Weak KAM Theory and Partial Differential Equations

Lawrence C. Evans
Department of Mathematics
University of California, Berkeley

1 Overview, KAM theory

These notes record and slightly modify my 5 lectures from the CIME conference on "Calculus of variations and nonlinear partial differential equations", held in Cetraro during the week of June 27 - July 2, 2005, organized by Bernard Dacorogna and Paolo Marcellini. I am proud to brag that this was the third CIME course I have given during the past ten years, the others at the meetings on "Viscosity solutions and applications" (Montecatini Terme, 1995) and on "Optimal transportation and applications" (Martina Franca, 2001).

My intention was (and is) to introduce some new PDE methods developed over the past 6 years in so-called "weak KAM theory", a subject pioneered by J. Mather and A. Fathi. Succinctly put, the goal of this subject is the employing of dynamical systems, variational and PDE methods to find "integrable structures" within general Hamiltonian dynamics.

My main references for most of these lectures are Fathi's forthcoming book [F5] (as well as his sequence of short notes [F1]–[F4]) and my paper [E-G1] with Diogo Gomes. A nice recent survey is Kaloshin [K]; and see also the survey paper [E6].

1.1 Classical theory

I begin with a quick recounting of classical Lagrangian and Hamiltonian dynamics, and a brief discussion of the standard KAM theorem. A good elementary text is Percival and Richards [P-R], and Goldstein [G] is a standard reference.

1.1.1 The Lagrangian viewpoint

DEFINITION. We are given a function $L : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ called the *Lagrangian*, $L = L(v, x)$. We regard $x \in \mathbb{R}^n$, $x = (x_1, \dots, x_n)$, as the *position* variable and $v \in \mathbb{R}^n$, $v = (v_1, \dots, v_n)$, as the *velocity*.

HYPOTHESES. We hereafter assume that

(i) there exist constants $0 < \gamma \leq \Gamma$ such that

$$\Gamma|\xi|^2 \geq \sum_{i,j=1}^{n} L_{v_i v_j}(v,x)\xi_i\xi_j \geq \gamma|\xi|^2 \qquad \text{(uniform convexity)} \tag{1.1}$$

for all ξ, v, x; and

(ii) the mapping

$$x \mapsto L(v,x) \text{ is } \mathbb{T}^n\text{-periodic} \qquad \text{(periodicity)} \tag{1.2}$$

for all x, where $\mathbb{T}^n = [0,1]^n$ denotes the unit cube in $\mathbb{R}^n$, with opposite faces identified.

DEFINITION. Given a curve $\mathbf{x} : [0,T] \to \mathbb{R}$, we define its *action* to be

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

where $\dot{} = \frac{d}{dt}$.

THEOREM 1.1 (Euler–Lagrange equation) *Suppose that $x_0, x_T \in \mathbb{R}^n$ are given, and define the admissible class of curves*

$$\mathcal{A} := \{\mathbf{y} \in C^2([0,T];\mathbb{R}^n) \mid \mathbf{y}(0) = x_0,\ \mathbf{y}(T) = x_T\}.$$

Suppose $\mathbf{x}(\cdot) \in \mathcal{A}$ and

$$A_T[\mathbf{x}(\cdot)] = \min_{\mathbf{y}\in\mathcal{A}} A_T[\mathbf{y}(\cdot)].$$

Then the curve $\mathbf{x}(\cdot)$ solves the Euler–Lagrange equations

$$-\frac{d}{dt}(D_v L(\dot{\mathbf{x}}, \mathbf{x})) + D_x L(\dot{\mathbf{x}}, \mathbf{x}) = 0 \qquad (0 \leq t \leq T). \tag{E-L}$$

A basic example. The Lagrangian $L = \frac{|v|^2}{2} - W(x)$ is the difference between the kinetic energy $\frac{|v|^2}{2}$ and the potential energy $W(x)$. In this case the Euler-Lagrange equations read

$$\ddot{\mathbf{x}} = -DW(\mathbf{x}).$$

$\square$

1.1.2 The Hamiltonian viewpoint. In view of the uniform convexity of $v \mapsto L(v,x)$, we can uniquely and smoothly solve the equation

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

for

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

DEFINITION. We define the *Hamiltonian*

$$H(p, x) := p \cdot \mathbf{v}(p, x) - L(\mathbf{v}(p, x), x). \tag{1.3}$$

Equivalently,

$$H(p.x) = \max_{v \in \mathbb{R}^n} (p \cdot v - L(v, x)). \tag{1.4}$$

THEOREM 1.2 (Hamiltonian dynamics) *Suppose* $\mathbf{x}(\cdot)$ *solves the Euler-Lagrange equations* (E-L). *Define*

$$\mathbf{p}(t) := D_v L(\mathbf{x}(t), \dot{\mathbf{x}}(t)).$$

Then the pair $(\mathbf{x}(\cdot), \mathbf{p}(\cdot))$ *solves Hamilton's equations*

$$\begin{cases} \dot{\mathbf{x}} = D_P H(\mathbf{p}, \mathbf{x}) \\ \dot{\mathbf{p}} = -D_x H(\mathbf{p}, \mathbf{x}). \end{cases} \tag{H}$$

Also,

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

this is conservation of energy.

Basic example again. For the Lagrangian $L = \frac{|v|^2}{2} - W(x)$, the corresponding Hamiltonian is $H(p, x) = \frac{|p|^2}{2} + W(x)$, the sum of the kinetic and potential energies. The Hamiltonian dynamics (H) then read

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{p} \\ \dot{\mathbf{p}} = -DW(\mathbf{x}); \end{cases}$$

and the total energy is conserved. □

1.1.3 Canonical changes of variables, generating functions. Let $u : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, $u = u(P, x)$, be a given smooth function, called a *generating function*. Consider the formulas

$$\begin{cases} p := D_x u(P, x) \\ X := D_P u(P, x). \end{cases} \tag{1.5}$$

Assume we can solve (1.5) globally for X, P as smooth functions of x, p, and vice versa:

$$
\begin{array}{rcl} X &=& \mathbf{X}(p,x) \\ P &=& \mathbf{P}(p,x) \end{array} \quad \Leftrightarrow \quad \begin{array}{rcl} x &=& \mathbf{x}(P,X) \\ p &=& \mathbf{p}(P,X). \end{array} \tag{1.6}
$$

We next study the Hamiltonian dynamics in the new variables:

THEOREM 1.3 (Change of variables) *Let* $(\mathbf{x}(\cdot), \mathbf{p}(\cdot))$ *solve Hamilton's equations* (H). *Define*

$$
\begin{array}{rcl} \mathbf{X}(t) &:=& \mathbf{X}(\mathbf{p}(t), \mathbf{x}(t)) \\ \mathbf{P}(t) &:=& \mathbf{P}(\mathbf{p}(t), \mathbf{x}(t)). \end{array}
$$

Then

$$
\begin{cases} \dot{\mathbf{X}} = D_P K(\mathbf{P}, \mathbf{X}) \\ \dot{\mathbf{P}} = -D_X K(\mathbf{P}, \mathbf{X}) \end{cases} \tag{K}
$$

for the new Hamiltonian

$$
K(P,X) := H(\mathbf{p}(P,X), \mathbf{x}(P,X)).
$$

DEFINITION. A transformation $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$,

$$
\Psi(p,x) = (P,X)
$$

is called *canonical* if it preserves the Hamiltonian structure. This means

$$
(D\Psi)^T J \; D\Psi = J
$$

for

$$
J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}
$$

See any good classical mechanics textbook, such as Goldstein [G], for more explanation. The change of variables $(p,x) \mapsto (P,X)$ induced by the generating function $u = u(P,x)$ is canonical.

1.1.4 Hamilton–Jacobi PDE. Suppose now our generating function $u = u(P,x)$ solves the *stationary Hamilton-Jacobi equation*

$$
H(D_x u, x) = \bar{H}(P), \tag{1.7}
$$

where the right hand side at this point of the exposition just denotes some function of P alone. Then

$$K(P,X) = H(p,x) = H(D_x u, x) = \bar{H}(P);$$

and so (K) becomes

$$\begin{cases} \dot{\mathbf{X}} = D\bar{H}(\mathbf{P}) \\ \dot{\mathbf{P}} = 0. \end{cases} \tag{1.8}$$

The point is that these dynamics are trivial. In other words, if we can canonically change variables $(p,x) \mapsto (P,X)$ using a generating function u that solves a PDE of the form (1.7), we can easily solve the dynamics in the new variables.

