore $(\mathbf{X}, \mathbf{P})$ satisfies Hamilton's equations in the new coordinates, which are

$$\dot{P} = D_X \bar{H}(P, X) \qquad \dot{X} = -D_P \bar{H}(P, X). \tag{2.9}$$

Note that we are looking at time-independent changes of coordinates. Thus in order to have (2.8) we must have

$$H(p, x) = \overline{H}(P(p, x), X(p, x)),$$

and so

$$pdx - PdX = dS.$$

Suppose now we can write the function S as a function of x and X, that is $S \equiv S(x, X)$. Then

$$p = D_x S \qquad P = -D_X S. \tag{2.10}$$

Consider now the inverse procedure. Given $S(x, X)$, suppose that (2.10) defines a change of coordinates (for this to happen locally it is sufficient, by the implicit function theorem that $\det D^2_{xX} S \neq 0$). Then, in these new coordinates we have (2.9). Since S determines (at least formally) the change of coordinates, we call it a *generating function.* ◀

Example 5. Consider the generating function $S(x, X) = xX$. Then the corresponding canonical transformation is $p = X$, $P = -x$, that is $(x, p) \mapsto (X, P) = (p, -x)$ and $\overline{H}(P, X) = H(-P, X)$. ◀

Suppose now that S, written as a function of (x, P), is:

$$S(x, P) = -PX + S_1(x, P).$$

Then (2.8) can be written as:

$$pdx + PdX + XdP - D_x S_1 dx - D_P S_1 dP = PdX,$$

that is,

$$p = D_x S_1 \qquad X = D_P S_1.$$

Example 6. Let $S_1(x, P) = xP$. Then $p = P$ and $X = x$, therefore S_1 generates the identity transformation. ◀

Exercise 17. *Assume now that S can be written as a function of X and p and that we have*

$$S(X, p) = px + S_2(X, p).$$

Determine the corresponding canonical transformation in terms of S_2.

Exercise 18. *Suppose that S can be written as a function of p and P with the following form:*

$$S(p, P) = px - PX + S_3(p, P).$$

Determine corresponding canonical transformation in terms of S_3.

Example 7. Consider the Hamiltonian

$$H \equiv H(p_x, p_y, x - y).$$

Choosing

$$S_1 = P_1(x + y) + P_2(x - y)$$

we obtain

$$p_x = P_1 + P_2 \qquad p_y = P_1 - P_2,$$

$$X_1 = x_1 + x_2 \qquad X_2 = x - y,$$

and

$$\overline{H}(P_1, P_2, X_1, X_2) \equiv \overline{H}(P_1, P_2, X_2) = H\left(\frac{P_1 + P_2}{2}, \frac{P_1 - P_2}{2}, X_2\right),$$

which does not depend on X_1 and, therefore, P_1, the total linear momentum is conserved. ◀

Example 8. Let $S_1(x, P)$ be a C^2 solution of the Hamilton-Jacobi equation

$$H(D_x S_1(x, P), x) = \overline{H}(P).$$

Suppose that

$$X = D_P S_1(x, P) \qquad p = D_x S_1(x, P)$$

defines implicitly a change of coordinates $(x, p) \mapsto (X, P)$. Assume that $\det D^2_{xP} S_1 \neq 0$. Then, if $(\mathbf{x}(t), \mathbf{p}(t))$ satisfy

$$\dot{\mathbf{x}} = -D_p H(\mathbf{p}, \mathbf{x}) \qquad \dot{\mathbf{p}} = D_x H(\mathbf{p}, \mathbf{x}),$$

in the new coordinates we have

$$\dot{\mathbf{X}} = -D_P \overline{H}(\mathbf{P}) \qquad \dot{\mathbf{P}} = 0.$$

◀

Example 9. Consider now a Hamiltonian $H(p, x)$ with one degree of freedom, that is $x \in \mathbb{R}$. We would like to construct a canonical change of coordinates such that the new Hamiltonian depends only on P. We will first construct the corresponding generating function. For that, suppose that there exists a generating function $S_1(x, P)$. Then

$$dS_1 = XdP + pdx.$$

Fix a value P. We will try to choose $S-1$ so that the new Hamiltonian $\overline{H}$ depends only on P, that is $H(p(P,X), x(P,X)) = \overline{H}(P)$. For each curve $\gamma = (\mathbf{x}(\cdot), \mathbf{p}(\cdot))$ such that P is constant, we have

$$dS_1 = pdx.$$

Therefore,

$$S_1(\mathbf{x}(T), P) - S_1(\mathbf{x}(0), P) = \int_0^T \mathbf{p}(t) \cdot \dot{\mathbf{x}}(t) dt.$$

In principle, from the equation $H(p,x) = \overline{H}(P)$ we can solve for p as a function of x and of the value $\overline{H}(P)$. In this case, the generating function is automatically determined as a function of $\overline{H}$ and of x. ◀

Example 10. Consider the Hamiltonian system with one degree of freedom:

$$H(p,x) = \frac{p^2}{2} + V(x),$$

with $V(x)$ 2π-periodic. For each value of $\overline{H}(P)$ we have (assuming for definiteness $p > 0$)

$$S_1(x,P) = \int_0^x \sqrt{2(\overline{H}(P) - V(y))} dy.$$

Therefore,

$$X = \int_0^x \frac{\partial}{\partial \overline{H}} \sqrt{2(\overline{H}(P) - V(y))} D_P \overline{H}(P) dy.$$

In principle, the function $\overline{H}(P)$ can be more or less arbitrary. To impose uniqueness it is convenient to require periodicity in the change of variables

$$X(0,P) = X(2\pi, P),$$

which implies

$$D_P \overline{H}(P) = \left[\frac{\partial}{\partial \overline{H}} \int_0^{2\pi} \sqrt{2\left[\overline{H}(P) - V(y)\right]} dy \right]^{-1}.$$

◀

Exercise 19. *Show that the polar coordinates change of variables* $(x, p) = (r\cos\theta, r\sin\theta)$ *is not canonical. Determine a function* $g(r)$ *such that* $(x, p) = (g(r)\cos\theta, g(r)\sin\theta)$ *is a canonical transformation (for* $r > 0$*).*

2.3.4 Other variational principles

In the case of Hamiltonian systems, as the next exercise shows, there exists an additional variational principle:

Exercise 20. *Show that the critical points* $(\mathbf{x}, \mathbf{p})$ *of the functional*

$$\int_0^T \frac{\mathbf{p}\dot{\mathbf{x}} - \mathbf{x}\dot{\mathbf{p}}}{2} + H(\mathbf{p}, \mathbf{x})$$

are solutions to the Hamilton equation

Unfortunately the functional of the previous exercise is not coercive in $W^{1,2}$ and may not have any minimizer. The Clarke duality principle (following exercise) is another variational principle for convex Hamiltonians which is coercive.

Exercise 21 (Clarke duality). *Let* $H(p, x) : \mathbb{R}^{2n} \to \mathbb{R}$ *be a* C^∞ *function, strictly convex and coercive, both in* x *and* p. *Let* $H^*(\dot{v}_x, \dot{v}_p) : \mathbb{R}^{2n} \to \mathbb{R}$ *be the total Legendre transform*

$$H^*(w_x, w_p) = \sup_{x,p} -w_x \cdot x - w_p \cdot p - H(p, x).$$

Let $(\mathbf{v}_x, \mathbf{v}_p)$ *be a critical point of*

$$\int_0^T \frac{1}{2}\left[\mathbf{v}_x \cdot \dot{\mathbf{v}}_p - \dot{\mathbf{v}}_p \cdot \dot{\mathbf{v}}_x\right] + H^*(\dot{\mathbf{v}}_x, \dot{\mathbf{v}}_p).$$

Show that

$$\mathbf{x} = -D_{\dot{\mathbf{v}}_x} H^*(\dot{\mathbf{v}}_x, \dot{\mathbf{v}}_p) \qquad \mathbf{p} = -D_{\dot{\mathbf{v}}_p} H^*(\dot{\mathbf{v}}_x, \dot{\mathbf{v}}_p)$$

is a solution of Hamilton's equations.

Exercise 22. *Apply the previous exercise to the Hamiltonian*

$$H(p,x) = \frac{p^2+x^2}{2}.$$

Example 11 (Maupertuis principle)**.** Consider a system with Lagrangian L and energy given by

$$E(\mathbf{x},\dot{\mathbf{x}}) = D_vL(\mathbf{x},\dot{\mathbf{x}})\dot{\mathbf{x}} - L(\mathbf{x},\dot{\mathbf{x}}).$$

Since the energy is conserved by the solutions of the Euler-Lagrange equation, the critical points of the action are also critical points of the functional

$$\int_0^T L + E = \int_0^T D_vL(\mathbf{x},\dot{\mathbf{x}})\dot{\mathbf{x}},$$

under the constraint that energy is conserved.

Obviously, in general it is hard to construct energy-preserving variations. We are going to illustrate, in an example, how to avoid this problem. Let L be the Lagrangian

$$L(x,v) = \frac{1}{2}g_{ij}v_iv_j - U(x).$$

Then,

$$E = \frac{1}{2}g_{ij}v_iv_j + U(x)$$

and

$$D_vLv = g_{ij}v_iv_j.$$

Thus we can write

$$D_vLv = 2\left(E - U(x)\right).$$

Therefore the functional can be rewritten as

$$M(\mathbf{x},E) = \int_0^T \sqrt{2\left(E-U(\mathbf{x})\right)}\sqrt{g_{ij}\dot{\mathbf{x}}_i\dot{\mathbf{x}}_j}dt \qquad (2.11)$$

The last term represents the arc length along the curve that connects $\mathbf{x}(0)$ to $\mathbf{x}(T)$. This integral is independent of the parametrization and therefore we can look at its critical points (without any constraint) which obviously depend on the parameter E. Then, once determined, in principle we can choose a parametrization of the curve that preserves the energy. The next exercise shows that such critical points are solutions to the Euler-Lagrange equation:

Exercise 23. *Let* $\mathbf{x}$ *be a critical point of* $M(\mathbf{x}, E_0)$ *parametrized in such a way that*

$$E(\mathbf{x}, \dot{\mathbf{x}}) = E_0.$$

Show that $\mathbf{x}$ *is a solution of the Euler-Lagrange equation.*

◀

2.4 Sufficient conditions

This section addresses a very classical topic in the calculus of variations, namely the study of conditions that ensure that a solution to the Euler-Lagrange equation is indeed a minimizer.

2.4.1 Existence of minimizers

In general, it is not possible to guarantee that a solution to the Euler-Lagrange is a minimizer of the action. However, for short time, the next theorem settles this issue.

Theorem 7 (Existence of minimizers)**.** *Let* $L(x, v)$ *be strictly convex in* v *satisfying*

$$|D^2_{xx}L| \leq C, \qquad |D^2_{xv}L| \leq C.$$

Let $\mathbf{x}(\cdot)$ be a solution to the Euler-Lagrange equation. Then, for T sufficiently small, $\mathbf{x}(\cdot)$ is a minimizer of the action over all C^1 functions $\mathbf{y}(\cdot)$ with the same boundary conditions $\mathbf{y}(0) = \mathbf{x}(0)$, $\mathbf{y}(T) = \mathbf{x}(T)$.

Proof. Observe that if f is a C^2 function then

$$f(1) = f(0) + f'(0) + \int_0^1 \int_0^s f''(r) dr ds.$$

Applying this identity to

$$f(r) = L((1-r)\mathbf{x} + r\mathbf{y}, (1-r)\dot{\mathbf{x}} + r\dot{\mathbf{y}}),$$

we obtain

$$\begin{aligned}
&\int_0^T L(\mathbf{y}, \dot{\mathbf{y}}) dt \\
&= \int_0^T \big[L(\mathbf{x}, \dot{\mathbf{x}}) + D_x L(\mathbf{x}, \dot{\mathbf{x}})(\mathbf{y} - \mathbf{x}) + D_v L(\mathbf{x}, \dot{\mathbf{x}})(\dot{\mathbf{y}} - \dot{\mathbf{x}}) \\
&+ \int_0^1 \int_0^s \Big[(\mathbf{y} - \mathbf{x})^T D^2_{xx} L((1-r)\mathbf{x} + r\mathbf{y}, (1-r)\dot{\mathbf{x}} + r\dot{\mathbf{y}})(\mathbf{y} - \mathbf{x}) \\
&+ 2(\mathbf{y} - \mathbf{x})^T D^2_{xv} L((1-r)\mathbf{x} + r\mathbf{y}, (1-r)\dot{\mathbf{x}} + r\dot{\mathbf{y}})(\dot{\mathbf{y}} - \dot{\mathbf{x}}) \\
&+ (\dot{\mathbf{y}} - \dot{\mathbf{x}})^T D^2_{vv} L((1-r)\mathbf{x} + r\mathbf{y}, (1-r)\dot{\mathbf{x}} + r\dot{\mathbf{y}})(\dot{\mathbf{y}} - \dot{\mathbf{x}}) \Big] dr ds \big] dt.
\end{aligned}$$

Since $\mathbf{x}(\cdot)$ satisfies the Euler-Lagrange equation and, by strict convexity, $D^2_{vv} L \geq \gamma$, we have

$$\begin{aligned}
&\int_0^T L(\mathbf{y}, \dot{\mathbf{y}}) dt \geq \int_0^T \big[L(\mathbf{x}, \dot{\mathbf{x}}) \\
&+ \int_0^1 \int_0^s \big((\mathbf{y} - \mathbf{x})^T D^2_{xx} L((1-r)\mathbf{x} + r\mathbf{y}, (1-r)\dot{\mathbf{x}} + r\dot{\mathbf{y}})(\mathbf{y} - \mathbf{x}) \\
&+ 2(\mathbf{y} - \mathbf{x})^T D^2_{xv} L((1-r)\mathbf{x} + r\mathbf{y}, (1-r)\dot{\mathbf{x}} + r\dot{\mathbf{y}})(\dot{\mathbf{y}} - \dot{\mathbf{x}}) \big) dr ds \\
&+ \gamma |\dot{\mathbf{y}} - \dot{\mathbf{x}}|^2 \big] dt.
\end{aligned}$$

The one-dimensional Poincaré inequality implies

$$\int_0^T |\mathbf{y}-\mathbf{x}|^2 dt \leq \frac{T^2}{2}\int_0^T |\dot{\mathbf{y}}-\dot{\mathbf{x}}|^2 dt,$$

that is,

$$\begin{aligned}
&\int_0^T \int_0^1 \int_0^s (\mathbf{y}-\mathbf{x})^T \cdot \\
&\quad \cdot D^2_{xx} L((1-r)\mathbf{x}+r\mathbf{y},(1-r)\dot{\mathbf{x}}+r\dot{\mathbf{y}})(\mathbf{y}-\mathbf{x}) dr ds dt \\
&\geq -CT^2 \int_0^T |\dot{\mathbf{y}}-\dot{\mathbf{x}}|^2,
\end{aligned}$$

and, for any ϵ,

$$\begin{aligned}
&\int_0^T \int_0^1 \int_0^s (\mathbf{y}-\mathbf{x})^T \cdot \\
&\quad \cdot D^2_{vx} L((1-r)\mathbf{x}+r\mathbf{y},(1-r)\dot{\mathbf{x}}+r\dot{\mathbf{y}})(\dot{\mathbf{y}}-\dot{\mathbf{x}}) dr ds dt \\
&\geq -\epsilon \int_0^T |\dot{\mathbf{y}}-\dot{\mathbf{x}}|^2 - \frac{C}{\epsilon}\int_0^T |\mathbf{y}-\mathbf{x}|^2 \\
&\geq -\left(\epsilon + \frac{CT^2}{\epsilon}\right)\int_0^T |\dot{\mathbf{y}}-\dot{\mathbf{x}}|^2.
\end{aligned}$$

Thus, choosing T sufficiently small and taking, $\epsilon = T$ we obtain

$$\int_0^T L(\mathbf{y},\dot{\mathbf{y}})dt \geq \int_0^T L(\mathbf{x},\dot{\mathbf{x}}) + \theta \int_0^T |\dot{\mathbf{y}}-\dot{\mathbf{x}}|^2,$$

for some $\theta > 0$. □

Exercise 24. *Prove the one-dimensional Poincaré inequality*

$$\int_0^T \phi^2 \leq \frac{T^2}{2}\int_0^T |\dot{\phi}|^2$$

for all C^1 function ϕ satisfying $\phi(0) = \phi(T) = 0$.

Exercise 25. *Suppose that the Lagrangian L instead of satisfying*

$$|D^2_{xx}L| \leq C, \qquad |D^2_{xv}L| \leq C,$$

as in theorem 7, satisfies

$$|D^2_{xx}L| \leq C(1+|v|^2), \qquad |D^2_{xv}L| \leq C(1+|v|).$$

Assume further that the curves $\mathbf{y}$ *are constrained to have bounded derivatives in L^2. Can you adapt theorem 7 to include this case?*

2.4.2 Existence and regularity of minimizers

In this section we assume that the Lagrangian $L(x,v)$ is C^∞, strictly convex in v, satisfies

$$-C+\theta|v|^2 \leq L(x,v) \leq C(1+|v|^2), \tag{2.12}$$

for $\theta > 0$, and that, for each fixed compact K and $x \in K$ we have

$$|D_xL(x,v)| \leq C_K(1+|v|^2), \quad \text{and} \qquad |D_vL(x,v)| \leq C_K(1+|v|).$$

Theorem 8. *Suppose L satisfies the previous assumptions. Then, for each $T>0$ and x_0, x_1 in $\mathbb{R}^n$, there exists a minimizer of $\mathbf{x} \in W^{1,2}[0,T]$ of*

$$\int_0^T L(\mathbf{x}, \dot{\mathbf{x}})ds \tag{2.13}$$

satisfying $\mathbf{x}(0) = x_0$, $\mathbf{x}(T) = x_1$.

Proof. Let $\mathbf{x}_n$ be a minimizing sequence. Then, using (2.12) we conclude that $\|\dot{\mathbf{x}}_n\|_{L^2}$ is uniformly bounded. Then, by Poincaré inequality, we conclude that

$$\sup_n \|\mathbf{x}_n\|_{W^{1,2}} < \infty.$$

By Morrey's theorem, the sequence $\mathbf{x}_n$ is equicontinuous and bounded (since $\mathbf{x}_n(0)$ is fixed), thus there exists, by Ascoli-Arzela theorem, a subsequence which converges uniformly. We can extract a further subsequence that converges weakly in $W^{1,2}$ to a function $\mathbf{x}$. We would like to prove that $\mathbf{x}$ is a minimum. To do that it is enough to prove that the functional is weakly lower semicontinuous, that is, that

$$\liminf_{n\to\infty} \int_0^T L(\mathbf{x}_n, \dot{\mathbf{x}}_n) \geq \int_0^T L(\mathbf{x}, \dot{\mathbf{x}}).$$

By convexity,

$$\begin{aligned}&\int_0^T L(\mathbf{x}_n, \dot{\mathbf{x}}_n)\\ &\geq \int_0^T L(\mathbf{x}_n, \dot{\mathbf{x}}_n) - L(\mathbf{x}, \dot{\mathbf{x}}_n) + L(\mathbf{x}, \dot{\mathbf{x}}) + D_v L(\mathbf{x}, \dot{\mathbf{x}})(\dot{\mathbf{x}}_n - \dot{\mathbf{x}})\end{aligned}$$

Because $\dot{\mathbf{x}}_n \rightharpoonup \dot{\mathbf{x}}$ we have

$$\int_0^T D_v L(\mathbf{x}, \dot{\mathbf{x}})(\dot{\mathbf{x}}_n - \dot{\mathbf{x}}) \to 0,$$

since $D_v L(\mathbf{x}, \dot{\mathbf{x}}) \in L^2$. From the uniform convergence of $\mathbf{x}_n$ to $\mathbf{x}$ we conclude that

$$\int_0^T L(\mathbf{x}_n, \dot{\mathbf{x}}_n) - L(\mathbf{x}, \dot{\mathbf{x}}_n) \to 0,$$

since

$$|L(\mathbf{x}_n, \dot{\mathbf{x}}_n) - L(\mathbf{x}, \dot{\mathbf{x}}_n)| \leq C_K |\mathbf{x}_n - \mathbf{x}|(1 + |\dot{\mathbf{x}}_n|^2).$$

□

Theorem 9. *Let* $\mathbf{x}$ *be a minimizer of* (2.13)*. Then* $\mathbf{x}$ *is a weak solution to the Euler-Lagrange equation, that is, for all* $\varphi \in C_c^\infty(0,T)$*,*

$$\int_0^T D_x L(\mathbf{x}, \dot{\mathbf{x}})\varphi + D_v L(\mathbf{x}, \dot{\mathbf{x}})\dot{\varphi} = 0. \tag{2.14}$$

Proof. To obtain this result, it is enough to prove that at $\epsilon = 0$,

$$\frac{d}{d\epsilon}\int_0^T L(\mathbf{x}+\epsilon\varphi, \dot{\mathbf{x}}+\epsilon\dot{\varphi})\Big|_{\epsilon=0} = \int_0^T \frac{d}{d\epsilon}L(\mathbf{x}+\epsilon\varphi, \dot{\mathbf{x}}+\epsilon\dot{\varphi})\Big|_{\epsilon=0},$$

that is, justify the exchange of the derivative with the integral.

By Morrey's theorem, since $\mathbf{x} \in W^{1,2}(0,T)$, we have $\|\mathbf{x}\|_{L^\infty} \leq C$. So $\mathbf{x} \in K$ for a suitable compact set K. Let $|\epsilon| < 1$. Observe that there exists a compact $\tilde{K} \supset K$ such that $\mathbf{x} + \epsilon\varphi \in \tilde{K}$ for all t. For almost every $t \in [0,T]$, the function

$$\epsilon \mapsto L(\mathbf{x}+\epsilon\varphi, \dot{\mathbf{x}}+\epsilon\dot{\varphi})$$

is a C^1 function of ϵ. Furthermore

$$|L(\mathbf{x}+\epsilon\varphi, \dot{\mathbf{x}}+\epsilon\dot{\varphi})| \leq C_{\tilde{K}}(1+|\dot{\mathbf{x}}+\epsilon\dot{\varphi}|^2) \leq C_{\tilde{K}}(1+|\dot{\mathbf{x}}|^2+|\dot{\varphi}|^2),$$

and,

$$\left|\frac{d}{d\epsilon}L(\mathbf{x}+\epsilon\varphi, \dot{\mathbf{x}}+\epsilon\dot{\varphi})\right| \leq C_{\tilde{K}}(1+|\dot{\mathbf{x}}|^2+|\dot{\varphi}|^2)(|\varphi|+|\dot{\varphi}|).$$

This estimate allows us to exchange the derivative with the integral. □

Exercise 26. *Show that the identity (2.14) also holds for* $\varphi \in W_0^{1,2}$.

Theorem 10. *If L satisfies (2.12) and is strictly convex, then the weak solutions to the Euler-Lagrange equation are C^2 and, therefore, classical solutions.*

Proof. Let $\mathbf{x} \in W^{1,2}(0,T)$ be a weak solution to the Euler-Lagrange equation. Define

$$\mathbf{p}(t) = p_0 + \int_t^T D_x L(\mathbf{x}, \dot{\mathbf{x}})ds,$$

with $p_0 \in \mathbb{R}^n$ to be chosen later. For each $\varphi \in C_c^\infty(0,T)$ taking values in $\mathbb{R}^n$ we have

$$\int_0^T \frac{d}{dt}(\mathbf{p}\cdot\varphi)dt = p\cdot\varphi\Big|_0^T = 0.$$

Thus,

$$\int_0^T -D_x L(\mathbf{x},\dot{\mathbf{x}})\varphi + \mathbf{p}\dot{\varphi}dt = 0.$$

Using the Euler-Lagrange equation in the weak form we conclude that

$$\int_0^T (\mathbf{p} + D_v L(\mathbf{x},\dot{\mathbf{x}}))\dot{\varphi}dt = 0,$$

which implies that $\mathbf{p} + D_v L$ is constant, that is,

$$\mathbf{p} = -D_v L(\mathbf{x},\dot{\mathbf{x}}),$$

choosing p_0 conveniently. Since $\mathbf{p}$ is continuous, by the previous identity, $\dot{\mathbf{x}} = -D_p H(\mathbf{p},\mathbf{x})$. Therefore, $\dot{\mathbf{x}}$ is continuous. Moreover, if $H(p,x)$ is the Hamiltonian associated to L, we have

$$\dot{\mathbf{p}} = D_x H(\mathbf{p},\mathbf{x}),$$

which shows that $\mathbf{p}$ is C^1. But, since

$$\dot{\mathbf{x}} = -D_p H(\mathbf{p},\mathbf{x}),$$

we have that $\dot{\mathbf{x}}$ is C^1 and, therefore, $\mathbf{x}$ is C^2. □

2.5 Symmetries and Noether theorem

Noether's theorem concerns variational problems which admit symmetries. By this theorem, associated to each symmetry there is a quantity that is conserved by the solutions of the Euler-Lagrange equation. In classical mechanics, for instance, translation symmetry yields conservation of linear momentum, to rotation symmetry corresponds conservation of angular momentum and time-invariance implies energy conservation.

2.5.1 Routh's method

We start the discussion of symmetries by considering a classical technique to simplify the Euler-Lagrange equations. Consider a Lagrangian of the form $L(\mathbf{x}, \dot{\mathbf{x}}, \dot{\mathbf{y}})$, that is, independent of the coordinate $\mathbf{y}$. Note that this corresponds to translation invariance in the coordinate y. The Euler-Lagrange equation shows that

$$\mathbf{p}_y = -D_{\dot{\mathbf{y}}} L(\mathbf{x}, \dot{\mathbf{x}}, \dot{\mathbf{y}})$$

is constant. We will explore this fact to simplify the Euler-Lagrange equations. We assume further that $w \mapsto L(\mathbf{x}, \dot{\mathbf{x}}, w)$ is strictly convex and superlinear. Then we define the partial Legendre transform with respect to $\dot{\mathbf{y}}$, Routh's function, as

$$R(\mathbf{x}, \dot{\mathbf{x}}, \mathbf{p}_y) = \sup_w -\mathbf{p}_y \cdot w - L(\mathbf{x}, \dot{\mathbf{x}}, w).$$

By convexity, the supremum is achieved at a unique point $w(\mathbf{x}, \dot{\mathbf{x}}, \mathbf{p}_y)$.

We have that

$$\mathbf{p}_y = -D_w L \qquad \dot{\mathbf{y}} = -D_{\mathbf{p}_y} R.$$

Note that, by the Euler-Lagrange equation

$$\dot{\mathbf{p}}_y = 0$$

and,

$$\begin{aligned}\frac{d}{dt}\frac{\partial R}{\partial \dot{\mathbf{x}}} - \frac{\partial R}{\partial \mathbf{x}} =& -\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{x}}} + \frac{\partial L}{\partial \mathbf{x}} - \frac{d}{dt}\left[\frac{\partial L}{\partial w}\frac{\partial w}{\partial \dot{\mathbf{x}}} + \mathbf{p}_y \frac{\partial w}{\partial \dot{\mathbf{x}}}\right] \\ &+ \frac{\partial L}{\partial w}\frac{\partial w}{\partial x} + \mathbf{p}_y \frac{\partial w}{\partial x} \\ =& \frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{x}}} - \frac{\partial L}{\partial \mathbf{x}} = 0.\end{aligned}$$

Therefore, since $\mathbf{p}_y$ is constant, we can solve these equations in the following way: for each fixed $\mathbf{p}_y$ consider the equation

$$\frac{d}{dt}\frac{\partial R}{\partial \dot{\mathbf{x}}} - \frac{\partial R}{\partial \mathbf{x}} = 0.$$

Once this equation is solved, determine $\dot{\mathbf{y}}$ through

$$\dot{\mathbf{y}} = -D_{\mathbf{p}_y} R(\mathbf{x}, \dot{\mathbf{x}}, \mathbf{p}_y).$$

Exercise 27. *Apply Routh's method to the Lagrangian*

$$L = \frac{\dot{\mathbf{x}}^2}{2} + \frac{\dot{\mathbf{y}}^2}{2} - U(x).$$

Exercise 28. *Apply Routh's method to the symmetric to in an external field which has as Lagrangian*

$$L = \frac{I_1}{2}(\dot{\theta}^2 + \dot{\varphi}^2 \sin^2\theta) + \frac{I_3}{2}(\dot{\psi} + \dot{\varphi}\cos\theta)^2 - U(\varphi, \theta).$$

Exercise 29. *Apply Routh's method to the spherical pendulum whose Lagrangian is:*

$$L = \frac{\dot{\theta}^2 \sin^2\varphi + \dot{\varphi}^2}{2} - U(\varphi).$$

2.5.2 Noether theorem

As a motivation for the definition of invariance of a Lagrangian with respect to a transformation group, observe that if $\phi : \mathbb{R}^n \to \mathbb{R}^n$ is a diffeomorphism and $\gamma : [0,T] \to \mathbb{R}^n$ is an arbitrary curve, then $\phi(\gamma)$ is another curve in $\mathbb{R}^n$ whose velocity is $D_x\phi(\gamma)\dot{\gamma}$. We say that a Lagrangian $L(x,v)$ is invariant under a transformation group $\phi_\tau(x)$ if for each $\tau \in \mathbb{R}$

$$L(x,v) = L(\phi_\tau(x), D_x\phi_\tau(x)v).$$

Theorem 11. *Let L be a Lagrangian invariant under a transformation group $\phi_\tau(x)$. Let $\mathbf{x}$ be a solution of the Euler-Lagrange equation. then*

$$D_vL(\mathbf{x}_\tau(T), \dot{\mathbf{x}}_\tau(T)) \left. \frac{d}{d\tau}\phi_\tau(\mathbf{x}(T))\right|_{\tau=0}$$

is independent of T.

Proof. Let $\mathbf{x}$ be a solution of the Euler-Lagrange equation and

$$\mathbf{x}_\tau(t) = \phi_\tau(\mathbf{x}(t)).$$

Then

$$\dot{\mathbf{x}}_\tau = D_x\phi_\tau(\mathbf{x}(t))\dot{\mathbf{x}}(t).$$

Consequently,

$$\int_0^T L(\mathbf{x}_\tau, \dot{\mathbf{x}}_\tau) \tag{2.15}$$

is constant in τ. Differentiation (2.15) with respect to τ we obtain

$$\int_0^T D_xL(\mathbf{x}_\tau, \dot{\mathbf{x}}_\tau)\frac{d\mathbf{x}_\tau}{d\tau} + D_vL(\mathbf{x}_\tau, \dot{\mathbf{x}}_\tau)\frac{d\dot{\mathbf{x}}_\tau}{d\tau} = 0.$$

Integrating by parts, using the Euler-Lagrange equation, and taking $\tau = 0$ we obtain

$$D_vL(\mathbf{x}_\tau(0), \dot{\mathbf{x}}_\tau(0)) \left. \frac{d}{d\tau}\phi_\tau(\mathbf{x}(0))\right|_{\tau=0}$$
$$= D_vL(\mathbf{x}_\tau(T), \dot{\mathbf{x}}_\tau(T)) \left. \frac{d}{d\tau}\phi_\tau(\mathbf{x}(T))\right|_{\tau=0}.$$

∎

Exercise 30. *Let $\omega \in \mathbb{R}^n$ and $L(x,v)$ be a Lagrangian satisfying, for all τ, $L(x+\omega\tau, v) = L(x,v)$. Show that $D_vL\cdot\omega$ is a constant of motion.*

Exercise 31. *Let $L(x,y,v_x,v_y) = \frac{v_x^2+v_y^2}{2} - \frac{x^2+y^2}{2}$. Show that L is invariant by rotations and, using Noether's theorem, that the angular momentum $xv_y - yv_x$ is a constant of motion.*

Theorem 12. *Suppose L is a Lagrangian which does not depend on t. Then the energy is conserved.*

Proof. Observe that

$$\int_h^{T+h} L(\mathbf{x}(t-h), \dot{\mathbf{x}}(t-h))dt$$

is independent on h. Differentiate with respect to h, integrate by parts using the Euler-Lagrange equation. □

Example 12. Consider the Lagrangian

$$L = \frac{\dot{\mathbf{x}}^2 + \dot{\mathbf{y}}^2}{2\mathbf{y}^2},$$

corresponding to the geodesic flow in the Lobatchewski plane. Identifying the upper semi-plane with $\{z \in \mathbb{C} : \Im(z) > 0\}$ and the points (x,y) with $z = x+iy$, the mapping

$$z \mapsto \frac{az+b}{cz+d}$$

defines an action of the group $SL(2,\mathbb{R})$, the group of matrices with unit determinant, in the Lobatchewski plane, which leaves the Lagrangian invariant. Use matrices of the form

$$A_1(\tau) = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}, \quad A_2(\tau) = \begin{bmatrix} e^{\tau} & 0 \\ 0 & e^{-\tau} \end{bmatrix} \text{ e } \quad A_3(\tau) = \begin{bmatrix} 1 & 0 \\ \tau & 1 \end{bmatrix},$$

we obtain the conservation laws

$$\frac{\dot{\mathbf{x}}}{\mathbf{y}^2}, \quad \frac{\mathbf{x}\dot{\mathbf{x}}+\mathbf{y}\dot{\mathbf{y}}}{\mathbf{y}^2} \text{ and } \frac{\dot{\mathbf{x}}(\mathbf{x}^2-\mathbf{y}^2)+2\dot{\mathbf{y}}\mathbf{x}\mathbf{y}}{\mathbf{y}^2}.$$

◀

Exercise 32. *Obtain the general law $F(\mathbf{x},\mathbf{y}) = 0$ of motion of a geodesic in the Lobatchewski plane.*

2.5.3 Monotonicity formulas

A sub-symmetry (resp. super-symmetry) of L is a one-parameter mapping $\phi_\tau(x)$ such that

$$\frac{d}{d\tau}L(\phi_\tau(x), D_x\phi_\tau(x)v)\bigg|_{\tau=0} \leq 0 \quad (\text{resp. } \geq 0).$$

A simple variation of the proof of Noether's theorem yields:

Theorem 13. *Let ϕ_τ be a sub-symmetry of L. Then*

$$\frac{d}{dt}\left[D_vL(\mathbf{x},\dot{\mathbf{x}})\,\frac{d}{d\tau}\phi_\tau(\mathbf{x})\bigg|_{\tau=0}\right] \leq 0,$$

with the opposite inequality for super-symmetries.

