EFFECTIVE HAMILTONIANS AND AVERAGING FOR HAMILTONIAN DYNAMICS I

BY

L. C. Evans [1] and D. Gomes [2]
Department of Mathematics
University of California
Berkeley, CA 94720

Abstract. This paper, building upon ideas of Mather, Moser, Fathi, E and others, applies PDE methods to understand the structure of certain Hamiltonian flows. The main point is that the "cell" or "corrector" PDE, introduced and solved in a weak sense by Lions, Papanicolaou and Varadhan in their study of periodic homogenization for Hamilton–Jacobi equations, formally induces a canonical change of variables, in terms of which the dynamics are trivial. We investigate to what extent this observation can be made rigorous in the case that the Hamiltonian is strictly convex in the momenta, given that the relevant PDE does not usually in fact admit a smooth solution.

1. Introduction.

This is the first of a projected series of papers that develop PDE techniques to understand certain aspects of Hamiltonian dynamics with many degrees of freedom.

1.1. Changing variables.

The basic issue is this. Given a smooth Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $H = H(p, x)$, we wish to examine the Hamiltonian flow

$$(1.1) \qquad \begin{cases} \dot{\mathbf{x}} = D_p H(\mathbf{p}, \mathbf{x}) \\ \dot{\mathbf{p}} = -D_x H(\mathbf{p}, \mathbf{x}) \end{cases}$$

under a canonical change of variables

$$(1.2) \qquad (p, x) \to (P, X),$$

[1] Supported in part by NSF Grant DMS-9424342.
[2] Supported by Praxis XXI-BD 5228/95.

where

$$\begin{cases} p = D_x u(P,x) \\ X = D_P u(P,x) \end{cases} \tag{1.3}$$

for a generating function $u : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $u = u(P,x)$. Here we write $D_x = (\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$ and $D_P = (\frac{\partial}{\partial P_1}, \dots, \frac{\partial}{\partial P_n})$. Assuming that we can find a smooth function u to solve the Hamilton–Jacobi type PDE

$$H(D_x u(P,x), x) = \bar{H}(P) \quad \text{in } \mathbb{R}^n, \tag{1.4}$$

and supposing as well that we can invert the relationships (1.3) to solve for P, X as smooth function of p, x, a calculation shows that we thereby transform (1.1) into the trivial dynamics

$$\begin{cases} \dot{\mathbf{X}} = D\bar{H}(\mathbf{P}) \\ \dot{\mathbf{P}} = 0. \end{cases} \tag{1.5}$$

In terms of mechanics, P is an "action" and X an "angle" or "rotation" variable, as for instance in Goldstein [Gd].

Of course we cannot really carry out this classical procedure in general, since the PDE (1.4) does not usually admit a smooth solution and, even if it does, the transformation (1.2), (1.3) is not usually globally defined. Only very special Hamiltonians are integrable in this sense.

1.2. Homogenization.

On the other hand, under some reasonable hypotheses we can in fact build appropriate *weak* solutions of (1.4), as demonstrated within another context in the classic-but-unpublished paper Lions–Papanicolaou–Varadhan [L-P-V]. These authors look at the initial value problem for the Hamilton–Jacobi PDE

$$\begin{cases} u^\varepsilon_t + H\left(D_x u^\varepsilon, \frac{x}{\varepsilon}\right) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u^\varepsilon = g & \text{on } \mathbb{R}^n \times \{t = 0\}, \end{cases} \tag{1.6}$$

under the primary assumption that the mapping $x \mapsto H(p,x)$ is $\mathbb{T}^n$-periodic, where $\mathbb{T}^n$ denotes the flat torus, that is, the unit cube in $\mathbb{R}^n$, with opposite faces identified. Consequently as $\varepsilon \to 0$, the nonlinearity in (1.6) is rapidly oscillating; and the problem is to understand the limiting behavior of the solutions u^ε. Lions et al. show under some mild additional hypotheses on the Hamiltonian that $u^\varepsilon \to u$, the limit function u solving a Hamilton–Jacobi PDE of the form

$$\begin{cases} u_t + \bar{H}(D_x u) = 0 & \text{in } \mathbb{R}^n \times (0,\infty) \\ u = g & \text{on } \mathbb{R}^n \times \{t = 0\}. \end{cases} \tag{1.7}$$

Here $\bar{H} : \mathbb{R}^n \to \mathbb{R}$, $\bar{H} = \bar{H}(P)$, is the *effective* (or *averaged*) *Hamiltonian*, and is built from H as follows.

1.3. How to construct $\bar{H}$.

First, consider for fixed $P \in \mathbb{R}^n$ the *cell* (or *corrector*) *problem*

$$\begin{cases} H(P + D_x v, x) = \lambda & \text{in } \mathbb{R}^n, \\ x \mapsto v \ \text{ is } \mathbb{T}^n\text{-periodic.} \end{cases}$$

As proved in Lions–Papanicolaou–Varadhan [L-P-V] (and recounted in [E2] and in Braides–Defranceschi [B-D, §16.2]), there exists a unique real number λ for which there exists a viscosity solution. We may then *define*

$$\bar{H}(P) := \lambda,$$

and so rewrite the foregoing as

$$\begin{cases} H(P + D_x v, x) = \bar{H}(P) & \text{in } \mathbb{R}^n, \\ x \mapsto v \ \text{ is } \mathbb{T}^n\text{-periodic.} \end{cases} \tag{1.8}$$

Once we set

$$u(P, x) := P \cdot x + v(P, x), \tag{1.9}$$

the PDE in (1.8) is just (1.4).

Remark. We pause here to draw attention to some simple observations relating the cell problem (1.8) and semiclassical approximations in quantum mechanics for periodic potentials. These comments are intended as further motivation.

Consider the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2}\Delta\psi + V\psi = E\psi \quad \text{in } \mathbb{R}^n, \tag{1.10}$$

where $\hbar$ is Planck's constant, $V : \mathbb{R}^n \to \mathbb{R}$ is a $\mathbb{T}^n$-periodic potential, and E is the energy corresponding to the eigenstate $\psi : \mathbb{R}^n \to \mathbb{C}$. A standard textbook procedure is to look for a solution having the *Bloch wave* form

$$\psi = e^{i\frac{P\cdot x}{\hbar}}\phi, \tag{1.11}$$

where $\phi : \mathbb{R}^n \to \mathbb{C}$ is $\mathbb{T}^n$-periodic. We further suppose ϕ to have the WKB–structure

$$\phi = ae^{i\frac{v}{\hbar}} \tag{1.12}$$

for periodic $a, v : \mathbb{R}^n \to \mathbb{R}$. Our substituting (1.11), (1.12) into (1.10) and taking real parts yields

$$\frac{1}{2}|P + D_x v|^2 + V(x) = E, \tag{1.13}$$

up to terms formally of size $O(\hbar)$.

Thus in the semiclassical limit $\hbar \to 0$, we *heuristically* obtain the cell problem (1.8) for the Hamiltonian $H(p, x) = \frac{1}{2}|p|^2 + V(x)$ and $\bar{H}(P) = E$. □

1.4. Questions, absolute minimizers.

The procedure outlined in §1.3 provides us with at least a theoretical construction of $\bar{H}$ and of a generating function u. Returning then to the comments in §1.1, we can now formulate these

Basic Questions. *To what extent can we employ $\bar{H}$ and u to understand the solutions of the Hamiltonian flow* (1.1)*? In particular, how is information about the dynamics "encoded" into $\bar{H}$?*

These are really hard issues, and to make at least a little progress we will need some additional hypotheses on both the Hamiltonian and the particular trajectories of the ODE we examine. Let us henceforth suppose that the mapping $p \mapsto H(p,x)$ is uniformly convex, in which case we can associate with H the *Lagrangian*

$$L(q,x) := \max_p (p \cdot q - H(p,x)).$$

Consider then a Lipschitz curve $\mathbf{x}(\cdot)$ which minimizes the associated action integral, meaning that

$$\int_0^T L(\dot{\mathbf{x}},\mathbf{x})\,dt \leq \int_0^T L(\dot{\mathbf{y}},\mathbf{y})\,dt \tag{1.14}$$

for each time $T > 0$ and each Lipschitz curve $\mathbf{y}(\cdot)$ with $\mathbf{x}(0) = \mathbf{y}(0)$, $\mathbf{x}(T) = \mathbf{y}(T)$. We call $\mathbf{x}(\cdot)$ a (one-sided) *absolute minimizer*. If we as usual define the *momentum*

$$\mathbf{p} := D_q L(\dot{\mathbf{x}},\mathbf{x}),$$

then $(\mathbf{x}(\cdot),\mathbf{p}(\cdot))$ satisfy Hamilton's ODE (1.1).

A discovery of Aubry [A], Mather [Mt 1-4], Fathi [F1–F3], Moser [Mo], E [EW2], etc., is that solutions of (1.1) corresponding to absolute minimizers are in a strong sense "better" than other solutions. Indeed, these authors have shown that the Hamiltonian dynamics are in some sense "integrable" for such special trajectories. The main goal of our work is to continue this analysis, with particular emphasis upon PDE methods (based upon viscosity solutions of (1.8)), applied to problems with many degrees of freedom.

1.5. Outline.

In §2 below we review the definition of the effective Hamiltonian H, introduce the corresponding effective Lagrangian $\bar{L}$, and recall the connections with the large time asymptotics of absolute minimizers $\mathbf{x}(\cdot)$.

We then rescale in time $\mathbf{x}(\cdot)$ and $\mathbf{p}(\cdot)$ in §3, and introduce certain Young measures $\{\nu_t\}_{t\geq 0}$ on phase space, which record the oscillations of the rescaled functions in asymptotic limits. These measures contain information about the Hamiltonian flow, and so our goal in subsequent sections is understanding their structure. In §4 we show that each $\nu = \nu_t$ is supported on the graph of the mapping $p = D_x u(P, x)$ and furthermore "stays away" from the discontinuities in $D_x u$.

In §5 we prove that u is well behaved on the support of σ, the projection of ν onto x-space. For this, we firstly derive the formal L^2-bound

$$\int_{\mathbb{T}^n} |D_x^2 u|^2 d\sigma \leq C \tag{1.15}$$

and then the L^∞-estimate

$$|D_x^2 u| \leq C \quad \sigma\text{-a.e..} \tag{1.16}$$

We rigorously establish some analogues of (1.15), (1.16), entailing difference quotients in the x-variables. As an application, we provide in §6 a new proof of Mather's theorem that ν is supported on an n-dimensional Lipschitz continuous graph.

Section 7 extends the techniques from §5 to establish what amounts to an L^2-estimate for the mixed second partial derivatives,

$$\int_{\mathbb{T}^n} |D_{xP}^2 u|^2 d\sigma \leq C D^2 \bar{H}(P). \tag{1.17}$$

More precisely, we prove a similar inequality involving difference quotients in the variable P. An application of this bound appears in §8, where we demonstrate the strict convexity of $\bar{H}$ in certain directions.

In §9 we draw some further deductions under the assumptions that $\bar{H}$ is differentiable at P and the components of $Q := D\bar{H}(P)$ are rationally independent.

A forthcoming companion paper [E-G2] addresses problems with time-dependent Hamiltonians. The primary new topics developed there include a weak interpretation of the "adiabatic invariance of the action" and a discussion of the Berry–Hannay geometric phase correction, computed in terms of effective Hamiltonians.

Our work is strongly related to some extremely interesting papers of Fathi [F1–F3], which develop his "weak KAM theory". We hope later to work out more clearly some of the connections with Fathi's discoveries.

