PERTURBATION THEORY FOR VISCOSITY SOLUTIONS OF HAMILTON-JACOBI EQUATIONS AND STABILITY OF AUBRY-MATHER SETS

DIOGO AGUIAR GOMES*

Abstract. In this paper we study the stability of integrable Hamiltonian systems under small perturbations, proving a weak form of the KAM/Nekhoroshev theory for viscosity solutions of Hamilton-Jacobi equations. The main advantage of our approach is that only a finite number of terms in an asymptotic expansion are needed in order to obtain uniform control. Therefore there are no convergence issues involved. An application of these results is to show that Diophantine invariant tori and Aubry-Mather sets are stable under small perturbations.

Key words. Hamiltonian Dynamics, KAM theory, Aubry-Mather sets, viscosity solutions

AMS subject classifications. 37J40, 49L25, 35C20

1. Introduction.

We consider Hamiltonians of the form

$$H_\epsilon(p, x) = H_0(p) + \epsilon H_1(p, x), \tag{1.1}$$

with H_0, H_1 smooth, $H_0(p)$ strictly convex and $H_1(p, x)$ bounded with bounded derivatives, and $\mathbb{Z}^n$ periodic in x. The objective of this paper is to understand the dependence on ϵ, of periodic viscosity solutions (for the definition of viscosity solution see section 3) of

$$H_\epsilon(P + D_x u^\epsilon, x) = \overline{H}_\epsilon(P), \tag{1.2}$$

and prove stability of Aubry-Mather sets [28], [29], [30], [26], [27], under small perturbations. The equation (1.2) has two unknowns, the function u^ϵ and the value of the effective Hamiltonian $\overline{H}_\epsilon(P)$. Given a viscosity solution of (1.2), the Mather set is a set invariant under the Hamiltonian dynamics

$$\dot{x} = -D_p H_\epsilon(p, x) \qquad \dot{p} = D_x H_\epsilon(p, x)$$

that is contained on the graph $(x, p) = (x, P + D_x u^\epsilon(x))$. For smooth solutions of (1.2) it corresponds to KAM tori.

We are given a reference value $P = P_0$, and we assume that for $\epsilon = 0$ the rotation vector $\omega_0 = D_P \overline{H}_0(P_0)$ satisfies Diophantine non-resonance conditions

$$|\omega_0 \cdot k| \geq \frac{C}{|k|^s}, \tag{1.3}$$

for some positive constant C and some real $s > 0$. Is well known that the KAM theory applies to all non-resonant vectors ω_0. In particular it implies that for sufficiently small perturbations the solution to (1.2) is smooth. However, our results hold even if the solution to (1.2) fails to be smooth.

In general, if one keeps the momentum P_0 fixed, for $\epsilon > 0$ the new rotation vector $D_P \overline{H}_\epsilon(P_0)$ may fail to be Diophantine. In particular an ergodic flow for $\epsilon = 0$ may

*Departamento de Matemática, Instituto Superior Técnico, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal (dgomes@math.ist.utl.pt). Supported in part by FCT (Portugal) through programs POCTI, POCTI/32931/MAT/2000, BPD 1531/2000

give rise to periodic orbits when $\epsilon > 0$. Therefore it is convenient to let the momentum P change with ϵ while keeping the rotation vector fixed. To that effect, given $N > 0$ we construct an approximate solution $\tilde{u}_N^\epsilon(x, P)$ of (1.2), that is $\tilde{u}_N^\epsilon(x, P)$ solves

$$H_\epsilon(P + D_x\tilde{u}^\epsilon, x) = \tilde{H}_\epsilon^N(P) + O(\epsilon^N + |P - P_0|^N).$$

for all P close to P_0. Using the implicit function theorem we can prove that there exists a new momentum $P_\epsilon = P_0 + O(\epsilon)$ such that $D_P\tilde{H}_\epsilon^N(P_\epsilon) = D_P\overline{H}_0(P_0)$. Therefore we are able to keep the rotation vector fixed (up to order N) under small perturbations.

The two main results of this paper are stated in theorems 3 and 4. In theorem 3, we show that for any $M > 0$ there exists $N(M)$ such that

$$|u^\epsilon(x, P_\epsilon) - \tilde{u}_N^\epsilon(x, P_\epsilon)| = O(\epsilon^M).$$

To show this, we first prove estimates along trajectories of the Hamiltonian flow. Then we extend these estimates to nearby points using a priori Lipschitz continuity of viscosity solutions. To that effect, we use the fact that rotation vector is kept unchanged so that we can take advantage of the Diophantine properties. These properties imply that the Hamiltonian flow takes at most time $O\left(\frac{1}{\epsilon^{Ms}}\right)$ to visit an arbitrary ϵ^M neighborhood of any point of the torus [5] (similar estimates were proved originally in [8], [9]). From this we can extend the estimates along trajectories to ϵ^M neighborhoods, using the Lipschitz continuity of viscosity solutions, and therefore obtain uniform control. The technique of the proof would break down if we choose $P_\epsilon = P_0$ because we do not have any control on the number theoretical properties of $D_P\overline{H}_\epsilon(P_0)$.

Finally, in theorem 4, we prove estimates for the derivatives of viscosity solutions, showing that

$$\operatorname{esssup} |D_x u^\epsilon - D_x\tilde{u}_N^\epsilon| = O(\epsilon^{M/2}).$$

Note that these estimates do not imply that u^ϵ is differentiable - therefore they apply even when the invariant tori that exist when $\epsilon = 0$ cease to exits and are replaced by Mather sets. Therefore this result implies the stability of Mather sets under small perturbations.