Proof. It suffices to observe that

$$\begin{aligned} 0 \geq \frac{d}{d\tau}\int_0^T L(\phi_\tau(\mathbf{x}), D_x\phi_\tau(\mathbf{x})\dot{\mathbf{x}})dt\bigg|_{\tau=0} \\ = \int_0^T D_xL(\mathbf{x},\dot{\mathbf{x}})\,\frac{d}{d\tau}\phi_\tau(\mathbf{x})\bigg|_{\tau=0} + D_vL(\mathbf{x},\dot{\mathbf{x}})\frac{d}{dt}\,\frac{d}{d\tau}\phi_\tau(\mathbf{x})\bigg|_{\tau=0} \\ = D_vL(\mathbf{x},\dot{\mathbf{x}})\frac{d}{d\tau}\phi_\tau(\mathbf{x})\bigg|_{\tau=0}\bigg|_0^T, \end{aligned}$$

which then implies the result. □

An application of this theorem is the following corollary:

Corollary 14. *Let $L(x,v)\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be Lagrangian admitting a strict sub-symmetry. Then the corresponding Euler-Lagrange equations cannot have periodic orbits.*

Next we present some additional examples and applications. Suppose, for some $y \in \mathbb{R}^n$ and $h \geq 0$, $L(x+hy,v) \leq L(x,v)$, then

$$\frac{d}{dt} D_v L(\mathbf{x},\dot{\mathbf{x}})y \leq 0.$$

Another simple example is the case in which $L(\lambda x, \lambda v)$ is increasing in λ. Then

$$\frac{d}{dt} D_v L(\mathbf{x},\dot{\mathbf{x}})\mathbf{x} \geq 0.$$

Consider the mapping $\phi_\tau(x) = x + \tau F(x)$, and assume that

$$\frac{d}{d\tau} L(x+\tau F(x), v + \tau D_x F v) \leq 0,$$

at $\tau = 0$. Then

$$\frac{d}{dt} D_v L(\mathbf{x},\dot{\mathbf{x}})F(\mathbf{x}) \leq 0.$$

Consider the case $L = \frac{|v|^2}{2}$, and $F = \nabla U$, for some concave function U. Then

$$\frac{d}{d\tau} \frac{|(I+\tau D^2 U)v|^2}{2}\bigg|_{\tau=0} = v^T D^2 U v \leq 0.$$

Thus

$$\frac{d}{dt} \nabla U \cdot v \leq 0,$$

that is

$$\frac{d^2}{dt^2} U(\mathbf{x}) \leq 0,$$

that is $U(\mathbf{x}(t))$ is a concave function.

Consider now a system of non-interacting n-particles, and set

$$U = \sum_{i \neq j} |x_i - x_j|.$$

Clearly U is a convex function. By the previous results we have

$$\frac{d^2}{dt^2} |\mathbf{x}_i - \mathbf{x}_j| \geq 0.$$

Exercise 33. *Consider a Lagrangian of the form*

$$e^{-\alpha t} L(x, v)$$

This Lagrangian is sub-invariant in time. Prove that

$$\frac{d}{dt} e^{-\alpha t} E(t) \geq 0,$$

where

$$E = D_v L(\mathbf{x}, \dot{\mathbf{x}}) \dot{\mathbf{x}} - L(\mathbf{x}, \dot{\mathbf{x}}).$$

In particular, show that this estimate yields exponential blow up of the energy. Also observe that the exponential blow up of the kinetic energy can also be bounded using simple estimates by $E(t) \leq C e^{\beta t}$.

2.6 Bibliographical notes

There is a very large literature on the topics of this chapter. The main references we have used were [Arn95] and [AKN97]. Two classical physics books on this subject are [Gol80] and [LL76]. On the more geometrical perspective, the reader may want to look at [dC92] (see also [dC79]) and [Oli02]. Some aspects of classical calculus of variations can be found in [Dac09] and the classical book [Bol61]. In what concerns symmetries, additional material can be consulted in [Olv93]. A very good reference in Portuguese is [Lop06].

3

Classical optimal control

In this chapter we begin the study of deterministic optimal control problems, and its connection with Hamilton-Jacobi equations. We start the discussion in the next section with the set up of the problem. Then we present some elementary properties and examples. The dynamic programming principle and Pontryangin maximum principles are discussed in sections 3.3 and 3.4, respectively.imum principles are discussed in sections 3.3 and 3.4, respectively. The Pontryagin maximum principle is the analog of the Euler-Lagrange equation for optimal control problems. Then, in section 3.5 we will show that if the value function V is differentiable, it satisfies the Hamilton-Jacobi partial differential equation

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

in which $H(p,x)$, the Hamiltonian, is the (generalized) Legendre transform of the Lagrangian L

$$H(p,x) = \sup_{v \in U} -p \cdot f(x,v) - L(x,v). \tag{3.1}$$

We end this chapter with a verification theorem, section 3.6 that establishes that a sufficiently smooth solution to the Hamilton-Jacobi equation is the value function.

3.1 Optimal Control

A typical problem in optimal control, whose study we begin now is the terminal value optimal control problem. For that let the control space be a closed convex subset U of $\mathbb{R}^m$. A control on an interval $I \subset \mathbb{R}$ is a measurable function $\mathbf{u} : I \to U$. Let $f : \mathbb{R}^n \times U \to \mathbb{R}^n$ be a continuous function, Lipschitz in x. For each control $\mathbf{u}$ we can consider the controlled dynamics

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}). \tag{3.2}$$

It is well known from ODE theory that, at least locally in time, equation (3.2) admits a unique solution $\mathbf{x}$ of the initial value problem $\mathbf{x}(t) = x$, for any (bounded) control $\mathbf{u}$.

We are given a running cost $L : \mathbb{R}^n \times U \to \mathbb{R}$ and a terminal cost $\psi : \mathbb{R}^n \to \mathbb{R}$. Given a terminal time T, the terminal value optimal control problem consists in determining the optimal trajectories $\mathbf{x}(\cdot)$ which minimize

$$J[\mathbf{u}; x, t] = \int_t^T L(\mathbf{x}, \mathbf{u}) ds + \psi(\mathbf{x}(t_1)),$$

among all bounded controls $\mathbf{u}(\cdot) : [t, t_1] \to \mathbb{R}^n$ and all solutions $\mathbf{x}$ of (3.2) satisfying the initial condition $\mathbf{x}(t) = x$.

The value function V is

$$V(x,t) = \inf J[\mathbf{u}; x, t] \tag{3.3}$$

in which the infimum is taken over all controls on $[t,T]$.

An important example is the "calculus of variations setting", where, $f(x,u) = u$, and the optimal trajectories $\mathbf{x}(\cdot)$, as we have shown, are solutions to the Euler-Lagrange equation

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

Furthermore, $\mathbf{p} = -D_v L(\mathbf{x}, \dot{\mathbf{x}})$ is a solution of Hamilton's equations:

$$\dot{\mathbf{x}} = -D_p H(\mathbf{p}, \mathbf{x}), \quad \dot{\mathbf{p}} = D_x H(\mathbf{p}, \mathbf{x}).$$

In the next chapter we will consider this problem under the light of optimal control and generalize the previous results.

Before considering the "calculus of variations setting" we study a simpler but important situation, the bounded control case. In this the control space U is a compact convex set.

Furthermore, we suppose that $L(x,u)$ is a bounded continuous function, convex in u. We suppose that the function $f(x,u)$ satisfies the following Lipschitz condition

$$|f(x,u) - f(y,u)| \leq C|x-y|.$$

To establish existence of optimal solutions we simplify even further by assuming that $f(x,u)$ has the form

$$f(x,u) = A(x)u + B(x), \tag{3.4}$$

where A and B are Lipschitz continuous functions.

3.2 Elementary properties

In this section we establish some elementary properties of the terminal value problem.

Proposition 15. *The value function V satisfies the following inequalities*

$$-\|\psi\|_\infty \leq V \leq c_1|T-t| + \|\psi\|_\infty.$$

Proof. The first inequality follows from $L \geq 0$. To obtain the second inequality it is enough to observe that

$$V \leq J(x,t;0) \leq c_1|T-t| + \|\psi\|_\infty.$$

□

Example 13 (Lax-Hopf formula)**.** Suppose that $L(x,v) \equiv L(v)$, L convex in v and coercive. Assume further that $f(x,v) = v$. By Jensen's inequality

$$\frac{1}{T-t}\int_t^T L(\dot{\mathbf{x}}(s)) \geq L\left(\frac{1}{T-t}\int_t^T \dot{\mathbf{x}}(s)\right) = L\left(\frac{y-x}{T-t}\right),$$

where $y = \mathbf{x}(T)$. Therefore, to solve the terminal value optimal control problem, it is enough to consider constant controls of the form $\mathbf{u}(s) = \frac{y-x}{T-t}$. Thus

$$V(x,t) = \inf_{y\in\mathbb{R}^n}\left[(T-t)L\left(\frac{y-x}{T-t}\right) + \psi(y)\right],$$

and, consequently, the infimum is a minimum. Thus Lax-Hopf formula gives an explicit solution to the optimal control problem. ◀

Exercise 34. *Let Q and A be $n\times n$ constant, positive definite, matrices. Let $L(v) = \frac{1}{2}v^TQv$ and $\psi(y) = \frac{1}{2}y^TAy$. Use Lax-Hopf formula to determine $V(x,t)$.*

Proposition 16. *Let $\psi_1(x)$ and $\psi_2(x)$ be continuous functions such that*

$$\psi_1 \leq \psi_2.$$

Let $V_1(x,t)$ and $V_2(x,t)$ be the corresponding value functions. Then

$$V_1(x,t) \leq V_2(x,t).$$

Proof. Fix $\epsilon > 0$. Then there exists an almost optimal control $\mathbf{u}^\epsilon$ and corresponding trajectory $\mathbf{x}^\epsilon$ such that

$$V_2(x,t) > \int_t^{t_1} L(\mathbf{x}^\epsilon(s), \mathbf{u}^\epsilon(s), s)ds + \psi_2(\mathbf{x}^\epsilon(t_1)) - \epsilon.$$

Clearly

$$V_1(x,t) \leq \int_t^{t_1} L(\mathbf{x}^\epsilon(s), \mathbf{u}^\epsilon(s), s)ds + \psi_1(\mathbf{x}^\epsilon(t_1)),$$

and therefore

$$V_1(x,t) - V_2(x,t) \leq \psi_1(\mathbf{x}^\epsilon(t_1)) - \psi_2(\mathbf{x}^\epsilon(t_1)) + \epsilon \leq \epsilon.$$

Since ϵ is arbitrary, this ends the proof. □

An important corollary is the continuity of the value function on the terminal value, with respect to the L^∞ norm.

Corollary 17. *Let $\psi_1(x)$ and $\psi_2(x)$ be continuous functions and $V_1(x,t)$ and $V_2(x,t)$ the corresponding value functions. Then*

$$\sup_x |V_1(x,t) - V_2(x,t)| \leq \sup_x |\psi_1(x) - \psi_2(x)|.$$

Proof. Note that

$$\psi_1 \leq \tilde{\psi}_2 \equiv \psi_2 + \sup_y |\psi_1(y) - \psi_2(y)|.$$

Let $\tilde{V}_2$ be the value function corresponding to $\tilde{\psi}_2$. Clearly,

$$\tilde{V}_2 = V_2 + \sup_y |\psi_1(y) - \psi_2(y)|.$$

By the previous proposition,

$$V_1 - \tilde{V}_2 \leq 0,$$

which implies

$$V_1 - V_2 \leq \sup_y |\psi_1(y) - \psi_2(y)|.$$

By reverting the roles of V_1 and V_2 we obtain the other inequality. ☐

3.3 Dynamic programming principle

The *dynamic programming principle*, that we prove in the next theorem, is simply a semigroup property that the evolution of the value function satisfies.

Theorem 18 (Dynamic programming principle)**.** *Suppose that* $t \leq t' \leq T$*. Then*

$$V(x,t) = \inf_{\mathbf{u}} \left[\int_t^{t'} L(\mathbf{x}(s), \mathbf{u}(s), s) ds + V(y, t') \right], \tag{3.5}$$

where $\mathbf{x}(t) = x$ *and* $\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u})$*.*

Proof. Denote by $\tilde{V}(x,t)$ the right hand side of (3.5). For fixed $\epsilon > 0$, let $\mathbf{u}^\epsilon$ be an almost optimal control for $V(x,t)$. Let $\mathbf{x}^\epsilon(s)$ be the corresponding trajectory trajectory, i.e., assume that

$$J(x, t; \mathbf{u}^\epsilon) \leq V(x,t) + \epsilon.$$

We claim that $\tilde{V}(x,t) \leq V(x,t)+\epsilon$. To check this, let $\mathbf{x}(\cdot)=\mathbf{x}^\epsilon(\cdot)$ and $y=\mathbf{x}^\epsilon(t')$. Then

$$\tilde{V}(x,t) \leq \int_t^{t'} L(\mathbf{x}^\epsilon(s),\mathbf{u}^\epsilon(s),s)ds + V(y,t').$$

Additionally,

$$V(y,t') \leq J(y,t';\mathbf{u}^\epsilon).$$

Therefore

$$\tilde{V}(x,t) \leq J(x,t;\mathbf{u}^\epsilon) \leq V(x,t)+\epsilon,$$

and, since ϵ is arbitrary, $\tilde{V}(x,t) \leq V(x,t)$.

To prove the opposite inequality, we will proceed by contradiction. Therefore, if $\tilde{V}(x,t) < V(x,t)$, we could choose $\epsilon > 0$ and a control $\mathbf{u}^\sharp$ such that

$$\int_t^{t'} L(\mathbf{x}^\sharp(s),\mathbf{u}^\sharp(s),s)ds + V(y,t') < V(x,t)-\epsilon,$$

where $\dot{\mathbf{x}}^\sharp = f(\mathbf{x}^\sharp,\mathbf{u}^\sharp)$, $\mathbf{x}^\sharp(t)=x$, and $y=\mathbf{x}^\sharp(t')$. Choose $\mathbf{u}^\flat$ such that

$$J(y,t';\mathbf{u}^\flat) \leq V(y,t')+\frac{\epsilon}{2}$$

Define $\mathbf{u}^\star$ as

$$\begin{cases} \mathbf{u}^\star(s)=\mathbf{u}^\sharp(s) \text{ for } s<t' \\ \mathbf{u}^\star(s)=\mathbf{u}^\flat(s) \text{ for } t'<s. \end{cases}$$

So, we would have

$$\begin{aligned} V(x,t)-\epsilon &> \int_t^{t'} L(\mathbf{x}^\sharp(s),\mathbf{u}^\sharp(s),s)ds + V(y,t') \geq \\ &\geq \int_t^{t'} L(\mathbf{x}^\sharp(s),\mathbf{u}^\sharp(s),s)ds + J(y,t';\mathbf{u}^\flat) - \frac{\epsilon}{2} = \\ &= J(x,t;\mathbf{u}^\star)-\frac{\epsilon}{2} \geq V(x,t)-\frac{\epsilon}{2}, \end{aligned}$$

which is a contradiction. □

3.4 Pontryagin maximum principle

In this section we assume the control space U is bounded and, without any proof, that there exists an optimal control $\mathbf{u}^*$ and corresponding optimal trajectory $\mathbf{x}^*$. We assume also that the terminal data ψ is differentiable. A detailed discussion on existence issues will be postponed until next chapter.

Let $r \in [t, t_1)$ be a point where $\mathbf{u}^*$ is strongly approximately continuous, i.e.,

$$\varphi(\mathbf{u}^*(r)) = \lim_{\delta \to 0} \frac{1}{\delta} \int_r^{r+\delta} \varphi(\mathbf{u}^*(s)) ds,$$

for all continuous functions φ. Note that almost any r is a point of approximate continuity, see [EG92]. Denote by Ξ_0 the fundamental solution of

$$\dot{\xi}_0 = D_x f(\mathbf{x}^*, \mathbf{u}^*) \xi_0, \tag{3.6}$$

with $\Xi_0(r) = I$.

Let $\mathbf{p}^*$ be given by

$$\mathbf{p}^*(r) = D_x \psi(\mathbf{x}_R(t_1)) \Xi_0(t_1) + \int_r^{t_1} D_x L(\mathbf{x}^*(s), \mathbf{u}^*(s), s) \Xi_0(s) ds. \tag{3.7}$$

Lemma 19 (Pontryagin maximum principle)**.** *Suppose that ψ is differentiable. Let $\mathbf{u}^*$ be an optimal control and $\mathbf{x}^*$ the corresponding optimal trajectory. Then, for almost all $r \in [t, t_1)$,*

$$\begin{aligned} &f(\mathbf{x}^*(r), \mathbf{u}^*(r)) \cdot \mathbf{p}^*(r) + L(\mathbf{x}^*(r), \mathbf{u}^*(r), r) \\ &\quad = \min_{v \in U} \left[f(\mathbf{x}^*, v) \cdot \mathbf{p}^*(r) + L(\mathbf{x}^*(r), v, r) \right]. \end{aligned} \tag{3.8}$$

Proof. Let $v \in U$. For almost all $r \in [t_0, t_1)$ $\mathbf{u}^*$ is strongly approximately continuous (see [EG92]). Let r be one of these points. Define

$$\mathbf{u}_\delta(s) = \begin{cases} v & \text{if} \quad r < s < r + \delta \\ \mathbf{u}^*(s) & \text{otherwise.} \end{cases}$$

Let

$$\mathbf{x}_\delta(s) = \begin{cases} \mathbf{x}^*(s) & \text{if} \quad t < s < r \\ \mathbf{x}^*(r) + \int_r^s f(\mathbf{x}_\delta^*, v) & \text{if} \quad r < s < r + \delta \\ \mathbf{x}^*(s) + \delta\xi_\delta & \text{if} \quad r + \delta < s < t_1, \end{cases}$$

where

$$\xi_\delta(r + \delta) = \frac{1}{\delta} \int_r^{r+\delta} \left[f(\mathbf{x}_\delta^*(s), v) - f(\mathbf{x}^*(s), \mathbf{u}^*(s)) \right] ds,$$

and $\mathbf{y}_\delta \equiv bx^*(s) + \delta\xi_\delta$ solves, for $r + \delta < s < t_1$,

$$\dot{\mathbf{y}}_\delta = f(\mathbf{y}_\delta, \mathbf{u}^*).$$

Observe that

$$\xi_0(r) = \lim_{\delta \to 0} \xi_\delta(r + \delta) = f(\mathbf{x}^*(r), v) - f(\mathbf{x}^*(r), \mathbf{u}^*(r)).$$

Furthermore, as $\delta \to 0$, ξ_δ converges to a solution ξ_0 of (3.6). Thus $\xi_0(s) = \Xi_0(s) \left(f(\mathbf{x}^*(r), v) - f(\mathbf{x}^*(r), \mathbf{u}^*(r)) \right)$.

Clearly

$$J(t, x; \mathbf{u}^*) \leq \int_t^{t_1} L(\mathbf{x}_\delta(s), \mathbf{u}_\delta(s), s) ds + \psi(\mathbf{x}^*(t_1) + \delta\xi_\delta).$$

This last inequality implies

$$\begin{aligned} &\frac{1}{\delta} \int_r^{r+\delta} \left[L(\mathbf{x}_\delta(s), v, s) - L(\mathbf{x}^*(s), \mathbf{u}^*(s), s) \right] ds + \\ &+ \frac{1}{\delta} \int_{r+\delta}^{t_1} \left[L(\mathbf{x}^*(s) + \delta\xi_\delta, \mathbf{u}^*(s), s) - L(\mathbf{x}^*(s), \mathbf{u}^*(s), s) \right] ds + \\ &+ \frac{1}{\delta} \left[\psi(\mathbf{x}^*(t_1) + \delta\xi_\delta) - \psi(\mathbf{x}^*(t_1)) \right] \geq 0. \end{aligned}$$

When $\delta \to 0$, the first term converges to

$$L(\mathbf{x}^*(r), v, r) - L(\mathbf{x}^*(r), \mathbf{u}^*(r), r),$$

since $\mathbf{u}^*$ is strongly approximately continuous. The second term converges to

$$\int_r^{t_1} D_x L(\mathbf{x}^*(s), \mathbf{u}^*(s), s)\xi_0(s)ds,$$

whereas the third one has the following limit:

$$D_x\psi(\mathbf{x}_R(t_1)) \cdot \xi_0(t_1)).$$

This implies that for almost all r r,

$$L(\mathbf{x}^*(r), v, r) - L(\mathbf{x}^*(r), \mathbf{u}^*(r), r) + \mathbf{p}^*(r) \cdot (f(\mathbf{x}^*(r), v) - f(\mathbf{x}^*(r), \mathbf{u}^*(r))) \geq 0.$$

Consequently

$$f(\mathbf{x}^*(r), \mathbf{u}^*(r)) \cdot \mathbf{p}^*(r) + L(\mathbf{x}^*(r), \mathbf{u}^*(r), r) = \min_{v \in U} \left[f(\mathbf{x}^*(r), v) \cdot \mathbf{p}^*(r) + L(\mathbf{x}_R(r), v, r) \right],$$

as required. □

3.5 The Hamilton-Jacobi equation

We now show that if the value function is differentiable then it is a solution to the Hamilton-Jacobi equation.

Proposition 20. *Suppose the value function is C^1. Let $r \in [t, t_1)$ be a point where $\mathbf{u}^*$ is strongly approximately continuous. Then*

$$\mathbf{p}^*(r) = D_x V(x, r).$$

Proof. Let $\mathbf{u}^*$ be an optimal control for the initial condition (x, r). For $y \in \mathbb{R}^n$ and $\delta > 0$ consider the solution

$$\dot{\mathbf{x}}_\delta = f(\mathbf{x}_\delta, \mathbf{u}^*),$$

with initial condition $\mathbf{x}_\delta(t) = x + \delta y$. Then

$$\left.\frac{\partial \mathbf{x}_\delta(s)}{\partial \delta}\right|_{\delta=0} = \Xi_0(s)y.$$

Since for all δ

$$V(x + \delta y, r) \leq \int_r^{t_1} L(\mathbf{x}_\delta, \mathbf{u}^*)ds + \psi(\mathbf{x}_\delta(t_1)),$$

by differentiating with respect to δ we obtain

$$D_x V(x, r)y = \int_r^{t_1} D_x L(\mathbf{x}, \mathbf{u}^*)\Xi_0(s)yds + D_x\psi(\mathbf{x}(t_1))\Xi_0(t_1)y,$$

which implies the result. □

Theorem 21. *Suppose the value function V is C^1. Then it solves*

$$-V_t + H(D_x V, x) = 0. \tag{3.9}$$

Proof. Consider an optimal trajectory $\mathbf{x}^*$

$$V(\mathbf{x}^*(t), t) = \int_t^{t_1} L(\mathbf{x}^*(s), \mathbf{u}^*(s))ds.$$

Then, by differentiating with respect to t we have

$$V_t(\mathbf{x}^*(t), t) + D_x V(\mathbf{x}^*(t), t) f(\mathbf{x}^*(t), \mathbf{u}^*(t)) + L(\mathbf{x}^*(t), \mathbf{u}^*(t)) = 0.$$

Which by Pontryagin maximum principle is equivalent to the Hamilton-Jacobi equation (3.9). □

Exercise 35. *Let $M(t), N(t)$ be $n \times n$ matrices with time-differentiable coefficients. Suppose that is N invertible. Let D be a $n \times n$ constant matrix. Consider the Lagrangian*

$$L(x, v) = \frac{1}{2} x^T M(t) x + \frac{1}{2} v^T N(t) v$$

and the terminal condition $\psi = \frac{1}{2} x^T D x$. Show that there exists a solution to the Hamilton-Jacobi with terminal condition ψ at $t = T$ of the form

$$V = \frac{1}{2} x^T P(t) x,$$

where $P(t)$ satisfies the Ricatti equation

$$\dot{P} = P^T N^{-1} P - M$$

and $P(T) = D$.

3.6 Verification theorem

In the last section of this chapter we will show that any sufficiently smooth solution to the Hamilton-Jacobi equation is the value function and it can be used to compute an optimal control.

Theorem 22. *Let $L(x, v)$ be a C^1 Lagrangian, strictly convex in v, and let $f(x, u)$ be a linear control law as in* (3.4)*, and H the generalized Legendre transform* (3.1) *of L. Let $\Phi(x, t)$ be a classical solution to the Hamilton-Jacobi equation*

$$-\Phi_t + H(D_x \Phi, x) = 0 \tag{3.10}$$

on the time interval $[0, T]$, with terminal data $\Phi(x, T) = \psi(x)$. Then, for all $0 \leq t \leq T$,

$$\Phi(x, t) = V(x, t),$$

where V is the value function.

Proof. Let $\mathbf{u}$ be a control on (t,T) and $\mathbf{x}$ be the corresponding solution to

$$\dot{\mathbf{x}} = f(\mathbf{x},\mathbf{u}),$$

with $\mathbf{x}(t) = x$. Then, using $\Phi(\mathbf{x}(T),T) = \psi(\mathbf{x}(T))$ we have

$$\begin{aligned}\psi(\mathbf{x}(T)) - \Phi(\mathbf{x}(t),t) &= \int_t^T \frac{d}{ds}\Phi(\mathbf{x}(s),s)ds\\ &= \int_t^T D_x\Phi(\mathbf{x}(s),s)\cdot f(\mathbf{x},\mathbf{u}) + \Phi_s(\mathbf{x}(s),s)ds.\end{aligned}$$

Adding $\int_t^T L(\mathbf{x}(s),\mathbf{u}(s))ds + \Phi(\mathbf{x}(t),t)$ to the above equality and taking the infimum over all controls $\mathbf{u}$, we obtain

$$\begin{aligned}&\inf\left(\int_t^T L(\mathbf{x}(s),\mathbf{u}(s))ds + \varphi\big(\mathbf{x}(T)\big)\right)\\ &\quad= \Phi(\mathbf{x}(t),t)\\ &\quad+ \inf\left(\int_t^T \Phi_s(\mathbf{x}(s),s) + L(\mathbf{x}(s),\mathbf{u}(s)) + D_x\Phi(\mathbf{x}(s),s)\cdot f(\mathbf{x},\mathbf{u})ds\right).\end{aligned}$$

Now recall that for any v,

$$-H(p,x) \le L(x,v) + p\cdot f(x,v),$$

therefore

$$\begin{aligned}&\inf\left(\int_t^T L(\mathbf{x}(s),\dot{\mathbf{x}}(s))ds + \varphi\big(\mathbf{x}(T)\big)\right)\\ &\quad\ge \Phi(\mathbf{x}(t),t) + \inf\left(\int_t^T \big(\Phi_s(\mathbf{x}(s),s) + H(D_x\Phi(\mathbf{x}(s),s),\mathbf{x}(s))\big)ds\right)\\ &\quad= \Phi(\mathbf{x}(t),t).\end{aligned}$$

Let $r(x,t)$ be uniquely defined (uniqueness follows from convexity) as

$$r(x,t) \in \operatorname{argmin}_{v\in U} L(x,v) + D_x\Phi(x,t)\cdot f(x,v). \tag{3.11}$$

A simple argument shows that r is a continuous function.

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

$$\dot{\mathbf{x}}(s) = f(\mathbf{x}, r(\mathbf{x}(s), s)),$$

with initial condition $\mathbf{x}(t) = x$. Note that since the right-hand side is continuous there is a solution, although it may not be unique. Then

$$\begin{aligned}
&\inf \left(\int_t^T L(\mathbf{x}(s), \dot{\mathbf{x}}(s)) ds + \varphi\big(\mathbf{x}(T)\big) \right) \\
&\quad \le \Phi(\mathbf{x}(t), t) + \int_t^T \Big(\Phi_s\big(\mathbf{x}(s), s\big) - H\big(D_x\Phi(\mathbf{x}(s), s), \mathbf{x}(s)\big) \Big) ds \\
&\quad = \Phi(\mathbf{x}(t), t),
\end{aligned}$$

which ends the proof. □

We should observe from the proof that (3.11) gives an optimal feedback law for the optimal control, provided we can find a solution to the Hamilton-Jacobi equation.

3.7 Bibliographical notes

The main references we have used on optimal control are [BCD97], [FS06], [Lio82], [Bar94], and [Eva98b].

4

Viscosity solutions

In this chapter we build upon the theory developed previously to study the terminal value problem and address a few questions that were not answered previously. The first one is the existence of optimal controls, both for bounded and unbounded control spaces. This is addressed, for the bounded control setting in section 4.1. In section 4.2 we give some technical results concerning subdifferentials and semiconcavity. Then, in section 4.3 we consider the issue of existence of controls and regularity of the value function in the calculus of variations setting. It is well known that first order partial differential equations such as the Hamilton-Jacobi equation may not admit classical solutions. Using the method of characteristics, the next exercise gives an example of non-existence of smooth solutions:

Exercise 36. *Solve, using the method of characteristics, the equation*

$$\begin{cases} u_t + u_x^2 = 0 & x \in \mathbb{R}, t > 0 \\ u(x,0) = \pm x^2. \end{cases}$$

It is therefore necessary to consider weak solutions to the Hamilton-Jacobi equation: viscosity solutions. In section §4.4 we develop the theory of viscosity solutions for Hamilton-Jacobi equations, and show that the value function is the unique viscosity solution of the Hamilton-Jacobi equation.

4.1 Optimal controls - bounded control space

We now give a proof of the existence of optimal controls for bounded control space. The unbounded case will be addressed in §4.3.

Lemma 23. *Letf is as in* (3.4) *a linear control law. Then J is weakly lower semicontinuous, with respect to weak-* convergence in L^∞.*

Proof. Let u_n be a sequence of controls such that $\mathbf{u}_n \overset{*}{\rightharpoonup} \mathbf{u}$ in $L^\infty[t, t_1]$. Then, by using Ascoli-Arzela theorem, we can extract a subsequence of $\mathbf{x}_n(\cdot)$ converging uniformly to $\mathbf{x}(\cdot)$. Furthermore, because the control law (3.4) is linear we have

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}).$$

We have

$$J(x, t; \mathbf{u}_n) = \int_t^{t_1} \left[L(\mathbf{x}_n(s), \mathbf{u}_n(s), s) - L(\mathbf{x}(s), \mathbf{u}_n(s), s)\right] ds + \\ + \int_t^{t_1} L(\mathbf{x}(s), \mathbf{u}_n(s), s) ds + \psi(\mathbf{x}_n(t_1)).$$

The first term, $\int_t^{t_1} \left[L(\mathbf{x}_n(s), \mathbf{u}_n(s), s) - L(\mathbf{x}(s), \mathbf{u}_n(s), s)\right] ds$, converges to zero. Similarly, $\psi(\mathbf{x}_n(t_1)) \to \psi(\mathbf{x}(t_1))$. Finally, the convexity of L implies

$$L(\mathbf{x}(s), \mathbf{u}_n(s), s) \geq L(\mathbf{x}(s), \mathbf{u}(s), s) + D_v L(\mathbf{x}(s), \mathbf{u}(s), s)(\mathbf{u}_n(s) - \mathbf{u}(s)).$$

Since $\mathbf{u}_n \rightharpoonup \mathbf{u}$,

$$\int_t^{t_1} D_v L(\mathbf{x}(s), \mathbf{u}(s), s)(\mathbf{u}_n(s) - \mathbf{u}(s)) ds \to 0.$$

Hence

$$\liminf J(x, t; \mathbf{u}_n) \geq J(x, t; \mathbf{u}),$$

that is, J is weakly lower semicontinuous. □

Using the previous result we can now state and prove our first existence result.

Lemma 24. *Suppose the control set U is bounded, closed and convex. There exists a minimizer $\mathbf{u}^*$ of J.*

Proof. Let $\mathbf{u}_n$ be a minimizing sequence, that is, such that

$$J(x, t; \mathbf{u}_n) \to \inf_{\mathbf{u} \in \mathcal{U}_R} J(x, t; \mathbf{u}).$$

Because this sequence is bounded in L^∞, by Banach-Alaoglu theorem we can extract a sequence $\mathbf{u}_n \overset{*}{\rightharpoonup} \mathbf{u}^*$. Clearly, we have $\mathbf{u}^* \in U$, by closeness and convexity. We claim now that

$$J(x, t; \mathbf{u}^*) = \inf_{\mathbf{u}} J(x, t; \mathbf{u}).$$

This just follows from the weak lower semicontinuity:

$$\inf_{\mathbf{u}} J(x, t; \mathbf{u}) \leq J(x, t; \mathbf{u}^*) \leq \liminf J(x, t; \mathbf{u}_n) = \inf_{\mathbf{u}} J(x, t; \mathbf{u}),$$

which ends the proof. □

Example 14 (Bang-Bang principle)**.** Consider the case of a bounded closed convex control space U and suppose the Lagrangian L is constant. Suppose $f(x,u) = Au + B$, for suitable constant matrices A and B, and that the terminal value ψ is convex.

In this setting we first observe that the set of all optimal controls is convex. As such it admits an extreme point $\mathbf{u}^*$. We claim that $\mathbf{u}^*$ takes values on ∂U.