Some other relevant papers include Mather [Mt1-4], Weinan E [EW1-2], Sobolevskii [So1-2], Mañé [Mn2-3], Jauslin–Kreiss–Moser [J-K-M], Iturriaga [I], Dias Carneiro [DC], Arisawa [Ar], etc. A good survey is Mather–Forni [M-F], and we have found Mañé's book [Mn1] to be very useful.

See Concordel [C1,C2], Chou–Duffin [C-D], Nussbaum[N], etc. for connections with nonlinear additive eigenvalue problems. Fathi [F4], Namah–Roquejoffre [N-F], Roquejoffre [R], Barles–Souganidis [B-S] and Fathi–Mather [F-M] discuss some related questions about large time asymptotics of solutions to Hamilton–Jacobi equations. Similar problems for stochastic homogenization have been studied by Rezakhanlou [Rz] and Souganidis [S].

There is also a large literature for time–dependent Hamiltonians with one degree of freedom. In this setting ordering properties for minimizing trajectories provide powerful tools unavailable in higher dimensions. See Mather–Forni [M-F], Aubry [A], Bangert [B2], etc. for more.

We are grateful to G. Barles, L. Barreira, M. Crandall, Weinan E, W. Oliva, D. Serre and A. Weinstein for interesting suggestions and for references.

2. Effective Hamiltonians and Lagrangians.

2.1. The Hamiltonian and Lagrangian.

As in the introduction, $\mathbb{T}^n$ denotes the standard flat torus.

Hypotheses on the Hamiltonian. Assume now that the given, smooth Hamiltonian $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $H = H(p, x)$, satisfies these conditions:

(i) *periodicity*:

$$\begin{cases} \text{For each } p \in \mathbb{R}^n, \text{ the mapping} \\ x \mapsto H(p,x) \text{ is } \mathbb{T}^n\text{-periodic.} \end{cases} \tag{2.1}$$

(ii) *strict convexity*:

$$\begin{cases} \text{There exist constants } \Gamma, \gamma > 0 \text{ such that} \\ \gamma|\xi|^2 \leq \sum_{i,j=1}^n \frac{\partial^2 H}{\partial p_i \partial p_j} \xi_i \xi_j \leq \Gamma |\xi|^2 \\ \text{for each } p, x, \xi \in \mathbb{R}^n. \end{cases} \tag{2.2}$$

The Lagrangian. We define the associated *Lagrangian* $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $L = L(q, x)$, by duality:

$$L(q,x) := \sup_p (p \cdot q - H(p,x)) \tag{2.3}$$

for $q \in \mathbb{R}^n$. In view of (2.1), (2.2) we see that L is smooth,

$$\begin{cases} \text{for each } q \in \mathbb{R}^n, \text{ the mapping} \\ x \mapsto L(q,x) \text{ is } \mathbb{T}^n\text{-periodic,} \end{cases} \tag{2.4}$$

$$
(2.5)\qquad \begin{cases} \text{there exist constants } \Gamma, \gamma > 0 \text{ such that} \\ \gamma|\xi|^2 \leq \sum_{i,j=1}^n \frac{\partial^2 L}{\partial q_i \partial q_j}(q,x)\xi_i\xi_j \leq \Gamma|\xi|^2 \\ \text{for all } q, x, \xi \in \mathbb{R}^n. \end{cases}
$$

We physically interpret x as position, p as momentum and q as velocity. The corresponding capital letters X, P, Q will likewise respectively denote position, momentum and velocity in new coordinates.

2.2. The effective Hamiltonian and Lagrangian.

As explained in the Introduction, we intend next to "average" H, following Lions, Papanicolaou, Varadhan [L-P-V]:

Theorem 2.1. (i) *For each $P \in \mathbb{R}^n$ there exists a unique real number, denoted $\bar{H}(P)$, such that the* cell problem

$$
(2.6)\qquad H(P + D_x v, x) = \bar{H}(P) \quad \text{in } \mathbb{R}^n
$$

has a $\mathbb{T}^n$-periodic, Lipschitz continuous solution v.

(ii) *In addition, there exists a constant α such that*

$$
(2.7)\qquad D_x^2 v \leq \alpha I \quad \text{in } \mathbb{R}^n
$$

in the distribution sense.

We call the function

$$
(2.8)\qquad \bar{H} : \mathbb{R}^n \to \mathbb{R}
$$

so defined the *effective* or *averaged Hamiltonian.*

Remarks. (i) We understand v to solve (2.6) in the sense of viscosity solutions. This means that if $\phi = \phi(x)$ is a smooth function and

$$
(2.9)\qquad \begin{cases} v - \phi \text{ has a maximum (minimum) at} \\ \text{a point } x_0 \in \mathbb{R}^n, \text{ then} \\ H(P + D\phi(x_0), x_0) \leq \bar{H}(P) \ (\geq \bar{H}(P)). \end{cases}
$$

We will in fact mostly need only that v is differentiable a. e. with respect to n-dimensional Lebesgue measure, and that v solves the PDE (2.6) at any point of differentiability.

(ii) The inequality (2.7) means that

$$
(2.10)\qquad \begin{cases} \text{the function } \bar{v}(x) := v(x) - \frac{\alpha}{2}|x|^2 \\ \text{is concave on } \mathbb{R}^n. \end{cases}
$$

(iii) If v is a solution of (2.6), we will hereafter often write

$$v = v(P, x) \qquad (P, x \in \mathbb{R}^n)$$

to emphasize the dependence on P. □

Given $\bar{H}$ as above, we define also the *effective Lagrangian*

$$\bar{L}(Q) := \sup_P (P \cdot Q - \bar{H}(P)) \tag{2.11}$$

for $Q \in \mathbb{R}^n$.

2.3. Properties of $\bar{H}$ and $\bar{L}$.

Proposition 2.2. *The mappings*

$$P \mapsto \bar{H}(P), \qquad Q \mapsto \bar{L}(Q)$$

are convex and real-valued. Furthermore, $\bar{H}$ and $\bar{L}$ are superlinear:

$$\lim_{|P|\to\infty} \frac{\bar{H}(P)}{|P|} = \lim_{|Q|\to\infty} \frac{\bar{L}(Q)}{|Q|} = +\infty. \tag{2.12}$$

Proof. 1. See Lions, Papanicolaou, Varadhan [L-P-V] (or [E2]) for a proof that $\bar{H}$ is convex. The convexity of $\bar{L}$ is immediate from (2.11).

2. In view of (2.2),

$$\bar{H}(P) \geq \alpha |P + D_x v|^2 - \beta \geq \alpha |P|^2 + 2\alpha P \cdot D_x v - \beta \text{ a.e.} \tag{2.13}$$

for appropriate constants $\alpha > 0$, $\beta \geq 0$. We integrate this inequality over $\mathbb{T}^n$ and recall v is periodic, to deduce

$$\bar{H}(P) \geq \alpha |P|^2 - \beta.$$

Thus $\bar{H}$ is superlinear, and in particular $\bar{L}(Q) < \infty$ for each Q. On the other hand, by construction $\bar{H}(P) < \infty$ for each P; whence the duality formula

$$\bar{H}(P) = \sup_Q (P \cdot Q - \bar{L}(Q))$$

implies $\bar{L}$ is superlinear. □

In later sections we will relate $\bar{H}, \bar{L}$ to appropriately rescaled minimizers of the action functionals, and for this will several times invoke the following results of Lions–Papanicolaou–Varadhan [L-P-V, §IV]. (See also E [EW1], Braides–Defranceschi [B-D, §16.2].)

Theorem 2.3. (i) *If* $\mathbf{X} : [0,T] \to \mathbb{R}^n$ *is a Lipschitz continuous curve and* $\mathbf{x}_\varepsilon(\cdot) \to \mathbf{X}(\cdot)$ *uniformly, then*

$$\int_0^T \bar{L}(\dot{\mathbf{X}})\,dt \leq \liminf \int_0^T L\left(\dot{\mathbf{x}}_\varepsilon, \frac{\mathbf{x}_\varepsilon}{\varepsilon}\right)\,dt. \tag{2.14}$$

(ii) *Define*

$$S_\varepsilon(x,y,t) := \inf\left\{\int_0^t L\left(\dot{\mathbf{x}}, \frac{\mathbf{x}}{\varepsilon}\right) ds \mid \mathbf{x}(t) = x,\ \mathbf{x}(0) = y\right\}, \tag{2.15}$$

for $x, y \in \mathbb{R}^n, t > 0$. *Then*

$$S_\varepsilon(x,y,t) \to t\bar{L}\left(\frac{x-y}{t}\right) \quad \text{as } \varepsilon \to 0, \tag{2.16}$$

uniformly on compact subsets of $\mathbb{R}^n \times \mathbb{R}^n \times (0,\infty)$.

3. Young measures.

Next we study the asymptotic behavior as $t \to \infty$ of certain curves that minimize the action.

3.1. Hamilton's ODE, rescalings.

Definition. A Lipschitz continuous curve $\mathbf{x} : [0,\infty) \to \mathbb{R}^n$ is called a (one-sided) *absolute minimizer* if

$$\int_0^T L(\dot{\mathbf{x}}, \mathbf{x})\,dt \leq \int_0^T L(\dot{\mathbf{y}}, \mathbf{y})\,dt \tag{3.1}$$

for each time $T > 0$ and each Lipschitz continuous curve $\mathbf{y} : [0,\infty) \to \mathbb{R}^n$ such that

$$\mathbf{x}(0) = \mathbf{y}(0),\ \mathbf{x}(T) = \mathbf{y}(T).$$

Given as above an absolutely minimizing curve $\mathbf{x}(\cdot)$, define the corresponding momentum

$$\mathbf{p}(t) := D_q L(\dot{\mathbf{x}}(t), \mathbf{x}(t)) \tag{3.2}$$

for $t \geq 0$. Then

$$\begin{cases} \dot{\mathbf{x}}(t) = D_p H(\mathbf{p}(t), \mathbf{x}(t)) \\ \dot{\mathbf{p}}(t) = -D_x H(\mathbf{p}(t), \mathbf{x}(t)) \end{cases} \tag{3.3}$$

for $t \geq 0$.

We wish to understand the pair $(\mathbf{x}(\cdot), \mathbf{p}(\cdot))$ for large times, and to this end introduce the *rescaled dynamics*

$$\begin{cases} \mathbf{x}_\varepsilon(t) := \varepsilon \mathbf{x}(t/\varepsilon), \ \mathbf{p}_\varepsilon(t) := \mathbf{p}(t/\varepsilon) \\ \mathbf{x}_\varepsilon(0) = \varepsilon \mathbf{x}(0), \ \mathbf{p}_\varepsilon(0) = \mathbf{p}(0). \end{cases}$$

It follows from (3.3) that

$$\begin{cases} \dot{\mathbf{x}}_\varepsilon(t) = D_p H\left(\mathbf{p}_\varepsilon(t), \frac{\mathbf{x}_\varepsilon(t)}{\varepsilon}\right) \\ \dot{\mathbf{p}}_\varepsilon(t) = -\frac{1}{\varepsilon} D_x H\left(\mathbf{p}_\varepsilon(t), \frac{\mathbf{x}_\varepsilon(t)}{\varepsilon}\right) \end{cases} \tag{3.4}$$

for $t \geq 0$.