Our results should be compared previous results. One result is KAM theory that states that for all Diophantine rotation vectors, u^ϵ is smooth and can be computed by summing a convergent series, for ϵ small enough. A result that is somewhat close to ours is the following (see [4], [31]): suppose P_0 is such that the corresponding frequency is Diophantine. Then, there exists a canonical transformation T_ϵ, defined in the domain $|P - P_0| \leq C\epsilon$ such that in the new coordinates, the new Hamiltonian can be written as

$$H_0(P) + \tilde{H}_\epsilon(P) + e^{-C/\epsilon 1/n}R(x, P),$$

that is the system is very close to integrable in an appropriate coordinate system. In the other extreme one has that the viscosity solution u^ϵ, which is Lipschitz, converges uniformly to a constant, provided the rotation number is non-resonant ($\omega_0 \cdot k \neq 0$, for all $k \in \mathbb{Z}^n$) [22]. The results in this paper apply to all Diophantine rotation vectors, provided the Hamiltonian is smooth enough so that one can construct $\tilde{u}_N^\epsilon$ to all orders, and yields an asymptotic representation, for small ϵ, of u^ϵ and its derivatives, even when there is not smoothness, or no convergence of the approximations.

The plan of this paper is the following: in section 2 we review the classical Linstedt series expansion for perturbations of non-resonant integrable Hamiltonians. In

section 3 we recall the necessary background from the theory of viscosity solutions and its relation to Aubry Mather theory. In section 4 we study the expansions for $\overline{H}_\epsilon$. Uniform estimates for viscosity solutions are discussed in section 5. In the last section we present some applications to the stability of viscosity solutions and Aubry-Mather sets, bootstrapping from L^∞ estimates for viscosity solutions to $W^{1,\infty}$ estimates.

Other examples of asymptotic expansions for Hamilton-Jacobi equations in the case at which the perturbation is a second order elliptic operator can be found in [18], [19] and [3]. I would like to thank the Referees, L. C. Evans, R. de laLlave, and C. Valls for many suggestions and comments on the original version of this paper.

2. Classical Perturbation Theory. In this section we review the classical perturbation theory for Hamiltonian systems using a construction equivalent to the Poincaré normal form near an invariant tori. Somewhat incorrectly, but following [1], we call it the Linstedt series method. Although these results are fairly standard, see [1], for instance, we present them in a more convenient form for our purposes.

Consider the Hamiltonian dynamics:

$$\begin{cases} \dot{\mathbf{x}} = -D_p H_\epsilon(\mathbf{p}, \mathbf{x}) \\ \dot{\mathbf{p}} = D_x H_\epsilon(\mathbf{p}, \mathbf{x}), \end{cases} \tag{2.1}$$

we use the convention that boldface $(\mathbf{x}, \mathbf{p})$ are trajectories of the Hamiltonian flow and not the coordinates (x, p). The Hamilton-Jacobi integrability theory suggests that we should look for functions $\overline{H}_\epsilon(P)$ and $u^\epsilon(x, P)$, periodic in x, solving the Hamilton-Jacobi equation:

$$H_\epsilon(P + D_x u^\epsilon, x) = \overline{H}_\epsilon(P). \tag{2.2}$$

Then, by performing the change of coordinates $(x, p) \leftrightarrow (X, P)$ determined by:

$$\begin{cases} X = x + D_P u^\epsilon \\ p = P + D_x u^\epsilon, \end{cases} \tag{2.3}$$

the dynamics (2.1) is simplified to

$$\begin{cases} \dot{\mathbf{X}} = -D_P \overline{H}(\mathbf{P}) \\ \dot{\mathbf{P}} = 0, \end{cases}$$

we use again the convention that boldface $(\mathbf{X}, \mathbf{P})$ are trajectories of the Hamiltonian flow and not the new coordinates (X, P).

If $\tilde{u}$ is an approximate solution to (2.2) satisfying

$$H_\epsilon(P + D_x \tilde{u}, x) = \overline{H}_\epsilon(P) + f(x, P), \tag{2.4}$$

then the change of coordinates (2.3) transforms (2.1) into

$$\begin{cases} \dot{\mathbf{X}} = -D_P \overline{H}_\epsilon(\mathbf{P}) - D_P f(\mathbf{X}, \mathbf{P}) \\ \dot{\mathbf{P}} = D_X f(\mathbf{X}, \mathbf{P}), \end{cases} \tag{2.5}$$

with the convention that $f(X, P) = f(x(X, P), P)$.

The KAM theory deals with constructing solutions of (2.2) by using an iterative procedure, a modified Newton's method, that yields an expansion

$$u^\epsilon = u_0 + \epsilon v_1 + \epsilon^2 v_2 \cdots .$$

The main technical point in KAM theory is to prove the convergence of these expansions. An alternate method that yields such an expansion is the Linstedt series [1]. However we should point out that whereas the KAM expansion is a convergent one, the Linstedt series may fail to converge. Nevertheless, since we will only need finitely many terms we will use a variation of the Linstedt series that we describe next.

