Hamiltonian ODE, Homogenization, and Symplectic Topology

Fraydoun Rezakhanlou
Department of Mathematics, UC Berkeley

January 4, 2021

Abstract

This article is based on a course the author gave in fall of 2018 at UC Berkeley, in connection with the MSRI program *Hamiltonian systems, from topology to applications through analysis*. In this article we explore the connection between the Hamiltonian ODEs and Hamilton-Jacobi PDEs, and give an overview of some of the existing techniques for the question of homogenization. We also discuss stochastic formulations of several classical problems in symplectic geometry.

1 Introduction

Hamiltonian systems of ordinary differential equations appear in celestial mechanics to describe the motion of planets. They are also used in statistical mechanics to model the dynamics of particles in a fluid, gas or many other microscopic models. It was known to Liouville that the flow of a Hamiltonian system preserves the volume. Poincaré observed that the Hamiltonian flows are *symplectic*; they preserve certain *symplectic area* of two dimensional surfaces. Various *Symplectic Rigidity Phenomena* offer ways to take advantage of the symplecticity of Hamiltonian flows.

Writing q and p for the position and momentum coordinates respectively, a Hamiltonian function $H(q,p)$ represents the total energy associated with the pair (q,p). We regard a Hamiltonian system associated with H *completely integrable* if there exists a symplectic change of coordinates $(q,p) \mapsto (Q,P)$, such that our Hamiltonian system in new coordinates is still Hamiltonian system that is now associated with a Hamiltonian function $\bar{H}(P)$. For completely integrable systems the coordinates of $P = P(q,p)$ are conserved and the set of (q,p) at which $P(q,p)$ takes a fixed vector is an invariant set for the flow of our system. These invariant sets are homeomorphic to tori in many classical examples of completely integrable systems. According to *Kolmogorov-Arnold-Moser (KAM) Theory*, many of the *invariant tori* survive when a completely integrable system is slightly perturbed. *Aubry-Mather Theory* constructs a family of invariant sets provided that the Hamiltonian function is convex in the momentum variable. These invariant sets lie on the graph of the gradient of certain scalar-valued functions. A. Fathi [F] uses *viscosity solutions* of the Hamilton-Jacobi PDE associated with the Hamiltonian function H to construct Aubry-Mather invariant measures (see also [B1]). Recently there have been several interesting works to understand the connection between Aubry-Mather Theory and *Symplectic Topology*. The hope is to use tools from Symplectic Topology to construct interesting invariant sets/measures for Hamiltonian systems associated with non-convex Hamiltonian functions.

Most of the aforementioned works on Hamiltonian systems are done when the Hamiltonian function is defined on the cotangent bundle of a compact manifold. A prime example is when $p, q \in \mathbb{R}^d$, with H periodic in q-variable, so that we may regard H as a function that is defined on $T^*\mathbb{T}^d = \mathbb{T}^d \times \mathbb{R}^d$. To go beyond the periodic case, we may take a Hamiltonian function that is *quasi-periodic* with respect to q. In fact there is a probabilistic generalization of quasi-periodic condition by selecting H randomly according to a probability measure $\mathbb{P}$ that is invariant with respect to spatial shifts: $\tau_a H(q,p) = H(q+a,p)$. As it turns out the Hamiltonian $\bar{H}$ can be obtained from H by a scaling limit that is called *Homogenization.*

In these notes we will explore the connection between Hamilton-Jacobi PDE, Homogenization, Hamiltonian ODE and Symplectic Topology.

1.1 Hamiltonian ODE

In Euclidean setting a Hamiltonian system associated with a C^2 Hamiltonian function $H : \mathbb{R}^{2d} \to \mathbb{R}$ is the ODE

$$\dot{x} = X_H(x) := J\nabla H(x), \tag{1.1}$$

where

$$J := \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix},$$

with I denoting the $d \times d$ identity matrix. Writing $x = (q,p)$ with $q,p \in \mathbb{R}^d$, the system (1.1) means

$$\dot{q} = H_p(q,p), \qquad \dot{p} = -H_q(q,p).$$

We write $\phi_t^H(x)$ for the flow of the vector field X_H. Poincaré discovered that the form

$$\left(\phi_t^H\right)^*(\bar{\lambda}) - \bar{\lambda},$$

is exact for $\bar{\lambda} = p \cdot dq$. As a consequence $\left(\phi_t^H\right)^*(\bar{\omega}) = \bar{\omega}$, where

$$\bar{\omega} := d\bar{\lambda} = \sum_{i=1}^d dp_i \wedge dq_i.$$

This means that if $A(x,t) = \left(d\phi_t^H\right)_x$, then

$$\bar{\omega}\big(A(x,t)v, A(x,t)w\big) = \bar{\omega}(v,w), \quad or \quad A(x,t)^* J A(x,t) = J.$$

More generally, we can define *Hamiltonian vector fields* on any *symplectic manifold*. By a symplectic manifold we mean a pair (M,ω) with M a smooth manifold, and ω a non-degenerate closed 2-form on M. Given a smooth function $H : M \to \mathbb{R}$, we define the vector field $X_H = X_H^\omega$ as the unique vector field such that

$$i_{X_H}\omega = -dH.$$

In particular, $\mathcal{L}_{X_H}\omega = 0$, which implies the following identity for its flow:

$$\left(\phi_t^H\right)^*\omega = \omega.$$

When $\omega = \bar{\omega}$, and $M = \mathbb{R}^{2d}$, we have $X_H^{\bar{\omega}} = J\nabla H$.

Given a vector field X on a manifold M, we write ψ_t^X for its flow. Given C^1 scalar-valued function $f : M \to \mathbb{R}$, we define its *Lie derivative* with respect to X by

$$\mathcal{L}_X f(x) = \frac{d}{dt} f\big(\psi_t(x)\big)\Big|_{t=0} = (df)_x\big(X(x)\big). \tag{1.2}$$

More generally, if $u(x,t) = f\big(\psi_t(x)\big)$, then

$$u_t = \mathcal{L}_X u.$$

where u_t denotes the partial derivative of u with respect to t. From this, we learn that a function $f \in C^1(M;\mathbb{R})$ is *conserved* along the flow of X iff $\mathcal{L}_X f = 0$. In the case of a Hamiltonian vector field $X = X_H$, the Lie derivative $\mathcal{L}_X f$ is the *Poisson bracket* of H and f:

$$\{H, f\} := \mathcal{L}_{X_H} f = (df)(X_H) = -\omega(X_f, X_H) = \omega(X_H, X_f).$$

1.2 Completely integrable systems

We may call a Hamiltonian ODE completely integrable if we have a sufficiently explicit formula for its solutions. One strategy to achieve this is by finding enough conservation laws. As it turns out, a Hamiltonian system on a manifold M is completely integrable if it has d many independent conservation laws that do not *interact* with each other. Note that if $f_1, \dots, f_k : M \to \mathbb{R}$ are C^2 functions such that $\{H, f_i\} = 0, i = 1, \dots, k$, then the set

$$M_P = \big\{x \in M : \ (f_1(x), \dots, f_k(x)) = P\big\},$$

is invariant for the flow:

$$x \in M_P \quad \Longrightarrow \quad \phi_t(x) \in M_P.$$

We recall a classical result of Liouville and Arnold (see for example [Ar]).

Theorem 1.1 *Assume that there are C^2 functions $f_1, \dots, f_d : M \to \mathbb{R}$ such that the following conditions hold:*

- $\{H, f_i\} = \{f_i, f_j\} = 0$ *for all i and j.*
- *For $P \in \mathbb{R}^d$, the corresponding set M_P is compact.*
- *For each $x \in M_P$, the vectors $X_{f_1}(x), \dots, X_{f_d}(x)$ are linearly independent.*

Then each such M_P is homeomorphic to a d-dimensional torus. Moreover, the motion of X_H on M_P is conjugate to a linear motion. In other words, there exists a symplectic diffeomorphism $\Psi : \mathbb{T}^d \times \mathbb{R}^d \to M$ such that $\Psi^{-1} \circ \phi_t^H \circ \Psi$ is the flow of a Hamiltonian ODE for which the Hamiltonian function is independent of position.

Remark 1.1(i) For an example, assume that $M = T^*\mathbb{T}^d = \mathbb{T}^d \times \mathbb{R}^d$, and consider a Hamiltonian function H that is independent of q. If we think of a torus as $[0,1]^d$ with $0 = 1$, then the motion is given by $x(t) = x + tv (\mathrm{mod}\, 1)$, for some vector $v = \nabla H(p) \in \mathbb{R}^d$. Depending on the vector v, we may have a periodic or *quasi-periodic* orbit. (The latter means that the closure of the orbit is a k-dimensional linear subtorus for some $k > 1$.)

(ii) The set M_P is an example of a *Lagrangian submanifold.* This means that $\dim M_P = d$ and $\omega \restriction_{M_P} = 0$. The latter follows from

$$\omega\big(X_{f_i}, X_{f_j}\big) = \{f_i, f_j\} = 0,$$

and the independence of $\big\{X_{f_i}(x)\big\}_{i=1}^{d}$, for every $x \in M_P$.

(iii) When $f_1 = H$, let us present a sketch of the proof of Arnold-Liouville's theorem. If we define $\phi_t : M \to M, t = (t_1, \dots, t_d) \in \mathbb{R}^d$ by

$$\phi_t(x) = \phi_{t_1}^{f_1} \circ \dots \circ \phi_{t_d}^{f_d},$$

then $\phi_t(M_P) \subseteq M_P$. On the other hand, if we pick some point $a \in M_P$ and set $\varphi(t) = \phi_t(a)$, then $\varphi : \mathbb{R}^d \to M_P$, and the set

$$\Gamma = \big\{t \in \mathbb{R}^d : \ \varphi(t) = \varphi(0) = a\big\},$$

is a subgroup of $(\mathbb{R}^d, +)$. Indeed the compactness of M_P and the linear independence of $\big\{X_{f_i}(x)\big\}_{i=1}^{d}$ imply that the subgroup Γ is discrete. That is, there are vectors $v_1, \dots, v_d$, such that

$$\Gamma = \big\{n_1 v_1 + \dots + n_d v_d : \ n_1, \dots, n_d \in \mathbb{Z}\big\}.$$

Hence the quotient $\mathbb{R}^d/\Gamma$ is a torus and the map φ yields a homeomorphism $\hat{\varphi} : \mathbb{R}^d/\Gamma \to M_P$. Moreover, assuming that $f_1 = H$, then ϕ_s^H is conjugate to the map $(t_1, \dots, t_d) \mapsto (t_1 + s, \dots, t_d)$. If we use the basis $(v_1, \dots, v_d)$ for $\mathbb{R}^d$, we can then show that ϕ_s^H is conjugate to a linear motion. Writing Q for the coordinates of $\mathbb{R}^d/\Gamma \equiv \mathbb{T}^d$, we have a homeomorphism $\Psi^P = \hat{\varphi} : \mathbb{T}^d \to M_P$. As we vary P, we obtain a map

$$\Psi : T^*\mathbb{T}^d = \mathbb{T}^d \times \mathbb{R}^d \to M.$$

We think of $\Psi(Q, P) = x$ as a parametrization of M. Setting $\bar{H}(P) = H(x) = H\big(\Psi(Q, P)\big)$, for $x \in M_P$, we obtain a new Hamiltonian function $\bar{H} : T^*\mathbb{T}^d \to \mathbb{R}$ that is independent of Q. The motion of $\hat{\phi}_t\big(Q(0), P(0)\big) := (Q(t), P(t))$ may be defined by

$$\hat{\phi}_t := \Psi^{-1} \circ \phi_t^H \circ \Psi.$$

We already know that $Q(t)$ is a linear motion and that $P(t) = P(0)$. We may regard this motion as a solution to the Hamiltonian ODE

$$\dot{Q} = \nabla \bar{H}(P), \qquad \dot{P} = 0.$$

$\square$

In summary, we have seen that for a completely integrable Hamiltonian ODE, we can find a change of coordinates that turns our system to a linear motion. That is, there exists a diffeomorphism Ψ such that

$$\phi_t^{\bar{H}} = \Psi^{-1} \circ \phi_t^H \circ \Psi, \qquad \bar{H} = H \circ \Psi, \tag{1.3}$$

for a Hamiltonian function H that is independent of position. Recall that both ϕ_t^H and $\phi_t^{\bar{H}}$ are symplectic. It is no surprise that the change of coordinates map Ψ is also symplectic. As the following Proposition indicates, a symplectic change of coordinates always transforms a Hamiltonian system to another Hamiltonian system.

Proposition 1.1 *Let (M, ω) and (M', ω') be two symplectic manifolds and assume that $\Psi : M' \to M$ is a diffeomorphism such that $\Psi^* \omega = \omega'$. Let $H : M \to \mathbb{R}$ be a Hamiltonian function on M, and let ϕ_t^H be the flow of X_H^ω. Then*

$$\hat{\phi}_t := \Psi^{-1} \circ \phi_t^H \circ \Psi,$$

is the flow of the vector field $X_{\bar{H}}^{\omega'}$ for $\bar{H} = H \circ \Psi$.

We refer to [MS], [HZ] or [R1] for an introduction to Symplectic Geometry.

1.3 Kolmogorov-Arnold-Moser (KAM) theory

We may take a small perturbation of a completely integrable system and wonder whether or not some of the invariant tori persist. It turns out that for a small perturbation, an invariant torus persists if the *action variable* $\nabla H(P)$ is sufficiently irrational (see for example [W]).

Theorem 1.2 *Assume that $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is of the form*

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

with $\det D^2 H_0 \neq 0$ *and K real analytic. Then for every $\tau, \gamma > 0$, there exists $\varepsilon_0 = \varepsilon_0(\tau, \gamma) > 0$ such that if $\nabla H^0(p)$ satisfies a* Diophantine condition *of the form*

$$n \in \mathbb{Z}^d \setminus \{0\} \qquad \Longrightarrow \left| n \cdot \nabla H^0(p) \right| \geq \gamma |n|^{-\tau},$$

then the vector field X_{H^ε} has a quasi-periodic orbit of velocity $\nabla H^0(p)$, whenever $|\varepsilon| \leq \varepsilon_0$.

Remark 1.2 It is worth mentioning that if we set

$$D(\gamma, \tau) = \left\{ v \in \mathbb{R}^d : \ |v \cdot n| \geq \gamma |n|^{-\tau} \ \text{ for all } \ n \in \mathbb{Z}^d \setminus \{0\} \right\},$$

then the set $D(\tau) = \cup_{\gamma>0} D(\gamma,\tau)$ is of full measure whenever $\tau > d-1$. This is because, the complement of $D(\gamma,\tau)$, restricted to a bounded set, has a volume of order $O(\gamma|n|^{-\tau-1})$, and

$$\sum_{n\neq 0} |k|^{-\tau-1} < \infty,$$

iff $\tau + 1 > d$. □

1.4 Generating function

Note that a Hamiltonian vector field is very special as it is fully determined by a scalar-valued function, namely its Hamiltonian function. As it turns out, the symplectic maps are also locally determined by scalar-valued functions known as *generating functions.* To explain this, take an $\bar{\omega}$-symplectic map $\psi(q,p) = (Q,P)$, and observe that since $\psi^*\bar{\omega} = \bar{\omega}$, we can find a scalar-valued function S such that

$$p \cdot dq - P \cdot dQ = dS. \tag{1.4}$$

Normally we think of S as a function of (q,p) or (Q,P). However, it is more convenient to think of S as a function of other pairs. For example under some non-degeneracy assumption (for example if $Q_p(q,p)$ is invertible so that we can locally solve $Q(q,p) = Q$ implicitly for $p = p(q,Q)$), we may regard $S = S(q,Q)$ as a function of the pair (q,Q). Under such circumstances, (1.4) implies

$$S_q(q,Q) = p, \quad -S_Q(q,Q) = P, \qquad \psi\big(q, S_q(q,Q)\big) = \big(Q, -S_Q(q,Q)\big). \tag{1.5}$$

The scalar-valued functions S is an example of a generating function for the symplectic map ψ. Since there are other type of generating functions that we may consider for a symplectic map, let us refer to S as a *generating function of type I* (in short GFI).

Alternatively, we may set $W = p \cdot q - S$, and regard W as a function of (Q,p) so that (1.4) means

$$W_p(Q,p) = q, \quad W_Q(Q,p) = P, \qquad \psi\big(W_p(Q,p), p\big) = \big(Q,\ W_Q(Q,p)\big).$$

The function W is another example of a generating function for the symplectic map ψ and we will refer to it as a *generating function of type II* (in short GFII). Another popular choice for a generating function is $W' = W'(q,P)$ that will be referred to as a *generating function of type III* (in short GFIII).

If ψ is the change of coordinates transformation of a completely integrable system, we have

$$\bar{H}(P) = H(q,p) = H\big(q, W'_q(q,P)\big).$$

This means that for each fixed P, the function $q \mapsto W'(q,P)$ is a solution to a *Hamilton-Jacobi Equation (HJE)* associated with H. Thinking of $\mathbb{T}^d \times \mathbb{R}^d$, as $T^*\mathbb{T}^d$, we interpret $W'_q(q,P)$ as a 1-form on the torus for each P. If we write $W'(q,P) = q \cdot P + w^P(q)$ and assume that $w^P : \mathbb{T}^d \to \mathbb{R}$, is periodic, then our HJE reads as

$$H\big(q, P + (dw^P)_q\big) = \bar{H}(P). \tag{1.6}$$

We think of $\alpha^P = P + dw^P$ as a closed 1-form that belongs to the cohomology class of the constant (closed) form P.

1.5 Weak KAM theory

In the classical KAM Theory, we consider a small perturbation of a non-degenerate Hamiltonian function $H_0(p)$ that depends on p only. We have learned that the majority of the invariant tori of unperturbed systems persist for a sufficiently small perturbation. However some invariant tori could be destroyed after a small perturbation. In fact Arnold constructed an example of a perturbed integrable system, in which chaotic orbits - resulting from the breaking of unperturbed KAM tori - coexist with the invariant tori of KAM theorem. This phenomenon is known as *Arnold diffusion.* A natural question is whether or not we can construct a family of invariant sets $\big(M_P : P \in \mathbb{R}^d\big)$ for perturbed systems that come from the invariant tori of the unperturbed system and still carry some of their features. Aubry and Mather constructed such family for the so-called *twist maps* (these maps are the analog of Hamiltonian systems when $d = 1$ and time is discrete). The generalization of Aubry-Mather invariant sets to higher dimensions was achieved by Mather, Mañé and Fathi. They prove the existence of interesting invariant (action-minimizing) sets, which generalize KAM tori, and which continue to exist even after KAM tori disappearance.

Aubry-Mather Theory replaces the condition of being close to an integrable Hamiltonian with the *Tonelli Condition.* We say that a Hamiltonian function $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ is Tonelli, if the following conditions are true:

- $H(q,p)$ is C^2, and the matrix $H_{pp}(q,p)$ is positive definite for every (q,p).
- $|p|^{-1}H(q,p) \to \infty$ as $|p| \to \infty$, uniformly in q.

According to Aubry-Mather and Mather-Mane-Fathi Theory, for each P, there exists a constant $\bar{H}(P)$, a Lipschitz function $w^P : \mathbb{T}^d \to \mathbb{R}$, and an invariant measure μ^P for ϕ^H such that

- The function w^P solves the HJE (1.6) in a suitable weak sense.
- The support of the measure μ^P is a subset of

$$M_P = \left\{ \big(q, P + (dw^P)_q\big) : \ q \in \mathbb{T}^d \right\}.$$

Note that we only require the function w^P to be Lipschitz and not everywhere differentiable. This is because the HJE (1.6) does no possess classical solutions in general. One remedy for this is to consider certain generalized solutions. In fact if we consider the so called *viscosity solutions*, then (1.6) always has at least one Lipschitz solution for each P. This was established by Lions, Papanicolaou and Varadhan [LPV] in 1987. We then modify the definition of M_P with

$$M_P = \left\{ \big(q, P + (dw^P)_q\big) : \ q \in \mathbb{T}^d, \ w^P \ \text{ differentiable at } \ q \right\}. \tag{1.7}$$

1.6 From torus to general closed manifolds

We may replace the torus with any sufficiently smooth manifold M in weak KAM theory. Now our Hamiltonian function H is a C^2 function on the cotangent bundle T^*M. The manifold T^*M carries a standard symplectic form $\omega = d\lambda$ with λ defined as

$$\lambda_{(q,p)}\big(a\big) = p_q\big((d\pi)_{(q,p)}a\big),$$

where $\pi : T^*M \to M$ is the projection $\pi(q,p) = q$ to the base point, and its derivative $(d\pi)_{(q,p)} : T_{(q,p)}T^*M \to T_qM$ projects onto tangent vectors. Recall that in the case of a torus, we know that the (1.6) has at least one solution by [LPV]. This existence result has been extended to arbitrary closed manifold and convex Hamiltonian by Albert Fathi [F].

Theorem 1.3 *Let M be a smooth closed manifold and assume that $H : T^*M \to \mathbb{R}$ is a Tonelli Hamiltonian. Then for every closed form α, there exists a unique constant $\bar{H}(\alpha)$, and a Lipschitz function $w : M \to \mathbb{R}$ such that w satisfies*

$$H\big(q, \alpha_q + (dw)_q\big) = \bar{H}(\alpha), \tag{1.8}$$

in viscosity sense.

Because of the uniqueness of $\bar{H}$, it is clear that if we add an exact form to α, the value of $\bar{H}$ does not change. Abusing the nota tion slightly, we may define $\bar{H}$ on the space $H^1(M)$ of the cohomology classes of 1-forms and write $\bar{H}([\alpha])$ in place of $\bar{H}(\alpha)$. Alternatively, for each $P \in H^1(M)$, we may fix a representative $\bar{\alpha}^P$ in class P and search for a Lipschitz $w^P : M \to \mathbb{R}$ such that $\alpha^P = \bar{\alpha}^P + dw^P$. Even when we fix the representative, the function w^P may not be unique. Given a choice of w^P, we define an invariant set M' by

$$M' = \left\{ \big(q, \bar{\alpha}^P_q + (dw^P)_q\big) : \ q \in M, \ w^P \ \text{ differentiable at } \ q \right\}. \tag{1.9}$$

1.7 From torus to stochastic Hamiltonian and homogenization

Weak KAM Theory à la Fathi employs the HJE (1.6) in order to construct interesting invariant measures for the corresponding Hamiltonian ODE. It turns out that HJE can be used to model certain *deterministic and stochastic growths.* More precisely, imagine that we have an interface that separates different phases and this interface is represented by a graph of function $u(\cdot,t):\mathbb{R}^d\to\mathbb{R}$ at time t. Suppose that the growth rate of this interface depends on the position q, and the inclination of the interface u_q. Mathematically speaking, u satisfies a HJE of the form

$$u_t + H\big(q, u_q(q,t)\big) = 0, \tag{1.10}$$

for a Hamiltonian function $H:\mathbb{R}^{2d}\to\mathbb{R}$. We think of (1.10) as the microscopic equation describing the evolution of the interface. If a large parameter n represents the ratio between the macro and micro scales, then

$$u^n(q,t) = n^{-1}u(nq,nt),$$

is the corresponding macroscopic height above that macro position q at the macro time t. We observe that u^n now solves

$$u^n_t + H^n\big(q, u^n_q(q,t)\big) = 0, \tag{1.11}$$

where

$$H^n(q,p) = (\gamma_n H)(q,p) := H(nq,p).$$

A *homogenization* occurs if the limit

$$\bar u(q,t) = \lim_{n\to\infty} u^n(q,t),$$

exists whenever the limit

$$g(q) := \lim_{n\to\infty} u^n(q,0),$$

exists. As it turns out, in many examples of interest, the limit $\bar u$ satisfies a simpler HJE of the form

$$\begin{cases} \bar u_t + \bar H(\bar u_q) = 0 \\ \bar u(q,0) = g(q). \end{cases} \tag{1.12}$$

In fact we may use (1.6) to guess that when H is periodic in q, then $\bar H$ that appears in (1.12) coincides with $\bar H$ that appears in (1.6). This is because if w^P is a periodic function that satisfies (1.6), and we choose $u(q,0) = P\cdot q + w^P(q)$ as the initial condition for (1.10), then $u(q,t) = P\cdot q - t\bar H(P) + w^P(q)$, and

$$\bar u(q,t) = \lim_{n\to\infty} u^n(q,t) = P\cdot q - t\bar H(P),$$

which solves (1.12).

We may wonder whether a weak KAM Theory can be achieved for $H : \mathbb{R}^{2d} \to \mathbb{R}$ that is not necessarily periodic. Let us denote by $\mathcal{H}$ the set of all C^1 Hamiltonian functions $H : \mathbb{R}^{2d} \to \mathbb{R}$. For homogenization question, there are two relevant group actions on $\mathcal{H}$, namely the spacial translation and scaling. More precisely we set

$$\tau_a H(q,p) = H(q+a,p), \qquad \gamma_n H(q,p) = H(nq,p),$$

for $a \in \mathbb{R}^d$ and $n \in \mathbb{R}^+$. We certainly have

$$\tau_a \circ \tau_b = \tau_{a+b}, \qquad \gamma_m \circ \gamma_n = \gamma_{mn}.$$

We are interested to know for what Hamiltonian $H \in \mathcal{H}$ we have weak KAM theory and homogenization. Let us make a comment on bounded continuous functions K of the position variable. For $K : \mathbb{R}^d \to \mathbb{R}$, we define the translation operator $\tau_a K(q) = K(q+a)$ as before. We note that if a function K is periodic in q, then the set

$$\big\{\tau_a K : \ a \in \mathbb{R}^d\big\},$$

is homeomorphic to a d-dimensional torus. More generally, let us take a function $\hat{K} : \mathbb{T}^N \to \mathbb{R}$, and a $N \times d$ matrix A. We then set $K(q) = \hat{K}(Aq)$, which is an example of a *quasi-periodic* function. In fact the closure of the set

$$\Gamma(K) := \{\tau_a K : \ a \in \mathbb{R}^d\},$$

with respect to the uniform topology is

$$\Gamma(\hat{K}) := \big\{\hat{K}(\cdot + b) : \ b \in \mathbb{R}^N\big\},$$

if the following condition holds:

$$n \in \mathbb{Z}^N \setminus \{0\} \quad \Longrightarrow \quad nA \neq 0.$$

In general a bounded continuous function $K : \mathbb{R}^d \to \mathbb{R}$ is called *almost periodic* if the set $\Gamma(K)$ is precompact in $C_b(\mathbb{R}^d)$ with respect to the uniform topology.

