M P E J

Mathematical Physics Electronic Journal

ISSN 1086-6655
Volume 12, 2006

Paper 3
Received: Nov 2, 2005, Revised: May 8, 2006, Accepted: Jul 15, 2006
Editor: K. Khanin

PERTURBATION THEORY AND DISCRETE HAMILTONIAN DYNAMICS

DIOGO A. GOMES AND CLAUDIA VALLS

Departamento de Matemática
Instituto Superior Técnico
Av. Rovisco Pais
1049-001 Lisboa, Portugal

Abstract. In this paper we discuss a weak version of KAM theory for symplectic maps which arise from the discretization of the minimal action principle. These maps have certain invariant sets, the Mather sets, which are the generalization of KAM tori in the non-differentiable case. These sets support invariant measures, the Mather measures, which are action minimizing measures. We generalize viscosity solution methods to study discrete systems. In particular, we show that, under non-resonance conditions, the Mather sets can be approximated uniformly, up to any arbitrary order, by finite perturbative expansions. We also present new results concerning the approximation of Mather measures.

1. Introduction

In this paper, we discuss perturbation methods for symplectic maps that arise from the discretization of the minimal action principle. The motivation for this work is the following: the phase space of close to integrable Hamiltonian system can be split into regular parts, composed of invariant tori in which the system displays very simple periodic or quasi-periodic behaviour, and in another part where the system may

This work was partially supported by the Center for Mathematical Analysis, Geometry, and Dynamical Systems and FCT - POCTI-POCI/FEDER, POCI/MAT/55745/2004, and SFRH/BPD/14404/2003.

exhibit irregular behaviour. As the system gets farther from the integrable, the invariant tori do not disappear altogether but shrink to certain sets which have a characteristic action minimizing property that allows to prove their existence by a variational argument. In these sets, one can define certain action minimizing measures, the Mather measures. To observe these sets and measures numerically one must do some type of discretization. In this paper, we discuss a time discretization scheme for autonomous Hamiltonian systems, and study the stability of the action minimizing sets and measures under small perturbations.

The Aubry-Mather theory [Mat89a], [Mat89b], studies invariant measures of Lagrangian systems which have special minimizing properties. In its origin was the study of discrete systems such as area-preserving twist diffeomorphisms [Mat79], [Mat81], [Mat82], [Mat91]. Recently, the techniques of viscosity solutions [Fat97a], [Fat97b], [Fat98a], [Fat98b], [E99], [EG01] have been used with success to study continuous Lagrangian systems and can be appropriately adapted to study discrete systems [Gom02].

One motivation to study symplectic maps instead of flows comes from the fact that by discretizing the minimal action principle from classical mechanics, one obtains certain maps, which are symplectic, and have the form

$$\begin{cases} \mathbf{p}_{n+1} - \mathbf{p}_n = hD_xH(\mathbf{p}_{n+1}, \mathbf{x}_n) \\ \mathbf{x}_{n+1} - \mathbf{x}_n = -hD_pH(\mathbf{p}_{n+1}, \mathbf{x}_n), \end{cases}$$

where h is the time step and H the Hamiltonian. Those maps are discrete analogs of Hamilton's equations

$$\begin{cases} \dot{\mathbf{p}} = D_xH(\mathbf{p}, \mathbf{x}) \\ \dot{\mathbf{x}} = -D_pH(\mathbf{p}, \mathbf{x}). \end{cases}$$

It is very important to notice that this discretization comes from variational principles, as we discuss in section 2, and has interesting geometric and analytic features as will be pointed out later in the paper.

If the Hamiltonian H only depends on p, the dynamics is very simple since $p_n = p_0$, and $x_n = x_0 - nhD_pH(p_0)$. In this case, the system is called integrable. We would like to study Hamiltonians that are close to integrable. That is, when H is of the form $H(p, x) = H_0(p) + O(\epsilon)$, for ϵ a small parameter.

There are certain measures, the Mather measures, characterized by a variational principle (see section 2), which are invariant under the discrete Hamilton's equations. Our main objective is to understand the dependence on ϵ of these measures. In continuous time, this problem was partially addressed in [Gom03]. A classical result in Mather's theory is that the support of these measures (the Mather set) is a Lipschitz graph. In [Gom03] was proved that this set can be well approximated by certain smooth functions. In this paper we extend substantially the results in [Gom03] as we prove new estimates for the Mather measures of the perturbed problem (see section 7). In particular we give a new expression for an approximate density for the Mather measure.

We start section 2 with a discussion of generating functions and formal integrability methods. These results motivate the methods used in the remaining sections to study the dependence on ϵ of the Mather sets. Without these guiding principles it would be extremely hard to proceed as the equations are extremely complex.

For each $P \in \mathbb{R}^n$, one can construct a Mather measure on $\mathbb{T}^n \times \mathbb{R}^n$ that is supported on a Lipschitz graph. This graph can be determined from a solution u of

$$\begin{aligned} u(x,P) - u(\hat{x},P) + hH(P + D_x u(\hat{x},P), x) - \\ - hD_x u(\hat{x}) D_p H(P + D_x u(\hat{x}), x) = h\overline{H}(P), \end{aligned} \tag{1}$$

in which $\hat{x}$ is determined implicitly by

$$\hat{x} - x = -hD_p H(P + D_x u(\hat{x}, P), x).$$

The existence of smooth solutions to this equation can be addressed using KAM theory. However, in general, there are no smooth solutions. Thus one has to consider the class of viscosity solutions. Viscosity solution methods, in particular, definition, existence, regularity, and basic properties, are discussed in section 3. Although our results look similar to KAM, they are valid even after KAM torus cease to exist.

Although viscosity solutions may not be smooth, one can develop a formal expansion $\tilde{u}(x,P)$ in power series in ϵ and $P - P_0$ of the solution $u(x,P)$, in a neighborhood of $\epsilon = 0$ and $P = P_0$. There is a solvability condition in order to construct those expansions, namely that the rotation vector $D_P H_0(P_0)$ is Diophantine. These expansions are constructed in section 4, and some technical estimates are proved in section 5.

When considering the Mather set for $\epsilon > 0$ one has two choices. The first one is, given a vector P_0 fixed, trying to compare the formal approximation of the Mather set with the Mather set itself. Unfortunately there are serious problem with this approach as, when $\epsilon > 0$, the Diophantine conditions may be destroyed. The approach we consider is the following: for $\epsilon > 0$ we construct a vector P_ϵ which keeps an approximate rotation vector fixed. Then, in section 6, we prove that the viscosity solution and its approximation at P_ϵ are close both in the supremum norm and its derivatives. To prove these estimates we use a technique similar to the one in [Gom03], which is based in first proving estimates along trajectories for finite time. Since ergodization times can be controlled in terms of the Diophantine properties of the rotation vector (see [BGW98], [DDG96], [Dum99]), we extend this estimate for all points using a-priori Lipschitz bounds for the viscosity solution and its approximation.

The idea of controlling a viscosity solution during a long time and then extending to nearby points has been used by other authors. For instance in the paper [Bes], a ergodization times techniques are used to study Hamilton-Jacobi equations perturbed by an elliptic operator. This problem is also studied, in a different setting, in [FS86a], and [FS86b]. We should note that our results are related to the ones in [BK87], as both imply the stability of Mather sets. However, in this last paper the problem addressed is the persistence of periodic orbits and their techniques are quite different, whereas in our paper we study the non-resonant case.

2. Discrete variational principles and integrability

In the variational formulation of classical mechanics, the trajectories $\mathbf{x}(t)$ of a system are minimizers, or at least critical points, of the action

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

with fixed endpoints, in which $L(x, v) : \mathbb{R}^{2n} \to \mathbb{R}$, the Lagrangian, is the difference between kinetic and potential energy. We assume that the Lagrangian is smooth, superlinear, strictly convex in v (that is, $D^2_{vv}L$ is bounded away from 0), $D^2_{vv}L$ is uniformly bounded, and that it is $\mathbb{Z}^n$-periodic in x, which means that for all $k \in \mathbb{Z}^n$,

$$L(x + k, v) = L(x, v).$$

This periodicity makes it convenient to look at the Lagrangian as a function from $\mathbb{T}^n \times \mathbb{R}^n$ to $\mathbb{R}$ ($\mathbb{T}^n$ is the n dimensional flat torus).

The minimizing trajectories are solutions to the Euler equation

$$-\frac{d}{dt}D_vL(\mathbf{x}, \dot{\mathbf{x}}) + D_xL(\mathbf{x}, \dot{\mathbf{x}}) = 0. \tag{2}$$

In applications, it is important to consider discrete versions of classical mechanics, for instance, for computational purposes. There are two alternatives to make this discretization, one is to discretize the

Euler-Lagrange equations (2), the other is to discretize the variational principle. They are not equivalent, and this last approach has several advantages. In fact, in the continuous setting there are certain invariant sets, the Mather sets, which are obtained using a variational principle. Using this discretization of the variational principle, one can construct Mather measures, see for instance [Gom02]. Furthermore, the map that is obtained this way has better geometrical properties as it preserves the symplectic structure.

The discretization of the variational problem can be done by means of the Euler method for the ODE

$$\dot{\mathbf{x}} = \mathbf{v}(t).$$

This yields the discrete dynamics

$$\mathbf{x}_{n+1} = \mathbf{x}_n + h\mathbf{v}_n,$$

in which h is the time-step.

The corresponding variational problem consists in minimizing the action

$$h\sum_{n=0}^{N} L(\mathbf{x}_n, \mathbf{v}_n),$$

among all choices of $\mathbf{v}_n$, $0 \leq n \leq N$, with fixed endpoints x_0 and x_N.