To see this, choose a time r and suppose that for some ϵ there is a set of positive measure in $[r, r+\epsilon]$ for which $\mathbf{u}^*$ is in the interior of U. Then there exists an L^∞ function ν supported on this set such that $\int_r^{r+\epsilon} d\nu = 0$, and such that $\mathbf{u}^* \pm \nu$ is an admissible control. By our assumptions it is also an optimal control. It is clear then that $\mathbf{u}^*$ is not an extreme point which is a contradiction. ◀

4.2 Sub and superdifferentials

Before proceeding with the general case of unbounded control spaces we will need to discuss some technical results concerning sub-differentials and semiconcavity.

Let $\psi : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. The *superdifferential* $D_x^+\psi(x)$ of ψ at x is the set of vectors $p \in \mathbb{R}^n$ such that

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

Consequently, $p \in D_x^+\psi(x)$ if and only if

$$\psi(x+v) \leq \psi(x) + p \cdot v + o(|v|),$$

as $|v| \to 0$. Similarly, the *subdifferential*, $D_x^- \psi(x)$, of ψ at x is the set of vectors p such that

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

Exercise 37. *Show that if $u : \mathbb{R}^n \to \mathbb{R}$ has a maximum (resp. minimum) at x_0 then $0 \in D^+u(x_0)$ (resp. $D^-u(x_0)$).*

We can regard these sets as one-sided derivatives. In fact, ψ is differentiable then

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

More precisely,

Proposition 25. *If both $D_x^- \psi(x)$ and $D_x^+ \psi(x)$ are non-empty then*

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

furthermore ψ is differentiable at x with $D_x \psi = p$. Conversely, if ψ is differentiable at x then

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

Proof. Suppose that $D_x^- \psi(x)$ and $D_x^+ \psi(x)$ are both non-empty. Then we claim that these two sets agree and have a single point p. To see this, take $p^- \in D_x^- \psi(x)$ and $p^+ \in D_x^+ \psi(x)$. Then

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

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

By subtracting these two identities

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

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

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

which implies $p^- = p^+ \equiv p$. Consequently

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

and so $D_x\psi = p$.

To prove the converse it suffices to observe that if ψ is differentiable then

$$\psi(x+v) = \psi(x) + D_x\psi(x)\cdot v + o(|v|).$$

□

Exercise 38. *Let ψ be a continuous function. Show that if x_0 is a local maximum of ψ then $0 \in D^+\psi(x_0)$.*

Proposition 26. *Let*

$$\psi : \mathbb{R}^n \to \mathbb{R}$$

be a continuous function. Then, if

$$p \in D_x^+\psi(x_0) \quad (resp.\ p \in D_x^-\psi(x_0)),$$

there exists a C^1 function ϕ such that

$$\psi(x) - \phi(x)$$

has a local strict maximum (resp. minimum) at x_0 and such that

$$p = D_x\phi(x_0).$$

On the other hand, if ϕ is a C^1 function such that

$$\psi(x) - \phi(x)$$

has a local maximum (resp. minimum) at x_0 then

$$D_x\phi(x_0) \in D_x^+\psi(x_0) \quad (resp.\ D_x^-\psi(x_0)).$$

Proof. By subtracting $p \cdot (x - x_0) + \psi(x_0)$ to ψ, we can assume, without loss of generality, that $\psi(x_0) = 0$ and $p = 0$. By changing coordinates, if necessary, we can also assume that $x_0 = 0$. Because $0 \in D_x^+\psi(0)$ we have

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

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

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

More precisely we can choose

$$\rho(x) = \max\{\frac{\psi}{|x|}, 0\}.$$

Let $\eta(r) = \max_{|x|\leq r}\{\rho(x)\}$. Then η is continuous, non decreasing and $\eta(0) = 0$. Let

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

The function ϕ is C^1 and satisfies $\phi(0) = D_x\phi(0) = 0$. Additionally, if $x \neq 0$,

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

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

To prove the second part of the proposition, suppose that the difference $\psi(x) - \phi(x)$ has a strict local maximum at 0. Without loss of generality, we can assume $\psi(0) - \phi(0) = 0$ and $\phi(0) = 0$. Then $\psi(x) - \phi(x) \leq 0$ or, equivalently,

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

Thus, by setting $p = D_x\phi(0)$, and using the fact that

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

we conclude that $D_x\phi(0) \in D_x^+\psi(0)$. The case of a minimum is similar. □

A continuous function f is *semiconcave* if there exists C such that

$$f(x+y) + f(x-y) - 2f(x) \leq C|y|^2.$$

Similarly, a function f is *semiconvex* if there exists a constant such that

$$f(x+y) + f(x-y) - 2f(x) \geq -C|y|^2.$$

Proposition 27. *The following statements are equivalent:*

1. *f is semiconcave;*
2. *$\tilde{f}(x) = f(x) - \frac{C}{2}|x|^2$ is concave;*
3. *for all λ, $0 \leq \lambda \leq 1$, and any y, z such that $\lambda y + (1-\lambda)z = 0$ we have*

$$\lambda f(x+y) + (1-\lambda)f(x+z) - f(x) \leq \frac{C}{2}(\lambda|y|^2 + (1-\lambda)|z|^2).$$

Additionally, if f is semiconcave, then

a. *$D_x^+ f(x) \neq \emptyset$;*

b. *if $D_x^- f(x) \neq \emptyset$ then f is differentiable at x;*

c. *there exists C such that, for each $p_i \in D_x^+ f(x_i)$ $(i = 0, 1)$,*

$$(x_0 - x_1)\cdot(p_0 - p_1) \leq C|x_0 - x_1|^2.$$

Remark. Of course analogous results hold for semiconvex functions.

Proof. Clearly $2 \implies 3 \implies 1$. Therefore, to prove the equivalence, it is enough to show that $1 \implies 2$. Subtracting $C|x|^2$ to f, we may assume $C = 0$. Also, by changing coordinates if necessary, it suffices to prove that for all x, y such that $\lambda x + (1-\lambda)y = 0$, for some $\lambda \in [0,1]$, we have:

$$\lambda f(x) + (1-\lambda)f(y) - f(0) \leq 0. \tag{4.1}$$

We claim now that the previous equation holds for each $\lambda = \frac{k}{2^j}$, with $0 \leq k \leq 2^j$. Clearly (4.1) holds for $j = 1$. We will proceed by induction on j. Suppose that (4.1) if valid for $\lambda = \frac{k}{2^j}$. We will show that it also holds for $\lambda = \frac{k}{2^{j+1}}$. If k is even, we can reduce the fraction and, therefore, we assume that k is odd, $\lambda = \frac{k}{2^{j+1}}$ and $\lambda x + (1-\lambda)y = 0$. Note that

$$0 = \frac{1}{2}\left[\frac{k-1}{2^{j+1}}x + \left(1 - \frac{k-1}{2^{j+1}}\right)y\right] + \frac{1}{2}\left[\frac{k+1}{2^{j+1}}x + \left(1 - \frac{k+1}{2^{j+1}}y\right)\right].$$

consequently,

$$\begin{aligned} f(0) \geq &\frac{1}{2}f\left(\frac{k-1}{2^{j+1}}x + \left(1 - \frac{k-1}{2^{j+1}}\right)y\right) + \\ &+ \frac{1}{2}f\left(\frac{k+1}{2^{j+1}}x + \left(1 - \frac{k+1}{2^{j+1}}\right)y\right) \end{aligned}$$

but, since $k-1$ and $k+1$ are even, $\tilde{k}_0 = \frac{k-1}{2}$ and $\tilde{k}_1 = \frac{k+1}{2}$ are integers. Therefore

$$f(0) \geq \frac{1}{2}f\left(\frac{\tilde{k}_0}{2^j}x + \left(1 - \frac{\tilde{k}_0}{2^j}\right)y\right) + \frac{1}{2}f\left(\frac{\tilde{k}_1}{2^j}x + \left(1 - \frac{\tilde{k}_1}{2^j}\right)y\right)$$

But this implies

$$f(0) \geq \frac{\tilde{k}_0 + \tilde{k}_1}{2^{j+1}}f(x) + \left(1 - \frac{\tilde{k}_0 + \tilde{k}_1}{2^{j+1}}\right)f(y).$$

From $\tilde{k}_0 + \tilde{k}_1 = k$ we obtain

$$f(0) \geq \frac{k}{2^{j+1}} f(x) + \left(1 - \frac{k}{2^{j+1}}\right) f(y).$$

Since the function f is continuous and the rationals of the form $\frac{k}{2^j}$ are dense in $\mathbb{R}$, we conclude that

$$f(0) \geq \lambda f(x) + (1-\lambda) f(y),$$

for each real λ, with $0 \leq \lambda \leq 1$.

To prove the second part of the proposition, observe that by proposition 25, $a \implies b$. To check a, i.e., that $D_x^+ f(x) \neq \emptyset$, it is enough to observe that if f is concave then $D_x^+ f(x) \neq \emptyset$. By subtracting $C|x|^2$ to f, we can reduce the problem to concave functions. Finally, if $p_i \in D_x^+ f(x_i)$ $(i = 0, 1)$ then

$$f(x_0) - \frac{C}{2}|x_0|^2 \leq f(x_1) - \frac{C}{2}|x_1|^2 + (p_1 - Cx_1) \cdot (x_0 - x_1),$$

and

$$f(x_1) - \frac{C}{2}|x_1|^2 \leq f(x_0) - \frac{C}{2}|x_0|^2 + (p_0 - Cx_0) \cdot (x_1 - x_0).$$

Therefore,

$$0 \leq (p_1 - p_0) \cdot (x_0 - x_1) + C|x_0 - x_1|^2,$$

and so $(p_0 - p_1) \cdot (x_0 - x_1) \leq C|x_0 - x_1|^2$. □

Exercise 39. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function. Show that if x_0 is a local maximum then $0 \in D^+ f(x_0)$.*

4.3 Calculus of variations setting

We now consider the calculus of variations setting and prove the existence of optimal controls. The main technical issue is the fact that the control space $U = \mathbb{R}^n$ is unbounded and therefore compactness arguments do not work directly. We will consider the calculus of variations setting, that is $f(x, u) = u$ and we will work under the following assumptions:

$$L(x, v) : \mathbb{R}^{2n} \to \mathbb{R},$$

$x \in \mathbb{R}^n$, $v \in \mathbb{R}^n$, is a C^∞ function, strictly convex em v, i.e., $D^2_{vv}L$ is positive definite, and satisfying the coercivity condition

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

for each (x, t); without loss of generality, we may also assume that $L(x, v, t) \geq 0$, by adding a constant if necessary. We will also assume that

$$L(x, 0, t) \leq c_1, \quad |D_x L| \leq c_2 L + c_3,$$

for suitable constants c_1, c_2 and c_3; finally we assume that there exists a function $C(R)$ such that

$$|D^2_{xx}L| \leq C(R), \quad |D_v L| \leq C(R)$$

whenever $|v| \leq R$. The terminal cost, ψ, is assumed to be a bounded Lipschitz function.

Example 15. Although the conditions on L are quite technical, they are fulfilled by a wide class of Lagrangians, for instance

$$L(x, v) = \frac{1}{2} v^T A(x) v - V(x),$$

where A and V are C^∞,bounded with bounded derivatives, and $A(x)$ is positive definite. ◀

Fortunately, the coercivity of the Lagrangian is enough to establish the existence of a-priori bounds on optimal controls.

Theorem 28. *Let $x \in \mathbb{R}^n$ and $t_0 \leq t \leq t_1$. Suppose that the Lagrangian $L(x, v)$ satisfies:*

A. L is C^∞, strictly convex in v (i.e., $D^2_{vv}L$ is positive definite), and satisfying the coercivity condition

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

uniformly in (x, t);

B. L is bounded by bellow (without loss of generality we assume $L(x, v) \geq 0$);

C. L satisfies the inequalities

$$L(x, 0) \leq c_1, \qquad |D_x L| \leq c_2 L + c_3$$

for suitable c_1, c_2, and c_3;

D. there exist functions $C_0(R), C_1(R) : \mathbb{R}^+ \to \mathbb{R}^+$ such that

$$|D_v L| \leq C_0(R), \qquad |D^2_{xx} L| \leq C_1(R)$$

whenever $|v| \leq R$.

Then, if ψ is a bounded Lipschitz function,

1. There exists $\mathbf{u}^* \in L^\infty[t, t_1]$ such that its corresponding trajectory $\mathbf{x}^*$, given by

$$\dot{\mathbf{x}}^*(s) = \mathbf{u}(s) \qquad \mathbf{x}^*(t) = x,$$

is optimal, that is it satisfies

$$V(x, t) = \int_t^{t_1} L(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s)) ds + \psi(\mathbf{x}^*(t_1)).$$

2. *There exists C, depending only on L, ψ and $t_1 - t$ but not on x or t such that $|\mathbf{u}(s)| < C$ for $t \leq s \leq t_1$. The optimal trajectory $\mathbf{x}^*(\cdot)$ is a $C^2[t, t_1]$ solution of the Euler-Lagrange equation*

$$\frac{d}{dt} D_v L - D_x L = 0 \tag{4.2}$$

with initial condition $\mathbf{x}^*(t) = x$.

3. *The adjoint variable $\mathbf{p}$, defined by*

$$\mathbf{p}(t) = -D_v L(\mathbf{x}^*, \dot{\mathbf{x}}^*), \tag{4.3}$$

satisfies the differential equation

$$\begin{cases} \dot{\mathbf{p}}(s) = D_x H(\mathbf{p}(s), \mathbf{x}^*(s)) \\ \dot{\mathbf{x}}^*(s) = -D_p H(\mathbf{p}(s), \mathbf{x}^*(s)) \end{cases}$$

with terminal condition

$$\mathbf{p}(t_1) \in D_x^- \psi(\mathbf{x}^*(t_1)).$$

Additionally,

$$(\mathbf{p}(s), H(\mathbf{p}(s), \mathbf{x}^*(s))) \in D^- V(\mathbf{x}^*(s), s)$$

for $t < s \leq t_1$.

4. *The value function V is Lipschitz, and so almost everywhere differentiable.*

5. *If $D^2_{vv} L$ is uniformly bounded, then for each $t < t_1$, $V(x, t)$ is semiconcave in x.*

6. *For $t \leq s < t_1$*

$$(\mathbf{p}(s), H(\mathbf{p}(s), \mathbf{x}^*(s))) \in D^+ V(\mathbf{x}^*(s), s)$$

and, therefore, $DV(\mathbf{x}^*(s), s)$ exists for $t < s < t_1$.

Proof. We will divide the proof into several auxiliary lemmas.

For $R > 0$, define $\mathcal{U}_R = \{\mathbf{u} \in \mathcal{U} : \|\mathbf{u}\|_\infty \leq R\}$. From lemma 24 there exists a minimizer $\mathbf{u}_R$ of J in $\mathcal{U}_R$. Then we will show that the minimizer $\mathbf{u}_R$ satisfies uniform estimates in R. Finally, we will let $R \to \infty$.

Let $\mathbf{p}_R$ be the adjoint variable given by the Pontryagin maximum principle. We now will try to estimate the optimal control $\mathbf{u}_R$ uniformly in R, in order to send $R \to \infty$.

Lemma 29. *Suppose ψ is differentiable. Then there exists a constant C, independent on R, such that*

$$|\mathbf{p}_R| \leq C.$$

Proof. Since ψ is Lipschitz and differentiable we have

$$|D_x\psi| \leq \|D_x\psi\|_\infty < \infty.$$

Therefore

$$|\mathbf{p}_R(s)| \leq \int_s^{t_1} |D_x L(\mathbf{x}_R(r), \mathbf{u}_R(r)|dr + \|D_x\psi\|_\infty.$$

Let V_R be the value function for the terminal value problem with the additional constraint of bounded control: $|v| \leq R$. From $|D_x L| \leq c_2 L + c_3$, it follows

$$|\mathbf{p}_R(s)| \leq C(V_R(t, x) + 1),$$

for an appropriate constant C. Proposition 15, shows that there exists a constant C, which does not depend on R, such that $V_R \leq C$. Therefore

$$|\mathbf{p}_R| \leq C.$$

□

As we will see, the uniform estimates for $\mathbf{p}_R$ yield uniform estimates for $\mathbf{u}_R$.

Lemma 30. *Let ψ be differentiable. Then there exists $R_1 > 0$ such that, for all R,*

$$\|\mathbf{u}_R\|_\infty \leq R_1.$$

Proof. Suppose $|p| \leq C$. Then, for each c_1, the coercivity condition on L implies that there exists R_1 such that, if

$$v \cdot p + L(x, v) \leq c_1$$

then $|v| \leq R_1$. But then,

$$\mathbf{u}_R(s) \cdot \mathbf{p}_R(s) + L(\mathbf{x}_R(s), \mathbf{u}_R(s)) \leq L(\mathbf{x}_R(s), 0) \leq c_1,$$

that is, $\|\mathbf{u}_R\|_\infty \leq R_1$. □

Since $\mathbf{u}_R$ is bounded independently of R, we have

$$V = J(x, t; \mathbf{u}_{R_0}),$$

for $R_0 > R_1$. Let $\mathbf{u}^* = \mathbf{u}_{R_0}$ and $\mathbf{p} = \mathbf{p}_{R_0}$.

Lemma 31 (Pontryagin maximum principle - II)**.** *If ψ is differentiable, optimal control $\mathbf{u}^*$ satisfies*

$$\mathbf{u}^* \cdot \mathbf{p} + L(\mathbf{x}^*, \mathbf{u}^*) = \min_v \left[v \cdot \mathbf{p} + L(\mathbf{x}^*, v)\right] = -H(\mathbf{p}, \mathbf{x}^*),$$

for almost all s and, therefore,

$$\mathbf{p} = -D_v L(\mathbf{x}^*, \mathbf{u}^*) \qquad \text{and} \qquad \mathbf{u}^* = -D_p H(\mathbf{p}, \mathbf{x}^*),$$

where $H = L^*$. Additionally, $\mathbf{p}$ satisfies the terminal condition

$$\mathbf{p}(t_1) = D_x \psi(\mathbf{x}^*(t_1)).$$

Proof. Clearly it is enough to choose R sufficiently large. □

Lemma 32. *Let ψ be differentiable. The minimizing trajectory $\mathbf{x}(\cdot)$ is C^2 and satisfies the Euler-Lagrange equation (4.2). Furthermore,*

$$\dot{\mathbf{p}} = D_x H(\mathbf{p}, \mathbf{x}^*) \qquad \dot{\mathbf{x}} = -D_p H(\mathbf{p}, \mathbf{x}^*).$$

Proof. By its definition $\mathbf{p}$ is continuous. We know that

$$\dot{\mathbf{x}}^*(s) = -D_p H(\mathbf{p}(s), \mathbf{x}^*(s)),$$

almost everywhere. Since the right hand side of the previous identity is continuous, the identity holds everywhere and, therefore, we conclude that $\mathbf{x}^*$ is C^1. Because $\mathbf{p}$ is given by the integral of a continuous function (3.7),

$$\mathbf{p}(r) = D_x \psi(\mathbf{x}^*(t_1)) + \int_r^{t_1} D_x L(\mathbf{x}^*(s), \mathbf{u}^*(s)) ds,$$

we conclude that $\mathbf{p}$ is C^1. Additionally,

$$\dot{\mathbf{x}}^* = -D_p H(\mathbf{p}, \mathbf{x}^*)$$

and, therefore, $\dot{\mathbf{x}}^*$ is C^1, which implies that $\mathbf{x}$ is C^2. We have also

$$\mathbf{p} = -D_v L(\mathbf{x}^*, \dot{\mathbf{x}}^*) \qquad \dot{\mathbf{p}} = -D_x L(\mathbf{x}^*, \dot{\mathbf{x}}^*),$$

from which it follows

$$\frac{d}{dt} D_v L(\mathbf{x}^*, \dot{\mathbf{x}}^*) - D_x L(\mathbf{x}^*, \dot{\mathbf{x}}^*) = 0. \tag{4.4}$$

Thus, since $D_x L(\mathbf{x}^*, \dot{\mathbf{x}}^*) = -D_x H(\mathbf{p}, \mathbf{x}^*)$, we conclude that

$$\dot{\mathbf{p}} = D_x H(\mathbf{p}, \mathbf{x}^*) \qquad \dot{\mathbf{x}}^* = -D_p H(\mathbf{p}, \mathbf{x}^*),$$

as required. □

In the case in which ψ is only Lipschitz and not C^1, we can consider a sequence of C^1 functions, $\psi_n \to \psi$ uniformly, such that

$$\|D_x\psi_n\|_\infty \leq \|D\psi\|_{L^\infty}.$$

for each ψ_n. Let

$$J_n(x,t;\mathbf{u}) = \int_t^{t_1} L(\mathbf{x}_n(s),\mathbf{u}_n(s))ds + \psi_n(\mathbf{x}_n(t_1)),$$

and $\mathbf{x}_n^*$, $\mathbf{u}_n^*$ are, respectively, the corresponding optimal trajectory and optimal control. Similarly, let $\mathbf{p}_n$ be the corresponding adjoint variable. Passing to a subsequence, if necessary, the boundary values $\mathbf{x}_n(t_1)$ and $\mathbf{p}_n(t_1)$ converge, respectively, for some x_0 and p_0. The optimal trajectories $\mathbf{x}_n^*$ and corresponding optimal controls $\mathbf{u}_n^*$ converge uniformly, by using Ascoli-Arzela theorem, to optimal trajectories and controls of the limit problem. Let

$$\mathbf{p}(s) = \lim_{n\to\infty} \mathbf{p}_n(s).$$

Then, for almost every s,

$$\mathbf{u}^* \cdot \mathbf{p}(s) + L(\mathbf{x}^*(s),\mathbf{u}^*(s)) = \inf_v \left[v\cdot \mathbf{p}(s) + L(\mathbf{x}^*(s), v)\right],$$

which implies

$$\mathbf{p}(s) = -D_v L(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s)),$$

for almost all s. But, in the previous equation both terms are continuous functions thus the identity holds for all s.

Lemma 33. *For $t < s \leq t_1$ we have*

$$(\mathbf{p}(s), H(\mathbf{p}(s),\mathbf{x}^*(s))) \in D^-V(\mathbf{x}^*(s), s).$$

Proof. Let $\mathbf{x}^*$ be an optimal trajectory and $\mathbf{u}^*$ the corresponding optimal control. For $r \leq t_1$ and $y \in \mathbb{R}^n$, define $x_r = \mathbf{x}^*(r)$ and consider the sub-optimal control

$$\mathbf{u}^\sharp(s) = \mathbf{u}^*(s) + \frac{y - x_r}{r - t},$$

whose trajectory we denote by $\mathbf{x}^\sharp$, $\mathbf{x}^\sharp(t) = x$. Note that $\mathbf{x}^\sharp(r) = y$.

We have

$$V(x,t) = \int_t^s L(\mathbf{x}^*(\tau), \mathbf{u}^*(\tau))d\tau + V(\mathbf{x}^*(s), s)$$

and, by the sub-optimality of $x^\sharp$,

$$V(\mathbf{x}^*(t), t) \leq \int_t^r L(\mathbf{x}^\sharp(\tau), \mathbf{u}^\sharp(\tau))d\tau + V(y, r).$$

This implies

$$V(\mathbf{x}^*(s), s) - V(y, r) \leq \phi(y, r),$$

with

$$\phi(y,r) = \int_t^r L(\mathbf{x}^\sharp(\tau), \mathbf{u}^\sharp(\tau))d\tau - \int_t^s L(\mathbf{x}^*(\tau), \mathbf{u}^*(\tau))d\tau.$$

Since ϕ is differentiable at y and r,

$$(-D_y\phi(\mathbf{x}^*(s), s), -D_r\phi(\mathbf{x}^*(s), s)) \in D^-V(\mathbf{x}^*(s), s).$$

Observe that

$$\mathbf{x}^\sharp(\tau) = \mathbf{x}^*(\tau) + \frac{y - x_r}{r - t}(\tau - t),$$

and, therefore,

$$D_y\phi(\mathbf{x}^*(s), s) = \int_t^s \left[D_x L \frac{\tau - t}{s - t} + D_v L \frac{1}{s - t}\right] d\tau.$$

Integrating by parts and using (4.4), we obtain

$$D_y\phi(\mathbf{x}^*(s), s) = D_v L(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s)) = -\mathbf{p}(s).$$

Similarly,

$$\begin{aligned} D_r\phi(y,r) = L(y, \mathbf{u}^\sharp(r)) + \int_t^s \Bigg[&-D_x L \frac{y - x_r}{(r-t)^2}(\tau - t) \\ &+ D_x L \frac{-\mathbf{u}^*(r)}{(r-t)}(\tau - t) - D_v L \frac{y - x_r}{(r-t)^2} + D_v L \frac{-\mathbf{u}^*(r)}{r - t}\Bigg] d\tau. \end{aligned}$$

Integrating by parts and evaluating at $y = \mathbf{x}^*(s)$, $r = s$, we obtain

$$D_r\phi(\mathbf{x}^*(s), s) = L(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s)) - \mathbf{u}^*(s)D_vL(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s)) \\ = -H(\mathbf{p}(s), \mathbf{x}^*(s)),$$

as we needed. □

Lemma 34. *The value function V is Lipschitz.*

Proof. Let $t < t_1$ be fixed and x, y arbitrary. We suppose first that $t_1 - t < 1$. Then

$$V(y,t) - V(x,t) \le J(y,t;\mathbf{u}^*) - V(x,t),$$

where $V(x,t) = J(x,t;\mathbf{u}^*)$. Therefore, there exists a constant C, depending only on the Lipschitz constant of ψ and of the supremum of $|D_xL|$, such that

$$V(y,t) - V(x,t) \le C|x-y|.$$

Suppose that $t_1 - t > 1$. Let

$$\begin{cases} \tilde{\mathbf{u}}(s) = \mathbf{u}^* + (x-y) \text{ if } t < s < t+1 \\ \tilde{\mathbf{u}}(s) = \mathbf{u}^*(s) \text{ if } t+1 \le s \le t_1. \end{cases}$$

Then

$$V(y,t) - V(x,t) \le J(y,t;\tilde{\mathbf{u}}) - V(x,t) \le C|x-y|,$$

where the constant C depends only on D_xL and on D_vL, and not on the Lipschitz constant of ψ. Reverting the roles of x and y we conclude

$$|V(y,t) - V(x,t)| \le C|x-y|.$$

Without loss of generality we can suppose that $t < \hat{t}$. Note that

$$|V(x,t) - V(\mathbf{x}^*(\hat{t}), \hat{t})| \le C|t - \hat{t}|.$$

To prove that V is Lipschitz in t it is enough to check that

$$|V(\mathbf{x}^*(\hat{t}),\hat{t}) - V(x,\hat{t})| \leq C|t-\hat{t}|. \tag{4.5}$$

But since $\dot{\mathbf{x}}^*$ is uniformly bounded

$$|\mathbf{x}^*(\hat{t}) - x| \leq C|t-\hat{t}|$$

thus, the previous Lipschitz estimate implies (4.5). □

Lemma 35. *V is differentiable almost everywhere.*

Proof. Since V is Lipschitz, the almost everywhere differentiability follows from Rademacher theorem. □

In general, the value function is Lipschitz and not C^1 or C^2. However we can prove an one-side estimate for second derivatives, i.e. that V is semiconcave.

Lemma 36. *Suppose that $|D^2_{xv}L|, |D^2_{vv}L| \leq C(R)$ whenever $|v| \leq R$. Then, for each $t < t_1$, $V(x,t)$ is semiconcave in x.*

Proof. Fix t and x. Choose $y \in \mathbb{R}^n$ arbitrary. We claim that

$$V(x+y,t) + V(x-y,t) \leq 2V(x,t) + C|y|^2,$$

for some constant C. Clearly,

$$\begin{aligned} &V(x+y,t) + V(x-y,t) - 2V(x,t) \\ &\leq \int_t^{t_1} \left[L(\mathbf{x}^*+\mathbf{y},\dot{\mathbf{x}}^*+\dot{\mathbf{y}}) + L(\mathbf{x}^*-\mathbf{y},\dot{\mathbf{x}}^*-\dot{\mathbf{y}}) - 2L(\mathbf{x}^*,\dot{\mathbf{x}}^*)\right] ds, \end{aligned}$$

where

$$\mathbf{y}(s) = y\frac{t_1-s}{t_1-t}.$$

Since $|D^2_{xx}L| \leq C_1(R)$,

$$L(\mathbf{x}^* + \mathbf{y}, \dot{\mathbf{x}}^* + \dot{\mathbf{y}}) \leq L(\mathbf{x}^*, \dot{\mathbf{x}}^* + \dot{\mathbf{y}}) + D_x L(\mathbf{x}^*, \dot{\mathbf{x}}^* + \dot{\mathbf{y}})\mathbf{y} + C|\mathbf{y}|^2$$

and, in a similar way for the other term. We also have

$$L(\mathbf{x}^*, \dot{\mathbf{x}}^* + \dot{\mathbf{y}}) + L(\mathbf{x}^*, \dot{\mathbf{x}}^* - \dot{\mathbf{y}}) \leq 2L(\mathbf{x}^*, \dot{\mathbf{x}}^*) + C|\dot{\mathbf{y}}|^2 + C|\mathbf{y}||\dot{\mathbf{y}}|.$$

Thus

$$L(\mathbf{x}^* + \mathbf{y}, \dot{\mathbf{x}}^* + \dot{\mathbf{y}}) + L(\mathbf{x}^* - \mathbf{y}, \dot{\mathbf{x}}^* - \dot{\mathbf{y}}) \leq 2L(\mathbf{x}^*, \dot{\mathbf{x}}^*) + C|\mathbf{y}|^2 + C|\dot{\mathbf{y}}|^2.$$

This inequality implies the lemma. ☐

Lemma 37. *We have*

$$(\mathbf{p}(s), H(\mathbf{p}(s), \mathbf{x}^*(s))) \in D^+V(\mathbf{x}^*(s), s)$$

for $t \leq s < t_1$*. Therefore* $DV(\mathbf{x}^*(s), s)$ *exists for* $t < s < t_1$*.*

Proof. Let $\mathbf{u}^*$ be an optimal control at (x, s) and let $\mathbf{p}$ be the corresponding adjoint variable. Define W by

$$W(y, r) = J\left(y, r; \mathbf{u}^* + \frac{\mathbf{x}^*(r) - y}{t_1 - r}\right) - V(x, s).$$

Hence, for each $y \in \mathbb{R}^n$ and $t \leq r < t_1$,

$$V(y, r) - V(x, s) \leq W(y, r),$$

with equality at $(y, r) = (x, s)$. Since W is C^1, it is enough to check that

$$D_y W(\mathbf{x}^*(s), s) = \mathbf{p}(s),$$

and

$$D_r W(\mathbf{x}^*(s), s) = H(\mathbf{p}(s), \mathbf{x}^*(s)).$$

The first identity follows from

$$D_y W(s, \mathbf{x}^*(s)) = \int_s^{t_1} D_x L \varphi + D_v L \frac{d\varphi}{d\tau} d\tau,$$

where $\varphi(\tau) = \frac{t_1 - \tau}{t_1 - s}$. Using the Euler-Lagrange equation

$$\frac{d}{dt} D_v L - D_x L = 0$$

and integration by parts we obtain

$$D_y W(s, \mathbf{x}^*(s)) = -D_v L(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s)) = \mathbf{p}(s).$$

On the other hand,

$$D_r W(s, \mathbf{x}^*(s)) = -L(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s)) + \int_s^{t_1} D_x L \phi + D_v L \frac{d\phi}{d\tau} d\tau,$$

where

$$\phi(\tau) = \frac{\tau - t_1}{t_1 - s} \dot{\mathbf{x}}^*(s).$$

Using again the Euler-Lagrange equation and integration by parts, we obtain

$$D_r W(s, \mathbf{x}^*(s)) = -L(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s), s) + D_v L(\mathbf{x}^*(s), \dot{\mathbf{x}}^*(s)) \dot{\mathbf{x}}^*(s),$$

or equivalently

$$D_r W(s, \mathbf{x}^*(s)) = H(\mathbf{p}(s), \mathbf{x}^*(s)).$$

The last part of the lemma follows from proposition 25. □

This ends the proof of the theorem. □

In what follows we prove that the value function is differentiable at points of uniqueness of optimal trajectory.

A point (x,t) is *regular* if there exists a unique optimal trajectory $\mathbf{x}^*(s)$ such that $\mathbf{x}^*(t)=x$ and

$$V(x,t)=\int_t^{t_1} L(\mathbf{x}^*(s),\dot{\mathbf{x}}^*(s))ds+\psi(\mathbf{x}^*(t_1)).$$

Theorem 38. *V is differentiable with respect to x at (x,t) if and only if (x,t) is a regular point.*

Proof. The next lemma shows that differentiability at a point x implies that x is a regular point:

Lemma 39. *If V is differentiable with respect to x at a point (x,t), then there exists a unique optimal trajectory*

Proof. Since V is differentiable with respect to x at (x,t), then any optimal trajectory satisfies

$$\dot{\mathbf{x}}^*(t)=-D_pH(\mathbf{p}(t),\mathbf{x}^*(t)),$$

since $\mathbf{p}(t)=D_xV(x)$. Therefore, once $D_xV(\mathbf{x}^*(t),t)$ is given, the velocity $\dot{\mathbf{x}}^*(t)$ is uniquely determined. The solution of the Euler-Lagrange equation (4.2) is determined by the initial condition and velocity: $\mathbf{x}^*(t)$ and $\dot{\mathbf{x}}^*(t)$. Thus, the optimal trajectory is unique. □

To prove the other implication we need an auxiliary lemma:

Lemma 40. *Let p such that*

$$\|D_xV(\cdot,t)-p\|_{L^\infty(B(x,2\epsilon))}\to 0$$

when $\epsilon\to 0$. Then V is differentiable with respect to x at (x,t) and $D_xV(x,t)=p$.