We regard the group $\big\{\tau_a : \ a \in \mathbb{R}^d\big\}$ as a d-dimensional dynamical system on $\mathcal{H}$. A probability measure $\mathbb{P}$ on $\mathcal{H}$ is *translation invariant and ergodic* if the following conditions are met:

- For every Borel set $\mathcal{A} \subset \mathcal{H}$, and $a \in \mathbb{R}^d$, we have $\mathbb{P}\big(\tau_a \mathcal{A}\big) = \mathbb{P}(\mathcal{A})$.
- If a Borel set $\mathcal{A}$ is invariant i.e., $\tau_a \mathcal{A} = \mathcal{A}$ for all $a \in \mathbb{R}^d$, then $\mathbb{P}(\mathcal{A}) \in \{0,1\}$.

We may wonder whether or not the weak KAM theory or homogenization are applicable to generic Hamiltonian functions in the support of an invariant ergodic measure. The hope is that Birkhoff Ergodic Theorem would make up for the lack of compactness that has played an essential role when we considered a cotangent bundle of a compact manifold in **1.6**.

1.8 Variational techniques

Homogenization questions and the existence of interesting invariant measures are closely related to the existence of special orbits of the Hamiltonian ODEs. Such existence questions also play central role in several recent developments in symplectic topology. (A prime example is *Floer Homology* that was formulated by Floer in order to prove *Arnold's conjecture.*) Hamilton discovered a variational description for the solutions of Hamiltonian systems. More specifically, we may reduce the existence of special orbits of (1.1) to the existence of a critical point of a suitable *action functional*. To explain this, let us assume that (M, ω) is a symplectic manifold with $\omega = d\lambda$. We also write Γ_T for the space of C^1 functions $x : [0, T] \to T^*M$. Given a Hamiltonian function $H : T^*M \times [0, T] \to \mathbb{R}$, we define $\mathcal{A} = \mathcal{A}_H : \Gamma_T \to \mathbb{R}$ by

$$\mathcal{A}(\gamma) = \mathcal{A}_H^T(\gamma) := \int_0^T \left[\lambda_{\gamma(t)}\big(\dot\gamma(t)\big) - H(\gamma(t), t)\right]\, dt. \tag{1.13}$$

The form $\lambda^H = \lambda - H\ dt$ is known as the *Poincaré-Cartan* form. We note that if we regard $d\lambda^H = \omega + dt \wedge dH$ as a form on $T^*M \times \mathbb{R}$, and $\hat X_H = \big(X_H, 1\big)$, then

$$i_{\hat X_H} d\lambda^H = i_{X_H}\omega + dH = 0.$$

Moreover, if we take a variation of a path with fixed end points, for example

$$w : [0, T] \times [0, \delta] \to T^*M, \qquad (t, \theta) \mapsto w(t, \theta),$$

with

$$w(t, 0) = \gamma(t), \qquad w(0, \theta) = w(0, 0), \qquad w(T, \theta) = w(T, 0), \qquad w_\theta(t, 0) = v(t),$$

then

$$\begin{aligned} -\frac{d}{d\theta}\int_{w(\cdot,\theta)} \lambda\big|_{\theta=0} &= \lim_{h\to 0} h^{-1}\left[\int_{w(\cdot,0)}\lambda - \int_{w(\cdot,h)}\lambda\right] = \lim_{h\to 0} h^{-1}\int_{w([0,T]\times[0,h])}\omega \\ &= \lim_{h\to 0} h^{-1}\int_0^h\int_0^T \omega_w\big(w_t, w_\theta\big)\ dt d\theta = \int_0^T \omega_\gamma\big(\dot\gamma, v\big)\ dt. \end{aligned}$$

(Note that the orientation of w must be compatible with $\gamma = w(\cdot, 0)$ for Stokes Theorem to apply.) This in turn implies

$$\frac{d}{d\theta}\ \mathcal{A}_H^T\big(w(\cdot,\theta)\big)\big|_{\theta=0} = -\int_0^T \big(i_{\dot\gamma}\omega + dH\big)_\gamma(v)\ dt = -\int_0^T \big(i_{\dot\gamma - X_H(\gamma)}\ \omega\big)_\gamma(v)\ dt. \tag{1.14}$$

Hence, if we restrict $\mathcal{A}$ to the set of curves with the same end points, then its *critical points* are the orbits of X_H. In fact the *critical values* of $\mathcal{A}$ solve the corresponding Hamilton-Jacobi

PDE. To explain this, first we argue that the action functional can be used to produce generating functions for ϕ_T^H. Indeed if we define $\lambda_H^T : T^*M \to \mathbb{R}$, by

$$\lambda_H^T(x) = \mathcal{A}_H(\eta_T^x), \qquad \text{where} \quad \eta_T^x(t) = \phi_t^H(x) \ \text{ for } \ t \in [0, T], \tag{1.15}$$

then λ_H^T is a generating function for ϕ_T^H.

Proposition 1.2 *For every $T \geq 0$ and any Hamiltonian H, we have*

$$d\lambda_H^T = \left(\phi_T^H\right)^* \lambda - \lambda. \tag{1.16}$$

Proof Set

$$A(x) = \int_{\eta_T^x} \lambda, \qquad B(x) = \int_0^T H\big(\eta_T^x(t), t\big)\, dt.$$

Take any $\big(\tau(\theta) : \ 0 \leq \theta \leq \delta\big)$ with $\tau(0) = x$ and $\dot{\tau}(0) = v \in T_xM$. Set $y(t,\theta) = \phi_{-t}^H\big(\tau(\theta)\big)$,

$$\Theta_h = \{y(t,\theta) : \ 0 \leq t \leq T, \ 0 \leq \theta \leq h\},$$

and use Stokes' theorem to assert that for $h \in (0, \delta)$,

$$h^{-1} \int_0^h \int_0^T \omega_y(y_t, y_\theta)\, dt d\theta = h^{-1} \int_{\Theta_h} d\lambda = h^{-1} \left[\int_{\eta^{\tau(0)}} \lambda - \int_{\eta^{\tau(h)}} \lambda + \int_{\varphi\circ\tau(\cdot)} \lambda - \int_{\tau(\cdot)} \lambda\right],$$
$$h^{-1} \int_0^h \int_0^T \left(i_{X_H}\omega\right)_y (y_\theta)\, dt d\theta = h^{-1} \left[\int_{\eta^{\tau(0)}} \lambda - \int_{\eta^{\tau(h)}} \lambda + \int_{\tau(\cdot)} (\varphi^*\lambda - \lambda)\right],$$

where $\varphi = \phi_T^H$. Sending $h \to 0$ yields

$$-(dB)_x(v) = -(dA)_x(v) + \big(\varphi^*\lambda - \lambda\big)_x(v).$$

This is exactly (1.16). $\square$

Let us assume that $M = \mathbb{R}^{2d}$, $\omega = \bar{\omega}$, and that H is a C^2 Hamiltonian function with D^2H uniformly bounded. With the aid of Proposition 1.2 we may define a GFI of ϕ_t^H by

$$S(q(t), t; q) := \int_0^t \big[p(s) \cdot \dot{q}(s) - H(q(s), p(s), s)\big]\, ds, \tag{1.17}$$

where $(q(s), p(s)) = \phi_s^H\big(q(0), p(0)\big)$. Hence

$$\phi_t^H\big(q, -S_q(Q,t;q)\big) = \big(Q, S_Q(Q,t;q)\big), \qquad q(0) = q, \quad q(t) = Q, \quad p(t) = S_Q(Q,t;q).$$

Differentiating both sides of (1.17) with respect to t yields

$$S_t(Q,t;q) + S_Q(Q,t;q) \cdot \dot{q} = p(t) \cdot \dot{q}(t) - H(q(t), p(t), t).$$

As a result,

$$S_t(Q,t;q) + H\left(Q, S_Q(Q,t;q), t\right) = 0. \tag{1.18}$$

Similarly if we set $W = S + q \cdot p$, and regard $W(Q,t;p)$ as a function of (Q,p), then

$$W(q(t),t;p(0)) = p(0) \cdot q(0) + \int_0^t \left[p(s) \cdot \dot{q}(s) - H(q(s),p(s),s)\right] ds.$$

Differentiating both sides with respect to t yields

$$W_t(q(t),t;p(0)) + W_Q(q(t),t;p(0)) \cdot \dot{q}(t) = p(t) \cdot \dot{q}(t) - H(q(t),p(t),t).$$

This yields

$$W_t(Q,t;p) + H\left(Q, W_Q(Q,t;p), t\right) = 0, \tag{1.19}$$

because $W_Q(q(t),t;p(0)) = p(t)$.

Remark 1.3(i) In particular, if H is 1-periodic in t, $T = 1$, and we define $\mathcal{A}$ on the space of 1-periodic paths (loops), then the critical points of $\mathcal{A}$ correspond to the periodic orbits of X_H. Floer uses the gradient flow equation

$$w_s = -\partial\mathcal{A}(w), \tag{1.20}$$

to prove the existence of periodic orbits by showing that

$$\lim_{s\to\infty} w(\cdot,s),$$

exists. Here the gradient is defined with respect to the L^2 inner product, which guarantees that the equation (1.20) is an elliptic (in fact Cauchy-Riemann type) PDE. One may use the elliptic regularity of the solutions to obtain the compactness of path w in a suitable Sobolev space.

(ii) When H is a Tonelli Hamiltonian, it is more convenient to work with an action functional that is expressed in terms of the Legendre transform of H. To explain this, let us assume that there exists a C^2 function $L : TM \to \mathbb{R}$, $L = L(q,v)$, that is convex in the velocity v, and that the transformation $\mathbb{L} : TM \to T^*M$,

$$\mathbb{L}(q,v) = \big(q, L_v(q,v)\big), \tag{1.21}$$

is a C^1 diffeomorphism with

$$p = L_v(q,v) \quad \text{iff} \quad v = H_p(q,p).$$

(Here we identify $\big(T_qM\big)^{**}$ with T_qM.) The *Lagrangian* function L and the Hamiltonian function H are related to each other by Legendre transform

$$L(q,v)=\sup_{p\in T_q^*M}\big(p(v)-H(q,p)\big),\qquad H(q,p)=\sup_{v\in T_qM}\big(p(v)-L(q,v)\big).$$

Moreover

$$H\circ\mathbb{L}(q,v)=L_v(q,v)(v)-L(q,v).$$

Given a C^1 path $\alpha:[0,T]\to M$, we may define,

$$\mathcal{L}(\alpha):=\int_0^T L(\alpha,\dot\alpha)\ dt.$$

Note that if $x(t)=\phi_t^H(a)$ is a solution of (1.1), then

$$\lambda_x(\dot x)-H(x)=p_q\big((d\pi)_x(\dot x)\big)-H(q,p)=p_q(\dot q)-H(q,p)=L(q,\dot q).$$

Hence

$$\mathcal{A}\big(x(\cdot)\big)=\int_0^T\big(\lambda_x(\dot x)-H(x)\big)\ dt=\int_0^T L(q,\dot q)\ dt=\mathcal{L}\big(q(\cdot)\big).$$

By a classical work of Tonelli, we have the following results:

- If we regard the action functional $\mathcal{L}$ as a function on paths $\alpha:[0,T]\to M$ with specified endpoints, then $\mathcal{L}$ has a minimizer $q(\cdot)$. As a consequence, this minimizer is a critical point of $\mathcal{L}$, and satisfies the Euler-Lagrange equation

$$\frac{d}{dt}L_v(q,\dot q)=L_q(q,\dot q). \tag{1.22}$$

- The corresponding path $x(t)=\mathbb{L}\big(q(t),\dot q(t)\big)$ satisfies the equation (1.1).

□

1.9 Discrete models

Any symplectic map ψ from a symplectic manifold to itself serves as an example of a discrete analog of a Hamiltonian flow. We will be mainly interested in those symplectic diffeomorphisms for which a global generating function exists. For example, we may assume that a generating function of the first kind exists, i.e., (1.5) holds for some $S(q,Q)$ (with a slight abuse of notation we use the letter S for our generating function as in Subsection 1.9). In the Euclidean setting, we may write $S(q,Q)=:L(q,Q-q)$. If $L(q,v)$ is bounded below and has a superlinear growth at infinity in the *velocity* variable v, we call the corresponding map ψ a twist map and the corresponding dynamical model is a generalization of the *Frenkel-Kontorova Model.* Given a sequence $\mathbf{q} = (q_0, q_1, \dots, q_n)$, we define its *action* by

$$\mathcal{A}(\mathbf{q}) = \sum_{i=1}^{n} S(q_{i-1}, q_i) = \sum_{i=1}^{n} L(q_{i-1}, q_i - q_{i-1}).$$

The critical points of $\mathcal{A}$ correspond to the orbits of ψ. As we will see in Section 2, we may use the minimizers of $\mathcal{A}$ to construct interesting orbits of ψ.

We may also consider a generating function $W(Q,p) = Q \cdot p - w(Q,p)$ of type III so that

$$\psi\big(Q - w_p(Q,p), p\big) = \big(Q, p - w_Q(Q,p)\big).$$

In other words,

$$Q = q + w_p(Q,p), \qquad P = p - w_Q(Q,p),$$

which should be regarded as a discrete analog of a Hamiltonian ODE, with the function w playing the role of the Hamiltonian function.

Example 1.1 *(Standard Map)* Consider the Hamiltonian function $H(q,p) = \frac{1}{2}|p|^2 + V(q)$ for a C^2 potential function $V : \mathbb{T}^d \to \mathbb{R}$. The corresponding Hamiltonian equations are

$$\dot{q} = p, \qquad \dot{p} = -\nabla V(q).$$

For a discrete version of these equations, we consider a map $\psi(q,p) = (Q,P)$ with

$$P = p - \nabla V(q), \qquad Q = q + P.$$

This corresponds to a symplectic map associated with the generating function

$$S(q,Q) = \frac{1}{2}|Q-q|^2 - V(q).$$

□

2 Twist Maps and Their Generalizations

The origin of the twist maps goes back to Poincaré's work on area-preserving maps on annulus that he encountered in his work on 3-body problem of celestial mechanics. Before embarking on studying twist maps, we give an overview of circle diffeomorphisms and their rotation numbers.

Definition 2.1(i) Regarding $\mathbb{S}^1$ as the interval $[0,1]$ with $0 = 1$, let $f : \mathbb{S}^1 \to \mathbb{S}^1$ be an orientation preserving homeomorphism. Its *lift* $F = \ell(f)$ is an increasing map $F : \mathbb{R} \to \mathbb{R}$ such that $f(x) = F(x) \pmod 1$, and F can be written as $F(x) = x + G(x)$, for a 1-periodic function $G : \mathbb{R} \to \mathbb{R}$. We may also regard G as a map on the circle: $g : \mathbb{S}^1 \to \mathbb{R}$, $g(x) = G(x)$ for $x \in [0, 1)$.

(ii) We define $\pi : \mathbb{R} \to \mathbb{S}^1$ by $\pi(x) = e^{2\pi i x}$. For f and F as in **(i)**, we define its rotation number

$$\rho(F) = \lim_{n\to\infty} n^{-1} F^n(x), \qquad \rho(f) = \pi\big(\rho(F)\big). \tag{2.1}$$

(iii) Given $\rho \in [0, 1)$, we write r_ρ for a rotation of the circle through the angle ρ. Its lift R_ρ is given by $R_\rho(x) = x + \rho$.

(iv) We write $D(\tau)$ for the set of numbers that satisfy a *Diophantine condition of type* τ. More precisely, $\rho \in D(\tau)$ if and only if there exists a positive constant c such that for every $r, s \in \mathbb{Z}$,

$$\left|\rho - \frac{r}{s}\right| \geq \frac{c}{|s|^\tau}.$$

□

Theorem 2.1 *(Poincaré) Let $f : \mathbb{S}^1 \to \mathbb{S}^1$ be an orientation preserving homeomorphism and write F for its lift. Then the following statements are true:*

(i) *The rotation number always exists and is independent of x.*

(ii) *f has a fixed point iff $\rho(f) = 0$.*

(iii) *$\pm\rho(F) > 0$ iff $\pm(F(x) - x) > 0$.*

(iv) *Let (r, s) be a pair of coprime positive integers. Then f has a (r, s)-periodic orbit (this means that $F^s(x) = F(x) + r$ for $F = \ell(f)$), iff $\rho(f) = r/s$.*

(v) *If $\rho(f) \notin \mathbb{Q}$, then the set $\Omega_\infty(x)$ of the limit points of the sequence $\big\{f^n(x) : \ n \in \mathbb{N}\big\}$ is independent of x, and is either $\mathbb{S}^1$ or nowhere dense.*

Proof We only prove **(i)** and refer to [KH] for the proof of the other parts.

By induction, we can readily show that if $F(x) = x + g(x)$ for a periodic function g, then $F^n(x) = x + G_n(x)$ for a periodic function G_n that is simply given by

$$G_n(x) = \sum_{i=0}^{n-1} G\big(F^i(x)\big) = \sum_{i=0}^{n-1} g\big(f^i(x)\big). \tag{2.2}$$

Since F^n is increasing, we learn that if $0 \le y \le x < 1$, then

$$x + G_n(x) = F^n(x) \ge F^n(y) = y + G_n(y), \quad \text{or} \quad G^n(y) - G^n(x) \le x - y < 1.$$

From this and 1-periodicity of G_n we deduce that $G^n(y) - G^n(x) < 1$ for all x and y. Hence

$$G_{m+n}(x) = G_m(x) + G_n\big(F^m(x)\big) \le G_m(x) + G_n(x) + 1.$$

This means that the sequence $\{a_n = G_n(x)\}$ is *almost* subadditive (more precisely, the sequence $\{a_n + 1\}$ is subadditive). From this we deduce

$$\rho(x) = \lim_{n\to\infty} n^{-1} G_n(x) = \lim_{n\to\infty} n^{-1}\big(F_n(x) - x\big) = \lim_{n\to\infty} n^{-1} F_n(x),$$

exists. From the last equality we learn that the limit ρ is non-decreasing, whereas the first equality implies that ρ is 1-periodic. This is possible only if $\rho(x)$ is a constant function. □

Theorem 2.2 *Let f and F be as in Theorem 2.1.*

(i) *(Denjoy) If $f \in C^1$ with f' a function of bounded variation, and $\rho = \rho(f) \notin \mathbb{Q}$, then there exists a homeomorphism h such that $f = h^{-1} \circ r_\rho \circ h$.*

(ii) *(Herman [H]) If $f \in C^{2+\alpha}$ with $\alpha \in [0,1)$, and $\rho(F) \in D(\tau)$ for some $\tau > 2$, then h in Part* **(i)** *is in $C^{1+\alpha}$. (See Definition 2.1*(**iv**) *for the definition of $D(\tau)$.)*

Remark 2.1(i) Let us write H, F, and $R_\rho(x) = x + \rho$, for the lifts of the maps h, f and r_ρ, respectively. Since the Lebesgue measure is invariant for R_ρ, and $F \circ H^{-1} = H^{-1} \circ R_\rho$, we learn that for any 1-periodic continuous function $\zeta : \mathbb{R} \to \mathbb{R}$,

$$\int \zeta \circ F \; dH = \int \zeta \; dH.$$

In other words, the measure μ with $\mu[0, x] = H(x)$ is invariant for f. Hence Part **(ii)** is equivalent to the statement that if $f \in C^{2+\alpha}$, then the dynamical system associated with f is (uniquely) ergodic with an invariant measure that has a C^α density with respect to Lebesgue measure.

(ii) In terms of the invariant measure, the rotation number can be express as

$$\rho(f) = \int g \; d\mu,$$

by (2.1), (2.2) and the Ergodic Theorem.

(iii) Define $\mathcal{F}$ to be the set of continuous increasing functions $F : \mathbb{R} \to \mathbb{R}$ such that

$$\sup_x |F(x) - x| < \infty.$$

Write $F(x) = x + G(x)$, and define a translation operator that translates G:

$$(\tau_a F)(x) = F(x+a) - a = x + G(x+a).$$

Let $\mathbb{P}$ be a τ-invariant ergodic probability measure on $\mathcal{F}$. Then one can show that there exists a constant $\rho(\mathbb{P})$ such that

$$\lim_{n\to\infty} n^{-1} F^n(x) = \rho(\mathbb{P}),$$

for $\mathbb{P}$-almost all choices of F. $\square$

We next study *cylinder maps.*

Definition 2.2(i) Let $\varphi : \mathbb{S}^1 \times [-1,1] \to \mathbb{S}^1 \times [-1,1]$, be an orientation preserving homeomorphism. Its *lift* $\ell(\varphi) = \Phi : \mathbb{R} \times [-1,1] \to \mathbb{R} \times [-1,1]$ is a homeomorphism such that

$$\varphi(x) = \Phi(x) \pmod 1,$$

and Φ can be written as $\Phi(q,p) = (q,0) + \Psi(q,p)$, for a continuous $\Psi : \mathbb{R} \times [-1,1] \to \mathbb{R} \times [-1,1]$, that is 1-periodic function in q-variable.

(ii) An orientation-preserving diffeomorphism $\varphi : \mathbb{S}^1 \times [-1,1] \to \mathbb{S}^1 \times [-1,1]$ is called a *twist map* if the following conditions are met:

(i) φ (or equivalently its lift Φ) is area-preserving.

(ii) If we define $\Phi^{\pm}$ by $(\Phi^{\pm}(q), \pm 1) = \Phi(q, \pm 1)$, then $\pm\big(\Phi^{\pm}(x) - x\big) > 0$. Equivalently, $\pm\rho\big(\Phi^{\pm}\big) > 0$.

$\square$

Our main result about twist maps is the following:

Theorem 2.3 *(Poincaré and Birkhoff) Any twist map has at least two fixed points.*

Poincaré established Theorem 2.3 provided that φ has a global generating function. Such a generating function exists if φ is a *monotone twist map.* To explain Poincare's argument, let us formulate a condition on $\Phi = \ell(\varphi)$ that would guarantee the existence of a global generating function $S(q, Q)$ for Φ.

Definition 2.3 A C^1 area-preserving map φ or its lift $\Phi(q, p) = (Q(q, p), P(q, p))$ is called *positive* twist if $Q_p(q, p) > 0$ for all (q, p). We say φ is *negative* twist if φ^{-1} is a positive twist. We say that φ is a *monotone* twist, if φ either positive or negative twist. $\square$

Proposition 2.1 *Let Φ be a monotone twist map. Then there exists a C^2 function $S : U \to \mathbb{R}$ with*

$$U = \{(q, q') : Q(q, -1) \leq q' \leq Q(q, +1)\}$$

such that

$$\Phi(q, -S_q(q, Q)) = (Q, S_Q(q, Q)).$$

Moreover

$$S(q + 1, Q + 1) = S(q, Q), \qquad S_{qQ} < 0. \tag{2.3}$$

Proof The image of the line segment $\{q\} \times [-1, 1]$ under Φ is a curve γ with parametrization $\gamma(p) = (Q(q, p), P(q, p))$. By the monotonicity, the relation $Q(q, p) = Q$ can be inverted to yield $p = p(q, Q)$ which is increasing in Q. The set $\gamma([-1, 1])$ can be viewed as a graph of the function

$$Q \mapsto P(q, p(q, Q))$$

with $Q \in [Q(q, -1), Q(q, +1)]$. The anti-derivative of this function yields $S(q, Q)$. This can be geometrically described as the area of the region Δ between the curve $\gamma([-1, 1])$, the line $P = -1$, and the vertical line $\{q\} \times [-1, 1]$. We now apply Φ^{-1} on this region. The line segment $\{Q\} \times [-1, 1]$ is mapped to a curve $\hat{\gamma}([-1, 1])$ which coincides with a graph of a function $q \mapsto p(q, Q)$. Since Φ is area preserving the area of $\Phi^{-1}(\Delta)$ is $S(q, Q)$. From this we deduce that $S_Q = -p$. Here we have used the fact that Φ^{-1} is a (negative) twist map; indeed if we write $\Phi^{-1}(Q, P) = (\hat{q}(Q, P), \hat{p}(Q, P))$, then

$$(\Phi^{-1})' = \begin{bmatrix} \hat{q}_Q & \hat{q}_P \\ \hat{p}_Q & \hat{p}_P \end{bmatrix} = \begin{bmatrix} Q_q & Q_p \\ P_q & P_p \end{bmatrix}^{-1} = \begin{bmatrix} P_p & -Q_p \\ -P_q & Q_q \end{bmatrix}$$

which implies that $\hat{q}_P = -Q_p < 0$.

The periodicity (2.3) is an immediate consequence of $\Phi(q + 1, p) = \Phi(q, p) + (1, 0)$;

$$\Phi(\{q + 1\} \times [-1, 1]) = \Phi(\{q\} \times [-1, 1]) + (1, 0).$$

As for the second assertion in (2.3), recall that $p(q, Q)$ is increasing in Q. Hence

$$S_{qQ} = -p_Q < 0.$$

□

We now show how the existence of a generating function can be used to prove the existence of fixed points.

Proof of Theorem 2.3 for a monotone twist map Define $L(q) = S(q, q)$. We first argue that a critical point of L corresponds to a fixed point of Φ. Indeed, if $L'(q^0) = 0$, then $S_q(q^0, q^0) + S_Q(q^0, q^0) = 0$. Since $\Phi(q^0, -S_q(q^0, q^0)) = (q^0, S_Q(q^0, q^0))$, we deduce that $\Phi(q^0, y^0) = (q^0, y^0)$ for $y^0 = -S_q(q^0, q^0) = S_Q(q^0, q^0)$. On the other hand, by (2.3), we have that $L(q + 1) = L(q)$. Either L is identically constant which yields a continuum of fixed points for Φ, or L is not constant. In the latter case, L has at least two distinct critical points, namely a maximizer and minimizer. These yield two distinct critical points of ϕ.

□

See for example [MS] for a proof of Theorem 2.3 for general twist maps.

To see Poincaré-Birkhoff's theorem within a larger context, we interpret it in the following way: since $0 \in \big(\rho(\Phi^-), \rho(\Phi^+)\big)$, then φ has at least two orbits in the interior of the cylinder that are associated with 0 rotation number, namely fixed points. In fact an analogous result is true for periodic orbits that is in the same spirit as Theorem 1.1**(iv)**.

Theorem 2.4 *(Birkhoff) Let $\varphi : \mathbb{S}^1 \times [-1, 1] \to \mathbb{S}^1 \times [-1, 1]$, be an area and orientation preserving C^1-diffeomorphism. If $r/s \in \big(\rho(\Phi^-), \rho(\Phi^+)\big)$ is a rational number with r and s coprime, then φ has at least two (r, s)-periodic orbits in the interior of $\mathbb{S}^1 \times [-1, 1]$.*

We may wonder whether a similar strategy as in the proof of Theorem 2.3 can be used to Prove Theorem 2.4 when φ is a monotone area-preserving map. Indeed if Φ is a monotone twist map, then we can associate with it a variational principle which is the discrete analog of the Principle of Least Action, as can be seen in the following proposition.

