Hamilton-Jacobi Equations, Viscosity Solutions and Asymptotics of Hamiltonian Systems

by

Diogo Aguiar Gomes

B.Sc. in Physics (Instituto Superior Técnico, Lisbon, Portugal) 1995

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Mathematics

in the

GRADUATE DIVISION
of the
UNIVERSITY of CALIFORNIA at BERKELEY

Committee in charge:

Professor Lawrence C. Evans, Chair
Professor Fraydoun Rezakhanlou
Professor Robert Littlejohn

Spring 2000

The dissertation of Diogo Aguiar Gomes is approved:

Chair Date

Date

Date

University of California at Berkeley

Spring 2000

Hamilton-Jacobi Equations, Viscosity Solutions and Asymptotics of Hamiltonian Systems

Abstract

Hamilton-Jacobi Equations, Viscosity Solutions and Asymptotics of Hamiltonian Systems

by

Diogo Aguiar Gomes
Doctor of Philosophy in Mathematics

University of California at Berkeley

Professor Lawrence C. Evans, Chair

The objective of this dissertation is to understand the relations between Hamiltonian dynamics and viscosity solutions of Hamilton-Jacobi equations. By combining ideas from classical mechanics with viscosity solution techniques we study the asymptotic behavior and invariant sets of Hamiltonian systems. Then we consider the problem of slowly varying Hamiltonians and provide a weak interpretation of both the adiabatic invariance of action and the Hannay angle. We apply measure and ergodic theory tools to characterize fine properties of Hamilton-Jacobi partial differential equations and show how ergodic properties of the Hamiltonian dynamics control the regularity of viscosity solutions. In particular, we prove a (sharp) partial uniqueness result for the time independent case, study the regularity of solutions and prove several estimates on difference quotients. Finally, we prove that the dual of a certain infinite dimensional linear programming problem, that is the core of Aubry-Mather theory, is equivalent to determining viscosity solutions of Hamilton-Jacobi equations.

To Alexandra.

Acknowledgements

My advisor, Professor Lawrence C. Evans, deserves special thanks by his guidance and enthusiasm about my research. Many parts of this thesis were discussed during our weekly meetings having benefited enormously from his suggestions and advice.

I am grateful to Professor Waldyr Oliva, who initiated me to research in Hamiltonian systems, by his support, encouragement and many comments and suggestions about my work.

Part of the last chapter of this dissertation was written at the University of Rome I, La Sapienza, while participating in the TMR program on viscosity solutions and applications. Besides having access to an extrordinary library, I beneficiated enormously from the knowledge and expertise of Professors M. Bardi (from U. of Padova), C. Dolcetta, and A. Siconolfi.

I have greatly profited from the comments and suggestions of many of my colleagues and friends, in particular L. Barreira, H. Bursztyn, J. Colliander, P. Duarte, D. Markiewicz, and R. L. Fernandes. I am also grateful to all the community of the Mathematics Department at Berkeley as well as at the Instituto Superior Técnico.

I have been supported by the program PRAXIS XXI while on leave from the Mathematics Department of Instituto Superior Técnico.

Chapter 1

Introduction

The objective of this dissertation is to understand the relations between Hamiltonian dynamics and viscosity solutions of Hamilton-Jacobi equations. By combining ideas from classical mechanics with viscosity solutions and control theory techniques we study the asymptotic behavior and invariant sets of Hamiltonian systems. Furthermore, we show how ergodic properties of the Hamiltonian dynamics control the regularity of viscosity solutions.

Building upon the pioneering works of A. Fathi [Fat97a], [Fat97b], [Fat98a], and [Fat98b], and W. E [E99], we recover and extend several results from Aubry-Mather theory [Mat89a], [Mat89b], [Mat91], [Mn92], [Mn96], using partial differential equations methods. Then we consider the problem of slowly varying Hamiltonians and provide a weak interpretation of both the adiabatic invariance of action and the Hannay angle. Additionally we apply measure and ergodic theory tools to characterize fine properties of Hamilton-Jacobi partial differential equations. In particular, we prove a (sharp) partial uniqueness result for the time independent case, study the regularity of solutions and prove several estimates on difference quotients (for different proofs, extensions of these results and applications consult [EG99a]). Finally, we show that the dual of a certain infinite dimensional linear programming problem, that is the core of Aubry-Mather theory, is equivalent to determining viscosity solutions of Hamilton-Jacobi equations.

In the chapters 2 to 4 we discuss background material: optimal control theory (chapter 2), classical mechanics and KAM theory (chapter 3), and homogenization of Hamilton-Jacobi equations (chapter 4). Then in the next two chapters we present the new results that we describe briefly bellow.

Classical Theory

Let $H = H(p, x)$ be smooth and periodic in $x \in \mathbb{R}$. Suppose that, for each P, there exists a constant $\overline{H}(P)$ and a function $u(x, P)$ (u periodic in x) solving a time independent Hamilton-Jacobi equation

$$H(P + D_x u, x) = \overline{H}(P). \tag{1.1}$$

Moreover, assume that both $\overline{H}(P)$ and $u(x, P)$ are smooth functions. Then, if the system of equations

$$p = P + D_x u \qquad X = x + D_P u \tag{1.2}$$

defines a smooth change of coordinates $X(x, p)$ and $P(x, p)$, the Hamilton equations

$$\dot{x} = D_p H \qquad \dot{p} = -D_x H, \tag{1.3}$$

can be rewritten as

$$\dot{X} = D_P \overline{H} \qquad \dot{P} = 0.$$

This last problem is trivial, thus the solution of the differential equation (1.3) is reduced to inverting the change of coordinates (1.2).

In the general case, the classical theory is unsatisfactory. Firstly, it is well known that the non-linear partial differential equation (1.1) need not have a smooth solution (smooth in x, for each P). Secondly, for fixed P there is not uniqueness, not even up to constants. Therefore, in general, smoothness in P cannot be guaranteed. Finally, even in the case in which u and $\overline{H}$ are smooth, the system of equations (1.2) may not have a solution or the functions $X(x, p)$ and $P(x, p)$ may not be smooth or globally defined.

Viscosity Solutions

In general, equation (1.1) does not admit smooth solutions. Consequently, we have to introduce appropriate weak solutions - viscosity solutions. A bounded uniformly continuous function $u(x)$ is a *viscosity solution* of (1.1) if for any smooth function $\varphi(x)$ such that $u - \varphi$ has a local maximum (respectively, local minimum) at some point x_0 then $H(P + D_x\varphi(x_0), x_0) \leq \overline{H}(P)$ (respectively, $\geq \overline{H}(P)$).

The existence of both $\overline{H}(P)$ and $u(x, P)$ viscosity solution of (1.1) was established by Lions, Papanicolaou and Varadhan [LPV88] in their study of homogenization for

Hamilton-Jacobi equations. Equation (1.1) was considered in order to study the limit of solutions V^ϵ of the terminal value problem

$$-V_t^\epsilon + H(D_x V^\epsilon, \frac{x}{\epsilon}) = 0, \quad V^\epsilon(x,1) = g(x), \quad \text{as } \epsilon \to 0.$$

It turns out that V^ϵ converges uniformly to a function V that solves

$$-V_t + \overline{H}(D_x V) = 0, \quad V(x,1) = g(x),$$

$\overline{H}(P)$ being characterized by:

Theorem 1 (Lions, Papanicolaou, Varadhan). *The function $\overline{H}(P)$ is the unique number λ for which the equation*

$$H(P + D_x u, x) = \lambda$$

has a periodic viscosity solution. Furthermore $\overline{H}(P)$ is convex in P.

Aubry-Mather Theory

J. Mather in [Mat89a], [Mat89b], and [Mat91] considered the problem of minimizing the functional

$$A[\mu] = \int L d\mu,$$

over the set of probability measures μ supported on $\mathbb{T}^n \times \mathbb{R}^n$ that are invariant under the flow associated with the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial L}{\partial v} - \frac{\partial L}{\partial x} = 0.$$

Here $L = L(v, x)$ is the Legendre transform of H, and $\mathbb{T}^n$ the n-dimensional torus, identified with $\mathbb{R}^n/\mathbb{Z}^n$ whenever convenient.

The main idea is to replace invariant sets by invariant probability measures. The supports of such measures are the analogs of the invariant sets defined by $P = constant$ given by the classical theory.

Theorem 2 (J. Mather). *For each rotation vector $\omega \in \mathbb{R}^n$ there exists an invariant probability measure μ supported on $\mathbb{T}^n \times \mathbb{R}^n$ such that*

$$\int v d\mu = \omega,$$

where (x, v) is the generic point on $\mathbb{T}^n \times \mathbb{R}^n$, and

$$\min \int L d\nu = \int L d\mu,$$

where the minimum is taken over all probability measures ν, invariant under the Euler-Lagrange equation and with $\int v d\nu = \omega$. Moreover, μ is supported on a Lipschitz graph $v = v(x)$. Furthermore, there exists a vector P such that

$$\min \int (L + Pv) d\eta = \int (L + Pv) d\mu,$$

where the minimum is taken over all probability measures η, invariant under the Euler-Lagrange equation.

The connection between Aubry-Mather theory and viscosity solutions of Hamilton-Jacobi equations was discovered, independently, by A. Fathi [Fat97a], [Fat97b], [Fat98a], [Fat98b], and W. E [E99]. Their main result is:

Theorem 3 (A. Fathi, W. E). *Suppose $u(x, P)$ is a viscosity solution of equation (1.1). Then there exists an invariant set $\mathcal{I}$ contained on the graph $\{(x, P + D_x u(x, P))\}$. Furthermore, $\mathcal{I}$ is a subset of a Lipschitz graph, i.e. $D_x u(x, P)$ is a Lipschitz function on $\pi(\mathcal{I})$, where $\pi(x, p) = x$ is the projection in the spatial coordinate. If $\overline{H}$ is differentiable at P, then any solution $(x(t), p(t))$ of (1.3) with initial conditions on $\mathcal{I}$ satisfies*

$$\lim_{t \to \infty} \frac{|x(t) - D_P \overline{H} t|}{t} = 0.$$

The proof that $\mathcal{I}$ is a Lipschitz graph is a modification of a related proof by J. Mather [Mat91]. In [E99] W. E considers time-dependent Hamiltonians with one degree of freedom for which more precise results can be proved.

New Results

Since $\overline{H}$ is convex, it is twice differentiable for almost every P. For such values of P, we can extend the previous theorem and give more precise asymptotics:

Theorem 4. *Suppose $(x(\cdot), p(\cdot))$ is a solution of the Hamiltonian equations with initial conditions on $\mathcal{I}$. Furthermore assume $\overline{H}$ is twice differentiable at P. Let*

$$\|x - y\| := \min_{k \in \mathbb{Z}^n} |x - y + k|,$$

i.e. the "periodic distance" between x and y. Then there exists a constant C such that

$$\frac{|x(t) - x(0) - D_P\overline{H}t|}{\sqrt{t}} \leq C\sqrt{\|x(t) - x(0)\|}.$$

If there exists a continuous function ω, with $\omega(0) = 0$, such that

$$|u(x, P) - u(x, P')| \leq \omega(|P - P'|),$$

then

$$|x(t) - x(0) - D_P\overline{H}t| \leq \min_\delta \frac{\|x(t) - x(0)\| \wedge \omega(\delta)}{\delta} + Ct\delta.$$

In particular, if $\omega(\delta) \leq C\delta$ then

$$|x(t) - x(0) - D_P\overline{H}t| \leq C.$$

Finally, if, for all $x \in \pi(\mathcal{I})$, $u(x, P)$ is uniformly differentiable in P, then

$$x(t) + D_P u(x(t), P) - x(0) - D_P u(x(0), P) - D_P\overline{H}t = 0.$$

A similar result holds for time dependent problems. We extend this theorem to slowly varying Hamiltonians $H(p, x, \epsilon t)$ and discuss the way homogenization theory for time dependent problems is related with adiabatic invariants and Hannay angles.

Then we consider a class of measures that can be obtained as weak limits of minimizing trajectories. Let $(x(t), p(t))$ be a trajectory with initial conditions on the invariant set $\mathcal{I}$. For $E \subset T^n \times \mathbb{R}^n$ define the measure

$$\mu_T(E) = \frac{1}{T}\int_0^T 1_E(x(t), p(t)).$$

Because μ_T is a probability measure and $p(t)$ is bounded, we can extract a weakly converging subsequence as $T \to \infty$ to some invariant measure μ. Since $\mathcal{I}$ is closed, this measure will be supported on $\mathcal{I}$. The relation between this measures and Mather measures is:

Theorem 5. *Suppose $\mu_T \rightharpoonup \mu$ as $T \to \infty$. Then*

$$\int (L + P \cdot v) d\mu = \inf_\nu \int (L + P \cdot v) d\nu,$$

where the infimum is taken over all probability measures that are invariant under the Euler-Lagrange equations. Consequently μ is a Mather measure.

Given a Mather measure μ, there exists a viscosity solution of (1.1) and a sequence of measures μ_T as before, such that $\mu_T \rightharpoonup \mu$.

Any Mather measure μ is supported on the graph $p = P + D_x u$, for any u, viscosity solution of (1.1).

The last part of the theorem implies that the union of the supports of all Mather measures - the Mather set - is a subset of the graph $p = P + D_x u$, for any u viscosity solution of (1.1). This statement is surprising because the solution may not be unique, even up to constants. Actually, Mather measures encode important information about the dynamics and are key tools in proving regularity results for viscosity solutions. An example of one of these results is the uniform continuity theorem:

Theorem 6. *Suppose μ is an ergodic Mather measure with $\mu|_{\operatorname{supp}(\mu)}$ uniquely ergodic with respect to the restricted flow. Assume $P_n \to P$. Then*

$$u(x, P_n) \to u(x, P)$$

uniformly on the support of $\sigma = \pi^* \mu$, provided an appropriate constant $C(P_n)$ is added to $u(x, P_n)$.

This theorem is particularly important because it shows that u is uniformly continuous in P on the support of μ. Therefore there exists $\omega(\cdot)$ such that, on $\operatorname{supp} \mu$, $|u(x, P) - u(x, P')| \leq \omega(|P - P'|)$, exactly the information needed in the statement of theorem 4 in order to get improved asymptotic estimates. The hypothesis on the theorem are sharp, in the sense that without unique ergodicity one may to construct counterexamples. In general, non uniqueness examples are obtained by combining viscosity solutions techniques with Mañe's results [Mn96].

Using difference quotients it is possible to prove additional regularity estimates on viscosity solutions:

Theorem 7. *Let u be a periodic viscosity solution of (1.1), and μ a Mather measure. Then*

$$\int |D_x u(x + y, P) - D_x u(x, P)| d\sigma = O(|y|^2),$$

with $\sigma = \pi^* \mu$. Furthermore, if $\overline{H}$ is twice differentiable at P then

$$\int |D_x u(x, P') - D_x u(x, P)|^2 d\sigma = O(|P - P'|^2).$$

Derivation and applications of such estimates are discussed in detail in a joint paper with L. C. Evans [EG99a]. Many of these results may be extended to the case of periodic time-dependent Hamiltonians, which is the subject of the second joint paper [EG99b]. Some of these results are discussed, with slightly different proofs, in chapter 6.

Finally, using Fenchel-Legendre duality, we show that Mather's problem is the dual of computing $\overline{H}$. More precisely:

Theorem 8 (Duality Formula).

$$\inf_{\substack{\mu \text{ invariant} \\ \int d\mu = 1}} \int (L + Pv) d\mu = \inf_{\substack{\varphi \in C^1, \\ \varphi \text{ periodic}}} \sup_x H(P + D_x\varphi, x) = \overline{H}(P),$$

This formula characterizes the minimum of Mather's problem and can be used to compute $\overline{H}$ numerically.

Chapter 2

Optimal Control

This chapter we review Optimal Control Theory and Viscosity Solutions for Hamilton-Jacobi-Belmann equations. Firstly we discuss the applications of Optimal Control Theory to Calculus of Variations. Secondly we derive the Hamilton-Jacobi-Belmann Equation and define viscosity solutions. Then we study an important tool to find fixed points for this equation - the additive eigenvalue problems. Finally we discuss the Infinite Horizon Discounted Cost problem.

The main references for this chapter are the books [BCD97] and [FS93]. For an introduction to Hamilton-Jacobi-Belmann equations and viscosity solutions the reader may want to consult [Lio82], [Bar94], or [Eva98]. The additive eigenvalue problems follows [Con95], [Con97], [Con96]. Other references include [CD87] and [Nus91].

2.1 Introduction to Optimal Control

The terminal cost problem in optimal control consists in minimizing a functional

$$J[u] = \int_t^{t_1} L(x, \dot{x}, s)ds + \psi(x(t_1)),$$

among all Lipschitz paths $x(\cdot)$, with initial condition $x(t) = x$ and satisfying the differential equation $\dot{x} = u$. To make explicit the dependence on the initial conditions we write $J[x, t; u]$ instead of $J[u]$.

To ensure that the minimum is achieved and is non-degenerate, we need to impose certain conditions on the *Lagrangian* L and *terminal cost* ψ.

Let $L : \mathbb{R}^{2n+1} \to \mathbb{R}$ (we write $L(x, v, t)$ with $x \in \mathbb{R}^n$, $v \in \mathbb{R}^n$ and $t \in \mathbb{R}$) be a smooth function, strictly convex in v, i.e., $D^2_{vv}L$ positive definite, satisfying the coercivity condition $\lim_{|v|\to\infty} \frac{L(x,v,t)}{|v|} = \infty$, for each (x, t); furthermore, suppose that $L(x, v, t) \geq 0$, by adding a constant, if necessary; assume also $L(x, 0, t) \leq c_1$, $|D_x L| \leq c_2 L + c_3$ for suitable constants c_1, c_2 and c_3; finally, suppose that there exists a positive function $C(R)$ such that $D^2_{xx}L \leq C(R)$, $|D_v L| \leq C(R)$ whenever $|v| \leq R$.

Example 1. Although the conditions above are somewhat technical, this class of functions is rather broad; for instance,

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

where A is a positive definite matrix and V a periodic smooth function, satisfies these hypothesis. L is the Lagrangian corresponding to a particle in a periodic time dependent potential. ◀

Assume ψ is a bounded Lipschitz function and $t_0 \leq t \leq t_1$. For each *control* $u \in L^\infty([t, t_1])$, $x \in \mathbb{R}^n$ let $x(\cdot)$ be the solution of the differential equation $\dot{x} = u$ with initial condition $x(t) = x$. Define the *cost functional* J by

$$J(x, t; u) = \int_t^{t_1} L(x(s), \dot{x}(s), t)ds + \psi(x(t_1)).$$

The integral term in the definition of J is called the running cost. Suppose we want to minimize J among all controls u in the *control space* $L^\infty[t, t_1; \mathbb{R}^n]$. Define the *value function* V by

$$V(x, t) = \inf_{u \in \mathcal{U}} J(x, t; u). \tag{2.1}$$

This function measures the smallest possible cost starting from a state x at time t. The next proposition gives some straightforward, yet useful, bounds on V.

Proposition 1. *The value function V satisfies the inequalities*

$$-\|\psi\|_\infty \leq V \leq c_1|t_1 - t| + \|\psi\|_\infty.$$

Proof. The first inequality follows from the fact that $L \geq 0$. To get the second inequality, simply observe that $V \leq J(x, t; 0) \leq c_1|t_1 - t| + \|\psi\|_\infty$. ■

For a general Lagrangian the infimum may not be attained by any control u. However, with the previous hypothesis on L and ψ, it is possible to prove that this infimum

is indeed a minimum. In the next example we prove this directly. The general case is the main theorem in this section.

Example 2 (Lax-Hopf formula). Suppose $L = L(v)$. By Jensen's inequality

$$\frac{1}{t_1 - t}\int_t^{t_1} L(\dot{x}(s)) \geq L\left(\frac{1}{t_1 - t}\int_t^{t_1} \dot{x}(s)\right) = L\left(\frac{y - x}{t_1 - t}\right),$$

where $y = x(t_1)$. Therefore it suffices to consider constant controls of the form $u(s) = \frac{y-x}{t_1-t}$. Hence

$$V(x,t) = \inf_{y\in\mathbb{R}^n}\left[(t_1 - t)L\left(\frac{y - x}{t_1 - t}\right) + \psi(y)\right],$$

and this inf is clearly a minimum. ◀

The *Dynamic Programming Principle* states that the evolution operator for the value function V satisfies a semigroup property.

Theorem 9 (Dynamic Programming Principle). *Suppose* $t_0 \leq t \leq t' \leq t_1$. *Then*

$$V(x,t) = \inf_{y\in\mathbb{R}^n}\inf_{x(\cdot)}\left[\int_t^{t'} L(x(s),\dot{x}(s),s)ds + V(y,t')\right], \tag{2.2}$$

where $x(t) = x$ *and* $x(t') = y$.

Proof. Let $\tilde{V}(x,t)$ denote the right hand side of (2.2). For fixed $\epsilon > 0$ let $x^\epsilon(s)$ be an almost optimal trajectory corresponding to $V(x,t)$, i.e.,

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

First we claim that $\tilde{V}(x,t) \leq V(x,t)+\epsilon$. To see this we may take $x(\cdot) = x^\epsilon(\cdot)$ and $y = x^\epsilon(t')$. Then

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

Also

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

Thus $\tilde{V}(x,t) \leq J(x,t;\dot{x}^\epsilon) \leq V(x,t) + \epsilon$, as required. Since ϵ is arbitrary, $\tilde{V}(x,t) \leq V(x,t)$. To prove the other inequality we will argue by contradiction. Indeed, if $\tilde{V}(x,t) < V(x,t)$, we could choose $\epsilon > 0$ and a control $u^\sharp$ such that

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

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

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

Define a control $u^\star(s) = u^\sharp(s)$ for $s < t'$ and $u^\star(s) = u^\flat(s)$ for $t' < s$. Then

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

which is a contradiction. ■

The evolution semigroup that maps the terminal cost to the value function is monotone, i.e., if $\psi_1 \leq \psi_2$ then the corresponding value functions satisfy $V_1 \leq V_2$. More precisely:

Proposition 2. *Suppose $\psi_1(x)$ and $\psi_2(x)$ are bounded Lipschitz functions with $\psi_1 \leq \psi_2$. 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 trajectory x^ϵ such that

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

Clearly

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

and so

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

Since ϵ is arbitrary this proves the proposition. ■

An important corollary to this proposition is that the value function depends continuously, in L^∞ norm, on the terminal data (L^∞ well-posedness).

Corollary 1 (L^∞ well-posedness). *Suppose $\psi_1(x)$ and $\psi_2(x)$ are bounded Lipschitz functions. Let $V_1(x,t)$ and $V_2(x,t)$ be 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 reversing the roles of V_1 and V_2 we get the other inequality. ■

As before, suppose $L(x,v,t)$ is convex in v, and satisfies the coercivity condition $\lim_{|v|\to\infty} \frac{L(x,v,t)}{|v|} = \infty$. The *Legendre transform* of L, denoted by L^*, is the function

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

Our definition of Legendre transform agrees with the definition in [FS93] or [BCD97] and is the usual in optimal control problems but it is different from the customary in classical mechanics. The latter one is $L^\sharp(p,x,t) = \sup_v v \cdot p - L(x,v,t)$, as defined, for instance, in [Arn99] or [Eva98]. The relation between them is obvious $L^*(p,x,t) = L^\sharp(-p,x,t)$. Our definition simplifies the definition of viscosity solutions for the terminal value problem.

It is customary to denote $L^*(p,x,t)$ by $H(p,x,t)$. This last function is called the *Hamiltonian* corresponding to the Lagrangian L. This transform has some important properties:

Proposition 3. *Suppose that $L(x,v,t)$ is convex and coercive in v. Let $H = L^*$. Then*

1. *$H(p,x,t)$ is convex in p;*
2. *$H^* = L$;*
3. *For each (x,t), $\lim_{|p|\to\infty} \frac{H(x,p,t)}{|p|} = \infty$;*
4. *Define v^* by the equation $p = -D_v L(x,v^*,t)$; Then*
$$H(p,x,t) = -v^* \cdot p - L((x,v^*,t);$$
5. *Similarly define p^* by the equation $v = -D_p H(x,p^*,t)$; Then*
$$L(x,v,t) = -v \cdot p^* - H((x,v^*,t);$$
6. *If $p = -D_v L(x,v,t)$ or $v = -D_p H(x,p,t)$ then $D_x L(x,v,t) = -D_x H(p,x,t)$.*

Proof. The first fact follows from the fact that the supremum of convex functions is convex. To prove the second one note that

$$H^*(x,w,t) = \sup_p -w \cdot p - H(p,x,t) = \sup_p \inf_v (v-w) \cdot p + L(x,v,t).$$

If we choose $v = w$ we get

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

The reverse inequality is obtained by observing that, since L is convex in v, there exists s such that $L(x, v, t) \ge L(x, w, t) + s \cdot (v - w)$ and so

$$H^*(x, w, t) \ge \sup_p \inf_v (p + s) \cdot (v - w) + L(x, w, t) \ge L(x, w, t),$$

by choosing $p = -s$. To prove the third observation note that

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

by choosing $v = -\lambda \frac{p}{|p|}$. If we let $|p| \to \infty$ we conclude

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

Since λ is arbitrary we get

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

To prove the fourth part note that, for fixed p, the function $-v \cdot p - L(x, v, p)$ is smooth and strictly convex. Therefore its unique minimum (which exists by coercivity and is unique by strict convexity) is achieved when $-p - D_v L = 0$. The fifth part is similar. Finally, to prove the last part, note that if $p = -D_v L(x, v, t)$ then $H(x, p, t) = -v \cdot p - L(x, v, t)$ if we differentiate with respect to x we conclude $D_x H = -D_x L$. ■

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

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

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

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

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

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

These sets are one-sided analog of derivatives. Indeed, if ψ is differentiable

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

More precisely,

Proposition 4. *If $D_x^-\psi(x), D_x^+\psi(x) \neq \emptyset$ then $D_x^-\psi(x) = D_x^+\psi(x) = \{p\}$ and ψ is differentiable at x with $D_x\psi = p$. Conversely, if ψ is differentiable at x then*

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

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

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

By subtraction we conclude

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

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

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

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

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

and so $D_x\psi = p$.