Proof. Since V is Lipschitz, it is differentiable almost everywhere. By Fubinni theorem, for almost every point with respect to the Lebesgue measure induced in S^{n-1}, V is differentiable $y = x + \lambda k$, with respect to the Lebesgue measure in $\mathbb{R}$. For these directions

$$\frac{V(y,t) - V(x,t) - p \cdot (y-x)}{|x-y|} = \int_0^1 \frac{(D_x V(x + s(y-x), t) - p) \cdot (y-x)}{|x-y|} ds.$$

Suppose $0 < |x-y| < \epsilon$. Then

$$\left| \frac{V(x,t) - V(y,t) - p \cdot (x-y)}{|x-y|} \right| \leq \|D_x V(\cdot, t) - p\|_{L^\infty(B(x,\epsilon))}.$$

In principle, the last identity only holds almost everywhere. However, for $y \neq x$, the left-hand side is continuous in y. consequently, the inequality holds for all $y \neq x$. Therefore, when $y \to x$,

$$\left| \frac{V(x,t) - V(y,t) - p \cdot (x-y)}{|x-y|} \right| \to 0,$$

which implies $D_x V(x,t) = p$. □

Suppose that V is not differentiable at (x,t). We claim that (x,t) is not regular. By contradiction, suppose that (x,t) is regular. Then if V fails to be differentiable, the previous lemma implies that for each p,

$$\|D_x V(\cdot, t) - p\|_{L^\infty(B(x,\epsilon))} \nrightarrow 0.$$

Thus, we could choose two sequences x_n^1 and x_n^2 such that $x_n^i \to x$ but whose corresponding optimal trajectories $\mathbf{x}_n^i$ satisfy

$$\lim \dot{\mathbf{x}}_n^1(t) \neq \lim \dot{\mathbf{x}}_n^2(t).$$

However, this shows that (x,t) is not regular. Indeed if (x,t) were regular, and x_n were any sequence converging to x, and $\mathbf{x}_n^*(\cdot)$ the corresponding optimal trajectory then

$$\dot{\mathbf{x}}_n^*(t) \to \dot{\mathbf{x}}^*(t).$$

If this were not true, by Ascoli-Arzela theorem, we could extract a convergent subsequence $\dot{\mathbf{x}}_{n_k}(\cdot) \to \dot{\mathbf{y}}(\cdot)$, and for which

$$\dot{\mathbf{x}}_{n_k}^*(t) \to v \neq \dot{\mathbf{x}}^*(t).$$

Let $\mathbf{y}(\cdot)$ be the solution to the Euler-Lagrange equation with initial condition $\mathbf{y}(t) = \mathbf{x}(t)$ and $\dot{\mathbf{y}}(t) = v$. Note that $\mathbf{x}_n^*(\cdot) \to \mathbf{y}(\cdot)$ and $\dot{\mathbf{x}}_n^*(\cdot) \to \dot{\mathbf{y}}(\cdot)$, uniformly in compact sets, and, therefore,

$$\begin{aligned} V(x,t) = \lim_{n\to\infty} V(x_n,t) &= \lim_{n\to\infty} J(x_n,t;\dot{\mathbf{x}}_n) \\ &= J(x,t;\dot{\mathbf{y}}) > J(x,t;\dot{\mathbf{x}}^*) = V(x,t), \end{aligned}$$

since the trajectory $\mathbf{y}$ cannot be optimal, by regularity, which is a contradiction. □

Remark. This theorem implies that all points of the form $(\mathbf{x}^*(s), s)$, in which $\mathbf{x}^*$ is and optimal trajectory are regular for $t < s < t_1$.

Exercise 40. *Show that in the optimal control "bounded control space" setting, the value function is Lipschitz continuous if the terminal cost is Lipschitz continuous.*

Exercise 41. *In the optimal control "bounded control space" setting, show that if ψ is Lipschitz, for any (x,t) there exists $\mathbf{p}$ such that*

$$(\mathbf{p}(s), H(\mathbf{p}(s), \mathbf{x}^*(s))) \in D^- V(\mathbf{x}^*(s), s)$$

for $t < s \le t_1$ and

$$(\mathbf{p}(s), H(\mathbf{p}(s), \mathbf{x}^*(s))) \in D^+ V(\mathbf{x}^*(s), s)$$

for $t \le s < t_1$.

4.4 Viscosity solutions

As we have mentioned before, in general the value function is not differentiable enough to be a classical solution to the Hamilton-Jacobi equation. Nevertheless, as we discuss now it is a solution to the Hamilton-Jacobi equation in an appropriate weak sense.

In this section we discuss the viscosity solutions in the calculus of variations setting. With small modifications our results hold for the bounded control setting, and therefore we have added exercises in which guide the reader into filling the gaps.

Theorem 41. *Consider the calculus of variations setting for the optimal control problem. Suppose that the value function V is differentiable at (x, t). Then, at this point, V satisfies the Hamilton-Jacobi equation*

$$-V_t + H(D_x V, x, t) = 0. \tag{4.6}$$

Proof. If V is differentiable at (x, t) then the result follows by using statement 6 in theorem 28. □

Exercise 42. *Show that* (4.6) *also holds in the "bounded control case" setting.* **Hint:** *use exercises 40 and 41.*

Corollary 42. *Consider the calculus of variations setting for the optimal control problem. Then the value function V satisfies the Hamilton-Jacobi equation almost everywhere.*

Proof. Since the value function V is differentiable almost everywhere, by theorem 28, theorem 41 implies this result. □

Exercise 43. *Show that the previous corollary also holds in the "bounded control case" setting.*

However, it is not true that a Lipschitz function satisfying the Hamilton-Jacobi equation almost everywhere is the value function of the terminal value problem, as shown in the next example.

Example 16. Consider the Hamilton-Jacobi equation

$$-V_t + |D_x V|^2 = 0$$

with terminal data $V(x,1) = 0$. The value function is $V \equiv 0$, which is a (smooth) solution of the Hamilton-Jacobi equation However, there are other solutions, for instance,

$$V(x,t) = \begin{cases} 0 & \text{if} \quad |x| \geq 1-t \\ |x| - 1 + t & \text{if} \quad |x| < 1-t \end{cases}$$

which satisfy the same terminal condition $t = 1$ and is solution almost everywhere. ◀

A bounded uniformly continuous function V is a *viscosity subsolution (resp. supersolution)* of the Hamilton-Jacobi equation (4.6) if for any C^1 function ϕ and any interior point $(x,t) \in$ argmax $V - \phi$ (resp. argmin) then

$$-D_t\phi + H(D_x\phi, x, t) \leq 0$$

(resp. ≥ 0) at (x,t). A bounded uniformly continuous function V is a *viscosity solution* of the Hamilton-Jacobi equation if it is both a sub and supersolution.

The value function is a viscosity solution of (4.6), although it may not be a classical solution. The motivation behind the definition of viscosity solution is the following: if V is differentiable and $(x,t) \in argmax V - \phi$ (or argmin) then $D_x V = D_x\phi$ and $D_t V = D_t\phi$, therefore we should have both inequalities. The specific choice of inequalities is related with the following parabolic approximation of the Hamilton-Jacobi equation

$$-D_t u^\epsilon + H(D_x u^\epsilon, x, t) = \epsilon \Delta u^\epsilon. \tag{4.7}$$

This equation arises naturally in optimal stochastic control (see chapter 6). The limit $\epsilon \to 0$ corresponds to the case in which the diffusion coefficient vanishes.

Proposition 43. *Let u^ϵ be a family of solutions of (4.7) such that, as $\epsilon \to 0$, the sequence $u^\epsilon \to u$ uniformly. Then u is a viscosity solution of (4.6).*

Proof. Suppose that $\phi(x,t)$ is a C^2 function such that $u - \phi$ has a strict local maximum at $(\overline{x}, \overline{t})$. We must show that

$$-D_t\phi + H(D_x\phi, \overline{x}, \overline{t}) \leq 0.$$

By hypothesis, $u^\epsilon \to u$ uniformly. Therefore we can find sequences $(\overline{x}_\epsilon, \overline{t}_\epsilon) \to (\overline{x}, \overline{t})$ such that $u^\epsilon - \phi$ has a local maximum at $(\overline{x}_\epsilon, \overline{t}_\epsilon)$. Therefore,

$$Du^\epsilon(\overline{x}_\epsilon, \overline{t}_\epsilon) = D\phi(\overline{x}_\epsilon, \overline{t}_\epsilon)$$

and

$$\Delta u^\epsilon(\overline{x}_\epsilon, \overline{t}_\epsilon) \leq \Delta\phi(\overline{x}_\epsilon, \overline{t}_\epsilon).$$

Consequently,

$$-D_t\phi(\overline{x}_\epsilon, \overline{t}_\epsilon) + H(D_x\phi(\overline{x}_\epsilon, \overline{t}_\epsilon), \overline{x}_\epsilon, \overline{t}_\epsilon) \leq \epsilon\Delta\phi(\overline{x}_\epsilon, \overline{t}_\epsilon).$$

It is therefore enough to take $\epsilon \to 0$ to end the proof. □

An useful characterization of viscosity solutions is given in the next proposition:

Proposition 44. *Let V be a bounded uniformly continuous function. Then V is a viscosity subsolution of (4.6) if and only if for each $(p, q) \in D^+V(x,t)$,*

$$-q + H(p, x, t) \leq 0.$$

Similarly, V is a viscosity supersolution if and only if for each $(p, q) \in D^-V(x,t)$,

$$-q + H(p, x, t) \geq 0.$$

Proof. This result is an immediate corollary of proposition 26. □

Example 17. In example 16 we have found two different almost everywhere solutions to

$$-V_t + |D_x V|^2 = 0$$

satisfying $V(x,T) = 0$.

It is easy to check that the value function $V = 0$ is viscosity solution (it is smooth, satisfies the equation and the terminal condition) and it agrees with the value function of the corresponding optimal control problem. The other solution, which is an almost everywhere solution is not a viscosity solution (check!).

Now we will show that the definition of viscosity solution is consistent with classical solutions.

Proposition 45. *Let V be a C^1 viscosity solution of (4.6). Then V is a classical solution.*

Proof. If V is differentiable then

$$D^+V = D^-V = \{(D_x V, D_t V)\}.$$

Since V is a viscosity solution, we obtain immediately

$$-D_tV + H(D_xV, x, t) \leq 0, \quad \text{and} \quad -D_tV + H(D_xV, x, t) \geq 0,$$

therefore $-D_tV + H(D_xV, x, t) = 0$. □

Theorem 46. *Let V be the value function of the terminal value problem. Then V is a viscosity solution to*

$$-V_t + H(D_xV, x) = 0.$$

Proof. Let $\varphi : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$, $\varphi(x,t)$, be a C^∞ function, and let $(x_0, t_0) \in \operatorname{argmax}(V - \varphi)$. We must show that

$$-\varphi_t(x_0, t_0) + H(D_x\varphi(x_0, t_0), x_0) \leq 0,$$

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

$$-\varphi_t(x_0, t_0) - v \cdot D_x\varphi(x_0, t_0) - L(x_0, v) \leq 0.$$

Fix $v \in \mathbb{R}^d$. Let $\mathbf{x} = x_0 + v(t - t_0)$. Then, for any $h > 0$

$$\int_{t_0}^{t_0+h} \varphi_t + vD_x\varphi(\mathbf{x}(s), s)ds = \varphi(\mathbf{x}(t_0 + h), t_0 + h) - \varphi(x_0, t_0)$$

$$\geq V(\mathbf{x}(t_0 + h), t_0 + h) - V(x_0, t_0) \geq -\int_{t_0}^{t_0+h} L(\mathbf{x}, \dot{\mathbf{x}})dt.$$

Dividing by h and letting $h \to 0$ we obtain the result.

Now let $(x_0, t_0) \in \operatorname{argmin}(V - \varphi)$. We must show that

$$-\varphi_t(x_0, t_0) + H(D_x\varphi(x_0, t_0), x_0) \geq 0,$$

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

$$-\varphi_t(x_0, t_0) - v \cdot D_x\varphi(x_0, t_0) - L(x_0, v) \geq 0.$$

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

$$-\varphi_t(x_0, t_0) - v \cdot D_x\varphi(x_0, t_0) - L(x_0, v) < -\theta,$$

for all v. Because the mapping $v \mapsto L$ is superlinear and φ is C^1, there exists a $R > 0$ and $r_1 > 0$ such that for all $(x, t) \in B_{r_1}(x_0, t_0)$ and all $v \in B_R^c(0) = \mathbb{R}^d \setminus B_R(0)$ we have

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

By continuity, for some $0 < r < r_1$ and all $(x, t) \in B_r(x_0, t_0)$ we have

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

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

Therefore, for any trajectory $\mathbf{x}$ with $\mathbf{x}(0) = x_0$ and any $T \geq 0$ such that the trajectory $\mathbf{x}$ stays near x_0 on $[t_0, t_0+T]$, i.e., $(\mathbf{x}(t), t) \in B_r(x_0, t_0)$ for $t \in [t_0, t_0 + T]$ we have

$$\begin{aligned} &V(\mathbf{x}(t_0+T), t_0+T) - V(x_0, t_0) \geq \varphi(\mathbf{x}(t_0+T), t_0+T) - \varphi(x_0, t_0) \\ &= \int_{t_0}^{t_0+T} \varphi_t(\mathbf{x}(t), t) + \dot{\mathbf{x}}(t) \cdot D_x\varphi(\mathbf{x}(t))\big) dt \\ &\geq \frac{\theta}{2} \int_{t_0}^{t_0+T} dt - \int_{t_0}^{t_0+T} L(\mathbf{x}, \dot{\mathbf{x}}) dt. \end{aligned}$$

This yields

$$V(x_0, t_0) \leq -\frac{\theta}{2} T + \int_{t_0}^{t_0+T} L(\mathbf{x}, \dot{\mathbf{x}}) dt + V(\mathbf{x}(t_0+T), t_0+T)$$

Choose $\epsilon < \frac{\theta T}{4}$. Let $\mathbf{x}^\epsilon$ be such that:

$$V(x_0, t_0) \geq \int_{t_0}^{t_0+T} L(\mathbf{x}^\epsilon, \dot{\mathbf{x}}^\epsilon) dt + V(\mathbf{x}^\epsilon(t_0+T), t_0+T) - \epsilon$$

This then yields a contradiction.

□

Exercise 44. *Show that the function $V(x,t)$ given by the Lax-Hopf formula is Lipschitz in x for each $t < t_1$, regardless of the smoothness of the terminal data (note, however that the constant depends on t).*

Exercise 45. *Use the Lax-Hopf formula to determine the viscosity solution of*

$$-u_t + u_x^2 = 0,$$

para $t < 0$ and $u(x,0) = \pm x^2 - 2x$.

Exercise 46. *Use the Lax-Hopf formula to determine the viscosity solution of*

$$-u_t + u_x^2 = 0,$$

for $t < 0$ and

$$u(x,0) = \begin{cases} 0 \text{ if } x < 0 \\ x^2 \text{ if } 0 \leq x \leq 1 \\ 2x - 1 \text{ if } x > 1. \end{cases}$$

4.5 Uniqueness of viscosity solutions

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

Lemma 47. *Let V be a viscosity solution of*

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

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

$$-\phi_t(x_0,t_0) + H(D_x\phi(x_0,t_0), x_0) \leq 0 \text{ (resp. } \geq 0) \quad \text{at} \quad (x_0,t_0). \tag{4.8}$$

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

Proof. Only the case $t_0 = 0$ requires proof since in the other case the maximum is interior and then the viscosity property (the definition of viscosity solution) yields the inequality. So suppose $(x_0, 0)$ is a strict maximum point. Consider

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

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

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

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

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

Analogously we obtain the opposite inequality for the case of local minimum, using $\tilde{\phi} = \phi - \frac{\epsilon}{t}$. □

Finally we establish uniqueness of viscosity solutions:

Theorem 48 (Uniqueness)**.** *Suppose H satisfies*

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

Then the value function is the unique viscosity solution to the Hamilton-Jacobi equation

$$-V_t + H(D_xV, x) = 0$$

that satisfies the terminal condition $V(x, T) = \psi(x)$.

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

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

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

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

When ϵ,λ are sufficiently small we have

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

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

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

and

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

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

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

as $\epsilon\to 0$.

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

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

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

Let ϕ be given by

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

Then, the difference

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

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

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

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

the difference

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

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

Therefore

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

and

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

Simplifying, we have

$$\lambda - 2\frac{t_{\epsilon,\lambda} - s_{\epsilon,\lambda}}{\epsilon^2} + H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} + 2\epsilon x_{\epsilon,\lambda}, x_{\epsilon,\lambda}) \leq 0, \tag{4.9}$$

and

$$-\lambda - 2\frac{t_{\epsilon,\lambda} - s_{\epsilon,\lambda}}{\epsilon^2} + H(2\frac{x_{\epsilon,\lambda} - y_{\epsilon,\lambda}}{\epsilon^2} - 2\epsilon y_{\epsilon,\lambda}, y_{\epsilon,\lambda}) \geq 0. \tag{4.10}$$

From (4.9) we gather that

$$H(2\frac{x_{\epsilon,\lambda}-y_{\epsilon,\lambda}}{\epsilon^2}+2\epsilon x_{\epsilon,\lambda},x_{\epsilon,\lambda})\leq-\lambda+\frac{o(1)}{\epsilon}. \tag{4.11}$$

By subtracting (4.9) to (4.10) we have

$$\begin{aligned}
&2\lambda\leq H(2\frac{x_{\epsilon,\lambda}-y_{\epsilon,\lambda}}{\epsilon^2}-2\epsilon y_{\epsilon,\lambda},y_{\epsilon,\lambda})-H(2\frac{x_{\epsilon,\lambda}-y_{\epsilon,\lambda}}{\epsilon^2}+2\epsilon x_{\epsilon,\lambda},x_{\epsilon,\lambda})\\
&\leq H(2\frac{x_{\epsilon,\lambda}-y_{\epsilon,\lambda}}{\epsilon^2}-2\epsilon y_{\epsilon,\lambda},y_{\epsilon,\lambda})-H(2\frac{x_{\epsilon,\lambda}-y_{\epsilon,\lambda}}{\epsilon^2}-2\epsilon y_{\epsilon,\lambda},x_{\epsilon,\lambda})\\
&+H(2\frac{x_{\epsilon,\lambda}-y_{\epsilon,\lambda}}{\epsilon^2}-2\epsilon y_{\epsilon,\lambda},x_{\epsilon,\lambda})-H(2\frac{x_{\epsilon,\lambda}-y_{\epsilon,\lambda}}{\epsilon^2}+2\epsilon x_{\epsilon,\lambda},x_{\epsilon,\lambda})\\
&\leq\left(C+CH(2\frac{x_{\epsilon,\lambda}-y_{\epsilon,\lambda}}{\epsilon^2}+2\epsilon x_{\epsilon,\lambda},x_{\epsilon,\lambda})\right)|x_{\epsilon,\lambda}-y_{\epsilon,\lambda}|\\
&+C\epsilon\left(\left|2\frac{x_{\epsilon,\lambda}-y_{\epsilon,\lambda}}{\epsilon^2}+2\epsilon x_{\epsilon,\lambda}\right|+\left|2\frac{x_{\epsilon,\lambda}-y_{\epsilon,\lambda}}{\epsilon^2}-2\epsilon y_{\epsilon,\lambda}\right|\right)|x_{\epsilon,\lambda}-y_{\epsilon,\lambda}|\\
&\leq\left(\frac{o(1)}{\epsilon}+C\right)(|x_{\epsilon,\lambda}-y_{\epsilon,\lambda}|+|t_{\epsilon,\lambda}-s_{\epsilon,\lambda}|)\to 0,
\end{aligned}$$

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

4.6 Bibliographical notes

The main references for this section are [FS06], [Lio82], [BCD97], [Bar94]. Introductory material can be found in [Eva98b]. A very nice introduction to viscosity solutions written in Portuguese is the book [LLF].

5

Stationary deterministic control

In this chapter we consider stationary control problems. These include the discounted cost infinite horizon as well as stationary control problems. The discounted cost infinite horizon problem, corresponds to the Hamilton-Jacobi equation is

$$\alpha u + H(Du, x) = 0,$$

where in stationary optimal control problem we have the Hamilton-Jacobi equation

$$H(D_x u, x) = \overline{H}.$$

Because two of the main applications we will be considering are homogenization problems and Aubry-Mather theory, we will consider only periodic problems and we will not discuss boundary conditions. Supplementary material can be found in [Bar94], for instance. To simplify, we will only consider the calculus of variations setting, however similar results hold for the bounded control setting.

5.1 Discounted cost infinite horizon

We will work in the calculus of variations setting, with a convex and superlinear Lagrangian $L(x, v)$. Let $\alpha > 0$ be the discount rate. We define the discounted cost function J_α, with discount rate α, as

$$J_\alpha(x; u) = \int_0^\infty L(\mathbf{x}(s), \dot{\mathbf{x}}(s)) e^{-\alpha s} ds.$$

In this case, the optimal trajectories $\mathbf{x}(\cdot)$ satisfy the differential equation

$$\dot{\mathbf{x}} = \mathbf{u},$$

with the initial condition $\mathbf{x}(0) = x$.

As before, the value function, u_α, is given by

$$u_\alpha(x) = \inf J_\alpha(x; \mathbf{u}),$$

where infimum is taken over all controls $\mathbf{u} \in L^\infty_{loc}$.

The dynamic programming principle in this case is

Proposition 49. *For each $t > 0$*

$$u_\alpha(x) = \inf_{\mathbf{x}(0)=x} \left[\int_0^t L(\mathbf{x}(s), \dot{\mathbf{x}}(s)) e^{-\alpha s} ds + e^{-\alpha t} u_\alpha(\mathbf{x}(t)) \right].$$

Proof. Observe that

$$\begin{aligned} u_\alpha(x) = \inf_{\mathbf{x}(0)=x} \Bigg[\int_0^t L(\mathbf{x}(s), \dot{\mathbf{x}}(s)) e^{-\alpha s} ds \\ + e^{-\alpha t} \int_t^\infty L(\mathbf{x}(s), \dot{\mathbf{x}}(s)) e^{-\alpha(s-t)} ds \Bigg] \\ \geq \inf_{\mathbf{x}(0)=x} \left[\int_0^t L(\mathbf{x}(s), \dot{\mathbf{x}}(s)) e^{-\alpha s} ds + e^{-\alpha t} u_\alpha(\mathbf{x}(t)) \right]. \end{aligned}$$

The other inequality is left as an exercise:

Exercise 47. *Show that*

$$u_\alpha(x) \leq \inf_{\mathbf{x}(0)=x} \left[\int_0^t L(\mathbf{x}(s), \dot{\mathbf{x}}(s)) e^{-\alpha s} ds + e^{-\alpha t} u_\alpha(\mathbf{x}(t)) \right].$$

□

Because of the dynamic programming principle, it is clear that

$$V(x,t) = e^{-\alpha t} u_\alpha(x)$$

is a viscosity solution of

$$-V_t + e^{-\alpha t} H(e^{\alpha t} D_x V, x) = 0.$$

This then implies

Corollary 50. u^α *is a viscosity solution of*

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

Furthermore

Corollary 51. *If* u_α *is differentiable then it is a solution of*

$$H(D_x u_\alpha, x) + \alpha u_\alpha = 0. \tag{5.1}$$

Exercise 48. *Show that the optimal trajectories for the discounted cost infinite horizon are solutions to the (negatively damped) Euler-Lagrange equation*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\mathbf{x}}} - \alpha \frac{\partial L}{\partial \dot{\mathbf{x}}} - \frac{\partial L}{\partial \mathbf{x}} = 0. \tag{5.2}$$

Note that if $\mathbf{x}(t)$ satisfies (5.2), the energy H may not be conserved

Example 18. Let $L(x, v) = \frac{v^2}{2} + \cos x$. Then (5.2) reads

$$\ddot{\mathbf{x}} - \alpha\dot{\mathbf{x}} + \sin \mathbf{x} = 0.$$

When $\alpha = 0$ the energy

$$H = \frac{\dot{\mathbf{x}}^2}{2} - \cos \mathbf{x}$$

is constant in time, but for $\alpha > 0$ we have

$$\frac{dH}{dt} = \alpha\dot{\mathbf{x}}^2.$$

Therefore, the energy increases in time unless $\dot{\mathbf{x}} = 0$. ◀

Proposition 52. *Suppose that* $\mathbf{x}(t)$ *satisfies (5.2). Then*

$$\frac{dH}{dt} = \alpha D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t)) \cdot \dot{\mathbf{x}}(t).$$

Proof. Let

$$\mathbf{p}(t) = -D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t)).$$

Then we have

$$\begin{aligned} \frac{dH}{dt} &= D_p H \cdot \dot{\mathbf{p}} + D_x H \cdot \dot{\mathbf{x}} \\ &= \dot{\mathbf{x}} \cdot (\alpha D_v L + D_x L) - D_x L \cdot \dot{\mathbf{x}} = \alpha D_v L \cdot \dot{\mathbf{x}}. \end{aligned}$$

□

We assume now that H is $\mathbb{Z}^n$ periodic in x. We prove some estimates that will be necessary in the next section to show that as $\alpha \to 0$, the solution u_α converges (up to constants) to a solution of

$$H(D_x u, x) = \overline{H}. \tag{5.3}$$

for some $\overline{H}$.

Theorem 53. *Let u_α be a viscosity solution to*

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

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

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

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

which yields that αu_α is uniformly bounded.

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

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

for all x, y. By contradiction, assume that for every $M > 0$ there exists x and y such that

$$u_\alpha(x) - u_\alpha(y) > M|x - y|.$$

Let $\varphi(x) = u_\alpha(y) + M|x - y|$. Then $u_\alpha(x) - \varphi(x)$ has a maximum at some point $x \ne y$. Therefore

$$\alpha u_\alpha(x) + H\left(M\tfrac{x-y}{|x-y|}, x\right) \le 0,$$

which by the coercivity of H yields a contradiction if M is sufficiently large. $\square$

Example 19. We can also use directly calculus of variations methods to show that the exists C, independent of α, such that

$$u_\alpha \leq \frac{C}{\alpha}.$$

Indeed, since $L(x,0)$ is bounded

$$u_\alpha(x) \leq J_\alpha(x,0) \leq \int_0^\infty L(x,0)e^{-\alpha s}ds \leq \frac{C}{\alpha}.$$

◀

5.2 Periodic problems

We now address stationary problems in the periodic setting. These are extremely important in homogenization problems, discussed in section 5.6 as well in Aubry-Mather theory, the subject of chapter 8.

Theorem 54. *(Stability of viscosity solutions) Assume that for $\alpha > 0$ function u^α is a viscosity solution for $H^\alpha(u, Du, x) = 0$. Let $H^\alpha \to H$ uniformly on compact sets, and $u^\alpha \to u$ uniformly. Then u is a viscosity solution for $H(u, Du, x) = 0$.*

Proof. Suppose $u - \varphi$ has a strict local maximum (resp. minimum) at a point x_0. Then there exists $x_\alpha \to x$ such that $u_\alpha - \varphi$ has a local maximum (resp. minimum) at x_α. Then

$$H^\alpha(u^\alpha(x_\alpha), D\varphi(x_\alpha), x_\alpha) \leq 0 \qquad (\text{resp. } \geq 0).$$

Letting $\alpha \to 0$ finishes the proof. □

As demonstrated in context of homogenization of Hamilton-Jacobi equations, in the classic but unpublished paper by Lions, Papanicolaou and Varadhan [LPV88], it is possible to construct, using the previous result, viscosity solutions to the stationary Hamilton-Jacobi equation

$$H(Du, x) = \overline{H}. \tag{5.4}$$

Theorem 55 (Lions, Papanicolao, Varadhan)**.** *There exists a number $\overline{H}$ and a function $u(x)$, $\mathbb{Z}^d$ periodic in x, that solves (5.4) in the viscosity sense.*

Proof. Since $u_\alpha - \min u_\alpha$ is periodic, equicontinuous, and uniformly bounded, it converges, up to subsequences, to a function u. Moreover $u_\alpha \leq \frac{C}{\alpha}$, thus αu_α converges uniformly, up to subsequences, to a constant, which we denote by $-\overline{H}$. Then, the stability theorem for viscosity solutions, theorem 54, implies that u is a viscosity solution of

$$H(Du, x) = \overline{H}.$$

□

Exercise 49. *Let $u : \mathbb{R} \to \mathbb{R}$ be continuous and piecewise differentiable (with left and right limits for the derivative at any point). Show that u is a viscosity solution of*

$$H(D_x u, x) = \overline{H}$$

if

1. *u satisfies the equation almost everywhere;*

2. *whenever $D_x u$ is discontinuous then $D_x u(x^-) > D_x u(x^+)$.*

5.3 Some examples

In this section we discuss some examples and explicit solutions in the periodic setting. We start with two linear problems and then follow with a non-linear example.

Example 20. Consider a linear (nonresonant) Hamiltonians

$$H(p, x) = \omega \cdot p + V(x, y). \tag{5.5}$$

Suppose u is a smooth viscosity solution of (5.4) for this Hamiltonian. The divergence theorem yields

$$\int_{\mathbb{T}^n} \omega \cdot D_x u = 0.$$

Therefore

$$\overline{H}(0) = \int_{\mathbb{T}^n} V, \tag{5.6}$$

and $\overline{H}(P) = \overline{H}(0) + \omega \cdot P$.

For the example

$$(1, \sqrt{2}) \cdot Du + \cos(2\pi x)$$

we obtained $D_P \bar{H} = (1, \sqrt{2})$ and $\bar{H}(0, 0) = 0$.

The Hamilton-Jacobi equation

$$(1, 1) \cdot Du + \cos(2\pi x),$$

is non-resonant because of the specific potential that we used, although the vector $(1, 1)$ is rationally dependent. ◀

Example 21. Linear resonant linear Hamiltonians (5.5) may fail to have a viscosity solutions. An example is

$$(0, 1) \cdot Du + \sin(2\pi x) = \overline{H}.$$

The formula (5.6) yields $\overline{H}(0) = 0$ if there were a solution of (5.4). However, we have

$$\inf_{\phi} \sup_{x} H(D_x\phi, x) = 1.$$

Let ϕ be an arbitrary periodic function. Set $x_0 = \frac{1}{4}$, so that

$$\sin(2\pi x_0) = 1.$$

Then $\phi(x_0, y)$ is a periodic function of y and so $D_y\phi(x_0, y) = 0$ at some $y = y_0$. Thus

$$\sup_{x} H(D_x\phi, x) \geq H(D_x\phi(x_0, y_0), x_0, y_0) = 1.$$

Example 22 (One dimensional pendulum)**.** The Hamiltonian corresponding to a one-dimensional pendulum with unit mass and unit length is

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

In this case, it is not difficult to determine explicitly the solution to the Hamilton-Jacobi equation

$$H(P + D_x u, x) = \overline{H}(P),$$

where P is a real parameter. In fact, for $P \in \mathbb{R}$ and almost every $x \in \mathbb{R}$, the solution $u(P, x)$ satisfies

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

consequently, $\overline{H}(P) \geq 1$ and, therefore,

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

Thus

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

where $|s(y)| = 1$. Since H is convex em p and u is a viscosity solution, the only possible discontinuities on the derivative of u are the ones that satisfy $D_x u(x^-) - D_x u(x^+) > 0$, see exercise 49. Therefore s can change sign from 1 to -1 at any point, however the jumps from -1 to 1 can only happen when

$$\sqrt{2(\overline{H}(P) + \cos 2\pi x)} = 0.$$

Since we are looking for 1-periodic solutions, there are only two cases to consider. The first, in which $\overline{H}(P) > 1$ and the solution is C^1 since $\sqrt{2(\overline{H}(P) + \cos 2\pi y)}$ never vanishes. In this case $\overline{H}(P)$ can be determined as from P through the equation

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

It is easy to check that this equation has a unique solution $\overline{H}(P)$ whenever

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

that is,

$$|P| > \frac{4}{\pi}.$$

The second case occurs whenever the last inequality does not hold, that is $\overline{H}(P) = 1$ and thus $s(x)$ can have discontinuities. In fact, $s(x)$ jumps from -1 to 1 whenever $x = \frac{1}{2} + k$, with $k \in \mathbb{Z}$, and there exists a point x_0 defined by the equation

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

such that $s(x)$ jumps from 1 to -1 at $x_0 + k$, $k \in \mathbb{Z}$. ◀

Exercise 50. *Let $\phi : \mathbb{T}^n \to \mathbb{R}$ be a C^1 function not identically constant. Show that there exist two distinct viscosity solutions of*

$$D_x u \cdot (D_x u - D_x \phi) = 0,$$

whose difference is not constant.

5.4 Regularity

In this section we establish a-priori Lipschitz and semiconcavity estimates for stationary Hamilton-Jacobi equation. We start by the Lipschitz estimates:

Theorem 56. *Let $H(p,x) : \mathbb{R}^{2n} \to \mathbb{R}$ be a continuous function satisfying*

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

Let $u : \mathbb{T}^d \to \mathbb{R}$ be a viscosity solution to

$$H(Du,x) = C.$$

Then u is Lipschitz, and the Lipschitz constant does not depend on u.

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

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

Then, it is enough to argue as in the proof of Theorem 53. □

Recall that a function u is semiconcave if there exists a constant C such that

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

We assume that $L(x,v)$ satisfies the following estimate

$$\begin{aligned} &L(x+\theta y, v+\eta y) - 2L(x,v) + L(x-\theta y, v-\eta y) \\ &\quad \leq (C + CL(x,v))(\theta^2+\eta^2)|y|^2. \end{aligned} \tag{5.7}$$

Proposition 57. *Consider the solution V be a solution to the value problem with any (bounded uniformly continuous) terminal data ψ at time T. Then V is semiconcave in x for each fixed time $t < T$.*

Proof. We will do the proof for $t = 0$. Fix $\epsilon > 0$. Let $\mathbf{x}$ be a trajectory such that

$$V(x,0) \geq \int_0^T L(\mathbf{x},\dot{\mathbf{x}})ds + \psi(\mathbf{x}(T)) - \epsilon.$$

Then we have

$$\int_0^T L(\mathbf{x},\dot{\mathbf{x}})ds \leq C,$$

for some constant uniformly bounded as $\epsilon \to 0$.