Proposition 2.2 *Let Φ be a monotone twist map with generating function S. Fix an integer $n \geq 2$.*

(i) *Given q and $Q \in \mathbb{R}$, define*

$$L(q_1, q_2, \ldots, q_{n-1}; q, Q) = \sum_{j=0}^{n-1} S(q_j, q_{j+1}),$$

with $q_0 = q$, and $q_n = Q$. Then $(q_1, q_2, \ldots, q_{n-1})$ is a critical point of $L(\cdot; q, Q)$ iff there exist $p_0, p_1, \ldots, p_n$ such that $\Phi(q_j, p_j) = (q_{j+1}, p_{j+1})$ for $j = 0, 1, 2, \ldots, n-1$.

(ii) *Given a positive integer r, define*

$$K(q_0, q_1, \ldots, q_{n-1}) = S(q_{n-1}, q_0 + r) + \sum_{j=0}^{n-2} S(q_j, q_{j+1}).$$

Then $(q_0, q_1, \dots, q_{n-1})$ *is a critical point of* K *if and only if there exist* $p_0, p_1, p_2, \dots, p_n$ *such that* $\Phi(q_j, p_j) = (q_{j+1}, p_{j+1})$ *for* $j = 0, \dots, n-1$*, with* $q_n = q_0 + r$*.*

Proof We only prove **(ii)** because **(i)** can be proved by a verbatim argument. Let $(q_0, \dots, q_{n-1})$ be a critical point and set $q_n = q_0 + r$. We also set $p_j = -S_q(q_j, q_{j+1})$. The result follows because if $P_j = S_Q(q_j, q_{j+1})$, then

$$K_{q_j} = p_j - P_{j-1}, \qquad \Phi(q_j, p_j) = (q_{j+1}, P_j),$$

for $j = 0, 1, 2, \dots, n-1$. $\square$

As we mentioned earlier, Theorem 2.4 for monotone twist maps can be established with the aid of Proposition 2.2. See for example [KH] or [G] for a reference.

Remark 2.2 Naturally we are led to the following question: Can we find an orbit of φ associated with such $\rho \in \big(\rho(\Phi^-), \rho(\Phi^+)\big)$? The answer to this question is affirmative and this is the subject of the *Aubry-Mather Theorem.* For any irrational $\rho \in \big(\rho(\Phi^-), \rho(\Phi^+)\big)$, there exists an invariant set on the cylinder that in some sense has the rotation number ρ. This invariant set q-projects onto either a Cantor-like subset of S^1 or the whole S^1. The invariant set lies on a graph of a Lipschitz function defined on S^1. These invariant sets are known as *Aubry–Mather sets.* $\square$

Arnold formulated an influential conjecture that is a vast generalization of Theorem 2.3 to higher dimensions. Given a Hamiltonian function $H : M \times \mathbb{R} \to \mathbb{R}$ on a closed symplectic manifold (M, ω), we may wonder whether or not the corresponding Hamiltonian vector field $X_H = X_H^\omega$ has T-periodic orbits for a given period T. *Arnold's Conjecture* offers a non-trivial lower bounds on the number of such periodic orbits. To convince that such a question is natural and important, let us examine this question when the Hamiltonian function is time-independent first. We note that for the autonomous X_H we can even find rest points (or constant orbits) and there is a one-one correspondence between the constant orbits of X_H and the critical points of H. We can appeal to the following classical theories in Algebraic Topology to obtain sharp universal lower bounds on the number of critical points of a smooth function on M where M is a smooth closed manifold. Let us write $Crit(H)$ for the set of critical points of $H : M \to \mathbb{R}$.

(i) According to Lusternik-Schnirelmann (LS) Theorem,

$$\sharp Crit(H) \geq c\ell(M), \tag{2.4}$$

where $c\ell(M)$ denotes the *cuplength* of M.

(ii) According to Morse Theory, for a Morse function H,

$$\sharp Crit(H) \geq \sum_k \beta_k(M), \tag{2.5}$$

where $\beta_k(M)$ denotes the k-th *Betti's number* of M.

According to Arnold's conjecture, the analogs of (2.4) and (2.5) should be true for the non-autonomous Hamiltonian functions provided that we count 1-periodic orbits of X_H in place of constant orbits. For the sake of comparison, we may regard (2.4) and (2.5) as a lower bound on the number of 0-periodic orbits when H is 0-periodic in t. In Arnold's conjecture, we replace 0-periodicity with 1-periodicity. Note that if H is 1-periodic in time, then $\phi_{t+1}^H(x) = \phi_t^H(x)$ for all t iff $\phi_1^H(x) = x$. To this end, we define

$$\mathbb{F}ix(H) := \big\{x \in M : \ \phi_1^H(x) = x\big\} =: Fix\big(\phi_1^H\big). \tag{2.6}$$

Arnold's Conjecture: Let (M, ω) be a closed symplectic manifold and let $H : M\times[0, \infty) \to \mathbb{R}$ be a smooth Hamiltonian function that is 1-periodic in the time variable. Then

$$\sharp\mathbb{F}ix(H) \geq c\ell(M). \tag{2.7}$$

Moreover, if $\varphi := \phi_1^H$ is non-degenerate in the sense that $\det(d\varphi - id)_x \neq 0$ for every $x \in Fix(\varphi)$, then

$$\sharp\mathbb{F}ix(H) \geq \sum_k \beta_k(M). \tag{2.8}$$

□

We now describe a strategy for tackling Arnold's conjecture under some additional conditions on M: We may establish the Arnold's conjecture by studying the set of critical points of $\mathcal{A}_H : \Gamma \to \mathbb{R}$, where Γ is the space of 1-periodic $x : \mathbb{S}^1 \to M$ and

$$\mathcal{A}_H(x(\cdot)) = \int_w \omega - \int_{\mathbb{S}^1} H(x(t), t)\ dt. \tag{2.9}$$

where $w : \mathbb{D} \to M$ is any extension of $x : \mathbb{S}^1 \to M$ to the unit disc $\mathbb{D}$. Note that the right-hand side of (2.9) would be independent of the extension w if the symplectic form ω is *aspherical* i.e., $\int_{f(\mathbb{S}^2)} \omega = 0$ for every smooth map $f : \mathbb{S}^2 \to M$. We may try to apply LS and Morse Theory to the functional $\mathcal{A}_H$ in order to get lower bounds on $\sharp\mathbb{F}ix(H)$. Of course we cannot apply either Morse Theorem (2.5) or LS Theorem (2.4) to $\mathcal{A}_H$ directly because Γ is neither compact nor finite-dimensional. However in the case of a torus or the cotangent bundle of a torus (namely $M = \mathbb{T}^d \times \mathbb{R}^d$), we may reduce the dimension to a finite (possibly very large) number by using *generalized generating functions* (see [G] for example). In fact, one can show that ϕ_t^H has a type II or III generating function (as we discussed in Subsections 1.8 and 1.9) provided that t is sufficiently small. We then use the group property of the flow to write

$$\varphi = \phi_1^H = \psi_1 \circ \cdots \circ \psi_N,$$

where each ψ_i has a generating function. This can be used to build a generalized generating function for φ à la Chaperon [C2]. We may establish Arnold's conjecture with the aid of generalized generating functions in some cases. Arnold's conjecture was established by Conley and Zehnder when $M = \mathbb{T}^{2d}$.

Theorem 2.5 *Assume that $\varphi = \phi_1^H$, for a smooth Hamiltonian function $H : \mathbb{T}^{2d} \times \mathbb{R} \to \mathbb{R}$ such that $H(x, t + 1) = H(x, t)$ for every $(x, t) \in \mathbb{T}^{2d} \times \mathbb{R}$. Then φ has at least $2d + 1$ fixed points.*

We first prove Theorem 2.5, when the map ϕ has a global generating function. Before embarking on this, we make some observations and state some definitions.

For our purposes, it is more convenient to think of the Hamiltonian function as a function $H : \mathbb{R}^{2d} \times \mathbb{R} \to \mathbb{R}$ that is 1-periodic in all the coordinates of (x, t). (With a slight abuse of notion, this Hamiltonian function is also denoted by H.) The flow of this Hamiltonian function is denoted by $\Phi_t^H : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$. Note that $\Phi := \Phi_1^H$ is a lift of φ of Theorem 2.5.

Definition 2.4(i) Let us write $\mathcal{H} = \mathcal{H}(\mathbb{R}^{2d})$ for the space of C^2 Hamiltonian functions $H : \mathbb{R}^{2d} \times \mathbb{R} \to \mathbb{R}$. For each $a = (b, c) \in \mathbb{R}^d \times \mathbb{R}^d$, we define

$$(\tau_b H)(q, p, t) = H(q+b, p, t), \quad (\eta_c H)(q, p, t) = H(q, p+c, t), \quad (\theta_a H)(q, p, t) = H(q+b, p+c, t).$$

(ii) We write $\mathcal{C}^1$ for the set of C^1 maps $\Phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$. We set $\mathcal{F}(\Phi) = \Phi - id$, where id denotes the identity map. We write $\mathcal{S}$ for the set of symplectic diffeomorphism $\Phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ and set $\tilde{\mathcal{S}} = \mathcal{F}\big(\mathcal{S}\big)$. For $a \in \mathbb{R}^{2d}$, the translation operators $\theta_a : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ and $\theta_a, \theta_a' : \mathcal{C}^1 \to \mathcal{C}^1$ are defined by

$$\theta_a(x) = x + a, \quad \theta_a F = F \circ \theta_a, \quad \theta_a' = \mathcal{F}^{-1} \circ \theta_a \circ \mathcal{F},$$

for $x \in \mathbb{R}^{2d}$ and $F \in \mathcal{C}^1$. Note that for $\Phi \in \mathcal{C}^1$,

$$\big(\theta_a' \Phi\big)(x) = \big(\theta_{-a} \circ \Phi \circ \theta_a\big)(x) = \Phi(x + a) - a.$$

(iii) Let Φ be a symplectic diffeomorphsim with

$$\Phi(q, p) = \big(Q(q, p), P(q, p)\big).$$

We say that Φ is *exact* if for every $p \in \mathbb{R}^d$, the map $q \mapsto Q(q, p)$ is a diffeomorphism of $\mathbb{R}^d$. We write $\hat{q}(Q, p)$ for the inverse:

$$Q(q, p) = Q \quad \Leftrightarrow \quad q = \hat{q}(Q, p).$$

We also set $\hat{P}(Q, p) = P\big(\hat{q}(Q, p), p\big)$, and

$$\widehat{\Phi}(Q, p) = \big(\hat{q}(Q, p), \hat{P}(Q, p)\big), \quad \widetilde{\Phi}(Q, p) = \big(\hat{P}(Q, p), \hat{q}(Q, p)\big).$$

$\square$

Proposition 2.3 **(i)** *We have* $\mathcal{F}\big(\theta'_a\Phi\big) = \theta_a\mathcal{F}(\Phi)$, *and*

$$\phi^{\theta_a H} = \theta_{-a} \circ \phi^H \circ \theta_a = \theta'_a \phi^H. \tag{2.10}$$

In particular, if H *is* 1*-periodic, i.e.,* $\theta_n H = H$*, for all* $n \in \mathbb{Z}^{2d}$*, and* $\Phi = \phi_1^H$*, then* $\mathcal{F}\big(\Phi\big)$ *is also* 1*-periodic.*

(ii) *For every exact* Φ*, and* $a \in \mathbb{R}^d$*, we have*

$$\widehat{\theta'_a\Phi} = \theta'_a\widehat{\Phi}.$$

In particular, if $\mathcal{F}(\Phi)$ *is* 1*-periodic, then so is* $\mathcal{F}\big(\widehat{\Phi}\big)$*.*

(iii) *Assume that* $\Phi \in \mathcal{S}$ *is exact. Then there exists a* C^2 *function* $W : \mathbb{R}^{2d} \to \mathbb{R}$ *such that* $\widetilde{\Phi} = \nabla W$*.*

(iv) *If* $\mathcal{F}(\Phi)$ *is* 1*-periodic, with*

$$\int_{\mathbb{T}^{2d}} \mathcal{F}(\Phi)(x)\ dx = 0,$$

then

$$W(Q,p) = Q \cdot p - w(Q,p),$$

for a function w *that is* 1*-periodic.*

Proof(i) The proof of $\mathcal{F}\big(\theta'_a\Phi\big) = \theta_a\mathcal{F}(\Phi)$ is straightforward and is omitted. The claim (2.10) is an immediate consequence of the fact that if $y(\cdot)$ is an orbit of $X_{\theta_a H}$, then $x(\cdot) = \theta_{-a}y(\cdot) = y(\cdot) - a$ is an orbit of X_H.

(ii) Fix $a = (b,c) \in \mathbb{R}^{2d}$. Let us define

$$\Phi^a(q,p) := (\theta'_a\Phi)(q,p) = \big(Q^a(q,p), P^a(q,p)\big), \qquad \widehat{\Phi}^a(Q,p) = \big(\hat{q}^a(Q,p), \hat{P}^a(Q,p)\big).$$

We certainly have

$$\begin{aligned} Q(q+b,p+c) - b = Q^a(q,p) = Q \quad &\Leftrightarrow \quad \hat{q}^a(Q,p) = q, \\ Q(q+b,p+c) = Q + b \quad &\Leftrightarrow \quad \hat{q}(Q+b,p+c) = q + b. \end{aligned}$$

Hence $\hat{q}^a(Q,p) = \hat{q}(Q+b,p+c) - b$. On the other hand

$$\begin{aligned} \hat{P}^a(Q,p) &= P^a\big(\hat{q}^a(Q,p),p\big) = P\big(\hat{q}^a(Q,p) + b, p + c\big) - c \\ &= P\big(\hat{q}(Q+b,p+c), p+c\big) - c = \hat{P}\big(Q+b,p+c\big) - c, \end{aligned}$$

as desired.

(iii) Since Φ is symplectic, we have

$$\begin{aligned} d\big(\hat{P}\cdot dQ + \hat{q}\cdot dp\big) &= d\big(\hat{P}\cdot dQ + d(p\cdot\hat{q}) - p\cdot d\hat{q}\big) \\ &= d\big(\hat{P}\cdot dQ - p\cdot d\hat{q}\big) = d\big(P\cdot dQ - p\cdot dq\big) = 0. \end{aligned}$$

Hence, there exists a function $W = W(Q,p)$ such that

$$dW = \hat{P}\cdot dQ + \hat{q}\cdot dp.$$

As a result, $\nabla W = \widetilde{\Phi}$.

(iv) Define

$$\hat{G} := \mathcal{F}\big(\widehat{\Phi}\big), \qquad w(Q,p) := Q\cdot p - W(Q,p), \qquad \widehat{\nabla} w := (w_p, w_Q).$$

We certainly

$$\begin{aligned} (Q,p) + \hat{G}(Q,p) &= \nabla W(Q,p) = \big(W_p(Q,p), W_Q(Q,p)\big) \\ &= \big(Q - w_p(Q,p), p - w_Q(Q,p)\big) = (Q,p) - \widehat{\nabla} w(Q,p). \end{aligned}$$

In summary, $\hat{\nabla} w = -G$. By **(ii)** we know that $\hat{G}$ is a 1-periodic function. We wish to show that w is also a 1-periodic function. For this, it suffices to show

$$\int_{[0,1]^{2d}} \hat{\nabla} w(y)\ dy = -\int_{[0,1]^{2d}} \hat{G}(y)\ dy = 0.$$

(Here $y = (Q,p)$.) To verify this, observe that if

$$A := (B,C) := \int_{[0,1]^{2d}} \hat{G}(y)\ dy, \qquad B, C \in \mathbb{R}^d,$$

then there exits a C^2 periodic function $v(Q,p)$ such that $\hat{G} - A = -\hat{\nabla} v$, or

$$\hat{P}(Q,p) = C + p - v_Q(Q,p), \qquad \hat{q} = B + Q - v_p(Q,p).$$

On the other hand, by assumption,

$$\begin{aligned} 0 &= \int_{[0,1]^{2d}} G(q,p)\ dqdp = \int_{[0,1]^{2d}} \big(Q(q,p) - q, P(q,p) - p\big)\ dqdp \\ &= \int_{[0,1]^{2d}} \big(Q - \hat{q}(Q,p), \hat{P}(Q,p) - p\big)\ \det\big(\hat{q}_Q(Q,p)\big)\ dQdp \\ &= \int_{[0,1]^{2d}} \big(v_p(Q,p) - B, C - v_Q(Q,p)\big)\ \det\big(I - v_{Qp}(Q,p)\big)\ dQdp \\ &= (-B,C) + \int_{[0,1]^{2d}} J\nabla v(Q,p)\ \det\big(I - v_{Qp}(Q,p)\big)\ dQdp, \end{aligned}$$

where I denotes the $(2d)\times(2d)$ identity matrix. We are done if we can show

$$\int_{[0,1]^{2d}} \nabla v(Q,p)\ \det\big(I - v_{Qp}(Q,p)\big)\ dQdp = 0. \tag{2.11}$$

The proof of this is left as an exercise. $\square$

Exercise Verify (2.11). $\square$

With the aid of Proposition 2.3, we can establish Theorem 2.5 when Φ (the lift of φ) is exact in the sense of Definition 2.4**(iii)**. The proof can be carried out in exactly the same way that we proved Theorem 2.3 for monotone twist maps. To go beyond exact maps, we first express $\Phi = \Phi_H^1$ as a finite composition of exact maps and use their generating functions to construct a (generalized) generating function of type II for Φ.

Proposition 2.4 *Let $\Phi_i, i = 1,\dots,k$, be k exact symplectic diffeomorphisms with generating functions $W^i(Q,p) = Q\cdot p - w^i(Q,p), i = 1,\dots,k$, respectively. Let $\Phi = \Phi_k\circ\dots\circ\Phi_1$.*

(i) *With $p_0 = p, q_k = Q$, and $\xi = \big(q_1, p_1, \dots, q_{k-1}, p_{k-1}\big)$, define*

$$W(Q,p;\xi) = \sum_{i=1}^{k} W^i(q_i, p_{i-1}) - \sum_{i=1}^{k-1} q_i\cdot p_i =: Q\cdot p + w(Q,p;\xi).$$

Then

$$W_\xi(Q,p;\xi) = 0 \quad\Longrightarrow\quad \Phi\big(W_p(Q,p;\xi),p\big) = \big(Q, W_Q(Q,p;\xi)\big). \tag{2.12}$$

In particular, if the full derivative ∇W of W with respect to its arguments Q,p and ξ vanishes at some point $(\bar Q,\bar p,\bar\xi)$, then $(\bar Q,\bar p)$ is a fixed point of Φ.

(ii) *Given $\mathbf{x} = (x_0,\dots,x_{k-1})$, $x_0 = (q_0,p_0),\dots,x_{k-1} = (q_{k-1},p_{k-1})$, define*

$$\mathcal{A}^k(\mathbf{x}) = \sum_{i=1}^{k} W^i(q_i,p_{i-1}) - \sum_{i=1}^{k} q_i\cdot p_i,$$

with $x_0 = x_k = (q_k,p_k)$. (In other words, $\mathcal{A}^k$ is defined for k-periodic sequences.) Then any critical point $\mathbf{x}$ of $\mathcal{A}^k$ yields an orbit $\Phi_i(x_{i-1}) = x_i, i = 1,\dots,k$. In particular $x_0 = x_k$ is a fixed point of Φ.

Proof(i) If we write $\hat q_{i-1} = W_p^i\big(q_i,p_{i-1}\big)$, and $\hat p_i = W_Q^i\big(q_i,p_{i-1}\big)$, then $\Phi_i\big(\hat q_{i-1},p_{i-1}\big) = \big(q_i,\hat p_i\big)$. On the other hand, for $i = 1,\dots,k-1$,

$$\begin{aligned} W_{q_i}(Q,p;\xi) &= \hat p_i - p_i, & W_{p_i}(Q,p;\xi) &= \hat q_i - q_i,\\ W_p(Q,p;\xi) &= W_p^1(q_1,p) = \hat q_0, & W_Q(Q,p;\xi) &= W_Q^k(Q,p_k). \end{aligned}$$

From this, we can readily deduce (2.12).

(ii) As in Part **(i)**,

$$\begin{aligned} \mathcal{A}^k_{q_i}(\mathbf{x}) &= \hat{p}_i - p_i, & \mathcal{A}^k_{p_i}(\mathbf{x}) &= \hat{q}_i - q_i, \\ \mathcal{A}^k_{q_k}(\mathbf{x}) &= \hat{p}_k - p_k, & \mathcal{A}^k_{p_0}(\mathbf{x}) &= \hat{q}_0 - q_0 \end{aligned}$$

for $i = 1, \dots, k-1$. Hence at a critical point we have $\Phi_i(x_{i-1}) = x_i$ for $i = 1, \dots, k$. This completes the proof. □

Remark 2.3 Note that $\mathcal{A}^k$ can be written as

$$\mathcal{A}^k(\mathbf{x}) = \sum_{i=1}^{k} \big(p_{i-1} \cdot (q_i - q_{i-1}) - w^i(q_i, p_{i-1})\big),$$

which is a discrete variant of (1.17). □

Proof of Theorem 2.5 *(Sketch)* For some sufficiently large k, we can find exact symplectic diffeomorphisms $\Phi_i, i = 1, \dots, k$, such that $\Phi = \Phi_k \circ \dots \circ \Phi_1$. In Proposition 2.4**(ii)**, we found a one-to-one correspondence between a fixed point x_0 of Φ, and a critical point $\mathbf{x} = (x_0, \dots, x_{k-1})$ of $\mathcal{A}^k$. Observe that from Proposition 2.3**(iv)** we know that $w^1, \dots, w^k$ are periodic. Let us write $\mathcal{A}^k = \mathcal{A}^k_0 - \mathbf{w}$, where

$$\mathcal{A}^k_0(\mathbf{x}) = \sum_{i=1}^{k} p_{i-1} \cdot (q_i - q_{i-1}), \qquad \mathbf{w}(x) = \sum_{i=1}^{k} w^i(q_i, p_{i-1}).$$

To ease the notation, let us write $z_i = x_i - x_{i-1} = (q_i', p_i')$, and $\mathbf{z} = (z_1, \dots, z_{k-1})$. Since $x_k = x_0$, we may rewrite $\mathcal{A}^k_0$ as

$$\mathcal{A}^k_0(\mathbf{x}) = \sum_{i=1}^{k} (p_{i-1} - p_0) \cdot (q_i - q_{i-1}) = \sum_{i=1}^{k} (p_{i-1}' + \dots + p_1') \cdot q_i',$$

$$\begin{aligned} \mathcal{A}^k_0(\mathbf{x}) &= -\sum_{i=1}^{k} (p_i - p_{i-1}) \cdot q_i = -\sum_{i=1}^{k} (p_i - p_{i-1}) \cdot (q_i - q_0) \\ &= -\sum_{i=1}^{k} (q_i' + \dots + q_1') \cdot p_i'. \end{aligned}$$

Using this, we can express $\mathcal{A}^k_0(\mathbf{x})$ as $2^{-1}B\mathbf{z} \cdot \mathbf{z}$, for a matrix $B = [B_{ij}]_{i,j=1}^{k-1}$, with each B_{ij} a $(2d) \times (2d)$ matrix. We may express B as

$$B = \begin{bmatrix} 0 & C \\ -D & 0 \end{bmatrix},$$

with both C and D invertible. Hence B is non-singular. Since for each $m \in \mathbb{Z}^{2d}$,

$$\mathcal{A}^k(x_0 + m, \dots, x_{k-1} + m) = \mathcal{A}^k(x_0, \dots, x_{k-1}),$$

we can write

$$\mathcal{A}^k(\mathbf{x}) = \frac{1}{2} B\mathbf{z} \cdot \mathbf{z} + \hat{\mathbf{w}}(x_0, \mathbf{z}),$$

for a bounded C^2 function $\hat{\mathbf{w}}(x_0, \mathbf{z})$ that is periodic in x_0. Let us $\mathbf{y} = (x_0, \mathbf{z})$, and $\mathcal{B}(\mathbf{y})$ for $\mathcal{A}^k(\mathbf{x})$ in these new coordinates. Observe that $\mathcal{B} : \mathbb{T}^{2d} \times \mathbb{R}^{2d(k-1)} \to \mathbb{R}$ is a bounded perturbation of a non-degenerate quadratic function $\mathbf{z} \mapsto 2^{-1} B\mathbf{z} \cdot \mathbf{z}$. We may study the set of critical points of $\mathcal{B}$ by analyzing the corresponding gradient flow $\dot{\mathbf{y}} = -\nabla \mathcal{B}(\mathbf{y})$. Equivalently,

$$\dot{\mathbf{z}} = B\mathbf{z} + \hat{\mathbf{w}}_{\mathbf{z}}(x_0, \mathbf{z}), \qquad \dot{x}_0 = \hat{\mathbf{w}}_{x_0}(x_0, \mathbf{z}). \tag{2.13}$$

If we write $\psi_t(\mathbf{y})$ for the flow of the equation (2.13), and X for the set $\mathbf{y}$ such that the corresponding orbit $\big(\psi_t(\mathbf{y}) : t \in \mathbb{R}\big)$ is bounded, then X inherits the topology of $\mathbb{T}^{2d}$. To explain this, observe that $X = \mathbb{T}^{2d} \times \{0\}$, when $\hat{\mathbf{w}} = 0$. In general, the projection map $(x_0, \mathbf{z}) \mapsto x_0$ from X to $\mathbb{T}^{2d}$ induces an injective map from the *Cech Homology* of $\mathbb{T}^{2d}$ to the Cech Homology of X. This allows us to deduce that (2.13) has at least $2d+1$ many constant solution (we refer to [HZ] for more details). □

Remark 2.4(i) The full proof of Theorem 2.5 as we sketched above can be found in [MS]. A similar proof has been used in [HZ] by studying critical points of the operator $\mathcal{A}_H$ of (1.13) directly.

(ii) A variant of Theorems 2.3 and 2.5 can be proved when the periodicity of $\Phi - id$ is replaced with almost periodicity, or even when $\Phi - id$ is selected randomly according to a translation invariant probability measure. See [PR1-2] for more details. □

Exercises(i) Let $b : \mathbb{R} \to \mathbb{R}$ be a positive 1-periodic function and write ϕ_t for the flow of the ODE $\dot{x} = b(x)$. Find the rotation number of this ODE by evaluating the following limit:

$$\lim_{t \to \infty} t^{-1} \big(\phi_t(x) - x\big).$$

Also, find a strictly increasing function $K : \mathbb{R} \to \mathbb{R}$ such that

$$K \circ \phi_t \circ K^{-1},$$

is a translation.