The analog of the Euler-Lagrange equations is

$$-\frac{D_vL(\mathbf{x}_{n+1}, \mathbf{v}_{n+1}) - D_vL(\mathbf{x}_n, \mathbf{v}_n)}{h} + D_xL(\mathbf{x}_{n+1}, \mathbf{v}_{n+1}) = 0. \tag{3}$$

In the continuous case, (2) can be written in a Hamiltonian form:

$$\begin{cases} \dot{\mathbf{p}} = D_xH(\mathbf{p}, \mathbf{x}) \\ \dot{\mathbf{x}} = -D_pH(\mathbf{p}, \mathbf{x}), \end{cases}$$

for $\mathbf{p} = -D_vL(\mathbf{x}, \dot{\mathbf{x}})$, and the Hamiltonian

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

Similarly, (3) can be written in the equivalent form [Gom02]

$$\begin{cases} \mathbf{p}_{n+1} - \mathbf{p}_n = hD_xH(\mathbf{p}_{n+1}, \mathbf{x}_n) \\ \mathbf{x}_{n+1} - \mathbf{x}_n = -hD_pH(\mathbf{p}_{n+1}, \mathbf{x}_n), \end{cases} \tag{4}$$

with

$$\mathbf{p}_{n+1} = -D_vL(\mathbf{x}_n, \mathbf{v}_n). \tag{5}$$

Note that the dynamics (4) is semi-explicit, that is, implicit in p and explicit in x. This may therefore constraint the size of h for which (4) defines a discrete flow, depending on bounds for the derivatives of H.

We consider the case in which the Lagrangian

$$L_\epsilon(x, v) : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{R},$$

is a perturbation of an integrable one $L_0(v)$. More precisely,

$$L_\epsilon = L_0(v) + \epsilon L_1(x, v), \tag{6}$$

in which ϵ is a small parameter. When $\epsilon = 0$, the dynamics (4) is very simple since v_n is constant. The Hamiltonian corresponding to (6) has an expansion of the form

$$H(p, x) = H_0(p) + \epsilon H_1(p, x) + \epsilon^2 H_2(p, x) + \cdots$$

This expression is a straightforward application of the implicit function theorem for the Legendre transform $p = -D_v L_\epsilon$.

As in the case of continuous flows (see [Arn89], [AKN97] or [Gol80]), one can use generating functions to change coordinates, and, in particular, there is a version of the Hamilton-Jacobi integrability for maps. In the next proposition we prove the main result on generating functions and change of coordinates, which serves to motivate our methods.

Theorem 2.1. *Let $(x, p) \in \mathbb{R}^{2n}$ be the original canonical coordinates and $(X, P) \in \mathbb{R}^{2n}$ be another coordinate system. Suppose there is a smooth function $S(x, P)$ such that*

$$p = D_x S(x, P) \qquad X = D_P S(x, P) \tag{7}$$

defines a global change of coordinates (this function is called a generating function). Additionally, assume that $D^2_{xP} S$ is non-singular, and suppose there is a smooth function $\overline{H}(P, X)$ such that, for h sufficiently small,

$$\begin{aligned} S(x, P) - S(\hat{x}, \hat{P}) + hH(D_x S(\hat{x}, \hat{P}), x) - \\ - hD_x S(\hat{x}, \hat{P}) D_p H(D_x S(\hat{x}, \hat{P}), x) + \\ + hD_P S(\hat{x}, \hat{P}) D_X \overline{H}(P, D_P S(\hat{x}, \hat{P})) = \\ = h\overline{H}(P, D_P S(\hat{x}, \hat{P})), \end{aligned} \tag{8}$$

in which

$$\hat{P} - P = hD_X \overline{H}(P, D_P S(\hat{x}, \hat{P})) \quad \hat{x} - x = -hD_p H(D_x S(\hat{x}, \hat{P}), x). \tag{9}$$

In the new coordinate system, the equations of motion (4) are

$$\begin{cases} \mathbf{X}_{n+1} - \mathbf{X}_n = -hD_P \overline{H}(\mathbf{P}_n, \mathbf{X}_{n+1}) \\ \mathbf{P}_{n+1} - \mathbf{P}_n = hD_X \overline{H}(\mathbf{P}_n, \mathbf{X}_{n+1}) \end{cases} \tag{10}$$

In particular, if $\overline{H}$ does not depend on X, these equations simplify to

$$
\begin{cases} \mathbf{X}_{n+1} - \mathbf{X}_n = -hD_P\overline{H}(\mathbf{P}_n) \\ \mathbf{P}_{n+1} - \mathbf{P}_n = 0. \end{cases} \tag{11}
$$

REMARK. 1. When we perform this change of coordinates we obtain a new dynamics that is semi-explicit as (4), but this time is implicit in X and explicit in P.

REMARK. 2. In the continuous case, given a Hamiltonian H and the generating function S, the new Hamiltonian is fully determined since $H(p,x) = \overline{H}(P,X)$. However, in our case, that is not immediate since (8) is in fact a partial differential equation for $\overline{H}$.

REMARK. 3. In this paper we will use mostly this theorem when $(x,p) \in \mathbb{T}^n \times \mathbb{R}^n$. In this case, S is identified with a function in the tangent space of the universal covering of $\mathbb{T}^n$.

PROOF. For simplicity, we will set $h = 1$ in the proof, by absorbing it into H (note however that we need h sufficiently small in order for the implicit expressions to be defined and smooth). Define $\hat{S} = S(\hat{x}, \hat{P})$, and we use the convention $D_x\hat{S} = (D_xS)(\hat{x}, \hat{P})$, as well as $D_P\hat{S} = (D_PS)(\hat{x}, \hat{P})$, to simplify the notation. By differentiating (8) with respect to x we obtain,

$$
\begin{aligned}
&D_xS - D_x\hat{S}D_x\hat{x} - D_P\hat{S}D_x\hat{P} + D_pHD_x\left(D_x\hat{S}\right) + D_xH - D_x\left(D_x\hat{S}\right)D_pH- \\
&- D_x\hat{S}D_x\left(D_pH\right) + D_x\left(D_P\hat{S}\right)D_X\overline{H} + D_P\hat{S}D_x\left(D_X\overline{H}\right) = \\
&\qquad = D_X\overline{H}D_x\left(D_P\hat{S}\right).
\end{aligned}
$$

Then, by canceling some terms we have

$$
\begin{aligned}
&D_xS - D_x\hat{S}D_x\hat{x} - D_P\hat{S}D_x\hat{P} + D_xH- \\
&\qquad - D_x\hat{S}D_x\left(D_pH\right) + D_P\hat{S}D_x\left(D_X\overline{H}\right) = 0.
\end{aligned}
$$

Observe that

$$
D_x\hat{x} = I - D_x(D_pH),
$$

and

$$
D_P\hat{S}D_x\left(-\hat{P} + D_X\overline{H}\right) = 0.
$$

Therefore

$$
D_xS - D_x\hat{S} + D_xH = 0. \tag{12}
$$

Differentiating (8) with respect to P we get

$$D_P S - D_x\hat{S} D_P\hat{x} - D_P\hat{S} D_P\hat{P} + D_p H D_P\left(D_x\hat{S}\right) - D_P\left(D_x\hat{S}\right) D_p H -$$
$$- D_x\hat{S} D_P\left(D_p H\right) + D_P\left(D_P\hat{S}\right) D_X\overline{H} + D_P\hat{S} D_P\left(D_X\overline{H}\right) =$$
$$= D_P\overline{H} + D_X\overline{H} D_P\left(D_P\hat{S}\right).$$

By canceling we get

$$D_P S - D_x\hat{S} D_P\hat{x} - D_P\hat{S} D_P\hat{P} -$$
$$- D_x\hat{S} D_P\left(D_p H\right) + D_P\hat{S} D_P\left(D_X\overline{H}\right) = D_P\overline{H}.$$

Note that

$$D_P\hat{x} = -D_P\left(D_p H\right),$$

therefore

$$D_P S - D_P\hat{S} D_P\hat{P} + D_P\hat{S} D_P\left(D_X\overline{H}\right) = D_P\overline{H}.$$

Since

$$D_P\hat{P} = I + D_P\left(D_X\overline{H}\right)$$

we have

$$D_P S - D_P\hat{S} = D_P\overline{H}. \tag{13}$$

Define $\hat{p} = D_x\hat{S}$, and $\hat{X} = D_P\hat{S}$. Now assume that $(\mathbf{x}_n, \mathbf{p}_n)$ are solutions to the dynamics (4). If we set $x = \mathbf{x}_n$ and $p = \mathbf{p}_n$, by the change of coordinates (7) we have, correspondingly, $X = \mathbf{X}_n$ and $P = \mathbf{P}_n$. Therefore, from (12) we get

$$\hat{p} = \mathbf{p}_{n+1}.$$

Thus, since $\hat{p} = D_x\hat{S}$ we have $\hat{x} = \mathbf{x}_{n+1}$ and $\hat{P} = \mathbf{P}_{n+1}$. This implies $\hat{X} = \mathbf{X}_{n+1}$, and so (13) reads

$$\mathbf{X}_n - \mathbf{X}_{n+1} = -D_P\overline{H}(\mathbf{P}_n, \mathbf{X}_{n+1}),$$

and we also have

$$\mathbf{P}_{n+1} - \mathbf{P}_n = D_X\overline{H}(\mathbf{P}_n, \mathbf{X}_{n+1}).$$

■

The previous theorem suggests that we should look for a generating function $S(x, P) = Px + u(x, P)$, periodic in x, and a new Hamiltonian $\overline{H}$ which only depends on P. If such a solution exist, the equations of motion reduce to (11) and we say that such a system is integrable. The function u should then satisfy the equation (1). As we will discuss later, see section 3, this equation will always admit a viscosity solution, which, however, may not be smooth.