Remark. Since $\frac{d}{dt} H\left(\mathbf{p}_\varepsilon(t), \frac{\mathbf{x}_\varepsilon(t)}{\varepsilon}\right) = 0$, we have $\sup_{t \geq 0} H\left(\mathbf{p}_\varepsilon(t), \frac{\mathbf{x}_\varepsilon(t)}{\varepsilon}\right) \leq C$ for some constant C, independent of ε. But $H(p, x) \geq \frac{\gamma}{2}|p|^2 - C$, and so

$$\sup_{t \geq 0} \{|\mathbf{p}_\varepsilon(t)|, |\dot{\mathbf{x}}_\varepsilon(t)|\} < \infty. \tag{3.5}$$

□

3.2. Recording oscillations.

We expect the functions $\mathbf{p}_\varepsilon(\cdot)$ and $\frac{\mathbf{x}_\varepsilon(\cdot)}{\varepsilon} (\text{mod } \mathbb{T}^n)$ to oscillate as $\varepsilon \to 0$, and so introduce measures on phase space to record these motions. Invoking for instance the methods from §1.E of [E1], we have

Proposition 3.1. *There exists a sequence $\varepsilon_k \to 0$ and for a.e. $t > 0$ a Radon probability measure ν_t on $\mathbb{R}^n \times \mathbb{T}^n$ such that*

$$\Phi\left(\mathbf{p}_{\varepsilon_k}(t), \frac{\mathbf{x}_{\varepsilon_k}(t)}{\varepsilon_k}\right) \rightharpoonup \bar{\Phi}(t) := \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} \Phi(p, x) \, d\nu_t(p, x) \tag{3.6}$$

for each bounded, continuous function

$$\Phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \ \Phi = \Phi(p, x),$$

such that $x \mapsto \Phi(p, x)$ is $\mathbb{T}^n$-periodic.

We call $\{\nu_t\}_{t \geq 0}$ *Young measures* associated with the dynamics (3.4).

Remark. The limit (3.6) means

$$\int_0^T \Phi\left(\mathbf{p}_{\varepsilon_k}, \frac{\mathbf{x}_{\varepsilon_k}}{\varepsilon_k}\right) \zeta \, dt \to \int_0^T \bar{\Phi} \zeta \, dt \tag{3.7}$$

for each $T > 0$ and each smooth function $\zeta : [0, T] \to \mathbb{R}$. □

Lemma 3.2. *The support of the measure ν_t is bounded, uniformly in t.*

This is clear from (3.5).

Lemma 3.3. *For each C^1 function Φ as above,*

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n}\{H,\Phi\}\,d\nu_t = 0 \tag{3.8}$$

for a.e. $t \geq 0$, where

$$\{H,\Phi\} := D_pH\cdot D_x\Phi - D_xH\cdot D_p\Phi \tag{3.9}$$

is the Poisson bracket.

The identity (3.8) means that the measure ν_t is invariant under the Hamiltonian flow (3.3).

Proof. We have

$$\begin{aligned}\frac{d}{dt}\Phi\left(\mathbf{p}_\varepsilon,\frac{\mathbf{x}_\varepsilon}{\varepsilon}\right) &= D_p\Phi\cdot\dot{\mathbf{p}}_\varepsilon + D_x\Phi\cdot\frac{\dot{\mathbf{x}}_\varepsilon}{\varepsilon}\\ &= \frac{1}{\varepsilon}\{H,\Phi\}\end{aligned}$$

according to (3.4). Take $\zeta:[0,T]\to\mathbb{R}$ to be smooth, with compact support. Then

$$\int_0^T\{H,\Phi\}\left(\mathbf{p}_\varepsilon,\frac{\mathbf{x}_\varepsilon}{\varepsilon}\right)\zeta\,dt = -\int_0^T\varepsilon\dot{\zeta}\Phi\left(\mathbf{p}_\varepsilon,\frac{\mathbf{x}_\varepsilon}{\varepsilon}\right)\,dt.$$

Sending $\varepsilon=\varepsilon_k\to 0$, we deduce (3.8). ☐

3.3. Convergence of trajectories, the action vector.

From (3.5), we conclude that the curves $\{\mathbf{x}_\varepsilon(\cdot)\}_{\varepsilon>0}$ are uniformly Lipschitz continuous. Hence we may assume (passing if necessary to a further subsequence) that

$$\mathbf{x}_{\varepsilon_k}\to\mathbf{X} \tag{3.10}$$

uniformly on compact subsets of $[0,\infty)$, where $\mathbf{X}:[0,\infty)\to\mathbb{R}^n$ is Lipschitz continuous, $\mathbf{X}(0)=0$.

Lemma 3.4. *We have*

$$\dot{\mathbf{X}}(t) = \mathbf{Q}(t) \quad \textit{for a.e. } t \geq 0, \tag{3.11}$$

for

$$\mathbf{Q}(t) := \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} D_p H(p,x)\, d\nu_t. \tag{3.12}$$

Proof. The limit (3.10) implies

$$\dot{\mathbf{x}}_{\varepsilon_k} \rightharpoonup \dot{\mathbf{X}};$$

whence (3.11), (3.12) follow from (3.4). □

Theorem 3.5. (i) *For a.e. time* $t \geq 0$

$$\bar{L}(\mathbf{Q}(t)) = \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} L(D_p H(p,x), x)\, d\nu_t. \tag{3.13}$$

(ii) *Furthermore, there exists* $P \in \mathbb{R}^n$ *such that*

$$P \in \partial \bar{L}(\mathbf{Q}(t)), \ \mathbf{Q}(t) \in \partial \bar{H}(P) \tag{3.14}$$

for a.e. $t \geq 0$.

Recall that if $\Phi : \mathbb{R}^n \to \mathbb{R}$ is convex, we write $y \in \partial\Phi(x)$ to mean

$$\Phi(x) + y \cdot (z - x) \leq \Phi(z) \quad \text{for all } z \in \mathbb{R}^n.$$

Remarks. (i) The point is that P does not depend on t. We call P an *action vector* for the rescaled trajectories $\{\mathbf{x}_\varepsilon(\cdot)\}_{\varepsilon>0}$.

(ii) The second assertion above can be restated

$$\begin{cases} \dot{\mathbf{X}} \in \partial \bar{H}(\mathbf{P}) \\ \dot{\mathbf{P}} = 0 \end{cases} \quad \text{for a.e. } t \geq 0,$$

and this formulation should be compared with (1.5).

(iii) The existence of P is also a consequence of the Pontryagin Maximum Principle; cf. Clarke [Cl]. □

Proof. 1. Let $y_\varepsilon := \mathbf{x}_\varepsilon(0) = \varepsilon \mathbf{x}(0) \to 0$. According to Theorem 2.3

$$S_{\varepsilon_k}(x, y_{\varepsilon_k}, t) \to t\bar{L}\left(\frac{x}{t}\right) \qquad (x \in \mathbb{R}^n, \ t > 0), \tag{3.15}$$

uniformly on compact subsets. But

$$S_\varepsilon(x, y_\varepsilon, t) = \inf\left\{\int_0^t L\left(\dot{\mathbf{x}}, \frac{\mathbf{x}}{\varepsilon}\right) ds \mid \mathbf{x}(t) = x,\ \mathbf{x}(0) = y_\varepsilon\right\},$$

and so

$$S_\varepsilon(\mathbf{x}_\varepsilon(t), y_\varepsilon, t) = \int_0^t L\left(\dot{\mathbf{x}}_\varepsilon, \frac{\mathbf{x}_\varepsilon}{\varepsilon}\right) ds, \tag{3.16}$$

since the curve $\mathbf{x}_\varepsilon(\cdot)$ is an absolute minimizer.

2. From (3.10), (3.15) we see that

$$S_{\varepsilon_k}(\mathbf{x}_{\varepsilon_k}(t), y_{\varepsilon_k}, t) \to t\bar{L}\left(\frac{\mathbf{X}(t)}{t}\right). \tag{3.17}$$

But then (3.16) implies that

$$L\left(\dot{\mathbf{x}}_{\varepsilon_k}, \frac{\mathbf{x}_{\varepsilon_k}}{\varepsilon_k}\right) \rightharpoonup \frac{d}{dt}\left(t\bar{L}\left(\frac{\mathbf{X}}{t}\right)\right). \tag{3.18}$$

Now

$$\frac{d}{dt}\left(t\bar{L}\left(\frac{\mathbf{X}}{t}\right)\right) \in \bar{L}\left(\frac{\mathbf{X}}{t}\right) + \partial\bar{L}\left(\frac{\mathbf{X}}{t}\right)\left(\dot{\mathbf{X}} - \frac{\mathbf{X}}{t}\right) \leq \bar{L}(\dot{\mathbf{X}}), \tag{3.19}$$

by convexity. Consequently, since $\dot{\mathbf{x}}_\varepsilon = D_pH\left(\mathbf{p}_\varepsilon, \frac{\mathbf{x}_\varepsilon}{\varepsilon}\right)$, we deduce from (3.18) that

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} L(D_pH(p,x),x)\, d\nu_t \leq \bar{L}(\dot{\mathbf{X}}(t)) \tag{3.20}$$

for a.e. $t > 0$.

Conversely, Theorem 2.3 implies

$$\int_a^b \bar{L}(\dot{\mathbf{X}}(t))\, dt \leq \lim_{\varepsilon\to 0}\int_a^b L\left(\dot{\mathbf{x}}_\varepsilon, \frac{\mathbf{x}_\varepsilon}{\varepsilon}\right)\, dt = \int_a^b\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} L(D_pH, x)\, d\nu_t\, dt$$

for all $0 \leq a < b < \infty$ and so

$$\bar{L}(\dot{\mathbf{X}}(t)) \leq \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} L(D_pH(p,x),x)\, d\nu_t$$

for a.e. t. This and (3.20) establish (3.13).

3. In particular,

$$\frac{d}{dt}\left(t\bar{L}\left(\frac{\mathbf{X}(t)}{t}\right)\right) = \bar{L}(\dot{\mathbf{X}}(t)) = \bar{L}(\mathbf{Q}(t)) \quad \text{a.e.;}$$

and so

$$\begin{aligned}\frac{1}{T}\int_0^T \bar{L}(\mathbf{Q}(t))\,dt &= \frac{1}{T}\int_0^T \frac{d}{dt}\left(t\bar{L}\left(\frac{\mathbf{X}(t)}{t}\right)\right)\,dt \\ &= \bar{L}\left(\frac{\mathbf{X}(T)}{T}\right) \\ &= \bar{L}\left(\frac{1}{T}\int_0^T \mathbf{Q}(t)\,dt\right).\end{aligned} \tag{3.21}$$

This identity, valid for each time $T > 0$, implies that $\{\mathbf{Q}(t)\}_{t\geq 0}$ lies a supporting domain of $\bar{L}$. This means that

$$P \in \partial\bar{L}(\mathbf{Q}(t)) \quad \text{for a.e. time } t \geq 0 \tag{3.22}$$

for some vector $P \in \mathbb{R}^n$. Equivalently, $\mathbf{Q}(t) \in \partial\bar{H}(P)$.

To confirm (3.22), fix a time $T > 0$, write $\bar{Q} := \frac{1}{T}\int_0^T \mathbf{Q}(t)\,dt$, and take any $P \in \partial\bar{L}(\bar{Q})$. Then owing to (3.21) we have

$$\bar{L}(\mathbf{Q}(t)) = \bar{L}(\bar{Q}) + P\cdot(\mathbf{Q}(t) - \bar{Q})$$

for a.e. time $0 \leq t \leq T$. Thus $\mathbf{Q}(t)$ is a minimizer of the convex function $\bar{L}(Q) - \bar{L}(\bar{Q}) - P\cdot(Q-\bar{Q})$, and so $P \in \partial\bar{L}(\mathbf{Q}(t))$, for a.e. time $0 \leq t \leq T$. Taking a sequence of times $T_k \to \infty$ and passing if necessary to a subsequence, we obtain a vector P satisfying (3.22). □

4. Structure of minimizing measures.

We next fix one of the Young measures ν_t and hereafter write $\nu = \nu_t$. Our goal is to understand the form of this measure, and in particular to describe its support.