(ii) Define $\tau_a b(x) = b(x+a)$, and write $\mathcal{B}$ for the set of uniformly positive Lipschitz function $b : \mathbb{R} \to \mathbb{R}$. Let $\mathbb{P}$ be a τ-invariant ergodic probability measure on $\mathcal{B}$. For each b, write $\phi_t(x; b)$ for the flow of the ODE $\dot{x} = b(x)$. Show that $\mathbb{P}$-almost surely, the limit

$$\lim_{t \to \infty} t^{-1} \big(\phi_t(x; b) - x\big),$$

exists for every x. Evaluate this limit. □

3 Discrete Type Hamilton-Jacobi Equation

In Section 2 we learned how the critical points of the action functional yield the orbits of the corresponding dynamical system. In this chapter we focus on the critical values of the action functional. We also examine how the stochasticity can play a role. We may choose the generating function randomly according to a probability law, or add some noise to the dynamics.

3.1 Frenkel-Kontorova model

Imagine that we have a sequence of symplectic maps $\big(\Phi_i:\ i\in\mathbb{N}\big)$ such that each Φ_i has a type I generating function $S^i(q,Q)$, so that

$$\Phi_i\big(q,S^i_q(q,Q)\big)=\big(Q,-S^i_Q(q,Q)\big).$$

We may define a dynamical system with orbits $(x_0,x_1,\dots,x_n,\dots)$;

$$x_{i+1}=\Phi_{i+1}(x_i),\qquad \text{or}\qquad x_n=\Phi_n\circ\cdots\circ\Phi_1(x_0).$$

If $\Phi_i=\Phi$ is independent of i, then we have an autonomous dynamical system with $x_n=\Phi^n(x_0)$. Under some type of non-degeneracy assumptions on the generating functions, we may regard our system as a second order dynamical system in q components. By this we mean that if $(x_n:\ n=0,1,\dots)$ is an orbit with $x_i=(q_i,p_i)$, then $(q_n:\ n=0,1,\dots)$ is an orbit of the dynamical system with the rule $q_n=F_n\big(q_{n-2},q_{n-1}\big)$, where F_n is implicitly defined by

$$S^{n-1}_Q\big(q_{n-2},q_{n-1}\big)+S^n_q\big(q_{n-1},q_n\big)=0. \tag{3.1}$$

Moreover, given q and Q, we can find an orbit $(q_0,\dots,q_n)$, with $q_0=q, q_n=Q$, iff $\big(q_1,\dots,q_{n-1}\big)$ is a critical point of

$$\mathcal{S}^n\big(q_1,\dots,q_{n-1};q,Q\big)=\sum_{i=1}^n S^i\big(q_{i-1},q_i\big).$$

For the construction of invariant measures, we may consider the following variation: given a continuous function $g:\mathbb{R}^d\to\mathbb{R}$, consider

$$\mathcal{S}^n\big(q_0,q_1,\dots,q_{n-1};g;Q\big)=g(q_0)+\mathcal{S}^n\big(q_1,\dots,q_{n-1};q_0,Q\big).$$

Given q and Q, a critical point of $\mathcal{S}^n\big(q_0,q_1,\dots,q_{n-1};g;Q\big)$ yields an orbit $(x_0,\dots,x_n)$ of our dynamical system with the properties

$$p_0=-S^1_q(q_0,q_1)=\nabla g(q_0),\qquad p_n=S^n_Q(q_{n-1},Q).$$

As we mentioned in Section 2, it is more convenient to write $S^i(q,Q) = L^i(q,Q-q)$. Because of some of the examples we have in mind, it is quite natural to assume that

$$\lim_{|v|\to\infty} \inf \inf_q |v|^{-1} L^i(q,v) = \infty. \tag{3.2}$$

Note that this condition is satisfied for a standard map associated with $L(q,v) = |v|^2/2 - V(q)$, for a bounded C^1 function V. Assuming (3.2) is valid for each S^i, we define two operators

$$(\mathcal{T}_i g)(Q) = \inf_q \big(g(q) + S^i(q,Q)\big), \quad (\widehat{\mathcal{T}}_i g)(q) = \sup_Q \big(g(Q) - S^i(q,Q)\big), \tag{3.3}$$

on the space Λ of Lipschitz functions $g : \mathbb{R}^d \to \mathbb{R}$. Note that if $S(q,Q)$ is a generating function for Φ, then $S'(q,Q) = -S(Q,q)$ is a generating function for Φ^{-1}. We will see later that $\mathcal{T}_i g \in \Lambda$ when $g \in \Lambda$. Observe

$$u_n(Q) := \big(\mathcal{T}_n \circ \cdots \circ \mathcal{T}_1\big)(g)(Q) = \inf_{q_0,\dots,q_{n-1}} \big(g(q_0) + \mathcal{S}^n\big(q_1,\dots,q_{n-1};q_0,Q\big)\big).$$

We regard

$$u_n = \mathcal{T}_n\big(u_{n-1}\big), \quad u_0 = g,$$

as a discrete variant of the (time inhomogeneous) HJE, where g is the initial data. Similarly,

$$u_{-n} = \widehat{\mathcal{T}}_n\big(u_{1-n}\big), \quad \hat{u}_0 = g,$$

is a discrete HJE with final condition $u_0 = g$. In particular, when $S^i = S$ is independent of i, we simply have $u_n = \mathcal{T}^n(g)$, and $u_n = \widehat{\mathcal{T}}^n(g)$, where

$$u(Q) := (\mathcal{T}g)(Q) = \inf_q \big(g(q) + S(q,Q)\big), \quad \hat{u}(q) := (\widehat{\mathcal{T}}g)(q) = \sup_Q \big(g(Q) - S(q,Q)\big). \tag{3.4}$$

If we assume that $L(q,v) = S(q,q+v)$ has a super linear growth at infinity, then the inf in (3.4) can be replace with min.

Assumption 3.1 There exists constants c_0, c_1 and $\delta > 0, \alpha > 1$ such that

$$\inf_q L(q,v) \ge \delta|v|^\alpha - c_0, \quad \sup_q L(q,0) \le c_1, \tag{3.5}$$
$$\sup_q \sup_{|v|\le\ell} |L(q+z,v) - L(q,v)| \le c_2(\ell)|z|.$$

Proposition 3.1 *Assume that* (3.5) *holds and that* $|g(q') - g(q)| \le \ell|q'-q|$ *for all* q, q'. *Then*

$$(\mathcal{T}g)(Q) = \min_{q:|Q-q|\le\ell'} \big(g(q) + S(q,Q)\big), \tag{3.6}$$

$$|u(Q') - u(Q)| \le \ell''|Q'-Q|, \tag{3.7}$$

for $\ell' = c_0 + c_1 + \big(\delta^{-1}(\ell+1)\big)^{\frac{1}{\alpha-1}}$, *and* $\ell'' = \ell + c_2(\ell')$.

Proof Observe

$$g(q) + S(q,Q) \geq g(Q) - \ell|Q-q| + \delta|Q-q|^\alpha - c_0.$$

Hence

$$g(Q) + S(Q,Q) \leq g(q) + S(q,Q),$$

if $c_0+c_1 \leq \delta|v|^\alpha - \ell|v|$, for $v = Q-q$. Then note that $\delta|v|^\alpha - \ell|v| \geq |v|$ if $|v| \geq \big(\delta^{-1}(\ell+1)\big)^{\frac{1}{\alpha-1}}$. This implies (3.6).

If $u(Q) = g(q) + L(q, Q-q)$ for some Q with $|Q-q| \leq \ell'$, then for $q' = q + Q' - Q$,

$$\begin{aligned} u(Q') &\leq g(q') + L(q', Q-q) \leq g(q) + L(q, Q-q) + \ell|Q'-Q| + c_2(\ell')|Q'-Q| \\ &= u(Q) + \big(\ell + c_2(\ell')\big)|Q'-Q|, \end{aligned}$$

which proves (3.7). □

We now describe some plausible applications of the operators $\mathcal{T}$ and $\widehat{\mathcal{T}}$ for finding invariant sets for the dynamical system associated with the transformation Φ. Recall that by Proposition 3.1, for each Q, there exists a point q such that

$$u(Q) := (\mathcal{T}g)(Q) = g(q) + S(q,Q).$$

Let us write $q = q(Q)$ for a minimizer in (3.4), which could be multivalued. If g is differentiable at q, then we have $\nabla g(q) + S_q(q,Q) = 0$, and if we write $A(q,Q) = g(q) + S(q,Q)$, then $A_q(q,Q) = 0$ when $q = q(Q)$. For now let us assume that the function $q(\cdot)$ is single-valued and differentiable at Q. Under such assumptions, u is differentiable at Q, and

$$\nabla u(Q) = A_q(q,Q)\ \nabla q(Q) + A_Q(q,Q) = S_Q(q,Q).$$

As a result,

$$\Phi\big(q, \nabla g(q)\big) = \big(Q, \nabla u(Q)\big). \tag{3.8}$$

This suggests that if U solves the discrete Hamilton-Jacobi Equation $\mathcal{T}(U) = U + c$ for a constant c (or equivalently $\nabla\mathcal{T}(U) = \nabla U$ at any differentiability point of U), then the set

$$Gr(U) = \big\{(q, \nabla U(q)) : \ U \ \text{ differentiable at } \ q\big\},$$

may serve as an invariant set for Φ. We will discuss the relevance of the equation $\mathcal{T}(U) = U + c$ and $\widehat{\mathcal{T}}(U) = U + c'$ to the question of homogenization in Section 5.

3.2 Type II generating function

If we consider a symplectic map with a type II generating function $W(Q,p) = Q\cdot p - w(Q,p)$, then a candidate for the action is

$$A(q,p;Q) = A(x;Q) = g(q) + W(Q,p) - q\cdot p = g(q) + (Q-q)\cdot p - w(Q,p).$$

Let us assume that both g and w are differentiable functions. Given Q, at any critical point $x=(q,p)$ of A we have

$$0 = A_q(q,p;Q) = \nabla g(q) - p, \qquad 0 = A_p(q,p;Q) = W_p(Q,p) - q.$$

Imagine that we can find a function $x(\cdot)$ such that $A_x\big(x(Q);Q\big) = 0$. If the function $x(\cdot)$ is differentiable at some $\bar{Q}$, then $u(Q) := A\big(x(Q);Q\big)$ is also differentiable at $\bar{Q}$, and

$$\nabla u(\bar{Q}) = A_x(x(\bar{Q});\bar{Q})(\nabla x)(\bar{Q}) + W_Q(\bar{Q},\bar{p}) = W_Q(\bar{Q},\bar{p}),$$

where $x(\bar{Q}) = (\bar{q},\bar{p})$. From this and $\Phi\big(W_p(\bar{Q},\bar{p}),\bar{p}\big) = \big(\bar{Q}, W_Q(\bar{Q},\bar{p})\big)$ we deduce

$$\Phi\big(\bar{q},\nabla g(\bar{q})\big) = \big(\bar{Q},\nabla u(\bar{Q})\big).$$

In the case of type I generating function, we simply take the minimum of the action when L is bounded below (see (3.4)). This is no longer the case for type II generating function. For example if Φ is a lift of a symplectic map on the torus, then w is periodic and A is a periodic perturbation of the quadratic function $A^0(x;Q); = (Q-q)\cdot p$. Since 0, the only critical point of A^0 is a saddle point, the best we can hope for is that given Q, the function $A(\cdot;Q)$ has a saddle point which is of the same type as the type 0 is for $A^0(\cdot;Q)$. Now imagine that we come up with a universal way of *selecting* a critical value of A no matter what g is. This critical value yields an operator

$$\mathcal{V}(g)(Q) = A\big(x(Q);Q\big),$$

where $x(Q)$ is our selected critical point. A solution to the equation $\mathcal{V}(U) = U + c$, for a constant c, may be used to construct invariant sets of the map Φ.

More generally, assume that $\Phi = \Phi_k \circ \cdots \circ \Phi_1$ and each Φ_i has a generating function $W^i(q_i,p_{i-1}) = q_i\cdot p_{i-1} - w^i(q_i,p_{i-1})$. Then Φ has a (generalized) generating function of the form

$$\begin{aligned} W(q_k,p_0;\xi) &= W\big(q_k,p_0;q_1,p_1,\dots,q_{k-1},p_{k-1}\big) \\ &= q_1\cdot p_0 + \sum_{i=2}^{k} p_{i-1}\cdot(q_i - q_{i-1}) - \sum_{i=1}^{k} w^i(p_{i-1},q_i). \end{aligned}$$

Recall that by (2.12),

$$W_\xi(q_k, p_0; \xi) = 0 \quad \implies \quad \Phi\big(W_{p_0}(q_k, p_0; \xi), p_0\big) = \big(q_k, W_{q_k}(q_k, p_0; \xi)\big)$$

Given an initial data g, we set

$$\begin{aligned} A(\xi'; q_k) = A\big(q_1, p_1, \dots, q_{k-1}, p_{k-1}; q_k\big) &= g(q_0) - p_0 \cdot q_0 + W(q_k, p_0; \xi) \\ &= g(q_0) + \sum_{i=1}^{k} \big(p_{i-1} \cdot (q_i - q_{i-1}) - w^i(p_{i-1}, q_i)\big), \end{aligned}$$

where $\xi' = (q_0, p_0, \xi)$. To study the orbits of the map Φ, we may search for a function $\xi'(q_k)$ such that $A_{\xi'}\big(\xi'(q_k); q_k\big) = 0$ for every q_k. Setting $u_k(q_k) = A(q_k; \xi'(q_k))$, we have

$$p_0 = \nabla g(q_0), \quad \Phi\big(q_0, p_0\big) = \big(q_k, \nabla u_k(q_k)\big),$$

provided that g is differentiable at q_0, and ξ' is differentiable at q_k. In Section 4, we will address the question of selecting the critical point ξ'.

3.3 Gibbs measures

There is a viscous variant of the discrete HJE that is related to the orbits (or rather realizations) of a Markov chain. Given $S : M \times M \to \mathbb{R}$, recall the action function $\mathcal{S}^n$ of Subsection 3.1. Instead of minimizing $\mathcal{S}^n$, we define a probability measure on M^{n-1} that favors states $\mathbf{q}^n = (q_1, \dots, q_{n-1})$ of lower *energy* $\mathcal{S}^n$. This measure depends on a positive parameter $\beta > 0$ that represents the *inverse temperature*. More precisely, we define a *Gibbs measure* $\mathbb{P}_n(\cdot) = \mathbb{P}_n(\cdot; q, Q; \beta)$ on M^{n-1} as

$$\mathbb{P}(d\mathbf{q}^n) = Z_n(q, Q)^{-1} \exp\big(-\beta \mathcal{S}^n(\mathbf{q}^n; q, Q)\big) \prod_{i=1}^{n-1} \nu(dq_i),$$

where $\nu(dq)$ is a finite *reference measure* (for example a volume form associated with a metric when M is a Riemannian manifold), and Z is the normalizing constant:

$$Z_n(q, Q) = \int_{M^{n-1}} \exp\big(-\beta \mathcal{S}^n(\mathbf{q}^n; q, Q)\big) \prod_{i=1}^{n-1} \nu(dq_i).$$

This constant is finite if for example

$$\sup_{a,b \in M} \int_M \exp\big(-\beta S^i(a, q) - \beta S^{i+1}(q, b)\big)\ \nu(dq) < \infty$$

for every i. For simplicity, let us assume that $S^i = S$ for all i. Now, if we attempt to normalize our measure inductively, we need to calculate

$$Z(q_{n-2}, Q) := \int_M \exp\big(-\beta S(q_{n-2}, q_{n-1}) - \beta S(q_{n-1}, Q)\big)\ \nu(dq_{n-1}),$$

which depends on q_{n-2}. Dividing the integrand by $Z(q_{n-2}, Q)$ would alter S. To avoid this, observe that if we replace $S(q, Q)$ with $S(q, Q) + u(Q) - u(q)$, then the corresponding Gibbs measure would not be affected (it only changes the normalizing constant). Motivated by this, we define

$$\mathcal{R}_\beta(h)(g)(Q) = \int_M e^{-\beta S(q,Q)} h(Q)\ \nu(dQ), \qquad \mathcal{R}^*_\beta(h)(g)(Q) = \int_M e^{-\beta S(q,Q)} h(q)\ \nu(dq).$$

The operator $\mathcal{R}^*_\beta$ is the adjoint of $\mathcal{R}_\beta$ with respect to the inner product

$$\langle h, k\rangle = \int_M hk\ d\nu.$$

The celebrated Krein-Rutman Theorem (an infinite-dimensional generalization of Perron-Frobenius Theorem) offers a way of modifying S so that we can normalize our measure inductively.

For simplicity, let us assume that M is a compact metric space.

Theorem 3.1 *The largest eigenvalue $\lambda'_\beta = e^{\beta\lambda_\beta}$ of $\mathcal{R}_\beta$ is positive and λ'_β satisfies $\lambda'_\beta \geq |\lambda'|$ for any other eigenvalue λ'. Moreover λ'_β is simple, and there exist functions $u_\beta, u^*_\beta : M \to \mathbb{R}$ such that*

$$\mathcal{R}_\beta\big(e^{\beta u_\beta}\big) = e^{\beta\lambda_\beta} e^{\beta u_\beta}, \qquad \mathcal{R}^*_\beta\big(e^{-\beta u^*_\beta}\big) = e^{\beta\lambda_\beta} e^{-\beta u^*_\beta}.$$

See for example [L] for a reference for a proof of Theorem 3.1 and Krein-Rutman Theorem. Motivated by Theorem 3.1, we set

$$\begin{aligned}
\hat{S}(q, Q) &:= S(q, Q) - (u_\beta(Q) - u_\beta(q)) + \lambda_\beta,\\
p(q, dQ) &:= p(q, Q)\ \nu(dQ) := \exp\big(-\beta\hat{S}(q, Q)\big)\ \nu(dQ).
\end{aligned}$$

By Theorem 3.1, the *kernel* $p(q, dQ)$ is a probability measure for each q. Using this kernel, we may define a Markov chain $\mathbf{q} = (q_0, q_1, \dots, q_n, \dots)$ such that

$$\mathbb{P}^q\big(q_n \in A \mid q_0, \dots, q_{n-1}\big) = \int_A p(q_{n-1}, dq_n), \qquad q_0 = q,$$

for every measurable set $A \subseteq M$. Here $\mathbb{P}^q$ is a probability measure on the set of sequences $\mathbf{q}$ with $q_0 = q$. Hence

$$\begin{aligned}\mathbb{P}^q\big(q_1 \in A_1, \dots, q_n \in A_n\big) &= \int_{A_1} \dots \int_{A_n} \prod_{i=1}^n p(q_{i-1}, dq_i) \\ &= \int_{A_1} \dots \int_{A_n} \exp\left(-\sum_{i=1}^n \beta \hat{S}(q_{i-1}, q_i)\right) \prod_{i=1}^n \nu(dq_i).\end{aligned}$$

Writing $\mathbb{P}_n^q(dq_1, \dots, dq_n)$ for the n-dimensional marginal of $\mathbb{P}^q$, we deduce

$$\mathbb{P}_n\big(dq_1, \dots, dq_{n-1}; q, Q\big) = \mathbb{P}_n^q\big(dq_1, \dots, dq_n \mid q_n = Q\big).$$

Also, if we define

$$\hat{\mathcal{T}}_\beta(g) = \beta^{-1} \log \mathcal{R}_\beta\big(e^{\beta g}\big),$$

then

$$u_n = \hat{\mathcal{T}}_\beta(u_{n-1}),$$

is a discrete analog of viscous HJE. Note that we always have $\hat{\mathcal{T}}_\beta(g) \leq \hat{\mathcal{T}}(g)$. In fact

$$\lim_{\beta \to \infty} \hat{\mathcal{T}}_\beta(g) = \hat{\mathcal{T}}(g),$$

if for example $\nu(U) > 0$ for every nonempty open subset U of M: If

$$U_\delta(q) = \big\{Q \in M : \ g(Q) - S(q, Q) \geq \hat{\mathcal{T}}(q) - \delta\big\},$$

for $q \in M$, and $\delta > 0$, then $U_\delta(q)$ is a nonempty open set, and

$$\hat{\mathcal{T}}_\beta(g)(q) \geq \hat{\mathcal{T}}(g)(q) - \delta + \beta^{-1} \log \nu\big(U_\delta(q)\big) \to \hat{\mathcal{T}}(g)(q) - \delta,$$

as $\beta \to \infty$.

In the same vein, we set

$$\mathcal{T}_\beta(g) = -\beta^{-1} \log \mathcal{R}_\beta^*\big(e^{-\beta g}\big),$$

so that

$$u_{-n} = \mathcal{T}_\beta(u_{1-n}),$$

is a discrete analog of backward viscous HJE. Also

$$\lim_{\beta \to \infty} \mathcal{T}_\beta(g) = \mathcal{T}(g). \tag{3.9}$$

We note

$$\hat{\mathcal{T}}(u_\beta) = u_\beta + \lambda_\beta, \qquad \mathcal{T}(u^*_\beta) = u^*_\beta - \lambda_\beta,$$

which is the analog of (3.7). Moreover, the eigenfunctions $e^{\beta u_\beta}$, and $e^{-\beta u^*_\beta}$, can be used to find an invariant measure for our Markov Chain. For this, observe that if we look for an invariant measure of the form $d\mu = Z^{-1}e^h \, d\nu$, the function h must satisfy

$$e^{h(Q)} = \int e^{h(q)} p(q,Q) \, \nu(dq) = e^{\beta(u_\beta(Q)-\lambda_\beta)} R^*_\beta \left(e^{h-\beta u_\beta}\right)(Q),$$

which holds, if we choose h so that $e^{h-\beta u_\beta} = e^{-\beta u^*_\beta}$. Hence for an invariant measure, we may choose a measure of the form

$$\mu(dq) = Z^{-1} e^{\beta(u_\beta - u^*_\beta)(q)} \, dq,$$

where Z is the normalizing constant.

As (3.9) indicates, the zero-temperature limit of our Gibbs measure $\mathbb{P}$ is associated with the Frenkel-Kontorova model of Subsection 3.1. We refer to Anantharaman [An] for some deep results regarding the type of limiting measure we obtain as $\beta \to \infty$.

4 Variational and Viscosity Solutions

In Subsection 1.8 we learned that critical points of the action functional $\mathcal{A}_H$ are the orbits of the Hamiltonian ODE associated with the Hamiltonian function H. This was used in Section 2 to prove the existence of periodic orbits. We also argued that the *critical values* of $\mathcal{A}_H$ yield solutions to HJE associated with the Hamiltonian function H. Though our derivations of the HJEs (1.18) and (1.19) were rather formal. For example, the derivation of (1.18), requires the existence a global C^1 generating function which is hardly the case. In this section, we focus on the HJE and try to figure out how generalized solutions can be constructed. Insisting on constructing a solution as a critical value of the action $\mathcal{A}_H$ would lead to the notion of *variational solutions* to HJEs. However, HJE also appears as a model of stochastic growth. Statistical mechanical considerations suggest an alternative strategy for constructing solutions: We may add a small viscous term to the HJE to guarantee the existence of a global solution, and then send the viscosity to 0. This yields the notion of *viscosity solutions*. Surprisingly viscosity solutions may differ from variational solutions when the Hamiltonian function is not convex in the momentum variables. As we saw in Subsections 1.5-1.6, solutions to HJE may be used to construct invariant measures for the corresponding Hamiltonian ODE. This has been the case for Tonelli Hamiltonians. For such Hamiltonians viscosity solutions coincide with variational solutions. One may hope to use variational solutions to come up with an analog of weak KAM theory for non convex Hamiltonian. Viterbo's work [V] settles the question of homogenization for such Hamiltonian functions.

4.1 Variational solutions

Let $\Phi : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ be a symplectic map with a generating function $W(Q,p) = Q\cdot p - w(Q,p)$. In Subsection 3.2 we learned that if g is a C^1 function, and

$$A\big(q_0, p_0, \dots, q_{n-1}, p_{n-1}; q_n; g\big) = g(q_0) + \sum_{i=1}^{k} \big(p_{i-1}\cdot(q_i - q_{i-1}) - w(q_i, p_{i-1})\big),$$

then a critical point of A yields an orbit $x_i = (q_i,p_i) = \Phi^i(x_0)$, $i = 1,\dots,n$, with $x_0 = (q_0,p_0)$, and $p_0 = \nabla g(q_0)$. Motivated by this, let us define

$$\mathcal{W}_n(x_0) = \sum_{i=1}^{n} \big(p_{i-1}\cdot(q_i - q_{i-1}) - w(q_i, p_{i-1})\big),$$

where $x_i = \Phi^i(x_0)$ for $i = 1,\dots,n$. In other words, $\mathcal{W}_n(x_0)$ denotes the action at time n of an orbit that starts from x_0. We then set

$$\mathcal{F}_n(g) = \big\{\big(Q, g(q) + \mathcal{W}_n\big(q, \nabla g(q)\big)\big) : \ q \in \mathbb{R}^d, \ \Phi^n(q, \nabla g(q)) = (Q,P) \ \text{ for some } P \in \mathbb{R}^d\big\}.$$

We may extend the definition of $\mathcal{F}_n$ to Lipschitz g. Recall that Λ denotes the set of Lipschitz functions $g : \mathbb{R}^d \to \mathbb{R}$.

Definition 4.1(i) Given $g \in \Lambda$, we write $\hat{\partial} g(q)$ for the set of vectors p such that there exists a sequence q_k for which the following conditions hold:

$$\nabla g(q_k) \quad \text{exists and} \quad q = \lim_{k \to \infty} q_k, \qquad p = \lim_{k \to \infty} \nabla g(q_k).$$

The convex hull of the set $\hat{\partial} g(q)$ is denoted by $\partial g(q)$.