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

Proposition 5. *Let ψ be continuous. Then, if $p \in D_x^+\psi(x_0)$ (resp. $p \in D_x^-\psi(x_0)$) there exists a C^1 function ϕ such that $\psi(x) - \phi(x)$ has a strict local maximum (resp. minimum) at x_0 and $p = D_x\phi(x_0)$. Conversely, if ϕ is a C^1 function such that $\psi(x) - \phi(x)$ has a local maximum (resp. minimum) at x_0 then $p = D_x\phi(x_0) \in D_x^+\psi(x_0)$ (resp. $D_x^-\psi(x_0)$).*

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

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

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

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

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

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

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

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

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

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

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

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

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

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

A continuous function f is said to be *semiconcave* provided there exists a constant C such that $f(x + y) + f(x - y) - 2f(x) \leq C|y|^2$. Semiconvex functions are defined analogously.

Proposition 6. *The following are equivalent:*

1. *f is semiconcave;*

2. *$\tilde{f}(x) = f(x) - C|x|^2$ is concave*

3. *if $\lambda y + (1 - \lambda)z = 0$, for $0 \leq \lambda \leq 1$, then*

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

.

Furthermore, if f is semiconcave

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

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

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

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

REMARK. A similar proposition holds for semiconvex functions.

PROOF. Clearly $2 \implies 3 \implies 1$. Thus, to prove the equivalence, it suffices to check that $1 \implies 2$. By subtracting $C|x|^2$ to f we may assume $C = 0$. By changing coordinates, it suffices to prove that that for any x, y such that $\lambda x + (1 - \lambda) y = 0$

$$\lambda f(x) + (1 - \lambda) f(y) - f(z) \le 0. \tag{2.3}$$

We claim that the previous equation holds for any $\lambda = \frac{k}{2^j}$, for any $0 \le k \le 2^j$. Clearly (2.3) holds when $j = 1$. Now we proceed with induction in j. Assume that (2.3) holds for $\lambda = \frac{k}{2^j}$. Then we claim that it holds with $\lambda = \frac{k}{2^{j+1}}$. If k is even we can reduce the fraction, therefore we may suppose that k is odd, $\lambda = \frac{k}{2^{j+1}}$ and $\lambda x + (1 - \lambda) y = 0$. Now note that

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

Thus

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

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. Hence

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

But this implies, that

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

Since $\tilde{k}_0 + \tilde{k}_1 = k$ we get

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

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

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

for any real $0 \le \lambda \le 1$.

To prove the second part of the proposition note that, by proposition 4, $a \implies b$. To check a, i.e., that $D_x^+ f(x) \neq \emptyset$, it suffices to observe that if f is convex then $D_x^+ f(x) \neq \emptyset$. By subtracting $C|x|^2$ to f, we can reduce to the convex case. Finally if $p_i \in D_x^+ f(x_i)$ $(i = 0, 1)$ then for

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

and

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

Thus

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

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

Theorem 10 (Main Theorem). *Suppose $x \in \mathbb{R}^d$ and $t_0 \leq t \leq t_1$. Assume $L(x, v, t)$ is a smooth function, strictly convex in v (i.e., $D^2_{vv} L$ positive definite), and satisfying the coercivity condition $\lim_{|v| \to \infty} \frac{L(x,v,t)}{|v|} = \infty$, for each (x, t). Furthermore suppose L bounded below (without loss of generality, we may take $L(x, v, t) \geq 0$); assume also $L(x, 0, t) \leq c_1$, $|D_x L| \leq c_2 L + c_3$ for suitable constants c_1, c_2, and c_3; finally suppose that there exists a positive functions $C_0(R), C_1(R)$ such that $|D_v L| \leq C_0(R)$ and $|D^2_{xx} L| \leq C_1(R)$ whenever $|v| \leq R$. Then for any bounded Lipschitz function ψ:*

1. *There exists a control $u^* \in L^\infty[t, t_1]$ such that the corresponding path x^*, defined by the initial value ODE*
$$\dot{x}^*(s) = u(s) \qquad x^*(t) = x,$$
satisfies
$$V(x, t) = \int_t^{t_1} L(x^*(s), \dot{x}^*(s), s) ds + \psi(x^*(t_1)).$$

2. *There exists a constant C, which depends only on L, ψ and $t_1 - t_0$ but not on x or t such that $|u(s)| < C$ for $t \leq s \leq t_1$. The optimal trajectory $x^*(\cdot)$ is a $C^2[t, t_1]$ solution of the Euler-Lagrange equation*
$$\frac{d}{dt} D_v L - D_x L = 0 \tag{2.4}$$
with initial condition $x(t) = x$.

3. *The adjoint variable P, defined by*

$$P(t) = -D_v L(x^*, \dot{x}^*, t)), \tag{2.5}$$

satisfies the differential equation

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

with terminal condition $P(t_1) \in D_x^- \psi(x^*(t_1))$. Additionally

$$(P(s), H(P(s), x^*(s), s)) \in D^- V(x^*(s), s)$$

for $t < s \leq t_1$.

4. *The value function V is Lipschitz continuous, thus differentiable almost everywhere.*

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

6. *$(P(s), H(P(s), x^*(s), s)) \in D^+ V(x^*(s), s)$ for $t \leq s < t_1$ and so $D_x V(x^*(s), s)$ exists for $t < s < t_1$.*

Proof. We divide the proof of the theorem into several lemmas.

For $R > 0$, define $\mathcal{U}_R = \{u \in \mathcal{U} : \|u\|_\infty \leq R\}$. First we prove the existence of a minimizer in $\mathcal{U}_R$. Then we give bounds on the minimizer u_R that do not depend on R. Finally we let $R \to \infty$.

Lemma 1. *J is weakly-$*$ lower semicontinuous.*

Proof. Suppose u_n is a sequence of controls such that $u_n \overset{*}{\rightharpoonup} u$ in $L^\infty[t, t_1]$. Then, by standard ODE theory, the corresponding sequence of trajectories $x_n(\cdot)$ converges uniformly to $x(\cdot)$. Thus

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

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

$$L(x(s), u_n(s), s) \geq L(x(s), u(s), s) + D_v L(x(s), u(s), s)(u_n(s) - u(s)).$$

Since $u_n \overset{*}{\rightharpoonup} u$

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

Thus

$$\liminf J(x, t; u_n) \geq J(x, t; u),$$

and so J is weakly-$*$ lower semicontinuous. ∎

Lemma 2. *There exists a minimizer u_R of J in $\mathcal{U}_R$.*

Proof. Take a minimizing sequence $u_n \in \mathcal{U}_R$ such that

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

This sequence is bounded in L^∞. Therefore by Banach-Alaoglu theorem we can extract a weakly-$*$ convergent sub-sequence $u_n \overset{*}{\rightharpoonup} u_R$. Clearly $u_R \in \mathcal{U}_R$. We claim that

$$J(x, t; u_R) = \inf_{u \in \mathcal{U}_R} J(x, t; u).$$

Indeed, by weakly-$*$ lower semicontinuity,

$$\inf_{u \in \mathcal{U}_R} J(x, t; u) \leq J(x, t; u_R) \leq \liminf J(x, t; u_n) = \inf_{u \in \mathcal{U}_R} J(x, t; u).$$

which proves the lemma. ∎

The next task is to prove bounds on the optimal control u_R that do not depend on R, so that we can let $R \to \infty$. First we assume ψ differentiable. Then we approximate ψ by smooth functions and pass to the limit.

Suppose $r \in [t, t_1)$ is a point where u_R is strongly approximately continuous, i.e.,

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

for all φ continuous. Define P_R by

$$P_R(r) = D_x\psi(x_R(t_1)) + \int_r^{t_1} D_x L(x_R(s), u_R(s), s)ds.$$

Lemma 3 (Pontryagin's maximum principle I). *Assume ψ is differentiable. Then for almost every $r \in [t, t_1)$*

$$u_R(r) \cdot P_R(r) + L(x_R(r), u_R(r), r) = \min_{|v| \leq R} \left[v \cdot P_R(r) + L(x_R(r), v, r) \right]. \tag{2.6}$$

PROOF. Choose v such that $|v| \leq R$. For $r \in [t_0, t_1)$ where u_R is strongly approximately continuous define

$$u_\delta(s) = \begin{cases} v & \text{if} \quad r < s < r + \delta \\ u_R(s) & \text{otherwise.} \end{cases}$$

Define

$$x_\delta(s) = \begin{cases} x_R(s) & \text{if} \quad t < s < r \\ x_R(r) + (s - r)v & \text{if} \quad r < s < r + \delta \\ x_R(s) + \delta\xi_\delta & \text{if} \quad r + \delta < s < t_1, \end{cases}$$

where

$$\xi_\delta = \frac{1}{\delta} \int_r^{r+\delta} [v - u_R(s)]\, ds.$$

Then

$$J(t, x; u_R) \leq J(t, x; u_\delta) = \int_t^{t_1} L(x_\delta(s), u_\delta(s), s) ds + \psi(x(t_1) + \delta\xi_\delta).$$

This inequality implies

$$\begin{aligned} &\frac{1}{\delta} \int_r^{r+\delta} [L(x_\delta(s), v, s) - L(x_R(s), u_R(s), s)]\, ds + \\ &+ \frac{1}{\delta} \int_{r+\delta}^{t_1} [L(x_R(s) + \delta\xi_\delta, u_R(s), s) - L(x_R(s), u_R(s), s)]\, ds + \\ &+ \frac{1}{\delta} [\psi(x_R(t_1) + \delta\xi_\delta) - \psi(x_R(t_1))] \geq 0. \end{aligned}$$

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

$$L(x_R(r), v, r) - L(x_R(r), u_R(r), r),$$

the second to

$$\left[\int_r^{t_1} D_x L(x_R(s), u_R(s), s) ds \right] \cdot (v - u_R(r)),$$

and the third to $D_x\psi(x_R(t_1)) \cdot (v - u_R(r))$. Thus this implies that for almost every r

$$L(x_R(r), v, r) - L(x_R(r), u_R(r), r) + P_R(r) \cdot (v - u_R(r)) \geq 0.$$

Therefore

$$u_R(r) \cdot P_R(r) + L(x_R(r), u_R(r), r) = \min_{|v| \leq R} [v \cdot P_R(r) + L(x_R(r), v, r)],$$

as required. ∎

Lemma 4. *Assume ψ is differentiable. Then there exists a constant C that does not depend on R such that $|P_R| \leq C$.*

Proof. Since ψ is Lipschitz and differentiable we have $|D_x\psi| \leq \|D_x\psi\|_\infty < \infty$. Therefore

$$|P_R(s)| \leq \int_s^{t_1} |D_x L(x_R(r), u_R(r), r| dr + \|D_x\psi\|_\infty$$

Since $|L_x| \leq c_2 L + c_3$,

$$|P_R(s)| \leq C(V_R(t,x) + 1),$$

for some appropriate constant C. We know, by proposition 1, that there exists a constant C_1, that does not depend on R, such that $V_R \leq C_1$. Therefore $|P_R| \leq C$. ∎

We will see that these bounds on P_R imply uniform bounds on u_R.

Lemma 5. *Assume ψ differentiable. There exists $R_1 > 0$ such that, for any R,*

$$\|u_R\|_\infty \leq R_1.$$

Proof. Suppose $|p| \leq C$. Then, for each c_1, the coercivity condition on L implies that there exists R_1 such that if $v \cdot p + L(x, v, s) \leq c_1$ we have $|v| \leq R_1$. But since

$$u_R(s) \cdot P_R(s) + L(x_R(s), u_R(s), s) \leq L(x_R(s), 0, s) \leq c_1,$$

we get $\|u_R\|_\infty \leq R_1$. ∎

Since u_R is bounded independently of R, we have $V = J(x, t; u_{R_0})$, for R_0 sufficiently large. Define $u^* = u_{R_0}$. Similarly define $P = P_{R_0}$.

Lemma 6 (Pontryagin's maximum principle II). *If ψ is differentiable, the optimal control u^* satisfies*

$$u^* \cdot P + L(x^*, u^*, s) = \min_v \left[v \cdot P + L(x^*, v, s)\right] = -H(P, x^*, s),$$

and so

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

where $H = L^*$. Furthermore P satisfies the terminal condition $P = D_x\psi(x^*(t_1))$.

Proof. Clearly it suffices to choose R sufficiently large. ∎

Lemma 7. *Assume ψ differentiable. The minimizing trajectory is C^2 and satisfies the Euler-Lagrange equations (2.4). Furthermore*

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

Proof. By its definition P is continuous. Since $\dot{x}^* = -D_p H(P(s), x^*, s)$, x^* is C^1. But then, since P itself is given by the integral of a continuous function,

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

we conclude that P is C^1. Additionally, $\dot{x}^* = -D_p H(P, x^*, s)$ and so $\dot{x}^*$ is C^1 which implies that x is C^2. We have

$$P = -D_v L \qquad \dot{P} = -D_x L,$$

which implies

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

Furthermore, since $D_x L(x^*, \dot{x}^*, t) = -D_x H(P, x^*, t)$, we conclude

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

as required. ∎

To take care of the general case where ψ is only Lipschitz, consider a sequence of smooth functions $\psi_n \to \psi$ uniformly, and with $\|D_x \psi_n\|_\infty \leq Lip(\psi)$. For each ψ_n let

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

Let x_n^* and u_n^* be, respectively, the optimal trajectory and control. Similarly define P_n to be the associated adjoint variable. By passing to a subsequence, if necessary, $x_n(t_1)$ and $P_n(t_1)$ converge, respectively, to some x_0 and P_0. The corresponding optimal trajectories x_n^* and optimal controls u_n^* also converge uniformly to optimal trajectories and controls for the limit problem. Define the adjoint variable $P(s) = \lim_{n \to \infty} P_n$ Then, for almost any s,

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

which implies

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

for almost every s. But, since in the previous equation both sides are continuous functions, the equality must in fact hold for all s.

Lemma 8. *For $t < s \leq t_1$, $(P(s), H(P(s), x^*(s), s)) \in D^- V(x^*(s), s)$.*

Proof. Fix an optimal trajectory x^* and associated optimal control u^*. Choose $r \leq t_1$ and $y \in \mathbb{R}^n$. Let $x_r = x^*(r)$. Define

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

such that $x^\sharp(r) = y$.

Observe that

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

and also

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

This implies that

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

with

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

Since ϕ is differentiable in y and r

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

Now note that

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

hence

$$D_y \phi(x^*(s), s) = \int_t^s D_x L \frac{\tau - t}{s - t} + D_v L \frac{1}{s - t},$$

and by integration by parts we get

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

Similarly,

$$D_r \phi(y, r) = L(x^*(s), \dot{x}^*(s), s) + \int_t^s -D_x L \frac{y - x_r}{(r - t)^2}(\tau - t) + D_x L \frac{-u^*(r)}{(r - t)}(\tau - t) - \\ - D_v L \frac{y - x_r}{(r - t)^2} + D_v L \frac{-u^*(r)}{r - t}.$$

Integrating by parts and evaluating at $y = x^*(s)$, $r = s$,

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

as required. ■

Lemma 9. *The value function V is Lipschitz continuous.*

Proof. Fix $t < t_1$ and choose x, y. Assume first that $t_1 - t < 1$. Then

$$V(y,t) - V(x,t) \leq J(y,t;u^*) - V(x,t),$$

where $V(x,t) = J(x,t;u^*)$. Thus there exists a constant C, depending on the supremum of D_xL and the Lipschitz constant of ψ such that

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

Now suppose $t_1 - t > 1$. Define $\tilde{u}(s) = u^* + (x-y)$ if $t < s < t+1$ and $\tilde{u}(s) = u^*(s)$ for $t + 1 \leq s \leq t_1$. Then

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

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

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

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

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

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

$$|V(x^*(\hat{t}), \hat{t}) - V(x, \hat{t})| \leq C|t - \hat{t}|. \tag{2.7}$$

But since $\dot{x}^*$ is uniformly bounded $|x^*(\hat{t}) - x| \leq C|t - \hat{t}|$ hence the Lipschitz property proved before implies (2.7). ■

Lemma 10. *V is differentiable almost everywhere.*

Proof. Since V is Lipschitz, differentiability almost everywhere follows from Lipschitz continuity by Rademacher theorem. ■

In general, the value function is only Lipschitz but we have a one-sided control on the second derivatives, namely semiconcavity.

Lemma 11. *Suppose $D^2_{vv}L$ is uniformly bounded. Then for each fixed $t < t_1$, $V(x,t)$ is semiconcave in x.*

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

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

for some constant C. Clearly

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

where $y(s) = y\frac{t_1-s}{t_1-t}$. Observe that

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

and similarly for the other term. Also we have

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

Thus

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

From this inequality implies the lemma. ■

Lemma 12. *We have $(P(s), H(P(s), x^*(s), s)) \in D^+V(x^*(s), s)$ for $t \leq s < t_1$. Therefore $D_xV(x^*(s), s)$ exists for $t < s < t_1$.*

Proof. Let u^* be an optimal control at (x,s) and let P be the corresponding adjoint variable. Define the function W by

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

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

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

with equality at (x,s). Since W is a continuously differentiable function, it suffices to check that

$$D_y W(s, x^*(s)) = P(s),$$

and

$$D_r W(s, x^*(s)) = H(P(s), x^*(s), s).$$

To prove the first identity note that

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

where $\varphi(\tau) = \frac{\tau - s}{t_1 - s}$. Using the Euler-Lagrange equation $\frac{d}{dt} D_v L - D_x L = 0$ and integration by parts we get

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

To prove the second identity note that

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

where

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

Again by applying the Euler-Lagrange equation and integration by parts we get

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

or equivalently

$$D_r W(s, x^*(s)) = H(P(s), x^*(s), s).$$

The final part of the lemma follows from proposition 4. ▪

This finishes the proof of this theorem. ■

Now we prove that whenever the optimal trajectory is unique the value function is differentiable along this trajectory. Consequently non-differentiability exists only when there are shocks, i.e., points where there is more than one minimizing trajectory.

A point (x, t) is called *regular* if there exists a unique trajectory $x^*(s)$ such that $x^*(t) = x$ and

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

It turns out that regularity is a necessary and sufficient condition for differentiability of the value function.

Theorem 11. *V is differentiable at (x, t) if and only if (x, t) is a regular point.*

Proof. Differentiability implies regularity is the next lemma.

Lemma 13. *If V is differentiable at a point (x, t) the optimal trajectory is unique.*

Proof. Since V is differentiable at (x, t) we have $\dot{x}^*(t) = -D_pH(x^*(t), p(t))$. Therefore, given the value of $D_xV(x^*(t), t)$, we can compute $\dot{x}^*(t)$. The solution of the Euler-Lagrange equation (2.4) is completely specified by the initial state $x^*(t)$ and velocity $\dot{x}^*(t)$. Thus it follows that the optimal trajectory is unique. ■

To prove the converse we need an auxiliary lemma.

Lemma 14. *Suppose there exists p such that $\|D_xV(\cdot, t) - p\|_{L^\infty(B(x,2\epsilon))} \to 0$ as $\epsilon \to 0$. Then V is differentiable at (x, t) and $D_xV(x, t) = p$.*

Remark. This continuity property of the derivative of the value function generalizes the regularity results from [JS87].

Proof. Since V is Lipschitz it is differentiable a.e., Fubbini's theorem implies that for almost every (with respect to Lebesgue measure on S^{n-1}) direction k, V is differentiable for almost every (with respect to Lebesgue measure on $\mathbb{R}$) $y = x + \lambda k$. For such directions

$$\frac{V(x,t) - V(y,t) - p \cdot (x-y)}{|x-y|} = \int_0^1 \frac{(D_xV(x + s(x-y), t) - p) \cdot (x-y)}{|x-y|} ds.$$

Suppose $\epsilon < |x - y| < 2\epsilon$. Then

$$\left| \frac{V(x,t) - V(y,t) - p \cdot (x-y)}{|x-y|} \right| \le \|D_xV(\cdot, t) - p\|_{L^\infty(B(x,2\epsilon))}.$$

In principle the last equality only holds almost everywhere. However the left-hand side is continuous. Therefore the inequality holds for all y. Thus when $y \to x$

$$\left| \frac{V(x,t) - V(y,t) - p \cdot (x-y)}{|x-y|} \right| \to 0$$

which implies $D_xV(x,t) = p$. ■

Suppose V is not differentiable at (x,t). We claim that (x,t) is not regular. Indeed, if V is not differentiable, the previous lemma implies that for any p,

$$\|D_xV(\cdot,t) - p\|_{L^\infty(B(x,2\epsilon))} \nrightarrow 0.$$

Thus, we can choose two sequences x_n^1 and x_n^2 such that $x_n^i \to x$ but

$$\lim\,(\dot{x}_n^1)^*(t) \neq \lim\,(\dot{x}_n^2)^*(t).$$

We claim that this implies that (x,t) is not regular. Indeed, if (x,t) is regular, x_n is any sequence converging to x, and $x_n^*(\cdot)$ an optimal trajectory corresponding to x_n then

$$\dot{x}_n^*(t) \to \dot{x}^*(t).$$

If this were not true we could extract a convergent subsequence $\dot{x}_{n_k}^*(t) \to v \neq \dot{x}^*(t)$. Let $y(\cdot)$ be a solution of the Euler-Lagrange equations with initial conditions $y(t) = x(t)$ and $\dot{y}(t) = v$. Note that $x_n^*(\cdot) \to y(\cdot)$ uniformly and so

$$V(x,t) = \lim_{n\to\infty} V(x_n,t) = \lim_{n\to\infty} J(x_n,t;\dot{x}_n) = J(x,t;\dot{y}) > J(x,t;\dot{x}^*) = V(x,t)$$

which is a contradiction. ■

Remark. This theorem implies that all points along $(x^*(s),s)$ with $t < s < t_1$ are regular.

2.2 Hamilton-Jacobi-Belmann Equation and Viscosity solutions

The dynamic programming principle (Theorem 9) has an infinitesimal counterpart - the *Hamilton-Jacobi-Belmann* (*HJB*) partial differential equation (*PDE*).

Theorem 12. *Suppose the value function V is differentiable at (x,t). Then, at this point, V solves the Hamilton-Jacobi-Belmann PDE*

$$-V_t + H(D_xV,x,t) = 0. \tag{2.8}$$

Proof. If V is differentiable at (x,t) then, by part 6 of theorem 10 we get the result. ■

Corollary 2. *The value function V solves the HJB equation a.e..*

Proof. Since the value function V is differentiable a.e. by theorem 10, theorem 12 implies the result. ■

However, it is not true that a Lipschitz function solving the HJB equation a.e. is the value function for some terminal data, as we show in the next example.

Example 3. Consider the HJB equation

$$-V_t + |D_x V|^2 = 0$$

with terminal condition $V(x, 1) = 0$. Clearly the value function is $V \equiv 0$ and it solves the HJB equation everywhere. However there are other solutions, for instance,

$$V(x,t) = \begin{cases} 0 \quad \text{if} \quad |x| \geq 1-t \\ |x| - 1 + t \quad \text{if} \quad |x| < 1-t \end{cases}$$

satisfies the same terminal condition and solves the equation a.e.. ◀

A bounded uniformly continuous function V is a *viscosity subsolution* (resp. *supersolution*) of the Hamilton-Jacobi-Belmann PDE (2.8) if for any smooth function ϕ such that $V - \phi$ has a local maximum (resp. minimum) at (x, t) then $-D_t\phi + H(D_x\phi, x) \leq 0$ (resp. ≥ 0) at (x, t). A bounded uniformly continuous function V is a *viscosity solution* of the HJB equation provided it is both a subsolution and a supersolution.

In one space dimension there is a simple criteria to decide whether a function V is or not a viscosity solution:

Proposition 7. *In one space dimension, a Lipschitz, piecewise smooth function V is a viscosity solution of (2.8) equation provided:*

1. *V solves the equation almost everywhere;*

2. *whenever $D_x V$ is discontinuous we have $D_x V(x^-, t) > D_x V(x^+, t)$.*

Although the value function is not, in general, a classical solution of (2.8) it is a viscosity solution. The idea behind this definition is that if V is smooth and $V - \phi$ has a maximum or minimum at (x, t) then $D_x V = D_x\phi$ and $V_t = \phi_t$. However, this definition makes sense even if V is not differentiable. Another motivation for the definition of viscosity solution is to consider the *parabolic approximation* of the equation (2.8)

$$-D_t u^\epsilon + H(D_x u^\epsilon, x, t) = \epsilon \Delta u^\epsilon. \tag{2.9}$$

This equation arises naturally is stochastic control theory (see [FS93]). The limit $\epsilon \to 0$ corresponds to the small noise limit.

