TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 366, Number 2, February 2014, Pages 903–929
S 0002-9947(2013)05936-3
Article electronically published on August 8, 2013

A STOCHASTIC EVANS-ARONSSON PROBLEM

DIOGO GOMES AND HÉCTOR SÁNCHEZ MORGADO

ABSTRACT. In this paper the stochastic version of the Evans-Aronsson problem is studied. Both for the mechanical case and two dimensional problems we prove the existence of smooth solutions. We establish that the corresponding effective Lagrangian and Hamiltonian are smooth. We study the limiting behavior and the convergence of the effective Lagrangian and Hamiltonian, Mather measures and minimizers. We end the paper with applications to stationary mean-field games.

1. INTRODUCTION

Given a Hamiltonian $H : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$, $P \in \mathbb{R}^d$, and $\varepsilon_1, \varepsilon_2 \geq 0$, the stochastic Evans-Aronsson problem consists in determining a solution u_ε of the following variational problem:

$$\bar{H}_\varepsilon(P) = \varepsilon_1 \log \inf_\phi \int e^{\frac{\varepsilon_2 \Delta\phi + H(x, P+D\phi)}{\varepsilon_1}} dx. \tag{1}$$

This problem was first considered in [W] as a generalization of the Evans-Aronsson problem in [E1]; see also [E2]. Unfortunately, some results of that paper depend on certain a priori bounds which are not completely clear. In fact, due to the second-order terms this variational problem is non-coercive and entails several technical difficulties which we believe were not addressed in that paper. We were able, however, to obtain detailed proofs for the special case of quadratic Hamiltonians in the momentum variable, as well as in dimensions 2 and 3.

Our original motivation was to completely understand the argument for the existence of smooth solutions for the minimization problem (1). However, as a result of our work we found an important connection between this class of problems and mean-field games, which is of independent interest and is described in section 9.

For mechanical Hamiltonians as well as in the two and three dimensional cases we establish the existence of smooth minimizers, smoothness of the effective Lagrangian and Hamiltonian functions, generalizing the results in [I-SM]. Furthermore we study the limiting behavior as $\varepsilon \to 0$. For the original Evans-Aronsson problem and $d = 1$ the convergence of solutions and corresponding measures was studied in [GISMY].

Mean-field games is a new class of problems introduced by Lions and Lasry [LL06a], [LL06b], [LL07a] and [LL07b] (see also [LLG10a], [LLG10b]), as well as independently by P. Caines and his co-workers [HMC06], [HMC07], which addresses

Received by the editors November 9, 2011 and, in revised form, June 1, 2012.

2010 *Mathematics Subject Classification.* Primary 49L99.

The first author was partially supported by CAMGSD-LARSys through FCT Program POCTI-FEDER and by grants PTDC/MAT/114397/2009, UTAustin/MAT/0057/2008, and UTA-CMU/MAT/0007/2009.

The second author is grateful to the Instituto Superior Técnico for its hospitality.

the behavior of individual rational agents in a large population. In addition to its interest in the study of Mather measures and viscosity solutions of Hamilton-Jacobi equations, these variational problems are also relevant in the study of stationary mean-field games. In fact, the Euler-Lagrange equation corresponding to (1) can be written as

$$\begin{cases} \varepsilon_2 \Delta u + H(x, P + Du) = \varepsilon_1 \ln m + \bar{H}_\varepsilon(P), \\ \varepsilon_2 \Delta m - \operatorname{div}(D_p H m) = 0, \end{cases}$$

which is an important example of a stationary mean-field game. In particular, the a priori estimates we obtained in this paper correspond immediately to new estimates for this mean-field game. In the case $\varepsilon_2 = 0$ some of the estimates we discuss in this paper were also proved in [E2].

The function $\bar{H}_\varepsilon(P)$ is related with the effective Hamiltonian $\bar{H}(P)$ for first-order Hamilton-Jacobi equations, which is the unique constant for which the equation

$$H(x, P + D\varphi(x)) = \bar{H}(P) \tag{2}$$

admits (periodic) viscosity solutions $\varphi : \mathbb{T}^d \to \mathbb{R}$. The constant $\bar{H}(P)$ is given by the min-max formula

$$\bar{H}(P) = \min_{\psi \in C^1(\mathbb{T}^d)} \max_x H(x, D\psi(x) + P). \tag{3}$$

Formally, as $\varepsilon \to 0$ we should have $\bar{H}_\varepsilon \to \bar{H}$. A related problem, formally corresponding to the case $\varepsilon_1 = 0$, $\varepsilon_2 > 0$, was studied in [G]. In particular

$$\check{H}_{\varepsilon_2}(P) = \min_{\psi \in C^2} \max_x H(x, D\psi(x) + P) + \varepsilon_2 \Delta\psi(x) \tag{4}$$

is the unique number $\check{H}_{\varepsilon_2}(P)$ such that

$$\varepsilon_2 \Delta\varphi(x) + H(x, D\varphi(x) + P) = \check{H}_{\varepsilon_2}(P) \tag{5}$$

admits viscosity (and also C^2) solutions $\varphi_{\varepsilon_2} : \mathbb{T}^d \to \mathbb{R}$. Moreover, the solution of (5) is unique up to addition of constants. This problem was further studied in [I-SM] where the differentiability properties of $\check{H}_{\varepsilon_2}(P)$ (also called the α function) and its Legendre transform $\check{L}_{\varepsilon_2}$ (also called the β function) were studied in detail.