(ii) Given $g \in \Lambda$, we set

$$\mathcal{F}_n(g) = \left\{ (q_n, g(q_0) + \mathcal{W}_n(q_0, p_0)) : \ q_0 \in \mathbb{R}^d, p_0 \in \partial g(q_0), \ \Phi^n(q_0, p_0) = (q_n, p_n) \right\}.$$

(iii) By a *variational solution* associated with Φ, we mean a collection of operators $\mathcal{V}_n = \mathcal{V}_n^w : \Lambda \to \Lambda, \ n \in \mathbb{N}$ with the following properties:

- $\mathcal{V}_n(g + c) = \mathcal{V}_n(g) + c$ for each n and every constant $c \in \mathbb{R}$.
- For $g, g' \in \Lambda$ with $g \leq g'$, we have $\mathcal{V}_n(g) \leq \mathcal{V}_n(g')$.
- For every $g \in \Lambda$, and $n \in \mathbb{N}$,

$$\left\{ \big(q, \mathcal{V}_n(g)(q)\big) : \ q \in \mathbb{R}^d \right\} \subseteq \mathcal{F}_n(g).$$

□

In the same fashion, variational solutions of the HJE (1.10) are defined. For this, let us assume that $H : \mathbb{R}^{2d} \to \mathbb{R}$ is a C^2 Hamiltonian function such that $D^2 H$ is uniformly bounded. For this H, the corresponding flow Φ^H is well-defined. Recall that for $\gamma : [0, t] \to \mathbb{R}^{2d}$, with $\gamma(s) = (q(s), p(s))$, the action is defined by

$$\mathcal{A}_t(\gamma) = \mathcal{A}_t^H(\gamma) = \int_0^t \big[p \cdot \dot{q} - H(\gamma) \big] \, ds.$$

Definition 4.2(i) We set $\phi^H_{[0,t]}(a)$ for the restriction of the flow $\phi^H_s(a)$ to the interval $[0, t]$. Given $a \in \mathbb{R}^{2d}$, we define

$$A_t^H(a) = \mathcal{A}_t^H\big(\phi^H_{[0,t]}(a)\big).$$

(ii) Given a Lipschitz function g, we set

$$\mathcal{F}_t(g) = \left\{ \big(q(t), g(q_0) + A_t^H(q_0, p_0)\big) : \ q_0 \in \mathbb{R}^d, p_0 \in \partial g(q_0), \ \phi_t^H(q_0, p_0) = (q(t), p(t)) \right\}.$$

(iii) By a *variational solution* of (1.10), we mean a collection of operators $\mathcal{V}_t = \mathcal{V}_t^H : \Lambda \to \Lambda, \ t \in [0, \infty)$ with the following properties:

- $\mathcal{V}_0$ is identity, and $\mathcal{V}_t(g+c) = \mathcal{V}_t(g) + c$ for each t and every constant $c \in \mathbb{R}$.
- For $g, g' \in \Lambda$ with $g \le g'$, we have $\mathcal{V}_t(g) \le \mathcal{V}_t(g')$.
- For every $g \in \Lambda$, and $t \in [0, \infty)$,
$$\big\{\big(q, \mathcal{V}_t(g)(q)\big) : \ q \in \mathbb{R}^d\big\} \subseteq \mathcal{F}_t(g).$$

□

When H is independent of q, then

$$\phi_t^H(q,p) = \big(q + t\nabla H(p), p\big), \qquad A_t^H(q,p) = t\big(p \cdot \nabla H(p) - H(p)\big).$$

As a result, $\mathcal{F}_t$ can simply be described as

$$\begin{aligned}\mathcal{F}_t(g) &= \big\{\big(q + t\nabla H(p), g(q) + t(p\cdot\nabla H(p) - H(p))\big) : \ q \in \mathbb{R}^d, \ p \in \partial g(q)\big\} \\ &= \big\{\big(Q, g(q) + p \cdot (Q-q) - tH(p)\big) : \ Q \in \mathbb{R}^d, \ Q - q = t\nabla H(p), \ \ p \in \partial g(q)\big\} \\ &= \big\{\big(Q, A^t(x;Q;g)\big) : \ Q \in \mathbb{R}^d, \ 0 \in \partial_x A^t(x;Q;g)\big\}\,, \end{aligned} \tag{4.1}$$

where $A^t(q,p;Q;g) = A^t(x;Q;g) = g(q) + p\cdot(Q-q) - tH(p)$.

Before examining some examples in dimension one, we define a type of discontinuity of u_q that will be relevant as we compare variational solutions with *viscosity solutions*.

Definition 4.3 Let $H : \mathbb{R} \to \mathbb{R}$ be a continuous function. We say that a pair of momenta (p^-, p^+) satisfies the *Oleinik Condition with respect to H*, if either $p^- > p^+$, and the graph of the restriction of H to $[p^+, p^-]$ is above the chord connecting $(p^-, H(p^-))$ to $(p^+, H(p^+))$, or $p^- < p^+$, and the graph of the restriction of H to $[p^-, p^+]$ is below the chord connecting $(p^-, H(p^-))$ to $(p^+, H(p^+))$. □

Example 4.1 Assume $d = 1$ and the Hamiltonian function H is independent of q, and that the initial condition is given by $g(q) = p^- q 1\!\!1(q \le 0) + p^+ q 1\!\!1(q \ge 0)$. Set

$$\bar{v}(p^-, p^+) := \frac{H(p^+) - H(p^-)}{p^+ - p^-}, \qquad v^\pm := H'(p^\pm).$$

As we will see in this example,

$$u(q,t) = \big(p^- q - tH(p^-)\big) 1\!\!1\big(q \le t\bar{v}\big) + \big(p^+ q - tH(p^+)\big) 1\!\!1\big(q \ge t\bar{v}\big), \tag{4.2}$$

provided that (p^-, p^+) satisfies the Oleinik Condition with respect to H. The solution (4.2) is an example of a *shock wave*. Our expression for the *shock speed* $\bar{v}$ is the celebrated *Rankine-Hugoniot Formula*. On the other hand, if H is concave, then the initial condition g results in solution that is an example of a *rarefaction wave*. The details of our claims follow.

Set $K(p) = pH'(p) - H(p)$. Recall

$$\mathcal{F}_t(g) = \left\{ \big(q + tH'(p), g(q) + tK(p)\big) : \ q \in \mathbb{R}, \ p \in \partial g(q) \right\}.$$

For example, if with $p^- > p^+$, then $\mathcal{F}_t(g) = \mathcal{F}_t^- \cup \mathcal{F}_t^0 \cup \mathcal{F}_t^+$, where

$$\begin{aligned}
\mathcal{F}_t^- &= \big\{ \big(q + tH'(p^-), p^- q + tK(p^-)\big) : \ q \le 0 \big\} = \big\{ \big(q, p^- q - tH(p^-)\big) : \ q \le tv_- \big\}, \\
\mathcal{F}_t^+ &= \big\{ \big(q + tH'(p^+), p^+ q + tK(p^+)\big) : \ q \ge 0 \big\} = \big\{ \big(q, p^+ q - tH(p^+)\big) : \ q \ge tv_+ \big\}, \\
\mathcal{F}_t^0 &= \big\{ \big(tH'(p), tK(p)\big) : \ p \in [p^+, p^-] \big\}.
\end{aligned}$$

Note

$$\mathcal{F}_t^\pm = t\mathcal{F}_1^\pm =: t\mathcal{F}^\pm, \qquad \mathcal{F}_t^0 = t\mathcal{F}_1^0 =: t\mathcal{F}^0.$$

Hence we only need to determine $\mathcal{F}^\cdot = \mathcal{F}_1^\cdot$. To analyze $\mathcal{F}^\cdot$ further, we examine several cases:

(i) Assume that H is strictly convex, or equivalently H' is increasing. We then set $L = K \circ (H')^{-1}$, which is simply the Legendre transform of H. Moreover $v^- > v^+$, and

$$\mathcal{F}^0 = \big\{ \big(v, L(v)\big) : \ v \in [v^+, v^-] \big\}.$$

Note that $\mathcal{F}^\pm$ are lines that intersect at the point $(\bar{v}, \bar{u})$ where $\bar{u} = p^\pm \bar{v} - H(p^\pm)$. Clearly the only continuous function $u(\cdot)$ such that the graph of u is a subset of $\mathcal{F}(g)$ is

$$u(q) = \big(p^- q - H(p^-)\big) \mathbb{1}\big(q \le \bar{v}\big) + \big(p^+ q - H(p^+)\big) \mathbb{1}\big(q \ge \bar{v}\big).$$

This yields the solution $u(q,1) = u(q)$ when $t = 1$. For general t we simply have (4.2). Observe that $g = \min\{g^-, g^+\}$, with $g^\pm(q) = qp^\pm$, and $\mathcal{V}_t(g) = \min\big\{\mathcal{V}_t(g^-), \mathcal{V}_t(g^+)\big\}$. This strong form of monotonicity is true for any pair of initial data $g^\pm$, and is a consequence of the convexity of H.

(ii) If H is strictly concave, then H' is decreasing. As before, we set $L = K \circ (H')^{-1}$, which is now concave . It may be defined by

$$L(v) = \min_{p \in [p^+, p^-]} (vp - H(p)).$$

Moreover, $v^- < v^+$, and

$$\mathcal{F}^0 = \big\{ \big(v, L(v)\big) : \ v \in [v^-, v^+] \big\}.$$

In fact $\mathcal{F}_t(g)$ is the graph of a function $\hat{u}(\cdot, t)$ that is given by

$$u(q,t) = \big(p^- q - tH(p^-)\big) \mathbb{1}\big(q \le tv^-\big) + \big(p^+ q - tH(p^+)\big) \mathbb{1}\big(q \ge tv^+\big) + tL(q) \mathbb{1}\big(tv^- \le q \le tv^+\big).$$

What we have is an example of a rarefaction wave.

(iii) We now relax the convexity assumption of part **(i)** to the Oleinik Condition. More precisely, we assume that the graph of $H : [p^+, p^-] \to \mathbb{R}$ lies below the chord connecting $(p^+, H(p^+))$ to $(p^-, H(p^-))$. We claim that under Oleinik Condition, the only possible u with its graph subset of $\mathcal{F}_1(g) = \mathcal{F}(g)$, is given by (4.2). For this, it suffices to show that no point of $\mathcal{F}^0$ can reach the set below the graph of u. Indeed by Oleinik Condition

$$\frac{H(p) - H(p^+)}{p - p^+} \le \bar{v} = \frac{H(p^+) - H(p^-)}{p^+ - p^-} \le \frac{H(p^-) - H(p)}{p^- - p},$$

for every $p \in [p^+, p^-]$. Hence

$$\begin{aligned}
\bar{v} \le q \quad &\Longrightarrow \quad \frac{H(p) - H(p^+)}{p - p^+} \le q \quad \Longrightarrow \quad p^+q - H(p^+) \le pq - H(p), \\
\bar{v} \ge q \quad &\Longrightarrow \quad \frac{H(p^-) - H(p)}{p^- - p} \ge q \quad \Longrightarrow \quad p^-q - H(p^-) \le pq - H(p).
\end{aligned}$$

As a result, we must have

$$u(q) \le \min_{p \in [p^+, p^-]} (pq - H(p)),$$

for every q. This means that the set $\mathcal{F}^0$ lies above the graph of u. On the other hand, if for some point $(H'(p), pH'(p) - H(p))$ lies on the graph of $\hat{u}$ for some $p \in [p^+, p^-]$, then

$$\text{either} \quad \bar{v} \le q = H'(p) = \frac{H(p) - H(p^+)}{p - p^+} \qquad \text{or} \quad \bar{v} \ge q = H'(p) = \frac{H(p^-) - H(p)}{p^- - p}.$$

By Oleinik Condition, we must have $\bar{v} = q$, which implies that the only possible intersection point between the graph of u and $\mathcal{F}^0$ is the corner point of the graph of u. This completes the proof of our claim.

(iv) Assume that $H(p^+) = H(p^-) = H'(p^-) = 0, H'(p^+) < 0$, and $H(p) < 0$ for every $p \in (p^+, p^-)$. We also assume that there exits $p_0 \in (p^+, p^-)$, such that H is convex in $[p^+, p_0]$, and that H is concave in the interval $[p_0, p^-]$. Clearly the Oleinik Condition is satisfied. We note that $\mathcal{F}^-$ ends at the origin, $\mathcal{F}^+$ passes through the origin, and $\mathcal{F}^0$ has two concave and convex pieces that are tangent to $\mathcal{F}^-$ and $\mathcal{F}^+$ respectively. The shock location is the origin, and $u(q, t) = g(q)$ for all $t \ge 0$. □

As Example 4.1 indicates, we may have a simple formula for the variational solution when H is convex in momentum variable. Note that the action can be expressed in terms of the Lagrangian because when $\dot{x} = J\nabla H(x)$ for $x = (q, p)$, then

$$p \cdot \dot{q} - H(q, p) = L(q, \dot{q}).$$

In fact in this case the variational solution is given by the Lax-Oleinik Formula. (See [Ro1] or [Ro2] for a reference.)

Theorem 4.1 *For a Tonelli Hamiltonian function H, we have*

$$\mathcal{V}_t^H(g)(Q) = \inf\left\{ g(q(0)) + \int_0^t L(q,\dot{q})\ ds :\ q(\cdot) \in C^1[0,t],\ q(t) = Q \right\}. \tag{4.3}$$

In particular if H is convex and independent of q, we may use (4.3) and (4.1) to write

$$\begin{aligned}\mathcal{V}_t^H(g)(Q) &= \inf_q \left(g(q) - tL\left(\frac{Q-q}{t}\right)\right) \\ &= \inf_q \sup_p \big(g(q) + p\cdot(Q-q) - tH(p)\big) = \inf_q \sup_p A^t(q,p;Q;g).\end{aligned} \tag{4.4}$$

This formula is not surprising; after all we are looking for a critical value of $A^t(\cdot;Q;g)$, which is a concave function in p. So it is natural to try a simple minimax critical value that happens to be finite when H is convex.

In fact if we set $t = 1$, then the role of q and p are of the same flavor. Because of this, we may wonder whether or not we have a simple formula for a variational solution when, for example g is concave. This is indeed the case as the following result confirms (See for example [Ro1]).

Theorem 4.2 *Assume that H is independent of q and has a superlinear growth as $|p| \to \infty$, and g is Lipschitz and concave. Then*

$$\mathcal{V}_t^H(g)(Q) = \inf_p \sup_q \big(g(q) + p\cdot(Q-q) - tH(p)\big). \tag{4.5}$$

The identity (4.5) is known as *Hopf's formula* and can be rewritten as

$$\mathcal{V}_t^H(g)(Q) = \inf_p \big(p\cdot Q - g^\dagger(p) - tH(p)\big) = \big(g^\dagger + tH\big)^\dagger(Q), \tag{4.6}$$

where we have used $\dagger$ for the Legendre transform:

$$g^\dagger(p) = \inf_q \big(p\cdot q - g(q)\big).$$

Note that $\big(g + tH\big)^\dagger$ is always well-defined and concave, even when H is not concave. If g is convex instead, then (4.5) and (4.6) change to

$$\mathcal{V}_t^H(g)(Q) = \sup_p \inf_q \big(g(q) + p\cdot(Q-q) - tH(p)\big) = \big(g^* + tH\big)^*(Q), \tag{4.7}$$

where we have used $*$ for the *other* Legendre transform:

$$g^*(p) = \sup_q \big(p\cdot q - g(q)\big).$$

Example 4.2(i) If the graph of H over $[p^+, p^-]$ consists of a collection of concave and convex pieces, then the set $\mathcal{F}^0$ is a union of the graphs of the Legendre transforms of such pieces. However, when $g(q) = \min\{p^-q, p^+q\}$ with $p^+ < p^-$, then g is concave, and the corresponding function u depends only the concave hull of the restriction of H to $[p^+, p^-]$. Indeed from (4.6), and the elementary fact that $g^\dagger(p) = -\infty 1\!\!1\big(p \notin [p^+, p^-]\big)$, we deduce

$$u(q, 1) = u(q) = \min_{p\in[p^+,p^-]} \big(pq - H(p)\big) = \min_{p\in[p^+,p^-]} \big(pq - \hat{H}(p)\big),$$

where $\hat{H}$ denotes the concave hull of the restriction of H to $[p^+, p^-]$. Note that the graph of H is below the chord connecting $\big(p^+, H(p^+)\big)$ to $\big(p^-, H(p^-)\big)$, iff the concave hull of the restriction of H to $[p^+, p^-]$ is this cord. If this is the case, then the Oleinik Condition is satisfied, and we have a shock.The solution is simply given by

$$u(q) = \min_{p\in[p^+,p^-]} \big(pq - H(p)\big) = \min\big\{p^-q - H(p^-), p^+q - H(p^+)\big\},$$

as in (4.2). In general the graph of u can have pieces that lie on $\mathcal{F}^0$. In order to have a feel for how complex u could be, imagine that there are points p_1, p_2, p_3 with $p^+ < p_1 < p_2 < p_3 < p^-$ such that $\hat{H} = H$ in the set $[p_1, p_2] \cup [p_3, p^-]$, and $\hat{H} \neq H$ in its complement. Then the graph of u would have two pieces of $\mathcal{F}^0$ associated with the intervals $[p_1, p_2]$ and $[p_3, p^-]$. More precisely we may express the graph of u as $F_1 \cup F_2 \cup F_3 \cup F_4$, where $F_1 = \mathcal{F}^-$, $F_4 \subset \mathcal{F}^+$, and

$$F_2 = \big\{(H'(p), K(p)) : \ p \in [p_3, p^-]\big\}, \qquad F_3 = \big\{(H'(p), K(p)) : \ p \in [p_1, p_2]\big\},$$

where $K(p) = pH'(p) - H(p)$. The momentum $u' = u_q$ consists of two rarefaction waves associated with F_2 and F_3 that are separated by a shock. The rarefaction F_3 is separated from F_4 by a shock.

(ii) Let us now assume that $p^- < p^+$. Then g is convex and we may apply (4.7) to assert

$$u(Q, 1) = u(Q) = \max_{p\in[p^-,p^+]} \big(pQ - H(p)\big) = \max_{p\in[p^-,p^+]} \big(pQ - \tilde{H}(p)\big),$$

where $\tilde{H}$ denotes the convex hull of H. In particular if the graph of the restriction of H to $[p^-, p^+]$ is above the chord connecting $(p^-, H(p^-))$ to $(p^+, H(p^+))$, then $H(p^\pm) = \tilde{H}(p^\pm)$, and

$$u(q, t) = \max\big\{qp^+ - H(p^+), qp^- - H(p^-)\big\}.$$

In other words, the *Oleinik Condition* is satisfied and we have a shock discontinuity. □

4.2 Viscosity solutions

We start with the definition of upper and lower derivatives:

Definition 4.4(i) Given a function $u : \mathbb{R}^k \to \mathbb{R}$, we write $\bar{\partial} u(z)$ for the set of vectors $a \in \mathbb{R}^k$ such that

$$\limsup_{h \to 0} |h|^{-1}\big(u(z+h) - u(z) - a \cdot h\big) \le 0.$$

Equivalently, $a \in \bar{\partial} u(z)$ iff there exists a C^1 function $\varphi : \mathbb{R}^k \to \mathbb{R}$ such that $\varphi(a) = u(a)$, $\nabla \varphi(z) = a$, and $u \le \varphi$. Similarly, $a \in \underline{\partial} u(z)$ iff

$$\liminf_{h \to 0} |h|^{-1}\big(u(z+h) - u(z) - a \cdot h\big) \ge 0.$$

Equivalently, $a \in \underline{\partial} u(z)$ iff there exists a C^1 function $\varphi : \mathbb{R}^k \to \mathbb{R}$ such that $\varphi(a) = u(a)$, $\nabla \varphi(z) = a$, and $u \ge \varphi$. □

Remark 4.1(i) Assume that $u : \mathbb{R}^k \to \mathbb{R}$ is continuous and there exists a C^1 surface Γ of codimension one such that u is C^1 on $\mathbb{R}^k \setminus \Gamma$. Near Γ, we write $u^\pm$ for the restriction of u on each side of Γ. We assume that $u^\pm$ are C^1 functions up to the boundary points on Γ. Pick a point on Γ. We wish to determine $\bar{\partial} u(a)$ in terms of $\nabla u^\pm(a)$. Assume that $v \in \bar{\partial} u(a) \neq \emptyset$. Let us write $T_a\Gamma$ for the tangent fiber at a to Γ, P_a for the orthogonal projection onto $T_a\Gamma$, and ν_a for the unit normal vector at a that points from $-$-side (on which u^- is defined) to the $+$-side (on which u^+ is defined). First take a smooth path $\gamma : (-\delta, \delta) \to \Gamma$ with $\gamma(0) = a$, $\dot{\gamma}(0) = \tau$. Using $v \in \bar{\partial} u(a)$, and

$$\left(\frac{d}{dt} u \circ \gamma\right)(0) = \nabla u^\pm(a) \cdot \tau,$$

we deduce that $\nabla u^\pm(a) \cdot \tau \le v \cdot \tau$. This also being also true for $-\tau \in T_a\Gamma$ implies that $\nabla u^\pm(a) \cdot \tau = v \cdot \tau$. Hence $\nabla u^+(a) - \nabla u^-(a)$ is orthogonal to $T_a\Gamma$. This is not surprising and follows from the continuity of u; since $u^+ = u^-$ on Γ, the τ-directional derivative of u^+ and u^- coincide whenever $\tau \in T_a\Gamma$. Now if we vary a in the direction of ν_a or $-\nu_a$, we deduce

$$\nabla u^+(a) \cdot \nu_a \le v \cdot \nu_a, \qquad \nabla u^-(a) \cdot (-\nu_a) \le v \cdot (-\nu_a).$$

Equivalently,

$$\nabla u^+(a) \cdot \nu_a \le v \cdot \nu_a \le \nabla u^-(a) \cdot \nu_a.$$

Hence, if $\bar{\partial} u(a) \neq \emptyset$, then $P_a \nabla u^+(a) = P_a \nabla u^-(a)$, $\nabla u^+(a) \cdot \nu_a \le \nabla u^-(a) \cdot \nu_a$, and

$$\bar{\partial} u(a) = \left\{ P_a \nabla u^\pm(a) + r\nu_a : \ r \in \left[\nabla u^+(a) \cdot \nu_a, \nabla u^-(a) \cdot \nu_a\right] \right\}.$$

Likewise, if $\underline{\partial} u(a) \neq \emptyset$, then $P_a \nabla u^+(a) = P_a \nabla u^-(a)$, $\nabla u^+(a) \cdot \nu_a \ge \nabla u^-(a) \cdot \nu_a$, and

$$\underline{\partial} u(a) = \left\{ P_a \nabla u^\pm(a) + r\nu_a : \ r \in \left[\nabla u^-(a) \cdot \nu_a, \nabla u^+(a) \cdot \nu_a\right] \right\}.$$

In summary, we always have $P_a \nabla u^+(a) = P_a \nabla u^-(a)$, and there are three possibilities:

$$\begin{aligned}
\nabla u^+(a) \cdot \nu = \nabla u^-(a) \cdot \nu \quad &\Longrightarrow \quad \bar{\partial} u(a) = \underline{\partial} u(a) = \{\nabla u^{\pm}(a)\}, \\
\nabla u^+(a) \cdot \nu < \nabla u^-(a) \cdot \nu \quad &\Longrightarrow \quad \bar{\partial} u(a) \neq \emptyset, \quad \underline{\partial} u(a) = \emptyset, \\
\nabla u^+(a) \cdot \nu > \nabla u^-(a) \cdot \nu \quad &\Longrightarrow \quad \bar{\partial} u(a) = \emptyset, \quad \underline{\partial} u(a) \neq \emptyset.
\end{aligned}$$

(ii) Let $u : \mathbb{R}^k \to \mathbb{R}$ be a Lipschitz function. Even though the function u is differentiable at almost all points, it is plausible that $\bar{\partial} u(a) \cup \underline{\partial} u(a) = \emptyset$ at some point $a \in \mathbb{R}^k$ (as an example, consider $u(x_1, x_2) = |x_1| - |x_2|$, and $a = (0,0)$). This would not be the case if u is *semi-convex/concave*. First observe that if for example u is convex, then

$$\underline{\partial} u(a) = \left\{ p \in \mathbb{R}^k : \ u(z) - u(a) - p \cdot (z - a) \geq 0 \ \text{ for all } \ z \in \mathbb{R}^k \right\},$$

which is always nonempty. We say a function u is *semi-convex*, if $w(z) = u(z) + \ell |z|^2$ is convex for some $\ell \geq 0$. For such a function

$$\underline{\partial} u(a) = \left\{ p - 2\ell a : \ p \in \underline{\partial} w(a) \right\},$$

which is also nonempty. In fact one can show that for a semi-convex function, we always have

$$\underline{\partial} u(a) = \partial u(a),$$

where ∂u was defined in Definition 4.1**(i)** (see for example Cannarsa-Sinestrari [CS] for a proof).

(iii) We can always approximate any Lipschitz function $u : \mathbb{R}^k \to \mathbb{R}$ by semi-convex/concave functions. For example, given $\delta > 0$, set

$$u^\delta(z) = \sup_y \left(u(y) - \delta^{-1} |z - y|^2 \right).$$

Then one can show that u^δ is always semi-convex, and

$$u(z) \leq u^\delta(z) \leq u(z) + \sup_{r > 0} \left(\ell r - \delta^{-1} r^2 \right) = u(z) + 4^{-1} \ell^2 \delta,$$

where ℓ is the Lipschitz constant of u. On the other hand if the supremum is achieved at $y_\delta(z)$, then for $p = 2\delta^{-1}\big(y_\delta(z) - z\big)$, we have

$$p \in \hat{\partial} u^\delta(z) \subseteq \underline{\partial} u^\delta(z), \qquad p \in \bar{\partial} u\big(y_\delta(z)\big).$$

We must have $|p| \leq \ell$ because for every a with $|a| = 1$, and $y = y_\delta(z)$,

$$-\ell \leq \delta^{-1}\big(u(y + \delta a) - u(y)\big) \leq p \cdot a.$$

In particular, $\big|y_\delta(z) - z\big| = O(\delta)$, which means that near each z, we can find y such that $\bar\partial u(y) \neq \emptyset$. In fact, it is well-known that there exists a one-to-one correspondence between the set of all maximizers $y_\delta(z)$, and the set $\hat\partial u^\delta(z)$ (see for example [CS]). As a result,

$$\cup_z \hat\partial u^\delta(z) \subseteq \cup_y \bar\partial u(y).$$

$\square$

Definition 4.5 We say a uniformly continuous function $u : \mathbb{R}^d \times [0,\infty) \to \mathbb{R}$ is a *viscosity solution* of (1.10) if every $(p,r) \in \bar\partial u(q,t), t > 0$ satisfies $r + H(q,p) \le 0$, and every $(p,r) \in \underline{\partial} u(q,t), t > 0$ satisfies $r + H(q,p) \ge 0$. $\square$

Remark 4.2(i) The theory of viscosity solutions offers a satisfactory notion of solution for the equation (1.10) for two major reasons:

- Under some natural and mild conditions on H, and for a given Lipschitz function $g : \mathbb{R}^d \to \mathbb{R}$, there exists a unique viscosity solution to (1.10) that satisfies the initial condition $g(q) = u(q,0)$. This allows us to define an operator $S_t^H g(q) := u(q,t)$ that enjoys the semigroup property $S_{t+s}^H = S_t^H \circ S_s^H$. See [E] for the proof of uniqueness. Later in Subsection 4.5, we use Game Theory to construct viscosity solutions.