We say that a vector $\omega \in \mathbb{R}^n$ is Diophantine if for all $k \in \mathbb{Z}^n \backslash \{0\}$, $|\omega \cdot k| \geq \frac{C}{|k|^s}$, for some $C, s > 0$. Let P_0 be such that $\omega_0 = D_P \overline{H}_0(P_0)$ is Diophantine. We look for an approximate solution of

$$H_\epsilon(P + D_x u^\epsilon(x, P), x) = \overline{H}_\epsilon(P),$$

valid for $P = P_0 + O(\epsilon)$. When $\epsilon = 0$, $\overline{H}_0(P) = H_0(P)$ and the solution u^0 is constant, for instance we may take $u^0 \equiv 0$. For $\epsilon > 0$ we have, formally, $u^\epsilon = O(\epsilon)$, and so we suggests the following approximation $\tilde{u}_N^\epsilon$ to u^ϵ:

$$\begin{aligned} \tilde{u}_N^\epsilon =& \epsilon v_1(x, P_0) + \epsilon(P - P_0) D_P v_1(x, P_0) + \epsilon^2 v_2(x, P_0) + \\ &+ \frac{1}{2}\epsilon(P - P_0)^2 D_{PP}^2 v_1(x, P_0) + \epsilon^2 (P - P_0) D_P v_2(x, P_0) + \\ &+ \epsilon^3 v_3(x, P_0) + \cdots , \end{aligned} \tag{2.6}$$

this expansion is carried out up to order $N-1$ in such a way that, formally $u^\epsilon - \tilde{u}_N^\epsilon = O(\epsilon^N)$. For example

$$\tilde{u}_1^\epsilon = 0, \quad \tilde{u}_2^\epsilon = \epsilon v_1, \quad \tilde{u}_3^\epsilon = \epsilon v_1 + \epsilon^2 v_2 + \epsilon(P - P_0) D_P v_1.$$

The functions v_i and $D_{P^k}^k v_i$ satisfy transport equations

$$D_p H_0(P_0) D_x w = f(\cdots),$$

for some suitable f, and can be solved inductively. For instance:

$$\overline{H}_1(P_0) = D_p H_0(P_0) D_x v_1 + H_1(P_0, x),$$

$$D_P \overline{H}_1(P_0) = D_p H_0(P_0) D_x (D_P v_1) + D_{pp}^2 H_0(P_0) D_x v_1 + D_p H_1(P_0, x),$$

and

$$\begin{gathered} \overline{H}_2(P_0) = \\ D_p H_0(P_0) D_x v_2 + \frac{1}{2} D_{pp}^2 H_0(P_0) D_x v_1 D_x v_1 + D_p H_1(P_0, x) D_x v_1. \end{gathered}$$

Note that the derivatives of v_i with respect to P, $D_{P^k}^k v_i$, are computed by solving appropriate transport equations, as is illustrated above for $D_P v_1$, and not by differentiating v_i. In fact v_i may not be defined for $P \neq P_0$. However if its derivative exists it satisfies a transport equation.

The constants $\overline{H}_1(P_0), D_P \overline{H}_1(P_0), \overline{H}_2(P_0)$... are uniquely determined by integral compatibility conditions, for example,

$$\overline{H}_1(P_0) = \int H_1(P_0, x) dx,$$

$$D_P\overline{H}(P_0)=\int D_pH_1(P_0,x)dx,$$

and

$$\overline{H}_2(P_0)=\int \frac{1}{2}D^2_{pp}H_0(P_0)D_xv_1D_xv_1+D_pH_1(P_0,x)D_xv_1dx.$$

If H is sufficiently smooth and ω_0 is non-resonant then these equations have smooth solutions that are unique up to constants. Finally one can check that

$$H_\epsilon(P+D_x\tilde{u}^\epsilon_N,x)=\tilde{H}^N_\epsilon(P)+O(\epsilon^N+|P-P_0|^N), \tag{2.7}$$

with

$$\tilde{H}^N_\epsilon(P)=\overline{H}_0(P_0)+\epsilon\overline{H}_1(P_0)+(P-P_0)D_P\overline{H}_0(P_0)+\epsilon^2\overline{H}_2(P_0)+\cdots,$$

and this expansion is carried up to order $N-1$ in such a way that formally

$$\overline{H}_\epsilon(P)=\tilde{H}^N_\epsilon(P)+O(\epsilon^N+|P-P_0|^N).$$

Consider the change of coordinates

$$\begin{cases} p=P+D_x\tilde{u}^\epsilon_N(x,P)\\ X=x+D_P\tilde{u}^\epsilon_N(x,P).\end{cases}$$

Then, by (2.4) and (2.5), (2.1) is transformed into:

$$\begin{cases} \dot{\mathbf{X}}=-D_P\overline{H}_\epsilon(\mathbf{P})+O(\epsilon^N+|\mathbf{P}-P_0|^{N-1})\\ \dot{\mathbf{P}}=O(\epsilon^N+|\mathbf{P}-P_0|^N).\end{cases}$$

3. Viscosity Solutions, Optimal Control and Mather Measures. In general (2.2) does not admit smooth classical solutions. However, (2.2) has viscosity solutions which are known to be the appropriate notion of weak solution for Hamilton-Jacobi equations. In this section we review the necessary background and in the rest of the paper we extend rigorously the classical perturbation procedure from the previous section to viscosity solutions.

For our purposes, a convenient definition of viscosity solution is the following: we say that a function u is a viscosity solution of (2.2) provided that it satisfies the following fixed point identity:

$$u(x)=\inf\int_0^t L(\mathbf{x},\dot{\mathbf{x}})+P\dot{\mathbf{x}}+\overline{H}(P)dt+u(\mathbf{x}(t)), \tag{3.1}$$

in which the infimum is taken over Lipschitz trajectories $\mathbf{x}(\cdot)$, with initial condition $\mathbf{x}(0)=x$, and $\overline{H}$ is the unique number for which (3.1) holds with $u(x)$ bounded. The Lagrangian L is the Legendre transform of the Hamiltonian:

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

This definition of viscosity solution is equivalent to the standard one [17], [2], as long as the Hamiltonian and Lagrangian L are strictly convex. Note that viscosity solutions do not have to be smooth. However, they are semiconcave, and therefore

twice differentiable almost everywhere. Furthermore they are differentiable along the optimal trajectory $\mathbf{x}(t)$. In fact, the optimal trajectory $\mathbf{x}(t)$ in (3.1), and the momentum $\mathbf{p}(t) = P + D_x u(\mathbf{x}(t))$ are solutions of (2.1), for all $t > 0$.