Our further deductions will be based entirely upon certain conclusions reached above. These are firstly that ν is a compactly supported Radon probability measure on $\mathbb{R}^n \times \mathbb{T}^n$, for which we define

$$Q := \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} D_pH(p,x)\,d\nu,$$

as in (3.12) above. In addition, we have

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} \{H,\Phi\}\,d\nu = 0 \tag{4.1}$$

for each C^1 function Φ that is $\mathbb{T}^n$-periodic, and furthermore

$$\bar{L}(Q) = \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} L(D_pH(p,x),x)\,d\nu. \tag{4.2}$$

These are, respectively, assertions (3.8) and (3.13) above.

Remarks. Our ν is therefore a *minimal measure* in the sense of Mather [Mt1], except that we work in phase space. The advantage is that the flow invariance condition (4.1) is fairly simple, and very useful, in the (p, x) variables. □

Notation. (i) We write $\mathcal{M} := \operatorname{spt}(\nu)$ and call $\mathcal{M}$ the *Aubry–Mather set.*

(ii) We denote by σ the *projection* of ν onto the x-variables. That is,

$$\sigma(E) := \nu(\mathbb{R}^n \times E)$$

for each Borel subset E of $\mathbb{T}^n$. □

Take now any $P \in \partial\bar{L}(Q)$ and let $v = v(P, x)$ be any viscosity solution of the corresponding cell problem

$$\begin{cases} H(P + D_x v, x) = \bar{H}(P) & \text{in } \mathbb{R}^n \\ x \mapsto v(P, x) \ \text{ is } \mathbb{T}^n\text{-periodic,} \end{cases} \tag{4.3}$$

satisfying the semiconcavity condition (2.7). We hereafter set

$$u(P, x) := P \cdot x + v(P, x).$$

4.1. Differentiability on the support of ν.

Theorem 4.1. (i) *The function u is differentiable in the variable x σ-a.e., and σ-a.e. point is a Lebesgue point for $D_x u$.*

(ii) *We have*

$$p = D_x u(P, x) \quad \nu\text{-a.e.}$$

(iii) *Furthermore,*

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} H(p, x)\, d\nu = \int_{\mathbb{T}^n} H(D_x u, x)\, d\sigma = \bar{H}(P); \tag{4.4}$$

and if $\bar{H}$ is differentiable at P,

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} D_p H(p, x)\, d\nu = \int_{\mathbb{T}^n} D_p H(D_x u, x)\, d\sigma = D\bar{H}(P).$$

Thus ν is supported on the graph $p = D_x u(P, x) = P + D_x v(P, x)$, which is single-valued σ-a.e. Also, the PDE (4.3) holds pointwise, σ-a.e.

Remarks. Formula (4.4) explicitly displays $\bar{H}$ as an average of H; but for this to be useful, we need to know more about the measure σ. We will later, in §9, discover a bit more about the structure of σ.

Observe also that from (4.4) we deduce

$$\bar{H}(P) = \bar{H}(\tilde{P}) \quad \text{if} \quad P, \tilde{P} \in \partial \bar{L}(Q). \tag{4.5}$$

Finally, compare assertion (ii) with the canonical change of variables (1.3). □

Proof. 1. To ease notation, we do not display the dependence of u on the variable P, and also write Du for $D_x u$.

Take η_ε to be a smooth, nonnegative, radial convolution kernel, supported in the ball $B(0,\varepsilon)$. Then set

$$u^\varepsilon := \eta_\varepsilon * u.$$

The strict convexity of H implies for all $p, q \in \mathbb{R}^n$ that

$$H(q,x) \geq H(p,x) + D_p H(p,x) \cdot (q-p) + \frac{\gamma}{2}|q-p|^2.$$

Take $q = Du(y)$, $p = Du^\varepsilon(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y) Du(y)\, dy$ in this expression, multiply by $\eta_\varepsilon(x-y)$, and then integrate with respect to y:

$$\begin{aligned} H(Du^\varepsilon(x), x) \leq & \int_{\mathbb{R}^n} \eta_\varepsilon(x-y) H(Du(y), x)\, dy \\ & - \frac{\gamma}{2} \int_{\mathbb{R}^n} \eta_\varepsilon(x-y) |Du(y) - Du^\varepsilon(x)|^2 dy. \end{aligned}$$

Since the PDE $H(D_x u, x) = \bar{H}(P)$ holds pointwise a.e., we conclude that

$$\beta_\varepsilon(x) + H(Du^\varepsilon(x), x) \leq \bar{H}(P) + C\varepsilon \tag{4.6}$$

for each $x \in \mathbb{T}^n$, where

$$\beta_\varepsilon(x) := \frac{\gamma}{2} \int_{\mathbb{R}^n} \eta_\varepsilon(x-y) |Du(y) - Du^\varepsilon(x)|^2\, dy. \tag{4.7}$$

2. Recalling again the strict convexity of H with respect to the variable p, we deduce

$$\begin{aligned} & \frac{\gamma}{2} \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} |Du^\varepsilon(x) - p|^2\, d\nu \\ & \leq \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} H(Du^\varepsilon(x), x) - H(p,x) - D_p H(p,x) \cdot (Du^\varepsilon(x) - p)\, d\nu. \end{aligned} \tag{4.8}$$

Now $Du^\varepsilon = P + Dv^\varepsilon$, where $v^\varepsilon = \eta_\varepsilon * v$ is periodic. Consequently

$$\int_{\mathbb{R}^n} \int_{\mathbb{T}^n} D_p H \cdot Dv^\varepsilon\, d\nu = 0,$$

according to (4.1). This observation and (4.6) imply

$$\tag{4.9} \begin{aligned} &\frac{\gamma}{2}\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} |Du^\varepsilon - p|^2\, d\nu + \int_{\mathbb{T}^n} \beta_\varepsilon\, d\sigma \\ &\qquad \le \bar{H}(P) - \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} H + D_pH\cdot(P-p)\, d\nu + C\varepsilon. \end{aligned}$$

Next, $P \in \partial \bar{L}(Q)$ implies

$$\bar{L}(Q) + \bar{H}(P) = P\cdot Q.$$

Furthermore

$$L(D_pH(p,x),x) + H(p,x) = D_pH(p,x)\cdot p.$$

Recalling that $Q = \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} D_pH\, d\nu$ and substituting into (4.9), we find

$$\tag{4.10} \begin{aligned} &\frac{\gamma}{2}\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} |Du^\varepsilon - p|^2\, d\nu + \int_{\mathbb{T}^n} \beta_\varepsilon\, d\sigma \\ &\qquad \le -\bar{L}(Q) + \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} L(D_pH,x)\, d\nu + C\varepsilon = C\varepsilon, \end{aligned}$$

according to (4.2).

3. Now send $\varepsilon \to 0$. Passing as necessary to a subsequence we deduce first from (4.10) that

$$\beta_\varepsilon \to 0 \qquad \sigma\text{-a.e.}$$

Thus σ-a.e. point x is a point of approximate continuity of Du, and Du is σ-measurable. Since $u = x\cdot P + v$ and v is semiconcave as a function of x (Theorem 2.1,(ii)), it follows that u is differentiable in x, σ-a.e. Thus

$$Du^\varepsilon \to Du$$

pointwise, σ-a.e., and so (4.10) in turn forces

$$p = Du(x) = P + Dv(x) \qquad \nu\text{-a.e.}$$

This proves assertion (ii), and (iii) follows then from the cell PDE. □

Remark. As a consequence of the foregoing proof, we have the identity

$$\tag{4.11} \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} D_pH(p,x)\cdot D_xv\ d\nu = \int_{\mathbb{T}^n} D_pH(D_xu,x)\cdot D_xv\ d\sigma = 0,$$

which we will need later. To confirm this, recall from above that

$$\int_{\mathbb{R}^n}\int_{\mathbb{T}^n} D_pH\cdot D_xv^\varepsilon\ d\nu = 0.$$

Since $D_xv^\varepsilon \to D_xv$ boundedly, ν-a.e., we can apply the Dominated Convergence Theorem. □

5. Derivative estimates in the variable x.

We devote this section to showing that our solution u of the cell problem is "smoother" on the support of σ than it may be at other points of $\mathbb{T}^n$. This is a sort of "partial regularity" assertion.

5.1. Formal L^2- and L^∞-estimates.

First of all, we provide for the reader some purely formal L^2 and L^∞ estimates for $D_x^2 u$ on the support of σ, calculations which provide motivation for the rigorous bounds obtained afterwards.

***L^2*-inequalities**. We assume for this that the generating function u is smooth, then differentiate the cell PDE twice with respect to x_i, and finally sum for $i = 1, \dots, n$:

$$\begin{aligned} H_{p_k p_l}(D_x u, x) u_{x_k x_i} u_{x_l x_i} &+ H_{p_k}(D_x u, x) u_{x_k x_i x_i} \\ &+ 2H_{p_k x_i}(D_x u, x) u_{x_k x_i} + H_{x_i x_i}(D_x u, x) = 0. \end{aligned}$$

The first term on the left is greater than or equal to $\gamma |D_x^2 u|^2$. Thus

$$\gamma \int_{\mathbb{T}^n} |D_x^2 u|^2 \, d\sigma + \int_{\mathbb{T}^n} D_p H \cdot D_x(\Delta_x u) \, d\sigma \leq C + C \int_{\mathbb{T}^n} |D_x^2 u| \, d\sigma.$$

Since $\Delta_x u = \Delta_x v$ is periodic, the second term on the left equals zero, according to (4.1). We consequently conclude

$$\int_{\mathbb{T}^n} |D_x^2 u|^2 \, d\sigma \leq C, \tag{5.1}$$

for some constant C depending only on H and P. □

***L^∞*-inequalities**. We can similarly differentiate the cell PDE twice in any unit direction ξ, to find

$$\begin{aligned} H_{p_k p_l}(D_x u, x) u_{x_k \xi} u_{x_l \xi} &+ H_{p_k}(D_x u, x) u_{x_k \xi \xi} \\ &+ 2H_{p_k \xi}(D_x u, x) u_{x_k \xi} + H_{\xi \xi}(D_x u, x) = 0, \end{aligned}$$

for $u_{\xi\xi} := \sum_{i,j=1}^n u_{x_i x_j} \xi_i \xi_j$. Take a nondecreasing, function $\Phi : \mathbb{R} \to \mathbb{R}$, and write $\phi := \Phi' \geq 0$. Multiply the above identity by $\phi(u_{\xi\xi})$, and integrate with respect to σ. After some simplifications, we find

$$\frac{\gamma}{2} \int_{\mathbb{T}^n} |D_x u_\xi|^2 \phi(u_{\xi\xi}) \, d\sigma + \int_{\mathbb{T}^n} D_p H \cdot D_x(\Phi(u_{\xi\xi})) \, d\sigma \leq C \int_{\mathbb{T}^n} \phi(u_{\xi\xi}) \, d\sigma.$$

Since $u_{\xi\xi} = v_{\xi\xi}$ is periodic, the second term on the left is zero. We select

$$\phi(z) = \begin{cases} 1 & \text{if } z \leq -\mu \\ 0 & \text{if } z > -\mu, \end{cases}$$

for a constant $\mu > 0$. Since $|D_x u_\xi|^2 \geq u_{\xi\xi}^2$, we conclude that $\sigma(\{u_{\xi\xi} \leq -\mu\}) = 0$ if μ is large enough. Because (2.10) provides the opposite estimate $u_{\xi\xi} \leq \alpha$, we thereby derive the formal bound

$$|u_{\xi\xi}| \leq C \quad \sigma\text{-a.e.,} \tag{5.2}$$

the constant C depending only upon known quantities. □

Remark. As the interested reader may wish to confirm, the foregoing derivations are especially transparent for the classical Hamiltonian

$$H(p, x) = \frac{1}{2}|p|^2 + V(x),$$

in which case the cell PDE (4.3) is the *eikonal equation*

$$\frac{1}{2}|D_x u|^2 + V(x) = \bar{H}(P)$$

and (4.1) corresponds to the *transport equation*

$$\operatorname{div}(\sigma D_x u) = 0.$$

A clear message is that these two PDE should be considered together as a pair, in accordance with formal semiclassical limits. (See the Remark in §1.3.) □

5.2. An L^2-estimate of difference quotients in x.

We now establish an analogue of estimate (5.1), with difference quotients replacing some of the derivatives.