DEFINITION. We call the Hamiltonian H *integrable* if there exists a *canonical* mapping

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

such that

$$(H \circ \Phi)(P,X) =: \bar{H}(P)$$

is a function only of P.

We call $P = (P_1, \dots, P_n)$ the *action variables* and $X = (X_1, \dots, X_n)$ the *angle variables*.

1.2 KAM Theory

A key question is whether we can in fact construct a canonical mapping Φ as above, converting to action-angle variables. This is in general impossible, since our PDE (1.7) will not usually have a solution u smooth in x and P; and, even if it does, it is usually not possible globally to change variables according to (1.5) and (1.6).

But KAM (Kolmogorov-Arnold-Moser) theory tells us that we can in fact carry out this procedure for a Hamiltonian that is an appropriate small perturbation of a Hamiltonian depending only on p. To be more specific, suppose our Hamiltonian H has the form

$$H(p,x) = H^0(p) + K^0_\varepsilon(p,x), \tag{1.9}$$

where we regard the term K^0_ε as a small perturbation to the integrable Hamiltonian H^0. Assume also that $x \mapsto K^0_\varepsilon(p,x)$ is $\mathbb{T}^n$-periodic.

1.2.1 Generating functions, linearization. We propose to find a generating function having the form

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

where v is small and periodic in x. Owing to (1.5) we would then change variable through the implicit formulas

$$\begin{cases} p = D_x u = P + D_x v \\ X = D_P u = x + D_P v. \end{cases} \tag{1.10}$$

We consequently must build v so that

$$H(D_x u, x) = \bar{H}(P), \tag{1.11}$$

the expression on the right to be determined.

Now according to (1.9), (1.10) and (1.11), we want

$$H^0(P + D_x v) + K_\varepsilon^0(P + D_x v, x) = \bar{H}(P). \tag{1.12}$$

We now make the informal assumption that K_ε^0, v, and their derivatives are $O(\varepsilon)$ as $\varepsilon \to 0$. Then

$$H^0(P) + DH^0(P) \cdot D_x v + K_\varepsilon^0(P, x) = \bar{H}(P) + O(\varepsilon^2). \tag{1.13}$$

Drop $O(\varepsilon^2)$ term and write

$$\omega(P) := DH^0(P).$$

Then

$$\omega(P) \cdot D_x v + K_\varepsilon^0(P, x) = \bar{H}(P) - H^0(P); \tag{1.14}$$

this is the formal linearization of our nonlinear PDE (1.12).

1.2.2 Fourier series. Next we use Fourier series to try to build a solution v of (1.14). To simplify a bit, suppose for this section that $\mathbb{T}^n = [0, 2\pi]^n$ (and not $[0, 1]^n$). We can then write

$$K_\varepsilon^0(P, x) = \sum_{k \in \mathbb{Z}^n} \hat{k}(P, k) e^{ik \cdot x}$$

for the Fourier coefficients

$$\hat{k}(P, k) := \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} K_\varepsilon^0(P, x) e^{-ik \cdot x} \, dx.$$

Let us seek a solution of (1.14) having the form

$$v(P, x) = \sum_{k \in \mathbb{Z}^n} \hat{v}(P, k) e^{ik \cdot x}, \tag{1.15}$$

the Fourier coefficients $\hat{v}(P,k)$ to be selected. Plug (1.15) into (1.14):

$$i\sum_{k\in\mathbb{Z}^n}(\omega(P)\cdot k)\hat{v}(P,x)e^{ik\cdot x}+\sum_{k\in\mathbb{Z}^n}\hat{k}(P,k)e^{ik\cdot x}=\bar{H}(P)-H^0(P).$$

The various terms agree if we *define*

$$\bar{H}(P):=H^0(P)+\hat{k}(P,0)$$

and set

$$\hat{v}(P,x):=\frac{i\hat{k}(P,k)}{\omega(P)\cdot k}\qquad(k\neq 0).$$

We have therefore derived the approximate solution

$$v(P,x)=i\sum_{k\neq 0}\frac{\hat{k}(P,k)}{\omega(P)\cdot k}e^{ik\cdot x},\tag{1.16}$$

assuming both that $\omega(P)\cdot k\neq 0$ for all nonzero $k\in\mathbb{Z}^n$ and that the series (1.16) converges.

1.2.3 Small divisors. To make rigorous use the foregoing calculations, we need first to ensure that $\omega(P)\cdot k\neq 0$, with some quantitative control:

DEFINITION. A vector $\omega\in\mathbb{R}^n$ is of type (L,γ) if

$$|k\cdot\omega|\geq\frac{L}{|k|^\gamma}\qquad\text{for all } k\in\mathbb{Z}^n,\ k\neq 0,\tag{1.17}$$

This is a "nonresonance condition". It turns out that most ω satisfy this condition for appropriate L,γ. Indeed, if $\gamma>n-1$, then

$$|\{\omega\in B(0,R)\mid|k\cdot\omega|\leq\frac{L}{|k|^\gamma}\text{ for some }k\in\mathbb{Z}^n\}|\to 0\quad\text{as } L\to 0.$$

1.2.4 Statement of KAM Theorem. We now explicitly put

$$K^0_\varepsilon(p,x)=\varepsilon K^0(p,x)$$

in (1.9), so that

$$H(p,x)=H^0(p)+\varepsilon K^0(p,x),$$

and assume

$$\begin{cases}\text{there exists } p^*\in\mathbb{R}^n\text{ such that}\\ \qquad\omega^*=DH^0(p^*)\\ \text{is type }(L,\gamma)\text{ for some }L,\gamma>0;\end{cases}\tag{H1}$$

(H2) $$D^2 H^0(p^*) \text{ is invertible;}$$

(H3) $$H^0, K^0 \text{ are real analytic.}$$

THEOREM 1.4 (KAM) *Under hypotheses* (H1)–(H3)*, there exists* $\varepsilon_0 > 0$ *such that for each* $0 < \varepsilon \leq \varepsilon_0$*, there exists* P^* *(close to* p^**) and a smooth mapping*

$$\Phi(P^*, \cdot) : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^n$$

such that for each $x_0 \in \mathbb{R}^n$

$$\Phi(P^*, x_0 + t\omega^*) =: (\mathbf{x}(t), \mathbf{p}(t))$$

solves the Hamiltonian dynamics (H)*.*

The idea of the proof is for $k = 0, 1, \ldots,$ iteratively to construct $\Phi^k, P_k, H^k, \varepsilon_k$ so that $P_0 = p^*$,

$$\omega^* = D\bar{H}^k(P_k) \quad \text{for all } k;$$
$$H^{k+1} := H^k \circ \Phi^k;$$

and

$$H^k(p, x) = \bar{H}^k(p) + K^k(p, x),$$

where $\|K^k\| \leq \varepsilon_k$ and the error estimates ε_k converge to 0 very rapidly. Then

$$P_k \to P^*, \quad \bar{H}^k(P_k) \to \bar{H}(P^*);$$

and

$$\Phi := \lim_{k \to \infty} \Phi^k \circ \Phi^{k-1} \circ \cdots \circ \Phi^1$$

is the required change of variables.

The full details of this procedure are very complicated. See Wayne's discussion [W] for much more explanation and for references to the vast literature on KAM.

REMARK. We have

$$\begin{cases} \mathbf{X}(t) = x_0 + t\omega^* \\ \mathbf{P}(t) \equiv P^* \end{cases} \tag{1.18}$$

for

$$\omega^* = D\bar{H}(P^*)$$

But note that, in spite of my notation, Φ is really only defined for $P = P^*$. □

2 Weak KAM Theory: Lagrangian Methods

Our goal in this and the subsequent sections is extending the foregoing classical picture into the large. The resulting, so-called "weak KAM theory" is a global and nonperturbative theory (but is in truth pretty weak, at least as compared with the assertions from the previous section). There are two approaches to these issues: the Lagrangian, dynamical systems methods (discussed in this section) and the nonlinear PDE methods (explained in the next section).

The following discussion follows Albert Fathi's new book [F5], which the interested reader should consult for full details of the proofs. Related expositions are Forni–Mather [Fo-M] and Mañé [Mn].

2.1 Minimizing trajectories

NOTATION. If $\mathbf{x} : [0,T] \to \mathbb{R}^n$ and

$$A_T[\mathbf{x}(\cdot)] \leq A_T[\mathbf{y}(\cdot)]$$

for all $\mathbf{y}(0) = \mathbf{x}(0)$, $\mathbf{y}(T) = \mathbf{x}(t)$, we call $\mathbf{x}(\cdot)$ a *minimizer* of the action $A_T[\cdot]$ on the time interval $[0,T]$.