First, let us quote an existence result [25]:

Theorem 1 (Lions, Papanicolao, Varadhan). *For each P there exists a number $\overline{H}(P)$ and a function $u(x,P)$, periodic in x, that solves (2.2) in the viscosity sense. Furthermore $\overline{H}(P)$ is convex in P and $u(x,P)$ is Lipschitz in x.* This theorem does not assert anything about uniqueness of the viscosity solution u. Indeed, such viscosity solutions are not unique even up to constants, see for instance [6]. However, as it was shown in [22], under certain hypothesis one can prove uniqueness and even continuity of the viscosity solution u with respect to parameters. These hypothesis can be formulated in terms of the ergodic properties of certain measures - Mather measures (see theorem 2) - that are invariant under the Hamiltonian dynamics.

The connection between classical mechanics and viscosity solutions is well known, and was explored by several authors, for instance [13], [14], [15], [16], [10], [11], [12], and [21]. One of the most basic results is the following:

Theorem 2 (A. Fathi, W. E). *Let u be a viscosity solution of (2.2).*

- *For each P there exists a set invariant under the dynamics (2.1) contained in the graph $(x, P + D_x u)$;*
- *There exists a probability measure $\mu(x,p)$ (Mather measure) invariant under (2.1) supported on this invariant set;*
- *This measure minimizes*

$$\int L(x,v) + Pv d\mu, \tag{3.2}$$

with $v = -D_p H(p,x)$, over all probability measures that are invariant under (2.1).

Conversely, any probability measure invariant under (2.1) that minimizes (3.2) is supported on the graph $(x, P + D_x u)$ for any viscosity solution u of (2.2). One of the main advantages of the previous theorem is that one can translate properties of viscosity solutions into properties of Mather sets or measures and vice-versa. Some properties of viscosity solutions are describe in the next theorem:

Proposition 1. *Suppose (x,p) is a point in the graph*

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

Then for all $t > 0$ the solution $(\mathbf{x}(t), \mathbf{p}(t))$ of (2.1) with initial conditions (x,p) belongs $\mathcal{G}$. Proof. The invariance of the graph for $t > 0$ is a consequence of the optimal control interpretation of viscosity solutions [17] and the reader may find a proof, for instance, in [21] or [20]. ■

Finally, an important identity [7] is:

Proposition 2 (Contreras, Iturriaga, Paternain, Paternain).

$$\overline{H}(P) = \inf_{\phi} \sup_{x} H(P + D_x \phi, x), \tag{3.3}$$

in which the infimum is taken over C^1 periodic functions ϕ.

This formula can be used to compute $\overline{H}(P)$ effectively [23], and, in conjunction with properties of viscosity solutions, to detect non-integrability of Hamiltonian systems [24].

4. Estimates for the effective Hamiltonian. We start this section by proving that $\tilde{H}^N_\epsilon$ is an asymptotic expansion to $\overline{H}_\epsilon$. Then we will use convexity techniques to prove estimates for the derivatives of viscosity solutions that we will use in the subsequent sections.

PROPOSITION 3. *Suppose we can construct an approximate solution as in (2.6). Then*

$$\overline{H}_\epsilon(P) \leq \tilde{H}^N_\epsilon(P) + O(\epsilon^N + |P - P_0|^N). \tag{4.1}$$

REMARK. Note that the error term is a function of x, but, by periodicity, it can be estimated uniformly by $O(\epsilon^N + |P - P_0|^N)$.

PROOF. The inf sup formula (3.3) implies

$$\overline{H}_\epsilon(P) \leq \sup_x H_\epsilon(P + D_x\tilde{u}^\epsilon_N, x).$$

By expanding this expression in Taylor series and taking the supremum we obtain the result. ∎

A converse inequality is also true:

PROPOSITION 4. *At any point at which $D_x u^\epsilon$ exists we have*

$$\begin{aligned}\overline{H}_\epsilon(P) \geq{}& H_\epsilon(P + D_x\tilde{u}^\epsilon_N, x) + D_pH_\epsilon(P + D_x\tilde{u}^\epsilon_N, x)(D_x\tilde{u}^\epsilon_N - D_xu^\epsilon) + \\ &+ \frac{\gamma}{2}|D_x\tilde{u}^\epsilon_N - D_xu^\epsilon|^2,\end{aligned} \tag{4.2}$$

for some positive constant $\gamma > 0$.

PROOF. Since $H_\epsilon(p, x)$ is strictly convex, there exists a constant $\gamma > 0$ such that

$$\begin{aligned}\overline{H}_\epsilon(P) &\geq H_\epsilon(P + D_xu^\epsilon, x) \geq \\ &\geq H_\epsilon(P + D_x\tilde{u}^\epsilon_N, x) + D_pH_\epsilon(P + D_x\tilde{u}^\epsilon_N, x)(D_x\tilde{u}^\epsilon_N - D_xu^\epsilon) + \\ &\quad + \frac{\gamma}{2}|D_x\tilde{u}^\epsilon_N - D_xu^\epsilon|^2.\end{aligned}$$

∎

The following corollary is going to be used in the next section:

COROLLARY 1. *If there exists an approximate solution as in proposition 3, then there exists a point x_0 for which*

$$|D_x\tilde{u}^\epsilon(x_0) - D_xu^\epsilon(x_0)| \leq C\epsilon^{N/2}.$$

PROOF. Since u^ϵ is semiconcave, $\tilde{u}^\epsilon - u^\epsilon$ is semiconvex. Therefore at the maximum x_0, $D_x(\tilde{u}^\epsilon - u^\epsilon) = 0$, that is, the derivative exists and is zero. Then the two previous propositions yield the desired result. ∎

COROLLARY 2. *If there exists an approximate solution as in proposition 3, then*

$$\overline{H}_\epsilon = \tilde{H}^N_\epsilon + O(\epsilon^N).$$

PROOF. It suffices to combine propositions 3 and 4 at the point x_0 given by the previous corollary. ∎

5. Uniform estimates. In this section we prove uniform estimates

$$\sup_x |u^\epsilon - \tilde{u}^\epsilon_N| = O(\epsilon^M),$$

under Diophantine conditions on the rotation vector ω_0 of the unperturbed problem. These results should be compared with the ones in [22] that show that unique ergodicity of the Mather measure implies uniform continuity of the viscosity solution with respect to the parameters.

First we construct an approximate solution $\tilde{u}^\epsilon_N$ for a sufficiently large but finite N using the Linstedt series expansion described in section 2. Then, by changing coordinates, we show that for long times there are uniform estimates along trajectories. Finally, by using Lipschitz estimates for u^ϵ and $\tilde{u}^\epsilon_N$ and the results by [8] on ergodization times we extend them to whole space.

Theorem 3. *Suppose the rotation vector:*

$$\omega_0 = D_P \overline{H}_0(P_0)$$

satisfies the Diophantine property (1.3). Furthermore suppose that for every N, there exists $P_\epsilon = P + O(\epsilon)$

$$D_P \tilde{H}^N_\epsilon(P_\epsilon) = \omega_0,$$

that is, the approximate rotation vector corresponding to P_ϵ is the original rotation vector ω_0. Let u^ϵ be a solution of

$$H_\epsilon(P_\epsilon + D_x u^\epsilon, x) = \overline{H}_\epsilon(P_\epsilon),$$

and $\tilde{u}^\epsilon_N$ be the corresponding approximate solution using a Linstedt series expansion up to order N.

Then for every M there exists $N(M)$ such that

$$\sup_x |u^\epsilon - \tilde{u}^\epsilon_N| = O(\epsilon^M).$$

Proof. Define P_ϵ by solving the equation

$$\omega_0 = D_P \tilde{H}^N_\epsilon(P_\epsilon),$$

that is,

$$\omega_0 = D_P \overline{H}_0(P_0) + \epsilon D_P \overline{H}_1(P_0) + (P_\epsilon - P_0) D^2_{PP} \overline{H}_0(P_0) + \cdots,$$

with expansion taken up to order $N-1$. Under the non-degeneracy condition $\det D^2_{pp}\overline{H}_0 \neq 0$ (which holds because $\overline{H}_0(P) = H_0(P)$ is strictly convex), the implicit function theorem yields a unique solution of the form

$$P_\epsilon = P_0 + \epsilon P_1 + \cdots$$

with $P_1 = -\left[D^2_{PP}\overline{H}_0(P_0)\right]^{-1} D_P \overline{H}_1(P_0)$.

Define the new coordinates (P, X) by

$$\begin{cases} p = P + D_x \tilde{u}^\epsilon_N(x, P) \\ X = x + D_P \tilde{u}^\epsilon_N(x, P). \end{cases}$$

To simplify notation we denote $X = \phi_\epsilon(x)$.

Let x_0 be the point given by corollary 1. Set

$$(\mathbf{x}(0), \mathbf{p}(0)) = (x_0, P_\epsilon + D_x u^\epsilon(x_0))$$

be the initial conditions for a trajectory $(\mathbf{x}(t), \mathbf{p}(t))$ of (2.1). Since $|D_x \tilde{u}^\epsilon(x_0) - D_x u^\epsilon(x_0)| \leq C\epsilon^{N/2}$, in the new coordinates we have

$$\mathbf{P}(0) = P_\epsilon + O(\epsilon^{N/2}).$$

Consider the Hamiltonian dynamics in the new coordinates (X, P), with initial condition $(\mathbf{X}(0), \mathbf{P}(0))$ (the value $\mathbf{X}(0)$ is not important)

$$\begin{cases} \dot{\mathbf{X}} = -D_P \tilde{H}_\epsilon(\mathbf{P}) + O(\epsilon^N + |\mathbf{P} - P_0|^{N-1}) \\ \dot{\mathbf{P}} = O(\epsilon^N + |\mathbf{P} - P_0|^N). \end{cases}$$

From this equation it follows that the momentum P in the new coordinates is conserved for long times:

Proposition 5.

$$\sup_{0 \leq t \leq \frac{1}{\epsilon^{N/2}}} |\mathbf{P}(t) - P_\epsilon| \leq O(\epsilon^{N/4}),$$

Proof. Note that

$$\begin{aligned} \frac{d}{dt}|\mathbf{P} - P_\epsilon|^2 &\leq C|\mathbf{P} - P_\epsilon|(\epsilon^N + |\mathbf{P} - P_\epsilon|^N) \leq \\ &\leq C\epsilon^N |\mathbf{P} - P_\epsilon|^2 + C\epsilon^N, \end{aligned}$$

as long as $|\mathbf{P} - P_\epsilon|^{N-1} \leq C\epsilon^N$. Note that for $N > 3$, and we are always assuming N large enough, we have $|\mathbf{P}(0) - P_\epsilon|^{N-1} \leq C\epsilon^N$. Thus Gronwall inequality implies

$$|\mathbf{P}(T) - P_\epsilon|^2 \leq e^{C\epsilon^N T} \left(|\mathbf{P}(0) - P_\epsilon|^2 + C\epsilon^N T \right).$$