Proposition 8. *Let u^ϵ be a family of solutions of (2.9). Suppose as $\epsilon \to 0$ the sequence $u^\epsilon \to u$ uniformly. Then u is a viscosity solution of (2.8).*

Proof. Suppose $\phi(x,t)$ is a smooth function such that $u-\phi$ has a strict local maximum at $(\overline{x},\overline{t})$. Then we must prove that

$$-D_t\phi + H(D_x\phi, \overline{x}, \overline{t}) \le 0.$$

Since, by hypothesis, $u^\epsilon \to u$ uniformly, we can find $(\overline{x}_\epsilon, \overline{t}_\epsilon) \to (\overline{x}, \overline{t})$ such that $u^\epsilon - \phi$ has a local maximum at $(\overline{x}_\epsilon, \overline{t}_\epsilon)$. Thus

$$Du^\epsilon(\overline{x}_\epsilon, \overline{t}_\epsilon) = D\phi(\overline{x}_\epsilon, \overline{t}_\epsilon)$$

and

$$\Delta u^\epsilon(\overline{x}_\epsilon, \overline{t}_\epsilon) \le \Delta\phi(\overline{x}_\epsilon, \overline{t}_\epsilon).$$

Thus

$$-D_t\phi(\overline{x}_\epsilon, \overline{t}_\epsilon) + H(D_x\phi(\overline{x}_\epsilon, \overline{t}_\epsilon), \overline{x}_\epsilon, \overline{t}_\epsilon) \le \epsilon\Delta\phi(\overline{x}_\epsilon, \overline{t}_\epsilon).$$

Thus it suffices to let $\epsilon \to 0$ to end the proof. ■

Another useful characterization of viscosity solutions is given in the next proposition.

Proposition 9. *Suppose V is a bounded uniformly continuous function. Then V is a viscosity subsolution of (2.8) if and only if for any $(p,q) \in D^+V(x,t)$, $-q + H(p,x;t) \le 0$. Similarly V is a viscosity supersolution if and only if for any $(p,q) \in D^-V(x,t)$, $-q + H(p,x;t) \ge 0$.*

Proof. This is a immediate corollary to proposition 5. ■

Example 4. In example 3 we found two solutions of the equation $-V_t + |D_xV|^2 = 0$ with the same terminal condition. It is easy to check that the value function $V = 0$ is a viscosity solution of this equation. However the other solution is not a viscosity solution; indeed take $\phi = t + \frac{x}{2}$. Then $V - \phi$ has a local minimum when $(x,t) = (1,0)$; however $-\phi_t + |D_x\phi|^2 = -\frac{3}{4} < 0$ which contradicts the definition of viscosity solution. ◄

This definition is consistent with the definition of classical solution:

Proposition 10. *A smooth viscosity solution of (2.8) is a classical solution.*

PROOF. If V is smooth then $D^+V = D^-V = \{(D_tV, D_xV)\}$. Since V is a viscosity solution we get immediately

$$-D_tV + H(D_xV, x, t) \leq 0, \quad \text{and} \quad -D_tV + H(D_xV, x, t) \geq 0,$$

thus $-D_tV + H(D_xV, x, t) = 0$. ■

We know that, in general, there are no smooth solutions of the Hamilton-Jacobi-Bellman equation. However, if we ask for solutions that solve the equation a.e., we loose uniqueness. Fortunately, viscosity solutions are not only the natural definition of solution, since the value function V is a viscosity solution, but also they are unique.

Theorem 13. *The value function V given by (2.1) is a viscosity solution of the terminal value problem for the Hamilton-Jacobi-Bellman PDE (2.8).*

PROOF. Before proving the theorem we need some definitions and a lemma. Define the action kernel $K(x, y; t, t_1)$ by the formula

$$K(x, y; t, t_1) = \inf_{x(t)=x, x(t_1)=y} \int_t^{t_1} L(x(s), \dot{x}(s), s)ds. \tag{2.10}$$

When t_1 is fixed we may write instead $K(x, y; t)$. Let T_{t,t_1} be the (nonlinear) operator defined by

$$T_{t,t_1}\psi(x) \equiv \inf_y K(x, y; t, t_1) + \psi(y).$$

This is a alternative way compute the value function associated with the terminal cost ψ.

Lemma 15. *If $\phi(x, t)$ is a smooth function then*

$$\lim_{h\to 0} \frac{T_{t,t+h}\phi(\cdot, t+h)(x) - \phi(x, t)}{h} = D_t\phi(x, t) - H(D_x\phi(x, t), x, t).$$

PROOF. It suffices to prove that, for h sufficiently small, $T_{t,t+h}\phi(\cdot, t+h)$ is a smooth function. Observe that the PDE

$$-W_t + H(D_xW, x, t) = 0,$$

with terminal condition $W(x, t+h) = \phi(x, t+h)$ has a unique classical solution (obtained using the method of characteristics), for $t \leq s \leq t+h$ and h sufficiently small (depending

on ϕ). The Pontryagin maximum principle implies that this solution is in fact the value function. The lemma follows easily using the characteristics method for this problem. ▪

Suppose ϕ is a smooth function such that $V - \phi$ has a local maximum at $(\overline{x}, \overline{t})$, with $V(\overline{x}, \overline{t}) = \phi(\overline{x}, \overline{t})$. Then $\phi \geq V$ and so

$$T_{\overline{t},\overline{t}+h}\phi(\cdot, \overline{t}+h)(\overline{x}) \geq T_{\overline{t},\overline{t}+h}V(\cdot, \overline{t}+h)(\overline{x}) = V(\overline{x}, \overline{t}) = \phi(\overline{x}, \overline{t}).$$

Thus

$$\lim_{h \to 0} \frac{T_{\overline{t},\overline{t}+h}\phi(\cdot, \overline{t}+h)(\overline{x}) - \phi(\overline{x}, \overline{t})}{h} \geq 0,$$

which implies

$$-D_t\phi(\overline{x}, \overline{t}) + H(D_x\phi(\overline{x}, \overline{t}), \overline{x}, \overline{t}) \leq 0.$$

The case where $V - \phi$ has a local minimum is similar. ■

Theorem 14 (Comparison principle). *Suppose U and V are, respectively, a viscosity subsolution and a viscosity supersolution of the Hamilton-Jacobi-Belmann equation (2.8). Furthermore assume $U(x, t_1) \leq V(x, t_1)$. Then*

$$U(x, t) \leq V(x, t).$$

Proof. See [FS93]. ■

Theorem 15 (Uniqueness). *The value function is the unique viscosity solution.*

Proof. Let V_1 and V_2 be two viscosity solutions with the same terminal cost. Then by the comparison principle we would have $V_1 \leq V_2$ and, by symmetry, $V_2 \leq V_1$. ■

2.3 Additive Eigenvalue Problems

In this section we study the time dependent Hamilton-Jacobi-Belmann equation using additive eigenvalues techniques (see [Con95], [Con97] and [Con96]). Other references for additive eigenvalue problems and applications include [CD87], and [Nus91].

Let V be the viscosity solution of the Hamilton-Jacobi-Belmann equation

$$-D_tV + H(D_xV, x, t) = 0,$$

with terminal condition $V(x, t_1) = \psi(x)$. Then

$$V(x, t) = \inf_y \left(K(x, y; t) + \psi(y)\right) = T_{t,t_1}\psi.$$

Assume that L is $1-$periodic in the t variable. Define

$$K_1(x,y) = K(x,y;t_1-1),$$

and

$$T_n = T_{t_1-n,t_1}.$$

In general, even when L is periodic in t, the solutions of the HJB equation (2.8) are not periodic in time. Indeed, if $L>0$, the value function is decreasing as t increases. Therefore T_n does not have fixed points. Thus, the natural question is whether there are functions $u(x)$ such that

$$T_n u(x) = u(x) + \overline{H}n,$$

for some value $\overline{H}$ and for all integers n.

Suppose T is a (possibly nonlinear) operator acting in suitable function space $\mathcal{C}$ containing the scalars. An *additive eigenvalue* and a corresponding *additive eigenvector* are, respectively, a number λ and a function v such that

$$Tv = v + \lambda.$$

The existence of $1-$periodic (in time) solutions of

$$-D_tV + H(D_xV, x, t) = \overline{H} \tag{2.11}$$

is equivalent to the existence of additive eigenvectors and eigenvalues for the operator T_1. The appropriate choice for the space $\mathcal{C}$, where T_1 is defined, turns out to be the space of functions $\{g : \mathbb{R}^n \to \mathbb{R} \mid g \text{ is continuous and } [0,1]^n - \text{periodic}\}$. It is not hard to prove that T_1 is well defined in this space and that $T_1 g$ is uniformly Lipschitz independently of g.

To prove the existence an additive eigenvalue in $\mathcal{C}$, consider the auxiliary operator

$$Lg(x) = T_1 g(x) - \min_{y\in\mathbb{R}^n} T_1 g(y).$$

Proposition 11. *Suppose L has a fixed point. Then T_1 has an additive eigenvalue and eigenvector.*

PROOF. Suppose g is a fixed point for L. Define $\lambda = \min_{y\in\mathbb{R}^n} T_1 g(y)$. Then $T_1 g = g + \lambda$. ∎

Theorem 16. *The operator T_1 has an additive eigenvalue $\overline{H}$ and corresponding eigenvector u. Furthermore the additive eigenvalue is unique.*

PROOF. The existence of a fixed point is a consequence of Schaefer's theorem (see [Eva98]). For the proof consult [EG99b].

To prove uniqueness of additive eigenvalue assume that we can find g_1, g_2,λ_1, and λ_2 such that

$$T_1 g_1 = g_1 + \lambda_1 \qquad T_1 g_2 = g_2 + \lambda_2,$$

with $\lambda_1 \neq \lambda_2$. Let $h = g_1 - g_2$ and choose $\overline{x}$ such that $h(\overline{x}) = \sup h$. Then, for some $\overline{y} \in \mathbb{R}^n$

$$g_1(\overline{y}) + K_1(\overline{x}, \overline{y}) = g_1(\overline{x}) + \lambda_1$$

and

$$g_2(\overline{y}) + K_1(\overline{x}, \overline{y}) \geq g_2(\overline{x}) + \lambda_2.$$

Hence

$$g_2(\overline{y}) - g_1(\overline{y}) \geq g_2(\overline{x}) - g_1(\overline{x}) + \lambda_2 - \lambda_1,$$

but since $h(\overline{y}) = \sup h$ we conclude

$$\lambda_2 - \lambda_1 \leq 0,$$

hence, by symmetry, $\lambda_2 = \lambda_1$. ■

REMARK. However the additive eigenvector need not be unique - see example 12.

2.4 Discounted Cost and Infinite Horizon

For optimizing long-time running costs is natural to consider the *discounted cost functional* J_α

$$J_\alpha(x; u) = \int_0^\infty L(x(s), \dot{x}(s)) e^{-\alpha s} ds,$$

where the trajectories $x(\cdot)$ satisfy the differential equation $\dot{x} = u$ with the initial condition $x(0) = x$, and $\alpha > 0$ is the discount rate. In financial applications, the discount rate may be the inflation rate or interest rate.

Define the value function $u_\alpha(x) = \inf J_\alpha(x; u)$, where the infimum is taken over controls $u \in L^\infty[0, +\infty)$.

Proposition 12. *There exists a constant C that does not depend on α such that*

$$u_\alpha \leq \frac{C}{\alpha}.$$

Proof. Since $L(x,0)$ is bounded

$$u_\alpha(x) \le J_\alpha(x,0) \le \int_0^\infty L(x,0)e^{-\alpha s}ds \le \frac{C}{\alpha}.$$

■

The analog of the dynamic programming principle is

Proposition 13. *For any $t > 0$*

$$u_\alpha(x) = \inf \int_0^t L(x(s),\dot{x}(s))e^{-\alpha s}ds + e^{-\alpha t}u_\alpha(x(t)).$$

Proof. Note that

$$u_\alpha(x) = \inf\left(\int_0^t L(x(s),\dot{x}(s))e^{-\alpha s}ds + e^{-\alpha s}\int_t^\infty L(x(s),\dot{x}(s))e^{-\alpha s}ds\right) \le$$

$$\le \inf \int_0^t L(x(s),\dot{x}(s))e^{-\alpha s}ds + e^{-\alpha s}u_\alpha(x(t)),$$

and the other inequality is trivial. ■

Proposition 14. *Suppose u_α is smooth. Then it satisfies the equation*

$$H(D_x u_\alpha, x) + \alpha u_\alpha = 0. \tag{2.12}$$

Proof. If u is smooth then at $t = 0$

$$\frac{d}{dt}(e^{-\alpha t}u_\alpha(x(t)) = -\alpha u_\alpha(x) + \dot{x}(0)\cdot D_x u_\alpha(x) = -L(x,\dot{x}(0)).$$

Hence

$$0 = \alpha u_\alpha(x) - \dot{x}(0)\cdot D_x u_\alpha(x) - L(x,\dot{x}(0)) \le \alpha u_\alpha(x) + H(D_x u_\alpha, x).$$

Conversely, for any other trajectory $y(\cdot)$ with $y(0) = 0$,

$$u_\alpha(x) \le \int_0^t L(y,\dot{y})e^{-\alpha s}ds + e^{-\alpha t}u_\alpha(y(t)).$$

Consequently,

$$0 \ge \alpha u_\alpha(x) - \dot{y}(0)\cdot D_x u_\alpha(x) - L(x,\dot{y}(0))$$

In particular if we choose $\dot{y}(0)$ conveniently we conclude

$$0 \ge \alpha u_\alpha(x) + H(D_x u_\alpha, x),$$

i.e., $\alpha u_\alpha(x) + H(D_x u_\alpha, x) = 0$. ■

In fact we may prove that 2.12 is satisfied in the viscosity sense (defined exactly as above).

Theorem 17. *u_α is a viscosity solution of (2.12).*

PROOF. Similar to the finite horizon case - see [FS93]. ■

Using the same techniques as before one can show that the optimal trajectories exist and satisfy the Euler-Lagrange equations with *negative dissipation*:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \alpha\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0. \tag{2.13}$$

If $x(t)$ solves (2.13) the energy given by the Hamiltonian H need not be conserved.

Example 5. Let $L = \frac{\dot{x}^2}{2} + \cos x$. The equations are

$$\ddot{x} - \alpha\dot{x} + \sin x = 0.$$

When $\alpha = 0$ the energy $H = \frac{\dot{x}^2}{2} - \cos x$ is conserved but for $\alpha > 0$

$$\frac{dH}{dt} = \alpha\dot{x}^2.$$

Thus the energy increases in time, unless $\dot{x} = 0$. ◀

Proposition 15. *Suppose $x(t)$ solves (2.13) and $p(t) = -D_vL$. Then*

$$\frac{dH}{dt} = \alpha D_vL \cdot v.$$

PROOF. We have

$$\frac{dH}{dt} = D_pH\dot{p} + D_xH\dot{x} = v(\alpha D_vL + D_xL) - D_xLv$$

which gives the result. ■

In an appropriate sense, this equation approximates, as $\alpha \to 0$, the *time independent* Hamilton-Jacobi equation

$$H(D_xu, x) = \overline{H}. \tag{2.14}$$

for some $\overline{H}$ that will not be necessarily 0.

Theorem 18. *There exists a function u and constants c_α such that $u_\alpha - c_\alpha \to u$ uniformly. Furthermore there exist a constant $\overline{H}$ such that $\alpha u_\alpha \to \overline{H}$. The limit function u is a viscosity solution of (2.14).*

PROOF. We just present a sketch of the proof - the details can be found in [LPV88] or [Con95].

Lemma 16. *Suppose there exist constants c_α such that $u_\alpha - c_\alpha \to u$ uniformly. Furthermore suppose $\alpha u_\alpha \to \overline{H}$. Then u is a viscosity solution of (2.14).*

PROOF. See [LPV88] or [Con95]. ∎

To prove convergence we need some estimates.

Lemma 17. *Suppose u_α solves (2.12). Then there is a constant C, which does not depend on α such that*

$$\|D_x u_\alpha\|_\infty < C$$

PROOF. Since $u_\alpha \geq 0$

$$H(D_x u^\alpha, x) \leq 0.$$

By coercivity of H we get the result. ∎

We may take $c_\alpha = \min u_\alpha$. Then the sequence $u_\alpha - c_\alpha$ is bounded and equicontinuous. Thus we can use Ascoli-Arzela theorem to extract a subsequence such that $u_\alpha - c_\alpha$ converges uniformly to some function u.

Finally we have to show that αu_α converges.

Lemma 18. *The sequence αu_α is uniformly bounded.*

PROOF. This follows proposition 12. ∎

From the previous lemma we can extract a subsequence such that, for some fixed point $\overline{x}$, $\alpha u_\alpha(\overline{x})$ converges. Since u_α is periodic and has a uniformly bounded Lipschitz constant we conclude that αu_α converges uniformly to a constant. ∎

Chapter 3

Hamiltonian Systems

In this chapter we review Hamiltonian dynamics and discuss the connections with Hamilton-Jacobi equations. In particular we study the integrability of Hamiltonian systems via classical Hamilton-Jacobi theory (generating functions theory), and perturbative methods (KAM theory). Finally, we discuss the theory of adiabatic invariants and its relations with homogenization theory.

3.1 Hamiltonian Systems

Let $H : \mathbb{R}^{2n+1} \to \mathbb{R}$ (we write $H(p, x, t)$ with $x \in \mathbb{R}^n$, $p \in \mathbb{R}^n$, and $t \in \mathbb{R}$) be a smooth function. The *Hamiltonian Ordinary Differential Equation* (Hamiltonian ODE) associated with the *Hamiltonian* H and *canonical coordinates* (p, x) is

$$\dot{x} = (D_pH)^T \qquad \dot{p} = -(D_xH)^T. \tag{3.1}$$

Remark. Generally one writes $\dot{x} = D_pH$ and $\dot{p} = -D_xH$. However, for certain purposes, it is convenient not to identify line vectors with column vectors.

We say that a set of coordinates (p, x) are canonical coordinates with respect to a ODE $(\dot{x}, \dot{p}) = (f, g)$ and Hamiltonian H if $f = D_pH$ and $g = -D_xH$.

Remark. The adjoint variable defined in equation (2.5) is a solution of a time reversed Hamiltonian ODE.

Usually, the Hamiltonian H represents the total energy of the system. If H is time independent, then the energy is conserved, as shown in the next proposition:

Proposition 16. *Suppose $(p(t), x(t))$ are solutions of (3.1). Then*

$$\frac{d}{dt}H(p(t), x(t), t) = D_t H(p(t), x(t), t),$$

in particular H is a constant of motion if it does not depend explicitly on time.

Proof. Suppose $(p(t), x(t))$ solve (3.1). Then

$$\frac{d}{dt}H(x(t), p(t)) = D_x H\dot{x} + D_p H\dot{p} + D_t H = D_x H(D_p H)^T - D_p H(D_x H)^T + D_t H,$$

thus $\frac{d}{dt}H(x(t), p(t)) = D_t H.$ ■

Example 6. The one dimensional pendulum corresponds to $H(p, x) = \frac{p^2}{2} - \cos x$. Then the equations of motion are

$$\dot{x} = p \qquad \dot{p} = -\sin x.$$

Since the energy is conserved, the solutions of these equations lie on level sets $H(p, x) = E$. Therefore $p(x) = \sqrt{2(E + \cos x)}$. ◀

In general, we may assume H to be time independent. Indeed, if H depends on t, define two extra variables (E, τ), canonically conjugated. Consider the Hamiltonian $K = H(p, x, \tau) + E$. Then

$$\dot{\tau} = \frac{\partial K}{\partial E} = 1 \qquad \dot{E} = -\frac{\partial K}{\partial \tau} = -\frac{\partial H}{\partial \tau}.$$

Thus $t = \tau$, up to additive constants, K is a constant of motion, and the equations for the remaining coordinates are left unchanged. However, in some applications it is convenient to have explicit time-dependence.

3.2 Hamilton-Jacobi Theory

When changing coordinates in a Hamiltonian system one must be careful because the special structure of the Hamiltonian ODE is not preserved under general change of coordinates. To overcame this problem we study the theory of generating functions.

Proposition 17. *Let (p, x) be the original canonical coordinates and (P, X) be another coordinate system. Suppose $S(x, P)$ is a smooth function such that*

$$p = (D_x S(x, P))^T \qquad X = (D_P S(x, P))^T$$

defines a change of coordinates. Furthermore assume that $D^2_{xP}S$ is non-singular. Let $\overline{H}(P,X,t) = H(p,x,t)$. Then, in the new coordinate system, the equations of motion are

$$\dot{X} = (D_P\overline{H})^T \qquad \dot{P} = -(D_X\overline{H})^T, \tag{3.2}$$

i.e., (P,X) are canonical coordinates. In particular, if $\overline{H}$ does not depend on X, these equations simplify to

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

If the generating function $S(x,P,t)$ depends on time define $\overline{H} = H + D_tS$ and the equations of motion are again in canonical form (3.2).

Proof. Observe that

$$-(D_xH)^T = \dot{p} = D^2_{xx}S(D_pH)^T + D^2_{Px}S\dot{P} + D^2_{tx}S,$$

and so

$$D^2_{Px}S\dot{P} = -\left[D^2_{xx}S(D_pH)^T + (D_xH)^T + D^2_{tx}S\right] \tag{3.3}$$

Since $\overline{H}(P, D_PS, t) = H(D_xS, x, t) + D_tS$,

$$D_X\overline{H}D^2_{xP}S = D_pHD^2_{xx}S + D_xH + D^2_{xt}S.$$

Transposing the previous equation and comparing with (3.3), using the fact that $D^2_{xP}S = (D^2_{Px}S)^T$ is non-singular and $D^2_{xx}S$ is symmetric,

$$\dot{P} = -(D_X\overline{H})^T.$$

We also have

$$\dot{X} = D_xX\dot{x} + D_PX\dot{P} + D_tX = D^2_{xP}S(D_pH)^T + D^2_{PP}S\dot{P} + D^2_{tP}S.$$

Again using the identity $\overline{H}(P, (D_PS)^T, t) = H((D_xS)^T, x, t) + D_tS$, we get

$$D_P\overline{H} + D_X\overline{H}D^2_{PP}S = D_pHD^2_{Px}S + D^2_{Pt}S.$$

Again by transposition,we get

$$\dot{X} = (D_P\overline{H})^T + (D^2_{PP}S)^T\left(\dot{P} + (D_X\overline{H})^T\right),$$

which implies $\dot{X} = (D_P\overline{H})^T$. ■

The function S in the previous proposition is called a *generating function.* There are several other types of generating functions which depend on other coordinates - see [Arn99], [AKN97] or [Gol80] for details. The systematic approach to generating functions requires exterior calculus and symplectic geometry.

Example 7. Let $H = \frac{(p+x)^2}{2}$. When $S = P \cdot x - \frac{x^2}{2}$, we get the change of coordinates:

$$p = D_x S = P - x \qquad X = D_P S = x.$$

In the new coordinates, $H(p, x) = \overline{H}(P) = \frac{P^2}{2}$ and so the equations of motion are

$$\dot{P} = 0 \qquad \dot{X} = P. \quad \blacktriangleleft$$

It turns out that, in principle, we can find generating functions that transform the original Hamiltonian $H(x, p)$ in a Hamiltonian $\overline{H}(P)$ depending only on P by solving a PDE. When this is possible we say that the Hamiltonian ODE is *completely integrable.*

Proposition 18. *Suppose $S(x, P)$ is a smooth generating function such that in the new coordinates (X, P), $\overline{H}(X, P) \equiv \overline{H}(P)$. Then S is a solution of the PDE*

$$H(D_x S, x) = \overline{H}(P). \tag{3.4}$$

Furthermore if H depends on time, $S(x, P, t)$ solves $D_t S + H(D_x S, x, t) = \overline{H}(P)$.

PROOF. If $p = D_x S$ then $H(D_x S, x) = \overline{H}(P)$. Clearly, the time dependent case is similar. ■

However, in general, the PDE (3.4) does not have global smooth solutions. Indeed, there are examples of systems that are not completely integrable (see [AKN97] or [Oli95], for instance), and therefore, in such examples, (3.4) cannot have a globally defined solution.

Note that in the last proposition we have, in general, two unknowns, S and $\overline{H}(P)$. Finding $\overline{H}(P)$ is as important as finding S! Finally note that if we allow $\overline{H}$ to depend both on X and P the equation that we have to solve is

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

Suppose $H(p, x)$ is $[0, 1]^n$ periodic in x. In general the solution of the Hamilton-Jacobi equation (3.4) may not be periodic. Even if we require that the canonical change of variables $X(x, p)$ and $P(x, p)$ be periodic in x the corresponding generating function S may

not be periodic. For instance, it may have the form $S = P \cdot x + u(x)$, where u is a periodic function.

Suppose for each P we can find $\overline{H}(P)$ such that there exists a periodic smooth solution u of the PDE

$$H(P + D_x u, x) = \overline{H}(P). \tag{3.5}$$

Define the generating function $S = P \cdot x + u$. Assume further that

$$p = P + D_x u \qquad Q = x + D_P u$$

defines a smooth change of coordinates. Then, in the new coordinates

$$\dot{P} = 0 \qquad \dot{Q} = D_P \overline{H}.$$

If we identify $Q \mod [0,1]^n$ the orbits will lie in a n dimensional tori. These are called *Lagrangian tori* (see [E99]). Moreover, the *rotation vector* $\omega \equiv \lim_{t\to\infty} \frac{x(t)}{t}$ of the orbits $x(t)$ exists and is

$$\omega = \lim_{t\to\infty} \frac{x(t)}{t} = \lim_{t\to\infty} \frac{Q(t)}{t} = D_P \overline{H},$$

since $D_P u$ is bounded (under smoothness and periodicity assumptions).