Clearly

$$V(x \pm y, 0) \leq \int_0^T L(\mathbf{x} \pm y\frac{T-s}{T}, \dot{\mathbf{x}} \mp \frac{y}{T})ds + \psi(\mathbf{x}(T)).$$

Therefore

$$\begin{aligned} &V(x+y,0) - 2V(x,0) + V(x-y,0) \\ &\quad \leq \epsilon + \int_0^T \Big[L(\mathbf{x} + y\frac{T-s}{T}, \dot{\mathbf{x}} - \frac{y}{T}) - 2L(\mathbf{x},\dot{\mathbf{x}}) \\ &\qquad + L(\mathbf{x} - y\frac{T-s}{T}, \dot{\mathbf{x}} + \frac{y}{T})\Big] ds \\ &\quad \leq C(1 + \int_0^T L(\mathbf{x},\dot{\mathbf{x}})ds)|y|^2 \leq C|y|^2. \end{aligned}$$

□

Proposition 58. *Let u be a viscosity solution of $H(D_x u, x) = 0$. Then u is semiconcave.*

Proof. Consider the Hamilton-Jacobi equation.

$$-V_t + H(D_x V, x) = 0 \tag{5.8}$$

with $V(x,T) = u(x)$. Then $V(x,t) = u(x)$ is a viscosity solution to (5.8). By the uniqueness result for viscosity solutions we have that $V = u$ is the value function for the terminal value problem with terminal cost $\psi = u$. But then the previous proposition implies semiconcavity. □

Corollary 59. *Let $u : \mathbb{T}^1 \to \mathbb{R}$ be a viscosity solution of (5.4). Then Du satisfies the following jump condition: $D_x u(x^-) - D_x u(x^+) > 0$.*

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

5.5 The effective Hamiltonian

Theorem 60. *Let H be convex in p. Let u be a viscosity sub-solution of $H(Du,x) = C$ and let $u^\epsilon = u * \eta_\epsilon$ be a standard smoothing. Then:*

$$H(Du^\epsilon(x), x) \leqslant C + O(\epsilon), \tag{5.9}$$

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

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

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

$|y| \sup |D_x H|$.

$$\begin{aligned} C &\geqslant \int \eta_\epsilon(y) H(Du(x-y), x-y) dy \\ &\geqslant \int \eta_\epsilon(y) H(Du(x-y), x) dy - O(\epsilon). \end{aligned}$$

Now, Jensen's inequality yields

$$\begin{aligned} &\int \eta_\epsilon(y) H(Du(x-y), x) dy \\ &\quad \geq H\big(\textstyle\int \eta_\epsilon(y) Du(x-y) dy, x\big) = H(Du^\epsilon(x), x), \end{aligned}$$

which completes the proof. □

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

Theorem 61. *Let H be convex in p, $\mathbb{Z}^d$-periodic in x and $\frac{\partial H}{\partial x}$ is bounded. Let C be a real number, such that $H(Du, x) = C$ has a viscosity solution u. Then*

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

Proof. Let u be a viscosity solution. Inequality (5.9) implies

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

To show the opposite inequality we take any smooth function $\varphi(x)$. Due to periodicity, the set of points where $u - \varphi$ achieves a local minimum is non empty. For example, one could consider a point of global minimum. Let x_φ be a local minimum point for $u - \varphi$. The definition of viscosity solution implies $H(D\varphi(x_\varphi), x_\varphi) \geqslant C$. Thus, for any smooth function $\varphi(x)$, we have

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

Taking infimum over φ completes the proof. □

5.6 Homogenization

Homogenization theory for Partial Differential Equations studies solutions with high frequency oscillations. Such rapid oscillations may represent small-scale or microscopic structure of a material. The main goal of this theory is to understand the limits as oscillations become more and more rapid. In this section we are interested in understanding he limit as $\epsilon \to 0$ of the viscosity solutions of the Hamilton-Jacobi-Belmann

$$-V_t^\epsilon + H\left(D_x V^\epsilon, \frac{x}{\epsilon}\right) = 0, \tag{5.10}$$

with terminal condition $V^\epsilon(x, T) = g^\epsilon(x)$.

We assume that $H(p, y)$ is smooth, strictly convex in p, bounded from below and $[0, 1]^n$-periodic in y. Furthermore, we suppose that $g^\epsilon \to g$ uniformly. To understand what should be the limit problem we start with some formal calculations. Many of the results in this chapter were proved by the first time in the "classical-yet-unpublished" paper [LPV88]. For more details about homogenization of Hamilton-Jacobi equations the reader should consult [Con95], [Con97], [Con96] or the book [BD98].

5.6.1 Formal calculations

Suppose $V^\epsilon \to V_0$ uniformly as $\epsilon \to 0$. Assume V^ϵ has the expansion $V^\epsilon(x,t) = V_0(x,t) + \epsilon V_1(\frac{x}{\epsilon}, t) + O(\epsilon^2)$, where V_1 is the first-order correction term to V_0. Then, by matching powers of ϵ, we find that

$$-\frac{\partial V_0}{\partial t}(\epsilon y, t) + H\left(D_x V_0 + D_y V_1, y\right) = O(\epsilon),$$

where $y = \frac{x}{\epsilon}$. Letting $\epsilon \to 0$ we deduce that V_1 should be a periodic solution of the cell problem

$$H(P + D_y u, y) = \overline{H}(P), \tag{5.11}$$

with $P = D_x V_0$ and $\overline{H}(P) = \frac{\partial V_0}{\partial t}$.

This formal calculations suggest that the viscosity solution V^ϵ converges to a viscosity solution V of

$$-V_t + \overline{H}(D_x V) = 0.$$

5.6.2 Convergence

Motivated by the previous computations we study the convergence of V^ϵ to some function V using viscosity solutions methods. Consider the cell problem (5.11). From theorem 61, we know that for each P there exists a unique function $\overline{H}(P)$ for which the equation

$$H(P + D_x u, x) = \overline{H}(P) \tag{5.12}$$

has a periodic viscosity solution. The function $\overline{H}(P)$ is called the *effective Hamiltonian.*

Theorem 62. *The viscosity solution V^ϵ of the terminal value problem (5.10) converges uniformly to the viscosity solution of*

$$-V_t + \overline{H}\left(D_x V\right) = 0 \tag{5.13}$$

with terminal value $V(x, T) = g(x)$.

Proof. By choosing a suitable subsequence $\epsilon \to 0$ we may assume $V^\epsilon \to V$ uniformly. Now we claim that V is a viscosity solution of (5.13). First we need to prove that if ϕ is a C^1 function such that $V - \phi$ has a strict local maximum at $(\hat{x}, \hat{t})$ then

$$-\phi_t(\hat{x}, \hat{t}) + \overline{H}(D_x\phi(\hat{x}, \hat{t})) \leq 0.$$

Assume this statement is false. Then there exists a maximum point $(\hat{x}, \hat{t})$ of $V - \phi$ and $\theta > 0$ such that

$$-\phi_t(\hat{x}, \hat{t}) + \overline{H}(D_x\phi(\hat{x}, \hat{t})) > \theta. \tag{5.14}$$

Let $u(y)$ be a viscosity solution of

$$H(D_x\phi(\hat{x}, \hat{t}) + D_y u(y), y) = \overline{H}(D_x\phi(\hat{x}, \hat{t})). \tag{5.15}$$

Define

$$\phi^\epsilon(x, t) = \phi(x, t) + \epsilon u(\frac{x}{\epsilon}).$$

We claim that in the viscosity sense

$$-\phi^\epsilon_t(x, t) + \overline{H}(D_x\phi^\epsilon(x, t)) \geq \frac{\theta}{3},$$

in some ball $B((\hat{x}, \hat{t}), r) \subset \mathbb{R}^{n+1}$ with radius $r > 0$, chosen small enough, depending only on the modulus of continuity of $D_x\phi$ and H. Indeed, let ψ be a C^1 function and suppose $\phi^\epsilon - \psi$ has a local minimum at $(x_1, t_1) \in B((\hat{x}, \hat{t}), r)$. Note that since ϕ^ϵ is Lipschitz this implies that $|D_x\psi(x_1, t_1)|$ and $|D_t\psi(x_1, t_1)|$ are bounded by a constant that depends only on the Lipschitz constant of ϕ^ϵ. Observe also that

$$u(\frac{x}{\epsilon}) - \eta(\frac{x}{\epsilon}, \frac{t}{\epsilon}) \geq u(\frac{x_1}{\epsilon}) - \eta(\frac{x_1}{\epsilon}, \frac{t_1}{\epsilon}),$$

where $\eta(x,t) = \frac{1}{\epsilon}\left[\psi(\epsilon x, \epsilon t) - \phi(\epsilon x, \epsilon t)\right]$, for $(x,t) \in B((\hat{x}, \hat{t}), r)$. Thus $u - \eta$ has a local minimum at $(\frac{x_1}{\epsilon}, \frac{t_1}{\epsilon})$. Since u is a viscosity solution of (5.15) then

$$H\left(D_x\phi(\hat{x}, \hat{t}) + D_y\eta\left(\frac{x_1}{\epsilon}, \frac{t_1}{\epsilon}\right), \frac{x_1}{\epsilon}\right) \geq \overline{H}(D_x\phi(\hat{x}, \hat{t})).$$

By adding $-D_t\phi(\hat{x}, \hat{t})$ to both sides and using (5.14) we conclude

$$-D_t\phi(\hat{x}, \hat{t}) + H\left(D_x\phi(\hat{x}, \hat{t}) + D_x\psi(x_1, t_1) - D_x\phi(x_1, t_1), \frac{x_1}{\epsilon}\right) \geq \theta$$

If r is chosen small enough (depending on the modulus of continuity of $D_x\phi$) then

$$-D_t\phi(x_1, t_1) + H(D_x\psi(x_1, t_1), \frac{x_1}{\epsilon}) \geq \frac{\theta}{2}$$

Since u does not depend on t,

$$-D_t\eta(\frac{x_1}{\epsilon}, \frac{t_1}{\epsilon}) = 0,$$

and so $D_t\psi(x_1, t_1) = D_t\phi(x_1, t_1)$. Thus

$$-D_t\psi(x_1, t_1) + H(D_x\psi(x_1, t_1), \frac{x_1}{\epsilon}) \geq \frac{\theta}{2}.$$

By having chosen r even small enough (depending on $|D_tH|$) one has

$$-D_t\psi(x_1, t_1) + H(D_x\psi(x_1, t_1), \frac{x_1}{\epsilon}) \geq \frac{\theta}{3}.$$

Hence ϕ^ϵ is a viscosity supersolution of

$$-D_tV + H(D_xV, \frac{x}{\epsilon}, t) = 0,$$

in $B((\hat{x},\hat{t}),r)$; also V^ϵ is a viscosity subsolution of the same equation. Thus, by the comparison principle,

$$V^\epsilon(\hat{x},\hat{t}) - \phi^\epsilon(\hat{x},\hat{t}) \leq \sup_{\partial B((\hat{x},\hat{t}),r)} (V^\epsilon - \phi^\epsilon)$$

which contradicts the assumption that $V - \phi$ has a local maximum at $(\hat{x},\hat{t})$.

The other part of the proof, when $V - \phi$ has a strict local minimum, is similar. □

5.7 Bibliographical notes

For stationary problems in deterministic control we suggest the references [BCD97], [Lio82] and [Bar94]. Concerning homogenization reader should consult [Con95], [Con97], [Con96] or the book [BD98], in addition to the original paper [LPV88]. New important developments in the homogenization theory were achieve in a series of papers by Lions and Souganidis, among others, see for instance [Sou99], [LS03]. Additional material, written in Portuguese, can be found in [LLF].

6

Stochastic optimal control

The objective of this chapter is to present an introduction to stochastic optimal control, the basic techniques and ideas, as well as some applications.

6.1 The set-up of stochastic optimal control

In this section we discuss the set-up of stochastic optimal control for the finite horizon terminal value problem.

Fix a time interval $[t, T]$. Let $(\Omega, \mathcal{F}, P)$ be a probability space in which it is defined a d-dimensional Brownian motion W_s. Let $\mathcal{F}_s$ be the filtration associated with W_s. For $t \leq s \leq T$ denote $\mathcal{B}_s$ the Borel σ-algebra on $[t, s]$. A mapping $f : \Omega \times [t, T] \to \mathbb{R}^m$ is called progressively measurable if for any $t \leq s \leq T$ the restriction of f to $\Omega \times [t, s]$ is $\mathcal{F}_s \times \mathcal{B}_s$ measurable.

As before, assume we are given a control space U, that is a convex closed subset of $\mathbb{R}^m$. In this chapter we assume further, for simplicity, that the control space U is bounded.

A control on an interval $I \subset \mathbb{R}_0^+$ is a progressively measurable process $\mathbf{u}_t : I \to U$. Let $f(x, u) : \mathbb{R}^n \times U \to \mathbb{R}^n$ and $\sigma(x, u) : \mathbb{R}^n \times U \to \mathbb{R}^{n \times d}$ be continuous functions satisfying the following Lipschitz condition:

$$|f(x, u) - f(y, u)| + |\sigma(x, u) - \sigma(y, u)| \leq K|x - y|,$$

for some suitable (uniform) constant K, and

$$|f(x, u)|, |\sigma(x, u)| \leq C$$

This Lipschitz condition ensures that for each bounded progressively measurable control $\mathbf{u}$ there exists a unique solution to the following stochastic differential equation, the control law,

$$d\mathbf{x} = f(\mathbf{x}, \mathbf{u})dt + \sigma(\mathbf{x}, \mathbf{u})dW_t. \tag{6.1}$$

The boundedness of f and σ, though not essential, simplifies some arguments.

In the case one needs to consider unbounded controls, a convenient condition that ensures the existence of a unique solution to (6.1)

$$E \int_0^T |f(x, \mathbf{u})| + |\sigma(x, \mathbf{u})|^2 dt$$

is uniformly bounded for $x \in \mathbb{R}^n$. This condition will not be necessary here because we assume U to be bounded.

As before, consider a running cost $L(x, u) : \mathbb{R}^n \times U \to \mathbb{R}$ and a terminal cost $\psi(x) : \mathbb{R}^n \to \mathbb{R}$ at time T. For each control $\mathbf{u}$ on (t, T) we define

$$J(x, t; \mathbf{u}) = E \int_t^T L(\mathbf{x}, \mathbf{u}) dt + \psi(\mathbf{x}(T)),$$

where $\mathbf{x}$ solves (6.1) with $\mathbf{x}(t) = x$. We assume further that $L(x, u)$ is uniformly bounded.

The finite horizon terminal value problem consists in determining the control $\mathbf{u}^*$ which minimizes $J(x, t, \mathbf{u})$. We define the value function to be

$$V(x, t) = \inf_{\mathbf{u}} J(x, t; \mathbf{u}). \tag{6.2}$$

Exercise 51 (Stochastic Lax-Hopf formula)**.** *Consider the following controlled dynamics:*

$$d\mathbf{x} = \mathbf{u} dt + \sigma dW_t,$$

where the control space is $U = \mathbb{R}^n$, and, as before, we are considering progressively measurable controls $\mathbf{u}$. Let $L(u)$ be convex and superlinear in u. Let ψ be a terminal cost bounded below. Show that the value function is

$$V(x, t) = \inf_{y \in \mathbb{R}^n} (T - t) L\left(\frac{y - x}{T - t}\right) + E\psi(y + W_{T-t}),$$

where W_s a n-dimensional Brownian motion.

6.2 Verification theorem

Motivated by the verification results in the deterministic optimal control setting, we will now prove a similar verification theorem for the stochastic optimal control problem.

Given two $n \times m$ matrices A and B, respectively, define $A : B \equiv \operatorname{tr} A^T B$. We now define the second order Hamiltonian $\mathcal{H}(M, p, x) : \mathbb{R}^{n\times n} \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ to be:

$$\mathcal{H}(M, p, x) = \sup_{u\in U} -\frac{1}{2}\sigma(x, u)\sigma^T(x, u) : M - f(x, u) \cdot p - L(x, u).$$

If we assume that the diffusion matrix $\sigma(x)$ is independent of u then

$$\mathcal{H}(M, p, x) = -\frac{1}{2}\sigma(x)\sigma^T(x) : M + H(p, x),$$

where H is the generalized Legendre transform of L as in (3.1). As before, in chapter 4, we also assume that the drift is linear on the control, that is

$$f(x, u) = A(x)u + B(x), \tag{6.3}$$

and that the Lagrangian $L(x, u)$ is a strictly convex function of u. In this case there exists a unique value $\mu(x, p)$ for which

$$\mathcal{H}(M, p, x) = -\frac{1}{2}\sigma(x)\sigma^T(x) : M - f(x, \mu(x, p)) \cdot p - L(x, \mu(x, p)).$$

Theorem 63. *Let $L(x, u)$, $f(x, u)$ and $\sigma(x, u)$ be as defined previously. Suppose that L is strictly convex in u. Assume that the diffusion matrix does not depend on u, and that the drift $f(x, u)$ is linear in u as in* (6.3)*. Let H be the generalized Legendre transform* (3.1) *of L. Let $\Phi(x, t)$ be any classical solution to the Hamilton-Jacobi equation*

$$-\Phi_t - \frac{1}{2}\sigma(x)\sigma^T(x) : D^2_{xx}\Phi + H(D_x\Phi, x) = 0 \tag{6.4}$$

on the time interval $[0, T]$, with terminal cost $\Phi(x, T) = \psi(x)$. Then, for all $0 \le t \le T$,

$$\Phi(x, t) = V(x, t),$$

where V is the value function, defined as in (6.2)*. Furthermore $\mathbf{u} = \mu(\mathbf{x}(t), D_x V(\mathbf{x}(t), t))$ is an optimal feedback control.*

REMARK. We should observe that this theorem in particular implies the uniqueness of classical solutions to the Hamilton-Jacobi equation.

Proof. Let $\mathbf{u}$ be a progressively measurable control and $\mathbf{x}$ the corresponding solution to (6.1) with initial condition $\mathbf{x}(t) = x$. Then, by Dynkin's formula

$$
\begin{aligned}
&E\Phi(\mathbf{x}(T), T) - \Phi(\mathbf{x}(t), t) \\
&= E\left(\int_t^T \Phi_t(\mathbf{x}(s), s) + f(\mathbf{x}, \mathbf{u}) D_x \Phi(\mathbf{x}(s), s)\right. \\
&\quad \left. + \frac{1}{2}\sigma(\mathbf{x})\sigma^T(\mathbf{x}) : D^2_{xx}\Phi(\mathbf{x}(s), s) ds\right).
\end{aligned}
$$

Adding $E\left(\int_t^T L(\mathbf{x}(s), \mathbf{u}(s)) ds\right) + \Phi(\mathbf{x}(t), t)$ to the above equality, using the fact that $\Phi(x, T) = \psi(x)$, and taking the infimum over all admissible controls $\mathbf{u}$, we obtain

$$
\begin{aligned}
&\inf E\left(\int_t^T L(\mathbf{x}, \mathbf{u}) ds + \psi(\mathbf{x}(T))\right) \\
&= \Phi(\mathbf{x}(t), t) + \inf E\left(\int_t^T \left[\Phi_t(\mathbf{x}, s) + L(\mathbf{x}, \mathbf{u})\right.\right. \\
&\quad \left.\left. + D_x\Phi(\mathbf{x}, s) \cdot f(\mathbf{x}, \mathbf{u}) + \frac{1}{2}\sigma(\mathbf{x})\sigma^T(\mathbf{x}) : D^2_{xx}\Phi(\mathbf{x}, s)\right] ds\right).
\end{aligned}
$$

Now recall that for any v,

$$
-H(p, x) \leq L(x, v) + p \cdot f(x, v).
$$

Therefore

$$\begin{aligned}
&\inf E\left(\int_t^T L(\mathbf{x},\mathbf{u})ds + \psi\big(\mathbf{x}(T)\big)\right)\\
&\quad\geq \Phi(\mathbf{x}(t),t)\\
&+\inf E\Big(\int_t^T \big(\Phi_t(\mathbf{x},s)+\frac{1}{2}\sigma(\mathbf{x})\sigma^T(\mathbf{x}):D^2\Phi(\mathbf{x},s)\\
&\qquad - H(D_x\Phi(\mathbf{x},s),\mathbf{x})\big)ds\Big)\\
&\quad= \Phi(\mathbf{x}(t),t).
\end{aligned}$$

Let $r(x,t)$ be uniquely defined as

$$r(x,t)\in \operatorname{argmin}_{v\in U} L(x,v)+D_x\Phi(x,t)\cdot f(x,v). \tag{6.5}$$

Define the progressively measurable control $\mathbf{u}=r(\mathbf{x},s)$. Consider the trajectory $\mathbf{x}$ given by solving the stochastic differential equation

$$d\mathbf{x}=f(\mathbf{x},\mathbf{u})ds+\sigma(\mathbf{x})dW_s,$$

with initial condition $\mathbf{x}(t)=x$. Then

$$\begin{aligned}
&\inf E\Big(\int_t^T L(\mathbf{x},\mathbf{u})ds+\psi\big(\mathbf{x}(T)\big)\Big)\\
&\quad\leq \Phi(\mathbf{x}(t),t)+E\left(\int_t^T \big[\Phi_s(\mathbf{x},s)+L(\mathbf{x},\mathbf{u})\right.\\
&\qquad\left.+D_x\Phi(\mathbf{x},s)\cdot f(\mathbf{x},\mathbf{u})+\frac{1}{2}\sigma(\mathbf{x})\sigma^T(\mathbf{x}):D^2_{xx}\Phi(\mathbf{x},s)\big]ds\right)\\
&\quad=\Phi(\mathbf{x}(t),t)\\
&+\inf E\Big(\int_t^T \big(\Phi_s(\mathbf{x},s)+\frac{1}{2}\sigma(\mathbf{x})\sigma^T(\mathbf{x}):D^2_{xx}\Phi(\mathbf{x},s)\\
&\qquad - H(D_x\Phi(\mathbf{x},s),\mathbf{x})\big)ds\Big)\\
&\quad=\Phi(\mathbf{x}(t),t).
\end{aligned}$$

which ends the proof. □

We should observe from the proof that (6.5) gives an optimal feedback law for the optimal control, provided we can find a solution to the Hamilton-Jacobi equation.

6.3 Continuity

We now prove that the value function is Lipschitz continuous. We start by proving some estimates concerning diffusions:

Lemma 64. *Let $f(x,t) : \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ and $\sigma(x,t)\mathbb{R}^n \times [0,T] \to \mathbb{R}^{n\times d}$ be (globally) Lipschitz in x. Let $\mathbf{x}_i$, $i = 1,2$, be solutions to the stochastic differential equation*

$$d\mathbf{x}_i = f(\mathbf{x}_i,t)dt + \sigma(\mathbf{x}_i,t)dW_t.$$

Then

$$E[|\mathbf{x}_1(t) - \mathbf{x}_2(t)|^2] \leq C|\mathbf{x}_1(0) - \mathbf{x}_2(0)|^2,$$

for all $0 \leq t \leq T$.

Proof. It suffices to compute

$$\begin{aligned}\frac{d}{dt}E|\mathbf{x}_1 - \mathbf{x}_2|^2 =& 2E[(\mathbf{x}_1 - \mathbf{x}_2)(f(\mathbf{x}_1,t) - f(\mathbf{x}_2,t))] \\ &+ E[\mathrm{tr}(\sigma(\mathbf{x}_1,t) - \sigma(\mathbf{x} - 2,t))(\sigma(\mathbf{x}_1,t) - \sigma(\mathbf{x}_2,t))^T] \\ \leq& CE[|\mathbf{x}_1 - \mathbf{x}_2|^2],\end{aligned}$$

and apply Gronwal's inequality. □

Theorem 65. *Assume that*

$$|L(x,u) - L(y,u)| \leq C|x - y|,$$

and

$$|\psi(x) - \psi(y)| \leq C|x - y|.$$

Then the value function is Lipschitz continuous both in x.

Proof. Let x and y be arbitrary points, and $t < T$. Fix $\epsilon > 0$. Let $\mathbf{u}^\epsilon$ be an almost optimal control for the point x, that is:

$$V(x,t) + \epsilon \geq E\left(\int_t^T L(\mathbf{x}^\epsilon, \mathbf{u}^\epsilon)ds + \psi(\mathbf{x}^\epsilon(T))\right).$$

Let $\mathbf{y}^\epsilon$ be a solution to

$$d\mathbf{y}^\epsilon = f(\mathbf{y}^\epsilon, \mathbf{u}^\epsilon)dt + \sigma(\mathbf{y}^\epsilon, \mathbf{u}^\epsilon)dW_t$$

with $\mathbf{y}^\epsilon(t) = y$. Then

$$\begin{aligned} V(y,t) - V(x,t) \leq& \epsilon + E\int_t^T (L(\mathbf{y}^\epsilon, \mathbf{u}^\epsilon) - L(\mathbf{x}^\epsilon, \mathbf{u}^\epsilon))ds \\ &+ \psi(\mathbf{y}^\epsilon(T)) - \psi(\mathbf{x}^\epsilon(T)) \\ &\leq E\int_t^T C|\mathbf{y}^\epsilon(s) - \mathbf{x}^\epsilon(s)| + C|\mathbf{y}^\epsilon(T) - \mathbf{x}^\epsilon(T)| \\ &\leq C|x - y|, \end{aligned}$$

using Cauchy-Schwartz inequality and lemma 64. □

6.4 Stochastic dynamic programming

As in deterministic control, the value function of a stochastic optimal control terminal value problem satisfies a semigroup property called the stochastic dynamic programming principle.

Theorem 66 (Dynamic programming principle)**.** *Suppose that* $t \leq t' \leq T$*. Then*

$$V(x,t) = \inf_{\mathbf{u}} E\left[\int_t^{t'} L(\mathbf{x}(s),\mathbf{u}(s))ds + V(\mathbf{x}(t'),t')\right], \tag{6.6}$$

where $\mathbf{x}(t) = x$ *and* $d\mathbf{x} = f(\mathbf{x},\mathbf{u})dt + \sigma(\mathbf{x},\mathbf{u})dW_t$*.*

REMARK. We will only sketch the proof of this theorem, a detailed proof can be found in [FS06]. We should also observe that the Dynamic programming principle also holds if instead of a fixed time t' one chooses a stopping time τ.

Proof. Denote by $\tilde{V}(x,t)$ the right hand side of (6.6). For fixed $\epsilon > 0$, let $\mathbf{u}^\epsilon$ be an almost optimal control for $V(x,t)$. Let $\mathbf{x}^\epsilon(s)$ be the corresponding trajectory trajectory, i.e., assume that

$$J(x,t;\mathbf{u}^\epsilon) \leq V(x,t) + \epsilon.$$

We claim that $\tilde{V}(x,t) \leq V(x,t) + \epsilon$. To check this, let $y = \mathbf{x}^\epsilon(t')$. Then

$$\tilde{V}(x,t) \leq \int_t^{t'} L(\mathbf{x}^\epsilon(s),\mathbf{u}^\epsilon(s))ds + V(y,t').$$

Additionally,

$$V(y,t') \leq J(y,t';\mathbf{u}^\epsilon).$$

Therefore

$$\tilde{V}(x,t) \leq J(x,t;\mathbf{u}^\epsilon) \leq V(x,t) + \epsilon,$$

and, since ϵ is arbitrary, $\tilde{V}(x,t) \leq V(x,t)$.

To prove the opposite inequality, we will proceed by contradiction. Therefore, if $\tilde{V}(x,t) < V(x,t)$, we could choose $\epsilon > 0$ and a control $\mathbf{u}^\sharp$ such that

$$E\int_t^{t'} L(\mathbf{x}^\sharp(s),\mathbf{u}^\sharp(s))ds + V(\mathbf{x}^\sharp(t'),t') < V(x,t) - \epsilon,$$

where $\dot{\mathbf{x}}^\sharp = f(\mathbf{x}^\sharp, \mathbf{u}^\sharp)$, $\mathbf{x}^\sharp(t) = x$. Choose $\mathbf{u}^\flat$ such that

$$J(\mathbf{x}^\sharp(t'), t'; \mathbf{u}^\flat) \leq V(\mathbf{x}^\sharp(t'), t') + \frac{\epsilon}{2}$$

Define $\mathbf{u}^\star$ as

$$\begin{cases} \mathbf{u}^\star(s) = \mathbf{u}^\sharp(s) \text{ for } s < t' \\ \mathbf{u}^\star(s) = \mathbf{u}^\flat(s) \text{ for } t' < s. \end{cases}$$

So, we would have

$$\begin{aligned} V(x,t) - \epsilon > E \int_t^{t'} L(\mathbf{x}^\sharp(s), \mathbf{u}^\sharp(s))ds + V(\mathbf{x}^\sharp(t'), t') \geq \\ \geq E \int_t^{t'} L(\mathbf{x}^\sharp(s), \mathbf{u}^\sharp(s))ds + J(\mathbf{x}^\sharp(t'), t'; \mathbf{u}^\flat) - \frac{\epsilon}{2} = \\ = J(x, t; \mathbf{u}^\star) - \frac{\epsilon}{2} \geq V(x,t) - \frac{\epsilon}{2}, \end{aligned}$$

which is a contradiction. □

6.5 The Hamilton-Jacobi equation

In this section we will establish that if the value function is smooth then it satisfies the Hamilton-Jacobi equation.

Theorem 67. *Let V be the value function to the terminal value problem. Suppose V is C^2. Then it solves the Hamilton-Jacobi equation*

$$-V_t + \mathcal{H}(D^2_{xx}V, D_xV, x) = 0.$$

Proof. Fix any constant control u^*. Then, by the dynamic programming principle

$$V(x,t) \leq E \int_t^{t+h} L(\mathbf{x}(s), u^*) + V(\mathbf{x}(t+h), t+h).$$

By using Itô's formula and dividing by h, as $h \to 0$ we obtain

$$\begin{aligned}0 &\le V_t + L(x, u^*) + f(x, u^*) \cdot D_x V + \frac{1}{2}\sigma\sigma^T D^2_{xx} V \\ &\le V_t - \mathcal{H}(D^2_{xx}V, D_x V, x),\end{aligned}$$

that is

$$-V_t + \mathcal{H}(D^2_{xx}V, D_x V, x) \le 0.$$

Suppose now that in fact the previous inequality were strict at a point (x_0, t_0), that is

$$-V_t(x_0, t_0) + \mathcal{H}(D^2_{xx}V(x_0, t_0), D_x V(x_0, t_0), x_0) = -\delta < 0.$$

Then in a neighborhood N of (x_0, t_0) we have

$$-V_t(x, t) + \mathcal{H}(D^2_{xx}V(x, t), D_x V(x, t), x) < -\frac{\delta}{2}.$$

Let $\mathbf{u}^*$ be an optimal control, and let τ be the exit time of N of the corresponding trajectory. Then, by Dynkin's formula,

$$E(V(\mathbf{x}(\tau), \tau)) - V(x_0, t_0) = E\int_{t_0}^{\tau} V_t + f \cdot D_x V + \frac{1}{2}\sigma\sigma^T : D^2_{xx} V dt,$$

and, by the stochastic dynamic programming principle applied to the stopping time τ,

$$V(x_0, t_0) = E\left(\int_{t_0}^{\tau} L(\mathbf{x}, \mathbf{u}^*)dt + V(\mathbf{x}(\tau), \tau)\right).$$

Thus

$$\begin{aligned}0 &= E\int_{t_0}^{\tau} L(\mathbf{x}, \mathbf{u}^*) + V_t + f \cdot D_x V + \frac{1}{2}\sigma\sigma^T : D^2_{xx} V dt \\ &\ge E\int_{t_0}^{\tau} V_t - \mathcal{H}(D^2_{xx}V(\mathbf{x}, t), D_x V(\mathbf{x}, t), \mathbf{x}) \ge \frac{\delta}{2} E\int_{t_0}^{\tau} dt,\end{aligned}$$

which is a contradiction. □

6.6 Viscosity Solutions

As before, we recall that a bounded uniformly continuous function $V(x,t)$ is a viscosity solution to the Euler-Lagrange equation

$$-V_t + \mathcal{H}(D^2V, DV, x) = 0$$

if for any smooth function $\psi(x,t)$ and any point $(x_0,t_0) \in \operatorname{argmax} V - \psi$ we have

$$-\psi_t(x_0,t_0) + \mathcal{H}(D^2\psi(x_0,t_0), D\psi(x_0,t_0), x_0) \leq 0,$$

with the opposite inequality for points in the argmin.

Theorem 68. *The value function is a viscosity solution to*

$$-V_t + \mathcal{H}(D^2V, DV, x) = 0.$$

Proof. Let $\varphi(x,t)$ be a smooth function and let $(x_0,t_0) \in \operatorname{argmin} V - \varphi$. Without loss of generality we may assume $V(x_0,t_0) = \varphi(x_0,t_0)$, and so $V(x,t) \geq \varphi(x,t)$ for all (x,t).