As we will discuss in sections 2 and §3, the stochastic Evans-Aronsson problem is the dual problem of the entropy penalized stochastic Mather problem, which consists in minimizing

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} L(x, v) d\mu + \varepsilon_1 S[\mu],$$

over all measures that satisfy a stochastic holonomy constraint:

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} vD\varphi + \varepsilon_2 \Delta\varphi d\mu(x, v) = 0,$$

for all $\varphi \in C^2(\mathbb{T}^d)$. Here L is the Legendre transform of H and S is a suitable entropy functional.

For $\varepsilon_2 = 0$ the minimization problem (1) is the Evans-Aronsson problem [E1], [E2] of minimizing

$$\int e^{\frac{H(x, P + D\phi)}{\varepsilon_1}} dx \tag{6}$$

over all $\phi \in C^1(\mathbb{T}^d)$. We define

$$\hat{H}_{\varepsilon_1}(P) = \varepsilon_1 \log \inf_{\phi} \int e^{\frac{H(x,P+D\phi)}{\varepsilon_1}} dx.$$

For $d = 1$ the convergence of minimizing φ and measure μ as $\varepsilon_1 \to 0$ was studied in [GISMY]. Minimizing measures for $\varepsilon_1 = \varepsilon_2 = 0$ are called Mather measures.

In this paper we study the problem (1) for two classes of Hamiltonians. The first case is the important mechanical setting, that is,

$$L(x, v) = \frac{|v|^2}{2} + P \cdot v - V(x),$$

to which there corresponds the Hamiltonian

$$H(x, p) = \frac{|P + p|^2}{2} + V(x). \tag{7}$$

For this mechanical setting we have the following main result:

Theorem 1. *Suppose H has the mechanical form (7). Then for each P there exists a unique smooth minimizer u and a unique minimizing measure μ. This measure is supported on the graph $(x, P + Du)$. Furthermore u is the first component of the unique solution of the system*

$$\begin{cases} \varepsilon_2 \Delta u + \frac{1}{2}|P + Du|^2 + V(x) & = \varepsilon_1 \frac{u-v}{2\varepsilon_2}, \\ -\varepsilon_2 \Delta v + \frac{1}{2}|P + Dv|^2 + V(x) & = \varepsilon_1 \frac{u-v}{2\varepsilon_2}. \end{cases} \tag{8}$$

Additionally, if we define

$$m_\mu = C e^{\frac{u-v}{2\varepsilon_2}},$$

where C is determined by the normalization condition $\int_{\mathbb{T}^d} m_\mu(x) dx = 1$, we have

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \psi(x, v) d\mu = \int_{\mathbb{T}^d} \psi(x, P + Du) m_\mu(x) dx.$$

The second case we study concerns general Hamiltonians that satisfy the following hypothesis:

- H is uniformly convex:

$$D^2_{pp} H \geq \gamma > 0; \tag{9}$$

- H has at most quadratic growth, i.e. there is a constant C such that

$$|D^2_{pp} H| \leq C; \tag{10}$$

  thus we have that

$$\frac{\gamma}{2}|p|^2 - C \leq H(p, x) \leq C|p|^2 + C$$

  and

$$|D_p H|^2 \leq CH + C;$$

- the derivatives of H also satisfy

$$|D_x H|^2 \leq C + CH, |D^2_{xx} H| \leq C + CH \quad \text{and} \quad |D^2_{xp} H|^2 \leq C + CH; \tag{11}$$

- if L is the Legendre transform of H

$$L(x,v) = \sup_{v\in\mathbb{R}^d} v\cdot p - H(x,p)$$

we suppose further that

$$L(x, D_pH(x,p)) \geq cH(x,p) - C. \tag{12}$$

Theorem 2. *Suppose $d \leq 3$ and H satisfies* (9)-(12)*. Then* (1) *admits a smooth minimizer.*

We further establish smooth dependence on P:

Theorem 3. *The generalized effective Lagrangian and Hamiltonian functions $\bar{L}_\varepsilon$, $\bar{H}_\varepsilon$ are smooth.*

Finally we obtain the following convergence result:

Theorem 4. *Suppose the Hamiltonian is either mechanical or $d \leq 3$, under assumptions* (9)-(12)*. We suppose all limits as $\varepsilon \to (0,0)$ are taken through any sequence with $\frac{\varepsilon_1}{\varepsilon_2}$ bounded. Then*

(1) *We have*

$$\lim_{\varepsilon\to(0,0)} \bar{H}_\varepsilon(P) = \bar{H}(P).$$

(2) *There is $u \in W^{1,2}$ such that $u_\varepsilon \rightharpoonup u$ in $W^{1,2}$. Moreover u is Lipschitz and*

$$H(x, P + Du) \leq \bar{H}(P).$$

(3) *For mechanical Hamiltonians, if as $\varepsilon \to (0,0)$ we also require $\frac{\varepsilon_1}{\varepsilon_2} \to 0$, then u is a viscosity solution of*

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

(4) *The Mather measure μ_ε converges to a Mather measure μ.*

This paper is structured as follows: we start in section 2 to discuss a possible motivation of the stochastic Evans-Aronsson problem, the entropy penalized stochastic Mather measures. The connection between these two problems is made clear in section 3 where we consider the dual problem. Then in section 4 we consider mechanical Hamiltonians and, using a generalized Hopf-Cole transformation, give an explicit characterization of the minimizing measures. Section 5 is dedicated to several a priori bounds which in particular show the existence of minimizers for general Hamiltonians in dimensions 2 and 3. Then we briefly discuss uniqueness and convexity in section 6 and the smoothness properties in section 7. The convergence as $\varepsilon \to 0$ is addressed in section 8. We end the paper, in section 9, with a discussion of the connection of this problem with a class of mean-field games.