3.3 KAM theory

We know that, in general, it is not possible to find global smooth solutions of the Hamilton-Jacobi PDE. However, in some trivial cases, e.g. $H \equiv H(p)$, one can find explicit solutions of the PDE. In this example, any function $u(P)$ and $\overline{H}(P) = H(P)$, solves

$$H(P + D_x u) = \overline{H}(P).$$

The Kolmogorov-Arnold-Moser (KAM) theory studies the case where H can be written as $H(p,x) = H_0(p) + \delta H_1(p,x)$, for δ sufficiently small. Using perturbative methods, it is possible to find (under generic assumptions) smooth solutions for the Hamilton-Jacobi PDE and thus prove the existence of Lagrangian tori.

Theorem 19 (KAM). *Suppose $H(p,x) = H_0(p) + \delta H_1(p,x)$ where H_0 and H_1 are smooth and H_1 is $[0,1]^n$ periodic in x. Assume that the non-degeneracy condition* $\det D^2_{pp} H_0 \neq 0$ *holds. Then, if δ is small enough, there exists a function $\overline{H}(P)$ defined on a set $\mathcal{P}_\delta$ such that for all $P \in \mathcal{P}_\delta$ there exists a smooth periodic solution $u(P,x)$ of the equation (3.5). Moreover, $|\mathcal{P}^c_\delta| \to 0$ as $\delta \to 0$.*

PROOF. For the proof of this theorem see [Gal83] and [MS88]. However, since the proof is quite technical the reader may want to read first the sketch in [Arn99] or [AKN97]. The main idea is to construct a sequence of approximate solutions of the Hamilton-Jacobi equation

$$H_0(D_xS) + \delta H_1(D_xS, x) = \overline{H}(P).$$

The first approximation is $S = x \cdot P + \delta S_1$, where S_1 satisfies

$$D_P H_0(P) D_x S_1 + H_1(P, x) = 0.$$

At this point one usually expands $H_1(P, x)$ in a multiple Fourier series and (under non-resonance conditions on $D_P H_0$) we can find a solution S_1. The generating function that is obtained gives us a change of coordinates that transforms the original Hamiltonian in a new Hamiltonian $\overline{H}_1$ with the form

$$\overline{H}^1(P, X) = \overline{H}_0^1(P) + \delta^2 \overline{H}_1^1(P, X; \delta).$$

By applying this procedure to $\overline{H}^1$ the terms up to δ^4 will be removed in $\overline{H}^2$. In general the error in $\overline{H}^n$ is δ^{2^n}. This is a superconvergent method, similar to Newton's method. Finally under non-resonance assumptions it is possible to prove that these approximations converge for most values of P. Indeed if $\det D^2_{pp} H_0 \neq 0$ the the set of values of P for which this iterative procedure may not converge may be dense but it will have a small measure. ■

Example 8. Let $H = \frac{p^2}{2} + \delta \cos x$. Then $S_1 = -\frac{\sin x}{P}$. This yields the generating function $S = xP - \delta \frac{\sin x}{P}$ and a new Hamiltonian

$$\overline{H}^1(P, X) = \frac{P^2}{2} + \frac{\delta^2 \cos^2(X)}{2P^2} + O(\delta^3). \quad ◀$$

Even in the case where the Hamiltonian does not have a small parameter it may be possible, by conveniently rescale the variables, to make appear one and apply KAM theory.

Example 9. Suppose

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

Rescale p to $\frac{p}{\sqrt{\delta}}$ and time to $t\sqrt{\delta}$. Then this system is equivalent to one with

$$H_\delta = \frac{|p|^2}{2} + \delta V(x).$$

To this system we can apply KAM theory. Since the invariant tori exist for all time, we can rescale back the time and show that in the original system there are KAM tori for $|p|$ large enough. ◀

The next theorem is a generalization of KAM theorem under weaker conditions and it is particularly useful in the time dependent case where the previous non-resonance condition is never satisfied.

Theorem 20 (Isoenergetic nondegeneracy). *Suppose*

$$H(p,x) = H_0(p) + \delta H_1(p,x)$$

where H_0 and H_1 are smooth and H_1 is $[0,1]^n$ periodic in x. Assume that

$$\det \begin{bmatrix} D^2_{pp}H_0 & (D_pH_0)^T \\ D_pH_0 & 0 \end{bmatrix} \neq 0.$$

Then, if δ is small enough, there exists a function $\overline{H}(P)$ defined on a set $\mathcal{P}_\delta$ such that for all $P \in \mathcal{P}_\delta$ there exists a smooth periodic solution $u(P,x)$ of the equation (3.5). Moreover, $|\mathcal{P}^c_\delta| \to 0$ as $\delta \to 0$.

Proof. The proof is similar to the KAM theory but the non-resonance condition has to be adapted (see [Gal83], [MS88], [Arn99] or [AKN97]). ■

This previous theorem is particularly useful when the perturbation is time dependent. Indeed, by using the standard way to pass from a time dependent to a time-independent, the condition in theorem 19 will always fail. However, in general, we can still apply the previous result.

We finish this section with an example due to Pöschel [Pös82]. This example illustrates the importance of Diophantine conditions and the differentiability properties of solutions of a Hamilton-Jacobi equation.

Example 10. Consider the linearized equation

$$P \cdot D_x u = f(x),$$

with f periodic. Assume that u and f can be written as

$$u(x) = \sum_{\alpha \in \mathbb{Z}^n} a_\alpha e^{i\alpha\cdot x} \qquad f(x) = \sum_{\alpha \in \mathbb{Z}^n} b_\alpha e^{i\alpha\cdot x}.$$

Then

$$iP \cdot \alpha a_\alpha = b_\alpha,$$

hence

$$a_\alpha = \frac{b_\alpha}{iP \cdot \alpha}.$$

The convergence of the Fourier series of u can be estimated in terms of the coefficients b_α and a Diophantine condition on P. More precisely, if

$$|P \cdot \alpha| \geq C|\alpha|^{-\gamma}, \tag{3.6}$$

we have

$$\sum_{\alpha \in \mathbb{Z}^n} |a_\alpha| \leq C \sum_{\alpha \in \mathbb{Z}^n} |b_\alpha||\alpha|^\gamma.$$

For instance if P is linearly dependent over the rationals, i.e. there exists a vector $\alpha \in \mathbb{Z}^n$ such that $\alpha \cdot P = 0$, then the convergence cannot be guaranteed since some coefficients can be infinite. However almost every P will satisfy a condition like (3.6). There are also values of P that, although not being linearly dependent over the rationals, do not satisfy (3.6). For those numbers the convergence may be a lot harder to prove or the series may not converge at all.

To compute the derivative $D_P u$, the Diophantine condition is even more important. Indeed, if we differentiate formally a_α in P we get

$$D_P a_\alpha = -\frac{b_\alpha \alpha}{i(P \cdot \alpha)^2}.$$

Thus in the absence of a Diophantine condition, convergence may be a lot harder to obtain. ◀

3.4 Invariant measures

As before, suppose u is a smooth periodic solution of

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

Assume further that we can make the change of coordinates

$$X = x + D_P u \qquad p = P + D_x u.$$

For fixed P the set defined by $p = P + D_x u$ is invariant. Now we turn our attention invariant measures supported on this set. First observe that since $\dot{X}$ is constant, the measure

$$\mu(A) = \int_A dX,$$

is a invariant probability measure, with respect to the dynamics

$$\dot{X} = D_P \overline{H}.$$

The change of coordinates formula yields

$$\mu(A) = \int_{X^{-1}(A)} \det(I + D^2_{xP} u) dx.$$

Therefore the measure ν

$$\nu(B) = \int_B \alpha(x) dx,$$

with $\alpha(x) = \det(I + D^2_{xP} u)$, is a invariant measure under the dynamics

$$\dot{x} = D_p H(P + D_x u, x). \tag{3.7}$$

Proposition 19. *The density α satisfies the nonlinear partial differential equation*

$$\nabla \left(\alpha(x) D_p H(P + D_x u, x) \right) = 0.$$

Proof. Since the measure ν is invariant we have, for any set B with smooth boundary ∂B,

$$\frac{d}{dt} \int_{\Theta_t(B)} \alpha(x) dx = 0$$

where $\Theta_t(x)$ is the flow corresponding to (3.7).

This implies

$$\int_{\partial B} D_p H(P + D_x, x) \alpha(x) dx = 0.$$

Using the divergence theorem we conclude

$$\int_B \nabla \left(\alpha(x) D_p H(D_x u, x) \right) dx = 0.$$

Since B is arbitrary, this proves the proposition. ■

One of the main questions in ergodic theory is whether invariant measures can be obtained by time averaging along trajectories. More precisely, is it true that for any Borel set B

$$\nu(B) = \lim_{t \to +\infty} \frac{1}{t} \int_0^t 1_B(x(t)), \tag{3.8}$$

at least for "generic" trajectories $x(t) = \Theta_t(x)$?

Proposition 20. *If the rotation vector $\omega = D_P\overline{H}$ is linearly independent over the rationals (i.e. $\omega \cdot v = 0$, for $v \in \mathbb{Z}^n$, implies $v = 0$) then for almost any initial point (3.8) holds for all Borel sets B. If ω is not incommensurable, define $\mathcal{F}_I$ to be the σ-algebra generated by the Θ-invariant sets. Then*

$$\lim_{t\to+\infty}\frac{1}{t}\int_0^t 1_B(x(t)) = \nu(B|\mathcal{F}_I),$$

where $\nu(B|\mathcal{F}_I)$ is the conditional probability of B given $\mathcal{F}_I$.

PROOF. This is a straightforward application of the ergodic theorem [Dur96]. ■

Finally we would like to point out that the Diophantine condition that appears in the KAM theorem has an interpretation in terms of invariant measures. Suppose we have a linear completely integrable Hamiltonian system with Hamiltonian $H(p)$. Then the orbits are straight lines. If D_pH is linearly independent over the rationals then the Lebesgue measure is the unique ergodic measure on the torus $[0,1]^n$. One could ask how long do we have to take T such that the measure μ_T defined by the ergodic average

$$\frac{1}{T}\int_0^T f(x(s))ds = \int f d\mu_T$$

gives a good approximation to the Lebesgue measure. It turns out that a Diophantine condition is exactly what is necessary for having a small value of T. This can be seen by considering $f_k(x) = e^{ik\cdot x}$ for $k \in \mathbb{Z}^n$. Then

$$\int f_k d\mu_T = \frac{e^{ik\cdot\omega T}-1}{iT\omega\cdot k}$$

and so, for many values of k this number can be rather large if ω is close to a resonance, whereas the average of f_k with respect to the Lebesgue measure is zero.

3.5 Adiabatic Invariants

Now we consider the case in which the Hamiltonian depends slowly on time, i.e., $H(p,x,\epsilon t)$, for some small parameter ϵ. Suppose H is periodic in time (without loss of generality we may assume period one, i.e. $H(p,x,t) = H(p,x,t+1)$). We are interested in the asymptotic properties for times $0 \le t \le \frac{1}{\epsilon}$, when $\epsilon \to 0$. In principle, we could try to find a periodic (in time) solution of the time dependent Hamilton-Jacobi equation

$$-D_tS + H(P + D_xS, x, \epsilon t) = \overline{H}_\epsilon(P) \tag{3.9}$$

and find a global change of coordinates. This approach has two main problems: one is that this equation may not have a smooth solution, the other is that even if a smooth solution exists it may be very hard to determine explicitly. Sometimes it is easier to solve, for fixed t, the equation

$$H(P + D_x S, x, \epsilon t) = \overline{H}(P, \epsilon t). \tag{3.10}$$

In fact, in the one dimensional case the latter equation can be solved explicitly (up to computing integrals and solving algebraic equations). Suppose that equation(3.10) has a smooth (in all variables) and periodic (in t) family of solutions. Furthermore assume $X = D_P S$ and $p = P + D_x S$ defines a smooth change of coordinates. Then we can write $P(t) = P(x(t), p(t))$ and $X(t) = X(x(t), p(t))$. The function P is called the action and X the corresponding angle. We say that the function P is an adiabatic invariant provided

$$\sup_{0 \le t \le \frac{1}{\epsilon}} |P(t) - P(0)| = O(\epsilon).$$

In general, we do not expect the function $X(t)$ to be conserved. The natural question turns out to be whether we can approximate $X(t)$ by

$$\overline{X}(t) = X(0) + \int_0^t D_P \overline{H}(P, \epsilon s) ds.$$

A reasonable conjecture would be that

$$\sup_{0 \le t \le \frac{1}{\epsilon}} |X(t) - \overline{X}(t)| = O(1).$$

The next example show that this is false.

Example 11. Let $L(x, \dot{x}, t) = \frac{(\dot{x} - c(t))^2}{2}$ where $|c(t)| \le 1$, $c(t) = 1$ for $t \in [1/8, 3/8]$ and $c(t) = -1$ for $t \in [5/8, 7/8]$.

More generally we may always add to the solutions $u(x, t)$ of (3.10) any periodic function $f(P, t)$. This yields another family of solutions of the same equation. The arbitrariness of f makes meaningless any comparison of $X(\frac{t}{\epsilon})$ with $\overline{X}(\frac{t}{\epsilon})$ unless t is integer. So the only thing we can expect to be true generally is that the limit $X(\frac{1}{\epsilon}) - \overline{X}(\frac{1}{\epsilon}) \to \Theta_H$, called the Hannay angle, exists. If this limit exists it can be computed by the formula

$$\Theta_H = \lim_{\epsilon \to 0} \int_0^{\frac{1}{\epsilon}} (D_P \overline{H}_\epsilon(P) - D_P \overline{H}(P, \epsilon s)) ds.$$

Since the PDE (3.10) may not have classical solutions, it might be impossible to define the action variable. Even when this is possible there is no rigorous proof that it is an adiabatic invariant or that the Hannay angle exits. However, it is commonly believed that these facts are true, at least for generic (in an appropriate sense) Hamiltonian systems and "most" initial conditions.

Sometimes is convenient to work with a rescaled form of (3.9), namely

$$-D_t S + H(P + D_x S, \frac{x}{\epsilon}, t) = \overline{H}(P),$$

with $0 < t < 1$. One would expect, as $\epsilon \to 0$, the oscillatory effects to be averaged. This is in fact true and is studied in the next chapter.

Chapter 4

Homogenization Theory

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.

The adiabatic invariants theory motivates the study of the limit as $\epsilon \to 0$ of viscosity solutions of the Hamilton-Jacobi-Belmann equation

$$-V_t^\epsilon + H\left(D_x V^\epsilon, \frac{x}{\epsilon}, t\right) = 0, \tag{4.1}$$

with terminal condition $V^\epsilon(x,T) = g^\epsilon(x)$. Such limit problem is a typical example of the applications of homogenization theory.

We assume that $H(p,y,t)$ 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].

4.1 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

$$-\frac{\partial V_0}{\partial t}(\epsilon y, t) + H\left(D_x V_0 + D_y V_1, y, t\right) = O(\epsilon),$$

where $y = \frac{x}{\epsilon}$. Letting $\epsilon \to 0$ we deduce that V_1 should be a periodic solution of the cell problem

$$H(P + D_y u, y, t) = \overline{H}(P, t), \tag{4.2}$$

with $P = D_x V_0$ and $\overline{H}(P, t) = \frac{\partial V_0}{\partial t}$.

This formal calculations suggest that the (viscosity) solution V^ϵ will converge to some function V that solves

$$-V_t + \overline{H}(D_x V, t) = 0.$$

4.2 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 (4.2). From theorem 18, we know that for each P and t (fixed) there exists a unique $\overline{H}(P, t)$ for which the equation

$$H(P + D_x u, x, t) = \overline{H}(P, t) \tag{4.3}$$

has a periodic viscosity solution. The function $\overline{H}(P, t)$ is called the *effective Hamiltonian.* Obviously, the viscosity solution of $H(P + D_x u, x, t) = \overline{H}(P, t)$ is not unique. In fact, given any solution $u(P, x, t)$, $v(P, x, t) = u(P, x, t) + f(P, t)$, for any arbitrary function f, is another solution. Even modulo functions of P and t the solution may not be unique (see Example 12). Also u may not be smooth as a function of t.

Note that the cell problem is the same problem that we have to solve to construct invariant tori in the Hamilton-Jacobi or KAM theory if the Hamiltonian is time-independent. When H depends on t the cell problem is used in the adiabatic invariants theory.

Theorem 21. *The viscosity solution V^ϵ of the terminal value problem (4.1) converges uniformly to the viscosity solution of*

$$-V_t + \overline{H}\left(D_x V, t\right) = 0 \tag{4.4}$$

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 (4.4). 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

$$-\phi_t(\hat{x}, \hat{t}) + \overline{H}(D_x\phi(\hat{x}, \hat{t}), \hat{t}) \le 0.$$

Assume this statement is false. Then there exists a maximum point $(\hat{x}, \hat{t})$ of $V - \phi$ and $\theta > 0$ such that

$$-\phi_t(\hat{x}, \hat{t}) + \overline{H}(D_x\phi(\hat{x}, \hat{t}), \hat{t}) > \theta. \tag{4.5}$$

Let $u(y)$ to be a viscosity solution of the problem

$$H(D_x\phi(\hat{x}, \hat{t}) + D_y u(y), y, \hat{t}) = \overline{H}(D_x\phi(\hat{x}, \hat{t}), \hat{t}). \tag{4.6}$$

Define

$$\phi^\epsilon(x, t) = \phi(x, t) + \epsilon u(\frac{x}{\epsilon}).$$

We claim that in the viscosity sense

$$-\phi^\epsilon_t(x, t) + \overline{H}(D_x\phi^\epsilon(x, t), t) \ge \frac{\theta}{3},$$

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

$$u(\frac{x}{\epsilon}) - \eta(\frac{x}{\epsilon}, \frac{t}{\epsilon}) \ge u(\frac{x_1}{\epsilon}) - \eta(\frac{x_1}{\epsilon}, \frac{t_1}{\epsilon}),$$

where $\eta(x, t) = \frac{1}{\epsilon}[\psi(\epsilon x, \epsilon t) - \phi(\epsilon x, \epsilon t)]$, 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 (4.6) then

$$H\left(D_x\phi(\hat{x}, \hat{t}) + D_y\eta\left(\frac{x_1}{\epsilon}, \frac{t_1}{\epsilon}\right), \frac{x_1}{\epsilon}, \hat{t}\right) \ge \overline{H}(D_x\phi(\hat{x}, \hat{t}), \hat{t}).$$

By adding $D_t\phi(\hat{x}, \hat{t})$ to both sides and using (4.5) we conclude

$$D_t\phi(\hat{x}, \hat{t}) + H(D_x\phi(\hat{x}, \hat{t}) + D_x\psi(x_1, t_1) - D_x\phi(x_1, t_1), \frac{x_1}{\epsilon}, \hat{t}) \ge \theta$$

If r is chosen small enough (depending on the modulus of continuity of $D_x\phi$) then

$$D_t\phi(x_1, t_1) + H(D_x\psi(x_1, t_1), \frac{x_1}{\epsilon}, \hat{t}) \ge \frac{\theta}{2}$$

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

$$D_t\psi(x_1,t_1)+H(D_x\psi(x_1,t_1),\frac{x_1}{\epsilon},\hat{t})\geq\frac{\theta}{2}.$$

By having chosen r even small enough (depending on $|D_tH|$) one has

$$D_t\psi(x_1,t_1)+H(D_x\psi(x_1,t_1),\frac{x_1}{\epsilon},t_1)\geq\frac{\theta}{3}.$$

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.■

4.3 Effective Hamiltonian

Since we are only interested in solving the cell problem (4.2) for fixed t, in this section we will assume, for simplicity of notation, that H does not depend on t.

Given a classical solution of the time independent Hamilton-Jacobi equation (4.3) it is easy to construct a classical solution of the Hamilton-Jacobi-Belmann PDE (2.8) by using separation of variables. The next proposition proves a generalization of this idea for viscosity solutions.

Proposition 21. *Let $u(x,P)$ be a periodic viscosity solution of $H(P+D_xu,x)=\overline{H}(P)$. Then $V(x,t)=u(x,P)+\overline{H}(P)(t-t_1)$ is a viscosity solution of the problem*

$$-V_t+H(P+D_xV,x)=0$$

with terminal condition $V(t_1,x)=u(P,x)$.

Proof. Suppose u is a viscosity solution of $H(P + D_x u, x) = \overline{H}(P)$ and ϕ a smooth function. Fix (x, t) and assume that $V - \phi$ has a local maximum at (x, t). Then, since V is differentiable in t, $\phi_t = V_t = \overline{H}(P)$; moreover $u(\cdot) - \phi(\cdot, t)$ has a local maximum at x thus

$$H(P + D_x\phi(x,t), x) \leq \overline{H}(P) = \phi_t,$$

or equivalently $-\phi_t + H(P + D_x\phi(x,t), x) \leq 0$. The other part of the proof, when $V - \phi$ has a local minimum, is similar. ■

Lemma 19. *For fixed P the Legendre transform of $H(P + \cdot, x)$ is $L(v, x) + P \cdot v$*

Proof. By the definition of Legendre transform,

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

as required. ■

Theorem 22. *The effective Hamiltonian can be computed explicitly by*

$$\overline{H}(P) = -\lim_{t_1 \to \infty} \inf_{x(\cdot)|x(0)=x} \frac{\int_0^{t_1} L(x(s), \dot{x}(s)) + P \cdot \dot{x}(s) ds + u(P, x(t_1))}{t_1}$$

where $u(P, x)$ is any periodic viscosity solution for the problem (4.3) and the inf *is taken over all Lipschitz trajectories with $x(0) = x$, for some fixed x.*

Proof. Since the value function is the unique viscosity solution of the Hamilton-Jacobi-Bellman terminal value problem we can use the representation formula:

$$V(x, t) = \inf_{x(\cdot)} \int_t^{t_1} L(x(s), \dot{x}(s)) + P \cdot \dot{x}(s) ds + u(P, x(t_1)),$$

since $[H(P + \cdot, x)^*](v) = L + P \cdot v$, by Lemma 19. But $V(x, t) = u(P, x) + \overline{H}(P)(t - t_1)$ by Proposition 21. Thus, setting $t = 0$, dividing by t_1 and letting $t_1 \to \infty$, we get the result. See also [Con95], for a different proof. ■

From this theorem follows immediately

Corollary 3. *For any bounded Lipschitz function ψ*

$$\overline{H}(P) = -\lim_{t \to \infty} \inf_{x(\cdot)} \frac{\int_0^t L(x(s), \dot{x}(s)) + P\dot{x}(s) ds + \psi(x(t))}{t}. \tag{4.7}$$

Proof. This follows from last theorem and proposition 1. ■

We conclude this section by stating some properties of $\overline{H}$.

Proposition 22. *The effective Hamiltonian $\overline{H}(P)$ is Lipschitz, convex and superlinear in P.*

PROOF. See [LPV88] or [Con95]. ■

Chapter 5

Invariant Sets

The objective of this chapter is to develop, using viscosity solutions methods, an analog of the Hamilton-Jacobi integrability methods (generating functions) and KAM theory.

In general, there are no smooth solutions of the problem

$$H(P + D_x u, x) = \overline{H}(P). \tag{5.1}$$

Therefore we cannot apply the classical theory of generating functions. KAM theory attempts to solve this problem for a class of Hamiltonians where perturbation methods can be applied. Our approach is quite different. We use the fact that, under mild assumptions on H, there exists a viscosity solution of (5.1)

Any viscosity solution u of (5.1) is differentiable almost everywhere. Thus one is led to study the analog of a KAM tori: the graph $(x, P + D_x u)$ with u differentiable at x. Such set turns out to be backwards invariant, not forward invariant. This, however, is sufficient to construct for each P an invariant set $\mathcal{I}$.

In general, u and $\overline{H}$ are not smooth thus we cannot make the canonical change of variables $X = x + D_P u$ and $p = P + D_x u$. But to the classical statement $\dot{X} = D_P \overline{H}$ corresponds a weak analog

$$\lim_{t\to\infty} \frac{x(t)}{t} = D_P \overline{H}(P),$$

for almost every value P.

These results were proved originally by Moser et al. [JKM99] and W. E [E99] in the one-dimensional case. Using a different formulation, the case of compact manifolds

was studied by A. Fathi [Fat97a], [Fat97b], [Fat98a], and [Fat98b]. We present an unified approach to these results in sections 5.1 - 5.3. Then we study the structure of the invariant set and viscosity solutions (section 5.4), prove new the asymptotic estimates (section 5.5), and apply our results to the study of adiabatic invariants (section 5.6).

Before developing the general theory, we consider an example where we can solve the time independent Hamilton-Jacobi equation and gain some insight over the problem.