Theorem 5.1. *There exists a constant C, depending only on H and P, such that*

$$\int_{\mathbb{T}^n} |D_x u(P, x+h) - D_x u(P, x)|^2 \, d\sigma \leq C|h|^2 \tag{5.3}$$

for $h \in \mathbb{R}^n$.

Remark. If $D_x u(P, x+h)$ is multivalued, we interpret (5.3) to mean

$$\int_{\mathbb{T}^n} |\boldsymbol{\xi} - D_x u|^2 \, d\sigma \leq C|h|^2 \tag{5.4}$$

for some σ-measurable selection $\boldsymbol{\xi} \in D_x u(P, \cdot + h)$. □

Proof. 1. To simplify notation we do not display the dependence of u on P, and just write Du for $D_x u$.

Fix $h \in \mathbb{R}^n$ and define the shifted function

$$\tilde{u}(\cdot) := u(\cdot + h).$$

Then

$$H(D\tilde{u}, x + h) = \bar{H}(P) \quad \text{ a.e. in } \mathbb{R}^n.$$

Mollifying as in the proof of Theorem 4.1, we have

$$H(D\tilde{u}^\varepsilon, x + h) \leq \bar{H}(P) + C\varepsilon \quad \text{ in } \mathbb{R}^n.$$

Therefore

$$H(D\tilde{u}^\varepsilon, x) - H(Du, x) \leq C\varepsilon + H(D\tilde{u}^\varepsilon, x) - H(D\tilde{u}^\varepsilon, x + h)$$

σ-a.e., and consequently

$$\begin{aligned} \frac{\gamma}{2} \int_{\mathbb{T}^n} |D\tilde{u}^\varepsilon - Du|^2 \, d\sigma &+ \int_{\mathbb{T}^n} D_p H(Du, x) \cdot (D\tilde{u}^\varepsilon - Du) \, d\sigma \\ &\leq C\varepsilon + \int_{\mathbb{T}^n} H(D\tilde{u}^\varepsilon, x) - H(D\tilde{u}^\varepsilon, x + h) \, d\sigma \\ &\leq C(\varepsilon + |h|^2) - \int_{\mathbb{T}^n} D_x H(D\tilde{u}^\varepsilon, x) \cdot h \, d\sigma. \end{aligned} \tag{5.5}$$

2. Since $D\tilde{u}^\varepsilon - Du = D\tilde{v}^\varepsilon - Dv$, the second term on the left hand side of (5.5) vanishes, in view of (4.1), (4.11). Therefore

$$\begin{aligned} \frac{\gamma}{2} \int_{\mathbb{T}^n} |D\tilde{u}^\varepsilon - Du|^2 \, d\sigma \leq C(\varepsilon + |h|^2) &- \int_{\mathbb{T}^n} D_x H(Du, x) \cdot h \, d\sigma \\ &+ C \int_{\mathbb{T}^n} |D\tilde{u}^\varepsilon - Du||h| \, d\sigma, \end{aligned}$$

and thus

$$\frac{\gamma}{4} \int_{\mathbb{T}^n} |D\tilde{u}^\varepsilon - Du|^2 \, d\sigma \leq C(\varepsilon + |h|^2) - \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} D_x H \cdot h \, d\nu.$$

However (4.1) implies the last term here is zero; whence

$$\int_{\mathbb{T}^n} |D\tilde{u}^\varepsilon - Du|^2 \, d\sigma \leq C(\varepsilon + |h|^2). \tag{5.6}$$

3. We send $\varepsilon \to 0$. Passing as necessary to a subsequence we have

$$D\tilde{u}^\varepsilon \rightharpoonup \boldsymbol{\xi} \;\text{ weakly in } L^2_\sigma$$

and

$$\int_{\mathbb{T}^n} |\boldsymbol{\xi} - Du|^2 \, d\sigma \leq C|h|^2. \tag{5.7}$$

4. To conclude, we must show

$$\boldsymbol{\xi} \in D\tilde{u} = Du(\cdot + h) \quad \sigma\text{-a.e.}, \tag{5.8}$$

which means that for σ-a.e. point x there exists a constant C such that

$$\tilde{u}(y) \leq \tilde{u}(x) + \boldsymbol{\xi} \cdot (y - x) + C|y - x|^2 \tag{5.9}$$

for all y. To confirm this, recall that $\tilde{u}$, and so also $\tilde{u}^\varepsilon$, are semiconcave:

$$\tilde{u}^\varepsilon(y) \leq \tilde{u}^\varepsilon(x) + D\tilde{u}^\varepsilon(x) \cdot (y - x) + C|y - x|^2$$

for all x, y. Take $g \in L^2_\sigma$, $g \geq 0$. Then fixing y and integrating the variable x with respect to σ, we find

$$0 \leq \int_{\mathbb{T}^n} (-\tilde{u}^\varepsilon(y) + \tilde{u}^\varepsilon(x) + D\tilde{u}^\varepsilon(x) \cdot (y - x) + C|y - x|^2) g(x) \, d\sigma(x).$$

Let $\varepsilon \to 0$ and note $\tilde{u}^\varepsilon \to \tilde{u}$ uniformly. We conclude

$$0 \leq \int_{\mathbb{T}^n} (-\tilde{u}(y) + \tilde{u}(y) + \boldsymbol{\xi} \cdot (y - x) + C|y - x|^2) g(x) \, d\sigma(x).$$

This inequality is true for all g as above; whence (5.7) holds for σ-a.e. point x and all y. □

5.3. L^∞-estimates of difference quotients in x.

We next refine the integration arguments above, to derive an L^∞ bound on second-order difference quotients. This will be a variant of the formal estimate (5.2) above.

Theorem 5.2. *There exists a constant C, depending only on H and P, such that*

$$|u(P, x + h) - 2u(P, x) + u(P, x - h)| \leq C|h|^2 \tag{5.10}$$

for all $h \in \mathbb{R}^n$ and each point $x \in \operatorname{spt}(\sigma)$.

Proof. 1. Take $h \neq 0$, and write

$$\tilde{u} = u(\cdot + h), \ \hat{u} = u(\cdot - h).$$

We as before consider the mollified functions $\tilde{u}^\varepsilon, \hat{u}^\varepsilon$, where we take

$$0 < \varepsilon \leq \eta |h|^2, \tag{5.11}$$

for small $\eta > 0$. As in the earlier proofs, we have

$$\begin{cases} H(D\tilde{u}^\varepsilon, x+h) \leq \bar{H}(P) + C\varepsilon, \\ H(D\hat{u}^\varepsilon, x-h) \leq \bar{H}(P) + C\varepsilon. \end{cases}$$

Therefore for σ-a.e. point x,

$$\begin{aligned} &H(D\tilde{u}^\varepsilon, x) - 2H(Du, x) + H(D\hat{u}^\varepsilon, x) \\ &\qquad \leq C\varepsilon + H(D\tilde{u}^\varepsilon, x) - H(D\tilde{u}^\varepsilon, x+h) + H(D\hat{u}^\varepsilon, x) - H(D\hat{u}^\varepsilon, x-h). \end{aligned}$$

Hence

$$\begin{aligned} &\frac{\gamma}{2}(|D\tilde{u}^\varepsilon - Du|^2 + |D\hat{u}^\varepsilon - Du|^2) + D_pH(Du, x) \cdot (D\tilde{u}^\varepsilon - 2Du + D\hat{u}^\varepsilon) \\ &\qquad \leq C(\varepsilon + |h|^2) + (D_xH(D\hat{u}^\varepsilon, x) - D_xH(D\tilde{u}^\varepsilon, x)) \cdot h, \end{aligned}$$

and consequently

$$\begin{aligned} &\frac{\gamma}{4}(|D\tilde{u}^\varepsilon - Du|^2 + |D\hat{u}^\varepsilon - Du|^2) \\ &\qquad + D_pH(Du, x) \cdot (D\tilde{u}^\varepsilon - 2Du + D\hat{u}^\varepsilon) \leq C(\varepsilon + |h|^2). \end{aligned}$$

2. Fix now a smooth, nondecreasing, function $\Phi : \mathbb{R} \to \mathbb{R}$, and write $\phi := \Phi' \geq 0$. Multiply the last inequality above by $\phi\left(\frac{\tilde{u}^\varepsilon - 2u + \hat{u}^\varepsilon}{|h|^2}\right)$, and integrate with respect to σ:

$$\begin{aligned} &\frac{\gamma}{4}\int_{\mathbb{T}^n}(|D\tilde{u}^\varepsilon - Du|^2 + |D\hat{u}^\varepsilon - Du|^2)\phi\left(\frac{\tilde{u}^\varepsilon - 2u + \hat{u}^\varepsilon}{|h|^2}\right) d\sigma \\ &+ \int_{\mathbb{T}^n} D_pH(Du, x) \cdot (D\tilde{u}^\varepsilon - 2Du + D\hat{u}^\varepsilon)\phi(\cdots)\, d\sigma \\ &\qquad \leq C(\varepsilon + |h|^2)\int_{\mathbb{T}^n}\phi(\cdots)\, d\sigma. \end{aligned} \tag{5.12}$$

Now the second term on the left hand side of (5.12) equals

$$|h|^2 \int_{\mathbb{R}^n}\int_{\mathbb{T}^n} D_pH(p, x) \cdot D_x\Phi\left(\frac{\tilde{u}^\varepsilon - 2u + \hat{u}^\varepsilon}{|h|^2}\right) d\nu \tag{5.13}$$

and thus is zero. (To see this, note from (4.1) that the expression (5.13) vanishes if we replace u by a mollified function u^δ. Let $\delta \to 0$, recalling the estimates in the proof of Theorem 4.1.)

So now dropping the above term from (5.12) and rewriting, we deduce

$$\begin{aligned}(5.14)\qquad &\int_{\mathbb{T}^n} |Du^\varepsilon(x+h) - Du^\varepsilon(x-h)|^2 \phi\left(\frac{u^\varepsilon(x+h) - 2u(x) + u^\varepsilon(x-h)}{|h|^2}\right) d\sigma \\ &\qquad \le C(\varepsilon + |h|^2) \int_{\mathbb{T}^n} \phi\left(\frac{u^\varepsilon(x+h) - 2u(x) + u^\varepsilon(x-h)}{|h|^2}\right) d\sigma.\end{aligned}$$

3. We confront now a technical problem, as (5.14) entails a mixture of first-order difference quotients for Du^ε and second-order difference quotients for u, u^ε. We can however relate these expressions, since u is semiconcave.