THEOREM 2.1 (Velocity estimate) *For each $T > 0$, there exists a constant C_T such that*

$$\max_{0 \leq t \leq T} |\dot{\mathbf{x}}(t)| \leq C_T \tag{2.1}$$

for each minimizer $\mathbf{x}(\cdot)$ on $[0,T]$.

Idea of Proof. This is a fairly standard derivative estimate for solutions of the Euler–Lagrange equation (E-L). □

2.2 Lax–Oleinik semigroup

DEFINITION. Let $v \in C(\mathbb{T}^n)$ and set

$$T_t^- v(x) := \inf\{v(\mathbf{x}(0)) + \int_0^t L(\dot{\mathbf{x}}, \mathbf{x})\, ds \mid \mathbf{x}(t) = x\}$$

We call the family of nonlinear operators $\{T_t^-\}_{t \geq 0}$ the *Lax–Oleinik semigroup.*

REMARK. The infimum above is attained. So in fact there exists a curve $\mathbf{x}(\cdot)$ such that $\mathbf{x}(t) = x$ and

$$T_t^- v(x) = v(\mathbf{x}(0)) + \int_0^t L(\dot{\mathbf{x}}, \mathbf{x})\, dt.$$

$\square$

THEOREM 2.2 (Regularization) *For each time $t > 0$ there exists a constant C_t such that*

$$[T_t^- v]_{Lip} \le C_t \tag{2.2}$$

for all $v \in C(\mathbb{T}^n)$.

Idea of Proof. Exploit the strict convexity of $v \mapsto L(v, x)$, as in similar arguments in Chapter 3 on my PDE text [E1]. $\square$

We record below some properties of the Lax–Oleinik semigroup acting on the space $C(\mathbb{T}^n)$, with max norm $\| \ \|$.

THEOREM 2.3 (Properties of Lax–Oleinik semigroup)

(i) $T_t^- \circ T_s^- = T_{t+s}^-$ (semigroup property).

(ii) $v \le \hat{v}$ *implies* $T_t^- v \le T_t^- \hat{v}$.

(iii) $T_t^-(v + c) = T_t^- v + c$.

(iv) $\|T_t^- v - T_t^- \hat{v}\| \le \|v - \hat{v}\|$ (nonexpansiveness).

(v) *For all* $v \in C(\mathbb{T}^n)$, $\lim_{t \to 0} T_t^- v = v$ *uniformly.*

(vi) *For all* v, $t \mapsto T_t^- v$ *is uniformly continuous.*

2.3 The Weak KAM Theorem

We begin with an abstract theorem about nonlinear mappings on a Banach space X:

THEOREM 2.4 (Common fixed points) *Suppose $\{\phi_t\}_{t \ge 0}$ is a semigroup of nonexpansive mappings of X into itself. Assume also for all $t > 0$ that $\phi_t(X)$ is precompact in X and for all $x \in X$ that $t \mapsto \phi_t(x)$ is continuous.*

Then there exists a point x^* such that

$$\phi_t(x^*) = x^* \qquad \text{for all times } t \ge 0.$$

Idea of Proof. Fix a time $t_0 > 0$. We will first show that ϕ_{t_0} has a fixed point.

Let $0 < \lambda < 1$. Then there exists according to the Contraction Mapping Theorem a unique element x_λ satisfying

$$\lambda \phi_{t_0}(x_\lambda) = x_\lambda.$$

By hypothesis $\{\phi_{t_0}(x_\lambda) \mid 0 < \lambda < 1\}$ is precompact. So there exist $\lambda_j \to 1$ for which $x_{\lambda_j} \to x$ and

$$\phi_{t_0}(x) = x.$$

Now let x_n be fixed point of $\phi_{1/2^n}$. Then x_n is fixed point also of $\phi_{k/2^n}$ for $k = 1, 2, \ldots$. Using compactness, we show that $x_{n_j} \to x^*$ and that $\phi_t(x^*) = x^*$ for all $t \geq 0$. $\square$

Next is an important result of Albert Fathi [F1]:

THEOREM 2.5 (Weak KAM Theorem) *There exists a function $v_- \in C(\mathbb{T}^n)$ and a constant $c \in \mathbb{R}$ such that*

$$T_t^- v_- + ct = v_- \qquad \textit{for all } t \geq 0. \tag{2.3}$$

Idea of proof. We apply Theorem 2.4 above. For this, let us write for functions $v, \hat{v} \in C(\mathbb{T}^n)$

$$v \sim \hat{v}$$

if $v - \hat{v} \equiv$ constant; and define also the equivalence class

$$[v] := \{\hat{v} \mid v \sim \hat{v}\}.$$

Set

$$X := \{[v] \mid v \in C(\mathbb{T}^n)\}$$

with the norm

$$\|[v]\| := \min_{a \in \mathbb{R}} \|v + a\mathbb{1}\|,$$

where $\mathbb{1}$ denotes is the constant function identically equal to 1.

We have $T_t : X \to X$, according to property (iii) of Theorem 2.3. Hence there exists a common fixed point

$$T_t[v]^* = [v]^* \qquad (t \geq 0).$$

Selecting any representative $v_- \in [v]^*$, we see that

$$T_t v_- = v_- + c(t)$$

for some $c(t)$. The semigroup property implies $c(t + s) = c(t) + c(s)$, and consequently $c(t) = ct$ for some constant c. $\square$

2.4 Domination

NOTATION. Let $w \in C(\mathbb{T}^n)$. We write

$$w \prec L + c \qquad (\text{“}w \text{ is dominated by } L + c\text{”})$$

if

$$w(\mathbf{x}(b)) - w(\mathbf{x}(a)) \leq \int_a^b L(\dot{\mathbf{x}}, \mathbf{x}) ds + c(b-a) \tag{2.4}$$

for all times $a < b$ and for all Lipschitz continuous curves $\mathbf{x} : [a, b] \to \mathbb{R}^n$.

REMARK.

$$w \prec L + c \text{ if and only if } w \leq ct + T_t^- w \text{ for all } t \geq 0.$$

□

THEOREM 2.6 (Domination and PDE)

(i) *If $w \prec L + c$ and the gradient $Dw(x)$ exists at a point $x \in \mathbb{T}^n$, then*

$$H(Dw(x), x) \leq c. \tag{2.5}$$

(ii) *Conversely, if w is Lipschitz continuous and $H(Dw, x) \leq c$ a.e., then*

$$w \prec L + c.$$

Idea of proof. (i) Select any curve $\mathbf{x}$ with $\mathbf{x}(0) = x$, $\dot{\mathbf{x}}(0) = v$. Then

$$\frac{w(\mathbf{x}(t)) - w(x)}{t} \leq \frac{1}{t} \int_0^t L(\dot{\mathbf{x}}, \mathbf{x}) \, ds + c.$$

Let $t \to 0$, to discover

$$v \cdot Dw \leq L(v, x) + c;$$

and therefore

$$H(Dw(x), x) = \max_v (v \cdot Dw - L(v, x)) \leq c.$$

(ii) If w is smooth, we can compute

$$\begin{aligned} w(\mathbf{x}(b)) - w(\mathbf{x}(a)) &= \int_a^b \frac{d}{dt} w(\mathbf{x}(t)) \, ds \\ &= \int_a^b Dw(\mathbf{x}) \cdot \dot{\mathbf{x}} \, dt \\ &\leq \int_a^b L(\dot{\mathbf{x}}, \mathbf{x}) + H(Dw(\mathbf{x}), \mathbf{x}) \, dt \\ &\leq \int_a^b L(\dot{\mathbf{x}}, \mathbf{x}) \, dt + c(b-a). \end{aligned}$$

See Fathi's book [F5] for what to do when w is only Lipschitz. $\square$

2.5 Flow invariance, characterization of the constant c

NOTATION. (i) We will hereafter write

$$T(\mathbb{T}^n) := \mathbb{R}^n \times \mathbb{T}^n = \{(v,x) \mid v \in \mathbb{R}^n,\ x \in \mathbb{T}^n\}$$

for the *tangent bundle* over $\mathbb{T}^n$, and

$$T^*(\mathbb{T}^n) := \mathbb{R}^n \times \mathbb{T}^n = \{(p,x) \mid p \in \mathbb{R}^n,\ x \in \mathbb{T}^n\}$$

for the *cotangent bundle*. We work in the tangent bundle for the Lagrangian viewpoint, and in the cotangent bundle for the Hamiltonian viewpoint.