Example 12. Consider the Hamiltonian corresponding to a pendulum,

$$H(p, x) = p^2 - \cos x.$$

We can find explicitly the solution of the corrector problem (5.1). Indeed, for each $P \in \mathbb{R}$ and a.e. $x \in \mathbb{R}$, the solution $u(P, x)$ satisfies

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

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

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

Thus

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

where $|s(y)| = 1$. Because u is a viscosity solution, the only possible discontinuities in the derivative of u are the ones that satisfy $D_x u(x^-) - D_x u(x^+) > 0$. Therefore s can change sign from 1 to -1 at any point but jumps from -1 to 1 can happen only when $\sqrt{2(\overline{H}(P) + \cos x)} = 0$. If we require 2π-periodicity there are two cases, first if $\overline{H}(P) > 1$ the solution is C^1 since $\sqrt{2(\overline{H}(P) + \cos y)}$ is never zero. These solutions correspond to invariant torus. In this case P and $\overline{H}(P)$ satisfy the equation

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

It is easy to check that this equation has a solution $\overline{H}(P)$ whenever

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

When this inequality fails, $\overline{H}(P) = 1$ and $s(x)$ can have a discontinuity. Indeed, $s(x)$ jumps from -1 to 1 when $x = \pi + 2k\pi$, with $k \in \mathbb{Z}$, and there is a point x_0 defined by the equation

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

in which $s(x)$ jumps from 1 to -1. In this last case the graph $(x, P + D_x u)$ is a backwards invariant set contained in the unstable manifold of the hyperbolic equilibria of the pendulum. The graph of $\overline{H}(P)$ has a flat spot near $P = 0$.

This example also shows that the cell problem does not have a unique solution. Indeed $\cos x$ is also 4π periodic. So if we look for 4π periodic solutions we find out that for $|P|$ small we can have two points where the derivative is discontinuous and we can choose freely one of them because our only constraint is periodicity. ◀

In the last example, the graph $(x, P + D_x u)$ is not invariant, at least for small $|P|$. However, this set is backwards invariant. The main idea is that by running the time backwards we eventually end up with an invariant set. This will be the construction used in the general case.

5.1 Invariant sets

In chapter 3 we proved that, given a smooth periodic solution of the time independent Hamilton-Jacobi equation

$$H(P + D_x u, x) = \overline{H}(P), \tag{5.2}$$

it is possible construct an invariant set: the graph $(x, P+D_x u)$. Usually this set is identified with a n dimensional torus. As in the previous example, we would like to extend this construction to viscosity solutions. Suppose that u is a periodic viscosity solution of (5.2). Then u is a Lipschitz function in x, and so, by Rademacher theorem, it is differentiable a.e.. Let $\mathcal{G}$ be the set given by

$$\mathcal{G} = \{(x, P + D_x u) : u \,\text{is differentiable at}\, x\}.$$

However, if we look at the previous example, we see that $\mathcal{G}$ is not an invariant set - it is only backwards invariant. The next proposition shows that this is a general fact. Let Ξ_t be the *backwards flow* (time-reversed flow) associated with the Hamiltonian ODE associated with H

$$\dot{p} = D_x H(p, x) \qquad \dot{x} = -D_p H(p, x). \tag{5.3}$$

Proposition 23. *$\mathcal{G}$ is backwards invariant under Ξ_t - more precisely, for all $t > 0$, we have $\Xi_t(\mathcal{G}) \subset \mathcal{G}$.*

Proof. Let u be a viscosity solution of (5.2). Consider the time dependent problem

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

with terminal condition $V(t_1, x) = u(P, x)$. The (unique) viscosity solution is

$$V(x,t) = u(x) + \overline{H}(P)(t - t_1).$$

If u is differentiable at a point x_0 then, by theorem 11, $(t, x) = (0, x_0)$ is a regular point. Thus there exists a unique trajectory $x^*(s)$ such that $x^*(0) = x_0$ and

$$V(x_0, 0) = \int_0^{t_1} L(x^*(s), \dot{x}^*(s)) + P \cdot \dot{x}^*(s) ds + u(x^*(t_1)).$$

By theorem 10, along this trajectory the value function V is differentiable. Recall that the adjoint variable is defined by

$$p^*(s) = P + D_x V(x^*(s), s).$$

We know that the pair (x^*, p^*) solves the backwards Hamilton ODE (5.3). Therefore

$$(x^*(s), P + D_x V(x^*(s), s)) = (x^*(s), p(s)) = \Xi_s(x, p(0)) = \Xi_s(x, P + D_x V(x, 0)).$$

This implies

$$\Xi_s(x, P + D_x u) \in \mathcal{G},$$

for all $0 < s < t_1$. Since t_1 is arbitrary the previous inclusion holds for any $s \geq 0$. ■

For periodic time-dependent problems we can consider an additive eigenvector $u(x)$ corresponding to the evolution semigroup associated with the time-dependent Hamilton-Jacobi equation

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

Then $\mathcal{G}$ will be invariant under the time one Ξ_1 backwards flow. Other construction is to consider the (time-one periodic) set $\hat{\mathcal{G}} = \{(s, \Xi_s(z)), s \in S^1, z \in \mathcal{G}\}$, where $S^1 = \mathbb{R}/\mathbb{Z}$ (looking at the time as periodic). Then this set is backwards invariant. .

Lemma 20. *If $\mathcal{G}$ is an invariant set then its closure $\overline{\mathcal{G}}$ is also invariant.*

Proof. Take a sequence $(x_n, p_n) \in \mathcal{G}$ and suppose this sequence converges to $(x, p) \in \overline{\mathcal{G}}$. Then, for any t, $\Xi_t(x_n, p_n) \to \Xi_t(x, p)$. This implies $\Xi_t(x, p) \in \overline{\mathcal{G}}$. ■

Define $\mathcal{G}_t = \Xi_t(\overline{\mathcal{G}})$. Note that $\mathcal{G}_t$ is, in general, a proper closed subset of $\overline{\mathcal{G}}$. Let

$$\mathcal{I} = \cap_{t>0} \mathcal{G}_t.$$

Theorem 23. *$\mathcal{I}$ is a nonempty closed invariant set for the Hamiltonian flow.*

PROOF. Since $\mathcal{G}_t$ is a family of compact sets with the finite intersection property, its intersection is nonempty. Invariance follows from its definition. ■

This theorem generalizes the original one dimensional case by Moser et al. [JKM99] and W. E [E99]. A. Fathi has a different characterization of the invariant set using backward and forward viscosity solutions [Fat97a], [Fat97b], [Fat98a], and [Fat98b].

In the proof of theorem 23 we do not need to use the closure of $\mathcal{G}$. Even if $z \in \overline{\mathcal{G}} \backslash \mathcal{G}$ we have $\Xi_t(z) \in \mathcal{G}$, for all $t > 0$. Indeed, by theorem 10, the only points in an optimal trajectory that may fail to be regular are the end points.

We will see later that whenever the equation 5.2 has a classical solution, for instance when KAM theory is applicable, the set $\mathcal{I}$ corresponds to an invariant torus and $\mathcal{G} = \mathcal{I}$. In general the set $\mathcal{G}$ can be though of as a generalized unstable manifold of the set $\mathcal{I}$.

5.2 Rotation vector

It turns out, as we explain next, that the dynamics in the invariant set $\mathcal{I}$ is particularly simple. Suppose there is a smooth (both in P and x) periodic solution of the time independent Hamilton-Jacobi equation (5.2). Define $X = x + D_P u$. Then, for trajectories with initial conditions on the set $p = P + D_x u$ we have

$$\dot{X} = D_P \overline{H}(P),$$

or, equivalently, $X(t) = X(0) + D_P \overline{H}(P)t$. Therefore the dynamics of the original Hamiltonian system can be completely determined (assuming that one can invert $X = x + D_P u$).

We would like to prove an analog of this fact for orbits in the invariant set $\mathcal{I}$. A simple observation is that, in the smooth case,

$$\lim_{t \to \infty} \frac{x(t)}{t} = D_P \overline{H}(P) \equiv \omega, \tag{5.4}$$

the vector ω is called the rotation vector. The next theorem shows that (5.4) holds, under more general conditions, for all trajectories with initial conditions in the invariant set $\mathcal{I}$, provided $D_P \overline{H}$ exists.

Theorem 24. *Suppose $\overline{H}(P)$ is differentiable for some P. Then, the trajectories $x(t)$ of the Hamiltonian flow with initial conditions on the invariant set $\mathcal{I}(P)$ satisfy*

$$\lim_{t\to\infty}\frac{x(t)}{t}=D_P\overline{H}(P).$$

Proof. Fix P and P' and choose any $(x,p)\in\mathcal{I}$. By the representation formula (4.7)

$$\overline{H}(P)=-\lim_{t\to\infty}\frac{\int_0^t L(x^*(s),\dot{x}^*(s))+P\cdot\dot{x}^*(s)ds+u(x^*(t),P)}{t},$$

for some optimal trajectory x^*. Furthermore

$$\overline{H}(P')=-\lim_{t\to\infty}\inf_{x(\cdot):x(0)=x}\frac{\int_0^t L(x(s),\dot{x}(s))+P'\cdot\dot{x}(s)ds+u(x(t),P)}{t}. \tag{5.5}$$

Thus

$$\overline{H}(P')\geq-\liminf_{t\to\infty}\frac{\int_0^t L(x^*(s),\dot{x}^*(s))+P'\cdot\dot{x}^*(s)ds+u(P,x^*(t))}{t}.$$

The right hand side is equal to

$$-\liminf_{t\to\infty}\frac{\int_0^t(P'-P)\cdot\dot{x}^*(s)ds}{t}+\overline{H}(P).$$

Therefore

$$\overline{H}(P')-\overline{H}(P)\geq\limsup_{t\to\infty}\frac{\int_0^t(P-P')\cdot\dot{x}^*(s)ds}{t}=\limsup_{t\to\infty}\frac{(P-P')\cdot x^*(t)}{t}.$$

This implies immediately that for any vector Ω

$$-D_P\overline{H}(P)\cdot\Omega\geq\limsup_{t\to\infty}\frac{\Omega\cdot x^*(t)}{t}.$$

Replacing Ω by $-\Omega$ yields

$$-D_P\overline{H}(P)\cdot\Omega\leq\liminf_{t\to\infty}\frac{\Omega\cdot x^*(t)}{t}.$$

Consequently

$$-D_P\overline{H}(P)=\lim_{t\to\infty}\frac{x^*(t)}{t}.$$

Now note that the optimal trajectory $x^*(s)$ with initial conditions $(x^*(0),p^*(0))\in\mathcal{I}$ solves the backwards Hamilton ODE. So, any solution $x(t)$ of the Hamilton ODE with initial conditions on $\mathcal{I}$ satisfies

$$D_P\overline{H}(P)=\lim_{t\to\infty}\frac{x^*(t)}{t},$$

as required. ■

Corollary 4. *Suppose $x(t)$ is a optimal trajectory with initial conditions in $\mathcal{I}$. Then, for any subsequence t_j such that*

$$\omega \equiv \lim_{j\to\infty} \frac{x(t_j)}{t_j}$$

exists,

$$\overline{H}(P') \geq \overline{H}(P) + (P - P') \cdot \omega,$$

i.e. $\omega \in D_P^- \overline{H}(P)$.

Proof. By taking $t_j \to +\infty$ instead of $t \to +\infty$ in (5.5) we get

$$\overline{H}(P') - \overline{H}(P) \geq (P - P') \cdot \omega,$$

which proves the result. ■

Remark. These proofs are also valid for time dependent systems, we just would have to consider integer instead of arbitrary times.

We say that rotation vector ω is incommensurable provided that the unique solution $v \in \mathbb{Q}^n$ of the equation

$$\omega \cdot v = 0$$

is $v = 0$.

It is easy to see that if the rotation number incommensurable then there are no periodic orbits and the trajectories may be dense (modulo $\mathbb{Z}^n$), for instance when the system is integrable, or they may have a complicated Cantor-set like structure [Mos86]. In dimension one if the rotation number is rational then there are periodic orbits. However the proof seems to depend heavily on the dimension.

5.3 Regularity of the Invariant Set

In the two previous sections we proved that for each P there exists an invariant set $\mathcal{I}$ in which the Hamiltonian dynamics is particularly simple. However, little information is available, so far, about the structure of this set itself. In this section we apply a method by J. Mather [Mat91] to show that the invariant set is a Lipschitz graph - in particular this implies improved regularity for viscosity solutions.

The set $\mathcal{I}$ is contained in the closure of a graph, namely $(x, P + D_x u)$, where u is a viscosity solution of the cell problem

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

In general the derivative $D_x u$ may be discontinuous (see example 12). However we will prove that this does not happen in $\mathcal{I}$. Define the projection $\pi : \mathcal{I} \to \mathbb{R}^n$ by $\pi(p, x) = x$.

Proposition 24. *$D_x u$ is continuous at all points $x \in \pi(\mathcal{I})$.*

Proof. This follows from the observation that $x \in \pi(\mathcal{I})$ is a regular point. ■

The objective of this section is to improve the previous result and prove the, perhaps surprising, result that the invariant set is a Lipschitz graph. To prove this result we start by quoting a lemma from [Mat91].

Lemma 21. *Suppose $K > 0$, then there exist $\epsilon, \delta, \eta, C > 0$ such that for any*

$$\alpha, \beta : [t_0 - \epsilon, t_0 + \epsilon] \to \mathbb{R}^n$$

solutions of the Euler-Lagrange equation with $|\dot\alpha(t)|, |\dot\beta(t)| \leq K$, $|\alpha(t_0) - \beta(t_0)| \leq \delta$ and

$$|\dot\alpha(t_0) - \dot\beta(t_0)| \geq C|\alpha(t_0) - \beta(t_0)|,$$

then there exist C^1 curves

$$a, b : [t_0 - \epsilon, t_0 + \epsilon] \to \mathbb{R}^n$$

with $a(t_0 - \epsilon) = \alpha(t_0 - \epsilon)$, $b(t_0 - \epsilon) = \beta(t_0 - \epsilon)$, $a(t_0 + \epsilon) = \beta(t_0 + \epsilon)$, $b(t_0 + \epsilon) = \alpha(t_0 + \epsilon)$, and

$$A(\alpha) + A(\beta) - A(a) - A(b) \geq \eta |\dot\alpha(t_0) - \dot\beta(t_0)|^2,$$

where

$$A(x) = \int_{t_0 - \epsilon}^{t_0 + \epsilon} L(x(s), \dot x(s), s)\, ds,$$

for any C^1 curve $x(\cdot)$.

Proof. The idea of the proof is to construct explicitly a and b from α and β. More precisely, let $\mu(t) = \frac{\alpha(t)+\beta(t)}{2}$, and define

$$a(t) = \mu(t) + \\ + \frac{-t + t_0 + \epsilon}{2\epsilon}\left[\alpha(t_0 - \epsilon) - \mu(t_0 - \epsilon)\right] + \frac{t - t_0 + \epsilon}{2\epsilon}\left[\beta(t_0 + \epsilon) - \mu(t_0 + \epsilon)\right],$$

and

$$b(t) = \mu(t) + \\ + \frac{-t + t_0 + \epsilon}{2\epsilon}\left[\beta(t_0 - \epsilon) - \mu(t_0 - \epsilon)\right] + \frac{t - t_0 + \epsilon}{2\epsilon}\left[\alpha(t_0 + \epsilon) - \mu(t_0 + \epsilon)\right].$$

Then by analyzing carefully $A(a)$ and $A(b)$ one can arrive to the desired inequality - for details consult [Mat91]. ■

Theorem 25. *The invariant set $\mathcal{I}$ is a Lipschitz graph, more precisely, there exists a constant C such that if $(x,p),(x',p') \in \mathcal{I}$ then $|p-p'| \leq C|x-x'|$.*

Proof. We will prove this theorem by contradiction. For simplicity of notation we consider the time-independent case. Assume the theorem is false. Then for any n, there exists $(x_n,p_n),(x_n',p_n') \in \mathcal{I}$ such that $|p_n - p_n'| \geq n|x_n - x_n'|$. Since p_n, p_n' are bounded this imply $|x_n - x_n'| \to 0$ and so $|p_n - p_n'| \to 0$ (by proposition 24). Define $v_n = D_pH(p_n,x_n)$, $v_n' = D_pH(p_n',x_n')$. Since H is strictly convex we have $|v_n - v_n'| \geq \theta n|x_n - x_n'|$, for some $\theta > 0$ and all n large. Indeed, since $x_n \to x_n'$

$$|v_n - v_n'| \geq |D^2_{pp}H(p_n - p_n')| - |D^2_{px}H(x_n - x_n')| + o(|x_n - x_n'|) \geq \theta n|x_n - x_n'|.$$

Let $\epsilon, \delta, \eta, C > 0$ be as in the previous lemma (with K is a uniform bound on $D_p(p,x)$ for $(x,p) \in \mathcal{I}$). Choose n so large that $\theta n > C$. Let α and β be the solutions of the Euler Lagrange equations that pass through x_n and x_n' when $t_0 = 0$, and with $\dot\alpha(0) = v_n$, $\dot\beta(0) = v_n'$. Then

$$u(\alpha(-\epsilon)) = A(\alpha) + u(\alpha(\epsilon)) \qquad u(\beta(-\epsilon)) = A(\beta) + u(\beta(\epsilon)),$$

and, because this trajectories are absolute minimizers

$$u(\alpha(-\epsilon)) \leq A(a) + u(\beta(\epsilon)) \qquad u(\beta(-\epsilon)) \leq A(b) + u(\alpha(\epsilon)).$$

By combining these equations we get

$$A(\alpha) + A(\beta) \leq A(a) + A(b).$$

Thus

$$0 < \eta|v_n - v_n'|^2 \leq A(\alpha) + A(\beta) - A(a) - A(b) \leq 0,$$

which is a contradiction. ■

5.4 Uniqueness and decomposition

In general, the solution of the cell problem (5.2) is not unique (see example 12). In this section we state conditions under which the viscosity solution of (5.2) is unique (up

to constants), proving a partial uniqueness result on the invariant set. Furthermore, we show that whenever the solution is not unique the invariant set $\mathcal{I}$ can be decomposed in a suitable way. In the next chapter we will study again this problem using measure theoretic tools and improve some of this results.

First consider the case in which there exists a smooth solution and the rotation vector satisfies non-resonance conditions. In this situation the viscosity solution is indeed unique.

Theorem 26. *Suppose that $\overline{H}$ is differentiable at P_0 and the rotation vector $D_P\overline{H}(P_0)$ is linearly independent over the rationals. Assume further that in a neighborhood of P_0 there exists a smooth (both in x and P) solution u_1 of the cell problem (5.2) such that the change of variables $X(x) = x + D_P u$ has inverse $x(X) = X + \phi(X)$ with ϕ continuous. Then, up to a constant, the viscosity solution of the cell problem is unique.*

Proof. Suppose u_2 is another viscosity solution of the cell problem. We may assume $u_1 \leq u_2$ with equality $u_1(x_0) = u_2(x_0)$ at a point x_0, by adding a constant if necessary. Let V_1 and V_2 be solutions of the time dependent problem

$$-D_t V_i + H(P + D_x V_i, x) = 0,$$

with terminal value $V_i(x, T) = u_i$, $i = 1, 2$. By monotonicity $V_1 \leq V_2$. To prove the other inequality, take an optimal trajectory $x_1^*(\cdot)$ for V_1 with $x(T) = x_0$. This is possible because u_1 is differentiable and therefore there there exists a unique optimal trajectory that passes through each point. Then

$$V_2(x_1^*(0), 0) \leq J_1(x_1^*(0), 0; \dot{x}_1^*),$$

since at $x_1^*(T) \equiv x_0$, $V_2(x_0, T) = V_1(x_0, T)$. But

$$J_1(x_1^*(0), 0; \dot{x}_1^*) = V_1(x_1^*(0), 0; \dot{x}_1^*)$$

, consequently, along this trajectory, $V_2 \leq V_1$. In particular, since the rotation vector is incommensurable and $x(X)$ is continuous the trajectory is dense. Thus $V_1 = V_2$ everywhere, and, consequently, $u_1 = u_2$. ■

Corollary 5. *Suppose $H_0(p) + \epsilon H_1(p, x)$ for ϵ sufficiently small. Then, if the conditions of the KAM theorem hold, the KAM tori agree with the the invariant sets obtained via viscosity solutions.*

Now we will study the case in which the cell problem has more than one solution (modulo constants). Let u_1 and u_2 be genuinely different solutions, i.e., $u_1 - u_2$ not constant. Define

$$U^+\{x : u_1(x) < u_2(x)\}, \qquad U^-\{x : u_1(x) > u_2(x)\}.$$

Let

$$\mathcal{G}^+ = \{(x,p) : x \in U^+, p = D_x u_1(x), u_1 \text{ differentiable at } x\},$$

and, similarly,

$$\mathcal{G}^- = \{(x,p) : x \in U^-, p = D_x u_2(x), u_2 \text{ differentiable at } x\}.$$

Proposition 25. *$\mathcal{G}^\pm$ are backwards invariant sets.*

Proof. It suffices to prove that $\mathcal{G}^-$ is backwards invariant since the other part of the proof is completely symmetric. Recall that the action kernel $K(x,y;t)$, defined in equation (2.10), is

$$K(x,y;t) = \inf_{x(t)=x, x(t_1)=y} \int_t^{t_1} L(x(s), \dot{x}(s), s) ds.$$

Choose any point $(x,p) \in \mathcal{G}^-$ Then, for any z

$$u_1(x) + \overline{H}(P)t \leq K(x,z;t) + u_1(z).$$

In particular we may choose Let $z = \Xi_t(x)$, where Ξ_t is the backwards Hamiltonian flow for u_2 starting at $(x,p) = (x, D_x u_2(x))$. Hence

$$u_2(x) + \overline{H}(P)t = K(x,z;t) + u_2(z).$$

Thus

$$0 < u_1(x) - u_2(x) \leq u_1(z) - u_2(z).$$

Therefore $u_1(z) > u_2(z)$. ∎

Theorem 27. *If u_1 and u_2 are viscosity solutions of the cell problem then also is $u = u_1 \wedge u_2$, where $(u_1 \wedge u_2)(x) = \min\{u_1(x), u_2(x)\}$.*

Proof. Note that u is a bounded Lipschitz function. Moreover

$$V(x,t) = u(x) + \overline{H}(P)t = \inf_y \left(K(x,y;t) + u(y)\right).$$

This implies that V is a viscosity solution of $-D_t V + H(P + D_x V, x) = 0$. But since $D_t V = \overline{H}(P)$ we get $H(P + D_x u, x) = \overline{H}(P)$. ∎

5.5 Asymptotics

In section 5.2 we proved that

$$\lim_{t\to\infty} \frac{x(t)}{t} \to D_P\overline{H},$$

for all trajectories in the invariant set $\mathcal{I}$. However such property does not give precise information about the difference $x(t) - D_P\overline{H}t$, for instance is it bounded as $t \to \infty$? In the integrable case we have $x(t) - D_P\overline{H}t = O(1)$, as $t \to \infty$. In this section we prove a new general estimate, namely $x(t) - D_P\overline{H}t = O(t^{1/2})$ using viscosity solutions (the best results so far were $x(t) - D_P\overline{H}t = O(t)$) If more assumptions are made in the regularity in P of $u(x, P)$ then this estimate can be improved.

Theorem 28. *Suppose $x(\cdot)$ is a solution of the Hamiltonian equations with initial conditions on $\mathcal{I}$. Furthermore assume $\overline{H}$ is twice differentiable at P. Let $\|x-y\| \equiv \min_{k\in\mathbb{Z}^n} |x-y+k|$, i.e., the "periodic distance" between x and y. Then there exists a constant C such that*

$$|x(t) - x(0) - D_P\overline{H}t| \le C\sqrt{\|x(t) - x(0)\|t}.$$

If there exists a continuous function ω, with $\omega(0) = 0$, such that

$$|u(x, P) - u(x, P')| \le \omega(|P - P'|),$$

then

$$|x(t) - x(0) - D_P\overline{H}t| \le \min_\delta \frac{C\|x(t) - x(0)\| \wedge \omega(\delta)}{\delta} + Ct\delta.$$

Finally if u is uniformly differentiable in P in $\pi(\mathcal{I})$ then

$$x(t) + D_P u(x(t), P) - x(0) - D_P u(x(0), P) - D_P\overline{H}t = 0.$$

Proof. Let $u(x, P)$ be the viscosity solution of $H(P + D_x u, x) = \overline{H}(P)$. Then, for some C^1 function $x^*(s)$ with $x^*(t) = y$ and $x^*(0) = x$

$$u(x, P) = u(y, P) + \int_0^t \left[L(x^*, \dot{x}^*) + P \cdot \dot{x}^* + \overline{H}(P)\right] ds.$$

For any other P'

$$u(x, P') \le u(y, P') + \int_0^t \left[L(x^*, \dot{x}^*) + P' \cdot \dot{x}^* + \overline{H}(P')\right] ds.$$

Therefore

$$u(x,P)-u(x,P')\geq u(y,P)-u(y,P')+\int_0^t\left[(P-P')\cdot\dot{x}^*+\overline{H}(P)-\overline{H}(P')\right]ds.$$

If $\overline{H}$ is twice differentiable (or at least $C^{1,1}$) at P

$$\overline{H}(P')\leq\overline{H}(P)-\omega\cdot(P'-P)+C|P'-P|^2,$$

where $\omega=-D_P\overline{H}(P)$. Thus

$$u(x,P)-u(x,P')+u(y,P')-u(y,P)\geq(P-P')\cdot\int_0^t[\dot{x}^*-\omega]-Ct|P'-P|^2.$$

The left hand side can be estimated by

$$u(x,P)-u(x,P')+u(y,P')-u(y,P)\leq C\|x-y\|.$$

If we let $|P-P'|=\sqrt{\frac{\|x-y\|}{t}}$ we get

$$|\int_0^t[\dot{x}^*-\omega]\,|\leq C\sqrt{\|x-y\|t}.$$

If there exists a continuous function ω, with $\omega(0)=0$, such that

$$|u(x,P)-u(x,P')|\leq\omega(|P-P'|)$$

then

$$u(x,P)-u(x,P')+u(y,P')-u(y,P)\leq C\|x-y\|\wedge\omega(|P-P'|).$$

Finally if u is uniformly differentiable in P we get

$$x(t)+D_Pu(x(t),P)-x(0)-D_Pu(x(0),P)-D_P\overline{H}t=0.$$

This last equality shows that whenever the change of coordinates

$$X=x+D_Pu$$

makes sense in the invariant set, we have $\dot{X}=D_P\overline{H}$. ∎

In the next chapter we will use measure-theoretic tools to show a modulus of continuity in P, $\omega(\delta)$, actually exists making more precise the previous theorem.