- Many stochastic interfaces in statistical mechanics can be described macroscopically by viscosity solutions of suitable HJEs (see for example [R2] and[R4]).

(ii) When H is convex in the momentum variable, then any semi-concave weak solution is also a viscosity solution. Simply because $\bar\partial u(z)$ is the convex hull of $\hat\partial u(z)$, and the set

$$A(q) := \big\{(p,r) : \ r + H(q,p) \le 0\big\},$$

is convex. $\square$

Exercise Assume that $d = 1$ and u is a (continuous) viscosity solution of (1.10). Let U be an open set in $\mathbb{R} \times (0,\infty)$ and assume that u is C^1 in $U \setminus \Gamma$, where

$$\Gamma = \big\{(a(t),t)) : \ t \in (t_0,t_1)\big\} \subset U,$$

with $a : (t_0,t_1) \to \mathbb{R}$ a C^1 function. Assume that $u = u^+$ and u^-, on the right and left side of Γ in U and both $u^\pm$ solve (1.10) classically. Use Remark 4.1 to show the following:

- $\dot a(t) = H\big[u_q^+(a(t),t), u_q^-(a(t),t)\big]$.

- The pair $\big(u_q^-(a(t),t), u_q^+(a(t),t)\big)$ satisfies the Oleinink Condition for every $t \in (t_1,t_2)$. $\square$

4.3 Viscosity solution versus variational solution

In Example 4.1**(i)**, **(iii)**, **(iv)**, and Example 4.2, we have variational solutions for which u_q has shock discontinuities. In all these examples, the jump discontinuity of u_q satisfies an Oleinik Condition. However it is known that in general Oleinik Condition could be violated for a variational solution. Several explicit examples have been discovered for such a violation. The following recent example is due to V. Roos [Ro1]. This example is constructed by performing a small perturbation to our Example 4.1**(iv)**.

Theorem 4.3 *Assume $d = 1$, $H \in C^2$ is independent of q, and that H'' is uniformly bounded. Assume that $p^+ < p^-$, $H(p^+) = H(p^-) = H'(p^-) = 0 > H''(p^-)$, and $H(p) < 0$ for every $p \in (p^+, p^-)$. Let $f \in C^2$ be a Lipschitz, strictly convex function such that f'' is uniformly bounded, and $f(0) = f'(0) = 0$. Assume that the initial condition g is of the form*

$$g(q) = p^- q 1\!\!1(q \le 0) + \big(p^+ q + f(q)\big) 1\!\!1(q \ge 0).$$

Then there exist $t_0 > 0$ and a continuous function $q : [0, t_0) \to \mathbb{R}$ such that $q(0) = 0$, and for every $t \in [0, t_0)$, there exists a point $q(t) > 0$ such that for every variational solution u, the function $u_q(q, t)$ is discontinuous at $q(t)$. Moreover the Oleinik Condition is violated at $q(t)$.

Proof *(Step 1)* As before, $\mathcal{F}_t(g) = \mathcal{F}_t^+ \cup \mathcal{F}_t^0 \cup \mathcal{F}_t^-$, where

$$\begin{aligned}
\mathcal{F}_t^+ &=: t\mathcal{G}_t = \big\{t\big(q + H'(g'(tq)), t^{-1}g(tq) + K(g'(tq))\big) :\ q \ge 0\big\},\\
\mathcal{F}_t^- &= \mathcal{F}^- = \big\{\big(q, qp^-\big) :\ q \le 0\big\},\\
\mathcal{F}_t^0 &= t\mathcal{F}^0 = \big\{t\big(H'(p), K(p)\big) :\ p \in [p^+, p^-]\big\}.
\end{aligned}$$

Note that the sets $\mathcal{F}^-$ and $\mathcal{F}^0$ are independent of f and coincide with what we had in Example 4.1**(iv)**. Let us write

$$\mathcal{F}^+ = \big\{\big(q, qp^+ - H(p^+)\big) :\ q \ge H'(p^+)\big\} = \big\{\big(q, qp^+\big) :\ q \ge H'(p^+)\big\},$$

which is what we get when $f = 0$ and $t = 1$.

We now examine the set $\mathcal{F}_t^+$. We claim that for $t \in (0, t_0)$, with

$$t_0 = \big[\sup_p |H''(p)|\ \sup_q |f''(q)|\big]^{-1}, \tag{4.8}$$

the set $\mathcal{F}_t^+$ is a graph of a convex function that is above $t\mathcal{F}^+$, and is tangent to $t\mathcal{F}^+$ at its end point. For convexity, observe that if

$$a(q) = q + H'(g'(tq)), \quad b(q) = t^{-1}g(tq) + K(g'(tq)),$$

then $a'(q) = 1 + tH''(g'(tq))g''(tq) = 1 + tH''(g'(tq))f''(tq) > 0$, and

$$b'(q) = g'(tq) + tg'(tq)H''(g'(tq))g''(tq) = g'(tq)a'(tq).$$

Hence the slope of $\mathcal{F}_t^+$ at the point $t(a(q), b(q))$ is $g'(tq)$. Since both a' and g' are increasing, $\mathcal{F}_t^+$ is convex. At $q = 0$ the slope is p^+, which means that the line $t\mathcal{F}^+$ is tangent to $\mathcal{F}_t^+$ at its end point $t(a(0), b(0))$, hence it lies above this line.

(Step 2) For small $\delta > 0$, the set

$$\hat{\mathcal{F}}_t^0 := t\hat{\mathcal{F}}^0 := \big\{t\big(H'(p), K(p)\big) : \ p \in [p^- - \delta, p^-]\big\} \subset \mathcal{F}_t^0,$$

is a graph of concave function that starts from the origin and lies below a line of slope p^- that passes through the origin. We claim that the set $\mathcal{F}_t^+$ will intersect $\hat{\mathcal{F}}_t^0$ at some point $t(a(q^t), b(q^t)), q^t > 0$, for small and positive t. To see this, let us compare the set $\mathcal{G}_t$ with $\hat{\mathcal{F}}^0$. The set $\mathcal{G}_t$ is above $\mathcal{F}^+$ and tangent to $\mathcal{F}^+$ at its end point. Moreover, since

$$g'(tq) = p^+ + f'(tq) = p^+ + o(1), \quad t^{-1}g(tq) = qp^+ + t^{-1}f(tq),$$

we have that $\mathcal{G}_t^+ \to \mathcal{F}^+$ as $t \to 0$. This guarantees that the sets $\mathcal{G}_t$ and $\hat{\mathcal{F}}^0$ intersect at a some point $\big(a(q^t), b(q^t)\big)$ near the origin for small $t > 0$, as desired.

(Step 3) The intersection point of the sets $\mathcal{F}_t^+$ and $\hat{\mathcal{F}}_t^0$ represents a corner of the variational solution $u(q,t)$ at $q = q(t) := ta(q^t)$. The left and right derivatives of $u(\cdot, t)$ at $q(t)$, are given by the slope of $\mathcal{F}_t^0$ and $\mathcal{F}_t^+$ at the point $t(a(q^t), b(q^t))$. The right derivative is given by $\tilde{p}^+ := g'(tq^t)$ as we showed in *Step 1*. To calculate the left derivative, take $\tilde{p}^- \in [p^- - \delta, p^-]$, such that $H'(\tilde{p}^-) = a(q^t)$. We then have

$$b(q^t) = K(\tilde{p}^-) = \tilde{p}^- H'(\tilde{p}^-) - H(\tilde{p}^-),$$

and the tangent vector to $\hat{\mathcal{F}}_t^0$ at $(a(q^t), b(q^t))$ is $\big(H''(\tilde{p}^-), \tilde{p}^- H''(\tilde{p}^-)\big)$, which has a slope $\tilde{p}^-$. It remains to show that the Oleinik Condition is violated for the left and right momenta $\tilde{p}^-$ and $\tilde{p}^+$.

(Final Step) For small t, we have $\tilde{p}^- = p^- + o(1)$, $\tilde{p}^+ = p^+ + o(1)$. So $\tilde{p}^- > \tilde{p}^+$. By $H'(\tilde{p}^-) = a(q^t) = q^t + H'(g'(tq^t))$, we know that $H'(\tilde{p}^+) = H'(\tilde{p}^-) - q^t$. Hence,

$$\begin{aligned}\tilde{p}^- H'(\tilde{p}^-) - H(\tilde{p}^-) &= b(q^t) = t^{-1}g(tq^t) + \tilde{p}^+ H'(\tilde{p}^+) - H(\tilde{p}^+) \\ &= t^{-1}g(tq^t) - \tilde{p}^+ q^t + \tilde{p}^+ H'(\tilde{p}^-) - H(\tilde{p}^+).\end{aligned}$$

Equivalently,

$$\begin{aligned}(\tilde{p}^- - \tilde{p}^+)H'(\tilde{p}^-) + H(\tilde{p}^+) - H(\tilde{p}^-) &= t^{-1}\big(g(tq^t) - g'(tq^t)tq^t\big) \\ &= t^{-1}\big(f(tq^t) - f'(tq^t)tq^t\big) =: t^{-1}\varphi(tq^t).\end{aligned}$$

We note that $\varphi(0) = 0$ and $\varphi'(q) < 0$ for $q > 0$ by convexity of f. As a result,

$$(\tilde{p}^- - \tilde{p}^+)H'(\tilde{p}^-) < H(\tilde{p}^-) - H(\tilde{p}^+). \tag{4.9}$$

This violates the Oleinik Condition because $\tilde{p}^+ < \tilde{p}^-$. □

Since the Oleinik Condition is always satisfied by the pair $\big(\hat{u}_q(q-,t), \hat{u}_q(q+,t)\big)$ at every discontinuity point (q,t) of $\hat{u}_q$, where $\hat{u}$ is a viscosity solution (see the Exercise at the end of Subsection 4.2), we deduce that the variational solution of Theorem 4.3 is not a viscosity solution. In Example 4.3 below, we make some additional assumptions on H, so that we can find a rather precise description for the viscosity solution $\hat{u}$ for small t, with H and g as in Theorem 4.3. This would allow us to show that the jump discontinuity of $\hat{u}_q$ occurs at a point $\hat{q}(t)$ such that $q(t) < \hat{q}(t)$ for small positive t. Moreover, $u(q,t) > \hat{u}(q,t)$ for $q \in (0, \hat{q}(t))$, and small positive t. The details follow.

Example 4.3 Let H and g be as in Theorem 4.3. Additionally, assume that H is concave near p^-, and for some $\delta, \delta_1, \delta_2 > 0$,

$$\big\{p \in [p^+, p^-] : \ H(p) \in [-\delta, 0]\big\} = [p^+, p^+ + \delta_1] \cup [p^- - \delta_2, p^-].$$

Choose $\delta^- \in (0, \delta_2], \delta^+ \in (0, \delta_1]$ such that for each $p \in [p^+, p^+ + \delta^+]$, there exists a unique $\psi(p) \in [p^- - \delta^-, p^-]$ such that $\psi(p^+) = p^-$, and

$$H(p) - H\big(\psi(p)\big) = H'\big(\psi(p)\big)\big(p - \psi(p)\big). \tag{4.10}$$

Let us write $\hat{u}$ for the viscosity solution with the initial condition g. We claim that $\hat{u}(\cdot, t)$ has a corner at some $\hat{q}(t)$ with the following properties: $\hat{q}(0) = 0$, and for small $t > 0$,

$$\hat{q}'(t) = H'\big(\hat{p}_-(t)\big), \quad \hat{p}_-(t) = \psi\big(\hat{p}_+(t)\big), \tag{4.11}$$

where $\hat{p}_\pm(t) = \hat{u}_q\big(\hat{q}(t)\pm, t\big)$ represent the left and right values of $\hat{u}_q$ at $\hat{q}(t)$. We now express $\hat{p}^+(t)$ in terms of $\hat{q}(t)$, so that the ODE (4.11) can be solved uniquely for the initial condition $\hat{q}(0) = 0$. For this, let us write $h : [\hat{p}_+, \infty) \to [0, \infty)$ for $(g')^{-1}$, so that $h(p_+) = 0$. Note if for some q, we have $\hat{q}(t) = q + tH'(g'(q))$, then $\hat{p}_+(t) = g'(q)$. Equivalently,

$$\hat{q}(t) = h(\rho) + tH'(\rho), \quad \hat{p}_+(t) = \rho.$$

Let us write $\ell(q,t)$ for the inverse of $\rho \mapsto h(\rho) + tH'(\rho)$, that is increasing and well-defined for small t. This gives us the formula

$$\hat{p}_+(t) = \ell\big(\hat{q}(t), t\big),$$

which allows us to express $\hat{p}_-(t)$ as a function of $\hat{q}(t)$. The function $\ell(q,t)$ can be expressed as $\ell = w_q$, where w solves the HJE with initial condition $g(q)$, $q \ge 0$, and our formula for ℓ is compatible with (4.7). In particular

$$\ell_t + H'(\ell)\ell_q = 0.$$

We note that $\hat{q}'(0) = 0$ but $\hat{q}'(t) > 0$ for $t > 0$ and small because $H'\big(p_-(t)\big) > 0$. On the other hand,

$$p'_+(t) = \ell_t\big(\hat{q}(t),t\big) + \ell_q\big(\hat{q}(t),t\big)\hat{q}'(t) = \ell_q\big(\hat{q}(t),t\big)\big(H'(\hat{p}_-(t)) - H'(\hat{p}_+(t))\big).$$

Since $\ell_q > 0, H'(\hat{p}_-(t)) > 0, H'(\hat{p}_+(t)) < 0$, we deduce that $\hat{p}_+(t)$ is increasing as a function of t. Since ψ is decreasing, we learn that $\hat{p}_-(t)$ is decreasing. On the other hand,

$$\hat{q}''(t) = H''\big(p^-(t)\big)p'_-(t) > 0,$$

for small t. This means that $\hat{q}$ is convex. This is how the viscosity solution for short times look like:

- For $Q \ge \hat{q}(t)$ we have $\hat{u}(Q,t) = g(h(\rho)) + tK(\rho)$, where $\rho = \ell(Q,t)$.
- For $Q \le 0$, we have $\hat{u}(Q,t) = p^-Q$.
- For $Q \in [0, \hat{q}(t)]$, we first set $Q(s,t) = \hat{q}(s) + (t-s)H'(\hat{p}_-(s))$, for $s \le t$. We note that $Q_s = (t-s)H''(\hat{p}_-(s))\hat{p}'_-(s) > 0$, so that $s \mapsto Q(s,t)$ is increasing with $Q(0,t) = 0, Q(t,t) = \hat{q}(t)$. Its inverse is denoted by $s(Q,t)$, and $\hat{u}(Q,t) = \hat{u}(\hat{q}(s),s) + (t-s)H'(\hat{p}_-(s))$, for $s = s(Q,t)$.

What we have constructed is a viscosity solution because it solves HJE outside the set $\big\{(\hat{q}(t),t) : \ t \in [0,\delta)\big\}$ for small δ, and on this set the Oleinik Condition is satisfied. It also coincides with g initially. So $\hat{u}$ must be the unique viscosity solution.

For comparison, let us write u for the variational solution which has a corner at $q(t)$ with the left and right momenta at $q(t)$ given by $\tilde{p}^{\pm}(t)$ as we discussed in the proof of Theorem 4.3. Indeed by (4.10) and (4.9),

$$\begin{aligned} &H\big(\hat{p}_+(t)\big) - H\big(\hat{p}_-(t)\big) - H'\big(\hat{p}_-(t)\big)\big(\hat{p}_+(t) - \hat{p}_-(t)\big) = 0, \\ &H\big(\tilde{p}^+(t)\big) - H\big(\tilde{p}^-(t)\big) - H'\big(\tilde{p}^-(t)\big)\big(\tilde{p}^+(t) - \tilde{p}^-(t)\big) < 0, \end{aligned}$$

for $t > 0$. In comparison,

$$\begin{aligned} \hat{p}_-(t) &= \psi\big(\hat{p}_+(t)\big), & \tilde{p}^-(t) &> \psi\big(\tilde{p}^+(t)\big), \\ \hat{q}'(t) &= H'\big(\hat{p}_-(t)\big) = H'\big(\psi(\hat{p}_+(t))\big), & q'(t) &= H\big[\tilde{p}^+(t), \tilde{p}^-(t)\big] < H'\big(\psi(\tilde{p}^+(t))\big). \end{aligned}$$

To the left and right of the discontinuity curve, both u_q and $\hat{u}_q$ are classical solutions that can be determined by the method of characteristics. Hence

$$\hat{p}_+(t) = \ell\big(\hat{q}(t), t\big), \qquad \tilde{p}^+(t) = \ell(q(t), t).$$

Hence,

$$\hat{q}'(t) = H'\left(\psi\big(\ell(\hat{q}(t), t)\big)\right), \qquad q'(t) < H'\left(\psi\big(\ell(q(t), t)\big)\right).$$

From this and $q(0) = \hat{q}(0) = 0$, we deduce that $\hat{q}(t) < q(t)$ for small $t > 0$. Note that $u(q,t) = \hat{u}(q,t)$ for $q \notin (0, \hat{q}(t))$. We claim that $\hat{u}(q,t) < u(q,t)$ if $q \in (0, \hat{q}(t))$, and t is small. As a preparation, we t show that if $\rho = u_q$ and $\hat{\rho} = \hat{u}_q$, then $\hat{\rho}(q,t) < \rho(q,t)$ for $q \in (0, \hat{q}(t))$. To verify this, we first consider the case $q \in (q(t), \hat{q}(t))$. For small t, $\rho(q,t) = \rho(q_0, 0) = g'(q_0)$ for some q_0 that is close to 0. Hence $\rho(q,t)$ is close to p^+. However, since such q is on the left side of the jump discontinuity for $\hat{\rho}$, we have $\hat{\rho}(q,t)$ is close to ρ^-, which is strictly larger than ρ^+. This implies that $\hat{\rho}(q,t) < \rho(q,t)$ for small t, and $q \in (q(t), \hat{q}(t))$. In the same fashion we can treat the case $q \in (0, q(t))$.

We are now ready to show that $\hat{u}(q,t) < u(q,t)$ if $q \in (0, \hat{q}(t))$, and t is small. Indeed for $q \in (0, \hat{q}(t))$,

$$\begin{aligned} u(q,t) = u(\hat{q}(t), t) - \int_q^{\hat{q}(t)} \rho(a,t)\ da &= \hat{u}(\hat{q}(t), t) - \int_q^{q(t)} \rho(a,t)\ da \\ > \hat{u}(\hat{q}(t), t) - \int_q^{\hat{q}(t)} \hat{\rho}(a,t)\ da &= \hat{u}(q,t), \end{aligned}$$

as desired. □

As we have seen in the proof of Theorem 4.3, we can easily calculate a solution for small times if the second derivative of the initial data is uniformly bounded.

Proposition 4.1 *Assume that D^2H and D^2g are uniformly bounded, and g is C^1 and Lipschitz. Write u and $\hat{u}$ for variational and viscosity solution with initial condition g. Then there exists $t_0 > 0$ (with t_0 depending only on the uniform bounds on D^2H and D^2g) such that for $t \in [0, t_0]$, we have*

$$u(Q,t) = \hat{u}(Q,t) = g(q(0)) + \int_0^t [p \cdot \dot{q} - H(q,p)]\ ds,$$

where $(q(s), p(s)) = \phi_s\big(q(0), \nabla g(q(0))\big)$ is the unique Hamiltonian orbit such that $q(t) = Q$.

The proof of Proposition 4.1 is rather straightforward and is carried out by showing that the map $a \mapsto q(a,t)$ is a homeomorphism for small t, where $q(a,t)$ is the q-component of $\phi_t\big(a, \nabla g(a)\big)$ (see [B1] for details).

We saw in Example 4.3 that for the initial condition of Theorem 4.3, the variational solution dominates the viscosity solution. This indeed is always true as the following result of Bernard [B2] confirms.

Theorem 4.4 *Assume that D^2H is uniformly bounded and g is Lipschitz. We also assume that $g \in C^2$, and that there exists a constant c_0 such that $D^2g(q) \le c_0 I$ for every $q \in \mathbb{R}^d$ (or more generally, g is semi-concave). Write $\hat{u}$ and u for viscosity and variational solution with initial condition g. Then there exists $t_1 > 0$ (with t_1 depending only on c_0 and the bound on D^2H) such that the following statements are true for $t \in [0, t_1]$:*

- **(i)** $\hat{u}(q,t) \le u(q,t)$.
- **(ii)** $u(q,t) = \inf \big\{ z : \ (q,z) \in \mathcal{F}_t(g) \big\}$.

4.4 Variational selectors

We now give a recipe for the construction of variational solutions in the discrete setting. A similar construction can be give for the continuous setting. We write Λ for the set of Lipschitz functions, and Λ_r for the set of $g \in \Lambda$ such that $|g(q) - g(q')| \le r|q - q'|$. Recall that a variational solution $u_n(Q)$ is a critical value of

$$A(\mathbf{x}_n; Q; g) = g(q_0) + \sum_{i=1}^{n} \big[p_{i-1} \cdot (q_i - q_{i-1}) - w(p_{i-1}, q_i) \big],$$

where $q_n = Q$, and $\mathbf{x}_n = (x_0, \dots, x_{n-1})$, with $x_i = (q_i, p_i) \in \mathbb{R}^{2d}$. We assume that $w : \mathbb{R}^{2d} \to \mathbb{R}$ is a C^1 and Lipschitz function. We may write $A = \ell + f$, where ℓ is a quadratic function and f is a Lipschitz function. Writing $\mathbf{x}_n = x = (q, p) \in \mathbb{R}^k$ for $k = 2nd$, then

$$\ell(x) = \frac{1}{2} Bx \cdot x = \sum_{i=1}^{n-1} p_{i-1} \cdot (q_i - q_{i-1}) - p_{n-1} \cdot q_{n-1},$$

where B is a matrix of the form

$$B = \begin{bmatrix} 0 & D \\ D^t & 0 \end{bmatrix},$$

where D is a matrix which has -1 on its main diagonal, 1 right below the main diagonal, and 0 elsewhere. As a result, ℓ is a non-degenerate quadratic form. Because of the very form of A, we make the following definition.

Definition 4.6(i) We write $\mathcal{Q}_k$ for the set of non-degenerate quadratic functions $\ell : \mathbb{R}^k \to \mathbb{R}$. In other words, $\ell(x) = \frac{1}{2} Bx \cdot x$ for a nonsingular symmetric matrix B. We write $\Omega_k(\ell; r)$ for the set of functions $F : \mathbb{R}^k \to \mathbb{R}$ such that $F = \ell + f$ for some $f \in \Lambda_r$. We write

$$\mathcal{Q} = \cup_{k=1}^{\infty} \mathcal{Q}_k, \qquad \Omega_k = \cup_{r=1}^{\infty} \cup_{\ell \in \mathcal{Q}_k} \Omega_k(\ell; r), \qquad \Omega = \cup_{k=1}^{\infty} \Omega_k.$$

(ii) We call $\mathcal{C} : \Omega \to \mathbb{R}$ a *variational selector* if it satisfies the following conditions:

(1) If $F \in \Omega$ and $F \in C^1$, then $\mathcal{C}(F) = F(\bar{x})$, for some $\bar{x}$ with $\nabla F(\bar{x}) = 0$.

(2) If $f_1, f_2 \in \Lambda$, with $f_1 \le f_2$, and $\ell \in \mathcal{Q}$, then $\mathcal{C}(\ell + f_1) \le \mathcal{C}(\ell + f_2)$.

(3) $\mathcal{C}(F + c) = \mathcal{C}(F) + c$, for every $F \in \Omega$ and $c \in \mathbb{R}$.

(4) If $F \in \Omega$ is bounded below, then $\mathcal{C}(F) = \min F$.

(5) If $\psi : \mathbb{R}^k \to \mathbb{R}^k$ is a Lipschitz smooth diffeomorphism, and $F \in \Omega_k$, then $\mathcal{C}(F) = \mathcal{C}(F \circ \psi)$.

(6) If $F \in \Omega_k$, $\ell' \in \mathcal{Q}_{k'}$, and $F'(x, y) = F(x) + \ell'(y)$, then $\mathcal{C}(F') = \mathcal{C}(F)$.

□

Once a variational selector is known, then we can use it to construct a variational solution by setting

$$\mathcal{V}_n(g)(Q) = \mathcal{C}\big(A(\cdot; Q; g)\big). \tag{4.12}$$

As we mentioned before we use Lusternik-Schnirelmann (LS) Theory to construct a selector (see for example [BC] for more details). Before we give a precise recipe for $\mathcal{C}$, we make some remarks:

Proposition 4.2 **(i)** *If $F \in \Omega_k(\ell; r)$, with $F = \ell + f$, $\ell(x) = 2^{-1}Bx \cdot x$, and $\nabla F(\bar{x}) = 0$, then*

$$|\bar{x}| \le r\delta(\ell)^{-1}, \qquad \textit{where} \qquad \delta(\ell) = \inf_{|x|=1} |Bx|.$$

(ii) *If $\ell + f = \ell' + f'$, for $f, f' \in \Lambda$, $\ell, \ell' \in \mathcal{Q}_k$, then $\ell = \ell'$, and $f = f'$.*

Proof(i) At a critical point $\bar{x}$ we have $B\bar{x} = -\nabla f(\bar{x})$, which implies

$$\delta(\ell)|\bar{x}| \le |B\bar{x}| = \big|\nabla f(\bar{x})\big| \le r,$$

as desired.

(ii) If $\ell + f = \ell' + f'$, then $\ell'' = f''$, where $\ell'' = \ell' - \ell$, $f'' = f - f'$. Since f'' is Lipschitz, then $\ell'' = 0$. In fact if $\ell''(x) = B''x \cdot x$, and v is an eigenvector of B'' associated with eigenvalue λ, then $\varphi(t) = \lambda|v|^2t^2$ must be Lipschitz in t, which is impossible unless $\lambda|v|^2 = 0$. □

LS Theory is normally applied to continuous maps $F : M \to \mathbb{R}$, for a compact manifold M. In our case the non-degeneracy of quadratic function ℓ makes up for the lack of compactness. A standard way to find a critical value of F is by designing a collection $\mathcal{F}$ of subsets of $\mathbb{R}^k$ such that

$$c(F, \mathcal{F}) = \inf_{A \in \mathcal{F}} \sup_A F,$$

is a critical value of F. This is guaranteed if the collection $\mathcal{F}$ satisfies the following property:

$$A \in \mathcal{F},\ t > 0 \quad \Longrightarrow \quad \varphi_t^F(A) \in \mathcal{F},$$

where φ_t^F denotes the flow of the vector field $-\nabla F$. To have a universal collection $\mathcal{F}$ that works for all F, we assume two properties for $\mathcal{F}$:

(1) If $A \in \mathcal{F}$, and φ is a homeomorphism, then $\varphi(A) \in \mathcal{F}$.