2. Entropy penalized stochastic Mather measures

As mentioned in the introduction, the stochastic Evans-Aronsson problem (1) is related through duality to a certain entropy penalized generalization of the Mather problem. In this section we discuss the set up of this problem. The duality will be considered in the following section.

Given a probability measure μ in $\mathbb{T}^d \times \mathbb{R}^d$ we define its push forward m_μ through the projection to $\mathbb{T}^d$ by

$$\int_{\mathbb{T}^d} \varphi(x)dm_\mu(x) = \int_{\mathbb{T}^d\times\mathbb{R}^d} \varphi(x)d\mu(x,v). \tag{13}$$

In the set $\mathcal{A}$ of probability measures m absolutely continuous with respect to the Lebesgue measure in $\mathbb{T}^d$ (which we identify as usual with $[0,1]^d$) the mapping

$$m \mapsto \int_{\mathbb{T}^d} \log m(x) dm(x) \equiv S^*[m]$$

is convex and lower semicontinuous. This mapping can be extended in a unique way as a convex lower semicontinuous mapping to the set of all probability measures on $\mathbb{T}^d$ by

$$\bar{S}[m] = \liminf_{m_n \in \mathcal{A}, m_n \rightharpoonup m} S^*[m_n].$$

Note that the map $\bar{S}$ is allowed to take the value $+\infty$. Furthermore, since $z \ln z \geq -1/e$, we have $\bar{S} \geq -1/e$. Finally define $S[\mu] = \bar{S}[m_\mu]$.

Let $L(x,v) : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ be a smooth function, bounded from below, strictly convex in v and with superlinear growth, i.e., we assume that there is a function $\gamma : \mathbb{R}^+ \to [1, +\infty)$ such that

$$\frac{L(x,v)}{\gamma(|v|)} \to +\infty, \qquad \frac{\gamma(|v|)}{|v|} \to +\infty,$$

as $|v| \to +\infty$. For $\varepsilon_1 \geq 0$, we consider the convex lower semicontinuous functional

$$A_L(\mu) = \int_{\mathbb{T}^d \times \mathbb{R}^d} L(x,v) d\mu + \varepsilon_1 S[\mu], \tag{14}$$

defined on the space of Borel probability measures μ in $\mathbb{T}^d \times \mathbb{R}^d$. Given $\varepsilon_2 \geq 0$, the entropy penalized stochastic Mather problem consists in minimizing (14) over all Borel probability measures μ which satisfy the following stochastic holonomy constraint: for all $\varphi \in C^2(\mathbb{T}^d)$,

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} (D\varphi(x) v + \varepsilon_2 \Delta\varphi(x)) d\mu = 0. \tag{15}$$

The original Mather problem [Ma], [M] corresponds to $\varepsilon_1 = \varepsilon_2 = 0$. Measures for which (15) holds are called stochastic holonomic (or simply stochastic) for the following reason: consider a controlled Markov diffusion

$$dX = v dt + \sigma dW,$$

where $W(t)$ is a d-dimensional Brownian motion on some probability space $(\Omega, \mathbb{F}, P)$, and $v(t)$ is a bounded progressively measurable (with respect to the filtration generated by the Brownian motion) control taking values in $\mathbb{R}^d$. Define a measure μ_T on $\mathbb{T}^d \times \mathbb{R}^d$ by

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} \phi d\mu_T = \frac{1}{T} E\left[\int_0^T \phi(X(t), v(t)) dt\right],$$

for $\phi \in C_c(\mathbb{T}^d \times \mathbb{R}^d)$. If μ is a weak limit of a sequence μ_{T_n}, $T_n \to \infty$, using Dynkin's formula we have that μ satisfies (15) with $\varepsilon_2 = \frac{\sigma^2}{2}$.

When $\varepsilon_2 = 0$, the non-stochastic case, (15) is Mañe's holonomic condition. The previous fact is the analogue to weak limits of measures supported in liftings of closed curves being holonomic. According to the results of Mather [Ma] and Mañé [M], for the case where $\varepsilon_1 = \varepsilon_2 = 0$, holonomic minimizing measures are invariant under the Euler-Lagrange flow. In this classical setting an important role is played by viscosity solutions of Hamilton-Jacobi equations. Indeed, if μ is a minimizing holonomic measure, then

$$\int_{\mathbb{T}^d\times\mathbb{R}^d} L d\mu = -\bar{H}(0),$$

and a similar identity holds for the stochastic $\varepsilon_2 > 0$, non-entropy penalized ($\varepsilon_1 = 0$) problem [G].