3. Viscosity solutions

In general, equation (1) may not admit smooth solutions. However, see [Gom02], theorem 3.1, one can prove the existence of viscosity solutions, which are the correct notion of weak solution. In this section we review the main facts and prove some preliminary estimates.

For our purposes, a convenient definition of viscosity solution is the following: we say that a function u is a viscosity solution of (1) provided that it satisfies the following fixed point identity:

$$u(x, P) = h \min \sum_{j=0}^{N-1} [L(\mathbf{x}_j, \mathbf{v}_j) + P\mathbf{v}_j + \overline{H}(P)] + u(\mathbf{x}_N, P), \tag{14}$$

in which the minimum is taken over trajectories $(\mathbf{x}_n, \mathbf{v}_n)$, $0 \leq n \leq N$, with initial condition $\mathbf{x}_0 = x$, and $\mathbf{x}_{n+1} = \mathbf{x}_n + h\mathbf{v}_n$. The next proposition shows that a smooth viscosity solution is a solution of (1).

Proposition 3.1. *Suppose u is smooth, periodic in x and satisfies*

$$u(x, P) = \min_v h[L(x, v) + Pv + \overline{H}(P)] + u(x + hv, P). \tag{15}$$

Then, u solves (1).

Proof. Since u is smooth and periodic, and L grows superlinearly in v, there exists at least one optimal velocity v^* in (15). v^* satisfies

$$D_v L(x, v^*) = -P - D_x u(x + hv^*, P).$$

Using the Legendre transform

$$L(x, v^*) + (P + D_x u(x + hv^*, P))v^* = -H(P + D_x u(x + hv^*, P), x),$$

and so

$$v^* = -D_p H(P + D_x u(x + hv^*, P), x).$$

Therefore, with $\hat{x} = x + hv^* = x - hD_p H(P + D_x u(\hat{x}, P), x)$,

$$u(x, P) - u(\hat{x}, P) + hH(P + D_x u(\hat{x}, P), x) - \\ - hD_x u(\hat{x}, P) D_p H(P + D_x u(\hat{x}, P), x) = h\overline{H}(P).$$

■

The optimal trajectory $\mathbf{x}_n$ and the momentum $\mathbf{p}_n$, as given by (5), are solutions of (4) for $n \geq 0$.

Let us start by quoting an existence result whose prove is given in [Gom02]:

Theorem 3.2. *For each $P \in \mathbb{R}^n$ there exists a unique number $\overline{H}(P)$ and a family of solutions $u(x, P)$, periodic in x, that solves (1) in the viscosity sense. Furthermore, $\overline{H}(P)$ is convex in P, and $u(x, P)$ is Lipschitz in x.*

Motivated by the formal change of variables that was discussed in the previous section, we would like to relate this weak solution with the dynamics (4) - the next theorem makes this connection.

Theorem 3.3. *Let u be a viscosity solution of (1). Then:*

- For each $P \in \mathbb{R}^n$, there exists at least one subset of $\mathbb{T}^n \times \mathbb{R}^n$, called Mather set, which is invariant under the dynamics (4) and is contained in the graph $(x, p) = (x, P + D_x u(x))$.

- There exists a probability measure $\mu(x, p)$ on $\mathbb{T}^n \times \mathbb{R}^n$ (discrete Mather measure) invariant under (4) supported on this invariant set.

- This measure minimizes

$$\int L(x, v) + Pv d\nu, \tag{16}$$

with $v = -D_p H(\hat{p}, x)$, and $\hat{p} - p = h D_x H(\hat{p}, x)$, over all probability measures ν on $\mathbb{T}^n \times \mathbb{R}^n$ that satisfy

$$\int \phi(x + hv) - \phi(x) d\nu = 0,$$

for all continuous function $\phi : \mathbb{T}^n \to \mathbb{R}$. Furthermore

$$-\overline{H} = \int L(x, v) + Pv d\mu, \tag{17}$$

where $\overline{H}$ is the unique number for which (1) admits a periodic viscosity solution.

One of the main points in the previous theorem is that one can translate properties of viscosity solutions into properties of Mather sets or measures and vice-versa.

In the next proposition we discuss some of the properties of viscosity solutions, and its relations with the dynamics (4).

Proposition 3.4. *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 $n \geq 0$, the solution $(\mathbf{x}_n, \mathbf{p}_n)$ of (4) with initial conditions (x, p) belongs to $\mathcal{G}$.

A further result that we need, taken also from [Gom02], is a representation formula for $\overline{H}$ as a minimax. This is the discrete analog of the minimax formula for flows proved in [CIPP98].

Proposition 3.5. *For each $P \in \mathbb{R}^n$,*

$$\overline{H}(P) = \inf_{\varphi} \sup_{(x,v)} \left[\frac{\varphi(x) - \varphi(x + hv)}{h} - L(x, v) - Pv \right], \tag{18}$$

in which the infimum is taken over continuous periodic functions φ.

Suppose that $L_\epsilon(x,v) = L_0(v) + \epsilon L_1(x,v)$. The last estimates in this section show that, even without non-resonance conditions, the viscosity solution of (1) has good bounds for the semiconcavity and semiconvexity constants of u. These estimates should be seen as weak estimates for the second derivatives of u. Roughly speaking we have "$D^2_{xx}u = O(\sqrt{\epsilon})$" in the Mather set.

Proposition 3.6. *Suppose L_ϵ is as in (6), with $L_0(v)$ smooth, strictly convex, with bounded second derivative and coercive, and $L_1(x,v)$ is smooth, superlinear, strictly convex in v, periodic in x and with uniformly bounded second derivatives. Let u be a viscosity solution of (1). Then, for any x and y we have*

$$u(x+y) - 2u(x) + u(x-y) \le C\sqrt{\epsilon}|y|^2.$$

Proof. Let $(\mathbf{x}_n, \mathbf{v}_n)$, with $\mathbf{x}_{n+1} = \mathbf{x}_n + h\mathbf{v}_n$, $0 \le n \le N-1$ be an optimal trajectory with $\mathbf{x}_0 = x$, such that

$$u(x) = u(\mathbf{x}_N) + h\sum_{n=0}^{N-1}[L_0(\mathbf{v}_n) + \epsilon L_1(\mathbf{x}_n, \mathbf{v}_n) + P\mathbf{v}_n + \overline{H}(P)].$$

Then,

$$u(x \pm y) \le u(\mathbf{x}_N) + h\sum_{n=0}^{N-1}\left[L_0\left(\mathbf{v}_n \mp \frac{y}{Nh}\right) + \epsilon L_1\left(\mathbf{x}_n \pm \frac{N-n}{N}y, \mathbf{v}_n \mp \frac{y}{Nh}\right) + \right.$$
$$\left. + P\left(\mathbf{v}_n \mp \frac{y}{Nh}\right) + \overline{H}(P)\right].$$

Therefore,

$$\begin{aligned} u(x+y) - 2u(x) + u(x-y) \le& \\ \le h\sum_{n=0}^{N-1}&\left[L_0\left(\mathbf{v}_n - \frac{y}{Nh}\right) - 2L_0(\mathbf{v}_n) + L_0\left(\mathbf{v}_n + \frac{y}{Nh}\right)\right] + \\ + h\epsilon\sum_{n=0}^{N-1}&\left[L_1\left(\mathbf{x}_n + \frac{N-n}{N}y, \mathbf{v}_n - \frac{y}{Nh}\right) - 2L_1(\mathbf{x}_n, \mathbf{v}_n) + \right. \\ &\left. + L_1\left(\mathbf{x}_n - \frac{N-n}{N}y, \mathbf{v}_n + \frac{y}{Nh}\right)\right]. \end{aligned}$$

Note that, since $D^2 L_0$ is bounded,

$$L_0\left(\mathbf{v}_n - \frac{y}{Nh}\right) - 2L_0(\mathbf{v}_n) + L_0\left(\mathbf{v}_n + \frac{y}{Nh}\right) \le C\frac{|y|^2}{N^2h^2}.$$

Also

$$L_1\left(\mathbf{x}_n+\frac{N-n}{N}y,\mathbf{v}_n-\frac{y}{Nh}\right)-2L_1(\mathbf{x}_n,\mathbf{v}_n)+ \\ +L_1\left(\mathbf{x}_n-\frac{N-n}{N}y,\mathbf{v}_n+\frac{y}{Nh}\right)\leq C\left(1+\frac{1}{N^2h^2}\right)|y|^2.$$

Consequently, for h small,

$$u(x+y)-2u(x)+u(x-y)\leq C\left[\frac{1}{Nh}+\epsilon hN+\frac{\epsilon}{hN}\right]|y|^2.$$

By choosing $N=O\left(\frac{1}{h\sqrt{\epsilon}}\right)$ we obtain

$$u(x+y)-2u(x)+u(x-y)\leq C\sqrt{\epsilon}|y|^2.$$

■

Proposition 3.7. *Suppose L_ϵ is as in (6), with $L_0(v)$ smooth, strictly convex, with bounded second derivative and coercive, and $L_1(x,v)$ is smooth, superlinear, strictly convex in v, periodic in x and with uniformly bounded second derivatives. Let u be a viscosity solution of (1). Then, if x is in the Mather set and y is arbitrary, we have*

$$u(x+y)-2u(x)+u(x-y)\geq -C\sqrt{\epsilon}|y|^2.$$

PROOF. Since x belongs to the Mather set, there is a trajectory $(\mathbf{x}_n,\mathbf{v}_n)$, with $\mathbf{x}_{n+1}=\mathbf{x}_n+h\mathbf{v}_n$, $0\leq n\leq N-1$ with $\mathbf{x}_N=x$, such that

$$u(\mathbf{x}_0)=u(x)+h\sum_{n=0}^{N-1}[L_0(\mathbf{v}_n)+\epsilon L_1(\mathbf{x}_n,\mathbf{v}_n)+P\mathbf{v}_n+\overline{H}(P)].$$