To see this, first of all define

$$(5.15)\qquad E_\varepsilon := \{x \in \operatorname{spt}(\sigma) \mid u^\varepsilon(x+h) - 2u(x) + u^\varepsilon(x-h) \le -\mu|h|^2\},$$

the large constant $\mu > 0$ to be fixed below. Now according to (2.10), the functions

$$(5.16)\qquad \bar{u}(x) := u(x) - \frac{\alpha}{2}|x|^2, \quad \bar{u}^\varepsilon(x) := u^\varepsilon(x) - \frac{\alpha}{2}|x|^2$$

are concave. Also a point $x \in \operatorname{spt}(\sigma)$ belongs to E_ε if and only if

$$(5.17)\qquad \bar{u}^\varepsilon(x+h) - 2\bar{u}(x) + \bar{u}^\varepsilon(x-h) \le -(\mu+\alpha)|h|^2.$$

Set

$$(5.18)\qquad f^\varepsilon(s) := \bar{u}^\varepsilon\left(x + s\frac{h}{|h|}\right) \qquad (-|h| \le s \le |h|).$$

Then f is concave, and

$$\begin{aligned}\bar{u}^\varepsilon(x+h) - 2\bar{u}^\varepsilon(x) + \bar{u}^\varepsilon(x-h) &= f^\varepsilon(|h|) - 2f^\varepsilon(0) + f^\varepsilon(-|h|) \\ &= \int_{-|h|}^{|h|} f^{\varepsilon''}(x)(|h| - |s|)\, ds \\ &\ge |h| \int_{-|h|}^{|h|} f^{\varepsilon''}(s)\, ds \qquad (\text{since } f^{\varepsilon''} \le 0) \\ &= |h|(f^{\varepsilon'}(|h|) - f^{\varepsilon'}(-|h|)) \\ &= (D\bar{u}^\varepsilon(x+h) - D\bar{u}^\varepsilon(x-h)) \cdot h.\end{aligned}$$

Consequently if $x \in E_\varepsilon$, this inequality and (5.17) together imply

$$2|\bar{u}^\varepsilon(x) - \bar{u}(x)| + |D\bar{u}^\varepsilon(x+h) - D\bar{u}^\varepsilon(x-h)||h| \ge (\mu+\alpha)|h|^2.$$

Now $|\bar{u}^\varepsilon(x) - \bar{u}(x)| \leq C\varepsilon$ on $\mathbb{T}^n$, since u is Lipschitz continuous. We may therefore take η in (5.11) small enough to deduce from the foregoing that

$$|D\bar{u}^\varepsilon(x+h) - D\bar{u}^\varepsilon(x-h)| \geq (\frac{\mu}{2} + \alpha)|h| \tag{5.19}$$

But then

$$|Du^\varepsilon(x+h) - Du^\varepsilon(x-h)| \geq (\frac{\mu}{2} - \alpha)|h| \tag{5.20}$$

4. Return now to (5.14). Taking $\mu > 2\alpha$ and

$$\phi(z) = \begin{cases} 1 & \text{if } z \leq -\mu \\ 0 & \text{if } z > -\mu, \end{cases}$$

we discover from (5.14) that

$$(\frac{\mu}{2} - \alpha)^2 |h|^2 \sigma(E_\varepsilon) \leq C(\varepsilon + |h|^2)\sigma(E_\varepsilon).$$

We fix μ so large that

$$(\frac{\mu}{2} - \alpha)^2 \geq C + 1,$$

to deduce

$$(|h|^2 - C\varepsilon)\sigma(E_\varepsilon) \leq 0.$$

Thus $\sigma(E_\varepsilon) = 0$ if η in (5.11) is small enough, and this means

$$u^\varepsilon(x+h) - 2u(x) + u^\varepsilon(x-h) \geq -\mu|h|^2$$

for σ-a.e. point x. Now let $\varepsilon \to 0$:

$$u(x+h) - 2u(x) + u(x-h) \geq -\mu|h|^2$$

σ-a.e. Since

$$u(x+h) - 2u(x) + u(x-h) \leq \alpha|h|^2$$

owing to the semiconcavity, we have

$$|u(x+h) - 2u(x) + u(x-h)| \leq C|h|^2$$

for σ-a.e. point x. As u is continuous, the same inequality obtains for all $x \in \operatorname{spt}(\sigma)$. $\square$

6. Application: Lipschitz estimates for the support of ν.

We next improve the second derivative bounds from the previous section, and then show as a simple consequence that $\operatorname{spt}(\nu)$ lies on a Lipschitz continuous graph.

Theorem 6.1. *(i) There exists a constant C, depending only on H and P, such that*

$$|u(P,y) - u(P,x) - D_x u(P,x) \cdot (y-x)| \leq C|x-y|^2 \tag{6.1}$$

for all $y \in \mathbb{T}^n$ and σ-a.e. point $x \in \mathbb{T}^n$.

(ii) Furthermore,

$$|D_x u(P,y) - D_x u(P,x)| \leq C|x-y| \tag{6.2}$$

for all $y \in \mathbb{T}^n$ and for σ-a.e. point $x \in \mathbb{T}^n$.

(iii) In fact, u is differentiable at each point $x \in \operatorname{spt}(\sigma)$, and estimates (6.1), (6.2) *hold for all $y \in \mathbb{T}^n, x \in \operatorname{spt}(\sigma)$.*

Remark. When $D_x u(P,y)$ is multivalued, (6.2) asserts

$$|\xi - D_x u(P,x)| \leq C|x-y|$$

for all $\xi \in D_x u(P,y)$. In particular, for multivalued $D_x u(P,y)$ we have the estimate

$$\operatorname{diam}(D_x u(P,y)) \leq C \operatorname{dist}(y, \operatorname{spt}(\sigma)),$$

providing a quantitative justification to the informal assertion that "$\operatorname{spt}(\sigma)$ misses the shocks in Du". □

Proof. 1. Fix $y \in \mathbb{R}^n$ and take any point $x \in \operatorname{spt}(\sigma)$ at which u is differentiable.

According to Theorem 5.2 with $h := y - x$, we have

$$|u(y) - 2u(x) + u(2x-y)| \leq C|x-y|^2. \tag{6.3}$$

By semiconcavity, we have

$$u(y) - u(x) - Du(x) \cdot (y-x) \leq C|x-y|^2, \tag{6.4}$$

and also

$$u(2x-y) - u(x) - Du(x) \cdot (2x-y-x) \leq C|x-y|^2. \tag{6.5}$$

Use (6.5) in (6.3):

$$u(y) - u(x) - Du(x) \cdot (y-x) \geq -C|x-y|^2.$$

This and (6.4) establish (6.1).

2. Estimate (6.2) follows from (6.1), as follows. Take x, y as above. Let z be a point to be selected later, with $|x - z| \le 2|x - y|$.

The semiconcavity of u implies that

$$u(z) \le u(y) + Du(y) \cdot (z - y) + C|z - y|^2. \tag{6.6}$$

Also,

$$u(z) = u(x) + Du(x) \cdot (z - x) + O(|x - z|^2), \quad u(y) = u(x) + Du(x) \cdot (y - x) + O(|x - y|^2),$$

according to (6.1). Insert these indentities into (6.6) and simplify:

$$(Du(x) - Du(y)) \cdot (z - y) \le C|x - y|^2.$$

Now take

$$z := y + |x - y| \frac{Du(x) - Du(y)}{|Du(x) - Du(y)|}$$

to deduce (6.2).

3. Now take any point $x \in \mathrm{spt}(\sigma)$, and fix y. There exist points $x_k \in \mathrm{spt}(\sigma)$ $(k = 1, \dots)$ such that $x_k \to x$ and u is differentiable at x_k. According to estimate (6.1)

$$|u(y) - u(x_k) - Du(x_k) \cdot (y - x_k)| \le C|x_k - y|^2 \quad (k = 1, \dots).$$

The constant C does not depend on k or y. Now let $k \to \infty$. Owing to (6.2) we see that $\{Du(x_k)\}$ converges to some vector η, for which

$$|u(y) - u(x) - \eta \cdot (y - x)| \le C|x - y|^2.$$

Consequently u is differentiable at x and $Du(x) = \eta$. □

As an application of these bounds, we show next that the set $\mathcal{M} = \mathrm{spt}(\nu)$ lies on an n-dimensional Lipschitz continuous graph. This theorem (in position-velocity variables) is due originally to Mather [Mt2].

Theorem 6.2. *There exists a constant C, depending only on P and H, such that*

$$|D_x u(P, x_1) - D_x u(P, x_2)| \le C|x_1 - x_2| \tag{6.7}$$

for σ-a.e. pair of points x_1, x_2.

Proof. In view of (6.2) we can extend the mapping $x \mapsto Du(x)$ to a uniformly Lipschitz function defined on all of $\mathbb{T}^n$. The support of ν lies on the graph of this mapping. □

7. Derivative estimates in the variable P.

We turn next to some bounds involving variations in P. These are rather subtle and involve the smoothness properties of $\bar{H}$. (See Pöschel [P, p. 656–657] for an explicit linear example, showing that u can be less well behaved in P than in x.)

7.1. A formal L^2-estimate.

As in §5.1, we begin with a simple, but unjustified, calculation that suggests the later proof. So for the moment suppose u and $\bar{H}$ are smooth, differentiate the cell PDE twice with respect to P_i, and sum on i:

$$H_{p_k p_l}(D_x u, x) u_{x_k P_i} u_{x_l P_i} + H_{p_k}(D_x u, x) u_{x_k P_i P_i} = \bar{H}_{P_i P_i}(P). \tag{7.1}$$

The first term on the left is greater than or equal to $\gamma |D^2_{xP} u|^2$. Consequently

$$\gamma \int_{\mathbb{T}^n} |D^2_{xP} u|^2 \, d\sigma + \int_{\mathbb{T}^n} D_p H \cdot D_x(\Delta_{P} u) \, d\sigma \leq \Delta \bar{H}(P),$$

where $\Delta \bar{H} = \Delta_P \bar{H}$ is the Laplacian of $\bar{H}$ in P. Since $\Delta_{P} u = \Delta_{P} v$ is periodic, the second term on the left equals zero. Therefore

$$\int_{\mathbb{T}^n} |D^2_{xP} u|^2 \, d\sigma \leq C \Delta \bar{H}(P). \tag{7.2}$$

7.2. An L^2-estimate of difference quotients in P.

We next provide a rigorous version of the foregoing calculation, replacing derivatives by difference quotients.

Theorem 7.1. *There exists a positive constant C, depending only on H, such that*

$$\int_{\mathbb{T}^n} |D_x u(\tilde{P}, x) - D_x u(P, x)|^2 \, d\sigma \leq C(\bar{H}(\tilde{P}) - \bar{H}(P) - Q \cdot (\tilde{P} - P)) \tag{7.3}$$

for all $\tilde{P} \in \mathbb{R}^n$.