(2) If $A \in \mathcal{F}$, and $A \subset B$, then $B \in \mathcal{F}$.

Note that the second property is harmless and can always be assumed because of the infimum over subsets of $A \in \mathcal{F}$ in the definition of c. Its raison d'être is the following alternative expression for $c(F, \mathcal{F})$:

$$c(F, \mathcal{F}) = \inf_{A \in \mathcal{F}} \sup_A F = \inf_{r \in \mathbb{R}} \big\{ r :\ M_r(F) \in \mathcal{F} \big\}, \tag{4.13}$$

where

$$M_r(F) = \big\{ x :\ F(x) < r \big\}.$$

Indeed if we write c and $\bar{c}$ for the left and right-hand sides of the second equality in (4.13), then for any $a > c$, we can find $A \in \mathcal{F}$ such that $\sup_A F < a$, which means that $A \subseteq M_a(F)$. This in turn implies that $M_a(F) \in \mathcal{F}$, which leads to $\bar{c} \le c$. In the same fashion, we can verify $c \le \bar{c}$.

It remains to design a family $\mathcal{F}$ such that **(1)** and **(2)** hold, and $c(F, \mathcal{F})$ is finite. Once such a family is found, we set $\mathcal{C}(F) = c(F, \mathcal{F})$. In view of (4.13), and property **(1)**, we my choose $\mathcal{F}$ the collection of sets with certain degree of topological complexity, so that $c(F, \mathcal{F})$ is the first r for which the sublevel set $M_r(F)$ reaches such complexity. We now describe the LS strategy. Write $\Omega_k^0(\ell, r_0)$ for the set of $F \in \Omega_k(\ell, r_0)$ such that $F(0) = 0$. Let us consider $F \in \Omega_k^0(\ell, r_0)$, and set $c_0 = r_0 \delta(\ell)^{-1}$, $c_1 = r_0 c_0$, so that

$$\nabla F(\bar{x}) = 0 \quad \Longrightarrow \quad |\bar{x}| \le c_0 \quad \Longrightarrow \quad |F(\bar{x})| \le c_1,$$

by Proposition 4.2**(i)**. Note that ℓ has a single critical point at the origin. Hence for $a < 0 < b$, the set $M_b(\ell)$ is topologically more complex than $M_a(\ell)$. Since F is a Lipschitz perturbation of ℓ, and all critical values of F are in the interval $[-c_1, c_1]$, we expect $M_{c_1}(F)$ to be topologically more complex than $M_{-c_1}(F)$. We wish to design a collection $\mathcal{F}$ that captures such complexity. *Relative Cohomology Classes* allow us to measure such complexities.

Definition 4.7 Given two open sets $A \subset B$, we write $\Lambda^j(B, A)$ for the set of closed j forms α in B such that the restriction of α to the set A is exact. We write $\alpha \sim \beta$ for two forms in $\Lambda^j(B, A)$ such that $\beta - \alpha$ is exact in B. We write $H^j(B, A)$ for the set of equivalent classes and $H^*(B, A)$ for the union of $H^j(B, A), j = 0, 1, \dots$. □

For example, for $a < 0 < b$, one can show that $H^*\big(M_b(\ell), M_a(\ell)\big)$ is the same as $H^*(D, \partial D)$, where D is a disc in $\mathbb{R}^{r^-}$, with r^- denoting the number of the negative eigenvalues of B. In fact the set $M_{-c_1}(F)$ is homeomorphic to $M_{-c_1}(\ell)$, and we may define

$$\mathcal{C}(F) = \inf\big\{r : \ H^*\big(M_r(F), M_{-c_1}(F)\big) \neq 0\big\} = \sup\big\{r : \ H^*\big(M_r(F), M_{-c_1}(F)\big) = 0\big\}.$$

Remark 4.3 More generally, we may take any $\alpha \in H^*\big(M_b(\ell), M_a(\ell)\big)$, and set

$$\begin{aligned}\mathcal{C}(F; \alpha) &= \inf\big\{r : \ \text{ the restriction of } \alpha \text{ to } M_r(F) \text{ is not exact}\big\}\\ &= \sup\big\{r : \ \text{ the restriction of } \alpha \text{ to } M_r(F) \text{ is exact}\big\}.\end{aligned}$$

We refer to [BC] for more details. □

4.5 Game theory

We now offer a way of constructing viscosity solutions via Game Theory that in spirit is close to our construction of variational solutions in Subsection 4.4. When $H(q, p)$ is convex in the momentum variable, then the variational solution is also a viscosity solution and (4.3) offers a control theoretical representation of the solution (see [CS] for a thorough discussion on the applications of (4.3)). When H is not convex in the momentum, a minimax type variational description does the job.

For our purposes, it is more convenient to solve the final value problem

$$\left\{\begin{array}{ll} u_t + H(q, u_q) = 0, & t < T, \\ u(q, T) = g(q). & \end{array}\right. \tag{4.14}$$

We assume that H is of the following form

$$H(q, p) = \inf_{z \in Z} \hat{H}(q, p; z) = \inf_{z \in Z} \sup_{v} \big(p \cdot v - \hat{L}(q, v; z)\big),$$

where Z is some measure space, $\hat{H}(q, p; z)$ is convex in p for each $z \in Z$, and we writing $\hat{L}(q, v; z)$ for its Legendre transform in the p-variable. We assume that there exist constants

$\eta_0 > 1, \delta_0 > 0$, and a_0 such that

$$
\begin{aligned}
&\hat{L}(q, v; z) \geq L_0(v) := \delta_0 |v|^{\eta_0} - a_0, \qquad \sup_{|v'| \leq 1} \hat{L}(q, v'; z) \leq a_0, \\
&\lim_{\delta \to 0} \sup_{z' \in Z} \sup_{|x| \leq 1} \sup_{|x - x'| \leq \delta} \left| \hat{H}(x'; z') - \hat{H}(x; z') \right| = 0, \\
&\sup_{z' \in Z} \sup_{q'} \sup_{|p'| \leq \ell} |\hat{H}_p(q', p'; z')| < \infty,
\end{aligned} \tag{4.15}
$$

for all $q, v \in \mathbb{R}^d$, $z \in Z$, and $\ell > 0$.

Definition 4.8 We write $V(t, T)$ for the set of bounded measurable maps $v : [t, T] \to \mathbb{R}^d$, and $Z(t, T)$ for the set of measurable maps $z : [t, T] \to Z$. We write $\Delta(t, T)$ for the set of *strategies*. By a strategy, we mean a map $\alpha : Z(t, T) \to V(t, T)$ such that if $t < s \leq T$, and $z = z'$ on $[t, s]$, then $\alpha[z] = \alpha[z']$ on $[t, s]$. □

We are now ready to offer a solution to (4.14). For $t \leq T$, set

$$
u(q, t) = \mathcal{S}_t^T(g)(q) = \sup_{\alpha \in \Delta(t,T)} \inf_{z \in Z(t,T)} \left[g(q(T)) - \int_t^T \hat{L}\big(q(\theta), \dot{q}(\theta); z(\theta)\big) \, d\theta \right], \tag{4.16}
$$

where $q(\cdot) = q(\cdot; t, q, \alpha[z])$ is uniquely specified by the requirements $q(t) = q$, and $\dot{q} = \alpha[z] =: v$. In other words, for $\theta \in [t, T]$,

$$
q(\theta) = q + \int_t^\theta \alpha[z](\theta') \, d\theta'.
$$

Note that we may write $\dot{q}(\theta) = \hat{H}_p\big(q(\theta), p(\theta); z(\theta)\big)$, where

$$
p(\theta) = \hat{L}_v\big(q(\theta), \alpha[z](\theta); z(\theta)\big).
$$

In terms of $p(\cdot)$, we have

$$
\hat{L}\big(q(\theta), \dot{q}(\theta); z(\theta)\big) = p(\theta) \cdot \dot{q}(\theta) - \hat{H}\big(q(\theta), p(\theta); z(\theta)\big).
$$

When H is not convex in p, the relationship $v = H_p(q, p)$ is no longer invertible in p for a given q. However, if we specify z, then we can invert $p \mapsto \hat{H}_p(q, p; z)$. The role of the path $q(\cdot)$ is the same as the characteristic. The optimal path still solves the Hamiltonian ODE locally, but it is allowed to have corners when we switch from one label z to another.

Theorem 4.5 *The function u as in* (4.16) *is a viscosity solution of* (4.14).

The main ingredient for the proof of Theorem 4.5 is the following *dynamic programming optimality condition*:

Theorem 4.6 *For $s \in [t, T]$, we have*

$$\mathcal{S}_t^T(g)(q) = \sup_{\alpha\in\Delta(t,s)} \inf_{z\in Z(t,s)} \left[\mathcal{S}_s^T(g)(q(s)) - \int_t^s \hat{L}\big(q(\theta), \dot{q}(\theta); z(\theta)\big)\ d\theta\right]. \tag{4.17}$$

Proof Fix q. We write $u(q,t)$ and $u'(q,t)$ for the left and right hand sides of (4.17) respectively. We carry out the proof in two steps.

First we pick $c < u'(q,t)$ and show that $c < u(q,t)$. Observe that since $c < u'(q,t)$, there exists $\beta \in \Delta[t,s]$ such that for all $y \in Z(t,s)$, we have

$$c < \mathcal{S}_s^T(g)(q(s)) - \int_t^s \hat{L}\big(q(\theta), \dot{q}(\theta); y(\theta)\big)\ d\theta,$$

with $q(\theta) = q + \int_t^\theta \beta[y](\theta')\ d\theta'$, for $\theta \in [t,s]$. Now given $a = q(s)$, we can find $\gamma_a \in \Delta(s,T)$ such that for every $w \in Z(s,T)$, we have

$$c < g(\rho(T)) - \int_s^T \hat{L}\big(\rho(\theta), \dot{\rho}(\theta); w(\theta)\big)\ d\theta - \int_t^s \hat{L}\big(q(\theta), \dot{q}(\theta); y(\theta)\big)\ d\theta, \tag{4.18}$$

where

$$\rho(\theta) = q(s) + \int_s^\theta \gamma_{q(s)}[w](\theta')\ d\theta' = q + \int_t^s \beta[y](\theta')\ d\theta' + \int_s^\theta \gamma_{q(s)}[w](\theta')\ d\theta',$$

for $\theta \in [s,T]$. We now construct $\alpha \in \Delta(t,T)$ as follows: Given $z \in Z(t,T)$, we set

$$\hat{\alpha}[z](\theta) = \begin{cases} \beta\big[z \restriction_{[t,s]}\big](\theta), & \theta \in [t,s] \\ \gamma_{\underline{q}(s)}\big[z \restriction_{[s,T]}\big](\theta), & \theta \in (s,T], \end{cases}$$

where $\underline{q}(s) = q + \int_t^s \beta\big[z \restriction_{[t,s]}\big](\theta)\ d\theta$. More generally, we define $\underline{q}(\cdot)$, as

$$\underline{q}(\theta) = q + \int_t^\theta \hat{\alpha}[z](\theta')\ d\theta',$$

for $\theta \in [t,T]$. Observe that (4.18) means

$$c < g(\underline{q}(T)) - \int_t^T \hat{L}\big(\underline{q}(\theta), \dot{\underline{q}}(\theta); z(\theta)\big)\ d\theta \le u(q,t),$$

for every $z \in Z(t,T)$. This completes the proof of $u' \leq u$.

We now turn to the proof of $u \leq u'$. Pick $c < u(q,t)$, and choose $\hat{\alpha} \in \Delta(t,T)$ such that for every $z \in Z(t,T)$

$$\begin{aligned} c &< g(\underline{q}(T)) - \int_t^T \hat{L}\big(\underline{q}(\theta), \dot{\underline{q}}(\theta); z(\theta)\big)\ d\theta \\ &= g(\underline{q}(T)) - \int_s^T \hat{L}\big(\underline{q}(\theta), \dot{\underline{q}}(\theta); z(\theta)\big)\ d\theta - \int_t^s \hat{L}\big(\underline{q}(\theta), \dot{\underline{q}}(\theta); z(\theta)\big)\ d\theta, \end{aligned}$$

where $\underline{q}(\theta) = q + \int_t^\theta \hat{\alpha}[z](\theta')\ d\theta'$, for $\theta \in [t,T]$. We then define $\beta \in \Delta(t,s)$ as follows: for every $y \in Z(t,s)$, we have $\beta[y] = \alpha[y']$, where $y' \in Z(t,T)$, is any extension of y. For this β, we wish to show that for every $y \in Z(t,s)$,

$$c < \mathcal{S}_s^T(g)(q(s)) - \int_t^s \hat{L}\big(q(\theta), \dot{q}(\theta); z(\theta)\big)\ d\theta,$$

where $q(\theta) = q + \int_t^\theta \beta[y](\theta')\ d\theta'$ for $\theta \in [t,s]$. Given $y \in Z(t,s)$, we need to come up with a family of strategies $\gamma_a \in \Delta(s,T)$ such that for every $w \in Z(s,T)$, we have

$$c < g(\rho(T)) - \int_s^T \hat{L}\big(\rho(\theta), \dot{\rho}(\theta); w(\theta)\big)\ d\theta - \int_t^s \hat{L}\big(q(\theta), \dot{q}(\theta); y(\theta)\big)\ d\theta,$$

where

$$\rho(\theta) = q(s) + \int_s^\theta \gamma_{q(s)}[w](\theta')\ d\theta'.$$

This is achieved by setting

$$\gamma_{q(s)}[w] = \alpha[y \oplus w],$$

where

$$(y \oplus w)(\theta) = \begin{cases} y(\theta), & \theta \in [t,s], \\ w(\theta) & \theta \in [s,T]. \end{cases}$$

$\square$

As our next step we show that we can always restrict α in (4.16) to those with bounded range:

Proposition 4.3 *If $g \in \Lambda_r$, then the supremum in* (4.16) *can be restricted to those α such that*

$$M(\alpha) := \sup_{z \in Z(t,T)} M(\alpha, z) := \sup_{z \in Z(t,T)} \left[\frac{1}{T-t} \int_t^T \big|\alpha[z](\theta)\big|^{\eta_0}\ d\theta\right]^{\frac{1}{\eta_0}} \leq C_0, \tag{4.19}$$

where

$$C_0 = C_0(r, \delta_0, \eta_0, a_0) = 2a_0 + \left(\frac{r+1}{\delta_0}\right)^{\frac{1}{\eta_0 - 1}}.$$

Proof Assume that $g \in \Lambda_r$. Write

$$A\big(q; \alpha, z(\cdot)\big) := g(q(T)) - \int_t^T \hat{L}\big(q(\theta), \dot{q}(\theta); z(\theta)\big)\ d\theta,$$

with $q(\cdot)$ as in (4.16). From $g \in \Lambda_r$, and (4.15),

$$\begin{aligned} A\big(q; \alpha, z(\cdot)\big) &\leq g(q) + r\left|\int_t^T \alpha[z]\ d\theta\right| + a_0(T-t) - \delta_0(T-t)M(\alpha, z)^{\eta_0} \\ &\leq g(q) + r(T-t)M(\alpha, z) + a_0(T-t) - \delta_0(T-t)M(\alpha, z)^{\eta_0}. \end{aligned} \tag{4.20}$$

On the other hand,

$$A\big(q; 0, z(\cdot)\big) = g(q) - \int_t^T \hat{L}\big(q, 0; z(\theta)\big)\ d\theta \geq g(q) - a_0(T-t),$$

by (4.15). In (4.16), we may ignore those α such that

$$\inf_{z \in Z(t,T)} A\big(q; \alpha, z(\cdot)\big) < g(q) - a_0(T-t). \tag{4.21}$$

Using (4.20), the inequality (4.21) would be, if that for some $z(\cdot) \in Z(t,T)$, we have

$$r(T-t)M(\alpha, z) + a_0(T-t) - \delta_0(T-t)M(\alpha, z)^{\eta_0} < -a_0(T-t)$$

Equivalently,

$$\delta_0 M(\alpha, z)^{\eta_0} - rM(\alpha, z) - 2a_0 > 0.$$

This inequality is valid if

$$M(\alpha, z) > C_0 := 2a_0 + \left(\frac{r+1}{\delta_0}\right)^{\frac{1}{\eta_0 - 1}}.$$

In summary, we may ignore those α such that

$$\sup_{z \in Z(t,T)} M(\alpha, z) > C_0.$$

We are done. $\square$

With the aid of (4.19), we can show the regularity of of $u = \mathcal{S}_t(g)$.

Theorem 4.7 *Assume that $g \in \Lambda_r$. Then the following statements are true:*
(i) *The value of $u(q,t) = \big(\mathcal{S}_t^T g\big)(q)$ depends only on the restriction of g to the set*

$$B_{C_0(T-t)}(q) := \big\{q' : \ |q'-q| \le C_0(T-t)\big\}.$$

(ii) *The value of $u(q,t) = \big(\mathcal{S}_t^T g\big)(q)$ depends only on the restriction of $\hat{H}$ to the set*

$$B_{C_0(T-t)}(q) \times \mathbb{R}^d \times Z = \big\{(q',p,z) \in \mathbb{R}^{2d} \times Z : \ |q'-q| \le C_0(T-t)\big\}.$$

(iii) *We have*

$$-a_0(T-t) \le u(q,t) - g(q) \le C_1(T-t), \tag{4.22}$$

where $C_1 = C_1(r) = a_0 + c_1 r^{\eta_1}$, for constants $\eta_1 = \eta_0/(\eta_0 - 1)$, and $c_1 = c_1(\delta_0, \eta_0)$.
(iv) *Assume that $s \in [t,T]$. Then*

$$-a_0(s-t) \le u(q,t) - u(q,s) \le C_1(s-t). \tag{4.23}$$

(v) *For every $t < T$, and $q, q' \in \mathbb{R}^d$, we have*

$$|u(q',t) - u(q,t)| \le (C_1 + a_0 + r)|q'-q|. \tag{4.24}$$

Proof(i) The dependence of u on the final data is of the form $g(q(T))$ with

$$|q(T) - q| = \left| \int_t^T \alpha[z] \ d\theta \right| \le C_0(T-t),$$

by(4.19).
(ii) The spatial dependence of $\hat{L}$ is $q(\theta)$ with $\theta \in [t,T]$. We are done because $|q(\theta) - q| \le C_0(T-t)$ by (4.19).
(iii) By choosing the strategy $\alpha = 0$ in the definition of u, and using (4.15) we get

$$u(q,t) \ge g(q) - a_0(T-t).$$

On the other hand, by $g \in \Lambda_r$ and (4.15),

$$\begin{aligned}
u(q,t) &\le g(q) + \sup_{\alpha \in \Delta(t,T)} \ \inf_{z \in Z(t,T)} \left[r|q(T) - q| - \int_t^T L_0\big(\dot{q}(\theta)\big) \ d\theta \right] \\
&\le g(q) + \sup_{\alpha \in \Delta(t,T)} \ \inf_{z \in Z(t,T)} \left[r|q(T) - q| - (T-t) L_0 \left(\frac{q(T) - q(t)}{T-t} \right) \right] \\
&= g(q) + \sup_Q \left[r|Q - q| - (T-t) L_0 \left(\frac{Q-q}{T-t} \right) \right] \\
&= g(q) + (T-t) \sup_{a \ge 0} \big[ra - \delta_0 a^{\eta_0} + a_0 \big] \\
&= g(q) + (T-t)\big[a_0 + c_1 r^{\eta_1}\big],
\end{aligned}$$

as desired.

(iv) Set $\delta = s - t$. From (4.17) and since $\hat{L}$ does not depend on time,

$$u(q,t) = \big(\mathcal{S}^T_{s-\delta} g\big)(q) = \big(\mathcal{S}^{T-\delta}_{s-\delta}\big(\mathcal{S}^T_{T-\delta} g\big)\big)(q) = \big(\mathcal{S}^T_s\big(\mathcal{S}^T_{T-\delta} g\big)\big)(q).$$

From this, $u(q,s) = \mathcal{S}^T_s g(q)$, and the contraction of the operator $\mathcal{S}^T_s$,

$$\inf\big(\mathcal{S}^T_{T-\delta} g - g\big) \leq u(q,t) - u(q,s) \leq \sup\big(\mathcal{S}^T_{T-\delta} g - g\big).$$

This and (4.22) yield (4.23).

(v) First we assume that $|q - q'| \geq T - t$. We then use (4.22) to write

$$\begin{aligned} u(q',t) - u(q,t) &\leq (C_1 + a_0)(T-t) + g(q') - g(q) \\ &\leq (C_1 + a_0)(T-t) + r|q' - q| \\ &\leq (C_1 + a_0 + r)|q' - q|. \end{aligned}$$

Hence

$$|q' - q| \geq T - t \qquad \Longrightarrow \qquad |u(q',t) - u(q,t)| \leq (C_1 + a_0 + r)|q' - q|. \tag{4.25}$$

On the other hand, when $\rho := |q - q'| < T - t$, we use (4.17) and (4.21) to write

$$\begin{aligned} u(q,t) &= \sup_{\alpha \in \Delta(t,t+\rho)} \inf_{z \in Z(t,t+\rho)} \left[u\big(q(t+\rho), t+\rho\big) - \int_t^{t+\rho} \hat{L}\big(q(\theta), \dot{q}(\theta); z(\theta)\big)\, d\theta \right] \\ &\geq \sup_{\alpha \in \Delta(t,t+\rho)} \inf_{z \in Z(t,t+\rho)} \left[u\big(q(t+\rho), t\big) - \int_t^{t+\rho} \hat{L}\big(q(\theta), \dot{q}(\theta); z(\theta)\big)\, d\theta \right] - C_1\rho. \end{aligned}$$

Pick a vector e and choose the constant strategy $\alpha[z] = e$ to assert

$$\begin{aligned} u(q,t) &\geq \inf_{z \in Z(t,t+\rho)} \left[u(q + \rho e, t) - \int_t^{t+\rho} \hat{L}\big(q + \theta e, e; z(\theta)\big)\, d\theta \right] - C_1\rho \\ &\geq u(q + \rho e, t) - (C_1 + a_0)\rho. \end{aligned}$$

We now choose $e = (q' - q)/|q' - q|$ to conclude

$$u(q,t) - u(q',t) \geq -(C_1 + a_0)\rho,$$

which yields

$$|q' - q| \leq T - t \qquad \Longrightarrow \qquad |u(q',t) - u(q,t)| \leq (C_1 + a_0)|q' - q|.$$

This and (4.25) yield (4.24). $\square$

Proof of Theorem 4.5 Fix (q_0, t_0), and assume that $\phi \in C^1$ with

$$u(q_0,t_0) = \phi(q_0,t_0), \quad u \le \phi, \quad p_0 = \phi_q(q_0,t_0), \quad r_0 = \phi_t(q_0,t_0). \tag{4.26}$$

Pick $\delta > 0$, and write $\Delta'(t_0, t_0+\delta)$ for the set of $\alpha \in \Delta(t_0, t_0+\delta)$ such that

$$M(\alpha) := \sup_{z\in Z(t_0,t_0+\delta)} \left[\delta^{-1}\int_{t_0}^{t_0+\delta} \big|\alpha[z](\theta)\big|^{\eta_0}\, d\theta\right]^{\frac{1}{\eta_0}} \le C_0.$$

By Theorem 4.6, and (4.19),

$$u(q_0,t_0) = \sup_{\alpha\in\Delta'(t_0,t_0+\delta)} \inf_{z\in Z(t_0,t_0+\delta)} \left[u\big(q(t_0+\delta), t_0+\delta\big) - \int_{t_0}^{t_0+\delta} \hat L\big(q(\theta),\dot q(\theta); z(\theta)\big)\, d\theta\right],$$

where $q(\theta) = q_0 + \int_{t_0}^{\theta} \alpha[z](\theta)\, d\theta$. To ease the notation, we write Δ'_δ and Z_δ for $\Delta'(t_0, t_0+\delta)$ and $Z(t_0, t_0+\delta)$. From this and our assumption (4.26) we deduce

$$\begin{aligned}
0 &\le \sup_{\alpha\in\Delta'_\delta} \inf_{z\in Z_\delta} \left[\phi\big(q(t_0+\delta), t_0+\delta\big) - \phi(q_0,t_0) - \int_{t_0}^{t_0+\delta} \hat L\big(q(\theta),\dot q(\theta); z(\theta)\big)\, d\theta\right] \\
&= \sup_{\alpha\in\Delta'_\delta} \inf_{z\in Z_\delta} \left[\int_{t_0}^{t_0+\delta} \Big(\phi_t(q(\theta),\theta) + \dot q(\theta)\cdot\phi_q(q(\theta),\theta) - \hat L\big(q(\theta),\dot q(\theta); z(\theta)\big)\Big)\, d\theta\right] \\
&\le \sup_{\alpha\in\Delta'_\delta} \inf_{z\in Z_\delta} \left[\int_{t_0}^{t_0+\delta} \Big(\phi_t(q(\theta),\theta) + \hat H\big(q(\theta),\phi_q(q(\theta),\theta); z(\theta)\big)\Big)\, d\theta\right] \\
&\le \sup_{\alpha\in\Delta'_\delta} \inf_{z\in Z} \left[\int_{t_0}^{t_0+\delta} \Big(\phi_t(q(\theta),\theta) + \hat H\big(q(\theta),\phi_q(q(\theta),\theta); z\big)\Big)\, d\theta\right],
\end{aligned} \tag{4.27}$$

where for the last inequality, we take the infimum over constant paths in $Z(t_0, t_0+\delta)$. On the other hand, since $M(\alpha) \le C_0$, for $\theta \in [t_0, t_0+\delta]$,

$$|q(\theta) - q_0| \le \int_{t_0}^{\theta} \big|\alpha[z](\theta')\big|\, d\theta' \le \int_{t_0}^{t_0+\delta} \big|\alpha[z](\theta)\big|\, d\theta \le \delta M(\alpha) \le C_0\delta, \tag{4.28}$$

where we used the Hölder's inequality for the third inequality.