(ii) Consider this initial-value problem for the Euler-Lagrange equation:

$$\begin{cases} -\frac{d}{dt}(D_v L(\dot{\mathbf{x}},\mathbf{x})) + D_x L(\dot{\mathbf{x}},\mathbf{x}) = 0 \\ \mathbf{x}(0) = x,\ \dot{\mathbf{x}}(0) = v. \end{cases} \tag{2.6}$$

We define the *flow map* $\{\boldsymbol{\phi}_t\}_{t\in\mathbb{R}}$ on $T(\mathbb{T}^n)$ by the formula

$$\boldsymbol{\phi}_t(v,x) := (\mathbf{v}(t), \mathbf{x}(t)), \tag{2.7}$$

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

DEFINITION. A probability measure μ on the tangent bundle $T(\mathbb{T}^n)$ is *flow invariant* if

$$\int_{T(\mathbb{T}^n)} \Phi(\boldsymbol{\phi}_t(v,x))\, d\mu = \int_{T(\mathbb{T}^n)} \Phi(v,x)\, d\mu$$

for each bounded continuous function Φ.

Following is an elegant interpretation of the constant c from the Weak KAM Theorem 2.5, in terms of action minimizing flow invariant measures:

THEOREM 2.7 (Characterization of c) *The constant from Theorem 2.5 is given by the formula*

$$c = -\inf\{\int_{T(\mathbb{T}^n)} L(v,x)\, d\mu \mid \mu \text{ flow invariant, probability measure}\}. \tag{2.8}$$

DEFINITIONS. (i) The *action* of a flow-invariant measure μ is

$$A[\mu] := \int_{T(\mathbb{T}^n)} L(v,x)\, d\mu.$$

(ii) We call μ a *minimizing* (or *Mather*) *measure* if

$$-c = A[\mu]. \tag{2.9}$$

Idea of proof. 1. Recall that

$$T_t^- v_- = v_- + ct \quad \text{and} \quad v_- \prec L + c.$$

Let us write

$$\boldsymbol{\phi}_s(v,x) = (\dot{\mathbf{x}}(s), \mathbf{x}(s))$$

for $(v,x) \in T(\mathbb{T}^n)$, and put

$$\pi(v,x) := x$$

for the projection of $T(\mathbb{T}^n)$ onto $\mathbb{T}^n$.

Then

$$v_-(\pi\boldsymbol{\phi}_1(v,x)) - v_-(\pi(v,x)) \leq \int_0^1 L(\boldsymbol{\phi}_s(v,x))\, ds + c. \tag{2.10}$$

Integrate with respect to a flow-invariant probability measure μ:

$$\begin{aligned} 0 &= \int_{T(\mathbb{T}^n)} v_-(\pi\boldsymbol{\phi}_1(v,x)) - v_-(\pi(v,x))\, d\mu \\ &\leq \int_0^1 \int_{T(\mathbb{T}^n)} L(\boldsymbol{\phi}_s(v,x))\, d\mu ds + c \\ &= \int_{\mathbb{T}(\mathbb{T}^n)} L(v,x)\, d\mu + c. \end{aligned}$$

Therefore

$$-c \leq \int_{T(\mathbb{T}^n)} L(v,x)\, d\mu \qquad \text{for all flow-invariant } \mu.$$

2. We must now manufacture a measure giving equality above. Fix $x \in \mathbb{T}^n$. Find a curve $\mathbf{x} : (-\infty, 0] \to \mathbb{R}^n$ such that $\mathbf{x}(0) = x$, $\dot{\mathbf{x}}(0) = v$; and for all times $t \leq 0$:

$$v_-(\mathbf{x}(0)) - v_-(\mathbf{x}(t)) = \int_t^0 L(\boldsymbol{\phi}_s(v,x))\, ds - ct.$$

Define for $t \geq 0$ the measure μ_t by the rule

$$\mu_t(\Phi) = \frac{1}{t}\int_{-t}^{0} \Phi(\boldsymbol{\phi}_s(v,x))\, ds$$

for each continuous function Φ. Then

$$\frac{v_-(x) - v_-(\mathbf{x}(-t))}{t} = \int_{T(\mathbb{T}^n)} L\, d\mu_t + c.$$

Sending $t_k \to -\infty$, we deduce that

$$d\mu_{t_k} \rightharpoonup d\mu$$

weakly as measures, for some flow invariant probability measure μ that satisfies

$$-c = \int_{T(\mathbb{T}^n)} L\, d\mu.$$

$\square$

2.6 Time-reversal, Mather set

It is sometimes convenient to redo our theory with time reversed, by introducing the *backwards Lax–Oleinik semigroup*:

DEFINITION.

$$T_t^+ v(s) := \sup\{v(\mathbf{x}(t)) - \int_0^t L(\dot{\mathbf{x}}, \mathbf{x}) ds \mid \mathbf{x}(0) = x\}.$$

As above, there exists a function $v_+ \in C(\mathbb{T}^n)$ such that

$$T_t^+ v_+ - ct = v_+ \qquad \text{for all } t \geq 0$$

for the same constant c described in Theorem 2.7.

DEFINITION. We define the *Mather set*.

$$\tilde{M}_0 := \overline{\bigcup_{\mu} \operatorname{spt}(\mu)},$$

the union over all minimizing measures μ as above. The *projected Mather set* is

$$M_0 := \pi(\tilde{M}_0).$$

One goal of weak KAM theory is studying the structure of the Mather set (and the related Aubry set), in terms of the underlying Hamiltonian dynamics. See Fathi [F5] for much more.

3 Weak KAM Theory: Hamiltonian and PDE methods

3.1 Hamilton–Jacobi PDE

In this section, we reinterpret the foregoing ideas in terms of the theory of *viscosity solutions* of nonlinear PDE:

THEOREM 3.1 (Viscosity solutions)

(i) *We have*

$$H(Dv_{\pm}, x) = c \qquad a.e.$$

(ii) *In fact*

$$\begin{cases} H(Dv_-, x) = c \\ -H(Dv_+, x) = -c \end{cases} \tag{3.1}$$

in sense of viscosity solutions.

Idea of Proof. 1. There exists a minimizing curve $\mathbf{x} : (\infty, 0] \to \mathbb{R}^n$ such that $\mathbf{x}(0) = x$, $\dot{\mathbf{x}}(0) = v$, and

$$v_-(x) = v_-(\mathbf{x}(-t)) + \int_{-t}^{0} L(\dot{\mathbf{x}}, \mathbf{x})\, ds + ct \qquad (t \geq 0).$$

If v is differentiable at x, then we deduce as before that

$$v \cdot Dv_-(x) = L(v, x) + c,$$

and this implies $H(Dv_-, x) \geq c$. But always $H(Dv_-, x) \leq 0$.

2. The assertions about $v_{\pm}$ being viscosity solutions follow as in Chapter 10 of my PDE book [E1]. □

3.2 Adding P dependence

Motivated by the discussion in Section 1 about the classical theory of canonical transformation to action-angle variables, we next explicitly add dependence on a vector P.

So select $P \in \mathbb{R}^n$ and define the *shifted Lagrangian*

$$\hat{L}(v, x) := L(v, x) - P \cdot v. \tag{3.2}$$

The corresponding Hamiltonian is

$$\hat{H}(p, x) = \max_v (p \cdot v - \hat{L}(v, x)) = \max_v ((p + P) \cdot v - L(v, x)),$$

and so

$$\hat{H}(p,x) = H(P+p,x). \tag{3.3}$$

As above, we find a constant $c(P)$ and periodic functions $v_\pm = v_\pm(P,x)$ so that

$$\begin{cases} H(P+Dv_-,x) = \hat{H}(Dv_-,x) = c(P) \\ -H(P+Dv_+,x) = -\hat{H}(Dv_+,x) = -c(P) \end{cases}$$

in the viscosity sense.

NOTATION. We write

$$\bar{H}(P) := c(P), \quad u_- := P\cdot x + v_-, \quad u_+ := P\cdot x + v_+.$$

Then

$$\begin{cases} H(Du_-,x) = \bar{H}(P) \\ -H(Du_+,x) = -\bar{H}(P) \end{cases} \tag{3.4}$$

in the sense of viscosity solutions.