Before we proceed, we need to set up additional notation. Denote by C_0^γ the set of continuous functions on $\mathbb{T}^d \times \mathbb{R}^d$ which satisfy

$$\frac{\phi(x,v)}{\gamma(|v|)} \to 0,$$

as $|v| \to +\infty$. C_0^γ is a Banach space under the norm

$$\|\phi\|_{C_0^\gamma} = \sup_{x,v} \frac{|\phi(x,v)|}{\gamma(|v|)}.$$

Let $\mathcal{M}$ be the space of Borel probability measures on $\mathbb{T}^d \times \mathbb{R}^d$ satisfying

$$\int \gamma(|v|) d\mu < \infty.$$

This set can be identified with the dual of C_0^γ and thus endowed with the weak topology. Set

$$\mathcal{N}_{\varepsilon_2} = cl\left\{\mu \in \mathcal{M} : \int (\varepsilon_2 \Delta\phi(x) + D\phi(x)v) d\mu = 0,\ \forall \phi \in C^2(\mathbb{T}^d)\right\}.$$

Given an element $[\omega] \in H_1(\mathbb{T}^d, \mathbb{R})$ we identify every representative ω of $[\omega]$ with the corresponding vector field. The adjoint δ of the exterior derivative d with respect to the flat metric is simply the divergence operator.

Define $\rho : \mathcal{N}_{\varepsilon_2} \to H_1(\mathbb{T}^d, \mathbb{R})$ by

$$\langle \rho(\mu), [\omega] \rangle = \int (\varepsilon_2 \operatorname{div} \omega(x) + \omega(x) \cdot v) d\mu. \tag{16}$$

It is an immediate consequence of the fact that μ belongs to $\mathcal{N}_\varepsilon$ that the integral in (16) does not depend on the representative of the cohomology class $[\omega]$. Indeed take $\tilde{\omega} = \omega + D\varphi$ for $\varphi \in C^2(\mathbb{T}^d)$. Then

$$\begin{aligned}&\int (\varepsilon_2 \operatorname{div} \omega(x) + \omega(x) \cdot v) d\mu - \int (\varepsilon_2 \operatorname{div} \tilde{\omega}(x) + \tilde{\omega}(x) \cdot v) d\mu \\ &= \int (\varepsilon_2 \Delta\varphi(x) + D\varphi(x)v) d\mu = 0.\end{aligned}$$

Thus, according to Poincaré duality, we can consider $\rho(\mu)$ as a homology class. The map ρ is onto; see Lemma 2.1 in [I-SM].

We now define the generalized effective Lagrangian and Hamiltonian functions for the stochastic version of the Evans-Aronsson problem.

The effective Lagrangian $\bar{L}_\varepsilon : H_1(\mathbb{T}^d, \mathbb{R}) \to \mathbb{R}$ is defined by

$$\bar{L}_\varepsilon(Q) = \inf\{A_L(\mu) : \mu \in \mathcal{N}_{\varepsilon_2}, \rho(\mu) = Q\}. \tag{17}$$

We will prove in Lemma 1 that $\bar{L}_\varepsilon$ is convex and in Corollary 1 that its Legendre transform $\bar{H}_\varepsilon : H^1(\mathbb{T}^d, \mathbb{R}) \to \mathbb{R}$ is given by (1).

3. The dual problem

We now relate the problem of minimizing the functional (14) to the one of minimizing

$$\int e^{\frac{\varepsilon_2\Delta\phi(x)+H(x,D\phi)}{\varepsilon_1}}dx, \tag{18}$$

over all functions $\phi\in C^2(\mathbb{T}^d)$ with $\int\phi(x)dx=0$ (to simplify we have set $P=0$ in (1)).

Proposition 1. *For any $\mu\in\mathcal{N}_{\varepsilon_2}$ and $\phi\in C^2(\mathbb{T}^d)$ we have*

$$A_L(\mu)\geq-\varepsilon_1\log\int_{\mathbb{T}^d}e^{\frac{\varepsilon_2\Delta\phi(x)+H(x,D\phi)}{\varepsilon_1}}dx. \tag{19}$$

Proof. Let $\mu\in\mathcal{N}_{\varepsilon_2}$ and $\phi\in C^2(\mathbb{T}^d)$. Let m_μ be given by (13). We have

$$\begin{aligned}\int L(x,v)d\mu &= \int\left(L(x,v)-D\phi(x)v-\varepsilon_2\Delta\phi(x)\right)d\mu\\ &\geq -\int\left(\varepsilon_2\Delta\phi(x)+H(x,D\phi)\right)dm_\mu,\end{aligned}$$

since, for all $(x,v)\in\mathbb{T}^d\times\mathbb{R}^d$, $L(x,v)-D\phi(x)v\geq-H(D\phi(x),x)$. Define

$$a(m,\phi)=\int(\varepsilon_2\Delta\phi(x)+H(x,D\phi)dm-\varepsilon_1\bar{S}[m]$$

for any Borel probability measure m in $\mathbb{T}^d$. Then

$$A_L(\mu)\geq-a(m_\mu,\phi).$$

Let

$$\lambda_\phi=\varepsilon_1\log\int e^{\frac{\varepsilon_2\Delta\phi(x)+H(x,D\phi)}{\varepsilon_1}}dx$$

and

$$m_\phi(x)=e^{\frac{\varepsilon_2\Delta\phi(x)+H(x,D\phi)-\lambda_\phi}{\varepsilon_1}}. \tag{20}$$

Then $a(m_\phi,\phi)=\lambda_\phi$.