Note that this identity implies that this trajectory achieves the minimum in (14). Thus,

$$u(\mathbf{x}_0)\leq u(x\pm y)+h\sum_{n=0}^{N-1}\left[L_0\left(\mathbf{v}_n\pm\frac{y}{Nh}\right)+\epsilon L_1\left(\mathbf{x}_n\pm\frac{n}{N}y,\mathbf{v}_n\pm\frac{y}{Nh}\right)+ \\ +P\left(\mathbf{v}_n\pm\frac{y}{Nh}\right)+\overline{H}(P)\right],$$

hence

$$
\begin{aligned}
u(x+y)-2u(x)+u(x-y) \geq & \\
\geq -h\sum_{n=0}^{N-1}\Big[L_0\Big(\mathbf{v}_n+\frac{y}{Nh}\Big)-2L_0(\mathbf{v}_n)+L_0\Big(\mathbf{v}_n-\frac{y}{Nh}\Big)\Big] - & \\
-h\epsilon\sum_{n=0}^{N-1}\Big[L_1\Big(\mathbf{x}_n+\frac{n}{N}y,\mathbf{v}_n+\frac{y}{Nh}\Big)-2L_1(\mathbf{x}_n,\mathbf{v}_n)+ & \\
+L_1\Big(\mathbf{x}_n-\frac{n}{N}y,\mathbf{v}_n-\frac{y}{Nh}\Big)\Big]. &
\end{aligned}
$$

Note that, since D^2L_0 is bounded,

$$
L_0\Big(\mathbf{v}_n+\frac{y}{Nh}\Big)-2L_0(\mathbf{v}_n)+L_0\Big(\mathbf{v}_n-\frac{y}{Nh}\Big)\leq C\frac{|y|^2}{h^2N^2}.
$$

Also

$$
\begin{aligned}
L_1\Big(\mathbf{x}_n+\frac{n}{N}y,\mathbf{v}_n+\frac{y}{Nh}\Big)-2L_1(\mathbf{x}_n,\mathbf{v}_n)+ & \\
+L_1\Big(\mathbf{x}_n-\frac{n}{N}y,\mathbf{v}_n-\frac{y}{Nh}\Big)\leq C\left[1+\frac{1}{N^2h^2}\right]|y|^2. &
\end{aligned}
$$

Therefore

$$
u(x+y)-2u(x)+u(x-y)\geq -C\left[\frac{1}{Nh}+h\epsilon N+\frac{\epsilon}{Nh}\right]|y|^2.
$$

By choosing $N=O\left(\frac{1}{h\sqrt{\epsilon}}\right)$ we obtain

$$
u(x+y)-2u(x)+u(x-y)\geq -C\sqrt{\epsilon}|y|^2.
$$

■

Theorem 3.8. *Suppose L_ϵ is as in (6), with $L_0(v)$ smooth, strictly convex, with bounded second derivative and coercive, and $L_1(x,v)$ is smooth, superlinear, strictly convex in v, periodic in x and with uniformly bounded second derivatives. Let u be a viscosity solution of (1). Then, if x is in the Mather set and y is arbitrary then*

$$
|u(x)-u(y)-D_xu(x)(y-x)|\leq C\sqrt{\epsilon}|x-y|^2.
$$

Proof. This follows from the proof of a similar theorem in [Gom02] by replacing the semiconcavity and (local) semiconvexity constants by $C\sqrt{\epsilon}$ which result from the two previous propositions. Note that the derivative in the formula makes sense since u is differentiable in the Mather set. ■

4. Formal Perturbation Theory

In this section we discuss formal perturbation methods using an analog of Linstead series. We outline the construction of formal expansions for the solutions $u^\epsilon(x,P)$ and $\overline{H}^\epsilon(P)$ of (1) that we denote, respectively, by $\tilde{u}_N^\epsilon(x,P)$ and $\tilde{H}_N^\epsilon(P)$. Those functions are given in power series of ϵ and $P-P_0$, and satisfy (1) up to order $O(\epsilon^N+|P-P_0|^N)$ in a neighborhood of $\epsilon=0$ and $P=P_0$. More precisely,

$$\begin{aligned}\tilde{u}_N^\epsilon(x,P)-\tilde{u}_N^\epsilon(\tilde{x},P)+hH(P+D_x\tilde{u}_N^\epsilon(\tilde{x},P),x)-\\ -hD_x\tilde{u}_N^\epsilon(\tilde{x},P)D_pH(P+D_x\tilde{u}_N^\epsilon(\tilde{x},P),x)=h\tilde{H}_N^\epsilon(P)\\ +O(\epsilon^N+|P-P_0|^N),\end{aligned} \tag{19}$$

in which the point $\tilde{x}$ is defined implicitly by

$$\tilde{x}-x=-hD_pH(P+D_x\tilde{u}_N^\epsilon(\tilde{x},P),x). \tag{20}$$

The main difficulty is that an approximate generating function may not yield a Hamiltonian $\overline{H}(P,X)$. However, by setting $S=Px+\tilde{u}_N^\epsilon(x,P)$, and performing the change of coordinates given by (7) the dynamics can still be written in a simpler Hamiltonian form, up to high order terms.

Proposition 4.1. *Suppose $\tilde{u}_N^\epsilon(x,P)$ is a solution of (19) for $P\in\mathbb{R}^n$ and $x\in\mathbb{T}^n$. Let*

$$S=Px+\tilde{u}_N^\epsilon(x,P),$$

and define new coordinates (X,P) by (7). Then

$$\begin{cases}\mathbf{X}_{n+1}-\mathbf{X}_n=-hD_P\tilde{H}_N^\epsilon(\mathbf{P}_n)+O(\epsilon^N+|\mathbf{P}_n-P_0|^{N-1})\\ \mathbf{P}_{n+1}-\mathbf{P}_n=O(\epsilon^N+|\mathbf{P}_n-P_0|^N).\end{cases} \tag{21}$$

Proof. For simplicity, as in the proof of theorem 2.1, we set $h=1$. By differentiating (19) with respect to x, and canceling the terms, as in theorem 2.1, we get

$$D_xS-D_x\hat{S}+D_xH=O(\epsilon^N+|P-P_0|^N), \tag{22}$$

in which $\hat{S}=S(\hat{x},P)$. Differentiating now with respect to P, and simplifying, we obtain

$$D_PS-D_P\hat{S}=D_P\tilde{H}_N^\epsilon(P)+O(\epsilon^N+|P-P_0|^{N-1}). \tag{23}$$

Define $\hat{p}=D_x\hat{S}$ and $\hat{X}=D_P\hat{S}$. If we assume that $(\mathbf{x}_n,\mathbf{p}_n)$ are solutions to the dynamics (4), and set $x=\mathbf{x}_n$ and $p=\mathbf{p}_n$, in the new coordinates the corresponding points are $X=\mathbf{X}_n$ and $P=\mathbf{P}_n$. Then, from (22), we get

$$\hat{p}=\mathbf{p}_{n+1}+O(\epsilon^N+|\mathbf{P}_n-P_0|^N).$$

Therefore, the inverse function theorem implies

$$\mathbf{P}_{n+1} - \mathbf{P}_n = O(\epsilon^N + |\mathbf{P}_n - P_0|^N).$$

Then $\hat{X} = \mathbf{X}_{n+1} + O(\epsilon^N + |\mathbf{P}_n - P_0|^{N-1})$. So, equation (23) reads

$$\mathbf{X}_n - \mathbf{X}_{n+1} = D_P \tilde{H}_N^\epsilon(\mathbf{P}_n) + O(\epsilon^N + |\mathbf{P}_n - P_0|^{N-1}).$$

■

The Linstead method consists in constructing solutions of (1) by using an iterative procedure that yields an expansion $\tilde{u}_N^\epsilon$ of the solution u^ϵ and $\tilde{H}_N^\epsilon$ of the Hamiltonian $\overline{H}^\epsilon$, as a power series in ϵ and $(P - P_0)$. Then, $\tilde{u}_N^\epsilon$ and $\tilde{H}_N^\epsilon$ satisfy (19).

Of course, there are some conditions that have to be satisfied in order to construct the approximated solution. These can be expressed in terms of the Diophantine properties of the vector P_0.

We say that a vector $\omega \in \mathbb{R}^n$ is Diophantine if

(24) $$\forall k \in \mathbb{Z}^n \backslash \{0\},\ m \in \mathbb{Z},\ \ |\omega \cdot k - m| \geq \frac{C}{|k|^s}, \text{ for some } C, s > 0.$$

We will assume that the vector $\omega_0 = D_P H_0(P_0)$ is Diophantine.