3.3 Lions–Papanicolaou–Varadhan Theory

3.3.1 A PDE construction of $\bar{H}$. We next explain an alternative, purely PDE technique, due to Lions–Papanicolaou–Varadhan [L-P-V], for finding the constant $c(P)$ as above and $v = v_-$.

THEOREM 3.2 (More on viscosity solutions) *For each vector $P \in \mathbb{R}^n$, there exists a unique real number $c(P)$ for which we can find a viscosity solution of*

$$\begin{cases} H(P+D_x v,x) = c(P) \\ \quad v \text{ is } \mathbb{T}^n\text{-periodic.} \end{cases} \tag{3.5}$$

REMARK. In addition v is semiconcave, meaning that $D^2 v \le CI$ for some constant C.
□

Idea of proof. 1. *Existence.* Routine viscosity solution methods assert the existence of a unique solution v^ε of

$$\varepsilon v^\varepsilon + H(P+D_x v^\varepsilon, x) = 0.$$

Since H is periodic in x, uniqueness implies v^ε is also periodic.

It is now not very difficult to derive the uniform estimates

$$\max |D_x v^\varepsilon|, |\varepsilon v^\varepsilon| \leq C.$$

Hence we may extract subsequences for which

$$v^{\varepsilon_j} \to v \text{ uniformly, } \varepsilon_j v^{\varepsilon_j} \to -c(P).$$

It is straightforward to confirm that v is a viscosity solution of $H(P + D_x v, x) = c(P)$.

2. *Uniqueness of* $c(P)$. Suppose also

$$H(P + D_x \hat{v}, x) = \hat{c}(P).$$

We may assume that $\hat{c}(P) > c(P)$ and that $\hat{v} < v$, upon adding a constant to v, if necessary. Then

$$\delta \hat{v} + H(P + D_x \hat{v}, x) > \delta v + H(P + D_x v, x)$$

in viscosity sense, if $\delta > 0$ is small. The viscosity solution comparison principle then implies the contradiction $\hat{v} \geq v$. □

NOTATION. In agreement with our previous notation, we write

$$\bar{H}(P) := c(P), \quad u = P \cdot x + v, \tag{3.6}$$

and call $\bar{H}$ the *effective Hamiltonian*. Then

$$H(D_x u, x) = \bar{H}(P). \tag{E}$$

We call (E) the *generalized eikonal* equation.

3.3.2 Effective Lagrangian. Dual to the effective Hamiltonian is the *effective Lagrangian*:

$$\bar{L}(V) := \max_P (P \cdot V - \bar{H}(P)) \tag{3.7}$$

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

We sometime write $\bar{L} = \bar{H}^*$ to record this Legendre transform. Then

$$P \in \partial \bar{L}(V) \quad \Leftrightarrow \quad V \in \partial \bar{H}(P) \quad \Leftrightarrow \quad P \cdot V = \bar{L}(V) + \bar{H}(P), \tag{3.8}$$

∂ denoting the (possibly multivalued) subdifferential of a convex function.

THEOREM 3.3 (Characterization of effective Lagrangian) *We have the formula*

$$\bar{L}(V) = \inf\{\int_{T(\mathbb{T}^n)} L(v,x)\,d\mu \mid \mu \textit{ flow invariant, } \int_{T(\mathbb{T}^n)} v\,d\mu = V\}. \tag{3.9}$$

Compare this with Theorem 2.7.

Proof. Denote by $\tilde{L}(V)$ the right hand side of (3.9). Then

$$\begin{aligned} -\bar{H}(P) &= \inf_{\mu}\{\int_{T(\mathbb{T}^n)} L(v,x) - P\cdot v\,d\mu\} \\ &= \inf_{\mu,V}\{\int_{T(\mathbb{T}^n)} L(v,x)\,d\mu - P\cdot V \mid \int_{T(\mathbb{T}^n)} v\,d\mu = V\} \\ &= \inf_{V}\{\tilde{L}(V) - P\cdot V\} \\ &= -\sup_{V}\{P\cdot V - \tilde{L}(V)\}. \end{aligned}$$

So $\tilde{L} = \bar{H}^* = \bar{L}$. □

3.3.3 Application: homogenization of nonlinear PDE

THEOREM 3.4 (Homogenization) *Suppose g is bounded and uniformly continous, and u^ε is the unique bounded, uniformly continuous viscosity solution of the initial-value problem*

$$\begin{cases} u^\varepsilon_t + H(Du^\varepsilon, \frac{x}{\varepsilon}) = 0 & (t>0) \\ u^\varepsilon = g & (t=0). \end{cases}$$

Then $u^\varepsilon \to u$ locally uniformly, where u solves the homogenized equation

$$\begin{cases} u_t + \bar{H}(Du) = 0 & (t>0) \\ u = g & (t=0). \end{cases}$$

Idea of Proof. Let ϕ be a smooth function and suppose that $u-\phi$ has a strict maximum at the point (x_0,t_0). Define the perturbed test function

$$\phi^\varepsilon(x,t) := \phi(x,t) + \varepsilon v\left(\frac{x}{\varepsilon}\right),$$

where v is a periodic viscosity solution of

$$H(P+Dv,x) = \bar{H}(P)$$

for $P = D\phi(x_0, t_0)$.

Assume for the rest of the discussion that v is smooth. Then ϕ^ε is smooth and $u^\varepsilon - \phi^\varepsilon$ attains a max at a point $(x_\varepsilon, t_\varepsilon)$ near (x_0, t_0). Consequently

$$\phi_t^\varepsilon + H\left(D\phi^\varepsilon, \frac{x_\varepsilon}{\varepsilon}\right) \leq 0.$$

And then

$$\phi_t + H\left(D\phi(x_\varepsilon, t_\varepsilon) + Dv\left(\frac{x_\varepsilon}{\varepsilon}\right), \frac{x_\varepsilon}{\varepsilon}\right) \approx \phi_t + \bar{H}(D\phi(x_0, t_0)) \leq 0.$$

The reverse inequality similarly holds if $u - \phi$ has a strict minimum at the point (x_0, t_0).

See my old paper [E2] for what to do when v is not smooth. □

3.4 More PDE methods

In this section we apply some variational and nonlinear PDE methods to study further the structure of the Mather minimizing measures μ and viscosity solutions $u = P \cdot x + v$ of the eikonal equation (E).

Hereafter μ denotes a Mather minimizing measure in $T(\mathbb{T}^n)$. It will be more convenient to work with Hamiltonian variables, and so we define ν to be the pushforward of μ onto the contangent bundle $T^*(\mathbb{T}^n)$ under the change of variables $p = D_v L(v, x)$.

For reference later, we record these properties of ν, inherited from μ:

(A) $$V = \int_{T^*(\mathbb{T}^n)} D_p H(v, x)\, d\nu$$

(B) $$\bar{L}(V) = \int_{T^*(\mathbb{T}^n)} L(D_p H(p, x), x)\, d\nu$$

(C) $$\int_{T^*(\mathbb{T}^n)} \{H, \Phi\}\, d\nu = 0 \qquad \text{for all } C^1 \text{ functions } \Phi,$$

where

$$\{H, \Phi\} := D_p H \cdot D_x \Phi - D_x H \cdot D_P \Phi.$$

is the *Poisson bracket*. Statement (A) is the definition of V, whereas statement (C) is a differential form of the flow-invariance, in the Hamiltonian variables.

NOTATION. We will write

$$\sigma = \pi_\# \nu$$

for the push-forward of ν onto $\mathbb{T}^n$.

Now select any $P \in \partial \bar{L}(V)$. Then as above construct a viscosity solution of the generalized eikonal PDE

$$\text{(E)} \qquad H(D_x u, x) = \bar{H}(P)$$

for

$$u = P \cdot x + v, \ v \text{ periodic}.$$

We now study properties of u in relation to the measures σ and ν, following the paper [E-G1].

THEOREM 3.5 (Regularity properties)
(i) *The function u is differentiable for σ a.-e. point $x \in \mathbb{T}^n$.*
(ii) *We have*

$$\text{(3.10)} \qquad p = D_x u \quad \nu - a.e.$$

(iii) *Furthermore,*

$$\text{(3.11)} \qquad \bar{H}(P) = \int_{T^*(\mathbb{T}^n)} H(p, x)\, d\nu.$$

REMARK. Compare assertion (ii) with the first of the classical formulas (1.5) □
.