Therefore, up to $T = \frac{1}{\epsilon^{N/2}}$ we have

$$|\mathbf{P}(t) - P_\epsilon|^2 \leq C\epsilon^{N/2},$$

for ϵ sufficiently small. ∎

Observe that $\phi_\epsilon(x) = X$ is a diffeomorphism for small ϵ. Let

$$U(X) = u^\epsilon(\phi_\epsilon^{-1}(X)) - \tilde{u}_N^\epsilon(\phi_\epsilon^{-1}(X)),$$

in particular

$$U(\mathbf{X}(t)) = u^\epsilon(\mathbf{x}(t)) - \epsilon v_1(\mathbf{x}(t)) - \epsilon(P_\epsilon - P_0) D_P v_1(\mathbf{x}(t)) - \cdots$$

Recall that $D_x u^\epsilon(\mathbf{x}(t)) = \mathbf{p}(t) - P_\epsilon$. Thus

$$\begin{aligned} \frac{d}{dt} U(\mathbf{X}(t)) &= (\mathbf{p}(t) - P_\epsilon) D_p H_\epsilon(\mathbf{p}(t), \mathbf{x}(t)) - \\ &- \epsilon D_x v_1(\mathbf{x}(t)) D_p H_\epsilon(\mathbf{p}(t), \mathbf{x}(t)) - \cdots = \\ &= (\mathbf{p}(t) - P_\epsilon - \epsilon D_x v_1 - \cdots) D_p H_\epsilon = \\ &= (\mathbf{P}(t) - P_\epsilon) D_p H_\epsilon(\mathbf{p}(t), \mathbf{x}(t)), \end{aligned}$$

since $\mathbf{p}(t) = \mathbf{P}(t) + \epsilon D_x v_1(\mathbf{x}(t)) + \cdots$. Therefore

$$\frac{d}{dt}U(\mathbf{X}(t)) = O(\epsilon^{N/4}),$$

for $0 \le t \le \frac{1}{\epsilon^{N/2}}$.

We may also add a convenient constant to u^ϵ in such a way that $U(\mathbf{X}(0)) = 0$ and so we obtain:

$$\sup_{0 \le t \le \frac{1}{\epsilon^{N/8}}} U(\mathbf{X}(t)) = O(\epsilon^{N/8}),$$

along the trajectory.

Since for small ϵ ϕ_ϵ is a diffeomorphism, U is a Lipschitz function. The Diophantine property implies that the flow

$$\dot{\mathbf{X}} = D_P \overline{H}_0(P_0) + O(\epsilon^{N/2})$$

the Hamiltonian flow takes at most time gets within a distance of ϵ^M of any point in time $T = O\left(\frac{1}{\epsilon^{Ms}}\right)$ (see [5], and also [8], [9]). Thus, if $M < \frac{N}{8s}$, we get for some $0 \le t \le \frac{1}{\epsilon^{N/8}}$ that

$$|X - \mathbf{X}(t)| \le C\epsilon^M,$$

and so

$$|U(X)| \le |U(X) - U(\mathbf{X}(t))| + |U(\mathbf{X}(t))| \le C\epsilon^M.$$

Because ϕ_ϵ is a diffeomorphism, the same estimate carries over for the difference $u^\epsilon - \tilde{u}_N^\epsilon$. ■

A final comment is that since $\tilde{u}_N^\epsilon - \tilde{u}_M^\epsilon = O(\epsilon^M)$, we also have

$$\sup_x |u^\epsilon - \tilde{u}_M^\epsilon| = O(\epsilon^M),$$

although we need the existence of $\tilde{u}_N^\epsilon$ to prove this estimate.

6. Applications - stability of Mather sets and regularity. This last section is dedicated to prove estimates on the derivatives $D_x u^\epsilon - D_x \tilde{u}^\epsilon$. Such estimates rely on the uniform estimates from the previous section. Since the Mather sets are supported on the graph $(x, P + D_x u)$ estimates on the derivatives show the stability of Mather sets.

Proposition 6. *Suppose $\omega_0 = D_P H_0(P_0)$ is Diophantine and (1.2) admits an approximate solution $\tilde{u}_N^\epsilon$. Then*

$$\frac{1}{T}\int_0^T \frac{\gamma}{2}|D_x u^\epsilon(\mathbf{x}(t)) - D_x \tilde{u}_N^\epsilon(\mathbf{x}(t))|^2 dt \le$$
$$\le C\epsilon^N + \frac{2}{T}\sup_x |u^\epsilon - \tilde{u}_N^\epsilon|,$$

in which the integral is taken along a trajectory $\mathbf{x}(\cdot)$ of

$$\dot{\mathbf{x}}(t) = -D_p H_\epsilon(P_\epsilon + D_x u^\epsilon(\mathbf{x}(t)), \mathbf{x}(t)).$$

PROOF. The strict convexity of H_ϵ together with corollary 2 implies

$$O(\epsilon^N) \geq D_p H_\epsilon(P_0 + D_x u^\epsilon(\mathbf{x}(t)), \mathbf{x}(t))(D_x \tilde{u}_N^\epsilon(\mathbf{x}(t)) - D_x u^\epsilon(\mathbf{x}(t))) + \\ + \frac{\gamma}{2}|D_x \tilde{u}_N^\epsilon(\mathbf{x}(t)) - D_x u^\epsilon(\mathbf{x}(t))|^2.$$

Integrating with respect to t and observing that

$$\int_0^T D_p H_\epsilon(P_0 + D_x u^\epsilon(\mathbf{x}(t)), \mathbf{x}(t))(D_x \tilde{u}_N^\epsilon(\mathbf{x}(t)) - D_x u^\epsilon(\mathbf{x}(t)))dt = \\ = -\int_0^T \dot{\mathbf{x}}(t)(D_x \tilde{u}_N^\epsilon(\mathbf{x}(t)) - D_x\ u^\epsilon(\mathbf{x}(t)))dt = \\ = -u^\epsilon(\mathbf{x}(0)) + \tilde{u}_N^\epsilon(\mathbf{x}(0)) + u^\epsilon(\mathbf{x}(T)) - \tilde{u}_N^\epsilon(\mathbf{x}(T)),$$

we obtain the result. ■

This proposition is the key to prove the main result of this paper that is discussed in the next theorem - pointwise estimates for first derivatives of viscosity solutions.