The convex function $t\mapsto t\log t$ has Legendre transform $s\mapsto e^{s-1}$. In particular this implies that $t\log t+1\geq t$, and so, for any probability measure m in $\mathbb{T}^d$ absolutely continuous with respect to Lebesgue measure, we obtain

$$\begin{aligned}&a(m,\phi)\leq a(m_\phi,\phi)\\ &+\int(\varepsilon_2\Delta\phi(x)+H(x,D\phi)-\varepsilon_1\log m_\phi-\varepsilon_1)(m(x)-m_\phi(x))dx.\end{aligned}$$

The convexity and an approximation argument shows that in fact the previous inequality holds for all measures not necessarily absolutely continuous with respect to Lebesgue measure. From the definition of m_ϕ, and since m and m_ϕ are probability measures, the second term on the rhs vanishes and then

$$\lambda_\phi=\sup_m a(m,\phi).$$

Therefore

$$\inf_\mu A_L(\mu)\geq-\varepsilon_1\log\inf_\phi\int e^{\frac{\varepsilon_2\Delta\phi(x)+H(x,D\phi)}{\varepsilon_1}}dx. \tag{21}$$

□

Proposition 2. *If the infimum I_ε on the rhs of* (21) *is achieved at some function $\varphi \in C^4(\mathbb{T}^d)$, then*

$$\varepsilon_2 \Delta m_\varphi - \operatorname{div}(m_\varphi D_p H(x, D\varphi)) = 0. \tag{22}$$

Conversely, if (22) *has a smooth solution φ, then*

$$\mu_\varepsilon(x, v) = \delta(v - D_p H(x, D\varphi(x))) m_\varphi(x) dx \tag{23}$$

is a minimizer of (14).

Proof. Equation (22) is simply the Euler-Lagrange equation for the minimizers in (21).

From definition (23) it follows that for all continuous functions $F : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ we have

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} F(x, v) d\mu_\varepsilon = \int_{\mathbb{T}^d} F(x, D_p H(x, D\varphi(x))) m_\varphi(x) dx.$$

Then, by (22),

$$A_L(\mu_\varepsilon) = -a(m_\varphi, \varphi) = -\varepsilon_1 \log \int e^{\frac{\varepsilon_2 \Delta\varphi(x) + H(x, D\varphi)}{\varepsilon_1}} dx, \tag{24}$$

and so inequality (21) is in fact an equality. □

Corollary 1. *The Legendre transform of $\bar{L}_\varepsilon$ is $\bar{H}_\varepsilon$.*

Proof. Because the Hamiltonian associated to the Lagrangian $L(x, v) - Pv$ is $H(x, P + p)$ and (21) is an equality, we have that

$$\sup_Q PQ - \bar{L}_\varepsilon(Q) = - \inf_{\mu \in \mathcal{N}_{\varepsilon_2}} A_{L - \langle P, \cdot \rangle}(\mu) = \varepsilon_1 \log \inf_\phi \int e^{\frac{\varepsilon_2 \Delta\phi(x) + H(x, P + D\phi)}{\varepsilon_1}} dx.$$

□

4. A Hopf-Cole type transform and existence of minimizers

In this section we address the minimization problem (1) for mechanical Hamiltonians of the form

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

where V is a smooth periodic function. Thanks to the special quadratic structure we are able to use a generalized Hopf-Cole transformation which makes it possible to prove existence of minimizers in any dimension. Similar Hopf-Cole transformations were used, for instance, in [GISMY] and [Ge1], though the existence results in this section, which are based on the a priori bound in Theorem 5 are, as far as we know, new.

We start this section with the following elementary example: set $\varepsilon_1 = \varepsilon_2 = 1$. Suppose $H(p, x) = \frac{1}{2}|p|^2 + V(x)$. Then, equation (22) is simply

$$\Delta m_\varphi - \operatorname{div}(m_\varphi D\varphi) = 0. \tag{25}$$

If we set $m_\varphi = e^\varphi$, we see that it is a solution, provided

$$e^\varphi = e^{\Delta\varphi(x) + H(x, D\varphi)}.$$

Taking logarithms on both sides we see that

$$\varphi = \Delta\varphi(x) + H(x, D\varphi).$$

Of course, in general, m is not a probability measure, but because (25) is homogeneous m can be normalized appropriately, which amounts to adding a constant to φ. This example motivates the generalized Hopf-Cole transformation, which will be discussed in what follows, that makes it possible to address the case where $P \neq 0$.