Idea of proof. 1. Since $p \mapsto H(p, x)$ is uniformly convex, meaning $D_p^2 H \geq \gamma I$ for some positive γ, we have

$$\text{(3.12)} \qquad \beta_\varepsilon(x) + H(Du^\varepsilon, x) \leq \bar{H}(P) + O(\varepsilon)$$

for the mollified function $u^\varepsilon := \eta_\varepsilon * u$ and

$$\begin{aligned} \beta_\varepsilon(x) &= \frac{\gamma}{2} \int_{\mathbb{T}^n} \eta_\varepsilon(x - y) |Du(y) - Du^\varepsilon(x)|^2 dy \\ &\approx \frac{\gamma}{2} \fint_{B(x,\varepsilon)} |Du - (Du)_{x,\varepsilon}|^2 dy. \end{aligned}$$

In this formula $(Du)_{x,\varepsilon}$ denotes the average of Du over the ball $B(x, \varepsilon)$.

2. Uniform convexity also implies

$$\frac{\gamma}{2}\int_{T^*(\mathbb{T}^n)} |Du^\varepsilon - p|^2 d\nu \le \int_{T^*(\mathbb{T}^n)} H(Du^\varepsilon, x) - H(p,x) - D_pH(p,x)\cdot(Du^\varepsilon - p)\, d\nu.$$

Now $Du^\varepsilon = P + Dv^\varepsilon$ and

$$\int_{T^*(\mathbb{T}^n)} D_pH \cdot Dv^\varepsilon d\nu = 0 \qquad \text{by property (C).}$$

So (3.12) implies

$$\begin{aligned}\frac{\gamma}{2}\int_{T^*(\mathbb{T}^n)} |Du^\varepsilon - p|^2 d\nu \;&\le\; \int_{T^*(\mathbb{T}^n)} \bar{H}(P) - H(p,x) - D_pH(p,x)\cdot(P-p) - \beta_\varepsilon(x)\, d\nu \\ &\qquad + O(\varepsilon) \\ &\le\; \bar{H}(P) - V\cdot P - \int_{T^*(\mathbb{T}^n)} \bar{H}(p,x) - D_pH(p,x)\cdot p - \beta_\varepsilon(x)\, d\nu \\ &\qquad + O(\varepsilon) \\ &=\; \bar{L}(V) - \int_{T^*(\mathbb{T}^n)} L(D_pH, x) - \beta_\varepsilon(x)\, d\nu + O(\varepsilon),\end{aligned}$$

according to (3.8). Consequently

$$\int_{T^*(\mathbb{T}^n)} |Du^\varepsilon - p|^2\, d\nu + \int_{\mathbb{T}^n} \beta_\varepsilon(x)\, d\sigma \le O(\varepsilon). \tag{3.13}$$

Recall also that u is semiconcave. Therefore (3.13) implies σ-a.e. point is Lebesgue point of Du, and so Du exists σ-a.e. It then follows from (3.13) that $Du^\varepsilon \to Du$ ν-a.e.

2. We have $\int_{T^*(\mathbb{T}^n)} H(p,x)\, d\nu = \int_{\mathbb{T}^n} H(Du,x)\, d\sigma = \bar{H}(P)$. □

REMARK. In view of (3.10), the flow invariance condition (C) implies

$$\int_{\mathbb{T}^n} D_pH(Du,x)\cdot D\phi\, d\sigma = 0$$

for all $\phi \in C^1(\mathbb{T}^n)$; and so the measure σ is a weak solution of the *generalized transport (*or *continuity) equation*

$$\operatorname{div}(\sigma D_pH(D_xu, x)) = 0. \tag{T}$$

□

3.5 Estimates.

Now we illustrate how our two key PDE, the generalized eikonal equation (E) for u and the generalized transport equation (T) for σ and u, together yield more information about the smoothness of u.

To simplify the presentation, we take the standard example

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

for this section. Then our eikonal and transport equations become

$$\text{(E)} \qquad \frac{1}{2}|Du|^2 + W(x) = \bar{H}(P)$$

$$\text{(T)} \qquad \operatorname{div}(\sigma Du) = 0.$$

If u is smooth, we could differentiate (E) twice with respect to x_k:

$$\sum_{i=1}^{n} u_{x_i} u_{x_i x_k x_k} + u_{x_i x_k} u_{x_i x_k} + W_{x_k x_k} = 0,$$

sum on k, and then integrate with respect to σ over $\mathbb{T}^n$:

$$\int_{\mathbb{T}^n} D_x u \cdot D_x(\Delta u)\, d\sigma + \int_{\mathbb{T}^n} |D_x^2 u|^2 d\sigma = -\int_{\mathbb{T}^n} \Delta W\, d\sigma.$$

According to (T) the first term equals zero.

This establishes the formal estimate

$$\text{(3.14)} \qquad \int_{\mathbb{T}^n} |D_x^2 u|^2\, d\sigma \le C;$$

and a related rigorous estimate involving difference quotients holds if u is not smooth: see [E-G1].

Reworking the proof using appropriate cutoff functions, we can derive as well the formal bound

$$\text{(3.15)} \qquad |D_x^2 u|^2 \le C \quad \sigma \text{ a.e.};$$

and a related rigorous estimate involving difference quotients is valid if u is not smooth. Again, see [E-G1] for the details. We thereby establish the inequality

$$\text{(3.16)} \qquad |Du(y) - Du(x)| \le C|x-y| \quad \text{for } x \in \operatorname{spt}(\sigma), \text{ a.e. } y \in \mathbb{T}^n.$$

In particular, even though Du may be multivalued, we can bound

$$\operatorname{diam}(Du(y)) \leq C \operatorname{dist}(y, \operatorname{spt}(\sigma))$$

for some constant C. This is a sort of quantitative estimate on how far the support of σ lies from the "shocks" of the gradient of u.

An application of these estimates is a new proof of Mather's regularity theorem for the support of the minimizing measures:

THEOREM 3.6 (Mather) *The support of μ lies on a Lipschitz graph in $T(\mathbb{T}^n)$, and the support of ν lies on a Lipschitz graph in $T^*(\mathbb{T}^n)$.*

REMARK. In addition, if u is smooth in x and P, we have the formal bound

$$\int_{\mathbb{T}^n} |D^2_{xP} u|^2 \, d\sigma \leq C. \tag{3.17}$$

A related rigorous estimate involving difference quotients holds if u is not smooth. As an application, we show in [E-G1] that if $\bar{H}$ is twice differentiable at P, then

$$|D\bar{H}(P) \cdot \xi| \leq C(\xi \cdot D^2\bar{H}(P)\xi)^{1/2}.$$

for all vectors ξ and some constant C. □

4 An alternative variational/PDE construction

4.1 A new variational formulation.

This section follows [E3], to discuss an alternate variational/PDE technique for discovering the structure of weak KAM theory.

4.1.1 A minimax formula. Our motivation comes from "the calculus of variations in the sup-norm", as presented for instance in Barron [B]. We start with the following observation, due to several authors:

THEOREM 4.1 (Minimax formula for $\bar{\mathbf{H}}$) *We have*

$$\bar{H}(P) = \inf_{v \in C^1(\mathbb{T}^n)} \max_{x \in \mathbb{T}^n} H(P + Dv, x). \tag{4.1}$$

Idea of proof. Write $u = P \cdot x + v$, where u is our viscosity solution of (E), and put $\hat{u} = P \cdot x + \hat{v}$, where $\hat{v}$ is any C^1, periodic function. Then convexity implies

$$H(Du, x) + D_p H(Du, x) \cdot (D\hat{u} - Du) \leq H(D\hat{u}, x).$$

Integrate with respect to σ:

$$\begin{aligned} \bar{H}(P) + \int_{\mathbb{T}^n} D_p H(Du, x) \cdot D(\hat{v} - v)\, d\sigma &\leq \int_{\mathbb{T}^n} H(D\hat{u}, x)\, d\sigma \\ &\leq \max_x H(P + D\hat{v}, x). \end{aligned}$$

Therefore (T) implies that

$$\bar{H}(P) \leq \inf_{\hat{v}} \max_x H(P + D\hat{v}, x).$$

Furthermore, if $v^\varepsilon := \eta^\varepsilon * v$, where η^ε is a standard mollifier, then

$$H(P + Dv^\varepsilon, x) \leq \bar{H}(P) + O(\varepsilon);$$

and consequently

$$\max_x H(P + Dv^\varepsilon, x) \leq \bar{H}(P) + O(\varepsilon).$$

□

REMARK. See Fathi–Siconolfi [F-S1] for the construction of a C^1 subsolution. □

4.1.2 A new variational setting. The minimax formula provided by Theorem 4.1 suggests that we may be able somehow to approximate the "max" above by an exponential integral with a large parameter.