THEOREM 4. *Let $M > 0$. Suppose $\omega_0 = D_P H_0(P_0)$ is Diophantine such that the cell problem (1.2) admits an approximate solution $\tilde{u}_N^\epsilon$ for N sufficiently large such that theorem 3 holds. Then*

$$\operatorname*{esssup}_x |D_x u^\epsilon - D_x \tilde{u}_N^\epsilon| \leq C\epsilon^{M/2}.$$

PROOF. Since $\sup_x |\tilde{u}_N^\epsilon - u^\epsilon| = O(\epsilon^M)$ we have

$$\int_0^1 \frac{\gamma}{2}|D_x u^\epsilon(\mathbf{x}(t)) - D_x \tilde{u}_N^\epsilon(\mathbf{x}(t))|^2 dt \leq C\epsilon^M, \tag{6.1}$$

with $\mathbf{x}(t)$ as in the previous theorem, and for any initial condition $\mathbf{x}(0) = x$.

Let G be the set of the points at which $D_x u^\epsilon$ exists and such that

$$|D_x u^\epsilon - D_x \tilde{u}_N^\epsilon| \leq C\epsilon^{M/2},$$

for some fixed constant C, and set

$$B = \{x \in G^c | u^\epsilon \text{ is differentiable at } x\}.$$

Since u^ϵ is Lipschitz then $(B \cup G)^c$, which is the set of points of non-differentiability of u^ϵ, is of zero Lebesgue measure.

Let x be a point for which $D_x u^\epsilon$ exists and $|D_x \tilde{u}_N^\epsilon - D_x u^\epsilon| > C\epsilon^{M/2}$. Define $p_x = P_\epsilon + D_x u^\epsilon(x)$. Let $(\mathbf{x}(t), \mathbf{p}(t))$ be the solution of (2.1) with initial conditions (x, p_x).

The estimate (6.1) implies that we may assume that there exists $0 < T < 1$ such that $\mathbf{x}(T) \in G$. Let $y = \mathbf{x}(T)$ and

$$\tilde{p}_y = P_\epsilon + D_x \tilde{u}_N^\epsilon(\mathbf{x}(T)), \quad p_y = P_\epsilon + D_x u^\epsilon(\mathbf{x}(T)).$$

Since $y \in G$ we have

$$|\tilde{p}_y - p_y| \leq C\epsilon^{M/2}.$$

Let $(\mathbf{x}(t), \tilde{\mathbf{p}}(t))$ be the solution of (2.1) with initial conditions $(y, \tilde{p}_y)$. Define $\tilde{x} = \mathbf{x}(-T)$, $p_{\tilde{x}} = \tilde{\mathbf{p}}(-T)$.

By standard ODE theory

$$|p_x - p_{\tilde{x}}| \le C\epsilon^{M/2} \qquad |x - \tilde{x}| \le C\epsilon^{M/2}.$$

Note that

$$\begin{aligned}
&|D_x u^\epsilon(x) - D_x \tilde{u}_N^\epsilon(x)| \\
&\qquad \le |P_\epsilon + D_x u^\epsilon(x) - p_{\tilde{x}}| + |p_{\tilde{x}} - P_\epsilon - D_x \tilde{u}_N^\epsilon(\tilde{x})| + \\
&\qquad\quad + |D_x \tilde{u}_N^\epsilon(\tilde{x}) - D_x \tilde{u}_N^\epsilon(x)|.
\end{aligned}$$

The first term is controlled by

$$|P_\epsilon + D_x u^\epsilon(x) - p_{\tilde{x}}| = |p_x - p_{\tilde{x}}| \le C\epsilon^{M/2}.$$

The last term is controlled by the Lipschitz constant of $\tilde{u}_N^\epsilon$

$$|D_x \tilde{u}_N^\epsilon(\tilde{x}) - D_x \tilde{u}_N^\epsilon(x)| \le C|x - \tilde{x}| \le C\epsilon^{M/2}.$$

Therefore it suffices to estimate $|p_{\tilde{x}} - P_\epsilon - D_x \tilde{u}_N^\epsilon(\tilde{x})|$. To that effect observe that from differentiating (2.7) with respect to x it follows

$$D_p H(P_\epsilon + D_x \tilde{u}_N^\epsilon, x) D_{xx}^2 \tilde{u}_N^\epsilon + D_x H(P_\epsilon + D_x \tilde{u}_N^\epsilon, x) = O(\epsilon^N). \tag{6.2}$$