We look for an expansion of the form

(25)
$$\begin{aligned}\tilde{u}_N^\epsilon(x, P) &= \epsilon v_1(x, P_0) + \epsilon(P - P_0) D_P v_1(x, P_0) + \epsilon^2 v_2(x, P_0) \\ &\quad + \frac{1}{2}\epsilon(P - P_0)^2 D^2_{PP} v_1(x, P_0) + \epsilon^2 (P - P_0) D_P v_2(x, P_0) + \dots \\ &= \sum_{j=1}^{N-1} \sum_{i=1}^{j} \frac{1}{(j-i)!} \epsilon^i (P - P_0)^{j-i} D^{j-i}_{Pj-i} v_i(x, P_0),\end{aligned}$$

with the notation that, for $N = 1$, $\tilde{u}_1^\epsilon(x, P) = 0$. Furthermore,

$$\begin{aligned}\tilde{H}_N^\epsilon(P) =&\tilde{H}_0(P_0) + \epsilon \tilde{H}_1(P_0) + (P - P_0) D_P \tilde{H}_0(P_0) + \\ &+ \epsilon^2 \tilde{H}_2(P_0) + \epsilon(P - P_0) D_P \tilde{H}_1(P_0) + \frac{(P - P_0)^2}{2} D^2_{PP} \tilde{H}_0(P_0) + \cdots\end{aligned}$$

We will try to choose the functions v_j in such a way that, formally,

$$u^\epsilon(x, P) - \tilde{u}_N^\epsilon(x, P) = O(\epsilon^N + |P - P_0|^N),$$

by matching powers ϵ and $(P - P_0)$ in both sides of (19). The first term arises from taking $\epsilon = 0$, and $P = P_0$ in (19). Then,

$$\tilde{H}_0(P_0) = H_0(P_0),$$

and the solution $\tilde{u}_0^\epsilon = 0$. The first order terms in ϵ yield

$$v_1(x, P_0) - v_1(x - h D_p H_0(P_0), P_0) + h H_1(P_0, x) = h \tilde{H}_1(P_0),$$

and this equation determines $v_1(x,P_0)$ and $\tilde{H}_1(P_0)$. Furthermore

$$D_P\tilde{H}_0(P_0) = D_P H_0(P_0).$$

The function $v_2(x,P_0)$ and $\tilde{H}_2(P_0)$ are determined by solving the equation

$$\begin{aligned} &v_2(x,P_0) - v_2(x - hD_pH_0(P_0),P_0)+ \\ &\quad + \frac{h}{2}(D_x v_1(x - hD_pH_0(P_0),P_0))^2 D^2_{pp}H_0(P_0)+ \\ &\quad + hD_pH_1(P_0)D_x v_1(x - hD_pH_0(P_0),P_0)+ \\ &\quad + hH_2(P_0,x) = h\tilde{H}_2(P_0). \end{aligned}$$

To obtain $D_P v_1(x,P_0)$ and $D_P\tilde{H}_1(P_0)$ we consider the equation

$$\begin{aligned} &D_P v_1(x,P_0) - D_P v_1(x - hD_pH_0(P_0),P_0)+ \\ &\quad + hD_pH_1(P_0,x) = hD_P\tilde{H}_1(P_0). \end{aligned}$$

In general, we will have to solve equations of the form

$$G_{\omega_0}u = f + \lambda,$$

in which the operator G is given by

$$G_{\omega_0}u = u(x) - u(x-\omega_0), \tag{26}$$

the function f can be computed in terms of functions that are already known, λ is the unique constant for which (26) has a solution, and $\omega_0 = D_P H_0(P_0)$.

This operator can be analyzed by using Fourier coefficients. Note that,

$$G_{w_0}e^{2\pi ikx} = e^{2\pi ikx}(1 - e^{2\pi ikw_0}).$$

Thus, if $\eta(x) = \sum_k \eta_k e^{2\pi ikx}$ and $u(x) = \sum_k u_k e^{2\pi ikx}$, the equation

$$G_{\omega_0}u(x) = \eta(x), \tag{27}$$

reduces formally to

$$u_k = \frac{\eta_k e^{-2\pi ikx}}{1 - e^{2\pi ik\omega_0}}.$$

If ω_0 is non-resonant, the equation in (27) can be solved formally in Fourier coefficients, since no denominator vanishes, except for $k = 0$. Moreover, in order for (27) to have a solution, we need $\eta_0 = 0$. So this implies

$$\lambda = \int f(x)dx. \tag{28}$$

Furthermore, under Diophantine conditions on ω_0, the series for the solution u converges, in appropriate function spaces, as long as η is smooth enough.

From (28) we have

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

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

and

$$\tilde{H}_2(P_0) = \int \frac{1}{2}(D_x v_1(x - hD_p H_0(P_0), P_0))^2 D^2_{pp} H_0(P_0)dx + \\ + \int D_p H_1(P_0) D_x v_1(x - hD_p H_0(P_0), P_0) + H_2(P_0, x)dx.$$

Therefore, by computing these expansions we obtain, formally, that

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

This identity will be made rigorous in the next section.

5. Estimates for the Effective Hamiltonian

In this section we prove that $\tilde{H}^\epsilon_N(P)$ is an asymptotic expansion to $\overline{H}_\epsilon(P)$, therefore proving rigorously some of the results from the previous section.

Proposition 5.1. *Let u^ϵ and $\overline{H}^\epsilon$ be solutions of (1). Assume that there is an approximate solution $\tilde{u}^\epsilon_N$ of u^ϵ and $\tilde{H}^\epsilon_N$ of $\overline{H}^\epsilon$ satisfying (19). Then,*

$$\overline{H}^\epsilon(P) \leq \tilde{H}^\epsilon_N(P) + O(\epsilon^N + |P - P_0|^N). \tag{29}$$

Proof. We set $h = 1$ for simplicity. The inf sup formula (18) implies that

$$\overline{H}^\epsilon(P) \leq \sup_{(x,v)} \left[\tilde{u}^\epsilon_N(x, P) - \tilde{u}^\epsilon_N(x + v, P) - L(x, v) - Pv\right].$$

So, the optimal point v^* is given by

$$D_x\tilde{u}^\epsilon_N(x + v^*, P) = -L_v(x, v^*) - P,$$

and so

$$L(x, v^*) = -H(P + D_x\tilde{u}^\epsilon_N(x + v^*, P), x) - (P + D_x\tilde{u}^\epsilon_N(x + v^*, P))v^*.$$

Therefore,

$$\sup_v \left[\tilde{u}_N^\epsilon(x,P) - \tilde{u}_N^\epsilon(x+v,P) - L(x,v) - Pv\right] =$$
$$= \tilde{u}_N^\epsilon(x,P) - \tilde{u}_N^\epsilon(\hat{x},P) + H(P + D_x\tilde{u}_N^\epsilon(\hat{x},P),x) -$$
$$- D_x\tilde{u}_N^\epsilon(\hat{x},P)D_pH(P + D_x\tilde{u}_N^\epsilon(\hat{x},P),x) =$$
$$= \tilde{H}_N^\epsilon(P) + O(\epsilon^N + |P-P_0|^N).$$

Thus by taking the supremum, we obtain the result. ■

Proposition 5.2. *Let u^ϵ and $\overline{H}^\epsilon$ be solutions of* (1)*. Assume that there is an approximate solution $\tilde{u}_N^\epsilon$ of u^ϵ and $\tilde{H}_N^\epsilon$ of $\overline{H}^\epsilon$ satisfying (19). Then,*

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

Remark. This proposition and the previous one, together, imply

$$\overline{H}^\epsilon(P) = \tilde{H}_\epsilon^N(P) + O(\varepsilon^N + |P-P_0|^N). \tag{30}$$

Proof. Suppose $\tilde{u}_N^\epsilon(x,P)$ and $\tilde{H}_N^\epsilon(P)$ solve (19). Using a compactness argument, we can construct a measure $\tilde{\mu}$ on $\mathbb{T}^n \times \mathbb{R}^n$ such that for all continuous function ϕ with compact support

$$\int \phi(x,v)d\tilde{\mu} = \lim_{M\to\infty} \frac{1}{M}\sum_{n=1}^{M} \phi(\tilde{\mathbf{x}}_n, \tilde{\mathbf{v}}_n),$$

in which $(\tilde{\mathbf{x}}_n, \tilde{\mathbf{v}}_n)$ are given by

$$\tilde{\mathbf{x}}_{n+1} - \tilde{\mathbf{x}}_n = h\tilde{\mathbf{v}}_n = -hD_pH(P + D_x\tilde{u}_N^\epsilon(\tilde{\mathbf{x}}_{n+1},P),\mathbf{x}_n),$$

and the limit is taken through an appropriate subsequence. Note that $\tilde{\mu}$ is a probability measure, and for all continuous function $\phi(x)$ we have

$$\int \phi(x+hv) - \phi(x)d\tilde{\mu} = 0,$$

and

$$v = -D_pH(P + D_x\tilde{u}_N^\epsilon(x+hv,P),x),$$

on the support of $\tilde{\mu}$. Note that

$$\tilde{u}_N^\epsilon(x,P) - \tilde{u}_N^\epsilon(x+hv,P) - L_0(v) - \epsilon L_1(x,v) - Pv = \tag{31}$$
$$= \tilde{H}_N^\epsilon(P) + O(\epsilon^N + |P-P_0|^N),$$

for $v = -D_pH(P + D_x\tilde{u}_N^\epsilon(x+hv,P),x)$, by an argument similar to the one in the previous proposition. Therefore, by integrating (31) with respect to $\tilde{\mu}$, we obtain

$$\int L_0(v) + \epsilon L_1(x,v) + Pvd\tilde{\mu} = -\tilde{H}_N^\epsilon(P) + O(\epsilon^N + |P-P_0|^N).$$

Then, equation (17) implies

$$\overline{H}^{\epsilon}(P) \geq \tilde{H}_N^{\epsilon}(P) + O(\epsilon^N + |P - P_0|^N).$$

■

Finally, we prove a very simple result which will be essential in the following section.

Lemma 5.3. *Let u^ϵ and $\overline{H}^\epsilon$ be solutions of (1). Assume that there is an approximate solution $\tilde{u}_N^\epsilon$ of u^ϵ and $\tilde{H}_N^\epsilon$ of $\overline{H}^\epsilon$ satisfying (19). Then there exists a point x_0 at which*

$$D_x u^\epsilon(x_0, P) = D_x \tilde{u}_N^\epsilon(x_0, P).$$

Proof. Since $u^\epsilon(x,P) - \tilde{u}_N^\epsilon(x,P)$ is a periodic semiconcave function of x, then it has a minimum at some point x_0. Thus, at x_0, the derivative of $u^\epsilon(x,P) - \tilde{u}_N^\epsilon(x,P)$ with respect to x, exits and is zero. ■

6. Uniform estimates

This section is dedicated to prove two main results. One is that the solution of (19) approximates uniformly the viscosity solution of (1). The other is that the derivatives of the approximate solution are uniformly close to the ones of the viscosity solution. Since the Mather set is supported on the graph $(x, v(x))$, with v given by

$$v(x) = -D_p H(P + D_x u^\epsilon(x + hv(x), P), x),$$

this implies stability of Mather sets.