To be precise, we fix a large constant k and then introduce the integral functional

$$I_k[v] := \int_{\mathbb{T}^n} e^{kH(P+Dv,x)}\, dx.$$

(My paper [E3] cites some relevant references, and see also Marcellini [M1], [M2] and Mascolo–Migliorini [M-M] for more about such problems with exponential growth.)

Let v_k be the unique minimizer among perodic functions, subject to the normalization that

$$\int_{\mathbb{T}^n} v^k\, dx = 0.$$

As usual, put $u^k := P \cdot x + v^k$. The Euler–Lagrange equation then reads

$$\text{div}(e^{kH(Du^k,x)} D_p H(Du^k, x)) = 0. \tag{4.2}$$

4.1.3 Passing to limits. We propose to study the asymptotic limit of the PDE (4.2) as $k \to \infty$. We will discover that the structure of weak KAM theory appears in the limit.

NOTATION. It will be convenient to introduce the normalizations

$$\sigma^k := \frac{e^{kH(Du^k,x)}}{\int_{\mathbb{T}^n} e^{kH(Du^k,x)}dx} \tag{4.3}$$

and

$$\bar{H}^k(P) := \frac{1}{k}\log(\int_{\mathbb{T}^n} e^{kH(Du^k,x)}dx). \tag{4.4}$$

Define also

$$d\mu^k := \delta_{\{v=D_pH(Du^k,x)\}}\sigma^k\, dx.$$

Passing as necessary to subsequences, we may assume

$$\begin{aligned} \sigma^k dx \rightharpoonup d\sigma &\quad \text{weakly as measures on } \mathbb{T}^n, \\ d\mu^k \rightharpoonup d\mu &\quad \text{weakly as measures on } T(\mathbb{T}^n), \\ u^k \to u &\quad \text{uniformly.} \end{aligned}$$

A main question is what PDE (if any) do u and σ satisfy?

THEOREM 4.2 (Weak KAM in the limit)

(i) *We have*

$$\bar{H}(P) = \lim_{k\to\infty} \bar{H}^k(P). \tag{4.5}$$

(ii) *The measure μ is a Mather minimizing measure.*

(iii) *Furthermore*

$$\begin{cases} H(Du,x) \le \bar{H}(P) & a.e. \\ H(Du,x) = \bar{H}(P) & \sigma\text{-}a.e. \end{cases} \tag{4.6}$$

(iv) *The measure σ is a weak solution of*

$$div(\sigma D_pH(Du,x)) = 0. \tag{4.7}$$

(v) *In addition, u is a viscosity solution of* Aronsson's PDE

$$-A_H[u] := -\sum_{i,j=1}^{n} H_{p_i}H_{p_j}u_{x_ix_j} - \sum_{i=1}^{n} H_{x_i}H_{p_i} = 0. \tag{4.8}$$

4.2 Application: nonresonance and averaging.

We illustrate some uses of the approximation (4.2) by noting first that our (unique) solution $u^k := P{\cdot}x+v^k$ is smooth in both x and P. We can therefore legitimately differentiate in both variables, and are encouraged to do so by the classical formulas (1.5).

4.2.1 Derivatives of $\bar{\mathbf{H}}^k$. We can also calculate the first and second derivatives in P of the smooth approximate effective Hamiltonian $\bar{H}^k$. Indeed, a direct computation (cf. [E3]) establishes the useful formulas

$$D\bar{H}^k(P) = \int_{\mathbb{T}^n} D_pH(Du^k, x)\, d\sigma^k \tag{4.9}$$

and

$$\begin{aligned} D^2\bar{H}^k(P) = k\int_{\mathbb{T}^n} &(D_pH(Du^k, x)D^2_{xP}u^k - D\bar{H}^k(P)) \otimes (D_pHD^2_{xP}u^k - D\bar{H})\, d\sigma^k \\ &+ \int_{\mathbb{T}^n} D^2_pH(Du^k, x)D^2_{xP}u \otimes D^2_{xP}u\, d\sigma^k, \end{aligned} \tag{4.10}$$

where for notational convenience we write

$$d\sigma^k := \sigma^k dx.$$

4.2.2 Nonresonance. Next, let us assume $\bar{H}$ is differentiable at P and that $V = D\bar{H}(P)$ satisfies the weak nonresonance condition

$$V \cdot m \neq 0 \quad \text{for all } m \in \mathbb{Z}^n,\ m \neq 0. \tag{4.11}$$

What does this imply about the limits as $k \to \infty$ of σ^k and u^k??

THEOREM 4.3 (Nonresonance and averaging) *Suppose* (4.11) *holds and also*

$$|D^2\bar{H}^k(P)| \leq C$$

for all large k.

Then

$$\lim_{k\to\infty} \int_{\mathbb{T}^n} \Phi(D_Pu^k)\sigma^k\, dx = \int_{\mathbb{T}^n} \Phi\, dX. \tag{4.12}$$

for all continuous, periodic functions $\Phi = \Phi(X)$.

INTERPRETATION. As discussed in Section 1, if we could really change to the action-angle variables X, P according to (1.5) and (1.6), we would obtain the linear dynamics

$$\mathbf{X}(t) = X_0 + tV = X_0 + tD\bar{H}(P).$$

In view of the nonresonance condition (4.11), it follows then that that

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \Phi(\mathbf{X}(t))\,dt = \int_{\mathbb{T}^n} \Phi(X)\,dX \tag{4.13}$$

for all continuous, periodic functions Φ.

Observe next from (4.2) that σ^k is a solution of the transport equation

$$\operatorname{div}(\sigma^k D_pH(Du^k, x)) = 0. \tag{4.14}$$

Now a direct calculation shows that if u is a smooth solution of (E), then

$$\hat{\sigma} := \det D^2_{xP} u$$

in fact solves the same transport PDE:

$$\operatorname{div}(\hat{\sigma} D_pH(Du, x)) = 0.$$

This suggests that maybe we somehow have $\sigma^k \approx \det D^2_{xP}u^k$ in the asymptotic limit $k \to \infty$. Since

$$\int_{\mathbb{T}^n} \Phi(D_Pu)\hat{\sigma}\,dx = \int_{\mathbb{T}^n} \Phi(X)\,dX$$

if the mapping $X = D_Pu(P, x)$ is one-to-one and onto, we might then hope that something like (4.12) is valid. □

Idea of proof. Recall (4.14) and note also that

$$w := e^{2\pi i m\cdot D_Pu^k} = e^{2\pi i m\cdot(x+D_Pv^k)}$$

is periodic (even though $D_Pu^k = x + D_Pv^k$ is not). Hence

$$\begin{aligned}
0 &= \int_{\mathbb{T}^n} D_pH(Du^k, x)\cdot D_xw\sigma^k\,dx \\
&= 2\pi i\int_{\mathbb{T}^n} e^{2\pi i m\cdot D_Pu^k} m\cdot D_pHD^2_{xP}u^k\sigma^k\,dx \\
&= 2\pi i\int_{\mathbb{T}^n} e^{2\pi i m\cdot D_Pu^k} m\cdot D\bar{H}^k(P)\sigma^k\,dx \\
&\qquad +2\pi i\int_{\mathbb{T}^n} e^{2\pi i m\cdot D_Pu^k} m(D_pHD_{xP}u^k - D\bar{H}^k)\sigma^k\,dx \\
&=: A + B.
\end{aligned}$$

Using formula (4.10), we can estimate

$$|B| \le C(\frac{1}{k}|D^2\bar{H}^k(P)|)^{1/2} = o(1) \qquad \text{as } k \to \infty.$$

Consequently $A \to 0$. Since

$$m \cdot D\bar{H}^k(P) \to m \cdot V \neq 0,$$

we deduce that

$$\int_{\mathbb{T}^n} e^{2\pi i m \cdot D_P u^k} \sigma^k \, dx \to 0.$$

This proves (4.12) for all finite trigonometric polynomials Φ, and then, by the density of such trig polynomials, for all continuous Φ. □

5 Some other viewpoints and open questions

This concluding section collects together some comments about other work on, and extending, weak KAM theory and about possible future progress.

• **Geometric properties of the effective Hamiltonian**.