5.6 Adiabatic Invariants

This final section is dedicated to study of slowly varying Hamiltonians of the form $H(p, x, \epsilon t)$, where ϵ is a small parameter. We prove convergence of the effective Hamiltonian to an "averaged Hamiltonian" as $\epsilon \to 0$. Then we relate this result with the Hannay correction for the angle variables.

Assume, as before, that $L(x, v, t)$ is a periodic time-dependent Lagrangian, i.e. $L(x + k, v, t + 1) = L(x, v, t)$, for $k \in \mathbb{Z}^n$. Suppose $\epsilon > 0$, and consider the Lagrangian $L(x, \dot{x}, \epsilon t)$.

By the results in section 2.3 we know that there exists a function $V(x, t)$ and a constant $\overline{H}_\epsilon(P)$ such that $V(x, 0) = V(x, \frac{1}{\epsilon})$ and

$$-D_t V + H(P + D_x V, x, \epsilon t) = \overline{H}_\epsilon(P).$$

It is reasonable to expect that, as $\epsilon \to 0$, the solution of this problem looks more and more like an average of time independent problems

$$H(P + D_x V, x, \epsilon t) = \overline{H}(P, \epsilon t).$$

The next proposition is a first step in this direction.

Proposition 26. *As $\epsilon \to 0$ we have $\overline{H}_\epsilon(P) \to \overline{H}_0(P)$ where*

$$\overline{H}_0(P) = \int_0^1 \overline{H}(P, s) ds.$$

More precisely, $|\overline{H}_\epsilon(P) - \overline{H}_0(P)| \leq C\epsilon^{\frac{1}{2}}$.

Proof. Let k be an integer and $\epsilon = \frac{1}{k^2}$. Divide the interval $[0, k^2]$ into equal k subintervals $I_j = [t_j, t_{j+1}]$, $0 \leq j < k$, and $t_j = jk$. Let u_ϵ be an additive eigenvalue for the time dependent terminal value

$$-D_t V + H(P + D_x V, x, \epsilon t) = 0, \qquad V(x, k^2) = u_\epsilon(x),$$

i.e., V satisfies $V(x, 0) = u_\epsilon(x) - k^2 \overline{H}_\epsilon(P)$. Then

$$-k^2 \overline{H}_\epsilon(P) + u_\epsilon(x^*(0)) = u_\epsilon\left(x^*\left(k^2\right)\right) + \int_0^{k^2} \left[L(x^*(s), \dot{x}^*(s), \epsilon s) + P \cdot \dot{x}^*(s)\right] ds,$$

for some minimizing curve x^*.

For $0 \leq j < k$,

$$V(x,t_j) = \inf_{x(\cdot)} \left[V(x(t_{j+1}),t_{j+1}) + \int_{t_j}^{t_{j+1}} [L(x(s),\dot{x}(s),\epsilon s) + P\cdot \dot{x}(s)]\, ds \right].$$

Along the optimal trajectory x^*,

$$V(x^*(t_j),t_j) = V(x^*(t_{j+1}),t_{j+1}) + \int_{t_j}^{t_{j+1}} [L(x^*(s),\dot{x}^*(s),\epsilon s) + P\cdot \dot{x}^*(s)]\, ds.$$

Let u_{j+1} be a periodic viscosity solution of

$$H(P + D_x u_{j+1}, x, \frac{t_{j+1}}{k^2}) = \overline{H}(P, \frac{t_{j+1}}{k^2}).$$

By adding an appropriate constant to u_{j+1} we may assume that

$$\sup_x |V(x,t_{j+1}) - u_{j+1}(x)| \leq C,$$

for some constant C that does not depend on k or j. Also

$$\sup_{x,|v|<R}\ \sup_{t_j\leq s\leq t_{j+1}} |L(x,v,\frac{s}{k^2}) - L(x,v,\frac{t_{j+1}}{k^2})| \leq \frac{C}{k},$$

here R is an upper bound for $|\dot{x}^*|$. Therefore

$$\sup_x |V(x,t_j) - u_{j+1}(x) + \overline{H}(P,\frac{t_{j+1}}{k^2})k| \leq C.$$

Thus

$$\sup_x |V(x,t_j) - V(x,t_{j+1}) + \overline{H}(P,\frac{t_{j+1}}{k^2})k| \leq C.$$

Hence

$$\sup_x |V(x,0) + \sum_{j=0}^{k-1} \overline{H}(P,\frac{t_{j+1}}{k^2})k| \leq Ck.$$

and so

$$\overline{H}_\epsilon(P) = \frac{1}{k}\sum_{j=0}^{k-1} \overline{H}(P,\frac{j+1}{k}) + O(\frac{1}{k}).$$

Note that $\overline{H}(P,t)$ is Lipschitz in t. Therefore

$$|\int_0^1 \overline{H}(P,t)dt - \frac{1}{k}\sum_{j=0}^{k-1} \overline{H}(P,\frac{j+1}{k})| \leq \frac{C}{k}.$$

Thus $|\overline{H}_\epsilon(P) - \overline{H}_0(P)| \leq C\epsilon^{\frac{1}{2}}$, as $\epsilon \to 0$. ■

For a PDE proof of the previous theorem consult [EG99b].

If $u(x,P;t)$ is Lipschitz in t we can improve the rate of convergence of $\overline{H}_\epsilon$ to $\overline{H}$.

Proposition 27. *Suppose $u(x,P;t)$ is Lipschitz and periodic in t, and, for each fixed t, a viscosity solution of*

$$H(P+D_xu,x,t)=\overline{H}(P,t).$$

Then $|\overline{H}_\epsilon(P)-\overline{H}_0(P)|\leq C\epsilon$.

PROOF. Let k be an integer and $\epsilon=\frac{1}{k}$. Divide the interval $[0,k]$ into k subintervals. For $0\leq j\leq k$ let $t_j=j$ and $u_j(x)=u(x,\epsilon t_j)$. Define $\gamma_j=\sum_{i=j+1}^k H(P,t_i)$.

Lemma 22. *For $0\leq j\leq k$*

$$|V(x,t_j)-u_j(x)+\gamma_j|\leq C(1+\frac{k-j}{k}).$$

PROOF. By adding a suitable constant to u we may assume

$$|V(x,t_k)-u(x,t_k)|\leq C,$$

so the assertion is true for $j=k$. Now note that

$$\sup_{x,|v|<R}\sup_{t_j\leq s\leq t_{j+1}}|L(x,v,\frac{s}{k})-L(x,v,\frac{t_{j+1}}{k})|\leq\frac{C}{k},$$

where R is an upper bound for $|\dot{x}^*|$ along optimal trajectories. Therefore

$$\sup_x|V(x,t_{j-1})-u_{j-1}(x)+\gamma_{j-1}|\leq$$
$$\leq\sup_x|V(x,t_j)-u_j(x)+\gamma_j|+\sup_x|u_j-u_{j-1}|+\frac{C}{k}.$$

The Lipschitz hypothesis on u implies

$$\sup_x|u_j-u_{j-1}|\leq\frac{C}{k},$$

this proves the lemma by induction. ∎

Hence

$$V(x,0)=\gamma_0+O(1),$$

and so $|\overline{H}_\epsilon-\overline{H}_0|\leq C\epsilon$. ∎

Now we proceed with some formal derivations. Our objective is to show that the difference $\overline{H}_\epsilon(P)-\overline{H}(P)$ encodes the Hannay correction for the angle variables.

Suppose

$$\frac{\overline{H}_\epsilon(P)-\overline{H}_0(P)}{\epsilon}\to\overline{H}_1(P).$$

Take an optimal trajectory x^ϵ. Then

$$\int_0^{\frac{t}{\epsilon}} L(\frac{x^\epsilon}{\epsilon}, \dot{x}^\epsilon, s) + P \cdot \dot{x}^\epsilon = \left(\frac{\overline{H}_0(P)}{\epsilon} + \overline{H}_1(P) + o(1)\right) t + O(1),$$

where the $o(1)$ term is a ϵ dependent term, as $\epsilon \to 0$ and the $O(1)$ is a time dependent term, as $t \to +\infty$. Thus, for P' close to P,

$$-\frac{P - P'}{t} \int_0^{\frac{t}{\epsilon}} \dot{x}^\epsilon \leq \frac{\overline{H}_0(P) - \overline{H}_0(P')}{\epsilon} + \overline{H}_1(P) - \overline{H}_1(P') + o(1) + O(\frac{1}{t}).$$

Let $t \to +\infty$ and then $P \to P'$ to get

$$-\lim_{t\to\infty} \frac{1}{t} \int_0^{\frac{t}{\epsilon}} \dot{x}^\epsilon = \frac{D_P \overline{H}_0(P)}{\epsilon} + D_P \overline{H}_1(P) + o(1).$$

The term $D_P \overline{H}_1$ is the analog of the Hannay angle (the minus sign appears due to the time reversal).

The main problem with this formal derivations is that our estimates yield bounds on H_1, at least when the cell problem admits a family of Lipschitz (in t) solutions. However we do not have any information concerning the differentiability of H_1 in P.

Chapter 6

Invariant Measures

The main goal of this chapter is to understand regularity properties of viscosity solutions of Hamilton-Jacobi equations in terms of certain probability measures (Mather measures) supported in the invariant sets defined in the previous chapter. These measures are important by themselves because they describe the asymptotic behavior of certain trajectories of Hamiltonian systems. We derive many of its properties using new viscosity solutions techniques.

J. Mather [Mat91] considered the problem of minimizing the functional

$$A[\mu] = \int L d\mu,$$

over the set of probability measures μ supported on $\mathbb{T}^n \times \mathbb{R}^n$ that are invariant under the flow associated with the Euler-Lagrange equation

$$\frac{d}{dt}\frac{\partial L}{\partial v} - \frac{\partial L}{\partial x} = 0.$$

Here $L = L(x, v)$ is the Legendre transform of H, and $\mathbb{T}^n$ the n-dimensional torus, identified with $\mathbb{R}^n/\mathbb{Z}^n$ whenever convenient.

One can add also the additional constraint

$$\int v d\mu = \omega,$$

restricting the class of admissible measures to the ones with an average rotation number ω. It turns out [Mn91] that this constrained minimization problem can be solved by adding a Lagrange multiplier term:

$$A_P[\mu] = \int L(x, v) + P v d\mu.$$

The main idea is that instead of studying invariant sets one should consider invariant probability measures. The supports of such measures correspond to the invariant sets (tori) defined by $P = constant$ given by the classical theory.

In this chapter we present a construction of such invariant measures using Young measures. Then by applying viscosity solution methods we prove important properties. Namely, we show that the support of this measures is contained in the invariant sets we obtained before. Also, given a minimizing measure, it is possible to associate to it a viscosity solution of the Hamilton-Jacobi equation that records all the information carried out by the measure. Regularity estimates for the solutions of the equation

$$H(P + D_x u, x) = \overline{H}(P) \tag{6.1}$$

are given, improving some of the results in the previous chapter. We also study the connection with the adiabatic invariants theory. Finally we clarify the connection between Aubry-Mather theory and the existence of viscosity solutions of (6.1) by proving that one problem is the dual of the other, using Fenchel-Legendre duality.

Many of the results in sections 6.2, 6.3, 6.6, and 6.7 as well as those concerning difference quotients appear also in the joint papers [EG99a] and [EG99b]. These results benefited immensely from the help of my advisor L. C. Evans.

6.1 Weak convergence

In this section we present some background material in weak convergence and Young measures. Details can be found in [Eva90], for instance.

Suppose f_k is a bounded sequence of real-valued functions defined on a bounded set $U \subset \mathbb{R}^n$. Then there exists a subsequence (still denoted by f_k) and a bounded function $f : U \to \mathbb{R}$ such that for any $\phi \in C(U)$

$$\lim_{k\to+\infty} \int_U \phi f_k = \int_U \phi f.$$

However we cannot conclude that for any ψ we have

$$\lim_{k\to+\infty} \int_U \psi(f_k) = \int_U \psi(f).$$

Even in simple examples the previous identity fails.

Example 13. Let $U = [0, 2\pi]$, $f_k(x) = \sin(kx)$. Then

$$\int_\Omega \phi f_k \to 0,$$

and so $f = 0$. However

$$\lim_{k\to+\infty} \int_U f_k^2 = \pi.$$

Theorem 29. *Suppose f_k is a bounded sequence of real-valued functions defined on a bounded set $U \subset \mathbb{R}^n$. Then there exists a measure μ defined on $\Omega = U \times \mathbb{R}^n$ such that for any $\psi \in C(\Omega)$*

$$\lim_{k\to+\infty} \int_U \psi(x, f_k(x)) = \int_\Omega \psi(x, y) d\mu(x, y).$$

Furthermore, for almost every x, there exists a probability measure ν_x such that

$$\int_\Omega \psi(x, y) d\mu(x, y) = \int_U \left(\int_{\mathbb{R}^n} \psi(x, y) d\nu_x(dy) \right) dx.$$

Proof. Consult [Eva90], for example. ■

The measures ν_x are called the *Young Measures* associated with the sequence f_k.

6.2 Young Measures for Hamiltonian Systems

Young measures are used to encode fast oscillations that may be lost under weak convergence. This is why they are natural objects to study Hamiltonian systems of the form $H(p, \frac{x}{\epsilon}, \frac{t}{\epsilon})$, as $\epsilon \to 0$. Here $H(p, y, \tau)$ is 1−periodic in τ, $\mathbb{Z}^n$ periodic in y, satisfying the usual coercivity, convexity, and smoothness hypothesis, and $\epsilon > 0$ is a (small) parameter.

Let $V(x, t)$ be a periodic solution (periodic both in x and t) of the Hamilton-Jacobi equation

$$-D_t V + H(P + D_x V, x, t) = \overline{H}(P).$$

For each ϵ, let $x^\epsilon(\cdot)$ be a minimizing trajectory for the optimal control problem and $p^\epsilon(\cdot)$ the corresponding adjoint variable. Then, for any s and t

$$V(x^\epsilon(s), s) = \int_s^t \left[L(x^\epsilon(r), \dot{x}^\epsilon(r), r) - P \cdot \dot{x}^\epsilon(r) - \overline{H}(P) \right] dr + V(x^\epsilon(t), t).$$

Theorem 30 (Mather measures). *For almost every $0 \le t \le 1$ there exists a measure (Mather measure) ν_t such that for any, smooth and periodic in y and τ, function $\Phi(p, y, \tau, t)$*

$$\Phi(p^\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}, t) \rightharpoonup \overline{\Phi}(t),$$

with $\overline{\Phi}(t) = \int \Phi(p, y, \tau, t) d\nu_t(p, y, \tau)$. *More precisely, for any smooth function* $\varphi(t)$

$$\int_0^1 \varphi(t)\Phi(p^\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}, t)dt \to \int_0^1 \varphi(t)\overline{\Phi}(t)dt,$$

as $\epsilon \to 0$ *(through some subsequence, if necessary).*

PROOF. In general, the sequence $(\frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon})$ is not bounded. However if we consider $\frac{x^\epsilon}{\epsilon} \mod \mathbb{Z}^n$ and $\frac{t}{\epsilon} \mod 1$, this sequence is clearly bounded, and since, by hypothesis, Φ is periodic this does not change the result. Also p^ϵ can be uniformly bounded independently of ϵ. Thus, by the results of the previous section, we can find Young measures ν_t with the required properties. ■

We now prove that these measures are supported on the invariant set.

Proposition 28. *Let* V *be a periodic (in* x *and* t*) solution of* $-D_tV + H(P + D_xV, x, t) = \overline{H}(P)$ *and* ν_t *an associated Mather measure. Then* $p = P + D_xV$ ν_t *a.e..*

PROOF. The measure ν_t was obtained as a weak limit of measures supported on the closure of $p = P + D_xV$, for some fixed V. Thus the support of the limiting measure should also be contained on the closure of $p = P + D_xV$. ■

Latter we will improve this result and show that $p = P + D_xV$, ν_t a.e. for any (periodic in x and t) solution of $-D_tV + H(P + D_xV, x, t) = \overline{H}(P)$.

It turns out that these measures have important special properties, not only they are invariant measures but also, in a suitable sense, action-minimizing measures. First we will prove the invariance property, and then, in section 6.4, address the minimizing property.

Theorem 31. *For any smooth periodic (in* y *and* τ*) function* $\phi(p, y, \tau)$

$$\int D_y\phi(p, y, \tau)D_pH(p, y, \tau) - D_p\phi(p, y, \tau)D_xH(p, y, \tau) + D_\tau\phi(p, y, \tau)d\nu_t = 0.$$

PROOF. Note that

$$\frac{d\phi(p_\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon})}{dt} = \frac{D_y\phi D_pH - D_p\phi D_xH + D_\tau\phi}{\epsilon}.$$

Let

$$\psi(p, y, \tau) = D_y\phi(p, y, \tau)D_pH(p, y, \tau) - D_p\phi(p, y, \tau)D_xH(p, y, \tau) + D_\tau\phi(p, y, \tau).$$

Then

$$\int_0^1 \psi(p_\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon})dt = O(\epsilon)$$

which implies

$$\int \psi d\nu_t = 0,$$

for almost every t. ■

Corollary 6. *The measure ν_t is invariant.*

PROOF. This is just a restatement the previous theorem since a measure ν is invariant provided

$$\frac{d}{d\mu}\int \phi(x(x;\tau+\mu), p(p;\tau+\mu), \tau+\mu)d\nu = 0,$$

for all C^1 functions $\phi(x,p,\tau)$, periodic in x and τ. ■

Now we prove a variant of theorem 31 for the case where ϕ only depends on x and is Lipschitz continuous. The main technical problem that we need to solve is to make sense of what $D_x\phi(x)$ means in the support of the measure. Although ϕ is differentiable almost everywhere with respect to Lebesgue measure, the measure ν_t may be supported exactly where the derivative does not exist.

We say a function $\psi : \mathbb{R}^n \to \mathbb{R}^n$ is a *version* of $D_x\phi$ if the graph of ψ is contained in the vertical convex hull of the closure of the graph of $D_x\phi$. More precisely if

$$\psi(x) \in \mathcal{D}_x\phi(x),$$

where

$$\mathcal{D}_x\phi(x) = \mathrm{co}\{p : p = \lim_{n\to\infty} D_x\phi(x_n), \text{ with } x_n \to x, \phi \text{ differentiable at } x_n\}.$$

Proposition 29. *Assume that ϕ has the property that if $x_n \to x$ and ϕ is differentiable at x and at each x_n then $D_x\phi(x_n) \to D_x\phi(x)$. Furthermore any version of $D_x\phi$ coincides with the derivative of ϕ at all points where ϕ is differentiable.*

PROOF. The hypothesis on ϕ implies immediately that

$$\mathcal{D}_x\phi(x) = \{D_x\phi(x)\},$$

if ϕ is differentiable at x. ■

The solutions of $H(P + D_xu, x) = \overline{H}(P)$ have this property but this is not true for general Lipschitz functions.

Theorem 32. *If $\phi : \mathbb{R}^n \to \mathbb{R}$ is a Lipschitz function then, for almost every t, there exists a version of $D_x\phi$ such that*

$$\int D_x\phi(y) D_pH(p, y, \tau) d\nu_t = 0.$$

Proof. Consider a sequence $\epsilon_n \to 0$ such that

$$\int_0^1 \varphi(p^{\epsilon_n}, \frac{x^{\epsilon_n}}{\epsilon_n}, \frac{t}{\epsilon_n}) dt \to \int_0^1 \int \varphi d\nu_t dt,$$

for all smooth, and periodic in y and τ functions φ. Let z_n be any sequence such that $z_n \to 0$. Then

$$\int_0^1 \varphi(p^{\epsilon_n}, z_n + \frac{x^{\epsilon_n}}{\epsilon_n}, \frac{t}{\epsilon_n}) dt \to \int_0^1 \int \varphi d\nu_t dt.$$

Since ϕ is differentiable almost everywhere, it is possible to choose $z_n \to 0$ such that, for each n, $D_x\phi(z_n + \frac{x^{\epsilon_n}}{\epsilon_n})$ is defined for almost every t. Now consider the sequence of vector-valued measures μ_n defined by

$$\int \zeta(p, y, \tau, t) d\mu_n = \int_0^1 D_x\phi(z_n + \frac{x^{\epsilon_n}}{\epsilon_n}) \cdot \zeta(p^{\epsilon_n}, \frac{x^{\epsilon_n}}{\epsilon_n}, \frac{t}{\epsilon_n}, t) dt,$$

for all vector valued smooth, and periodic in y and τ functions ζ. Since $D_x\phi$ is bounded, we can extract subsequence, also denoted by μ_n, that converges weakly to a vector measure μ. Also $d\mu = d\eta_t dt$, for some family of vector measures η_t.

Then, for almost every t, $\eta_t << \nu_t$, in the sense that for any set A, $\nu_t(A) = 0$ implies that the vector $\eta_t(A) = 0$. Therefore, by Radon-Nikodym theorem, we have $d\eta_t = \psi_t d\nu_t$, for some $L^1(\nu_t)$ function ψ_t. Now it is clear that ψ_t is, for each (x, p), in $\mathcal{D}_x\phi$, and so it is a version of $D_x\phi$.

Finally, to see that

$$\int \psi_t D_pH(p, y, \tau) d\nu_t = 0,$$

we just have to observe that

$$\int D_pH(p, y, \tau) d\mu_n = O(\epsilon_n)$$

and so

$$0 = \int D_pH(p, y, \tau) d\eta_t = \int \psi_t D_pH(p, y, \tau) d\nu_t,$$

for almost every t. ■

6.3 Viscosity Solutions - I

Assume that V^ϵ is a periodic viscosity solution of the Hamilton-Jacobi equation

$$-D_t V^\epsilon + H(P + D_x V^\epsilon, \frac{x}{\epsilon}, \frac{t}{\epsilon}) = \overline{H}(P)$$

with terminal data $V^\epsilon(x,1) = \epsilon V(\frac{x}{\epsilon}, 1)$. Let $u(x) = V(x,1)$ be an additive eigenvalue for the terminal value problem $-D_t V + H(P + D_x V, x, t) = 0$, $V(x,1) = u(x)$. Let x^ϵ be an optimal trajectory.

Proposition 30. *There exists a function $X(\cdot)$ such that, as $\epsilon \to 0$, $x^\epsilon \to X(\cdot)$ uniformly; thus $\dot{x}^\epsilon \rightharpoonup \dot{X}$. Additionally*

$$L\left(\dot{x}^\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}\right) \rightharpoonup -P \cdot \dot{X} - \overline{H}(P).$$

Furthermore

$$-P \cdot \dot{X} - \overline{H}(P) = \overline{L}(\dot{X}),$$

where

$$\int L d\nu_t = \overline{L}(\dot{X}),$$

and $\overline{L} = \overline{H}^*$. Finally

$$\dot{X}(t) = \int D_p H d\nu_t.$$

Proof. Since x^ϵ is uniformly bounded and equicontinuous, there exists a function $X(\cdot)$ such that

$$x^\epsilon(t) \to X(t),$$

uniformly as $\epsilon \to 0$ (possibly after extracting a subsequence $\epsilon_k \to 0$). And consequently

$$\dot{x}^\epsilon(t) \rightharpoonup \dot{X}(t).$$

Observe that

$$V^\epsilon(x^\epsilon(t), t) \to 0,$$

uniformly as $t \to 0$. Thus

$$D_t V^\epsilon(x^\epsilon(t), t) + D_x V^\epsilon(x^\epsilon(t), t) \cdot \dot{x}^\epsilon(t) \rightharpoonup 0.$$

Along minimizing trajectories

$$D_x V^\epsilon(x^\epsilon(t), t) \cdot \dot{x}^\epsilon = -H\left(P + p^\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}\right) - L\left(\dot{x}^\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}\right) - P \cdot \dot{x}^\epsilon,$$

and

$$D_t V^\epsilon = H\left(P + p^\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}\right) - \overline{H}(P).$$

Therefore

$$-\overline{H}(P) - L\left(\dot{x}^\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}\right) - P \cdot \dot{x}^\epsilon \rightharpoonup 0,$$

or equivalently

$$L\left(\dot{x}^\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}\right) \rightharpoonup -\overline{H}(P) - P \cdot \dot{X}.$$

Thus

$$\overline{L}(\dot{X}) \equiv \int L d\nu_t = -\overline{H}(P) - P \cdot \dot{X}.$$

To prove $\overline{L} = \overline{H}^*$ it suffices to check that for any $\tilde{P}$

$$\overline{L}(\dot{X}) \geq -\overline{H}(\tilde{P}) - \tilde{P} \cdot \dot{X}.$$

Take $\tilde{V}^\epsilon$ to be a periodic viscosity solution of

$$-D_t \tilde{V}^\epsilon + H\left(\tilde{P} + D_x \tilde{V}^\epsilon, \frac{x}{\epsilon}, \frac{t}{\epsilon}\right) = \overline{H}(\tilde{P}),$$

with $\tilde{V}^\epsilon \to 0$ uniformly. Then, for almost every $y \in \mathbb{R}^n$,

$$D_t \tilde{V}^\epsilon(\epsilon^2 y + x^\epsilon(t), t) + D_x \tilde{V}^\epsilon(\epsilon^2 y + x^\epsilon(t), t) \cdot \dot{x}^\epsilon(t) \rightharpoonup 0,$$

where x^ϵ is the optimal trajectory for V^ϵ. Now note that

$$D_x \tilde{V}^\epsilon(\epsilon^2 y + x^\epsilon(t), t) \cdot \dot{x}^\epsilon \geq -H\left(\tilde{P} + D_x \tilde{V}^\epsilon, \epsilon y + \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}\right) - L\left(\dot{x}^\epsilon, \epsilon y + \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}\right) - \tilde{P} \cdot \dot{x}^\epsilon,$$

and

$$D_t \tilde{V}^\epsilon = H\left(\tilde{P} + D_x \tilde{V}^\epsilon, \epsilon y + \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}\right) - \overline{H}(\tilde{P}).$$

Thus

$$\overline{L}(\dot{X}) \geq -\overline{H}(\tilde{P}) - \tilde{P} \cdot \dot{X},$$

by passing to the limit. Finally note that

$$\dot{x}^\epsilon = -D_p H\left(p^\epsilon, \frac{x^\epsilon}{\epsilon}\right).$$

Thus

$$\dot{x}^\epsilon \rightharpoonup \int D_p H d\nu_t = \dot{X}.$$

■

Corollary 7. *Suppose $\overline{H}(P)$ is differentiable. Then*

$$\lim_{t\to\infty} \frac{x(t)}{t} = -D_P\overline{H}(P),$$

for all minimizing trajectories.