For $\varepsilon_1, \varepsilon_2 > 0$ we consider the problem of minimizing

$$\int_{\mathbb{T}^d} e^{\frac{\varepsilon_2 \Delta u + \frac{|P+Du|^2}{2} + V(x)}{\varepsilon_1}} dx. \tag{26}$$

Let u be a smooth minimizer of (26) and $m = e^{\frac{\varepsilon_2 \Delta u + H(P+Du,x)}{\varepsilon_1}}$. As discussed in the previous section, m solves the Euler-Lagrange equation

$$\varepsilon_2 \Delta m - \operatorname{div}((P+Du)m) = 0. \tag{27}$$

Proposition 3. *Let u and v be periodic solutions to*

$$\begin{cases} \varepsilon_2 \Delta u + \frac{1}{2}|P+Du|^2 + V(x) & = \varepsilon_1 \frac{u-v}{2\varepsilon_2}, \\ -\varepsilon_2 \Delta v + \frac{1}{2}|P+Dv|^2 + V(x) & = \varepsilon_1 \frac{u-v}{2\varepsilon_2}. \end{cases} \tag{28}$$

Then

$$m = e^{\frac{u-v}{2\varepsilon_2}}$$

solves (27).

Proof. We have

$$\begin{aligned} \varepsilon_2 \Delta m &= m \left[\frac{1}{2} \Delta u - \frac{1}{2} \Delta v + \frac{|Du - Dv|^2}{4\varepsilon_2} \right] \\ &= m \Delta u + \frac{m}{4\varepsilon_2} \left[|P+Du|^2 - |P+Dv|^2 + |Du-Dv|^2 \right] \\ &= (P+Du) \cdot Dm + m\Delta u = \operatorname{div}((P+Du)m). \end{aligned}$$

□

From the previous proposition we obtain the following lower bound:

Corollary 2. *Let u and v be periodic solutions to* (28). *Then the difference $\frac{\varepsilon_1}{\varepsilon_2}(u-v)$ is bounded from below; hence for fixed ε_1, m is also bounded from below.*

Proof. By adding the two equations in (28) we conclude that

$$\varepsilon_2 \Delta(u-v) + \frac{|P+Du|^2}{2} + \frac{|P+Dv|^2}{2} + 2V(x) = \frac{\varepsilon_1}{\varepsilon_2}(u-v).$$

At a point of minimum of $u - v$ we have $\Delta(u-v) \geq 0$ and so at this point we have

$$2V(x) \leq \frac{\varepsilon_1}{\varepsilon_2}(u-v).$$

Hence

$$\inf \frac{\varepsilon_1}{\varepsilon_2}(u-v) \geq 2 \inf V(x).$$

□

We have the following L^1 upper bound:

Corollary 3. *Let u and v be periodic solutions to* (28). *Then $\frac{\varepsilon_1}{2\varepsilon_2} \int u - v$ is bounded from above by $\frac{1}{2}|P|^2 + \sup V(x)$.*

Proof. Because (u, m) solve the necessary conditions of optimality, they are minimizers. Using $u \equiv 0$ in (26) we have

$$\int m \leq \int e^{\frac{|P|^2+2V(x)}{2\varepsilon_1}} dx \leq e^{\frac{|P|^2+2\sup V(x)}{2\varepsilon_1}}.$$

Consequently, from Jensen's inequality we conclude that

$$\int \frac{u-v}{2\varepsilon_2} \leq \frac{|P|^2+2\sup V(x)}{2\varepsilon_1}.$$

□

We now prove a variant of the classical Bernstein estimate for this system which gives

Theorem 5. *Let u and v be periodic solutions to* (28)*. Then*

$$\sup |Du| + |Dv| \leq C + C\frac{\varepsilon_1}{\varepsilon_2}. \tag{29}$$

Proof. Differentiate both equations in (28) with respect to x_i, and multiply, respectively, the first equation by u_{x_i} and the second by v_{x_i}. This yields

$$\begin{cases} \varepsilon_2 u_{x_i}\Delta u_{x_i} + u_{x_i}(P+Du)\cdot Du_{x_i} + V_{x_i}(x)u_{x_i} & = \varepsilon_1 u_{x_i}\frac{u_{x_i}-v_{x_i}}{2\varepsilon_2}, \\ -\varepsilon_2 v_{x_i}\Delta v_{x_i} + v_{x_i}(P+Dv)\cdot Dv_{x_i} + V_{x_i}(x)v_{x_i} & = \varepsilon_1 v_{x_i}\frac{u_{x_i}-v_{x_i}}{2\varepsilon_2}, \end{cases}$$

which by adding over i and rearranging gives the following identities:

$$\begin{cases} \varepsilon_2\Delta\frac{|Du|^2}{2} + (P+Du)\cdot D\frac{|Du|^2}{2} - \varepsilon_2|D^2u|^2 + DV\cdot Du & = \varepsilon_1 Du\frac{D(u-v)}{2\varepsilon_2}, \\ -\varepsilon_2\Delta\frac{|Dv|^2}{2} + (P+Dv)D\frac{|Dv|^2}{2} + \varepsilon_2|D^2v|^2 + DV\cdot Dv & = \varepsilon_1 Dv\frac{D(u-v)}{2\varepsilon_2}. \end{cases}$$

The first equation applied at a point of maximum of $|Du|^2$ yields

$$-\varepsilon_2|D^2u|^2 + DV\cdot Du \geq \varepsilon_1 Du\cdot\frac{D(u-v)}{2\varepsilon_2}.$$