Concordel in [C1], [C2] initiated the systematic study of the geometric properties of the effective Hamiltonian $\bar{H}$, but many questions are still open.

Consider, say, the basic example $H(p,x) = \frac{|p|^2}{2} + W(x)$ and ask how the geometric properties of the periodic potential W influence the geometric properties of $\bar{H}$, and vice versa. For example, if we know that $\bar{H}$ has a "flat spot" at its minimum, what does this imply about W?

It would be interesting to have some more careful numerical studies here, as for instance in Gomes-Oberman [G-O].

• **Nonresonance**.

Given the importance of nonresonance assumptions for the perturbative classical KAM theory, it is a critically important task to understand consequences of this condition for weak KAM in the large. Theorem 4.3 is a tiny step towards understanding this fundamental issue. See also Gomes [G1], [G2] for some more developments.

• **Aubry and Mather sets**.

See Fathi's book [F5] for a more detailed discussion of the Mather set (only mentioned above), the larger Aubry set, and their dynamical systems interpretations. One fundamental question is just how, and if, Mather sets can act as global replacements for the classical KAM invariant tori.

• **Weak KAM and mass transport**.

Bernard and Buffoni [B-B1], [B-B2], [B-B3] have rigorously worked out some of the fascinating interconnections between weak KAM theory and optimal mass transport theory (see Ambrosio [Am] and Villani [V]). Some formal relationships are sketched in my expository paper [E5].

• **Stochastic and quantum analogs**.

My paper [E4] discusses the prospects of finding some sort of quantum version of weak KAM, meaning ideally to understand possible connections with solutions of Schrödinger's equation in the semiclassical limit $h \to 0$. This all of course sounds good, but it is currently quite unclear if any nontrivial such connections really exist.

N. Anantharaman's interesting paper [An] shows that a natural approximation scheme (independently proposed in [E4]) gives rise to Mather minimizing measures which additionally extremize an entropy functional.

Gomes in [G3] and Iturriaga and Sanchez-Morgado in [I-SM] have discussed stochastic versions of weak KAM theory, but here too many key questions are open.

• **Nonconvex Hamiltonians**.

It is, I think, very significant that the theory [L-P-V] of Lions, Papanicolaou and Varadhan leads to the existence of solutions to the generalized eikonal equation (E) even if the Hamiltonian H is nonconvex in the momenta p: all that is really needed is the coercivity condition that $\lim_{|p|\to\infty} H(p,x) = \infty$, uniformly for $x \in \mathbb{T}^n$. In this case it remains a major problem to interpret $\bar{H}$ in terms of dynamics.

Fathi and Siconolfi [F-S2] have made great progress here, constructing much of the previously discussed theory under the hypothesis that $p \mapsto H(p,x)$ be *geometrically quasiconvex*, meaning that for for each real number λ and $x \in \mathbb{T}^n$, the sublevel set $\{p \mid H(p,x) \leq \lambda\}$ is convex.

The case of Hamiltonians which are coercive, but nonconvex and nonquasiconvex in p, is completely open.

• **Aronsson's PDE**.

I mention in closing one final mystery: does Aronsson's PDE (4.8) have anything whatsoever to do with the Hamiltonian dynamics? This highly degenerate, highly nonlinear elliptic equation occurs quite naturally from the variational construction in Section 4, but to my knowledge has no interpretation in terms of dynamical systems. (See Yu [Y] and Fathi–Siconolfi [F-S3] for more on this strange PDE.)

References

[Am] L. Ambrosio, Lecture notes on optimal transport problems, in *Mathematical Aspects of Wvolving Interfaces* 1–52, Lecture Notes in Mathematics 1812, Springer, 2003.

[An] N. Anantharaman, Gibbs measures and semiclassical approximation to action-minimizing measures, Transactions AMS, to appear.

[B] E. N. Barron, Viscosity solutions and analysis in L^∞, in *Nonlinear Analysis, Differential Equations and Control*, Dordrecht, 1999.

[B-B1] P. Bernard and B. Buffoni, The Monge problem for supercritical Mañé potentials on compact manifolds, to appear.

[B-B2] P. Bernard and B. Buffoni, Optimal mass transportation and Mather theory, to appear.

[B-B3] P. Bernard and B. Buffoni, The Mather-Fathi duality as limit of optimal transportation problems, to appear.

[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 homogenization of Hamilton-Jacobi equations II: eikonal equations, Proc. Roy. Soc. Edinburgh 127 (1997), 665–689.

[E1] L. C. Evans, *Partial Differential Equations*, American Mathematical Society, 1998. Third printing, 2002.

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

[E3] L. C. Evans, Some new PDE methods for weak KAM theory, Calculus of Variations and Partial Differential Equations, 17 (2003), 159–177.

[E4] L. C. Evans, Towards a quantum analogue of weak KAM theory, Communications in Mathematical Physics, 244 (2004), 311–334.

[E5] L. C. Evans, Three singular variational problems, in *Viscosity Solutions of Differential Equations and Related Topics*, Research Institute for the Matematical Sciences, RIMS Kokyuroku 1323, 2003.

[E6] L. C. Evans, A survey of partial differential equations methods in weak KAM theory, Communications in Pure and Applied Mathematics 57 (2004), 445–480.

[E-G1] L. C. Evans and D. Gomes, Effective Hamiltonians and averaging for Hamiltonian dynamics I, Archive Rational Mech and Analysis 157 (2001), 1–33.

[E-G2] L. C. Evans and D. Gomes, Effective Hamiltonians and averaging for Hamiltonian dynamics II, Archive Rational Mech and Analysis 161 (2002), 271–305.

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

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

[F3] A. Fathi, Orbites heteroclines 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

[F5] A. Fathi, *The Weak KAM Theorem in Lagrangian Dynamics* (Cambridge Studies in Advanced Mathematics), to appear.

[F-S1] A. Fathi and A. Siconolfi, Existence of C^1 critical subsolutions of the Hamilton-Jacobi equation, Invent. Math. 155 (2004), 363–388.

[F-S2] A. Fathi and A. Siconolfi, PDE aspects of Aubry-Mather theory for quasiconvex Hamiltonians, Calculus of Variations and PDE 22 (2005), 185–228.

[F-S3] A. Fathi and A. Siconolfi, Existence of solutions for the Aronsson–Euler equation, to appear.

[Fo-M] G. Forni and J. Mather, Action minimizing orbits in Hamiltonian systems, in *Transition to Chaos in Classical and Quantum Mechanics*, Lecture Notes in Math 1589, edited by S. Graffi, Springer, 1994.

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

[G1] D. Gomes, Viscosity solutions of Hamilton-Jacobi equations and asymptotics for Hamiltonian systems, Calculus of Variations and PDE 14 (2002), 345–357.

[G2] D. Gomes, Perturbation theory for Hamilton-Jacobi equations and stability of Aubry-Mather sets, SIAM J. Math. Analysis 35 (2003), 135–147.

[G3] D. Gomes, A stochastic analog of Aubry-Mather theory, Nonlinearity 10 (2002), 271-305.

[G-O] D. Gomes and A. Oberman, Computing the effective Hamiltonian using a variational approach, SIAM J. Control Optim. 43 (2004), 792–812.

[I-SM] R. Iturriaga and H. Sanchez-Morgado, On the stochastic Aubry-Mather theory, to appear.

[K] V. Kaloshin, Mather theory, weak KAM and viscosity solutions of Hamilton-Jacobi PDE, preprint.

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

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

[M1] P. Marcellini, Regularity for some scalar variational problems under general growth conditions, J Optimization Theory and Applications 90 (1996), 161–181.

[M2] P. Marcellini, Everywhere regularity for a class of elliptic systems without growth conditions, Annali Scuola Normale di Pisa 23 (1996), 1–25.

[M-M] E. Mascolo and A. P. Migliorini, Everywhere regularity for vectorial functionals with general growth, ESAIM: Control, Optimization and Calculus of Variations 9 (2003), 399-418.

[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.

[P-R] I. Percival and D. Richards, *Introduction to Dynamics*, Cambridge University Press, 1982.

[V] C. Villani, *Topics in Optimal Transportation*, American Math Society, 2003.

[W] C. E. Wayne, An introduction to KAM theory, in *Dynamical Systems and Probabilistic Methods in Partial Differential Equations*, 3–29, Lectures in Applied Math 31, American Math Society, 1996.

[Y] Y. Yu, L^∞ *Variational Problems, Aronsson Equations and Weak KAM Theory*, thesis, University of California, Berkeley, 2005.