Then, by the dynamic programming principle

$$\varphi(x_0,t_0) = V(x_0,t_0) = E\int_{t_0}^{t_0+h} L(\mathbf{x},\mathbf{u})dt + V(\mathbf{x}(t_0+h), t_0+h)$$
$$\geq E\int_{t_0}^{t_0+h} L(\mathbf{x},\mathbf{u})dt + \varphi(\mathbf{x}(t_0+h), t_0+h).$$

Using Dynkin's formula we then conclude that

$$0 \geq E\int_{t_0}^{t_0+h} L(\mathbf{x},\mathbf{u}) + \varphi_t + fD_x\varphi + \frac{1}{2}\sigma\sigma^T : D^2_{xx}\varphi dt$$
$$\geq E\int_{t_0}^{t_0+h} \varphi_t - \mathcal{H}(D^2\varphi, D\varphi, x).$$

by sending $h \to 0$ we conclude that

$$-\varphi_t(x_0,t_0) + \mathcal{H}(D^2_{xx}\varphi(x_0,t_0), D_x\varphi(x_0,t_0), x_0) \geq 0.$$

To obtain the second inequality, suppose $(x_0,t_0) \in \operatorname{argmax} V - \varphi$. As before, without loss of generality we may assume $V(x_0,t_0) = \varphi(x_0,t_0)$, and so $V(x,t) \leq \varphi(x,t)$ for all (x,t).

Then, by the dynamic programming principle

$$\begin{aligned}\varphi(x_0,t_0) = V(x_0,t_0) &\leq E\int_{t_0}^{t_0+h} L(\mathbf{x},\mathbf{u}^*)dt + V(\mathbf{x}(t_0+h), t_0+h)\\ &\leq E\int_{t_0}^{t_0+h} L(\mathbf{x},\mathbf{u}^*)dt + \varphi(\mathbf{x}(t_0+h), t_0+h),\end{aligned}$$

where $\mathbf{u}^*$ is a constant control. This then implies, by sending $h \to 0$,

$$0 \leq L(x_0,\mathbf{u}^*) + \varphi_t + fD_x\varphi + \frac{1}{2}\sigma\sigma^T : D^2_{xx}\varphi dt,$$

and so

$$-\varphi_t(x_0,t_0) + \mathcal{H}(D^2_{xx}\varphi(x_0,t_0), D_x\varphi(x_0,t_0), x_0) \leq 0.$$

□

6.7 Applications to Financial Mathematics

In this last section we present an application of stochastic optimal control, namely Merton's optimal investment problem. To model an investment problem we consider two types of assets: a bond, with a constant continuously compounded interest rate r, whose price evolution satisfies

$$dp = rpdt,$$

and a stock whose time evolution is modeled through a geometric Brownian motion with drift

$$dS = S(\mu dt + \sigma dW_t),$$

where μ and σ are constant parameters.

At any moment a large investor is allowed to have a fraction π of his wealth invested in stocks and a corresponding fraction $1 - \pi$ in bonds. We should note that because short selling of both stocks and bonds is allowed in fact we have $-\infty < \pi < +\infty$ rather than $0 \leq \pi \leq 1$.

The wealth process for this investor is simply

$$d\mathbf{x} = \mathbf{x}\left[(r + (\mu - r)\pi)dt + \sigma\pi dW_t\right].$$

Let U be a concave function, which for definiteness we take $U(x) = x^\gamma$, for $0 < \gamma < 1$. The objective of the investor is to allocate its portfolio so that to maximize

$$V(x,t) = \sup_\pi EU(\mathbf{x}(T)),$$

for some fixed terminal time T, where the supremum is taken over progressively measurable processes π satisfying $E\int_t^T |\pi|^2 < \infty$, and $\mathbf{x}(t) = x$.

From the results discussed before in this chapter, the function V is a viscosity solution to the following Hamilton-Jacobi equation

$$V_t + \mathcal{H}(D^2_{xx}V, D_xV, x) = 0,$$

where

$$\mathcal{H}(M,p,x) = \sup_{\pi\in\mathbb{R}} \left[x(r+(\mu-r)\pi)p + \frac{1}{2}\sigma^2\pi^2x^2M \right].$$

Note that because we are dealing with a maximization a few signs had to be exchanged.

By the homogeneity of the problem, it is easy to check that

$$V(x,t) = x^\gamma V(1,t) \equiv x^\gamma h(t).$$

Then a simple computation yields that h solves

$$0 = h' + \gamma h \sup_{\pi} \left[\gamma(r+(\mu-r)\pi) + \frac{1}{2}\sigma^2\pi^2\gamma(\gamma-1) \right].$$

This yields $h(t) = e^{a(T-t)}$ for some constant a and we also conclude that

$$\pi = \frac{\mu - r}{(1-\gamma)\sigma^2}$$

is the optimal control, which is constant in time.

6.8 Bibliographical notes

Two main references on stochastic optimal control are the books [FR75] and [FS06]. Additional material can also be found in the nice lecture notes [Tou].

7

Differential Games

This chapter is a brief introduction to deterministic differential games and its connection with viscosity solutions of Hamilton-Jacobi equations.

7.1 Dynamic programming principle

Consider a problem where two players have conflicting objectives. Each of them partially controls a dynamical system, and one of the players wants to maximize a pay-off functional, whereas the other one wishes to minimize the same pay-off functional. To set-up this problem, let U^+ and U^- be two convex closed subsets of, respectively, $\mathbb{R}^{m_+}$ and $\mathbb{R}^{m_-}$. The $+$ sign stands for the controls or variables available for the maximizing player, whereas the $-$ sign corresponds to the minimizing player.

Consider a differential equation

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}^+, \mathbf{u}^-), \tag{7.1}$$

where $\mathbf{u}^\pm$ are controls for the two players taking values on $U^\pm$. To simplify, we suppose that $U^\pm$ are compact sets, that f is globally bounded and satisfies the Lipschitz estimate

$$|f(x, u^+, u^-) - f(y, u^+, u^-)| \leq C|x - y|.$$

Let T be a terminal time. To each pair of controls $(\mathbf{u}^+, \mathbf{u}^-)$ on (t, T), consider the corresponding solution to (7.1) with initial condition $\mathbf{x}(t) = x$. We are given a running cost $L(x, \mathbf{u}^+, \mathbf{u}^-)$ and a terminal cost $\psi(x)$. Associated to the controls and these costs we define the cost

$$J[x, t; \mathbf{u}^+, \mathbf{u}^-] = \int_t^T L(\mathbf{x}, \mathbf{u}^+, \mathbf{u}^-) ds + \psi(\mathbf{x}(T)),$$

where $\mathbf{x}$ solves (7.1) with the initial condition $\mathbf{x}(t) = x$. The objective of the $+$ player is to maximize this cost, whereas the $-$ player wishes to minimize this cost. Of course the players are not allowed to foresee the future and we must therefore discuss the appropriate strategies.

Denote by $\mathcal{U}^\pm([t, T])$ the set of all mappings from $[t, T]$ into $U^\pm$. A non-anticipating strategy $\mu^\pm$ is a mapping

$$\mu^\pm : \mathcal{U}^\mp([t, T]) \to \mathcal{U}^\pm([t, T])$$

such that for any $\mathbf{u}^\mp, \tilde{\mathbf{u}}^\mp \in \mathcal{U}^\mp([t, T])$ and any $t < s < T$ such that, for all $t \leq \tau \leq s$,

$$\mathbf{u}^\mp(\tau) = \tilde{\mathbf{u}}^\mp(\tau)$$

we have

$$\mu^\pm(\mathbf{u}^\mp)(\tau) = \mu^\pm(\tilde{\mathbf{u}}^\mp)(\tau),$$

for all $t \leq \tau \leq s$. Denote by $\Lambda^\pm$ the set of all non-anticipating strategies.

The upper V^+ value functions are defined to be

$$V^+(x,t) = \sup_{\mu^+ \in \Lambda^+([t,T])} \inf_{u^- \in \mathcal{U}^-([t,T])} J(x,t;\mu^+(\mathbf{u}^-),\mathbf{u}^-),$$

whereas the lower value function is

$$V^-(x,t) = \inf_{\mu^- \in \Lambda^-([t,T])} \sup_{u^+ \in \mathcal{U}^+([t,T])} J(x,t;\mathbf{u}^+,\mu^-(\mathbf{u}^+)).$$

Theorem 69 (Dynamic programming principle)**.** *For any $t' < T$ we have*

$$V^+(x,t) = \sup_{\mu^+ \in \Lambda^\pm([t,t'])} \inf_{u^- \in \mathcal{U}^-([t,t'])} \int_t^{t'} L(\mathbf{x},\mu^+(\mathbf{u}^-),\mathbf{u}^-)ds + V^+(\mathbf{x}(t'),t').$$

Note that a similar result holds for the lower value, with a identical proof.

Proof. Define

$$\tilde{V}(x,t) = \sup_{\mu^+ \in \Lambda^+([t,t'])} \inf_{u^- \in \mathcal{U}^-([t,t'])} \int_t^{t'} L(\mathbf{x},\mu^+(\mathbf{u}^-),\mathbf{u}^-)ds + V^+(\mathbf{x}(t'),t').$$

Fix $\epsilon > 0$ and choose $\mu_\epsilon^+ \in \Lambda^\pm([t,t'])$ so that

$$\tilde{V}(x,t) \leq \inf_{u^- \in \mathcal{U}^-([t,t'])} \int_t^{t'} L(\mathbf{x},\mu_\epsilon^+(\mathbf{u}^-),\mathbf{u}^-)ds + V^+(\mathbf{x}(t'),t') + \epsilon.$$

Choose now $\tilde{\mu}_\epsilon^+ \in \Lambda^+([t',T])$ so that

$$V(\mathbf{x}(t'),t') \leq \inf_{u^- \in \mathcal{U}^-([t',T])} \int_t^{t'} L(\mathbf{x},\tilde{\mu}_\epsilon^+(\mathbf{u}^-),\mathbf{u}^-)ds + \psi(\mathbf{x}(T)) + \epsilon.$$

By considering the concatenation of the non-anticipating strategies μ_ϵ^+ and $\tilde{\mu}_\epsilon^+$ we obtain a non-anticipating strategy $\bar{\mu}_\epsilon^+$ such that

$$\begin{aligned}\tilde{V}(x,t) \leq &\inf_{u^- \in \mathcal{U}^-([t,T])} \int_t^T L(\mathbf{x}, \bar{\mu}_\epsilon^+(\mathbf{u}^-), \mathbf{u}^-)ds + \psi(\mathbf{x}(T)) + 2\epsilon \\ &\leq V^+(x,t) + 2\epsilon.\end{aligned}$$

Sending $\epsilon \to 0$ we obtain $\tilde{V} \leq V^+$.

To obtain the opposite inequality, fix again $\epsilon > 0$ and choose a non-anticipating strategy $\bar{\mu}_\epsilon^+$ so that

$$V^+(x,t) \leq \inf_{u^- \in \mathcal{U}^-([t,T])} \int_t^T L(\mathbf{x}, \bar{\mu}_\epsilon^+(\mathbf{u}^-), \mathbf{u}^-)ds + \psi(\mathbf{x}(T)) + \epsilon.$$

Note that

$$\inf_{u^- \in \mathcal{U}^-([t',T])} \int_{t'}^T L(\mathbf{x}, \bar{\mu}_\epsilon^+(\mathbf{u}^-), \mathbf{u}^-)ds + \psi(\mathbf{x}(T)) \leq V^+(\mathbf{x}(t'), t').$$

Therefore

$$\begin{aligned}V^+(x,t) \leq &\inf_{u^- \in \mathcal{U}^-([t,T])} \int_t^{t'} L(\mathbf{x}, \bar{\mu}_\epsilon^+(\mathbf{u}^-), \mathbf{u}^-)ds + V^+(\mathbf{x}(t'), t') + \epsilon \\ &\leq \tilde{V}(x,t) + \epsilon.\end{aligned}$$

□

7.2 Viscosity solutions

We define the upper and lower Hamiltonians to be, respectively

$$H^+(p,x) = \sup_{u^+ \in U^+} \inf_{u^- \in U^-} -p \cdot f(u^+, u^-, x) - L(x, u^+, u^-),$$

and

$$H^-(p,x) = \inf_{u^- \in U^-} \sup_{u^+ \in U^+} -p \cdot f(u^+, u^-, x) - L(x, u^+, u^-).$$

In general $H^- \geq H^+$ and the inequality may not be strict.

Before stating and proving the main result of this section, we will prove two auxiliary results.

Lemma 70. *Suppose φ satisfies*

$$-\varphi_t + H^+(D_x\varphi, x) \leq -\theta,$$

at a point (x,t) and for some $\theta > 0$. Then, for all h sufficiently small there exists $\mu^+ \in \Lambda^+([t, t+h])$ such that for all $u^- \in \mathcal{U}^-([t, t+h]$ we have

$$\int_t^{t+h} [L(\mathbf{x}, \mu^+(\mathbf{u}^-), \mathbf{u}^-) + f(\mathbf{x}, \mu^+(\mathbf{u}^-), \mathbf{u}^-) D_x\varphi(\mathbf{x}(s), s)$$
$$+\varphi_t(\mathbf{x}(s), s)]\, ds \geq h\frac{\theta}{2}.$$

Proof. For the proof of this lemma, consult [BCD97]. □

Lemma 71. *Suppose φ satisfies*

$$-\varphi_t + H^+(D_x\varphi, x) \geq \theta,$$

at a point (x,t) and for some $\theta > 0$. Then, for all h sufficiently small and any $\mu^+ \in \Lambda^+([t, t+h])$ there exists $u^- \in \mathcal{U}^-([t, t+h]$ such that

$$\int_t^{t+h} [L(\mathbf{x}, \mu^+(\mathbf{u}^-), \mathbf{u}^-) + f(\mathbf{x}, \mu^+(\mathbf{u}^-), \mathbf{u}^-) D_x\varphi(\mathbf{x}(s), s)$$
$$+\varphi_t(\mathbf{x}(s), s)]\, ds \leq -h\frac{\theta}{2}.$$

Proof. For the proof of this lemma, consult [BCD97]. □

We should observe that analogous results for H^- to lemmas 70 and refaxl2 can be established in exactly the same way.

Theorem 72. *The upper and lower values are viscosity solutions to the Isaacs-Bellman-Hamilton-Jacobi equation*

$$-V_t^{\pm} + H(D_x V^{\pm}, x) = 0,$$

with the terminal value $V(x, T) = \psi(x)$.

Proof. We will do the proof for the upper value V^+ as the case of the lower value is similar. Suppose $V^+ - \varphi$ has a strict local maximum at (x_0, t_0) but, by contradiction, there exists $\theta > 0$ such that

$$-\varphi_t^+ + H^+(D_x\varphi^+, x) \geq \theta.$$

Using the dynamic programming principle and the local maximum property, we have

$$\sup_{\mu^+ \in \Lambda^{\pm}([t_0, t_0+h])} \inf_{u^- \in \mathcal{U}^-([t_0, t_0+h])} \int_{t_0}^{t_0+h} L(\mathbf{x}, \mu^+(\mathbf{u}^-), \mathbf{u}^-) ds + \varphi(\mathbf{x}(t_0+h), t_0+h) - \varphi(x_0, t_0) \geq 0$$

This then contradicts lemma 71.

Now suppose $V^+ - \varphi$ has a strict local minimum at (x_0, t_0) but, by contradiction, there exists $\theta > 0$ such that

$$-\varphi_t^+ + H^+(D_x\varphi^+, x) \leq -\theta.$$

Using the dynamic programming principle and the local maximum property, we have

$$\sup_{\mu^+ \in \Lambda^{\pm}([t_0, t_0+h])} \inf_{u^- \in \mathcal{U}^-([t_0, t_0+h])} \int_{t_0}^{t_0+h} L(\mathbf{x}, \mu^+(\mathbf{u}^-), \mathbf{u}^-) ds + \varphi(\mathbf{x}(t_0+h), t_0+h) - \varphi(x_0, t_0) \leq 0$$

This then contradicts lemma 70. □

7.3 Bibliographical notes

The main reference for this chapter is the book [BCD97]. The reader may also want to consult [FS06] (the second edition of the book) for additional material.

8

Aubry Mather theory

This chapter is dedicated to the study of duality theory, relaxation of optimal control problems and applications to viscosity solutions and Aubry-Mather theory.

8.1 Model problems

In this section we discuss certain minimization problems which involve linear objective functions under linear constraints, that is, infinite dimensional linear programming problems. Surprisingly, there are deep relations between these problems and certain nonlinear partial differential equations.

8.1.1 Mather problem

Let $\mathbb{T}^d$ be the d-dimensional standard torus identified whenever convenient with $\mathbb{R}^d/\mathbb{Z}^d$. Consider a Lagrangian $L(x,v)$, $L : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$, smooth in both variables, strictly convex and superlinear in the velocity v. As discussed in chapter 2, the minimal action principle of classical mechanics asserts that the trajectories $\mathbf{x}(t)$ of mechanical systems are critical points or minimizers of the action

$$\int_0^T L(\mathbf{x}, \dot{\mathbf{x}}) ds. \tag{8.1}$$

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

$$\frac{d}{dt} D_v L(\mathbf{x}, \dot{\mathbf{x}}) - D_x L(\mathbf{x}, \dot{\mathbf{x}}) = 0. \tag{8.2}$$

A very important technique in calculus of variations is the relaxation method, which consists in enlarging the class of solutions so that existence of solutions is almost trivial. Of course there is then the problem of establishing that a relaxed solution somehow corresponds to a solution to the original problem. Mather's problem is a relaxed version of the minimal action principle of classical mechanics and consists in minimizing the action

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} L(x,v) d\mu(x,v) \tag{8.3}$$

among a suitable class of probability measures $\mu(x,v)$. Originally, in [Mat91], this minimization was performed over all measures invariant under the Euler-Lagrange equations (8.2). However, as realized by [Mn96], it is more convenient to consider a larger class of measures, the holonomic measures. It turns out that both problems are equivalent as any holonomic minimizing measure is automatically invariant under the Euler-Lagrange equations. In what follows, we will define this class of measures and provide the motivation for it.

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

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

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

$$\int \psi d\mu_{\mathbf{x}} = \lim_{T \to \infty} \int \psi d\mu_{\mathbf{x}}^T,$$

where the limit is taken through an appropriate subsequence.

Let $\gamma(v)$ be a continuous function, $\gamma : \mathbb{R}^d \to \mathbb{R}$, such that

$$\inf \frac{\gamma(v)}{1 + |v|} > 0,$$

and $\lim_{|v| \to \infty} \frac{\gamma(v)}{1+|v|} = \infty$. A measure μ in $\mathbb{T}^d \times \mathbb{R}^d$ is admissible if

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \gamma(v) d\mu < \infty.$$

An admissible measure μ on $\mathbb{T}^d \times \mathbb{R}^d$ is called holonomic if for all $\varphi \in C^1(\mathbb{T}^d)$ we have

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} v \cdot D\varphi d\mu = 0. \tag{8.4}$$

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

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

therefore $\mu_{\mathbf{x}}$ is holonomic.

Mather's problem consists in minimizing (8.3) over all probability measures that satisfy (8.4). As pointed out before, however, this problem was introduced by Mañe in [Mn96] in his study of Mather's original problem [Mat91].

8.1.2 Stochastic Mather problem

In the framework of stochastic optimal control one is led to replace deterministic trajectories by stochastic processes. Suppose that $\mathbf{x}(t)$ is a stochastic process satisfying the stochastic differential equation

$$d\mathbf{x} = \nu dt + \sigma dW,$$

in which ν is a bounded, progressively measurable process, $\sigma > 0$ and W a $n-$dimensional Brownian motion. One would like to minimize the average action

$$\frac{1}{T} E \int_0^T L(\mathbf{x}, \nu) dt.$$

As before, one can associate to these stochastic processes, probability measures μ in $\mathbb{T}^n \times \mathbb{R}^n$ defined as

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \phi(x, v) d\mu = \lim_{T \to \infty} \frac{1}{T} \int_0^T \phi(\mathbf{x}(t), \nu(t)) dt,$$

in which the limit is taken through an appropriate subsequence.

The analog for stochastic processes to the fundamental theorem of calculus is Dynkin's formula. This formula applied to $\varphi(\mathbf{x}(t))$, states that

$$E\left[\varphi(\mathbf{x}(T)) - \varphi(x)\right] = E \int_0^T \nu D_x \varphi(\mathbf{x}(t)) + \frac{\sigma^2}{2} \Delta \varphi(\mathbf{x}(t)) dt.$$

This identity implies

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} vD_x\varphi(x) + \frac{\sigma^2}{2}\Delta\varphi(x)d\mu = 0,$$

for all $\varphi(x) : \mathbb{T}^n \to \mathbb{R}$, C^2.

The *stochastic Mather problem* [Gom02] consists in minimizing

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

over all probability measures μ on $\mathbb{T}^n \times \mathbb{R}^n$ that satisfy the stochastic holonomy constraint

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} vD_x\varphi(x) + \frac{\sigma^2}{2}\Delta\varphi(x)d\mu = 0,$$

for all $\varphi(x) : \mathbb{T}^n \to \mathbb{R}$ of class C^2.

8.1.3 Discrete Mather problem

Also interesting is the discrete case, in which the trajectories are replaced by sequences (x_n, v_n) that satisfy $x_{n+1} = x_n + v_n$. In this case, if the sequence v_n is globally bounded, for instance, we can construct a measure μ in $\mathbb{T}^n \times \mathbb{R}^n$ through

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \phi(x, v)d\mu = \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \phi(x_n, v_n),$$

in which the limit is take through an appropriate subsequence.

For any continuous functions $\varphi : \mathbb{T}^n \to \mathbb{R}$ we have

$$\sum_{n=1}^{N} \varphi(x_n + v_n) - \varphi(x_n) = \varphi(x_{N+1}) - \varphi(x_1).$$

Thus

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \left[\varphi(x+v)-\varphi(x)\right] d\mu = 0.$$

Therefore, we define *Mather's discrete problem*, which consists in minimizing

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

over all probability measures μ in $\mathbb{T}^n\times\mathbb{R}^n$ that satisfy the discrete holonomy constraint:

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \left[\varphi(x+v)-\varphi(x)\right] d\mu = 0,$$

for all continuous function $\varphi:\mathbb{T}^n\to\mathbb{R}$.

8.1.4 Generalized Mather problem

Let $U\subset\mathbb{R}^m$ be a non-empty closed convex set. Assume that, for some $k\geq 0$ (usually $k=0,1,2$) there exists a linear operator $A^v: C^k(\mathbb{T}^n)\to C(\mathbb{T}^n\times U)$, which satisfies the following two conditions: the first one is that for each fixed $\varphi\in C^k(\mathbb{T}^n)$ we have

$$|A^v\varphi|\leq C_\varphi(1+|v|),$$

uniformly in $\mathbb{T}^n\times U$, which of course, if U is bounded means simply that $|A^v\varphi|$ is bounded; the second condition is that for $\varphi\in C^k(\mathbb{T}^n)$ the mapping $(x,v)\mapsto A^v\varphi$ is continuous in $\mathbb{T}^n\times U$.

We assume that there exists another operator B defined in $C^k(\mathbb{T}^n)$ which satisfies the following compatibility conditions with A^v:

$$A^v\kappa = B\kappa, \tag{8.5}$$

for any $\kappa \in \mathbb{R}$, and that, for any given probability measure ν on $\mathbb{T}^n$, there exists a probability measure μ_ν in $\mathbb{T}^n \times U$ such that

$$\int_{\mathbb{T}^n \times U} A^v \varphi d\mu_\nu = \int_{\mathbb{T}^n} B\varphi d\nu, \tag{8.6}$$

for all $\varphi \in C^k(\mathbb{T}^n)$.

The Lagrangian $L(x, v) : \mathbb{T}^n \times U \to \mathbb{R}$ is continuous and convex in v, bounded below, and, either U is bounded, and no further hypothesis are required, or if U is unbounded we assume the superlinear growth condition in v, that is, uniformly in x

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

The generalized Mather problem consists in minimizing

$$\int_{\mathbb{T}^n \times U} L(x, v) d\mu, \tag{8.7}$$

over all probability measures μ in $\mathbb{T}^n \times U$ that satisfy the constraint

$$\int_{\mathbb{T}^n \times U} A^v \varphi d\mu = \int_{\mathbb{T}^n} B\varphi d\nu, \tag{8.8}$$

for all functions $\varphi : \mathbb{T}^n \to \mathbb{R}$ with appropriate regularity.

8.2 Some informal computations

In Mather's problem, both in the deterministic and in stochastic cases, the constraint

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} vD_x\varphi(x) + \frac{\sigma^2}{2} \Delta\varphi(x) d\mu = 0,$$

($\sigma \geq 0$) is linear in v. Additionally, the Lagrangian is strictly convex in v. This implies that minimizing measure has support in a graph $(x, \bar{v}(x))$. In fact, if the minimizing measure $\mu(x, v)$ were not support in a graph, we could replace it by another measure $\tilde{\mu}$ given by

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \phi(x, v) d\tilde{\mu}(x, y) = \int_{\mathbb{T}^n} \phi(x, \bar{v}(x)) d\theta(x),$$

where

$$\bar{v}(x) = \int_{\mathbb{R}^n} v\mu(x, v) dv,$$

and

$$\int_{\mathbb{T}^n} \psi(x) d\theta(x) = \int_{\mathbb{T}^n \times \mathbb{R}^n} \psi(x) \mu(x, v) dv,$$

for all $\psi \in C(\mathbb{T}^n)$. Thus

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} v D_x \varphi(x) + \frac{\sigma^2}{2} \Delta \varphi(x) d\tilde{\mu} = 0.$$

Additionally, the convexity of L in v implies

$$\int L d\tilde{\mu} \leq \int L d\mu.$$

If L is strictly convex, the inequality is strict unless $v = \bar{v}(x)$, μ almost everywhere.

In conclusion:

Theorem 73. *Let $L(x, v)$ be strictly convex in v and μ a minimizing measure for Mather's problem (deterministic or stochastic). Then μ it is supported in a graph*

$$(x, v) = (x, \bar{v}(x)).$$

Additionally the projection θ of μ in the coordinate x satisfies

$$-\nabla \cdot (\bar{v}(x)\theta(x)) + \frac{\sigma^2}{2} \Delta \theta = 0,$$

and the distribution sense.

In order to simplify the presentation we are going to assume that $L = \frac{|v|^2}{2} - U(x)$. Using formally Lagrange multipliers, we conclude that Mather's problem is equivalent to the problem without constraints

$$\min_{\theta, v(x)} \int_{\mathbb{T}^n} \left(\frac{|v|^2}{2} - U(x) + vD_x\varphi + \frac{\sigma^2}{2}\Delta\varphi + \overline{H} \right) \theta dx.$$

The function φ corresponds to the Lagrange multiplier for the holonomy condition and $\overline{H}$ to the constraint $\int_{\mathbb{T}^n} \theta = 1$.

To obtain the Euler-Lagrange equation, we make the following variations

$$v \to v + \epsilon w, \qquad \theta \to \theta + \epsilon\eta.$$

This implies

$$v = -D_x\varphi(x),$$

and

$$\frac{|v|^2}{2} - U(x) + vD_x\varphi + \frac{\sigma^2}{2}\Delta\varphi + \overline{H} = 0.$$

Therefore

$$-\frac{\sigma^2}{2}\Delta\varphi + H(D_x\varphi, x) = \overline{H}, \tag{8.9}$$

with

$$H(p, x) = \frac{|p|^2}{2} + U(x).$$

As an application, we are going to prove an estimate for the second derivatives of the solution of the Hamilton-Jacobi equation. In order to keep the presentation as elementary as possible we assume that the dimension is 1. We further assume that the solution to equation (8.9) is twice differentiable in x:

$$-\frac{\sigma^2}{2}\Delta(\varphi_{xx}) + D_x\varphi D_x(\varphi_{xx}) + |D_x\varphi_x|^2 + U_{xx} = 0.$$

Since $v=-D_x\varphi$ we have

$$\int -\frac{\sigma^2}{2}\Delta(\varphi_{xx})+D_x\varphi D_x(\varphi_{xx})d\mu=0,$$

and therefore

$$\int |D^2\varphi|^2 d\mu \leq C.$$

Mather's problem is an infinite dimensional linear programming problem, and we can use duality, as we will discuss in section 8.3, to gain a better understanding of the problem. For the stochastic Mather problem, the dual is given by

$$\inf_\phi \sup_x -\frac{\sigma^2}{2}\Delta\phi+H(D_x\phi,x).$$

The duality theory implies that the value of this infimum is

$$-\int L d\mu.$$

On the other hand, this value is also the unique number $\overline{H}$ for which

$$-\frac{\sigma^2}{2}\Delta u+H(D_x u,x)=\overline{H}$$

has a periodic solution u. If we assume the existence of a smooth solution to the Hamilton-Jacobi equation we can check this fact directly. To do so, let u be a solution of (8.9) then

$$\inf_\phi \sup_x -\frac{\sigma^2}{2}\Delta\phi+H(D_x\phi,x)\leq \sup_x -\frac{\sigma^2}{2}\Delta u+H(D_x u,x)=\overline{H}.$$

Additionally, for each periodic function ϕ, $u-\phi$ has a minimum at a point x_0. At this point, $D_x u=D_x\phi$, and $\Delta u\geq\Delta\phi$. Therefore

$$\begin{aligned}\sup_x -\frac{\sigma^2}{2}\Delta\phi+H(D_x\phi,x) &\geq -\frac{\sigma^2}{2}\Delta\phi(x_0)+H(D_x\phi,x_0)\\ &\geq -\frac{\sigma^2}{2}\Delta u(x_0)+H(D_x u,x_0)=\overline{H}.\end{aligned}$$

8.3 Duality

In this section we make rigorous some of the previous discussion by considering duality theory. The main tool is the Legendre-Fenchel-Rockefellar theorem, whose proof will be presented in what follows, our proof is based in the one presented in [Vil03].

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

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

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

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

Theorem 74 (Fenchel-Legendre-Rockafellar)**.** *Let E be a locally convex topological vector space over $\mathbb{R}$ with dual E'. Let $h : E \to (-\infty,+\infty]$ be a convex function and $g : E \to [-\infty,+\infty)$ a concave function. Then, if there exists a point x_0 where both g and h are finite and at least one of them is continuous,*

$$\min_{y\in E'}\left[h^*(y) - g^*(y)\right] = \sup_{x\in E}\left[g(x) - h(x)\right]. \tag{8.10}$$

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

Proof. First we show the "$\geq$" inequality in (8.10). Recall that

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

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

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

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

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

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

$$g(x_1) - h(x_2) + (y, x_1 - x_2) \leq \lambda, \tag{8.11}$$

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

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

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

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

and

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

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

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

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

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

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

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

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

This is equivalent to (8.11) and thus we completed the proof. □

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

8.4 Generalized Mather problem

The generalized Mather problem is an infinite dimensional linear programming problem. We will use Fenchel-Legendre-Rockafellar's theorem to compute the dual problem.

Let $\Omega = \mathbb{T}^n \times U$. If U is bounded, set $\gamma = 1$, otherwise, let γ be a function $\gamma(v) : \Omega \to [1, +\infty)$ satisfying

$$\lim_{|v| \to +\infty} \frac{L(x, v)}{\gamma(v)} = +\infty, \qquad \lim_{|v| \to +\infty} \frac{|v|}{\gamma(v)} = 0.$$

Let $\mathcal{M}$ be the set of Radon measures in Ω with weight γ, that is,

$$\mathcal{M} = \left\{ \mu \text{ signed measure in } \Omega \text{ with } \int_\Omega \gamma d|\mu| < \infty \right\}.$$

The set $\mathcal{M}$ is the dual of the set $C_{\gamma,0}(\Omega)$ of continuous functions ϕ that satisfy

$$\|\phi\|_\gamma = \sup_\Omega \left| \frac{\phi}{\gamma} \right| < \infty, \tag{8.12}$$

if U is bounded, and, if U is unbounded, satisfy both (8.12) and

$$\lim_{|v|\to\infty} \frac{\phi(x,v)}{\gamma(v)} = 0.$$

Let

$$\mathcal{M}_1 = \left\{ \mu \in \mathcal{M} : \int_\Omega d\mu = 1, \mu \geq 0 \right\},$$

and

$$\mathcal{M}_2 = \operatorname{cl} \left\{ \mu \in \mathcal{M} : \int_\Omega A^v \varphi d\mu = \int_\Omega B\varphi d\nu, \forall \varphi(x) \in C^k(\mathbb{T}^n) \right\},$$

in which k is the degree of differentiability needed on φ so that $A^v\varphi$ is well defined, and the closure cl is taken in the weak topology.

For $\phi \in C_{\gamma,0}(\Omega)$ let

$$h(\phi) = \sup_{(x,v)\in\Omega} (-\phi(x,v) - L(x,v)).$$

Since h is the supremum of convex functions, it is also a convex function, and, as was shown in [Gom02], it is also continuous with respect to uniform convergence in $C_{\gamma,0}(\Omega)$. Consider the set

$$\mathcal{C} = \operatorname{cl} \left\{ \phi : \phi = A^v \varphi, \varphi \in C^k(\mathbb{T}^n) \right\},$$

where cl denotes the closure in $C_{\gamma,0}$. Since A^v is a linear operator, $\mathcal{C}$ is a convex set.

Let ν be a fixed probability measure on $\mathbb{T}^n$, and let μ_ν as in (8.6). Define

$$g(\phi) = \begin{cases} -\int \phi d\mu_\nu & \text{if } \phi \in \mathcal{C}, \\ -\infty & \text{otherwise.} \end{cases}$$

As $\mathcal{C}$ is a closed convex set, g is concave and upper semicontinuous. Note that if $\phi = A^v\varphi$, then $\int \phi d\mu_\nu = \int B\varphi d\nu$.

We claim that the dual of

$$\sup_{\phi \in C_0^\gamma(\Omega)} g(\phi) - h(\phi) \tag{8.13}$$

is the generalized Mather problem .

We start by computing the Legendre transforms of h and g.