Remark. Recall that $Q = \int_{\mathbb{R}^n} \int_{\mathbb{T}^n} D_p H(p, x) \, d\nu = \int_{\mathbb{T}^n} D_p H(D_x u, x) \, d\sigma$ and that $Q \in \partial \bar{H}(P)$. In (7.3), $u(\tilde{P}, x) = \tilde{P} \cdot x + v(\tilde{P}, x)$ and $v = v(\tilde{P}, x)$ is any viscosity solution of the cell problem

$$\begin{cases} H(\tilde{P} + D_x v, x) = \bar{H}(\tilde{P}) \text{ in } \mathbb{R}^n \\ x \mapsto v(\tilde{P}, x) \ \text{ is } \mathbb{T}^n\text{-periodic.} \end{cases} \tag{7.4}$$

If $D_x u(\tilde{P}, x)$ is multivalued, we interpret (7.3) to mean

$$\int_{\mathbb{T}^n} |\tilde{\boldsymbol{\xi}} - D_x u(P, x)|^2 \, d\sigma \leq C(\bar{H}(\tilde{P}) - \bar{H}(P) - Q \cdot (\tilde{P} - P))$$

for some σ-measurable selection $\tilde{\boldsymbol{\xi}} \in D_x u(\tilde{P}, \cdot)$. □

Proof. 1. Write $\tilde{v}(\cdot) = v(\tilde{P}, \cdot)$, $\tilde{u} = x \cdot \tilde{P} + \tilde{v}$. Mollifying, we have

$$H(D\tilde{u}^\varepsilon, x) \leq \bar{H}(\tilde{P}) + C\varepsilon. \tag{7.5}$$

Therefore for σ almost every point

$$\begin{aligned} \frac{\gamma}{2}|D\tilde{u}^\varepsilon - Du|^2 + D_p H(Du, x) \cdot (D\tilde{u}^\varepsilon - Du) &\leq H(D\tilde{u}^\varepsilon, x) - H(Du, x) \\ &\leq \bar{H}(\tilde{P}) - \bar{H}(P) + C\varepsilon. \end{aligned} \tag{7.6}$$

Observe that $D\tilde{u}^\varepsilon - Du = \tilde{P} - P + (D\tilde{v}^\varepsilon - Dv)$ and

$$\int_{\mathbb{R}^n} \int_{\mathbb{T}^n} D_p H \cdot (D\tilde{v}^\varepsilon - Dv) \, d\nu = 0.$$

Consequently (7.6) yields

$$\begin{aligned} \frac{\gamma}{2} \int_{\mathbb{T}^n} |D\tilde{u}^\varepsilon - Du|^2 \, d\sigma &\leq \bar{H}(\tilde{P}) - \bar{H}(P) - \int_{\mathbb{T}^n} D_p H(Du, x) \cdot (\tilde{P} - P) \, d\sigma + C\varepsilon \\ &= \bar{H}(\tilde{P}) - \bar{H}(P) - Q \cdot (\tilde{P} - P) + C\varepsilon. \end{aligned} \tag{7.7}$$

Let $\varepsilon \to 0$. □

Remark. For use later, we record the estimate

$$\limsup_{\varepsilon \to 0} \int_{\mathbb{T}^n} \beta_\varepsilon \, d\sigma \leq \bar{H}(\tilde{P}) - \bar{H}(P) - Q \cdot (\tilde{P} - P), \tag{7.8}$$

for

$$\beta_\varepsilon(x) := \frac{\gamma}{2} \int_{\mathbb{T}^n} \eta_\varepsilon(x - y) |D_x u(\tilde{P}, y) - D_x u^\varepsilon(\tilde{P}, x)|^2 \, dy. \tag{7.9}$$

To see this, note that as in the proof of Theorem 4.1 we can replace (7.5) by the stronger inequality

$$\beta_\varepsilon(x) + H(D\tilde{u}^\varepsilon, x) \leq \bar{H}(\tilde{P}) + C\varepsilon.$$

□

Corollary 7.2. (i) *For each $P \in \mathbb{R}^n$, we have*

$$\int_{\mathbb{T}^n} |D_x u(\tilde{P}, x) - D_x u(P, x)|^2 \, d\sigma \leq O(|\tilde{P} - P|) \quad \text{as } \tilde{P} \to P.$$

(ii) *If $\bar{H}$ is differentiable at P,*

$$\int_{\mathbb{T}^n} |D_x u(\tilde{P}, x) - D_x u(P, x)|^2 \, d\sigma \leq o(|\tilde{P} - P|) \quad \text{as } \tilde{P} \to P.$$

(iii) *If $\bar{H}$ is twice-differentiable at P,*

$$\int_{\mathbb{T}^n} |D_x u(\tilde{P}, x) - D_x u(P, x)|^2 \, d\sigma \leq O(|\tilde{P} - P|^2) \quad \text{as } \tilde{P} \to P.$$

8. Application: strict convexity of $\bar{H}$ in certain directions.

The next estimate allows us to deduce certain strict convexity properties of $\bar{H}$.

Theorem 8.1. (i) *There exists a positive constant C such that for each $R \in \mathbb{R}^n$, we have*

$$-R \cdot \tilde{Q}, R \cdot \hat{Q} \leq C \left(\liminf_{t \to 0^+} \frac{\bar{H}(P + tR) - 2\bar{H}(P) + \bar{H}(P - tR)}{t^2} \right)^{1/2}, \tag{8.1}$$

where $\tilde{Q}, \hat{Q} \in \partial \bar{H}(P)$.

(ii) *In particular, if $\bar{H}$ is twice differentiable at P, then*

$$|D\bar{H}(P) \cdot R| \leq C(R \cdot D^2\bar{H}(P)R)^{1/2} \tag{8.2}$$

for each $R \in \mathbb{R}^n$.

Proof. 1. Fix $R \in \mathbb{R}^n$, $t > 0$, and take

$$\tilde{u} = u(P + tR, \cdot), \ \hat{u} = u(P - tR, \cdot).$$

Then for σ-a.e. point x:

$$H(D\tilde{u}^\varepsilon, x) - 2H(Du, x) + H(D\hat{u}^\varepsilon, x) \leq \bar{H}(P + tR) - 2\bar{H}(P) + \bar{H}(P - tR) + C\varepsilon.$$

Similarly to the proof in §5.2, we deduce

$$\int_{\mathbb{T}^n} |D\tilde{u}^\varepsilon - Du|^2 + |D\hat{u}^\varepsilon - Du|^2 \, d\sigma \leq C(\bar{H}(P + tR) - 2\bar{H}(P) + \bar{H}(P - tR)) + C\varepsilon. \tag{8.3}$$

2. Since $H(D\tilde{u}^\varepsilon, x) \leq \bar{H}(P + tR) + C\varepsilon$, we have

$$\begin{aligned} \bar{H}(P) - \bar{H}(P + tR) &\leq \int_{\mathbb{T}^n} H(Du, x) - H(D\tilde{u}^\varepsilon, x) \, d\sigma + C\varepsilon \\ &\leq C \left(\int_{\mathbb{T}^n} |Du - D\tilde{u}^\varepsilon|^2 \, d\sigma \right)^{1/2} + C\varepsilon. \end{aligned} \tag{8.4}$$

Likewise,

$$\bar{H}(P) - \bar{H}(P - tR) \leq C \left(\int_{\mathbb{T}^n} |Du - D\hat{u}^\varepsilon|^2 \, d\sigma \right)^{1/2} + C\varepsilon. \tag{8.5}$$

Combining (8.3)–(8.5), sending $\varepsilon \to 0$, and recalling the convexity of $\bar{H}$, we find

$$-t\tilde{Q}(t) \cdot R, \ t\hat{Q}(t) \cdot R \leq C(\bar{H}(P + tR) - 2\bar{H}(P) + \bar{H}(P - tR))^{1/2}$$

for any

$$\tilde{Q}(t) \in \partial\bar{H}(P+tR),\ \hat{Q}(t) \in \partial\bar{H}(P-tR).$$

Taking any $t_k \to 0$, we may assume $\tilde{Q}(t_k) \to \tilde{Q}$, $\hat{Q}(t_k) \to \hat{Q}$ with $\tilde{Q}, \hat{Q} \in \partial\bar{H}(P)$. Estimate (8.1) follows. □

Remarks. (i) From (8.1) we deduce that $\bar{H}$ is strictly convex in any direction R which is not tangent to the level set $\{\bar{H} = \bar{H}(P)\}$, provided $\bar{H}$ is differentiable at P. (Compare this assertion with Iturriaga [I].)

(ii) More generally, if $\bar{H}(P) > \min_{\mathbb{R}^n} \bar{H}$, and so $0 \notin \partial\bar{H}(P)$, there exists an open convex cone of directions R in which $\bar{H}$ is strictly convex at P.

Therefore the graph of $\bar{H}$ can contain an n-dimensional flat region only possibly at its minimum value. This can in fact happen, even though H is uniformly convex in the variable p: see Lions–Papanicolaou–Varadhan [L-P-V] or Braides–Defranceschi [B-D, p. 149]. Consult Concordel [C1,C2] for more. Physically, a flat region at the minimum of $\bar{H}$ corresponds to "nonballistic" trajectories for the dynamics.

(iii) See also Bangert [B] and Weinan E [EW2] for an example showing that the level sets of $\bar{H}$ can have corners and/or flat parts. □

9. Application: averaging in the variable X.

Assume for this section that $\bar{H}$ is differentiable at P and furthermore that $Q = D\bar{H}(P)$ satisfies the *nonresonance condition*:

$$(9.1) \qquad Q \cdot k \neq 0 \quad \text{for each vector } k \in \mathbb{Z}^n, k \neq 0.$$

Notation. For $h > 0$, we write the vector of difference quotients

$$(9.2) \qquad D_P^h u(P,x) := \left(\dots, \frac{u(P+he_l,x) - u(P,x)}{h}, \dots \right),$$

for $e_l := (0, \dots, 1, \dots, 0)$, the 1 in the l^{th}-position. □

Theorem 9.1. *Suppose $Q = D\bar{H}(P)$ satisfies* (9.1). *Then*

$$(9.3) \qquad \lim_{h \to 0} \int_{\mathbb{T}^n} \Phi(D_P^h u(P,x))\, d\sigma = \int_{\mathbb{T}^n} \Phi(X)\, dX$$

for each continuous, $\mathbb{T}^n$-periodic function Φ.

Proof. 1. Let $u_l(\cdot) := u(P+he_l, \cdot)$, and $u_l^\varepsilon := \eta_\varepsilon * u_l$, for $l = 1, \dots, n$.

Since H is smooth, we have for all p, q lying in a compact subset of $\mathbb{R}^n$ that

$$H(q,x) = H(p,x) + D_pH(p,x)\cdot(q-p) + R, \quad \text{with } |R| \leq C|q-p|^2.$$

Take $q = Du_l(y)$, $p = Du_l^\varepsilon(x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y) Du_l(y)\,dy$, multiply by $\eta_\varepsilon(x-y)$, and integrate with respect to y:

$$H(Du_l^\varepsilon(x), x) = \int_{\mathbb{R}^n} \eta_\varepsilon(x-y) H(Du_l(y), x)\,dy - \int_{\mathbb{R}^n} \eta_\varepsilon(x-y) R\,dy. \tag{9.4}$$

Furthermore the PDE $H(Du_l, x) = \bar{H}(P+he_l)$ holds pointwise a.e., and so we can conclude that

$$H(Du_l^\varepsilon, x) = \bar{H}(P+he_l) + \gamma_\varepsilon^l, \tag{9.5}$$

where the error term is estimated by

$$|\gamma_\varepsilon^l| \leq C(\varepsilon + \beta_\varepsilon^l)$$

for

$$\beta_\varepsilon^l(x) := \frac{\gamma}{2} \int_{\mathbb{R}^n} \eta_\varepsilon(x-y) |Du_l(y) - Du_l^\varepsilon(x)|^2\,dy.$$

2. We introduce the partially smoothed vector of difference quotients

$$D_P^h u^\varepsilon(P, x) := \left(\dots, \frac{u_l^\varepsilon - u}{h}, \dots \right), \tag{9.6}$$

and take then a vector of integers $k = (k_1, \dots, k_n)$, $k \neq 0$.

Next, observe that the function

$$e^{2\pi i k \cdot D_P^h u^\varepsilon} = e^{2\pi i k \cdot x} e^{2\pi i k \cdot D_P^h v^\varepsilon}$$

is $\mathbb{T}^n$-periodic, even though $D_P^h u^\varepsilon$ is not periodic. Hence

$$\begin{aligned} 0 &= \int_{\mathbb{T}^n} D_p H(Du, x) \cdot D_x \left(e^{2\pi i k \cdot D_P^h u^\varepsilon} \right) d\sigma \\ &= 2\pi i \int_{\mathbb{T}^n} e^{2\pi i k \cdot D_P^h u^\varepsilon} \sum_{l=1}^n k_l D_p H(Du, x) \cdot D_x \left(\frac{u_l^\varepsilon - u}{h}\right) d\sigma. \end{aligned} \tag{9.7}$$

3. Now (9.5) implies

$$H(Du_l^\varepsilon, x) - H(Du, x) = \bar{H}(P+he_l) - \bar{H}(P) + \gamma_\varepsilon^l.$$

Consequently

$$D_p H(Du, x) \cdot D(u_l^\varepsilon - u) = \bar{H}(P+he_l) - \bar{H}(P) + \Gamma_\varepsilon^l,$$

where

$$|\Gamma_\varepsilon^l| \leq C(\varepsilon + \beta_\varepsilon^l + |Du_l^\varepsilon - Du|^2) \tag{9.8}$$

for $l = 1, \dots, n$.