These result should be though of as the discrete analogs to the ones in [Gom03] for the continuous problem. The main idea is that the approximated solution, built using the formal expansion, is very close to the viscosity solution along an optimal trajectory. Under non-resonance conditions, these trajectories get close to any point in the torus in finite time. Therefore, since u^ϵ and $\tilde{u}_N^\epsilon$ are Lipschitz functions, we can extend the estimate to every point. Then we bootstrap these estimates for estimates on the derivatives.

Theorem 6.1. *Suppose the rotation vector*

$$\omega_0 = D_P H_0(P_0)$$

satisfies the Diophantine property (24). Assume that ϵ is small enough. Then, for every M there exists a vector

$$P_\epsilon = P_0 + O(\epsilon)$$

and N such that

$$D_P \tilde{H}_N^\epsilon(P_\epsilon) = \omega_0.$$

Furthermore,

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

in which $u^\epsilon(x, P_\epsilon)$ *is any viscosity solution of*

$$u^\epsilon(x, P_\epsilon) - u^\epsilon(\hat{x}, P_\epsilon) + hH(P_\epsilon + D_x u^\epsilon(\hat{x}, P_\epsilon), x) - \\ - \quad hD_x u^\epsilon(\hat{x}, P_\epsilon) D_p H(P_\epsilon + D_x u^\epsilon(\hat{x}, P_\epsilon), x) = h\overline{H}^\epsilon(P_\epsilon), \tag{32}$$

with

$$\hat{x} - x = -hD_p H(P_\epsilon + D_x u^\epsilon(\hat{x}, P_\epsilon), x),$$

and $\tilde{u}^\epsilon_N(x, P_\epsilon)$, $\tilde{H}^\epsilon_N(P_\epsilon)$ *satisfy (19), to which an appropriate constant has been added for normalization.*

Remark. Note that this theorem is still valid even if there is no uniqueness for viscosity solution of (32).

Proof. Define P_ϵ by solving the equation

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

that is

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

in which the expansion in terms of $P - P_0$ and ϵ is taken up to order $N - 1$. Under the strict convexity assumption for $\overline{H}_0(P) = \tilde{H}_0(P) = H_0(P)$, we have

$$\det D^2_{PP} \tilde{H}_0(P_0) \neq 0,$$

and so, the implicit function theorem yields a unique solution of the form

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

for ϵ small enough, and

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

Define the new coordinates (P, X) by (7), that is

$$\begin{cases} p = P + D_x \tilde{u}^\epsilon_N(x, P) \\ X = x + D_P \tilde{u}^\epsilon_N(x, P). \end{cases} \tag{33}$$

To simplify the notation we denote $X = \phi(x, P)$. Let x_0 be the point given by Lemma 5.3. Let

$$(\mathbf{x}_0, \mathbf{p}_0) = (x_0, P_\epsilon + D_x u^\epsilon(x_0, P))$$

be the initial conditions for a trajectory $(\mathbf{x}_n, \mathbf{p}_n)$ of (4). In the new coordinates, we have

$$\mathbf{P}_0 = P_\epsilon.$$

Also the dynamics is transformed into

$$\begin{cases} \mathbf{X}_{n+1} - \mathbf{X}_n = -hD_P\tilde{H}_N^\epsilon(\mathbf{P}_n) + O(\epsilon^N + |\mathbf{P}_n - P_0|^{N-1}) \\ \mathbf{P}_{n+1} - \mathbf{P}_n = O(\epsilon^N + |\mathbf{P}_n - P_0|^N). \end{cases} \tag{34}$$

From this equation, it is clear that $\mathbf{P}_n$ stays close to P_ϵ for long times. The next lemma proves a quantitative estimate that reflects this idea.

Lemma 6.2.

$$\sup_{0 \le n \le \frac{1}{\epsilon^{N/4}}} |\mathbf{P}_n - P_\epsilon| \le O(\epsilon^{N/8}).$$

Proof. Note that for any $\delta > 0$, we have

$$\begin{aligned} |\mathbf{P}_{n+1} - P_\epsilon|^2 - |\mathbf{P}_n - P_\epsilon|^2 &= |\mathbf{P}_{n+1} + \mathbf{P}_n - 2P_\epsilon||\mathbf{P}_{n+1} - \mathbf{P}_n| \le \\ &\le \delta|\mathbf{P}_{n+1} + \mathbf{P}_n - 2P_\epsilon|^2 + \frac{1}{\delta}|\mathbf{P}_{n+1} - \mathbf{P}_n|^2 \le \\ &\le 2\delta\left[|\mathbf{P}_{n+1} - P_\epsilon|^2 + |\mathbf{P}_n - P_\epsilon|^2\right] + \frac{1}{\delta}|\mathbf{P}_{n+1} - \mathbf{P}_n|^2. \end{aligned}$$

Moreover,

$$|\mathbf{P}_{n+1} - \mathbf{P}_n|^2 \le C\epsilon^{2N} + C\epsilon^{2N}|\mathbf{P}_n - P_\epsilon|^2,$$

as long as

$$|\mathbf{P}_n - P_\epsilon|^{2N-2} \le C\epsilon^{2N}. \tag{35}$$

We will show that this inequality is always satisfied for large N, and for the range of values n that will be used in our estimates.

By choosing $\delta = \epsilon^N$ and setting

$$a_n = |\mathbf{P}_n - P_\epsilon|^2,$$

we have

$$a_{n+1} - a_n \le C\epsilon^N(a_{n+1} + a_n) + C\epsilon^N.$$

To obtain the final estimate, we need an auxiliary lemma:

Lemma 6.3. *Suppose a_n is a sequence such that $a_0 = 0$, and*

$$(1 - C\epsilon^N)a_{n+1} \le (1 + C\epsilon^N)a_n + C\epsilon^N.$$

Then, for all $0 \le n \le \frac{1}{3C\epsilon^N}$ and ϵ small enough, we have

$$a_n \le \sqrt{C}\epsilon^{N/2} n e^{2n\sqrt{C}\epsilon^{N/2}}.$$

Proof. Set $\alpha = C\epsilon^N$. We will proceed by induction over n. For $n = 0$, the estimate is clear. Therefore, we assume it holds for some n and we will prove it for $n + 1$. We have

$$a_{n+1} \le \frac{1+\alpha}{1-\alpha}a_n + \frac{\alpha}{1-\alpha} \le (1 + 3\alpha)a_n + 2\alpha,$$

for α sufficiently small. Then, by the induction hypothesis,

$$a_{n+1} \leq (1+3\alpha)\sqrt{\alpha}ne^{2\sqrt{\alpha}n}+2\alpha \leq \sqrt{\alpha}e^{2\sqrt{\alpha}(n+1)}+\sqrt{\alpha}(ne^{2\sqrt{\alpha}n}+2\sqrt{\alpha}ne^{2\sqrt{\alpha}n}).$$

Since $1 + 2\sqrt{\alpha} \leq e^{2\sqrt{\alpha}}$, we get

$$a_{n+1} \leq \sqrt{\alpha}(n+1)e^{2\sqrt{\alpha}(n+1)}.$$

■

We should note that both the proof and the result of the previous lemma are not sharp, however they are sufficient for our purposes. In fact, this previous lemma implies that

$$|\mathbf{P}_n - P_\epsilon| \leq C\epsilon^{N/8},$$

for all $0 \leq n \leq \frac{C}{\epsilon^{N/4}}$, and therefore (35) is also satisfied. ■

Observe that $X = \psi(x) \equiv \phi(x, P_\epsilon)$ is a diffeomorphism, for small ϵ. Let

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

Define $\tilde{\mathbf{X}}_n = \psi(\mathbf{x}_n)$. Then, we have

$$\begin{aligned}
&U(\tilde{\mathbf{X}}_{n+1}) - U(\tilde{\mathbf{X}}_n) = \\
&\quad = u^\epsilon(\mathbf{x}_{n+1}, P_\epsilon) - u^\epsilon(\mathbf{x}_n, P_\epsilon) - \tilde{u}_N^\epsilon(\mathbf{x}_{n+1}, P_\epsilon) + \tilde{u}_N^\epsilon(\mathbf{x}_n, P_\epsilon) = \\
&\quad = -h\overline{H}^\epsilon(P_\epsilon) + hH(P_\epsilon + D_x u^\epsilon(\mathbf{x}_{n+1}, P_\epsilon), \mathbf{x}_n) - \\
&\qquad - hD_x u^\epsilon(\mathbf{x}_{n+1}, P_\epsilon) D_p H(P_\epsilon + D_x u^\epsilon(\mathbf{x}_{n+1}, P_\epsilon), \mathbf{x}_n) + \\
&\qquad + h\tilde{H}_N^\epsilon(P_\epsilon) - hH(P_\epsilon + D_x \tilde{u}_N^\epsilon(\tilde{\mathbf{x}}_{n+1}, P_\epsilon), \mathbf{x}_n) + \\
&\qquad + hD_x \tilde{u}_N^\epsilon(\tilde{\mathbf{x}}_{n+1}, P_\epsilon) D_p H(P_\epsilon + D_x \tilde{u}_N^\epsilon(\tilde{\mathbf{x}}_{n+1}, P_\epsilon), \mathbf{x}_n) + \\
&\qquad + \tilde{u}_N^\epsilon(\tilde{\mathbf{x}}_{n+1}, P_\epsilon) - \tilde{u}_N^\epsilon(\mathbf{x}_{n+1}, P_\epsilon) + O(\epsilon^N),
\end{aligned}$$

in which

$$\tilde{\mathbf{x}}_{n+1} - \mathbf{x}_n = -hD_p H(P_\epsilon + D_x \tilde{u}_N^\epsilon(\tilde{\mathbf{x}}_{n+1}, P_\epsilon), \mathbf{x}_n).$$