Proposition 75. *We have*

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

and

$$g^*(\mu) = \begin{cases} 0 & \textit{if} \quad \mu \in \mathcal{M}_2 \\ -\infty & \textit{otherwise.} \end{cases}$$

Proof. By its definition

$$h^*(\mu) = \sup_{\phi \in C_0^\gamma(\Omega)} \left(-\int \phi d\mu - h(\phi) \right).$$

First we show that if μ is non-positive then $h^*(\mu) = \infty$.

Lemma 76. *If* $\mu \not\geq 0$ *then* $h^*(\mu) = +\infty$.

Proof. If $\mu \not\geq 0$ we can choose a sequence of non-negative functions $\phi_n \in C_0^\gamma(\Omega)$ such that

$$\int -\phi_n d\mu \to +\infty.$$

Therefore, since

$$\sup -\phi_n - L \leq 0,$$

we have $h^*(\mu) = +\infty$. □

Lemma 77. *If* $\mu \geq 0$ *then*

$$h^*(\mu) \geq \int L d\mu + \sup_{\psi \in C_0^\gamma(\Omega)} \left(\int \psi d\mu - \sup \psi \right).$$

Proof. Let L_n be a sequence of functions in $C_0^\gamma(\Omega)$ increasing pointwisely to L. Any ϕ in $C_0^\gamma(\Omega)$ can be written as $\phi = -L_n - \psi$, for some ψ in $C_0^\gamma(\Omega)$. Therefore

$$\sup_{\phi \in C_0^\gamma(\Omega)} \left(-\int \phi d\mu - h(\phi) \right) =$$
$$= \sup_{\psi \in C_0^\gamma(\Omega)} \left(\int L_n d\mu + \int \psi d\mu - \sup(L_n + \psi - L) \right).$$

Since

$$\sup (L_n - L) \leq 0,$$

we have

$$\sup(L_n + \psi - L) \leq \sup \psi.$$

Therefore

$$\sup_{\phi \in C_0^\gamma(\Omega)} \left(-\int \phi d\mu - h(\phi) \right)$$
$$\geq \sup_{\psi \in C_0^\gamma(\Omega)} \left(\int L_n d\mu + \int \psi d\mu - \sup \psi \right).$$

By the monotone convergence theorem

$$\int L_n d\mu \to \int L d\mu.$$

Thus,

$$\sup_{\phi \in C_0^\gamma(\Omega)} \left(-\int \phi d\mu - h(\phi) \right) \geq \int L d\mu + \sup_{\psi \in C_0^\gamma(\Omega)} \left(\int \psi d\mu - \sup \psi \right),$$

as required. □

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

$$\sup_{\psi \in C_0^\gamma(\Omega)} \left(\int \psi d\mu - \sup \psi \right) \geq \sup_{\alpha \in \mathbb{R}} \alpha \left(\int d\mu - 1 \right) = +\infty,$$

by choosing $\psi = \alpha$, constant. Therefore $h^*(\mu) = +\infty$.

When $\int d\mu = 1$, the previous lemma implies

$$h^*(\mu) \geq \int L d\mu,$$

by choosing $\psi = 0$.

Additionally, for each ϕ

$$\int (-\phi - L) d\mu \leq \sup(-\phi - L),$$

if $\int d\mu = 1$. Therefore

$$\sup_{\phi \in C_0^\gamma(\Omega)} \left(-\int \phi d\mu - h(\phi) \right) \leq \int L d\mu.$$

In this way,

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

Let μ_ν be such that

$$\int A^v \varphi d\mu_\nu = \int B\varphi d\nu,$$

for all $\varphi \in C^k(\mathbb{T}^n)$. We can write any measure $\mu \in \mathcal{M}_2$ as a sum of $\mu_\nu + \hat{\mu}$, with

$$\int A^v \varphi d\hat{\mu} = 0,$$

for all $\varphi \in C^k(\mathbb{T}^n)$. By continuity, it follows

$$\int \phi d\hat{\mu} = 0,$$

for all $\phi \in \mathcal{C}$. Furthermore, for any $\mu \notin \mathcal{M}_2$, there exists $\hat{\phi} \in \mathcal{C}$ such that

$$\int \hat{\phi} d(\mu - \mu_\nu) \neq 0.$$

Thus

$$g^*(\mu) = \inf_{\phi \in \mathcal{C}} - \int \phi d\mu + \int \phi d\mu_\nu = \begin{cases} 0 & \text{if } \mu \in \mathcal{M}_2 \\ -\infty & \text{otherwise.} \end{cases}$$

□

Theorem 78.

$$\sup_{\phi \in C_{\gamma,0}(\Omega)} (g(\phi) - h(\phi)) = \min_{\mu \in \mathcal{M}} (h^*(\mu) - g^*(\mu)). \tag{8.14}$$

NOTE 1: $\min_{\mu \in \mathcal{M}} (h^*(\mu) - g^*(\mu)) = \min_{\mu \in \mathcal{M}_1 \cap \mathcal{M}_2} \int L d\mu$.

NOTE 2: It is part of the theorem that the right-hand side of (8.14) is a minimum, and therefore there exists a generalized Mather measure.

Proof. The set $g > -\infty$ is non-empty, and, in this set, h is a continuous function as proved in [Gom02]. Then the result follows from Fenchel-Legendre-Rockafellar's Theorem, see, for instance [Vil03]. $\square$

Let
$$\mathcal{H}(\varphi, x) = \sup_v -L(x, v) - A^v\varphi.$$
As an example, suppose $A^v\varphi = \Delta\varphi + vD_x\varphi$. Then
$$\mathcal{H}(\varphi, x) = -\Delta\varphi + H(D_x\varphi, x).$$
The result in Theorem 78 can then be restated in the more convenient identity:
$$\min_\mu \int L d\mu = -\inf_\varphi \sup_x \left[\mathcal{H}(\varphi, x) + \int B\varphi d\nu\right], \tag{8.15}$$
where the minimum on the left-hand side is taken over all measures μ that satisfy (8.8), and the infimum on the right-hand side is taken over all $\varphi \in C^k(\mathbb{T}^n)$.

In the remaining of this section we consider Mather's classical problem $A^v\varphi = vD_x\varphi$ and $B = 0$.

Theorem 79. *Let $A^v\varphi = vD_x\varphi$. Let $H^\star$ given by*
$$H^\star = -\sup_{\phi\in C_0^\gamma(\Omega)} (h_2(\phi) - h_1(\phi)).$$
Then
$$H^\star = \inf\{\lambda : \exists\varphi \in C^1(\mathbb{T}^n) : H(D_x\varphi, x) < \lambda\}.$$

Proof. It is enough to observe that
$$H^\star = \inf_{\varphi\in C^1(\mathbb{T}^n)} \sup_{(x,v)\in\Omega} -vD_x\varphi - L = \inf_{\varphi\in C^1(\mathbb{T}^n)} \sup_{x\in\mathbb{T}^n} H(D_x\varphi, x).$$
$\square$

We recall that from theorem 61 $H^\star$ is the unique value for which

$$H(D_x u, x) = H^\star$$

admits a periodic viscosity solution.

8.4.1 Regularity

Now we present (with small adaptations) the regularity results for viscosity solutions in the support of the Mather measures by [EG01].

Lemma 80. *Let μ be a minimizing holonomic measure. Then*

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

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

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

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

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

that is,

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

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

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

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

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

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

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

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

because we can use identity $v = -D_p H(p, x)$ if $p = -D_v L(x, v)$.

Theorem 81. *Let u be any viscosity solution of* (5.4)*, and let μ be any minimizing holonomic measure. Then $\bar{\mu}$-almost everywhere, $D_x u(x)$ exists and $p = D_x u(x)$, $\tilde{\mu}$-almost everywhere.*

Proof. Let u be any viscosity solution of (5.4). Let η_ϵ be a standard mollifier, $u^\epsilon = \eta_\epsilon * u$. By strict uniform convexity there exists $\gamma > 0$ such that for any $p, q \in \mathbb{R}^d$ and any $x \in \mathbb{T}^d$ we have

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

By Theorem 56, any viscosity solution of (5.4), and in particular u, is Lipschitz.

Recall that, by Rademacher's theorem [Eva98a], a locally Lipschitz function is differentiable Lebesgue almost everywhere. Using

$p = D_x u(y)$ and $q = D_x u^\epsilon(x)$, conclude that for every point x and for Lebesgue almost every point y:

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

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

$$\begin{aligned}&H(D_x u^\epsilon(x), x) + \frac{\gamma}{2}\int_{\mathbb{R}^d} \eta_\epsilon(x-y)|D_x u^\epsilon(x) - D_x u(y)|^2 dy \\ &\quad \leq \int_{\mathbb{R}^d} \eta_\epsilon(x-y) H(D_x u(y), x) dy \leq \overline{H} + O(\epsilon).\end{aligned}$$

Let

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

Now observe that

$$\begin{aligned}&\frac{\gamma}{2}\int_{\mathbb{T}^d\times\mathbb{R}^d} |D_x u^\epsilon(x) - p|^2 d\tilde{\mu} \\ &\leq \int_{\mathbb{T}^d\times\mathbb{R}^d} \left[H(D_x u^\epsilon(x), x) - H(p, x) - D_p H(p, x)(D_x u^\epsilon(x) - p)\right] d\tilde{\mu} \\ &\leq \int_{\mathbb{T}^d\times\mathbb{R}^d} H(D_x u^\epsilon(x), x) d\tilde{\mu} - \overline{H},\end{aligned}$$

because

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

and

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

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

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

Thus, for $\bar{\mu}$-almost every point x, $\beta_\epsilon(x) \to 0$. Therefore, $\bar{\mu}$-almost every point is a point of approximate continuity of $D_x u$ (see [EG92], p. 49). Since u is semiconcave (Proposition 58), it is differentiable at points of approximate continuity. Furthermore

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

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

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

□

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

Corollary 82. *Let η_ϵ be a standard mollifier, $u^\epsilon = \eta_\epsilon * u$. Then*

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

as $\epsilon \to 0$.

As a Corollary we formulate an equivalent form of Theorem 81.

Corollary 83. *Let u be any viscosity solution of (5.4), and let μ be any minimizing holonomic measure. Then μ-almost everywhere, $D_x u(x)$ exists and*

$$D_v L(v, x) = D_x u(x) \qquad \mu - \textit{almost everywhere.} \tag{8.16}$$

and

$$D_x L(v, x) = -D_x H(D_x u(x), x) \qquad \mu - \textit{almost everywhere.} \tag{8.17}$$

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

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

implies the μ – almost everywhere identity

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

Thus (8.16) follows directly from Theorem 81.

Using (8.16) and the identity $D_x L(v,x) = -D_x H(D_v L(v,x),x)$, we arrive at (8.17). □

We observe that from the previous corollary it also follows

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

Indeed,

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

We have

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

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

$$\left| \int_{\mathbb{T}^d} D_p H(D_x u, x) \left(D_x u - D_x u^\epsilon\right) \right|$$
$$\leq \delta \int_{\mathbb{T}^d} |D_p H(D_x u, x)|^2 d\bar{\mu} + \frac{1}{\delta} \int_{\mathbb{T}^d} \left|D_x u - D_x u^\epsilon\right|^2 d\bar{\mu}.$$

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

Theorem 84. *Let u be any viscosity solution of (5.4), and let μ be any minimizing holonomic measure. Then*

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

Proof. Applying Theorem 60 we have

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

By Theorem 81 the derivative $D_x u(x)$ exists $\bar{\mu}$ almost everywhere. We recall a viscosity solution satisfies (5.4) in classical sense at all points of differentiability. Thus $H(D_x u(x), x) = \overline{H}$ for $\bar{\mu}$ almost all points x. Now observe that

$$\begin{aligned} C\epsilon &\geq H(D_x u^\epsilon(x+h), x+h) - H(D_x u(x), x) \\ &= H(D_x u^\epsilon(x+h), x+h) - H(D_x u^\epsilon(x+h), x) \\ &\quad + H(D_x u^\epsilon(x+h), x) - H(D_x u(x), x) \end{aligned}$$

The term

$$\begin{aligned} &H(D_x u^\epsilon(x+h), x+h) - H(D_x u^\epsilon(x+h), x) \\ &\quad = D_x H(D_x u^\epsilon(x+h), x)h + O(h^2) \\ &\quad = D_x H(D_x u(x), x)h + O(h^2 + h|D_x u^\epsilon(x+h) - D_x u(x)|) \\ &\quad \geq D_x H(D_x u(x), x)h + O(h^2) - \frac{\gamma}{4}|D_x u^\epsilon(x+h) - D_x u(x)|^2. \end{aligned}$$

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

$$\begin{aligned} &H(D_x u^\epsilon(x+h), x) - H(D_x u, x) \\ &\quad \leq C\epsilon - D_x H(D_x u(x), x)h + \frac{\gamma}{4}|D_x u^\epsilon(x+h) - D_x u(x)|^2 + Ch^2. \end{aligned}$$

Since

$$\begin{aligned} &H(D_x u^\epsilon(x+h), x) - H(D_x u, x) \\ &\quad \geq \frac{\gamma}{2}|D_x u^\epsilon(x+h) - D_x u(x)|^2 \\ &\quad + D_p H(D_x u, x)(D_x u^\epsilon(x+h) - D_x u(x)) \end{aligned}$$

we have

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

By (8.17) and Lemma 80 it follows

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

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

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

Finally, we claim that $\xi(x) = D_x u(x+h)$ for $\bar{\mu}$ almost all x. This follows from Theorem 81 and the fact that for $\bar{\mu}$ almost all x we have $\xi(x) \in D_x^- u(x+h)$, where D_x^- stands for the subdifferential. To see this, observe that by Proposition 58 u is semiconcave, therefore u^ϵ are uniformly semiconcave, that is

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

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

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

By passing to the limit we have that

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

for all y and $\bar{\mu}$-almost all x, that is, $\xi(x) \in D_x^- u(x+h)$ for $\bar{\mu}$-almost all x. □

Lemma 85. *Let u be any viscosity solution of* (5.4)*, and let μ be any minimizing holonomic measure. Let $\psi : \mathbb{T}^d \times \mathbb{R} \to \mathbb{R}$ be a smooth function. Then*

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

Proof. Clearly we have

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

By the uniform convergence of u^ϵ to u, and $L^2_{\bar{\mu}}$ convergence of $D_x u^\epsilon$ to $D_x u$, see Corollary 82, we get the result. □

Theorem 86. *Let u be any viscosity solution of* (5.4)*, and let μ be any minimizing holonomic measure. Then, for $\bar{\mu}$ almost every x and all $h \in \mathbb{R}^d$,*

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

Proof. Let $h \neq 0$ and define

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

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

$$0 < \epsilon \leq \eta|h|^2, \tag{8.18}$$

for small $\eta > 0$. We have

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

For $\bar{\mu}$-almost every point x, $Du(x)$ exists and therefore

$$H(Du(x), x) = \overline{H},$$

so we have

$$\begin{aligned} &H(D\tilde{u}^\epsilon, x) - 2H(Du, x) + H(D\hat{u}^\epsilon, x) \\ &\quad \le 2C\epsilon + H(D\tilde{u}^\epsilon, x) - H(D\tilde{u}^\epsilon, x+h) \\ &\quad\quad + H(D\hat{u}^\epsilon, x) - H(D\hat{u}^\epsilon, x-h). \end{aligned}$$

Hence

$$\begin{aligned} &\frac{\gamma}{2}(|D\tilde{u}^\epsilon - Du|^2 + |D\hat{u}^\epsilon - Du|^2) \\ &\quad\quad + D_pH(Du, x)\cdot(D\tilde{u}^\epsilon - 2Du + D\hat{u}^\epsilon) \\ &\quad\quad \le C(\epsilon + |h|^2) + (D_xH(D\hat{u}^\epsilon, x) - D_xH(D\tilde{u}^\epsilon, x))\cdot h. \end{aligned}$$

Using the inequality

$$\begin{aligned} &\left|\big(D_xH(p,x) - D_xH(q,x)\big)\cdot h\right| \\ &\quad \le \left\|\tfrac{\partial^2 H}{\partial p \partial x}\right\| |p-q|\,|h| \le \tfrac{\gamma}{4}|p-q|^2 + \tfrac{1}{\gamma}\left\|\tfrac{\partial^2 H}{\partial p\partial x}\right\|^2 |h|^2, \end{aligned}$$

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

$$\begin{aligned} &\frac{\gamma}{4}(|D\tilde{u}^\epsilon - Du|^2 + |D\hat{u}^\epsilon - Du|^2) \\ &+ D_pH(Du, x)\cdot(D\tilde{u}^\epsilon - 2Du + D\hat{u}^\epsilon) \\ &\quad\quad \le C(\epsilon + |h|^2). \end{aligned}$$

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

$$\begin{aligned} &\frac{\gamma}{4}\int_{\mathbb{T}^d}(|D\tilde{u}^\epsilon - Du|^2 + |D\hat{u}^\epsilon - Du|^2)\phi\left(\frac{\tilde{u}^\epsilon - 2u + \hat{u}^\epsilon}{|h|^2}\right) d\bar{\mu} \qquad (8.19) \\ &\quad + \int_{\mathbb{T}^d} D_pH(Du, x)\cdot(D\tilde{u}^\epsilon - 2Du + D\hat{u}^\epsilon)\phi(\cdots)\, d\bar{\mu} \\ &\quad\quad\quad \le C(\epsilon + |h|^2)\int_{\mathbb{T}^d}\phi(\cdots)\, d\bar{\mu}. \end{aligned}$$

Now the second term on the left hand side of (8.19) equals

$$|h|^2 \int_{\mathbb{R}^d} \int_{\mathbb{T}^d} D_p H(p,x) \cdot D_x \left[\Phi \left(\frac{\tilde{u}^\epsilon - 2u + \hat{u}^\epsilon}{|h|^2} \right) \right] d\bar{\mu} \tag{8.20}$$

and thus, by Lemma 85 it vanishes. So now dropping the above term from (8.19) and rewriting, we deduce

$$\begin{aligned} &\int_{\mathbb{T}^d} |Du^\epsilon(x+h) - Du^\epsilon(x-h)|^2 \phi \left(\tfrac{u^\epsilon(x+h) - 2u(x) + u^\epsilon(x-h)}{|h|^2} \right) d\bar{\mu} \\ &\qquad \le C(\epsilon + |h|^2) \int_{\mathbb{T}^d} \phi \left(\tfrac{u^\epsilon(x+h) - 2u(x) + u^\epsilon(x-h)}{|h|^2} \right) d\bar{\mu}. \end{aligned} \tag{8.21}$$

We confront now a technical problem, as (8.21) entails a mixture of first-order difference quotients for Du^ϵ and second-order difference quotients for u, u^ϵ. We can however relate these expressions, since u is semiconcave.

To see this, first of all define

$$E_\epsilon := \{x \in \operatorname{supp}(\mu) \mid u^\epsilon(x+h) - 2u(x) + u^\epsilon(x-h) \le -\kappa|h|^2\}, \tag{8.22}$$

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

$$\bar{u}(x) := u(x) - \frac{\alpha}{2}|x|^2, \quad \bar{u}^\epsilon(x) := u^\epsilon(x) - \frac{\alpha}{2}|x|^2 \tag{8.23}$$

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

$$\bar{u}^\epsilon(x+h) - 2\bar{u}(x) + \bar{u}^\epsilon(x-h) \le -(\kappa + \alpha)|h|^2. \tag{8.24}$$

Set

$$f^\epsilon(s) := \bar{u}^\epsilon \left(x + s\frac{h}{|h|} \right) \qquad (-|h| \le s \le |h|). \tag{8.25}$$

Then f is concave, and

$$\begin{aligned}\bar{u}^\epsilon(x+h) - 2\bar{u}^\epsilon(x) + \bar{u}^\epsilon(x-h) &= f^\epsilon(|h|) - 2f^\epsilon(0) + f^\epsilon(-|h|) \\ &= \int_{-|h|}^{|h|} f^{\epsilon''}(x)(|h| - |s|)\,ds \\ &\geq |h| \int_{-|h|}^{|h|} f^{\epsilon''}(s)\,ds \quad (\text{since } f^{\epsilon''} \leq 0) \\ &= |h|\left[f^{\epsilon'}(|h|) - f^{\epsilon'}(-|h|)\right] \\ &= (D\bar{u}^\epsilon(x+h) - D\bar{u}^\epsilon(x-h)) \cdot h.\end{aligned}$$

Consequently, if $x \in E_\epsilon$, this inequality and (8.24) together imply

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

Now $|\bar{u}^\epsilon(x) - \bar{u}(x)| \leq C\epsilon$ on $\mathbb{T}^d$, since u is Lipschitz continuous. We may therefore take η in (8.18) small enough to deduce from the foregoing that

$$|D\bar{u}^\epsilon(x+h) - D\bar{u}^\epsilon(x-h)| \geq (\frac{\kappa}{2} + \alpha)|h|. \tag{8.26}$$

But then

$$|Du^\epsilon(x+h) - Du^\epsilon(x-h)| \geq (\frac{\kappa}{2} - \alpha)|h|. \tag{8.27}$$

Return now to (8.21). Taking $\kappa > 2\alpha$ and

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

The inequality (8.21) was derived for smooth functions ϕ. However, by replacing ϕ in (8.21) by a sequence ϕ_n of smooth functions increasing pointwise to ϕ, and using the monotone convergence theorem, we conclude that (8.21) holds for this function ϕ. Then we discover from (8.21) that

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

We fix κ so large that

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

to deduce

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

Thus $\bar{\mu}(E_\epsilon) = 0$ if η in (8.18) is small enough, and this means

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

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

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

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

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

owing to the semiconcavity, we have

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

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

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

Theorem 87. *Let u be any viscosity solution of* (5.4)*, and let μ be any minimizing holonomic measure. Then for $\bar{\mu}$-almost every x, $D_x u(x)$ exists and for Lebesgue almost every y*

$$|D_x u(x) - D_x u(y)| \leq C|x-y|. \tag{8.28}$$

Proof. First we show that

$$|u(y) - u(x) - (y - x) \cdot D_x u(x)| \leq C|x - y|^2. \tag{8.29}$$

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

According to Theorem 86 with $h := y - x$, we have

$$|u(y) - 2u(x) + u(2x - y)| \leq C|x - y|^2. \tag{8.30}$$

By semiconcavity, we have

$$u(y) - u(x) - Du(x) \cdot (y - x) \leq C|x - y|^2, \tag{8.31}$$

and also

$$u(2x - y) - u(x) - Du(x) \cdot (2x - y - x) \leq C|x - y|^2. \tag{8.32}$$

Use (8.32) in (8.30):

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

This and (8.31) establish (8.29).

Estimate (8.28) follows from (8.29), as follows. Take x, y as above. Let z be a point to be selected later, with $|x - z| \leq 2|x - y|$. The semiconcavity of u implies that

$$u(z) \leq u(y) + Du(y) \cdot (z - y) + C|z - y|^2. \tag{8.33}$$

Also,

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

according to (8.29). Insert these identities into (8.33) and simplify:

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

Now take

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

to deduce (8.28).

Now take any point $x \in \text{supp}(\bar{\mu})$, and fix y. There exist points $x_k \in \text{supp}(\bar{\mu})$ $(k = 1, \dots)$ such that $x_k \to x$ and u is differentiable at x_k. According to estimate (8.29)

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

The constant C does not depend on k or y. Now let $k \to \infty$. Owing to (8.28) we see that $\{Du(x_k)\}$ converges to some vector η, for which

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

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

It follows from Theorem 87 that function

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

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

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

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

8.4.2 Holonomy variations

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

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

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

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

$$\dot{x}_k(x,v) = \xi_k(x)\,, \quad \dot{v}_k(x,v) = v_i \frac{\partial \xi_k}{\partial x_i}(x)\,. \tag{8.34}$$

Lemma 88. *The flow of* (8.34) *is given by*

$$X_k(t,x,v) = \Phi_k(t,x)\,, \qquad V_k(t,x,v) = v_s \frac{\partial \Phi_k}{\partial x_s}(t,x). \tag{8.35}$$

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

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

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

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

A simple computations yields

$$\begin{aligned}\frac{\partial}{\partial t} V_k(t,x,v) &= v_s \frac{\partial}{\partial x_s}\big(\tfrac{\partial}{\partial t}\Phi_k(t,x)\big) = v_s \frac{\partial}{\partial x_s}\Big(\xi_k\big(\Phi(t,x)\big)\Big) \\ &= v_s \left.\frac{\partial \xi_k}{\partial x_i}\right|_{\Phi(t,x)} \left.\frac{\partial \Phi_i}{\partial x_s}\right|_{(t,x)} = V_i(t,x,v) \left.\frac{\partial \xi_k}{\partial x_i}\right|_{\Phi(t,x)},\end{aligned}$$

which is the desired identity. □

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

$$\psi_t(x,v) = \psi\big(X(t,x,v), V(t,x,v)\big). \tag{8.36}$$

Thus the flow (8.35) generates the flow on space of functions $\psi(x,v)$ given by (8.36). Consider the set

$$\mathcal{C} = \Big\{\psi \in C_0^\gamma(\mathbb{T}^d \times \mathbb{R}^d) : \psi(x,v) = v \cdot D_x\varphi(x), \text{ for some } \varphi \in C^1(\mathbb{T}^d)\Big\}. \tag{8.37}$$

Lemma 89. *The set* $\mathcal{C}$, *defined in* (8.37), *is invariant under the flow given by* (8.36).

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

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

where ψ_t is given by (8.36). Indeed,

$$\begin{aligned}\psi_t(x,v) = V_k(t,x,v) \left.\frac{\partial g}{\partial x_k}\right|_{X(t,x,v)} &= v_s \left.\frac{\partial g}{\partial x_k}\right|_{\Phi(t,x)} \left.\frac{\partial \Phi_k}{\partial x_s}\right|_{(t,x)} \\ &= v_s \frac{\partial}{\partial x_s}\Big(g\big(\Phi(t,x)\big)\Big) = v_s \frac{\partial}{\partial x_s} g_t(x),\end{aligned}$$

and so the Lemma is proved. □

The flow on functions (8.36) generates the flow on measures: $(t,\mu)\mapsto\mu_t$, where

$$\int \psi d\mu_t = \int \psi_t d\mu. \tag{8.38}$$

Lemma 90. *The flow* (8.38) *preserves the holonomy constraint.*

Proof. Let μ be a holonomic measure. We have to prove that μ_t is also a holonomic, i.e., $\int \psi d\mu_t = 0$ for any $\psi \in \mathcal{C}$. This is clear since the flow (8.36) preserves the set $\mathcal{C}$. □

Theorem 91. *Let μ be a minimizing measure for the action* (8.3), *subject to the holonomy constraint. Then for any C^1 vector field $\xi : \mathbb{T}^d \to \mathbb{R}^d$ we have*

$$\int \frac{\partial L}{\partial x_s}\xi_s + \frac{\partial L}{\partial v_s} v_k \tfrac{\partial}{\partial x_k}\xi_s d\mu = 0. \tag{8.39}$$

Proof. Let μ_t be the flow generated from μ by (8.38). Relation (8.39) expresses the fact $\frac{d}{dt}\big(\int L(x,v)d\mu_t\big)\Big|_{t=0} = 0$. □

8.4.3 Invariance

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

In what follows $(\ )^{-1}_{js}$ denotes the j,s entry of the inverse matrix. We will only use this notation for symmetric matrices, thus, this notation will not lead to any ambiguity. Before stating and proving the main Theorem of this section, we will prove an auxiliary lemma.

Lemma 92. *Let μ be a minimal holonomic measure. Let $v^\epsilon(x)$ be any smooth function. Let $\phi(x,v)$ be any smooth compactly supported function. Then*

$$\begin{aligned}
&\int v_k \frac{\partial \phi}{\partial x_k}\big(x, v^\epsilon(x)\big) d\mu \\
&+ \int \frac{\partial \phi}{\partial v_j}\big(x, v^\epsilon(x)\big) M \left(\frac{\partial L}{\partial x_s}(x,v) - v_k \frac{\partial^2 L}{\partial x_k \partial v_s}\big(x, v^\epsilon(x)\big) \right) d\mu \\
&= \int v_k \frac{\partial}{\partial x_k}\Big(\phi\big(x, v^\epsilon(x)\big)\Big) d\mu - \int v_k \frac{\partial}{\partial x_k}\Big(\frac{\partial L}{\partial v_s}\big(x, v^\epsilon(x)\big)\dot{X}_s^\epsilon\Big) d\mu \\
&+ \int v_k \Big(\frac{\partial L}{\partial v_s}\big(x, v^\epsilon(x)\big) - \frac{\partial L}{\partial v_s}(x,v)\Big) \frac{\partial}{\partial x_k}\Big(\dot{X}_s^\epsilon\Big) d\mu,
\end{aligned} \tag{8.40}$$

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

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

and

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

Remark. We will only use this lemma for the case when v^ϵ is the standard smoothing of the function $\boldsymbol{v}(x)$, that is, $v^\epsilon = \eta_\epsilon * \boldsymbol{v}$, where η_ϵ is a standard mollifier.

Proof. This Lemma is based on Theorem 91. In this proof and bellow v^ϵ stands for the function $v^\epsilon(x)$. We have:

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

Rewrite the last term:

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

Plug these two lines into (8.40). And therefore we reduce (8.40) to

$$\begin{aligned}
\int \dot{X}_s^\epsilon(x) \left(\frac{\partial L}{\partial x_s}(x, v) - v_k \left(\frac{\partial^2 L}{\partial x_k \partial v_s}(x, v^\epsilon) + \frac{\partial^2 L}{\partial v_s \partial v_q}(x, v^\epsilon) \frac{\partial v_q^\epsilon}{\partial x_k} \right) \right) d\mu \\
= - \int v_k \frac{\partial}{\partial x_k} \left(\frac{\partial L}{\partial v_s}(x, v^\epsilon) \dot{X}_s^\epsilon \right) d\mu \\
+ \int v_k \left(\frac{\partial L}{\partial v_s}(x, v^\epsilon) - \frac{\partial L}{\partial v_s}(x, v) \right) \frac{\partial}{\partial x_k} \left(\dot{X}_s^\epsilon \right) d\mu. \quad (8.41)
\end{aligned}$$

Using the chain rule in the LHS and the Leibniz rule in the RHS we further reduce (8.41) to

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

Noting the cancellation of the term $\int v_k \dot{X}_s^\epsilon \frac{\partial}{\partial x_k} \left(\frac{\partial L}{\partial v_s}(x, v^\epsilon) \right) d\mu$, we see that the last identity is equivalent to (8.39) with $\xi_s(x) = \dot{X}_s^\epsilon(x)$. □

We will need the following result concerning invariant measures under a flow:

Lemma 93. *Let μ be a measure on a manifold M. Let χ be a smooth vector field on M. The measure μ is invariant with respect to the flow generated by the vector field χ if and only if for any smooth compactly* *supported function* $\xi : M \to \mathbb{R}$ *we have*

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

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

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

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

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

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

and coincides with μ.

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

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

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

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

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

We have

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

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

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

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

Theorem 94. *Let μ be a minimizing holonomic measure. Then μ is invariant under the Euler-Lagrange flow.*

Proof. By Lemma 93 we have to prove that for any smooth compactly supported function $\phi(x, v)$

$$\int v_k \frac{\partial \phi}{\partial x_k} + \frac{\partial \phi}{\partial v_j} \left(\frac{\partial^2 L}{\partial^2 v}\right)^{-1}_{js} \left[\frac{\partial L}{\partial x_s} - v_k \frac{\partial^2 L}{\partial x_k \partial v_s}\right] d\mu = 0, \qquad (8.42)$$

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

The idea of the proof is first to rewrite (8.42) in an equivalent form and then apply an approximation argument. Since μ is supported by the graph $v = \boldsymbol{v}(x)$ we will change the x, v arguments with $x, \boldsymbol{v}(x)$ for the following four types of functions $\frac{\partial \phi}{\partial x_k}$, $\frac{\partial \phi}{\partial v_j}$, $\left(\frac{\partial^2 L}{\partial^2 v}\right)^{-1}_{js}$, and $\frac{\partial^2 L}{\partial x_k \partial v_s}$, occurring in (8.42):

$$\int v_k \frac{\partial \phi}{\partial x_k}\big(x, \boldsymbol{v}(x)\big) d\mu \\ + \int \frac{\partial \phi}{\partial v_j}\big(x, \boldsymbol{v}(x)\big) M \left(\frac{\partial L}{\partial x_s}(x, v) - v_k \frac{\partial^2 L}{\partial x_k \partial v_s}\big(x, \boldsymbol{v}(x)\big)\right) d\mu = 0, \tag{8.43}$$

where

$$M = \left(\frac{\partial^2 L}{\partial^2 v}\right)^{-1}_{js}.$$

To complete the proof of the theorem, we use Lemma 92. The first and second integrals in the RHS of (8.40) are zero due to the holonomy constraint. The third integral in the RHS of (8.40) tends to zero as $\epsilon \to 0$, because $|v^\epsilon(x) - \boldsymbol{v}(x)| < c\epsilon$ and therefore $|v^\epsilon(x) - v| < c\epsilon$ μ-a.e., and because $\dot{X}^\epsilon_s$ is uniformly Lipschitz and hence $\partial_{x_k} \dot{X}^\epsilon_s$ is uniformly bounded. Therefore the LHS of (8.40) tends to zero as $\epsilon \to 0$.