PROOF. Since

$$\lim_{t\to\infty} \frac{x(t)}{t} = \dot{X},$$

it suffices to observe that $\dot{X} = -D_P\overline{H}$. This happens because $\overline{L}^* = \overline{H}$ and $\overline{L}(\dot{X}) = -P\cdot\dot{X} - \overline{H}(P)$. ■

Using viscosity solutions methods, we prove that the support of these measures is contained in the closure of the graph of $P+D_xu$, where u is any solution of $H(P+D_xu,x) = \overline{H}(P)$.

Theorem 33. *Let u be any periodic viscosity solution of $H(P+D_xu,x) = \overline{H}(P)$. Then for almost every t, u is differentiable ν_t a.e. and $p = P + D_xu$.*

PROOF. By strict convexity of H, there exists γ such that

$$\gamma|p-q|^2 + H(p,x) + D_pH(p,x)\cdot(q-p) \leq H(q,x).$$

Choose q to be a version of $P + D_xu$ (we will write, for simplicity, $q = P + D_xu$). Then

$$\gamma\int|p-(P+D_xu)|^2 d\nu_t \leq \int H(P+D_xu,x)d\nu_t - \int H(p,x)d\nu_t - \\ - \int D_pH(p,x)(P+D_xu-p)d\nu_t.$$

but the right-hand side integrates to zero because

$$H(P+D_xu,x) \leq \overline{H}(P)$$

everywhere (since we are using a version of D_xu and by convexity of H),

$$\int pD_pH - Hd\nu_t = \overline{L} \qquad P\int D_pHd\nu_t = -P\dot{X}.$$

Thus $\overline{H} + \overline{L} + P\dot{X} = 0$, and

$$\int D_pHD_xud\nu_t = 0.$$

To show that we do not really need a version of D_xu and instead u is differentiable on the support of ν, note that any version of $P + D_xu$ is in the vertical convex hull of the

closure of the graph of $P + D_x u$. However any point (x, p) in the closure of the graph of $P + D_x u$

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

but if the version of $\tilde{p} = P + D_x u$ were a strict linear combination of points in the closure of the graph of $P + D_x u$

$$H(\tilde{p}, x) < \overline{H}(P),$$

because H is strictly convex. This implies that for almost everywhere in the support of ν_t, $p = P + D_x u$, where $D_x u$ is in the closure of the derivative of u. To prove differentiability it suffices to observe that the backwards flow maps any point into a point where u is differentiable. Since the measure is invariant we conclude that u is ν_t almost everywhere differentiable. ■

Observe that by taking minimizing trajectories of the Lagrangian $L(\frac{x}{\epsilon}, \dot{x})$ we constructed flow invariant measures supported on the graph $p = P + D_x u$. Since there are several solutions (even modulo constants) of the equation

$$H(P + D_x u, x) = \overline{H}(P),$$

the previous result is indeed a partial uniqueness result for the solutions of this equation, i.e., if u_1 and u_2 are two different solutions we have $D_x u_1 = D_x u_2$ ν_t almost everywhere, for any such measure ν.

One could think that the set $\mathcal{I}$ could be identified with the union of the supports of all possible measures ν_t. However this is not true, see example 12, and $\mathcal{I}$ may be in fact larger.

6.4 Connection with Aubry-Mather theory

Next we show that Mather's definition of minimizing invariant measures [Mat91] is equivalent to our definition.

We say that a measure μ is a minimizing measure with rotation number ω if

$$\int L d\mu = \inf_{\nu} \int L d\nu,$$

where the infimum is taken over all probability measures ν invariant under the Euler-Lagrange equations, and satisfying

$$\int v d\nu = \omega.$$

It can be proved that for each ω there exists a P and a measure μ that minimizes the functional

$$\int L + Pvd\mu.$$

In [Mn96] these measures were obtained as weak limits of minimizing trajectories for the Lagrangian $L(x,v) + Pv$.

Theorem 34. *Suppose μ a Mather measure, associated with a periodic viscosity solution of*

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

Then μ minimizes

$$\int L + P \cdot vd\eta,$$

over all invariant probability measures η.

PROOF. If the claim were false, there would be an invariant probability measure ν such that

$$-\overline{H} = \int L + Pvd\mu > \int L + Pvd\nu = -\lambda.$$

We may assume that ν is ergodic, otherwise choose an ergodic component of ν for which the previous inequality holds. Take a generic point (x,v) in the support of ν and consider the projection $x(s)$ of its orbit. Then

$$u(x(0)) - \overline{H}(P)t \leq \int_0^t L(x(s), \dot{x}(s)) + P \cdot \dot{x}(s)ds + u(x(t)).$$

As $t \to \infty$

$$\frac{1}{t}\int_0^t L(x(s), \dot{x}(s)) + P \cdot \dot{x}(s)ds \to -\lambda,$$

by the ergodic theorem. Hence

$$-\overline{H} \leq -\lambda,$$

which is a contradiction. ■

Next we prove that any Mather measures (as defined originally by Mather) is "embedded" in a viscosity solution of a Hamilton-Jacobi equation. To do so we quote a theorem from [Mn96].

Theorem 35. *Suppose $\mu(P)$ is a ergodic minimizing measure. Then there exists a Lipschitz function $W : \operatorname{supp}(\mu) \to \mathbb{R}$ and a constant $\overline{H}(P) > 0$ such that*

$$-L - Pv = \overline{H}(P) + D_x W v + D_p W D_x H.$$

By taking Was initial condition (interpreting W as a function of x alone instead of (x, p) - which is possible because $\operatorname{supp} \mu$ is a Lipschitz graph) we can embed this minimizing measure in a viscosity solution. More precisely we have:

Theorem 36. *Suppose $\mu(P)$ is a ergodic minimizing measure. Then there exists a viscosity solution u of the cell problem*

$$H(P + D_x u, x) = \overline{H}(P)$$

such that $u = W$ on $\operatorname{supp}(\mu)$. Furthermore, for almost every $x \in \operatorname{supp}(\mu)$ the measures ν_t obtained by taking minimizing trajectories that pass trough x coincides with μ.

Proof. Consider the terminal value problem $V(x, 0) = W(x)$ if $x \in \operatorname{supp}(\mu)$ and $V(x, 0) = +\infty$ elsewhere, with

$$-D_t V + H(P + D_x V, x) = \overline{H}(P).$$

Then, for $x \in \operatorname{supp}(\mu)$ and $t > 0$

$$V(x, -t) = W(x).$$

Also if $x \notin \operatorname{supp}(\mu)$ then

$$V(x, -t) \leq V(x, -s),$$

if $s < t$. Hence, as $t \to \infty$ the function $V(x, -t)$ decreases pointwise. Since V is bounded and uniformly Lipschitz in x it must converge uniformly (because V is periodic) to some function u. Then u will be a viscosity solution of

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

Since $u = W$ on the support of μ, the second part of the theorem is a consequence of the ergodic theorem. ∎

In the case in which, for the same P, there are several different measures ν_t, we could use several functions $W_1 \ldots W_k$ as initial condition to construct a viscosity solution of $H(P + D_x u, x) = \overline{H}(P)$. By adding constants to W_i we may change this solution. This explains the non-uniqueness.

6.5 Regularity Estimates

Suppose u is a classical solution of

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

If $\det D^2_{xP} u \neq 0$,

$$X = x + D_P u \qquad P = p + D_x u$$

defines, at least locally, a smooth change of coordinates. If u is a viscosity solution, the meaning of $D_P u$ is not at all clear, even less the one of $\det D^2_{xP} u$. Nevertheless, in the previous chapter, many classical facts were recovered by using viscosity techniques. Therefore one would expect these expressions to be defined in some weak sense. In this section we will try to get regularity estimates for u and obtain more precise information about the dynamics. We start by studying the continuity of u in P in the invariant set. This estimates will improve the partial uniqueness result from the previous chapter. Then we will develop the L^2 and L^∞ regularity theory for higher derivatives (for a more complete discussion, consult [EG99a])

Fix P and let $(x, p) \in \mathcal{I}$. Assume that $x(\cdot)$ is a minimizing trajectory with initial condition $x(0) = x$. By adding a suitable constant to $u(\cdot, P')$ we may assume $u(x(0), P) = u(x(0), P')$.

Proposition 31. *Suppose ν is a Mather measure. Let $P_n \to P$. Then there exists a point x in the support of $\pi^*\nu$ (the projection of ν in the x coordinate) such that for any fixed T*

$$\sup_{0 \le t \le T} |u(x^*(t), P) - u(x^*(t), P_n)| \to 0,$$

as $n \to \infty$, provided $u(x, P_n) = u(x, P)$.

Proof. First we prove an auxiliary lemma

Lemma 23. *There exist sequences on the support of $\pi^*\nu$, $x_n \to x$, $\tilde{x}_n \to x$, $p_n \to p$ and $\tilde{p}_n \to p$, where (x, p) and (x_n, p_n) are optimal pairs for P and $(\tilde{x}_n, \tilde{p}_n)$ are optimal pairs for P_n.*

Remark. the nontrivial point of the lemma is that the limits of p_n and $\tilde{p}_n$ are the same.

Proof. Take a generic point (x_0, p_0) in the support of ν. Let $x^*(t)$ be the optimal trajectory for P with initial condition (x_0, p_0). Then for all $t > 0$

$$H(P + D_x u(x^*(t), P), x^*(t)) = \overline{H}(P).$$

Also, for almost every y,

$$H(P_n + D_x u(x^*(t) + y, P_n), x^*(t)) = \overline{H}(P_n) + O(|y|),$$

for almost every t. Choose y_n with $|y_n| \leq |P - P_n|$ such that the previous identity holds. By strict convexity of H in p, we get

$$\dot{x}^*(t)\xi + \theta\xi^2 \leq C|P_n - P|,$$

where

$$\xi = [P - P_n + D_x u(x^*(t), P) - D_x u(x^*(t) + y_n, P_n)],$$

and

$$\dot{x}^*(t) = -D_p H(P + D_x u(x^*(t), P), x^*(t)).$$

Note that

$$\left|\frac{1}{T}\int_0^T \dot{x}^*(t)\xi\right| | \leq |P - P_n| + \frac{|u(x^*(0), P) - u(x^*(T), P)|}{T} + \\ + \frac{|u(x^*(0) + y_n, P_n) - u(x^*(T) + y_n, P_n)|}{T}.$$

Therefore we may choose T so that

$$\left|\frac{1}{T}\int_0^T \dot{x}^*(t)\xi\right| \leq 2|P - P_n|.$$

Thus

$$\frac{1}{T}\int_0^T |P + D_x u(x^*(t), P) - P_n - D_x u(x^*(t) + y_n, P_n)|^2 \leq C|P - P_n|.$$

Hence we may choose $0 \leq t_n \leq T$ for which

$$|P + D_x u(x^*(t_n), P) - P_n - D_x u(x^*(t_n) + y_n, P_n)|^2 \leq C|P - P_n|.$$

Let $x_n = x^*(t_n)$, $\tilde{x}_n = x^*(t_n) + y_n$, and

$$p_n = P + D_x u(x^*(t_n), P) \qquad \tilde{p}_n = P_n + D_x u(x^*(t_n) + y_n, P).$$

By extracting a subsequence, if necessary we may assume $x_n \to x$, $\tilde{x}_n \to x$, etc. ■

To see that the lemma implies the proposition, let $x_n^*(t)$ be the optimal trajectory for P with initial conditions (x_n, p_n). Similarly let $\tilde{x}_n^*(t)$ be the optimal trajectory for P_n with initial conditions $(\tilde{x}_n, \tilde{p}_n)$. Then

$$u(x_n, P) = \int_0^t L(x_n^*, \dot{x}_n^*) + P \cdot \dot{x}_n^* + \overline{H}(P) ds + u(x_n^*(t), P),$$

and

$$u(\tilde{x}_n, P_n) = \int_0^t L(\tilde{x}_n^*, \dot{\tilde{x}}_n^*) + P_n \cdot \dot{\tilde{x}}_n^* + \overline{H}(P_n) ds + u(\tilde{x}_n^*(t), P_n).$$

On $0 \leq t \leq T$ both x_n^* and $\tilde{x}_n^*$ converge uniformly and, since by hypothesis

$$u(x_n, P), u(\tilde{x}_n, P_n) \to u(x, P),$$

we conclude that $u(\tilde{x}_n^*(t), P_n) - u(x_n^*(t), P) \to 0$ uniformly on $0 \leq t \leq T$. ■

Theorem 37. *Suppose ν is an ergodic Mather measure with $\nu|_{\mathrm{supp}(\nu)}$ uniquely ergodic with respect to the restricted flow. Assume $P_n \to P$. Then*

$$u(x, P_n) \to u(x, P),$$

uniformly on the support of $\pi^*\nu$, provided that an appropriate constant $C(P_n)$ is added to $u(x, P_n)$.

Proof. Fix $\epsilon > 0$. We need to show that if n is sufficiently large then

$$\sup_{x \in \mathrm{supp}(\nu)} |u(x, P_n) - u(x, P)| < \epsilon.$$

Choose M such that $\|D_x u(x, P)\|, \|D_x u(x, P_n)\| \leq M$. Let $\delta = \frac{\epsilon}{8M}$. Cover $\mathrm{supp}\, \nu$ with finitely many balls B_i with radius $\leq \delta$. Choose x as in the previous proposition. Let $x^*(t)$ be the optimal trajectory for P with initial condition x. Then there exists T_δ and $0 \leq t_i \leq T_\delta$ such that $x_i = x^*(t_i) \in B_i$. Choose n sufficiently large such that

$$\sup_{0 \leq t \leq T_\delta} |u(x^*(t), P) - u(x^*(t), P_n)| \leq \frac{\epsilon}{2}.$$

Then, on each y in B_i

$$\begin{aligned} |u(y, P) - u(y, P_n)| \leq |u(y, P) - u(y_i, P)| + |u(y_i, P) - u(y_i, P_n)| + \\ + |u(y_i, P_n) - u(y, P_n)| \leq 4M\delta + \frac{\epsilon}{2} \leq \epsilon. \end{aligned}$$

■

In [Mn96] Mañé proved that up arbitrarily small generic perturbations of the Lagrangian L all Mather measures are uniquely ergodic, therefore unique ergodicity seems to be a reasonable hypothesis. Furthermore the non-uniqueness implies that some ergodic-type hypothesis is necessary.

Corollary 8. *Let ν be as in the previous theorem. Suppose that $\overline{H}(P)$ is differentiable at P. Then there exists an increasing function function $\omega(\delta)$ with $\omega(0) = 0$ such that*

$$|x(t) - x(0) - D_P\overline{H}t| \leq \min_{\delta} \frac{\|x(t) - x(0)\| \wedge \omega(\delta)}{\delta} + Ct_1\delta,$$

for all (generic) trajectories of the Hamilton equations with initial conditions in the support of ν.

PROOF. It suffices to apply theorem 28 with

$$\omega(\delta) = \sup_{|P-P'|\leq\delta} \sup_{x\in\operatorname{supp}\nu} |u(x, P) - u(x, P')|.$$

■

Finally, we compute an explicit formula for $D_P u$ along trajectories (provided this derivative exists).

Proposition 32. *Suppose $\overline{H}(P)$ is differentiable in P and $u(x(0), P) = u(x(0), P')$. If $D_P u$ exists along the trajectory $x(\cdot)$ then*

$$D_P u = -\int_0^t (\dot{x} - D_P\overline{H}(P))ds.$$

PROOF. Note that

$$-\int_0^t (\dot{x} - D_P\overline{H})ds \in D_P^- u(x(t), P)$$

and when the derivative exists it must coincide with a unique point in $D_P^- u$. ■

A formal calculation yields

$$D_{Px}^2 u = \int_0^t \frac{\partial \dot{x}}{\partial x} ds$$

In the invariant set, $p = P + D_x u$ is a Lipschitz function. Thus we may be able to compute the derivative $D_{Px}^2 u$ since $\frac{\partial \dot{x}}{\partial x}$ "should" be defined almost everywhere.

This observation motivates the study of the regularity of the derivatives of u in the invariant set. Our approach is set via $L^2(\nu_t)$-type regularity estimates.

Theorem 38. *Suppose $\overline{H}$ is twice differentiable at P. Then (with $\sigma = \pi^*\nu$)*

$$\int |D_x u(x, P') - D_x u(x, P)|^2 d\sigma \leq C|P - P'|^2,$$

Proof. Let $u(x,P)$ be a family of solutions of

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

If by $D_x u(x,P')$ we denote a version of $D_x u(x,P')$ then

$$H(P' + D_x u(x,P'), x) \leq \overline{H}(P').$$

By subtraction and strict convexity of H we get

$$D_p H(P + D_x u(x,P), x)(P' + D_x u(x,P') - P - D_x u(x,P)) + \\ + \theta |P - P' + D_x u(x,P) - D_x u(x,P')|^2 \leq \overline{H}(P') - \overline{H}(P).$$

Then

$$\int D_p H(P + D_x u(x,P), x) \left(D_x u(x,P') - D_x u(x,P) \right) d\sigma = 0,$$

by theorem 32 . Also

$$(P' - P) \int D_p H(P + D_x u(x,P), x) d\sigma = D_P \overline{H}(P)(P' - P).$$

Thus

$$\theta \int |P - P' + D_x u(x,P) - D_x u(x,P')|^2 d\sigma = \overline{H}(P') - \overline{H}(P) - D_P \overline{H}(P)(P' - P).$$

If $\overline{H}(P)$ is twice differentiable at P then

$$\int |P - P' + D_x u(x,P) - D_x u(x,P')|^2 d\sigma \leq C|P - P'|^2.$$

Applying to this estimate a weighted Cauchy inequality we end the proof. ■

Now we turn our attention to $L^2(\sigma)$-type regularity estimates for differential quotients in the x variable.

Theorem 39. *Suppose u is a periodic viscosity solution of*

$$H(P + D_x u, x) = \overline{H}(P),$$

and ν and $\sigma = \pi^* \nu$ measures as in the previous theorem. Then

$$\int |D_x u(x + y, P) - D_x u(x,P)| d\sigma \leq C|y|^2.$$

PROOF. Without loss of generality, assume $P = 0$ and $\overline{H}(P) = 0$. Then, for any version of $D_x u(x+y)$

$$H(D_x u(x+y), x+y) \leq 0.$$

Subtracting $H(D_x u(x), x)$ yields

$$H(D_x u(x+y), x+y) - H(D_x u(x), x) \leq 0.$$

Thus

$$\theta |D_x u(x+y) - D_x u(x)|^2 + D_p H(D_x u(x), x)(D_x u(x+y) - D_x u(x)) \leq \\ H(D_x u(x+y), x) - H(D_x u(x+y), x+y). \quad (6.2)$$

Note that

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

Integrating (6.2) with respect to σ we conclude

$$\theta \int |D_x u(x+y) - D_x u(x)|^2 d\sigma \leq C|y|^2 - \int D_x H(D_x u(x+y), x) y d\sigma.$$

Since

$$\int D_x H(D_x u(x+y), x) y d\sigma = \int \left[D_x H(D_x u(x+y), x) - D_x H(D_x u(x), x) \right] y d\sigma,$$

we get

$$\frac{\theta}{2} \int |D_x u(x+y) - D_x u(x)|^2 d\sigma \leq C|y|^2,$$

using a standard weighted Cauchy inequality. ■

In the setting of KAM theory one considers small perturbations of integrable Hamiltonians

$$H(p, x) = H_0(p) + \lambda H_1(p, x), \quad (6.3)$$

with H_1 bounded with bounded first and second derivatives, the previous theorem yields

Corollary 9. *Suppose H is of the form (6.3). Then, for λ sufficiently small,*

$$\int |D_x u(x+y, P) - D_x u(x, P)|^2 d\sigma \leq C\lambda |y|^2.$$

PROOF. Repeating the calculations in the proof of the previous theorem with a Hamiltonian like (6.3) proves this bound. ■

6.6 Young Measures and Adiabatic Invariants

In this section we survey extensions of the previous results for time-dependent Hamiltonians. Since the proofs are similar we will omit them and refer the reader to [EG99b].

Let $V_\epsilon(x,t)$ be a periodic (both in x and t) solution of the Hamilton-Jacobi equation

$$-D_t V_\epsilon + H(P + D_x V_\epsilon, \frac{x}{\epsilon}, t) = \overline{H}_\epsilon(P).$$

Assume $x^\epsilon(\cdot)$ is a minimizing trajectory. Then, for any s and t

$$V_\epsilon(x^\epsilon(s), s) = \int_s^t \left[L(x^\epsilon(r), \dot{x}^\epsilon(r), r) - P \cdot \dot{x}^\epsilon(r) - \overline{H}_\epsilon(P) \right] dr + V_\epsilon(x^\epsilon(t), t).$$

We proceed now, as in section 6.2, to construct a family of Young measures associated with these trajectories.

Theorem 40. *For almost every $0 \leq t \leq 1$ there exists a measure ν_t such that for any smooth periodic (both in y, τ and t) function $\Phi(p, y, \tau, t)$*

$$\Phi(p^\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}, t) \rightharpoonup \overline{\Phi}(t),$$

with $\overline{\Phi}(t) = \int \Phi(p, y, \tau, t) d\nu_t(p, y, \tau)$. More precisely, for any smooth function $\varphi(t)$

$$\int_0^1 \varphi(t) \Phi(p^\epsilon, \frac{x^\epsilon}{\epsilon}, \frac{t}{\epsilon}, t) dt \to \int_0^1 \varphi(t) \overline{\Phi}(t) dt,$$

as $\epsilon \to 0$.

Proof. The proof is the same as in theorem 30. ∎

Now we will show that each of this measures is invariant for the flow induced by the Hamiltonian $H(p, y, \tau, T)$, for fixed $0 < T < 1$.

Theorem 41. *For any smooth periodic (in y and τ) function $\phi(p, y, \tau)$*

$$\int D_y \phi(p, y, \tau) D_p H(p, y, \tau, t) - D_p \phi(p, y, \tau) D_x H(p, y, \tau, t) + D_\tau \phi(p, y, \tau) d\nu_t = 0.$$

Proof. The proof is exactly the same as in theorem 31. ∎

Finally we have the analog of theorem 32:

Theorem 42. *If $\phi : \mathbb{R}^n \to \mathbb{R}$ is a Lipschitz function then, for almost every t, there exists a version of $D_x \phi$ such that*

$$\int D_x \phi(y) D_p H(p, y, \tau, t) d\nu_t = 0.$$

Proof. The proof is exactly the same as in theorem 32. ∎

6.7 Viscosity Solutions - II

Take V_ϵ, as before, to be a $[0,\epsilon]^n \times [0,1]$ periodic viscosity solution of the Hamilton-Jacobi equation

$$-D_t V_\epsilon + H(P + D_x V_\epsilon, \frac{x}{\epsilon}, t) = \overline{H}_\epsilon(P).$$

Let x^ϵ be an optimal trajectory.