Then by combining

$$\begin{aligned}
&\frac{d}{dt} \frac{|\tilde{\mathbf{p}}(t) - P_\epsilon - D_x \tilde{u}_N^\epsilon(\tilde{\mathbf{x}}(t)|^2}{2} = \\
&\quad = (\tilde{\mathbf{p}}(t) - P_\epsilon - D_x \tilde{u}_N^\epsilon(\tilde{\mathbf{x}}(t)) \left[D_x H(\tilde{\mathbf{p}}(t), \tilde{\mathbf{x}}(t)) + \right. \\
&\quad\quad \left. + D_{xx}^2 \tilde{u}_N^\epsilon(\tilde{\mathbf{x}}(t) \right] D_p H(\tilde{\mathbf{p}}(t), \tilde{\mathbf{x}}(t))),
\end{aligned}$$

with (6.2) we obtain

$$\frac{d}{dt} \frac{|\tilde{\mathbf{p}}(t) - P_\epsilon - D_x \tilde{u}_N^\epsilon(\tilde{\mathbf{x}}(t)|^2}{2} \le C|\tilde{\mathbf{p}}(t) - P_\epsilon - D_x \tilde{u}_N^\epsilon(\tilde{\mathbf{x}}(t)|^2 + O(\epsilon^N).$$

Then Gronwall inequality yields:

$$|p_{\tilde{x}} - P_\epsilon - D_x \tilde{u}_N^\epsilon(\tilde{x})|^2 \le C\epsilon^N.$$

■

REFERENCES

[1] 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].

[2] 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.

[3] U. Bessi, *Smooth approximation of Mather sets*, Preprint.

[4] F. Bonetto, G. Gallavotti, G. Gentile, and V. Mastropietro, *Lindstedt series, ultraviolet divergences and Moser's theorem*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 26 (1998), pp. 545–593.

[5] J. Bourgain, F. Golse, and B. Wennberg, *On the distribution of free path lengths for the periodic Lorentz gas*, Comm. Math. Phys., 190 (1998), pp. 491–508.
[6] M. Concordel, *Periodic homogenization of Hamilton-Jacobi equations*, PhD thesis, UC Berkeley, 1995.
[7] G. Contreras, R. Iturriaga, G. P. Paternain, and M. Paternain, *Lagrangian graphs, minimizing measures and Mañé's critical values*, Geom. Funct. Anal., 8 (1998), pp. 788–809.
[8] H. S. Dumas, *Ergodization rates for linear flow on the torus*, J. Dynam. Differential Equations, 3 (1991), pp. 593–610.
[9] H. S. Dumas, L. Dumas, and F. Golse, *On the mean free path for a periodic array of spherical obstacles*, J. Statist. Phys., 82 (1996), pp. 1385–1407.
[10] W. E, *Aubry-Mather theory and periodic solutions of the forced Burgers equation*, Comm. Pure Appl. Math., 52 (1999), pp. 811–828.
[11] L. C. Evans and D. Gomes, *Effective Hamiltonians and averaging for Hamiltonian dynamics. I*, Arch. Ration. Mech. Anal., 157 (2001), pp. 1–33.
[12] ———, *Effective Hamiltonians and averaging for Hamiltonian dynamics II*, Arch. Ration. Mech. Anal., http://dx.doi.org/10.1007/s002050100181 (2002).
[13] A. Fathi, *Solutions KAM faibles conjuguées et barrières de Peierls*, C. R. Acad. Sci. Paris Sér. I Math., 325 (1997), pp. 649–652.
[14] ———, *Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens*, C. R. Acad. Sci. Paris Sér. I Math., 324 (1997), pp. 1043–1046.
[15] ———, *Orbite hétéroclines et ensemble de Peierls*, C. R. Acad. Sci. Paris Sér. I Math., 326 (1998), pp. 1213–1216.
[16] ———, *Sur la convergence du semi-groupe de Lax-Oleinik*, C. R. Acad. Sci. Paris Sér. I Math., 327 (1998), pp. 267–270.
[17] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, Springer-Verlag, New York, 1993.
[18] W. H. Fleming and P. E. Souganidis, *Asymptotic series and the method of vanishing viscosity*, Indiana Univ. Math. J., 35 (1986), pp. 425–447.
[19] ———, *Erratum: "Asymptotic series and the method of vanishing viscosity"*, Indiana Univ. Math. J., 35 (1986), p. 925.
[20] D. Gomes, *Hamilton-Jacobi Equations, Viscosity Solutions and Asymptotics of Hamiltonian Systems*, Ph.D. Thesis, Univ. of California at Berkeley, 2000.
[21] ———, *Viscosity solutions of Hamilton-Jacobi equations, and asymptotics for Hamiltonian systems*, To appear in J. Cal. of Var., (2000).
[22] ———, *Regularity theory for Hamilton-Jacobi equations*, to appear in Journal of Differential Equations, (2002).
[23] D. Gomes and A. Oberman, *Computational methods for effective Hamiltonians*, Preprint, (2002).
[24] ———, *Converse KAM theory*, In preparation, (2002).
[25] P. L. Lions, G. Papanicolao, and S. R. S. Varadhan, *Homogeneization of Hamilton-Jacobi equations*, Preliminary Version, (1988).
[26] R. Mañé, *On the minimizing measures of Lagrangian dynamical systems*, Nonlinearity, 5 (1992), pp. 623–638.
[27] ———, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity, 9 (1996), pp. 273–310.
[28] J. N. Mather, *Minimal action measures for positive-definite Lagrangian systems*, in IXth International Congress on Mathematical Physics (Swansea, 1988), Hilger, Bristol, 1989, pp. 466–468.
[29] ———, *Minimal measures*, Comment. Math. Helv., 64 (1989), pp. 375–394.
[30] ———, *Action minimizing invariant measures for positive definite Lagrangian systems*, Math. Z., 207 (1991), pp. 169–207.
[31] A. Morbidelli and A. Giorgilli, *Superexponential stability of KAM tori*, J. Statist. Phys., 78 (1995), pp. 1607–1617.