From this and the continuity of $\hat H$ as in (4.15),

$$\phi_t(q(\theta),\theta) + \hat H\big(q(\theta),\phi_q(q(\theta),\theta); z\big) \le \phi_t(q_0,t_0) + \hat H\big(q_0,\phi_q(q_0,t_0); z\big) + c_1(\delta),$$

for a constant $c_1(\delta)$ such that $c_1(\delta) \to 0$ as $\delta \to 0$. This and (4.27) imply

$$\begin{aligned}
0 &\le \delta \sup_{\alpha\in\Delta'_\delta} \inf_{z\in Z} \left[\phi_t(q_0,t_0) + \hat H\big(q_0,\phi_q(q_0,t_0); z\big) + c_1(\delta)\right] \\
&= \delta \inf_{z\in Z} \left[r_0 + \hat H(q_0,p_0; z) + c_1(\delta)\right] = \delta\left[r_0 + H(q_0,p_0) + c_1(\delta)\right].
\end{aligned}$$

We divide both sides by δ and send $\delta \to 0$ to arrive at $0 \le r_0 + H(q_0, p_0)$, as desired. (Note that since we are solving a backward HJE, this is the correct inequality.)

We next assume that $\phi \in C^1$ is Lipschitz with

$$u(q_0, t_0) = \phi(q_0, t_0), \quad u \ge \phi, \quad p_0 = \phi_q(q_0, t_0), \quad r_0 = \phi_t(q_0, t_0).$$

After a repetition of what we did above, we now have

$$(4.29) \quad 0 \ge \sup_{\alpha \in \Delta'_\delta} \inf_{z \in Z_\delta} \left[\int_{t_0}^{t_0+\delta} \Big(\phi_t(q(\theta), \theta) + \dot{q}(\theta) \cdot \phi_q(q(\theta), \theta) - \hat{L}\big(q(\theta), \dot{q}(\theta); z(\theta)\big) \Big) \, d\theta \right].$$

We now make a selection for α. In principle, we wish to solve the ODE

$$\dot{q}(\theta) = v(q(\theta), \theta; z(\theta)) := \hat{H}_p\big(q(\theta), \phi_q(q(\theta), \theta); z(\theta)\big), \qquad q(t) = q,$$

for a given $z(\cdot) \in Z(t, T)$, and use the solution to define $\alpha[z](\theta) = v(q(\theta), \theta; z(\theta))$. Choosing such a strategy in (4.29) allows us to deduce

$$\begin{aligned} 0 &\ge \inf_{z \in Z_\delta} \left[\int_{t_0}^{t_0+\delta} \Big(\phi_t(q(\theta), \theta) + \hat{H}\big(q(\theta), \phi_q(q(\theta), \theta); z(\theta)\big) \Big) \, d\theta \right] \\ &\ge \left[\int_{t_0}^{t_0+\delta} \big(\phi_t(q(\theta), \theta) + H\big(q(\theta), \phi_q(q(\theta), \theta)\big) \big) \, d\theta \right]. \end{aligned}$$

Again using (4.15) and (4.28) we know

$$\phi_t(q(\theta), \theta) + H\big(q(\theta), \phi_q(q(\theta), \theta)\big) \ge r_0 + H(q_0, p_0) - c_1(\delta),$$

for some constant $c_1(\delta)$ satisfying $c_1(\delta) \to 0$ if $\delta \to 0$. As a result,

$$0 \ge \delta \left[r_0 + H(q_0, p_0) + c_1(\delta) \right].$$

We divide both sides by δ and send $\delta \to 0$ to arrive at $0 \ge r_0 + H(q_0, p_0)$, as desired. $\square$

Remark 4.4 Theorem 4.5 was established by Evans and Souganidis [ES] for more general games. For our presentation we have chosen a game that is more in line with our definition of variational solutions. In fact, [ES] assumes that the analog of the set Z is bounded. Under such an assumption the bound on $M(\alpha)$ becomes trivial and Proposition 4.3 is no longer needed. Though, the results of [ES] are applicable only for bounded Hamiltonian functions. $\square$

5 Homogenization

In Subsection 1.7, we discussed the homogenization phenomenon and its connection to weak KAM theory. In this section we explore the question of homogenization more closely. Several approaches have been developed to establish the homogenization for HJEs and their viscous variants that we now review:

1. The earliest homogenization for HJE was carried out in Lions, Papanicolaou and Varadhan [LPV] when the Hamiltonian function is periodic in position variable. This is achieved by solving (1.6) for w^P, for every $P \in \mathbb{R}^d$. Regarding the graph of a solution to HJE as an evolving interface separating different phases, the graph of $P + \nabla w^P$, is a realization of an invariant measure associated with the inclination P. Fathi [F] extends [LPV] from $\mathbb{T}^d \times \mathbb{R}^d$ to the cotangent bundles of compact manifolds provided that the Hamiltonian function is Tonelli.

2. The homogenization for the variational solutions in the periodic setting (i.e., when $M = \mathbb{T}^d \times \mathbb{R}^d$) has been established by Viterbo [V]. The *homogenized Hamiltonian function* $\bar{H}$ (see (1.12)) that Viterbo obtains for the variational solutions differs from what Lions et al. [LPV] obtains in the viscosity setting. Viterbo uses his homogenized Hamiltonian function to address questions in symplectic geometry.

3. For Tonelli Hamiltonians, Lax-Oleinik formula (4.3) allowed Souganidis [S] and Rezakhanlou-Tarver [RT] to establish the homogenization when the Hamiltonian function is selected according to a shift invariant probability measure. The evolution of the interface (which is the graph of a random *height function*) is a classical example of a stochastic growth model (see for example [R2]). The homogenization in this case (as many other stochastic growth models) can be shown with the aid of the *Subadditive Ergodic Theorem* (see [S] and [RT]).

4. Homogenization for a viscous HJE with $H(x,p) = |p|^2 + V(q)$ for a potential function V is equivalent to the large deviation principle (LDP) for a Brownian motion with killing (see for example Sznitman [Sz]). This suggests using LDP ideas (see for example [R3]) to establish homogenization (see [KRV]).

5. A probability measure on the space Hamiltonians yields a probability measure on the set of semigroup associated with the corresponding HJEs. Homogenization question can be formulated as a dynamical system problem for a group of transformations that are defined on the set of HJ semigroups. This approach was initiated in [R4]. □

For the rest of this section, we explain the approaches **3.** and **4.** for the Frenkel-Kontorova (FK) model of Subsection 3.1 (for the part of our presentation, we follow [M]).

Let us write $\mathcal{L}$ for the set of maps $S : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ such that the map $L(q, v) = S(q, q+v)$ satisfies Assumption 3.1. We equip $\mathcal{L}$ with the topology of L^∞_{loc} that is metrizable. For the question of homogenization, we define an operator that turns a microscopic height function $g : \mathbb{R}^d \to \mathbb{R}$ to a macroscopic height function. Its inverse does the opposite:

$$(\Gamma_n g)(q) = n^{-1} g(nq), \qquad \big(\Gamma_n^{-1} g\big)(q) = \big(\Gamma_{n^{-1}} g\big)(q) = ng\big(n^{-1}q\big).$$

We think of g as an initial macroscopic height function. Its growth is governed microscopically by the operators $\mathcal{T}$ or $\widehat{\mathcal{T}}$ of (3.3). The macroscopic height function after one macroscopic time step (which is n microscopic time steps) is given by $u_n = u_n^S := \widetilde{\mathcal{T}}_n^S g$, where the operator $\widetilde{\mathcal{T}}_n^S$ is defined as

$$\widetilde{\mathcal{T}}_n^S := \Gamma_n \circ \left(\widehat{\mathcal{T}}_S\right)^n \circ \Gamma_n^{-1}.$$

A *homogenization* occurs if the limit

$$\bar{\mathcal{T}}(g) := \lim_{n\to\infty} u_n^S, \tag{5.1}$$

exists for every Lipschitz function g. In the stochastic setting, we wish to establish the homogenization for almost all choices of S with respect to a probability measure that is defined on the set $\mathcal{L}$. This probability measure is assumed to be translation invariant and ergodic with respect to a natural notion of translation that will be defined shortly.

We may write

$$\begin{aligned} u_n^S(q) &= \sup_{q_1,\dots,q_n} \left[g\big(n^{-1}q_n\big) - n^{-1}\big(S(nq,q_1) + S(q_1,q_2) + \cdots + S(q_{n-1},q_n)\big)\right] \\ &= \sup_Q \left[g(Q) - n^{-1} S_n(nq, nQ)\right], \end{aligned} \tag{5.2}$$

where

$$S_n(q,Q) = \inf_{q_1,\dots,q_{n-1}} \big(S(q,q_1) + S(q_1,q_2) + \cdots + S(q_{n-1},Q)\big).$$

To display the dependence of S_n on the generating function S, let us write $S_n(q,Q;S)$ for $S_n(q,Q)$. We also define define the translations (in position variable q) as

$$\tau_a S(q,Q) = S(q+a, Q+a) = L(q+a, Q-q), \qquad \tau_a g(q) = g(q+a).$$

Observe

$$\widetilde{\mathcal{T}}_n^{\tau_a S} = \tau_a \circ \widetilde{\mathcal{T}}_n^S \circ \tau_{-a}, \qquad \text{or} \qquad \tau_a\big(\widetilde{\mathcal{T}}_n^S g\big) = \widetilde{\mathcal{T}}_n^{\tau_a S}(\tau_a g). \tag{5.3}$$

We are now ready to formulate our stochastic homogenization question:

Homogenization Problem: Let $\mathbb{P}$ be a probability measure on the set $\mathcal{L}$ that is invariant with respect to the translation group $\{\tau_a : \ a \in \mathbb{R}^d\}$. Show that the limit (5.1) exists almost surely with respect $\mathbb{P}$. Study the properties of the limit $\bar{\mathcal{T}}$ in terms of the underlying measure $\mathbb{P}$. □

Recall that our probability measure $\mathbb{P}$ is concentrated on the set of $S(q,Q) = L(q, Q-q)$ with L satisfying (3.1). This brings us two useful properties for the sequence u_n:

$$u_n^S(q) = \sup_{|Q-q|\le \ell(r)} \big[g(Q) - n^{-1}S_n(nq, nQ)\big], \tag{5.4}$$

$$\lim_{\delta\to 0} \sup_{S\in\mathcal{L}} \sup_{|q|\le c} \sup_{n} \big|u_n^S(q+\delta) - u_n^S(q;S)\big| = 0, \tag{5.5}$$

for every $c>0$. The proofs of these properties are similar to the proofs of Theorem 4.7**(i)** and **(v)**, and are omitted. From (5.6) we can readily deduce the compactness of the sequence u_n is L^∞_{loc}. For the rest of this section, we describe two strategies that can be employed to prove the existence of a pointwise limit for the sequence u_n^S.

If we set $K_n(Q;S) = S_n(0,Q;S)$, we then have

$$S_n(q,Q;S) = K_n(Q-q;\tau_q S),$$

and the following subadditivity of K_n:

$$K_{m+n}(Q+Q';S) \le K_m(Q;S) + K_n\big(Q';\tau_Q S\big).$$

As a consequence

$$K_{m+n}((m+n)Q;S) \le K_m(mQ;S) + K_n\big(nQ;\tau_{mQ} S\big).$$

This subadditivity can be used to establish the homogenization with the aid of the Subadditive Ergodic Theorem (we refer to [S] and [RT] for more details). More precisely, the Subadditive Ergodic Theorem can guarantee the large n limit of $n^{-1}K_n(nQ;S)$ exists almost surely. The disadvantage of this approach is that it does not offer much information about the limit.

We now turn to approach **4.** This approach is based on the following intuition that we partially discussed in Section 3: If for some C^1 Lipschitz function U, and a constant c, we have $\widehat{\mathcal{T}}(U) = U + c$, then $\Phi\big(q, \nabla U(q)\big) = \big(Q, \nabla U(Q)\big)$, for the corresponding symplectic map Φ. Relationship between q and $Q = F(q)$ is that Q is a critical point of $A(Q;q) = U(Q) - S(q,Q)$. So, $F(q)$ is implicitly given by

$$\nabla U(F(q)) = S_Q(q, F(q)). \tag{5.6}$$

For such a function U, the set $Gr(U)$ is invariant for Φ. Moreover, the q-component of the flow associated with the restriction of Φ to the set $Gr(U)$ can be fully determined in terms of the function $F : \mathbb{R}^d \to \mathbb{R}^d$. In fact in approach **1.**, we show that such solutions U exist. If we can show that for each $P \in \mathbb{R}^d$, there exists a solution $U = U^P$ such that $U(q) = P \cdot q + o(|q|)$, as $|q| \to \infty$, then we are in a position to establish our homogenization as in [LPV]. However, in general a solution U^P may not exist for every P in the stochastic setting. Nonetheless the intuition behind such (equilibrium-like) solutions would allow us to design a strategy that consists of three steps (this should be compared with ideas coming from LDP [R3]).

Step 1 *(Lower Bound)* To simplify our presentation, we first assume that the function $L(q, v) = S(q, q + v)$ is 1-periodic in q. Motivated by (5.6), we pick any continuous function $f : \mathbb{T}^d \to \mathbb{T}^d$, and write $F : \mathbb{R}^d \to \mathbb{R}^d$ for its lift. In particular, we can write $F(q) = q + G(q)$, with G a periodic function. We select $q_i = F^i(q_0)$ with $q_0 = a$ in (5.1). Note

$$n^{-1}q_n = n^{-1}\sum_{i=0}^{n-1} G\big(F^i(q_0)\big), \qquad \sum_{i=0}^{n-1} S(q_i, q_{i+1}) = \sum_{i=0}^{n-1} S^F\big(F^i(q_0)\big),$$

where $S^F(q) = S(q, F(q)) = L(q, G(q)) =: L^G(q)$, which is also periodic. Recall that we only need to study $w_n(S) = u_n(0)$. As a result $u_n(0)$ is close to $u_n\big(n^{-1}a\big)$. We certainly have

$$u_n\big(n^{-1}a\big) \geq g\left(n^{-1}a + n^{-1}\sum_{i=0}^{n-1} G\big(F^i(a)\big)\right) - n^{-1}\sum_{i=0}^{n-1} S^G\big(F^i(a)\big). \tag{5.7}$$

We wish to find the limit of the right-hand side of (5.7). Since both G and S^F are periodic, we may regard them as functions that are defined on the torus; with a slight abuse of notation, we write $G, S^F : \mathbb{T}^d \to \mathbb{R}$, so that we can write $G \circ F^i = G \circ f^i$, and $S^F \circ F^i = S^F \circ f^i$. Now if we pick any ergodic invariant measure for f, then we have

$$\lim_{n\to\infty} n^{-1}\sum_{i=0}^{n-1} G\big(F^i(a)\big) = \int G \, d\mu, \qquad \lim_{n\to\infty} n^{-1}\sum_{i=0}^{n-1} S^F\big(F^i(a)\big) = \int L^G \, d\mu, \tag{5.8}$$

for μ almost all choices of a. From this we obtain

$$\liminf_{n\to\infty} u_n(0) = \liminf_{n\to\infty} u_n\big(n^{-1}a\big) \geq g\left(\int G \, d\mu\right) - \int L^G \, d\mu.$$

This being true for any such pair (F, μ), we deduce

$$\liminf_{n\to\infty} u_n(0) \geq \sup_{(F,\mu)\in\mathcal{M}} \left[g\left(\int G \, d\mu\right) - \int S^F \, d\mu\right] = \sup_v \big[g(v) - \hat{L}(v)\big],$$

where $\mathcal{M}$ is the set of pairs (F, μ) such that μ is an ergodic invariant measure for the corresponding map f, and

$$\hat{L}(v) = \inf_{(F,\mu)\in\mathcal{M}} \left\{ \int S(q, F(q))\ \mu(dq) : \int (F(q) - q)\ \mu(dq) = v \right\}. \tag{5.9}$$

Using (5.2), it is not hard to replace 0 with any $q \in \mathbb{R}^d$ to obtain

$$\liminf_{n\to\infty} u_n(q) \geq \sup_{(F,\mu)\in\mathcal{M}} \left[g\left(\int G\ d\mu \right) - \int S^F\ d\mu \right] = \sup_v \big[g(q+v) - \hat{L}(v) \big]. \tag{5.10}$$

In the stochastic setting, we have a probability measure $\mathbb{P}$ on $\mathcal{L}$ that is τ invariant and ergodic. Here we equip $\mathcal{L}$ with the topology of local uniform convergence and $\mathcal{P}$ is a Radon measure with respect to this topology. We take any bounded continuous function $G : \mathcal{L} \to \mathbb{R}^d$. Out of this, we define a map $F(\cdot; S) : \mathbb{R}^d \to \mathbb{R}^d$, by

$$F(q; S) = q + G\big(\tau_q S\big).$$

We then use the sequence $q_n = F^n(a)$, to obtain a lower bound. Indeed, if we set

$$T = T_G : \mathcal{L} \to \mathcal{L}, \qquad T(S) = \tau_{G(S)} S, \qquad L^G(S) = S\big(0, G(S)\big),$$

then

$$q_n = F^n(a) = \sum_{i=0}^{n-1} G\big(T^i(\tau_a S)\big), \qquad \sum_{i=0}^{n-1} S(q_i, q_{i+1}) = \sum_{i=0}^{n-1} L^G\big(T^i(\tau_a S)\big),$$

To apply the Ergodic Theorem, we pick any T-invariant ergodic measure measure μ so that (5.8) is true. Moreover, if μ is absolutely continuous with respect to $\mathbb{P}$, then we also have (5.10), provided that the supremum is taken over pairs (G, μ) such that μ is T_G ergodic and invariant, and $\mu << \mathbb{P}$.

Step 2 *(Upper Bound)* Let us assume that the initial condition is of the form $g_p(q) = q \cdot p$ for some $p \in \mathbb{R}^d$. Let us write $\mathcal{U}$ for the set of continuous functions $w : \mathbb{R}^d \to \mathbb{R}$, such that

$$\lim_{|q|\to\infty} |q|^{-1} w(q) = 0.$$

We then define

$$\bar{H}(p; w) = \sup_{q,Q} \big(w(Q) - w(q) + p \cdot (Q - q) - S(q, Q) \big).$$

For any $w \in \mathcal{U}$, we use (5.4) to produce an upper bound for the large n limit of u_n as follows:

$$\begin{aligned} u_n(q) = \mathcal{T}_n^S g_p(q) &\leq \sup_{|Q-q|\leq \ell(|p|)} \big[g_p\big(Q\big) - \big(Q - q\big) \cdot p - n^{-1}\big(w(nQ) - w(nq)\big) \big] + H(p; w) \\ &= q \cdot p + H(p; w). \end{aligned}$$

As a result,

$$\limsup_{n\to\infty} u_n(q) \le q\cdot p + \inf_{w\in\mathcal{U}} \bar H(p;w) =: q\cdot p + \bar H(p) = \sup_v \big[g_p(q+v) - \bar L(v)\big],$$

where $\bar L$ is the Legendre transform of $\bar H$.

When H is periodic in q, we have a candidate for what the minimizing $w \in \mathcal{U}$ is namely, the solution $w = w^p$ of the equation (1.6). Writing $\mathcal{U}_0$ for the set of continuous 1-periodic functions, we may write

$$\bar H(p) = \inf_{w\in\mathcal{U}_0} \bar H(p;w).$$

The point is that a more restrictive infimum in the definition of $\bar H$ makes it easier when we try to match our upper bound with our lower bound in **Step 1.** We can also be more selective in the stochastic setting by choosing the type of w that have τ-stationary gradient. For example see [KRV] for more details.

Step 3 ($\hat L = \bar L$) To establish homogenization, it remains to show that the upper and lower limits of **Steps 1** and **2** coincide. This may be achieved by an introduction of a *Lagrange multiplier*, and an application of *Minimax Principle*. We explain this in the periodic case. Also, we simplify our presentation by replacing the set $\mathcal{M}$ with a larger set $\mathcal{M}'$. The set $\mathcal{M}'$ is the set of pairs (F,μ) such that μ is an invariant measure for the corresponding map f (we dropped the ergodicity requirement so that our choice of Lagrange multiplier simplifies). We also set

$$\hat L'(v) = \inf_{(F,\mu)\in\mathcal{M}'} \left\{ \int S(q,F(q))\ \mu(dq) :\ \int (F(q)-q)\ \mu(dq) = v \right\},$$

which is what we get as we replace $\mathcal{M}$ with $\mathcal{M}'$ in (5.9). If we write $\hat H'$ for the Legendre transform of $\hat L'$;

$$\hat H(p) := \sup_v \big(p\cdot v - \hat L(v)\big),$$

then we can show that $\hat H' = \bar H$:

$$\begin{aligned}
\hat H'(p) &= \sup_{(F,\mu)\in\mathcal{M}'} \left(\int \big((F(q)-q)\cdot p - S(q,F(q))\big)\mu(dq) \right) \\
&= \sup_F \sup_\mu \inf_{w\in\mathcal{U}_0} \left(\int \big((F(q)-q)\cdot p - S(q,F(q))\big)\mu(dq) + \int \big(w(F(q)) - w(q)\big)\ \mu(dq) \right) \\
&= \inf_{w\in\mathcal{U}_0} \sup_F \sup_\mu \left(\int \big((F(q)-q)\cdot p - S(q,F(q))\big)\mu(dq) + \int \big(w(F(q)) - w(q)\big)\ \mu(dq) \right) \\
&= \inf_{w\in\mathcal{U}_0} \sup_F \sup_q \big((F(q)-q)\cdot p - S(q,F(q)) + w(F(q)) - w(q)\ \big) \\
&= \inf_{w\in\mathcal{U}_0} \sup_Q \sup_q \big((Q-q)\cdot p - S(q,Q) + w(Q) - w(q)\big) = \bar H(p).
\end{aligned}$$

$\square$

Acknowledgments The author wishes to express his sincere gratitude to Albert Fathi and two anonymous referees whose numerous comments, remarks and suggestions contributed greatly to these notes.

References

[An] N. Anantharaman, On the zero-temperature or vanishing viscosity limit for certain Markov processes arising from Lagrangian dynamics, J. Eur. Math. Soc. **6**, 207-276 (2004).

[Ar] V. I. Arnold. Mathematical Methods of Classical Mechanics. Graduate Texts in Mathematics, 60. Springer-Verlag, New York, 1989.

[B1] P. Bernard, The Lax-Oleinik semi-group: a Hamiltonian point of view. Proc. Roy. Soc. Edinburgh Sect. A, **142**, 1131-1177 (2012).

[B2] P. Bernard, Semi-concave singularities and the Hamilton-Jacobi equation, Regul. Chaotic Dyn., **18**, 674-685 (2013).

[BC] O. Bernardi, and F. Cardin, On C^0-variational solutions for Hamilton-Jacobi equations. Discrete Contin. Dyn. Syst. **31**, 385–406 (2011).

[C1] M. Chaperon, Une idée du type géodésiques brisées pour les systemes hamiltoniens. C. R. Acad. Sci. Paris Sér. I Math., **298**, 293296 (1984).

[C2] M. Chaperon. Familles Génératrices. Cours donné a lécole d'été Erasmus de Samos, 1990.

[C3] M. Chaperon, Lois de conservation et géométrie symplectique. C. R. Acad. Sci. Paris Sér. I Math., **312**, 345-348 (1991).

[CS] P. Cannarsa and C. Sinestrari. Semiconcave Functions, HamiltonJacobi Equations, and Optimal Control. Progress in Nonlinear Differential Equations and Their Applications, **58**, Birkhauser, Boston, 2004.

[E] L. C. Evans. Partial Differential Equations. Graduate studies in mathematics. American Mathematical Society, 2010.

[ES] L. C. Evans and P. E. Souganidis, Differential Games and Representation Formulas for Solutions of Hamilton-Jacobi-Isaacs Equations, Indiana University Mathematics Journal **33**, 773–797 (1984).

[F] A. Fathi. The Weak KAM Theorem in Lagrangian Dynamics. Lecture Notes.

[G] C. Golé. Symplectic Twist Maps: Global Variational Techniques. World Scientific Publishers, 2001.

[H] M. R. Herman, Sur la conjugasion différentiable des difféomorphismes du cercle à des rotations, Publ. Math. Inst. Hautes Etudes Sci., **49**, 5–234 (1979).

[HZ] H. Hofer and E. Zehnder. Symplectic Invariants and Hamiltonian dynamics. Birkhauser Advanced Texts: Birkhauser Verlag, Basel, 1994.

[KH] A. Katok and B. Hasselblatt. Introduction to the Modern Theory of Dynamical Systems. Encyclopedia of Mathematics and Its Applications 54, Cambridge University Press, 1995.

[KRV] E. Kosygina, F. Rezakhanlou, and S.R.S. Varadhan, Stochastic homogenization of Hamilton-Jacobi-Bellman Equations. Communications on Pure and Applied Mathematics, **59**, 1489-1521 (2006).

[L] P. D. Lax. Functional Analysis. Pure and Applied Mathematics (New York). Wiley-Interscience, New York, 2002.

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

[M] I. Matic. Homogenizations and Large Deviations. PhD Thesis, UC Berkeley, 2010.

[MS] D. McDuff and D. Salamon. Introduction to Symplectic Topology. Third edition. Oxford Graduate Texts in Mathematics. Oxford University Press, Oxford, 2017.

[PR1] A. Pelayo and F. Rezakhanlou, The Poincare–Birkhoff Theorem in random dynamics, Trans. Amer. Math. Soc. **370**, 601-639 (2018).

[PR2] A. Pelayo and F. Rezakhanlou, in preparation.

[R1] F. Rezakhanlou. Lectures on Symplectic Geometry.
//https:math.berkeley.edu/rezakhan/symplectic.pdf

[R2] F. Rezakhanlou. Stochastic Growth and KPZ Equation.
//https:math.berkeley.edu/rezakhan/SGKPZ.pdf

[R3] F. Rezakhanlou. Lectures on Large Deviation Principle.
//https:math.berkeley.edu/rezakhan/LD.pdf

[R4] F. Rezakhanlou, Continuum Limit for Some Growth Models II, Ann. Probab. **29**, 1329-1372 (2001).

[RT] F. Rezakhanlou and J. E. Tarver, Homogenization for stochastic Hamilton-Jacobi equations. *Arch. Ration. Mech. Anal.,* **151**, 277–309 (2000).

[Ro1] V. Roos, Variational and viscosity solutions of the Hamilton-Jacobi equation, https://tel.archives-ouvertes.fr/tel-01635263, 2018.

[Ro2] V. Roos, Variational and viscosity operators for the evolutive Hamilton-Jacobi equation, arXiv:1801.09889 (2018).

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

[Sz] A.-S. Sznitman. Brownian Motion, Obstacles and Random Media. Springer Monographs in Mathematics, 1998.

[V] C. Viterbo, Symplectic Homogenization. arXiv:0801.0206v3 (2014).

[W] C. E. Wayne, An introduction to KAM theory. Dynamical systems and probabilistic methods in partial differential equations (Berkeley, CA, 1994), 329, Lectures in Appl. Math., **31**, Amer. Math. Soc., Providence, RI, (1996).