Observing that

$$|\Delta u|^2 \leq d|D^2u|^2,$$

and using (28) we obtain

$$\begin{aligned} \left|\frac{|P+Du|^2}{2} + V(x) - \varepsilon_1\frac{u-v}{2\varepsilon_2}\right|^2 &= \varepsilon_2^2|\Delta u|^2 \\ &\leq d\varepsilon_2 DV\cdot Du - \frac{d\varepsilon_1}{2}Du\cdot D(u-v). \end{aligned} \tag{30}$$

From Corollary 3 $\inf \frac{\varepsilon_1}{\varepsilon_2}(u-v)$ is bounded by above and then

$$\sup\frac{\varepsilon_1}{\varepsilon_2}|u-v| \leq C + \frac{\varepsilon_1}{\varepsilon_2}\sup|D(u-v)|. \tag{31}$$

This, together with Corollary 2 and (30), implies that

$$\sup|Du|^2 \leq C + C(1+\frac{\varepsilon_1}{\varepsilon_2})(\sup|Du| + \sup|Dv|). \tag{32}$$

Similarly, for the second equation we obtain

$$\sup|Dv|^2 \leq C + C(1+\frac{\varepsilon_1}{\varepsilon_2})(\sup|Du| + \sup|Dv|). \tag{33}$$

Adding (32) and (33) yields (29). □

Corollary 4. *Let u and v be periodic solutions of* (28)*. Then all Sobolev norms $W^{m,p}$ are bounded a priori, for $m \geq 0$ and $1 \leq p < \infty$.*

Proof. Because u and v are bounded from below and Lipschitz, they are bounded. This then shows that Δu and Δv are in L^∞ and so $u, v \in W^{2,p}$ for all $1 < p < \infty$. Then $|Du|^2$ and $|Dv|^2$ belong to $W^{1,p}$ for any $1 < p < \infty$. Hence Δu and Δv are in $W^{1,p}$, for all $1 < p < \infty$. So $u, v \in W^{3,p}$ for all $1 \leq p < \infty$. By iteration we get the result for all Sobolev norms. $\square$

We now address existence and uniqueness of minimizers and present the proof of Theorem 1, which is based upon the previous a priori estimates and the continuation method.

Proof. Consider $H_\lambda(x,p) = \frac{1}{2}|p+P|^2 + \lambda V(x)$. For $\lambda = 0$, $u_\varepsilon^\lambda = 0$ is a minimizer. We claim that the set Λ of $\lambda \in [0,1]$ for which there exists a minimizer u_ε^λ satisfying the a priori bounds in Theorem 5 and Corollary 4 is simultaneously open and closed. Then it is clear that we do have a minimizer. The proof that Λ is open is a standard application of the implicit function theorem, since equation (22) is an elliptic (fourth order) equation - this implicit function application will be described in detail in section 7. To prove that Λ is closed, observe that if u_ε^λ is a minimizer, then

$$f_\varepsilon^\lambda = \frac{1}{\varepsilon_1}[\varepsilon_2 \Delta u_\varepsilon^\lambda + \frac{1}{2}|P + Du_\varepsilon^\lambda|^2 + \lambda V(x)] \tag{34}$$

satisfies

$$\varepsilon_2(\Delta f_\varepsilon^\lambda + |Df_\varepsilon^\lambda|^2) = \Delta u_\varepsilon^\lambda + (P + Du_\varepsilon^\lambda)Df_\varepsilon^\lambda. \tag{35}$$

Let $v_\varepsilon^\lambda = u_\varepsilon^\lambda - 2\varepsilon_2 f_\varepsilon^\lambda$; from equations (34), (35)

$$\begin{aligned}
\frac{1}{2}|P + Dv_\varepsilon^\lambda|^2 &= \frac{1}{2}|P + Du_\varepsilon^\lambda|^2 - 2\varepsilon_2(P + Du_\varepsilon^\lambda)Df_\varepsilon^\lambda + 2\varepsilon_2^2|Df_\varepsilon^\lambda|^2 \\
&= \varepsilon_1 f_\varepsilon^\lambda - \lambda V(x) - \varepsilon_2 \Delta u_\varepsilon^\lambda + 2\varepsilon_2(\Delta u_\varepsilon^\lambda - \varepsilon_2 \Delta f_\varepsilon^\lambda) \\
&= \varepsilon_1 f_\varepsilon^\lambda - \lambda V(x) + \varepsilon_2 \Delta v_\varepsilon^\lambda.
\end{aligned}$$

Thus u_ε^λ and v_ε^λ solve (28) with V replaced by λV. By Corollary 4 all Sobolev norms $W^{m,p}$ of u_ε^λ and v_ε^λ are uniformly bounded. So any convergent subsequence of elements of Λ contains a further subsequence whose corresponding sequences $u_\varepsilon^\lambda, v_\varepsilon^\lambda$ converge uniformly, along with all derivatives to a solution of (28) (with V replaced by λV). $\square$

5. A priori bounds for the general case

In this section we discuss additional a priori bounds for the solutions of (22) for non-mechanical Hamiltonians satisfying (9)-(12). This section is divided into two parts. In the first part we discuss a priori bounds that are valid in any space dimension, namely Propositions 4, 5, and 6. In the second part we prove additional estimates, which are necessary to establish existence of smooth minimizers, which are only valid in dimension $d \leq 3$. By the generality of assumptions (9)-(12) we can take $P = 0$ and let $\bar{H}_\varepsilon = \bar{H}_\varepsilon(0)$. Let u be a minimizer of (1). We define the probability measure

$$m = e^{\frac{\varepsilon_2 \Delta u + H(x,Du) - \bar{H}_\varepsilon}{\varepsilon_1}}. \tag{36}$$