Therefore

$$\begin{aligned}
&U(\tilde{\mathbf{X}}_{n+1}) - U(\tilde{\mathbf{X}}_n) \leq \\
&\quad \leq C|D_x u^\epsilon(\mathbf{x}_{n+1}, P_\epsilon) - D_x \tilde{u}_N^\epsilon(\mathbf{x}_{n+1}, P_\epsilon)| + C|\tilde{\mathbf{x}}_{n+1} - \mathbf{x}_{n+1}| + O(\epsilon^N).
\end{aligned}$$

Now, observe that

$$\mathbf{p}_{n+1} = \mathbf{P}_{n+1} + D_x \tilde{u}_N^\epsilon(\mathbf{x}_{n+1}, \mathbf{P}_{n+1}),$$

together with

$$\mathbf{p}_{n+1} = P_\epsilon + D_x u^\epsilon(\mathbf{x}_{n+1}, P_\epsilon),$$

yields

$$
\begin{aligned}
&|D_x u^\epsilon(\mathbf{x}_{n+1}, P_\epsilon) - D_x \tilde{u}_N^\epsilon(\mathbf{x}_{n+1}, P_\epsilon)| \leq \\
&\quad \leq |D_x u^\epsilon(\mathbf{x}_{n+1}, P_\epsilon) - D_x \tilde{u}_N^\epsilon(\mathbf{x}_{n+1}, \mathbf{P}_{n+1})| + \\
&\quad\quad + |D_x \tilde{u}_N^\epsilon(\mathbf{x}_{n+1}, \mathbf{P}_{n+1}) - D_x \tilde{u}_N^\epsilon(\mathbf{x}_{n+1}, P_\epsilon)| \leq \\
&\quad \leq C|\mathbf{P}_{n+1} - P_\epsilon|.
\end{aligned}
$$

Moreover, the equation

$$y - \mathbf{x}_n = -hD_pH(P + D_x\tilde{u}_N^\epsilon(y, P), \mathbf{x}_n)$$

defines y as smooth function of $\mathbf{x}_n$ and P. Therefore,

$$|\tilde{\mathbf{x}}_{n+1} - \mathbf{x}_{n+1}| \leq C|\mathbf{P}_{n+1} - P_\epsilon|.$$

Thus, by using Lemma 6.2 we have

$$U(\tilde{\mathbf{X}}_{n+1}) - U(\tilde{\mathbf{X}}_n) \leq O(\epsilon^{N/8}),$$

for all $0 \leq n \leq \frac{C}{\epsilon^{N/4}}$. We may add a constant to u^ϵ in such a way that $U(\tilde{X}_0) = 0$, and thus

$$\sup_{0 \leq n \leq \frac{C}{\epsilon^{N/16}}} U(\tilde{\mathbf{X}}_n) = O(\epsilon^{N/16}).$$

The Diophantine property implies that the map

$$\mathbf{Y}_{n+1} - \mathbf{Y}_n = -hD_PH_0(P_0)$$

has an ergodization time of order $O\left(\frac{C}{\delta^r}\right)$, for some exponent r depending on the Diophantine exponent s in (24). That is, given δ and any Y, there exists $0 \leq n \leq O\left(\frac{C}{\delta^r}\right)$ such that $|Y - \mathbf{Y}_n| \leq \delta$. Consider the map given by the Hamiltonian dynamics which, in the new coordinates for $0 \leq n \leq \frac{C}{\epsilon^{N/16}}$, reads

$$\mathbf{X}_{n+1} - \mathbf{X}_n = -hD_PH_0(P_0) + O(\epsilon^{N-1}).$$

Then given ϵ and any X, there is a $0 \leq n \leq \frac{C}{\epsilon^{rM}}$ such that

$$|\mathbf{X}_n - X| \leq \epsilon^M,$$

provided $M < \frac{N}{16r}$. Furthermore, we have

$$\mathbf{X}_n = \phi(\mathbf{x}_n, \mathbf{P}_n),$$

and

$$\tilde{\mathbf{X}}_n = \phi(\mathbf{x}_n, P_\epsilon).$$

So, for $0 \leq n \leq \frac{C}{\epsilon^{N/16}}$, Lemma 6.2 implies

$$|\mathbf{X}_n - \tilde{\mathbf{X}}_n| \leq O(\epsilon^{N/8}).$$

Consequently, the sequence $\tilde{\mathbf{X}}_n$ satisfies that, given ϵ and any X, there is $0 \leq n \leq \frac{C}{\epsilon^{rM}}$ such that

$$|\tilde{\mathbf{X}}_n - X| \leq C\epsilon^M.$$

Since U is a Lipschitz function, by choosing $\tilde{\mathbf{X}}_n$ as in the previous formula

$$|U(X)| \leq |U(X) - U(\tilde{\mathbf{X}}_n)| + |U(\tilde{\mathbf{X}}_n)| \leq C\epsilon^M.$$

The same estimate carries over to $u^\epsilon(x, P_\epsilon) - \tilde{u}^\epsilon_N(x, P_\epsilon)$, as ψ is a diffeomorphism. ■

Remark. One should observe that

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

implies

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

although this last estimate requires the existence of $\tilde{u}^\epsilon_N$.

Theorem 6.4. *Let $M > 0$ and u^ϵ, $\overline{H}^\epsilon$ be solutions of (1). Suppose $\omega_0 = D_pH_0(P_0)$ is Diophantine, ϵ is sufficiently small and there is an approximate solution $\tilde{u}^\epsilon_N$ of u^ϵ and $\tilde{H}^\epsilon_N$ of $\overline{H}^\epsilon$ satisfying (19) for N sufficiently large so that Theorem 6.1 holds. Then,*

$$\operatorname*{esssup}_x |D_x u^\epsilon(\hat{x}, P_\epsilon) - D_x\tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon)| \leq C\epsilon^{M/2},$$

in which $\hat{x}$ and $\tilde{x}$ are defined, respectively by (9) and (20).

Proof. Subtracting (1) to (19), and using (30), we have

$$\begin{aligned} O(\epsilon^N) =& \tilde{u}^\epsilon_N(x, P_\epsilon) - \tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon) - u^\epsilon(x, P_\epsilon) + u^\epsilon(\hat{x}, P_\epsilon) + \\ &+ hH(P_\epsilon + D_x\tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon), x) - hH(P_\epsilon + D_x u^\epsilon(\hat{x}, P_\epsilon), x) - \\ &- hD_x\tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon) D_pH(P_\epsilon + D_x\tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon), x) + \\ &+ hD_x u^\epsilon(\hat{x}, P_\epsilon) D_pH(P_\epsilon + D_x u^\epsilon(\hat{x}, P_\epsilon), x). \end{aligned}$$

Moreover, by strict convexity, we have

$$\begin{aligned} H(P_\epsilon + D_x\tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon), x) \geq& H(P_\epsilon + D_x u^\epsilon(\hat{x}, P_\epsilon), x) + \\ &+ D_pH(P_\epsilon + D_x u^\epsilon(\hat{x}, P_\epsilon), x)(D_x\tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon) - D_x u^\epsilon(\hat{x}, P_\epsilon)) + \\ &+ \gamma|D_x\tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon) - D_x u^\epsilon(\hat{x}, P_\epsilon)|^2. \end{aligned}$$

Therefore, we get

$$\begin{aligned} O(\epsilon^N) \geq& \tilde{u}^\epsilon_N(x, P_\epsilon) - \tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon) - u^\epsilon(x, P_\epsilon) + u^\epsilon(\hat{x}, P_\epsilon) + \\ &+ hD_x\tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon)\left[D_pH(P_\epsilon + D_x u^\epsilon(\hat{x}, P_\epsilon), x) - D_pH(P_\epsilon + D_x\tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon), x)\right] + \\ &+ \gamma|D_x\tilde{u}^\epsilon_N(\tilde{x}, P_\epsilon) - D_x u^\epsilon(\hat{x}, P_\epsilon)|^2. \end{aligned}$$

Since

$$h\left[D_pH(P_\epsilon + D_x u^\epsilon(\hat{x}, P_\epsilon), x) - D_pH(P_\epsilon + D_x\tilde{u}_N^\epsilon(\tilde{x}, P_\epsilon), x)\right] = \tilde{x} - \hat{x},$$

and

$$-\tilde{u}_N^\epsilon(\tilde{x}, P_\epsilon) - D_x\tilde{u}_N^\epsilon(\tilde{x}, P_\epsilon)(\hat{x} - \tilde{x}) \geq -\tilde{u}_N^\epsilon(\hat{x}, P_\epsilon) - C\epsilon|\tilde{x} - \hat{x}|^2,$$

with

$$|\tilde{x} - \hat{x}|^2 \leq C|D_x\tilde{u}_N^\epsilon(\tilde{x}, P_\epsilon) - D_x u^\epsilon(\hat{x}, P_\epsilon)|^2,$$

we obtain,

$$\begin{aligned} O(\epsilon^N) \geq\ & \tilde{u}_N^\epsilon(x, P_\epsilon) - \tilde{u}_N^\epsilon(\hat{x}, P_\epsilon) - u^\epsilon(x, P_\epsilon) + u^\epsilon(\hat{x}, P_\epsilon) + \\ & + \tilde{\gamma}|D_x\tilde{u}_N^\epsilon(\tilde{x}, P_\epsilon) - D_x u^\epsilon(\hat{x}, P_\epsilon)|^2, \end{aligned}$$

for some constant $\tilde{\gamma}$, as long as ϵ is small enough. Therefore, the previous theorem yields

$$|D_x\tilde{u}_N^\epsilon(\tilde{x}, P_\epsilon) - D_x u^\epsilon(\hat{x}, P_\epsilon)|^2 \leq O(\epsilon^M).$$

■

We should point out that this last theorem shows that the Mather sets can be approximated through a perturbative method. In fact, since the Mather set is supported in

$$(x, -D_pH(P_\epsilon + D_x u^\epsilon(\hat{x}, P_\epsilon), x),$$

the result implies that this graph is approximated uniformly by

$$(x, -D_pH(P_\epsilon + D_x\tilde{u}_N^\epsilon(\tilde{x}, P_\epsilon), x).$$

7. Approximate Mather measures

For integrable Hamiltonian systems, as discussed in Theorem 2.1, one can change coordinates by (7). In these new coordinates, the Lebesgue measure dX is invariant by the flow. Thus, by the change of coordinates formula, the measure

$$\det(I + D^2_{xP}u)dx$$

is invariant for the original dynamics. The objective of this section is to show that Mather measures can be approximated by

$$\det(I + D^2_{xP}\tilde{u}_N^\epsilon)dx.$$

As the Mather measure in general is singular, we do not assert that it has a density approximated by the above expression but we claim that the mass in small boxes are comparable.