Therefore

$$D_pH(Du,x)\cdot D_x(\frac{u_l^\varepsilon - u}{h}) = Q_l + \left(\frac{\bar H(P+he_l)-\bar H(P)}{h} - Q_l + \frac{1}{h}\Gamma_\varepsilon^l\right). \tag{9.9}$$

4. Insert (9.9) into (9.7), and then estimate

$$\begin{aligned}\left|(Q\cdot k)\int_{\mathbb{T}^n} e^{2\pi i k\cdot D_P^h u^\varepsilon}\, d\sigma\right| &\leq \frac{C\varepsilon}{h} + C\sum_{l=1}^n \left(\frac{\bar H(P+he_l)-\bar H(P)}{h} - Q_l\right)\\ &\quad + \frac{C}{h}\sum_{l=1}^n \int_{\mathbb{T}^n} \beta_\varepsilon^l + |Du_l^\varepsilon - Du|^2\, d\sigma \quad \text{by (9.8)}\\ &\leq \frac{C\varepsilon}{h} + C\sum_{l=1}^n \left(\frac{\bar H(P+he_l)-\bar H(P)}{h} - Q_l\right)\\ &\quad + \frac{C}{h}\sum_{l=1}^n \int_{\mathbb{T}^n} \beta_\varepsilon^l\, d\sigma,\end{aligned} \tag{9.10}$$

the last inequality following from (7.7) in the proof of Theorem 7.1.

Next, send $\varepsilon \to 0$, and remember (7.8):

$$\left|(Q\cdot k)\int_{\mathbb{T}^n} e^{2\pi i k\cdot D_P^h u}\, d\sigma\right| \leq C\sum_{l=1}^n \left(\frac{\bar H(P+he_l)-\bar H(P)}{h} - Q_l\right).$$

Since $Q_l = \bar H_{P_l}(P)$ and $Q\cdot k \neq 0$, we conclude that

$$\lim_{h\to 0}\int_{\mathbb{T}^n} e^{2\pi i k\cdot D_P^h u}\, d\sigma = 0$$

for all $k \in \mathbb{Z}^n, k \neq 0$. Because any continuous, $\mathbb{T}^n$-periodic funtion Φ can be uniformly approximated by trigonometric polynomials, this implies assertion (9.3). □

Remarks. (i) Recalling the formal change of variables (1.3), we interpret (9.3) to assert

$$\text{“ } d\sigma = |\det D_{xP}^2 u|\; dx \text{ ”} \tag{9.11}$$

in some weak sense, provided (9.1) holds. See [E-G2, §5.1] for related formal computations.

(ii) Theorem 9.1 provides a partial, but rigorous, interpretation of the following heuristics.

Suppose that our generating function u is smooth, and induces the global change of variables $(p,x) \to (P,X)$ by (1.3). The the dynamics (1.1) become (1.5); that is,

$$\begin{cases}\dot{\mathbf{X}} = D\bar H(\mathbf{P})\\ \dot{\mathbf{P}} = 0.\end{cases}$$

Consequently $\mathbf{X}(t) = Qt + X_0, \mathbf{P}(t) \equiv P$. In view therefore of the nonresonance condition (9.1), we have

$$\lim_{T\to\infty} \frac{1}{\lambda T} \int_0^{\lambda T} \Phi(\mathbf{X}(t))\, dt = \int_{\mathbb{T}^n} \Phi(X)\, dX.$$

for each $\lambda > 0$. However

$$\begin{aligned}
\frac{1}{\lambda T} \int_0^{\lambda T} \Phi(\mathbf{X}(t))\, dt &= \frac{1}{\lambda T} \int_0^{\lambda T} \Phi(D_P u(P, \mathbf{x}(t)))\, dt \\
&= \frac{1}{\lambda} \int_0^{\lambda} \Phi(D_P u(P, \frac{\mathbf{x}_\varepsilon(t)}{\varepsilon}))\, dt \quad \text{for } \varepsilon = \frac{1}{T} \\
&\to \frac{1}{\lambda} \int_0^{\lambda} \int_{\mathbb{T}^n} \Phi(D_P u(P, x))\, d\sigma_t\, dt.
\end{aligned}$$

Consequently

$$\int_{\mathbb{T}^n} \Phi(D_P u(P, x))\, d\sigma_t = \int_{\mathbb{T}^n} \Phi(X)\, dX$$

for all $t \geq 0$. $\square$

References

[Ar] M. Arisawa, *Multiscale homogenization for first-order Hamilton–Jacobi–Bellman equations*, Advances in Differential Equations (to appear).

[A] S. Aubry, *The twist map, the extended Frenkel–Kantorova model and the devil's straircase*, Physica D **7** (1983), 240–258.

[B1] V. Bangert, *Minimal geodesics*, Ergodic Theory and Dyn. Systems **10** (1989), 263–286.

[B2] V. Bangert, *Geodesic rays, Busemann functions and monotone twist maps*, Calculus of Variations **2** (1994), 49–63.

[B-S] G. Barles and P. E. Souganidis, *On the long time behavior of solutions of Hamilton–Jacobi equations*, J. Math Analysis (to appear).

[B-D] A. Braides and A. Defranceschi, *Homogenization of Multiple Integrals*, Oxford Univ. Press, 1998.

[C-D] W. Chou and R. J. Duffin, *An additive eigenvalue problem of physics related to linear programming*, Advances in Appl. Math **8** (1987), 486–498.

[Cl] F. Clarke, *Optimization and Nonsmooth Analysis*, Wiley-Interscience, 1983.

[C1] M. Concordel, *Periodic homogenization of Hamilton-Jacobi equations I: additive eigenvalues and variational formula*, Indiana Univ. Math. J. **45** (1996), 1095–1117.

[C2] M. Concordel, *Periodic homogenisation of Hamilton-Jacobi equations II: eikonal equations*, Proc. Roy. Soc. Edinburgh **127** (1997), 665–689.

[DC] M. J. Dias Carneiro, *On minimizing measure of the action of autonomous Lagrangians*, Nonlinearity **8** (1995), 1077–1085.

[EW1] Weinan E, *A class of homogenization problems in the calculus of variations*, Comm Pure and Appl Math **44** (1991), 733–754.

[EW2] Weinan E, *Aubry–Mather theory and periodic solutions of the forced Burgers equation*, Comm Pure and Appl Math **52** (1999), 811–828.

[E1] L. C. Evans, *Weak Convergence Methods for Nonlinear Partial Differential Equations*, American Math Soc, 1990.

[E2] L. C. Evans, *Periodic homogenization of certain fully nonlinear PDE*, Proc Royal Society Edinburgh **120** (1992), 245–265.

[E-G2] L. C. Evans and D. Gomes, *Effective Hamiltonians and Averaging for Hamiltonian Dynamics II*, to appear.

[F1] A. Fathi, *Théorème KAM faible et théorie de Mather sur les systèmes lagrangiens*, C. R. Acad. Sci. Paris Sr. I Math. **324** (1997), 1043–1046.

[F2] A. Fathi, *Solutions KAM faibles conjuguées et barrières de Peierls*, C. R. Acad. Sci. Paris Sr. I Math. **325** (1997), 649–652.

[F3] A. Fathi, *Orbites hétéroclines et ensemble de Peierls*, C. R. Acad. Sci. Paris Sr. I Math. **326** (1998), 1213–1216.

[F4] A. Fathi, *Sur la convergence du semi-groupe de Lax-Oleinik*, C. R. Acad. Sci. Paris Sr. I Math. **327** (1998), 267–270.

[F-M] A. Fathi and J. Mather, *Failure of convergence of the Lax–Oleinik semigroup in the time–periodic case*, preprint (2000).

[Gd] H. Goldstein, *Classical mechanics* (2nd ed.), Addison-Wesley, 1980.

[G] D. Gomes, *Hamilton–Jacobi Equations, Viscosity Solutions and Asymptotics of Hamiltonian Systems*, Ph.D. Thesis, University of California, Berkeley (2000).

[J-K-M] H. R. Jauslin, H. O. Kreiss and J. Moser, *On the forced Burgers equation with periodic boundary conditions*, preprint (1998).

[L-P-V] P.-L. Lions, G. Papanicolaou, and S. R. S. Varadhan, *Homogenization of Hamilton–Jacobi equations*, unpublished, circa 1988.

[I] R. Iturriaga, *Minimizing measures for time-dependent Lagrangians*, Proc. London Math Society **73** (1996), 216–240.

[Mn1] R. Mañé, *Global Variational Methods in Conservative Dynamics*, Instituto de Matemática Pura e Aplicada, Rio de Janeiro.

[Mn2] R. Mañé, *On the minimizing measures of Lagrangian dynamical systems*, Nonlinearity **5** (1992), 623–638.

[Mn3] R. Mañé, *Generic properties and problems of minimizing measures of Lagrangian systems*, Nonlinearity **9** (1996), 273–310.

[Mt1] J. Mather, *Minimal measures*, Comment. Math Helvetici **64** (1989), 375–394.

[Mt2] J. Mather, *Action minimizing invariant measures for positive definite Lagrangian systems*, Math. Zeitschrift **207** (1991), 169–207.

[Mt3] J. Mather, *Differentiability of the minimal average action as a function of the rotation number*, Bol. Soc. Bras. Math **21** (1990), 59–70.

[Mt4] J. Mather, *Variational construction of connecting orbits*, Ann. Inst. Fourier, Grenoble **43** (1993), 1347–1386.

[M-F] J. Mather and G. Forni, *Action minimizing orbits in Hamiltonian systems*, Transition to Chaos in Classical and Quantum Mechanics, Lecture Notes in Math 1589 (S. Graffi, ed.), Sringer, 1994.

[Ms] J. Moser, *Recent developments in the theory of Hamiltonian systems*, SIAM Review **28** (1986), 459–486.

[N-F] G. Namah and J.-M. Roquejoffre, *Comportement asymptotique des solutions d'une classe d'equations paraboliques and de Hamilton–Jacobi*, C. R. Acad. Sci. Paris Sr. I Math. **324** (1997), 1367–1370.

[N] R. Nussbaum, *Convergence of iterates of a nonlinear operator arising in statistical mechanics*, Nonlinearity **4** (1991), 1223–1239.

[P] J. Pöschel, *Integrability of Hamiltonian systems on Cantor sets*, Comm. Pure Appl. Math **35** (1982), 653–696.

[Rz] F. Rezakhanlou, *Central limit theorem for stochastic Hamilton–Jacobi equations*, preprint (1998).

[R] J.-M. Roquejoffre, *Comportement asymptotique des solutions d'equations de de Hamilton–Jacobi monodimensionnelles*, C. R. Acad. Sci. Paris Sr. I Math. **326** (1998), 185–189.

[So1] A. N. Sobolevskii, *Periodic solutions of the Hamilton–Jacobi equation with periodic forcing term*, Russian Math Surveys **53** (1998), 1375–1376.

[So2] A. N. Sobolevskii, *Periodic solutions of the Hamilton–Jacobi equation with a periodic nonhomogeneous term and Aubry–Mather theory*, Sbornik: Mathematics **190** (1999), 1487–1504.

[S] P. E. Souganidis, *Stochastic homogenization of Hamilton–Jacobi equations and some applications*, Asymptotic Analysis **20** (1999), 1-11.