Theorem 43. *There exists a function $X(\cdot)$ such that, as $\epsilon \to 0$, $x^\epsilon(t) \to X(t)$ uniformly in $t \in [0,1]$; thus $\dot{x}^\epsilon \rightharpoonup \dot{X}$. Also $\overline{H}_\epsilon(P) \to \overline{H}(P)$ and $V_\epsilon \to V$, where V is a periodic viscosity solution of the equation*

$$-D_t V + \overline{H}(P + D_x V, t) = \overline{H}(P).$$

In the previous equation, $\overline{H}(P,t)$ denotes, for each t, the unique value for which the equation

$$H(P + D_y u, y, t) = \overline{H}(P,t)$$

has a periodic viscosity solution $u(y)$. The function, V does not depend on x and so it solves the (ordinary) differential equation

$$-D_t V + \overline{H}(P,t) = \overline{H}(P),$$

in the viscosity sense. In particular this implies

$$\overline{H}(P) = \int_0^1 \overline{H}(P,t)dt.$$

Furthermore $V_\epsilon(x_\epsilon(t),t) \to V(X(t),t) = V(t)$, uniformly in t, and

$$D_t V_\epsilon + D_x V_\epsilon \cdot \dot{x}^\epsilon \rightharpoonup D_t V + D_x V \cdot \dot{X} = D_t V.$$

Additionally

$$L\left(\dot{x}^\epsilon, \frac{x^\epsilon}{\epsilon}, t\right) \rightharpoonup -P \cdot \dot{X} - \overline{H}(P,t).$$

Finally,

$$-P \cdot \dot{X} - \overline{H}(P,t) = \overline{L}(\dot{X},t),$$

where

$$\int L d\nu_t = \overline{L}(\dot{X},t),$$

and $\overline{L} = \overline{H}^*$.

Proof. Since x^ϵ is uniformly bounded and equicontinuous there exists a function $X(\cdot)$ such that

$$x^\epsilon(t) \to X(t),$$

uniformly as $\epsilon \to 0$ (possibly after extracting a subsequence $\epsilon_k \to 0$). And consequently

$$\dot{x}^\epsilon(t) \rightharpoonup \dot{X}(t).$$

Since V_ϵ is bounded (possibly after subtracting a constant c_ϵ) and equicontinuous we may find a function V and a subsequence $\epsilon \to 0$ such that

$$V_\epsilon(x^\epsilon(t), t) \to V,$$

uniformly as $t \to 0$, and by passing to a further subsequence, if necessary, we may assume $\overline{H}_\epsilon(P) \to \overline{H}(P)$. By the results from chapter 4, we know that V must be a viscosity solution of

$$-D_t V + \overline{H}(P + D_x V, t) = \overline{H}(P).$$

Since V_ϵ has period ϵ the limit V does not depend on x. Thus

$$D_t V_\epsilon(x^\epsilon(t), t) + D_x V_\epsilon(x^\epsilon(t), t) \cdot \dot{x}^\epsilon(t) \rightharpoonup D_t V = \overline{H}(P, t) - \overline{H}(P).$$

Along minimizing trajectories

$$D_x V_\epsilon(x^\epsilon(t), t) \cdot \dot{x}^\epsilon = -H\left(P + p^\epsilon, \frac{x^\epsilon}{\epsilon}, t\right) - L\left(\dot{x}^\epsilon, \frac{x^\epsilon}{\epsilon}, t\right) - P \cdot \dot{x}^\epsilon,$$

and

$$D_t V_\epsilon = H\left(P + p^\epsilon, \frac{x^\epsilon}{\epsilon}, t\right) - \overline{H}_\epsilon(P).$$

Therefore we conclude that

$$-\overline{H}(P, t) - L\left(\dot{x}^\epsilon, \frac{x^\epsilon}{\epsilon}, t\right) - P \cdot \dot{x}^\epsilon \rightharpoonup 0,$$

or equivalently

$$L\left(\dot{x}^\epsilon, \frac{x^\epsilon}{\epsilon}, t\right) \rightharpoonup -\overline{H}(P, t) - P \cdot \dot{X}.$$

Thus

$$\overline{L} \equiv \int L d\nu_t = -\overline{H}(P, t) - P \cdot \dot{X}.$$

Finally we claim that $\overline{L} = \overline{H}^*$. To see this, it suffices to check that for any $\tilde{P}$

$$\overline{L}(\dot{X}) \geq -\overline{H}(\tilde{P}) - \tilde{P} \cdot \dot{X}.$$

Take $\tilde{V}_\epsilon$ to be a periodic viscosity solution of

$$-D_t\tilde{V}_\epsilon + H\left(\tilde{P} + D_x\tilde{V}_\epsilon, \frac{x}{\epsilon}, t\right) = \overline{H}_\epsilon(\tilde{P}).$$

Then, for almost every $y \in \mathbb{R}^n$,

$$D_t\tilde{V}_\epsilon(\epsilon^2 y + x^\epsilon(t), t) + D_x\tilde{V}_\epsilon(\epsilon^2 y + x^\epsilon(t), t) \cdot \dot{x}^\epsilon(t) \rightharpoonup D_t\tilde{V},$$

where x^ϵ is the optimal trajectory for V_ϵ and $\tilde{V} = \lim_{\epsilon\to 0}\tilde{V}_\epsilon$. Now note that

$$D_x\tilde{V}_\epsilon(\epsilon^2 y + x^\epsilon(t), t) \cdot \dot{x}^\epsilon \geq$$
$$\geq -H\left(\tilde{P} + D_x\tilde{V}_\epsilon, \epsilon y + \frac{x^\epsilon}{\epsilon}, t\right) - L\left(\dot{x}^\epsilon, \epsilon y + \frac{x^\epsilon}{\epsilon}, t\right) - \tilde{P}\cdot\dot{x}^\epsilon,$$

and

$$D_t\tilde{V}_\epsilon = H\left(\tilde{P} + D_x\tilde{V}_\epsilon, \epsilon y + \frac{x^\epsilon}{\epsilon}, t\right) - \overline{H}_\epsilon(\tilde{P}).$$

Thus

$$\overline{L} \geq -\overline{H}(\tilde{P}, t) - \tilde{P}\cdot\dot{X},$$

by passing to the limit. ■

Using viscosity solutions methods we prove that the support of these measures is contained in the closure of the graph of $P + D_x u$, where u is a solution of $H(P + D_x u, x, t) = \overline{H}(P, t)$ and $D_x u$ is an appropriate version of the derivative of u.

Theorem 44. *Let u be any viscosity solution of $H(P + D_x u, x, t) = \overline{H}(P, t)$. Then, for almost every t, $p = P + D_x u$, ν_t a.e., where $D_x u$ is a version of $D_x u$.*

Proof. By strict convexity of H, there exists γ such that

$$\gamma|p - q|^2 + H(p, x, t) + D_p H(p, x, t)\cdot(q - p) \leq H(q, x, t).$$

Choose q to be a version of $P + D_x u$ (an will denote, as usual $q = P + D_x u$). Then

$$\gamma\int |p - (P + D_x u)|^2 d\nu_t \leq \int H(P + D_x u, x, t)d\nu_t - \int H(p, x, t)d\nu_t -$$
$$- \int D_p H(p, x, t)(P + D_x u - p)d\nu_t.$$

but the right-hand side integrates to zero because

$$H(P + D_x u, x, t) \leq \overline{H}(P, t)$$

everywhere (since we are using a version of $D_x u$ and H is convex),

$$\int pD_pH - Hd\nu_t = \overline{L} \qquad P\int D_pHd\nu_t = -P\cdot\dot{X},$$

thus $\overline{H}+\overline{L}+P\cdot\dot{X}=0$, and

$$\int D_pHD_xud\nu_t = 0.$$

■

6.8 Convex duality

In the papers [FV88], [FV89] and [Fle89], the Fenchel duality theorem [Roc66] is used to analyze optimal control problems. This motivated us to investigate the duality relation between Aubry-Mather theory and viscosity solutions of Hamilton-Jacobi equations. In particular, we found that the problem of finding an Aubry-Mather measure is the dual of computing $\overline{H}$.

Let $\Omega = T^n\times\mathbb{R}^n$, where T^n is the n dimensional torus, identified, when convenient, with $[0,1]^n$. A pair $(x,v)=z$ represents a generic point $z\in\Omega$, with $x\in T^n$ and $v\in\mathbb{R}^n$. Choose a function $\gamma\equiv\gamma(|v|):\Omega\to[1,+\infty)$ satisfying

$$\lim_{|v|\to+\infty}\frac{L(v,x)}{\gamma(v)}=+\infty \qquad \lim_{|v|\to+\infty}\frac{|v|}{\gamma(v)}=0.$$

Let $\mathcal{M}$ be a set of weighted Radon measures on Ω, i.e.,

$$\mathcal{M}=\{\mu \text{ signed measure on } \Omega \text{ with } \int\gamma d|\mu|<\infty\}.$$

Note that $\mathcal{M}$ is the dual of $C_0^\gamma(\Omega)$, i.e., the set of continuous functions ϕ with

$$\|\phi\|_\gamma=\sup_\Omega|\frac{\phi}{\gamma}|<\infty, \qquad \lim_{|z|\to\infty}\frac{\phi(z)}{\gamma(v)}\to 0.$$

Define

$$\mathcal{M}_1=\{\mu\in\mathcal{M}:\int d\mu=1,\mu\geq 0\}$$

and

$$\mathcal{M}_2=\mathrm{cl}\{\mu\in\mathcal{M}:\int vD_x\varphi d\mu=0,\ \forall\varphi(x)\in C^1(T^n)\}.$$

The set $\mathcal{M}_2$ is the "measure theoretic" analog of the set of closed curves on T^n. Indeed, if $\theta:[0,1]\to T^n$ is a piecewise smooth closed curve, define a measure μ_θ by

$$\int_\Omega fd\mu_\theta=\int_0^1 f(\theta(t),\dot{\theta}(t))dt.$$

Clearly μ_θ belongs to $\mathcal{M}_2$, and since $\mathcal{M}_2$ is a linear space, it contains all linear combinations of measures of this form.

For $\phi \in C_0^\gamma(\Omega)$ define

$$h_1(\phi) = \sup_{(x,v)\in\Omega} (-\phi(x,v) - L(x,v)).$$

Let $\mathcal{C}$ be defined by

$$\mathcal{C} = \mathrm{cl}\{\phi : \phi = vD_x\varphi, \varphi \in C^1(T^n)\},$$

here cl denotes the closure in C_0^γ. We can think of the elements in $\mathcal{C}$ as generalized closed differential forms; indeed if $\theta : [0,1] \to T^n$ is a piecewise smooth closed curve and $\phi \in \mathcal{C}$ then

$$\int \phi d\mu_\theta = 0.$$

Define

$$h_2(\phi) = \begin{cases} 0 & \text{if } \phi \in \mathcal{C} \\ -\infty & \text{otherwise.} \end{cases}$$

We will show that the problem of computing

$$\sup_{\phi \in C_0^\gamma(\Omega)} h_2(\phi) - h_1(\phi) \tag{6.4}$$

is the dual problem of computing Aubry-Mather measures.

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

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

for $y \in E'$. Similarly, for concave, upper semicontinuous functions $g : E \to (-\infty, +\infty]$ let

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

The Rockafellar-Fenchel duality theorem states that

$$\sup_x g(x) - f(x) = \inf_y f^*(y) - g^*(y), \tag{6.5}$$

more precisely,

Theorem 45 (Rockafellar). *Let E be a locally convex Hausdorff topological vector space over $\mathbb{R}$ with dual E^*. Suppose $h : E \to (-\infty, +\infty]$ is convex and lower semicontinuous, $g : E \to [-\infty, +\infty)$ is concave and upper semicontinuous. Then (6.5) holds, provided that either h or g is continuous at some point where both functions are finite.*

PROOF. See [Roc66]. ■

We now compute the Legendre-Fenchel transforms of h_1 and h_2 in order to apply this theorem to problem (6.4).

Proposition 33. *We have*

$$h_1^*(\mu) = \begin{cases} \int L d\mu & \text{if} \quad \mu \in \mathcal{M}_1 \\ +\infty & \text{otherwise,} \end{cases}$$

and

$$h_2^*(\mu) = \begin{cases} 0 & \text{if} \quad \mu \in \mathcal{M}_2 \\ -\infty & \text{otherwise.} \end{cases}$$

PROOF. Recall that

$$h_1^*(\mu) = \sup_{\phi \in C_0^\gamma(\Omega)} \left(-\int \phi d\mu - h_1(\phi) \right).$$

First we will show that if μ is non-positive then $h_1^*(\mu) = \infty$.

Lemma 24. *If $\mu \ngeq 0$ then $h_1^*(\mu) = +\infty$.*

PROOF. If $\mu \ngeq 0$ we can choose a sequence of positive functions $\phi_n \in C_0^\gamma(\Omega)$ such that

$$\int -\phi_n d\mu \to +\infty.$$

Thus, since

$$\sup -\phi_n - L \leq 0,$$

we have $h_1^*(\mu) = +\infty$. ■

Lemma 25. *If $\mu \geq 0$ then*

$$h_1^*(\mu) \geq \int L d\mu + \sup_{\psi \in C_0^\gamma(\Omega)} \left(\int \psi d\mu - \sup \psi \right).$$

Proof. Let L_n be a sequence of functions in $C_0^\gamma(\Omega)$ increasing pointwise to L. Any function ϕ in $C_0^\gamma(\Omega)$ can be written as $\phi = -L_n - \psi$, for some ψ also in $C_0^\gamma(\Omega)$. Thus

$$\sup_{\phi \in C_0^\gamma(\Omega)} \left(-\int \phi d\mu - h_1(\phi) \right) = \sup_{\psi \in C_0^\gamma(\Omega)} \left(\int L_n d\mu + \int \psi d\mu - \sup(L_n + \psi - L) \right).$$

Since

$$\sup L_n - L \leq 0,$$

we have

$$\sup(L_n + \psi - L) \leq \sup \psi.$$

Thus

$$\sup_{\phi \in C_0^\gamma(\Omega)} \left(-\int \phi d\mu - h_1(\phi) \right) \geq \sup_{\psi \in C_0^\gamma(\Omega)} \left(\int L_n d\mu + \int \psi d\mu - \sup(\psi) \right).$$

By the monotone convergence theorem $\int L_n d\mu \to \int L d\mu$. Therefore

$$\sup_{\phi \in C_0^\gamma(\Omega)} \left(-\int \phi d\mu - h_1(\phi) \right) \geq \int L d\mu + \sup_{\psi \in C_0^\gamma(\Omega)} \left(\int \psi d\mu - \sup(\psi) \right),$$

as required. ∎

If $\int L d\mu = +\infty$ then $h_1^*(\mu) = +\infty$. 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 taking $\psi = \alpha$, constant. Therefore $h_1^*(\mu) = +\infty$.

If $\int d\mu = 1$ the previous lemma implies

$$h_1^*(\mu) \geq \int L d\mu,$$

by taking $\psi = 0$.

Also, for any ϕ

$$\int (-\phi - L) d\mu \leq \sup(-\phi - L),$$

if $\int d\mu = 1$. Hence

$$\sup_{\phi \in C_0^\gamma(\Omega)} \left(-\int \phi d\mu - h_1(\phi) \right) \leq \int L d\mu.$$

Thus

$$h_1^*(\mu) = \begin{cases} \int L d\mu & \text{if} \quad \mu \in \mathcal{M}_1 \\ +\infty & \text{otherwise.} \end{cases}$$

If $\mu \notin \mathcal{M}_2$ then there exists $\hat{\phi} \in \mathcal{C}$ such that

$$\int \hat{\phi} d\mu \neq 0.$$

Therefore

$$\inf_{\phi \in \mathcal{C}} - \int \phi d\mu \leq \inf_{\alpha \in \mathbb{R}} \alpha \int \hat{\phi} d\mu = -\infty.$$

If $\mu \in \mathcal{M}_2$ then $\int \phi d\mu = 0$, for all $\phi \in \mathcal{C}$. Therefore

$$h_2^*(\mu) = \inf_{\phi \in \mathcal{C}} \int \phi d\mu = \begin{cases} 0 & \text{if } \mu \in \mathcal{M}_2 \\ -\infty & \text{otherwise.} \end{cases}$$

■

The Fenchel-Rockafellar duality theorem states that

$$\sup_{\phi \in C_0^\gamma(\Omega)} (h_2(\phi) - h_1(\phi)) = \inf_{\mu \in \mathcal{M}} (h_1^*(\mu) - h_2^*(\mu)).$$

provided on the set $h_2 > -\infty$, h_1 is continuous, which is true because h_1 is continuous, as shown in the next lemma.

Lemma 26. *h_1 is continuous.*

Proof. Suppose $\phi_n \to \phi$ in C_0^γ. Then $\|\phi_n\|_\gamma$ and $\|\phi\|_\gamma$ are bounded uniformly by some constant C. The growth condition on L implies that there exists $R > 0$ such that

$$\sup_{\Omega} -\hat{\phi} - L = \sup_{T^n \times B_R} -\hat{\phi} - L,$$

for all $\hat{\phi}$ in $C_0^\gamma(\Omega)$ with $\|\hat{\phi}\|_\gamma < C$. On B_R, $\phi_n \to \phi$ uniformly and so

$$\sup_{\Omega} -\phi_n - L \to \sup_{\Omega} -\phi - L.$$

■

Denote by $H^\star$ the value

$$H^\star = - \sup_{\phi \in C_0^\gamma(\Omega)} (h_2(\phi) - h_1(\phi))$$

Theorem 46. *We have*

$$H^\star = \inf\{\lambda : \exists \varphi \in C^1(T^n) : H(D_x\varphi, x) < \lambda\}.$$

Proof. Note that

$$H^\star = \inf_{\varphi\in C^1(T^n)} \sup_{(x,v)\in\Omega} -vD_x\varphi - L = \inf_{\varphi\in C^1(T^n)} \sup_{x\in T^n} H(D_x\varphi, x).$$

■

Let Γ denote the set of piecewise smooth closed curves $\theta : [0,1] \to T^n$. Define

$$\mathcal{M}_3 = \operatorname{cl}\operatorname{span}\{\mu_\theta : \theta \in \Gamma\}.$$

The next task is to prove that the value $H^\star$, computed by considering a infimum over measures in $\mathcal{M}_2$, is the same as the value obtained using measures in $\mathcal{M}_3$.

$$\overline{H} = -\inf_{\mu\in\mathcal{M}_3}(h_1^*(\mu) - h_2^*(\mu)).$$

Recall that another characterization of $\overline{H}$ is that $\overline{H}$ is the unique value for which the equation

$$H(D_x u, x) = \overline{H}$$

has a periodic viscosity solution.

Theorem 47. *$H^\star$ is the unique value for which the equation*

$$H(D_x u, x) = H^\star$$

has a periodic viscosity solution.

Proof. First suppose u is a periodic viscosity solution of

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

We claim that there is no smooth function ψ with

$$H(D_x\psi, x) < \overline{H}.$$

Indeed, if this were false, we could choose a point x_0 at which $u - \psi$ has a local minimum. At this point we would have

$$H(D_x\psi, x_0) \geq \overline{H},$$

by the viscosity property, which is a contradiction. Hence $H^\star \geq \overline{H}$.

To prove the other inequality consider a standard mollifier η_ϵ and define $u_\epsilon = \eta_\epsilon \star u$. Then

$$H(D_x u_\epsilon, x) \leq \overline{H} + h(\epsilon, x),$$

where

$$h(\epsilon, x) = \sup_{|p| \leq R} \sup_{|x-y| \leq \epsilon} |H(p, x) - H(p, y)|,$$

where R is a bound on the Lipschitz constant of u. Let $H^\epsilon = \overline{H} + \sup_x h(\epsilon, x)$. Thus u_ϵ satisfies

$$H(D_x u_\epsilon, x) \leq H^\epsilon.$$

Thus $H^\star \leq \lim_{\epsilon \to 0} H^\epsilon = \overline{H}$. Hence $H^\star = \overline{H}$. ■

Corollary 10. *We have*

$$\inf_{\mu \in \mathcal{M}_2} (h_1^*(\mu) - h_2^*(\mu)) = \inf_{\mu \in \mathcal{M}_3} (h_1^*(\mu) - h_2^*(\mu)).$$

Proof. Our previous results show that we can construct a probability measure μ on $\mathcal{M}_3$ such that

$$\int L d\mu = \overline{H} = \inf_{\mu \in \mathcal{M}_2} (h_1^*(\mu) - h_2^*(\mu)).$$

Since $\mathcal{M}_3 \subset \mathcal{M}_2$ this completes the proof. ■

Bibliography

[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].

[Arn99] V. I. Arnold. *Mathematical methods of classical mechanics.* Springer-Verlag, New York, 199? 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] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations.* Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.

[BD98] Andrea Braides and Anneliese Defranceschi. *Homogenization of multiple integrals.* The Clarendon Press Oxford University Press, New York, 1998.

[CD87] Weiren Chou and R. J. Duffin. An additive eigenvalue problem of physics related to linear programming. *Adv. in Appl. Math.*, 8(4):486–498, 1987.

[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.

[Dur96] Richard Durrett. *Probability: theory and examples.* Duxbury Press, Belmont, CA, second edition, 1996.

[E99] Weinan E. Aubry-Mather theory and periodic solutions of the forced Burgers equation. *Comm. Pure Appl. Math.*, 52(7):811–828, 1999.

[EG99a] L. C. Evans and D. Gomes. Effective hamiltonians and averaging for hamiltonian dynamics i. *Preprint*, 1999.

[EG99b] L. C. Evans and D. Gomes. Effective hamiltonians and averaging for hamiltonian dynamics ii. *Preprint*, 1999.

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

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

[Fat97a] Albert Fathi. Solutions KAM faibles conjuguées et barrières de Peierls. *C. R. Acad. Sci. Paris Sér. I Math.*, 325(6):649–652, 1997.

[Fat97b] Albert Fathi. Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens. *C. R. Acad. Sci. Paris Sér. I Math.*, 324(9):1043–1046, 1997.

[Fat98a] Albert Fathi. Orbite hétéroclines et ensemble de Peierls. *C. R. Acad. Sci. Paris Sér. I Math.*, 326:1213–1216, 1998.

[Fat98b] Albert Fathi. Sur la convergence du semi-groupe de Lax-Oleinik. *C. R. Acad. Sci. Paris Sér. I Math.*, 327:267–270, 1998.

[Fle89] Wendell H. Fleming. Generalized solutions and convex duality in optimal control. In *Partial differential equations and the calculus of variations, Vol. I*, pages 461–471. Birkhäuser Boston, Boston, MA, 1989.

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

[FV88] Wendell H. Fleming and Domokos Vermes. Generalized solutions in the optimal control of diffusions. In *Stochastic differential systems, stochastic control theory and applications (Minneapolis, Minn., 1986)*, pages 119–127. Springer, New York, 1988.

[FV89] Wendell H. Fleming and Domokos Vermes. Convex duality approach to the optimal control of diffusions. *SIAM J. Control Optim.*, 27(5):1136–1155, 1989.

[Gal83] Giovanni Gallavotti. *The elements of mechanics.* Springer-Verlag, New York, 1983. Translated from the Italian.

[Gol80] Herbert Goldstein. *Classical mechanics.* Addison-Wesley Publishing Co., Reading, Mass., second edition, 1980. Addison-Wesley Series in Physics.

[JKM99] H. R. Jauslin, H. O. Kreiss, and J. Moser. On the forced Burgers equation with periodic boundary conditions. In *Differential equations: La Pietra 1996 (Florence)*, pages 133–153. Amer. Math. Soc., Providence, RI, 1999.

[JS87] R. Jensen and P. E. Souganidis. A regularity result for viscosity solutions of Hamilton-Jacobi equations in one space dimension. *Trans. Amer. Math. Soc.*, 301(1):137–147, 1987.

[Lio82] Pierre-Louis Lions. *Generalized solutions of Hamilton-Jacobi equations.* Pitman (Advanced Publishing Program), Boston, Mass., 1982.

[LPV88] P. L. Lions, G. Papanicolaou, and S. R. S. Varadhan. Homogeneization of hamilton-jacobi equations. *Preliminary Version*, 1988.

[Mat89a] John N. Mather. Minimal action measures for positive-definite Lagrangian systems. In *IXth International Congress on Mathematical Physics (Swansea, 1988)*, pages 466–468. Hilger, Bristol, 1989.

[Mat89b] John N. Mather. Minimal measures. *Comment. Math. Helv.*, 64(3):375–394, 1989.

[Mat91] John N. Mather. Action minimizing invariant measures for positive definite Lagrangian systems. *Math. Z.*, 207(2):169–207, 1991.

[Mn91] Ricardo Mañé. *Global Variational Methods in Conservative Dynamics.* CNPQ, Rio de Janeiro, 1991.

[Mn92] Ricardo Mañé. On the minimizing measures of Lagrangian dynamical systems. *Nonlinearity*, 5(3):623–638, 1992.

[Mn96] Ricardo Mañé. Generic properties and problems of minimizing measures of Lagrangian systems. *Nonlinearity*, 9(2):273–310, 1996.

[Mos86] Jürgen Moser. Recent developments in the theory of Hamiltonian systems. *SIAM Rev.*, 28(4):459–485, 1986.

[MS88] K. R. Meyer and D. G. Saari, editors. *Hamiltonian dynamical systems*, Providence, RI, 1988. American Mathematical Society.

[Nus91] Roger D. Nussbaum. Convergence of iterates of a nonlinear operator arising in statistical mechanics. *Nonlinearity*, 4(4):1223–1240, 1991.

[Oli95] Waldyr Oliva. *Integrability problems in Hamiltonian systems.* IST - Lecture Notes, Lisbon, 1995.

[Pös82] Jürgen Pöschel. The concept of integrability on Cantor sets for Hamiltonian systems. *Celestial Mech.*, 28(1-2):133–139, 1982.

[Roc66] R. T. Rockafellar. Extension of Fenchel's duality theorem for convex functions. *Duke Math. J.*, 33:81–89, 1966.