But the LHS of (8.40) also tends to the LHS of (8.43) as $\epsilon \to 0$. Indeed, since $\boldsymbol{v}(x)$ is a Lipschitz vector field we have

$$v^\epsilon(x) \to \boldsymbol{v}(x) \quad \text{(uniformly)}$$

and

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

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

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

and
$$\frac{\partial}{\partial x}\Big(\Psi\big(x, v^\epsilon(x)\big)\Big) \text{ is uniformly bounded.}$$
Also note that for μ almost all (x, v) we have $v = \boldsymbol{v}(x)$. Therefore the Theorem is proved. □

8.5 Generalized Viscosity solutions

At this point we must consider a suitable class of weak solutions to the equation
$$\mathcal{H}(u, x) = \lambda, \tag{8.44}$$
the viscosity solutions. The motivation to use viscosity solutions is the following: as we mentioned before, in the case that $U = \mathbb{R}^n$ and $A^v\varphi = vD_x\varphi$, the operator $\mathcal{H}$ is simply the Hamilton-Jacobi operator $H(D_xu, x)$. For first and second-order Hamilton-Jacobi equations, viscosity solution is the right notion of solution to (8.44), and it is therefore natural to extend it to our setting. To this end, we say that a function u is a viscosity solution of (8.44) if the following property holds: for any $\varphi \in C^k(\mathbb{T}^n)$ and any maximizer x_0 of $u - \varphi$ (resp. minimizer) such that $\varphi(x_0) = u(x_0)$, we have
$$\mathcal{H}(\varphi, x_0) \leq \lambda \quad (\text{resp.} \quad \geq).$$

We should remark that, in this generalized setting, there is no uniqueness of the number λ for which (8.44) admits a viscosity solution, even less uniqueness of viscosity solutions. For instance, uniqueness of λ it is false for the equation
$$-u + H(D^2u, Du, x) = \lambda,$$

which either admits viscosity solutions for all values of λ, or does not have any viscosity solution. Uniqueness of viscosity solution also fails as the following elementary example shows: let $\psi : \mathbb{T}^n \to \mathbb{R}$ be a non-constant C^1 function. Then

$$Du(Du - D\psi) = 0 \tag{8.45}$$

has two different solutions $u = 0$ and $u = \psi$.

Theorem 95. *Let $\lambda \in \mathbb{R}$ and let $u : \mathbb{T}^n \to \mathbb{R}$ be a corresponding viscosity solution to the Hamilton-Jacobi equation* (8.44)*. Assume that the operator B is monotone, that is, $\varphi_1 \geq \varphi_2$ implies $B\varphi_2 \geq B\varphi_1$, for all continuous functions φ_1, φ_2, and, furthermore, that it is continuous with respect to uniform convergence, that is, $\varphi_n \to \varphi$ uniformly, implies $B\varphi_n \to B\varphi$ uniformly. In addition, suppose that there exist C^k functions u_ϵ such that*

$$\mathcal{H}(u_\epsilon, x) \leq \lambda + O(\epsilon),$$

and, finally, that $u_\epsilon \to u$ uniformly. Then

$$\inf_\varphi \sup_x \mathcal{H}(\varphi, x) + \int B\varphi d\nu = \lambda + \int Bud\nu.$$

Proof. Let λ and u be as in the statement of the theorem. Then, for any smooth function $\varphi : \mathbb{T}^n \to \mathbb{R}$, there exists a point x_φ at which $u - \varphi$ has a minimum. Clearly,

$$\begin{aligned}
&\inf_\varphi \sup_x \mathcal{H}(\varphi, x) + \int_{\mathbb{T}^n} B\varphi d\nu \\
&\geq \inf_\varphi \mathcal{H}(u(x_\varphi) + \varphi - \varphi(x_\varphi), x_\varphi) + \int_{\mathbb{T}^n} B\varphi d\nu \\
&\quad + \int_{\mathbb{T}^n} B(u(x_0) - \varphi(x_0)) d\nu,
\end{aligned}$$

where we have used (8.5) to add the constant $u(x_0)-\varphi(x_0)$ to φ. By the viscosity property we conclude that

$$\begin{aligned}&\inf_\varphi \sup_x \mathcal{H}(\varphi,x)+\int_{\mathbb{T}^n} B\varphi d\nu\\ &\quad\geq \lambda+\inf_\varphi\int_{\mathbb{T}^n} B\varphi d\nu+\int_{\mathbb{T}^n} B(u(x_\varphi)-\varphi(x_\varphi))d\nu.\end{aligned}$$

Since x_φ is a minimum of $u-\varphi$, we have $u(x)-\varphi(x)\geq u(x_\varphi)-\varphi(x_\varphi)$, which implies $u(x)\geq\varphi(x)+u(x_\varphi)-\varphi(x_\varphi)$. This inequality yields, from the monotonicity of B, that

$$B\left[\varphi+u(x_\varphi)-\varphi(x_\varphi)\right]\geq Bu,$$

and thus

$$\inf_\varphi\sup_x\mathcal{H}(\varphi,x)+\int_{\mathbb{T}^n}B\varphi d\nu\geq\lambda+\int_{\mathbb{T}^n}Bud\nu.$$

To establish the reverse inequality, we use the sequence u_ϵ to obtain:

$$\begin{aligned}\inf_\varphi\sup_x\mathcal{H}(\varphi,x)+\int B\varphi d\nu&\leq\liminf_{\epsilon\to0}\sup_x\mathcal{H}(u_\epsilon,x)+\int Bu_\epsilon d\nu\\ &\leq\lambda+\liminf_{\epsilon\to0}\int Bu_\epsilon d\nu.\end{aligned}$$

□

As an example, consider the discounted Mather problem, in this case $\sigma=0$, $f(x,v)=v$ and

$$A^v\varphi=-\alpha\varphi+vD_x\varphi,$$

and

$$B\varphi=-\alpha\varphi.$$

The corresponding Hamilton-Jacobi equation is

$$\alpha u^\alpha + H(D_x u^\alpha, x) = 0, \tag{8.46}$$

and the value of the generalized Mather problem is given by

$$\alpha \int u^\alpha d\nu.$$

Therefore, if we set $\nu = \delta_y(x)$ we have

$$u^\alpha(y) = \int u^\alpha d\nu = \frac{1}{\alpha} \sup_\varphi \inf_x \left[-\alpha\varphi - H(D_x\varphi, x) + \alpha\varphi(y)\right],$$

which is a representation formula for the value of any viscosity solution of (8.46) and, in particular, implies uniqueness of solution of (8.46).

Lemma 96. *Let $\lambda \in \mathbb{R}$, and assume that there exists a viscosity solution of*

$$\mathcal{H}(u, x) = \lambda.$$

Furthermore, suppose that for all sufficiently small $\epsilon > 0$, there exist C^k functions u_ϵ such that

$$\mathcal{H}(u_\epsilon, x) \leq \lambda + O(\epsilon).$$

Suppose further, that $\int Bu_\epsilon d\nu - \int Bud\nu = O(\epsilon)$. Let μ be a minimizing measure with trace ν, and $\tilde{\mu}$ any probability measure. Then

$$\int Ld\mu \leq \int L + A^v u_\epsilon d\tilde{\mu} - \int Bu_\epsilon d\nu + O(\epsilon).$$

Proof. Since pointwise $L + A^v u_\epsilon \geq -\mathcal{H}(u_\epsilon, x)$, for any probability measure $\tilde{\mu}$

$$\int L + A^v u_\epsilon d\tilde{\mu} \geq -\mathcal{H}(u_\epsilon, x) \geq -\lambda + O(\epsilon).$$

Furthermore,

$$\int B u_\epsilon d\nu = \int B u d\nu + O(\epsilon).$$

Thus we conclude that

$$\int L + A^v u_\epsilon d\tilde{\mu} - \int B u_\epsilon d\nu \geq -\lambda + \int B u d\nu + O(\epsilon) = \int L d\mu + O(\epsilon).$$

□

8.6 Support of generalized Mather measures

The next result concerns the approximation of the support of minimizing measures.

Theorem 97. *Suppose that for all sufficiently small ϵ there exists a C^k function u_ϵ that satisfies*

$$\mathcal{H}(u_\epsilon, x) \leq \lambda + O(\epsilon).$$

Assume further that U is convex and that

$$L(x, v) + A^v u_\epsilon$$

is strictly convex in v. Let

$$v_\epsilon(x) = \operatorname{argmin} L(x, v) + A^v u_\epsilon, \tag{8.47}$$

and let μ be a minimizing measure. Then

$$\int |v - v_\epsilon(x)|^2 d\mu = O(\epsilon).$$

Remark. Since U is convex and $v \mapsto L(x,v) + A^v u_\epsilon$ is strictly convex its argmin is single valued, and thus v_ϵ is well defined.

Proof. Since v_ϵ is a minimizer, the strict convexity hypothesis implies $L(x,v) + A^v u_\epsilon \geq L(x,v_\epsilon) + A^{v_\epsilon} u_\epsilon + \theta|v - v_\epsilon|^2$. Thus

$$\begin{aligned}\int L d\mu &= \int L(x,v) + A^v u_\epsilon d\mu - \int B u_\epsilon d\nu \\ &\geq \int L(x,v_\epsilon) + A^{v_\epsilon} u_\epsilon + \theta|v - v_\epsilon|^2 d\mu - \int B u_\epsilon d\nu \\ &\geq -\lambda + O(\epsilon) + \int \theta|v - v_\epsilon|^2 d\mu - \int B u_\epsilon d\nu \\ &\geq O(\epsilon) + \int L d\mu + \theta \int |v - v_\epsilon|^2 d\mu,\end{aligned}$$

where in the last inequality we have used Lemma 96. □

Corollary 98. *Under the hypothesis of the previous theorem, let $\lambda \in \mathbb{R}$, and suppose that there exists a C^k solution $u : \mathbb{T}^n \to \mathbb{R}$ to*

$$\mathcal{H}(u,x) = \lambda,$$

and let μ be a corresponding Mather measure. Then

$$v \in \operatorname{argmin}\left[A^v u + L(x,v)\right]$$

μ almost everywhere.

8.7 Perturbation problems

This last section of this chapter is dedicated to the study of perturbations of the generalized Mather problem and its applications to the study of singular perturbations for viscosity solutions of Hamilton-Jacobi equations. In particular, we would like to understand which are the possible limits of regularizations of the Hamilton-Jacobi equation in situations where there is no uniqueness of solution.

8.7.1 Regular perturbation problems

In this section we establish a selection criterion for certain problems in which the perturbation arises in the holonomy constraint. One of the main applications is the study of vanishing discount rate problem for Hamilton-Jacobi equations, see Corollary 102.

Assume that

$$L_\epsilon(x,v) = L_0(x,v) + \epsilon L_1(x,v), \tag{8.48}$$

satisfying the hypothesis in the previous section, uniformly in ϵ, and that, additionally,

$$0 \leq L_1 \leq C L_0,$$

for some $C > 0$.

The linear operators that we consider have the form

$$A^v_\epsilon \varphi = A^v_0 \phi + \epsilon A^v_1 \varphi, \tag{8.49}$$

with corresponding boundary operators

$$B_\epsilon \varphi = B_0 \phi + \epsilon B_1 \varphi. \tag{8.50}$$

We say that the perturbation terms A^v_1 and B_1 are regular, if for any sequence φ_ϵ converging uniformly to φ we have

$$\int A^v_1 \varphi_\epsilon d\mu \to \int A^v_1 \varphi d\mu,$$

and

$$\int B_1 \varphi_\epsilon d\nu \to \int B_1 \varphi d\nu.$$

Now we state our main result:

Theorem 99. *Let $L_\epsilon(x,v)$ be as in (8.48), and consider operators of the form (8.49), (8.50) such that the perturbation terms are regular.*

Fix a probability trace measure ν, and let μ_ϵ be generalized Mather measures. We further assume that there exist viscosity solutions of

$$\mathcal{H}_\epsilon(u_\epsilon, x) = \lambda_\epsilon,$$

and corresponding C^k functions v_ϵ satisfying

$$\mathcal{H}_\epsilon(v_\epsilon, x) \leq \lambda_\epsilon + O(\epsilon^2),$$

and such that, as $\epsilon \to 0$, $v_\epsilon, u_\epsilon \to u$, uniformly in $\mathbb{T}^n$, for some viscosity solution of the limiting problem. Suppose $\mu_\epsilon \rightharpoonup \mu$. Then μ is a Mather measure for the limiting problem. Furthermore, we assume that for each solution $\tilde{u}$ of the limiting problem, there are C^k functions $\tilde{u}_\epsilon$ such that

$$\mathcal{H}_0(\tilde{u}_\epsilon, x) \leq \lambda_0 + O(\epsilon^2).$$

Then, for any viscosity solution $\tilde{u}$ of the limiting problem, we have

$$\int L_1 + A_1^v \tilde{u} d\mu - \int B_1 \tilde{u} d\nu \leq \int L_1 + A_1^v u d\tilde{\mu} - \int B_1 u d\nu.$$

Proof. Since, for any minimizing measure $\tilde{\mu}$ for the limit problem we have

$$\int L_0 d\mu \leq \lim_{\epsilon\to 0} \int L_0 d\mu_\epsilon \leq \lim_{\epsilon\to 0} \int L_\epsilon d\mu_\epsilon \leq \lim_{\epsilon\to 0} \int L_\epsilon d\tilde{\mu} = \int L_0 d\tilde{\mu}.$$

Since by a simple limiting argument

$$\int A_0^v \varphi d\mu = \int B_0 \varphi d\nu,$$

it follows that μ is a minimizing measure.

Note that

$$\begin{aligned}\int L_\epsilon d\mu_\epsilon &= \int L_\epsilon + A_\epsilon^v v_\epsilon d\mu_\epsilon - \int B_\epsilon v_\epsilon d\nu \\ &\leq O(\epsilon^2) + \int L_\epsilon + A_\epsilon^v v_\epsilon d\tilde{\mu} - \int B_\epsilon v_\epsilon d\nu \\ &= \int L_0 d\tilde{\mu} + \epsilon \left[\int L_1 + A_1^v v_\epsilon d\tilde{\mu} - \int B_1 v_\epsilon d\nu \right] + O(\epsilon^2).\end{aligned}$$

Similarly,

$$\begin{aligned}\int L_0 d\tilde{\mu} &= \int L_0 + A_0^v \tilde{u}_\epsilon d\tilde{\mu} - \int B_0 \tilde{u}_\epsilon d\nu \\ &\leq O(\epsilon^2) + \int L_\epsilon + A_\epsilon^v \tilde{u}_\epsilon d\mu_\epsilon - \int B_\epsilon \tilde{u}_\epsilon d\nu \\ &\quad - \epsilon \int L_1 + A_1^v \tilde{u}_\epsilon d\mu_\epsilon + \epsilon \int B_1 \tilde{u}_\epsilon d\nu \\ &= \int L_\epsilon d\mu_\epsilon - \epsilon \left[\int L_1 + A_1^v \tilde{u}_\epsilon d\mu_\epsilon - \int B_1 \tilde{u}_\epsilon d\nu \right] + O(\epsilon^2).\end{aligned}$$

Thus we conclude that

$$\int L_1 + A_1^v \tilde{u}_\epsilon d\mu_\epsilon - \int B_1 \tilde{u}_\epsilon d\nu \leq \int L_1 + A_1^v v_\epsilon d\tilde{\mu} - \int B_1 v_\epsilon d\nu + O(\epsilon), \tag{8.51}$$

and then the result in the theorem follows from sending $\epsilon \to 0$. ☐

Two elementary corollaries to the previous theorem are:

Corollary 100. *Suppose* $A_1^v = 0$, $B_1 = 0$, *and*

$$L_\epsilon(x, v) = \frac{|v|^2}{2} + P \cdot v + \epsilon U(x).$$

If P *is rationally dependent then the Mather measures at* $\epsilon = 0$, *are not unique, as their ergodic components are supported on lower* *dimensional tori or periodic orbits. In this case, the limiting Mather measure minimizes*

$$\int U(x)d\mu,$$

among all possible Mather measures.

Corollary 101. *Suppose $A_1^v = 0$, $B_1 = 0$. Let $P \in \mathbb{R}^n$, and*

$$L_\epsilon(x, v) = L_0(x, v) + \epsilon P \cdot v.$$

Assume that at $\epsilon = 0$ there are Mather measures with different rotation numbers

$$Q[\mu] = \int v d\mu.$$

Then the limiting measure minimizes the functional

$$P \cdot Q[\mu],$$

among all possible Mather measures.

Also as a further corollary to the previous theorem we have the following selection criterion for the discounted Mather measure problem:

Corollary 102. *Suppose u_ϵ is the unique viscosity solution to the Hamilton-Jacobi equation*

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

Consider a probability trace measure ν on $\mathbb{T}^n$ and the corresponding discounted Mather measure μ_ϵ. Let $\langle f \rangle$ denote

$$\langle f \rangle = f - \int f d\nu.$$

Suppose $\langle u_\epsilon \rangle \to u$ and $\mu_\epsilon \rightharpoonup \mu$. Let $\tilde{u}$ and $\tilde{\mu}$ be, respectively, any viscosity solution or Mather measure for the $\epsilon = 0$ problem. Then

$$\int \langle u \rangle d\tilde{\mu} \leq \int \langle \tilde{u} \rangle d\mu.$$

As an example, consider a one dimensional Hamiltonian

$$H(p,x) = \frac{p^2}{2} + U(x),$$

with the potential $U(x)$, $1/2$ periodic. Suppose further that the potential has a maximum at 0. In this case, one can verify directly that the viscosity solutions u_ϵ to the discounted problem are also 1/2-periodic. However, when $\epsilon = 0$, there are stationary solutions which are 1-periodic, and not 1/2-periodic.

Consider the Mather measure $\mu = \frac{1}{2}\left[\delta_0(x) + \delta_{1/2}(x)\right]\delta(v)$. It is easy to see that μ is a Mather measure. Consider now the trace measure $d\nu = dx$, and corresponding Mather measures μ_ϵ. Let u be the unique 1/2 periodic solution, which is given by

$$u(x) = \begin{cases} u(x) = \frac{1-\cos(2\pi x)}{\pi} & \text{for } x \in [0,1/4] \cup [3/4,1] \\ u(x) = \frac{1+\cos(2\pi x)}{\pi} & \text{for } x \in [1/4,3/4]. \end{cases}$$

We also have that $\tilde{\mu} = \left[\lambda\delta_0(x) + (1-\lambda)\delta_{1/2}\right]\delta(v)$ is a Mather measure. Thus, for any viscosity solution $\tilde{u}$, we have

$$\lambda u(0) + (1-\lambda)u(1/2) - \int u(x)dx \leq \frac{\tilde{u}(0) + \tilde{u}(1/2)}{2} - \int \tilde{u}(x)dx. \tag{8.52}$$

In the case of $U(x) = \cos(4\pi x)$, if we choose $\tilde{u}$ to be the only C^1 solution, we have,

$$D\tilde{u} = \begin{cases} \sqrt{2(1-\cos(4\pi x))} & 0 < x < 1/2 \\ -\sqrt{2(1-\cos(4\pi x))} & 1/2 < x < 1. \end{cases}$$

Then, $D\tilde{u} = 2\sin(2\pi x)$. We have $u(0) = u(1/2) = \tilde{u}(0) = 0$. A simple computation yields $\tilde{u}(1/2) = \frac{2}{\pi}$, and $\int u = \frac{\pi - 2}{\pi^2}$, $\int \tilde{u} = \frac{1}{\pi}$. Thus, the inequality (8.52) is strict and reads $\frac{2-\pi}{\pi^2} < \frac{1}{\pi} - \frac{1}{\pi} = 0$, which rules out $\tilde{u}$ as a possible limit.

As a second example, let $\psi : \mathbb{T}^n \to \mathbb{R}$ be a smooth function and consider the Lagrangian $L(x,v) = \frac{|v-D\psi(x)|^2}{4}$. The corresponding Hamiltonian is $H(p,x) = p\cdot(p - D\psi(x))$. The discounted Hamilton-Jacobi equation is then

$$\epsilon u_\epsilon + Du_\epsilon \cdot (Du_\epsilon - D\psi) = 0,$$

which has a unique solution $u_\epsilon = 0$. When $\epsilon = 0$ there are several solutions, for instance, $u = 0$ and $\tilde{u} = \psi$. Our selection criterion reads then

$$\int \langle 0 \rangle d\tilde{\mu} \leq \int \langle \psi \rangle d\mu.$$

The minimizing measures when $\epsilon = 0$ are supported at the critical points of ψ. This criterion rules out the possibility of these measures being supported at the maximizers of ψ, unless ν is itself supported there (in which case one would have $\langle \psi \rangle \geq 0$ everywhere).

8.7.2 Vanishing viscosity problems

In many important problems such as the vanishing viscosity problem [AIPSM04] the perturbations are not regular. However, the other hypothesis on theorem 99, namely, the existence of viscosity solutions, approximate supersolutions and convergence of these to the corresponding solutions of the limiting problem still holds. Therefore we still have inequality (8.51). Thus the main problem consists studying the limit $\epsilon \to 0$ in (8.51).

In the vanishing viscosity problem we have $B \equiv 0$, and

$$A^v_\epsilon \varphi = vD_x\varphi + \epsilon\Delta\varphi.$$

Formally, as $\epsilon \to 0$ we obtain

$$\int \Delta\tilde{u} d\mu \leq \int \Delta u d\tilde{\mu}.$$

In general, the previous inequality may not make sense. However, in some examples is still possible to make the proof of theorem 99 go through. As an example, consider the case in which

$$\mathcal{H}_\epsilon(u,x) = -\epsilon\Delta u + D_x u \cdot (D_x u - D_x\psi),$$

where ψ is an arbitrary smooth function in $\mathbb{T}^n$.

When $\epsilon = 0$, the Hamilton-Jacobi equation

$$D_x u \cdot (D_x u - D_x\psi) = 0$$

has two solutions (up to constants) $u = 0$ and $u = \psi$. However, for $\epsilon > 0$ there exists only one solution of

$$-\epsilon\Delta u_\epsilon + D_x u_\epsilon \cdot (D_x u_\epsilon - D_x\psi) = 0,$$

which is (up to constants) $u_\epsilon = 0$. Clearly, in this case the proof of 99 goes through as its is possible to take the limit as $\epsilon \to 0$ in (8.51). Thus any limiting Mather measure will satisfy

$$\int \Delta\psi d\mu \leq 0. \tag{8.53}$$

There is a nice interpretation of this result, that we describe next. The Lagrangian for this system is simply

$$L(x,v) = \frac{|v - D_x\psi|^2}{4}.$$

Therefore, when $\epsilon = 0$ the minimizing measures are invariant measures with respect to the gradient flow

$$\dot{\mathbf{x}} = D_x\psi(\mathbf{x}),$$

and therefore it should be supported in the critical points of ψ. Equation (8.53) means that, in average, these points should be maximizers. In fact, this can be proved directly, see [AIPSM04] for a different proof technique and related results, or by observing that in this case the projection in the x coordinate of the stochastic Mather measure has density $\theta(x) = e^{-\frac{\psi}{\epsilon}}$.

8.8 Bibliographical notes

In what concerns Mather measures the main references are, of course, the papers by Mather [Mat91] and Mañe [Mn96]. The connection of Mather measures with viscosity solutions was first observed by A. Fathi [Fat97a, Fat97b, Fat98a, Fat98b], and subsequent papers. Some PDE methods using viscosity solution techniques were first introduced in [EG01] and [EG02]. The reader is also advised to look at the forthcoming book by A. Fathi [Fat], as well as [CI99]. Generalized Mather measures were discussed in [Gom08] as a key tool to understand perturbation problems. Different techniques are also discussed in [AIPSM04]. An important problem related to Aubry-Mather theory is optimal transportation. A key reference for this problem is the book [Vil03].

9

Monotone semigroups

In this last chapter we summarize without proof some results concerning a viscosity solution characterization of monotone semigroups.

9.1 Monotone semigroups

We will follow here the approach in [Bit01], which simplifies the original work [AGLM93].

Assume we are given a family of operators $\mathcal{T}_t$ $(t \geq 0)$ mapping a subset X of continuous functions in $\mathbb{R}^n$ to itself.

We suppose that $\mathcal{T}_t$ is monotone in the following sense: if $f \leq g$ then, for any t we have

$$\mathcal{T}_t f \leq \mathcal{T}_t g.$$

We suppose also that $\mathcal{T}_t$ satisfies the following semigroup property: $\mathcal{T}_{t_1} \circ \mathcal{T}_{t_2} = \mathcal{T}_{t_1+t_2}$, for all $t_1, t_2 \geq 0$.

The translation operator τ is defined by $\tau_y \psi(x) = \psi(x+y)$, for any $\psi : \mathbb{R}^n \to \mathbb{R}$. We assume that the domain of functions X where $\mathbb{T}$ is defined satisfies the following hypothesis:

H1 $C_c^\infty(\mathbb{R}^n) \subset X$;

H2 for all $f \in X$ and all $y \in \mathbb{R}^n$, $\tau_y f \in X$.

H3 for any $f \in X$ there exists $\tilde{f} \in C^\infty \cap X$ such that $f \leq \tilde{f}$.

To describe the hypothesis that the operator will satisfy, we will need the following notation: for any sequence $d = (d_k)$ of positive reals we define

$$Q_d = \{\eta \in C_c^\infty(\mathbb{R}^n), \|D^\alpha \eta\|_\infty \leq d_k, |\alpha| \leq k\}.$$

I Continuity: for every $\psi \in X$ the function $(t, x) \mapsto \mathcal{T}_t[\psi](x)$ is continuous and for all $b > a \geq 0$ there exists $C = C(a, b, \psi)$ such that
$$|\mathcal{T}_t \psi| \leq C,$$
for any $t \in [a, b]$.

II Locality: for every $\psi_1, \psi_2 \in C^\infty(\mathbb{R}^n) \cap X$ and any fixed $x \in \mathbb{R}^n$, and $r > 0$, such that $\psi_1 = \psi_2$ in the ball $B(x, r)$ then
$$\mathcal{T}_{t-h,t}\psi_1 - \mathcal{T}_{t-h,t}\psi_2 = o(h),$$
as $h \to 0^+$.

III Regularity: for any sequence of positive numbers $d = (d_k)$, any compact set $K \subset \mathbb{R}^n$ and for every $\psi \in C^\infty(\mathbb{R}^n) \cap X$ there exists a function $m_{K,f,d}(\cdot) : \mathbb{R}^+ \to \mathbb{R}^+$ such that $m_{K,f,d}(0^+) = 0$,

$$|\mathcal{T}_t[\psi + \lambda\eta] - \mathcal{T}_t[\psi] - \lambda\eta(x)| \leq m_{K,f,d}(\lambda)t,$$

for any $(x, \eta) \in K \times Q_d$ and any $\lambda, t \geq 0$.

IV Translation: for any compact subset $K \subset \mathbb{R}^n$ and every $\psi \in C_c^\infty(K)$, there exists a function $n_{K,\psi} : \mathbb{R}^+ \to \mathbb{R}^+$, with $n_{K,\psi}(0^+) = 0$ such that

$$|\tau_y \mathcal{T}_t[\psi](x) - \mathcal{T}_t[\tau_y \psi](x)| \leq n_{K,\psi}(|y|)t,$$

for any $x \in K$ and $t \geq 0$.

Theorem 103. *Let X be a subspace of $C(\mathbb{R}^n)$ for which (H1)-(H3) hold. Let $\mathcal{T}_t$ be a monotone semigroup satisfying (I)-(IV). Then there exists a continuous function $F : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^{n \times n} \to \mathbb{R}$ such that, for all $f \in X$, $u(x,t) = \mathcal{T}_t f$ is a continuous viscosity solution to*

$$u_t + F(x, u, Du, D^2u) = 0,$$

with $u(x, 0) = f(x)$.

9.2 Bibliographical notes

The main reference for this chapter is the paper [Bit01], in addition to the paper [AGLM93].

Bibliography

[AGLM93] Luis Alvarez, Frédéric Guichard, Pierre-Louis Lions, and Jean-Michel Morel. Axioms and fundamental equations of image processing. *Arch. Rational Mech. Anal.*, 123(3):199–257, 1993.

[AIPSM04] N. Anantaraman, R. Iturriaga, P. Padilla, and H. Sanchez-Morgado. Physical solutions of the Hamilton-Jacobi equation. *preprint*, 2004.

[AKN97] V. I. Arnold, V. V. Kozlov, and A. I. Neishtadt. *Mathematical aspects of classical and celestial mechanics.* Springer-Verlag, Berlin, 1997. Translated from the 1985 Russian original by A. Iacob, Reprint of the original English edition from the series Encyclopaedia of Mathematical Sciences [*Dynamical systems. III*, Encyclopaedia Math. Sci., 3, Springer, Berlin, 1993; MR 95d:58043a].

[Arn95] V. I. Arnold. *Mathematical methods of classical mechanics.* Springer-Verlag, New York, 1995. Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein, Corrected reprint of the second (1989) edition.

[Bar94] Guy Barles. *Solutions de viscosité des équations de Hamilton-Jacobi.* Springer-Verlag, Paris, 1994.

[BCD97] Martino Bardi and Italo Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.

[BD98] A. Braides and A. Defranceschi. *Homogenization of Multiple Integrals.* Oxford Univ. Press, 1998.

[Bit01] Samuel Biton. Nonlinear monotone semigroups and viscosity solutions. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 18(3):383–402, 2001.

[Bol61] Oskar Bolza. *Lectures on the calculus of variations.* 2nd ed. Chelsea Publishing Co., New York, 1961.

[CI99] Gonzalo Contreras and Renato Iturriaga. *Global minimizers of autonomous Lagrangians.* 22^{o} Colóquio Brasileiro de Matemática. [22nd Brazilian Mathematics Colloquium]. Instituto de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 1999.

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

[Con96] Marie C. Concordel. Periodic homogenization of Hamilton-Jacobi equations: additive eigenvalues and variational formula. *Indiana Univ. Math. J.*, 45(4):1095–1117, 1996.

[Con97] Marie C. Concordel. Periodic homogenisation of Hamilton-Jacobi equations. II. Eikonal equations. *Proc. Roy. Soc. Edinburgh Sect. A*, 127(4):665–689, 1997.

[Dac09] Bernard Dacorogna. *Introduction to the calculus of variations*. Imperial College Press, London, second edition, 2009. Translated from the 1992 French original.

[dC79] Manfredo Perdigão do Carmo. *Geometria riemanniana*, volume 10 of *Projeto Euclides [Euclid Project]*. Instituto de Matemática Pura e Aplicada, Rio de Janeiro, 1979.

[dC92] Manfredo Perdigão do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

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

[EG01] L. C. Evans and D. Gomes. Effective Hamiltonians and averaging for Hamiltonian dynamics. I. *Arch. Ration. Mech. Anal.*, 157(1):1–33, 2001.

[EG02] L. C. Evans and D. Gomes. Effective Hamiltonians and averaging for Hamiltonian dynamics. II. *Arch. Ration. Mech. Anal.*, 161(4):271–305, 2002.

[Eva98a] L. C. Evans. *Partial differential equations*. American Mathematical Society, Providence, RI, 1998.

[Eva98b] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1998.

[Fat] Albert Fathi. *Weak Kam Theorem in Lagrangian Dynamics*. To appear.

[Fat97a] A. Fathi. Solutions KAM faibles conjuguées et barrières de Peierls. *C. R. Acad. Sci. Paris Sér. I Math.*, 325(6):649–652, 1997.

[Fat97b] A. Fathi. Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(9):1043–1046, 1997.

[Fat98a] A. Fathi. Orbite hétéroclines et ensemble de Peierls. *C. R. Acad. Sci. Paris Sér. I Math.*, 326:1213–1216, 1998.

[Fat98b] A. Fathi. Sur la convergence du semi-groupe de Lax-Oleinik. *C. R. Acad. Sci. Paris Sér. I Math.*, 327:267–270, 1998.

[FR75] Wendell H. Fleming and Raymond W. Rishel. *Deterministic and stochastic optimal control.* Springer-Verlag, Berlin, 1975. Applications of Mathematics, No. 1.

[FS06] Wendell H. Fleming and H. Mete Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability.* Springer, New York, second edition, 2006.

[Gol80] H. Goldstein. *Classical mechanics.* Addison-Wesley Publishing Co., Reading, Mass., second edition, 1980. Addison-Wesley Series in Physics.

[Gom02] D. Gomes. A stochastic analogue of Aubry-Mather theory. *Nonlinearity*, 15(3):581–603, 2002.

[Gom08] Diogo Aguiar Gomes. Generalized Mather problem and selection principles for viscosity solutions and Mather measures. *Adv. Calc. Var.*, 1(3):291–307, 2008.

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

[Lio82] Pierre-Louis Lions. *Generalized solutions of Hamilton-Jacobi equations*. Pitman (Advanced Publishing Program), Boston, Mass., 1982.

[LL76] L. D. Landau and E. M. Lifshitz. *Course of theoretical physics. Vol. 1*. Pergamon Press, Oxford, third edition, 1976. Mechanics, Translated from the Russian by J. B. Skyes and J. S. Bell.

[LLF] Helena Nussenzveig Lopes and Milton Lopes Filho. *Uma introdução a soluções de viscosidade para equações de Hamilton-Jacobi*.

[Lop06] Artur O. Lopes. *Introdução à mecânica clássica*. Ed. USP, São Paulo, 2006.

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

[LS03] Pierre-Louis Lions and Panagiotis E. Souganidis. Correctors for the homogenization of Hamilton-Jacobi equations in the stationary ergodic setting. *Comm. Pure Appl. Math.*, 56(10):1501–1524, 2003.

[Mat91] J. N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.*, 207(2):169–207, 1991.

[Mn96] Ricardo Mañé. Generic properties and problems of minimizing measures of Lagrangian systems. *Nonlinearity*, 9(2):273–310, 1996.

[Oli02] Waldyr Muniz Oliva. *Geometric mechanics*, volume 1798 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002.

[Olv93] Peter J. Olver. *Applications of Lie groups to differential equations*, volume 107 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1993.

[Sou99] P. E. Souganidis. Stochastic homogenization of Hamilton-Jacobi equations and some applications. *Asymptot. Anal.*, 20(1):1–11, 1999.

[Tou] N. Touzi. *Stochastic control and application to Finance*. Scuola Normale Superiore, Pisa. Special Research Semester on Financial Mathematics.

[Vil03] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.