Theorem 7.1. *Let $M > 0$. Assume $\omega_0 = D_p H_0(P_0)$ is Diophantine, and there exists functions $\tilde{u}_N^\epsilon$ and $\tilde{H}^\epsilon$ satisfying (19) for N large enough. Let μ_ϵ be any Mather measure corresponding to P_ϵ (see theorem 6.1) . Consider push-forward ν_ϵ of μ_ϵ by the map (33).*

Let k be an arbitrary integer and ϵ sufficiently small. Then, there exists a partition of $\mathbb{T}^n$ in boxes $\{B_\eta^+\}$ of size $\frac{1}{k}$ such that

$$\nu_\epsilon(B_\eta^+) \geq \frac{k^n}{(k+1)^n}|B_\eta^+|,$$

for all η.

Additionally, there exists another partition of $\mathbb{T}^n$ in boxes $\{B_\eta^-\}$ of size $\frac{1}{k}$ such that

$$\nu_\epsilon(B_\eta^-) \leq \frac{k^n}{(k-1)^n}|B_\eta^-|,$$

for all η.

In both cases, k can be taken polynomially large in ϵ.

Proof. To prove the first estimate, let k be an integer and consider a partition of $\mathbb{T}^n$ in boxes of size $\frac{1}{k+1}$ (for the partition we use semi-open boxes, in such a way that they are pairwise disjoint). Obviously, for some box B, we have $\nu_\epsilon(B) \geq |B|$. Now consider an additional partition of $\mathbb{T}^n$ in boxes $\{B_\eta^+\}$ of size $\frac{1}{k}$ such that the center of one of the boxes coincides with the center of B. As in Theorem 6.1, consider the ergodization time for the linear flow associated with a covering of $\mathbb{T}^n$ of radius smaller that $\frac{1}{2k^2}$. For ϵ sufficiently small and every η, there exists a time T_η, smaller than the ergodization time T, such that the image of B under the map (34) lies in the interior of the corresponding box B_η^+. Therefore, by the invariance of the Mather measure under this map, we have

$$\nu_\epsilon(B_\eta^+) \geq |B| = \frac{k^n}{(k+1)^n}|B_\eta^+|.$$

To justify that k can be taken polynomially large in ϵ, suppose $k = \epsilon^{-\alpha}$. The ergodization time for the linear flow corresponding to $\epsilon^{2\alpha}$ is of order $T = O(\epsilon^{-2\alpha r})$. For sufficiently large N and small ϵ, the image of the box B under time $T_\eta < T$ lies in the interior of B_η^+.

The proof of the second estimate is similar: one considers a cover of $\mathbb{T}^n$ in boxes of size $\frac{1}{k-1}$. One of these boxes B has $\nu_\epsilon(B) \leq |B|$. Consider an additional partition of $\mathbb{T}^n$ in boxes $\{B_\eta^-\}$ of size $\frac{1}{k}$ such that the center of one of the boxes coincides with the center of B. As before, we obtain

$$\nu_\epsilon(B_\eta^-) \le |B| = \frac{k^n}{(k-1)^n}|B_\eta^-|.$$

■

As a corollary to this theorem we can state:

Corollary 7.2. *Let $M > 0$, J be a positive integer and f be a Lipschitz function on $\mathbb{T}^n$. Assume $\omega_0 = D_p H_0(P_0)$ is Diophantine, and there exists functions $\tilde{u}_N^\epsilon$ and $\tilde{H}^\epsilon$ satisfying (19) for N large enough. Let μ_ϵ be any Mather measure corresponding to P_ϵ (see theorem 6.1). Consider push-forward ν_ϵ of μ_ϵ by the map (33). Then,*

$$\int_{\mathbb{T}^n} f(X)dX = \int_{\mathbb{T}^n} f(X)d\nu_\epsilon + O(\epsilon^J).$$

Proof. We can assume f to be a positive function. Consider a partition in boxes B_η^+ as in the previous theorem. Then we have

$$\begin{aligned}\int f d\nu_\epsilon &\ge \sum_\eta \nu_\epsilon(B_\eta^+) \min_{B_\eta^+} f \\ &\ge \frac{k^n}{(k+1)^n}\sum_\eta |B_\eta^+| \min_{B_\eta^+} f \\ &\ge O(\frac{1}{k}) + \frac{k^n}{(k+1)^n}\sum_\eta \int_{B_\eta^+} f,\end{aligned}$$

since f is Lipschitz. Therefore

$$\int f d\nu_\epsilon \ge O(\frac{1}{k}) + \frac{k^n}{(k+1)^n}\int_{\mathbb{T}^n} f \ge \int f + O(\epsilon^J),$$

taking $k = \epsilon^{-J}$, and using the fact that f is bounded.

The other inequality is similar, using the partition B_η^- given by the previous theorem. ■

From this corollary we conclude that for any Lipschitz function $f(x,p)$ and any positive J we can choose N large enough such that

$$\int_{\mathbb{T}^n} f(x, P_\epsilon + D_x u^\epsilon) d\mu_\epsilon = \int_{\mathbb{T}^n} f(x, P_\epsilon + D_x \tilde{u}_N^\epsilon) \det(I + D_{xP}^2 \tilde{u}_N^\epsilon) dx + O(\epsilon^J),$$

as $\epsilon \to 0$.

References

[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].

[Arn89] V. I. Arnold. *Mathematical methods of classical mechanics.* Springer-Verlag, New York, 1989. Translated from the 1974 Russian original by K. Vogtmann and A. Weinstein.

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

[BGW98] Jean Bourgain, François Golse, and Bernt Wennberg. On the distribution of free path lengths for the periodic Lorentz gas. *Comm. Math. Phys.*, 190(3):491–508, 1998.

[BK87] David Bernstein and Anatole Katok. Birkhoff periodic orbits for small perturbations of completely integrable Hamiltonian systems with convex Hamiltonians. *Invent. Math.*, 88(2):225–241, 1987.

[CIPP98] G. Contreras, R. Iturriaga, G. P. Paternain, and M. Paternain. Lagrangian graphs, minimizing measures and Mañé's critical values. *Geom. Funct. Anal.*, 8(5):788–809, 1998.

[DDG96] H. S. Dumas, L. Dumas, and F. Golse. On the mean free path for a periodic array of spherical obstacles. *J. Statist. Phys.*, 82(5-6):1385–1407, 1996.

[Dum99] H. S. Dumas. Filling rates for linear flow on the torus: recent progress and applications. In *Hamiltonian systems with three or more degrees of freedom (S'Agaró, 1995)*, volume 533 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 335–339. Kluwer Acad. Publ., Dordrecht, 1999.

[E99] Weinan E. Aubry-Mather theory and periodic solutions of the forced Burgers equation. *Comm. Pure Appl. Math.*, 52(7):811–828, 1999.

[EG01] L. C. Evans and D. Gomes. Effective Hamiltonians and averaging for Hamiltonian dynamics. I. *Arch. Ration. Mech. Anal.*, 157(1):1–33, 2001.

[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.

[FS86a] W. H. Fleming and P. E. Souganidis. Asymptotic series and the method of vanishing viscosity. *Indiana Univ. Math. J.*, 35(2):425–447, 1986.

[FS86b] W. H. Fleming and P. E. Souganidis. Erratum: "Asymptotic series and the method of vanishing viscosity". *Indiana Univ. Math. J.*, 35(4):925, 1986.

[Gol80] Herbert Goldstein. *Classical mechanics.* Addison-Wesley Publishing Co., Reading, Mass., second edition, 1980. Addison-Wesley Series in Physics.

[Gom02] D. Gomes. Aubry-Mather theory for multidimensional maps. *preprint*, 2002.

[Gom03] D. Gomes. Perturbation theory for viscosity solutions of Hamilton-Jacobi equations and stability of Aubry-Mather sets. *to appear SIAM J. Math. Analysis*, 2003.

[Mat79] John N. Mather. Area preserving twist homeomorphism of the annulus. *Comment. Math. Helv.*, 54(3):397–404, 1979.

[Mat81] John N. Mather. Invariant subsets for area preserving homeomorphisms of surfaces. In *Mathematical analysis and applications, Part B*, pages 531–562. Academic Press, New York, 1981.

[Mat82] John N. Mather. Existence of quasiperiodic orbits for twist homeomorphisms of the annulus. *Topology*, 21(4):457–467, 1982.

[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. Variational construction of orbits of twist diffeomorphisms. *J. Amer. Math. Soc.*, 4(2):207–263, 1991.