We observe that because m is a probability measure, Jensen's inequality implies the following estimate:

$$\int H(x, Du) \leq \bar{H}_\varepsilon, \tag{37}$$

this in particular implies that any minimizer u of (1) is a priori bounded in $W^{1,2}$. The previous inequality, together with setting $\phi = 0$ in (1) yields the following elementary estimate:

$$\min_{x,p} H(x,p) \leq \bar{H}_\varepsilon \leq \max_x H(x,0). \tag{38}$$

Proposition 4. *Let u be a minimizer of* (1) *and m as in* (36)*. We have*

$$\varepsilon_2^2 \int_{\mathbb{T}^d} |D \ln m|^2 \leq C.$$

Proof. We start by multiplying the Euler-Lagrange equation

$$\varepsilon_2 \Delta m - \operatorname{div}(D_p H m) = 0$$

by $\frac{1}{m}$. Then, by integrating by parts and using elementary estimates, we obtain

$$\varepsilon_2 \int_{\mathbb{T}^d} \frac{|Dm|^2}{m^2} \leq \frac{C}{\varepsilon_2} \int_{\mathbb{T}^d} |D_p H|^2 \leq \frac{C}{\varepsilon_2},$$

since $\int_{\mathbb{T}^d} H \leq C$, by (37). □

Corollary 5. *We have*

$$\| \ln m \|_{W^{1,2}} \leq \frac{C}{\varepsilon_1} + \frac{C}{\varepsilon_2}.$$

Proof. Let $f = \ln m$, then

$$\varepsilon_2 \Delta u + H(x, Du) = \varepsilon_1 f + \bar{H}_\varepsilon. \tag{39}$$

Integrating (39) and using $\int m = 1$ and Jensen's inequality, we obtain

$$\min H - \bar{H}_\varepsilon \leq \int H(x, Du) - \bar{H}_\varepsilon \leq \varepsilon_1 \int f \leq 0.$$

By Poincaré inequality we have

$$\int f^2 - \left(\int f \right)^2 = \int \left(f - \int f \right)^2 \leq C \int |Df|^2.$$

Thus, using (38), we have

$$\int f^2 \leq \frac{C}{\varepsilon_1^2} + \frac{C}{\varepsilon_2^2}. \tag{40}$$

□

Proposition 5. *Let u be a minimizer of* (1) *and m as in* (36)*. We have, uniformly in ε,*

$$\int_{\mathbb{T}^d} H m \leq C.$$

Proof. Observe that

$$\varepsilon_2\Delta u + H(x, Du) - D_pH(x, Du)Du + D_pH(x, Du)Du = \varepsilon_1 \ln m + \bar{H}_\varepsilon.$$

Integrating with respect to m yields

$$\int H(x, Du) - D_pH(x, Du)Dudm \geq \bar{H}_\varepsilon,$$

since $\int m \ln m \geq 0$ by Jensen's inequality. Because

$$H(x, p) - D_pH(x, p)p = -L(x, D_pH(x, p)),$$

we have

$$\int L(x, D_pH(x, Du))m \leq C,$$

which then coupled with (12) yields the desired estimate. □

Proposition 6. *Let u be a minimizer of* (1) *and m as in* (36). *Then*

$$\|\sqrt{m}\|_{H^1} \leq \frac{C}{\sqrt{\varepsilon_1}},$$

$$\left(\int m^{\frac{2^*}{2}}\right)^{\frac{2}{2^*}} \leq \frac{C}{\varepsilon_1},$$

$$\int |D^2u|^2 m \leq C,$$

and

$$\int H^2 m \leq C.$$

Proof. Let $f = \ln m$. Then, applying the Laplacian Δ to (39) we get

$$\varepsilon_2\Delta\Delta u + \Delta_x H + 2D^2_{p_k x_i}Hu_{x_k x_i} + D^2_{p_k p_l}Hu_{x_k x_i}u_{x_l x_i} + D_pHD\Delta u = \varepsilon_1\Delta f.$$

Multiplying by m and integrating by parts yields

$$\int \varepsilon_1|Df|^2 m + D^2_{p_k p_l}Hu_{x_k x_i}u_{x_l x_i}m + \Delta_x Hm + 2D^2_{p_k x_i}Hu_{x_k x_i}m = 0.$$

Thus

$$\begin{aligned}\int \varepsilon_1|Df|^2 m + \frac{\gamma}{2}\int |D^2u|^2 m &\leq \int |D^2_{xx}H|m + C\int |D^2_{px}H|^2 m\\ &\leq C + C\int Hm \leq C.\end{aligned}$$

This then implies

$$\int |D^2u|^2 m \leq C.$$

Because

$$\|\sqrt{m}\|^2_{H^1} = 1 + \int |D\sqrt{m}|^2 = 1 + \frac{1}{4}\int |Df|^2 m \leq \frac{C}{\varepsilon_1}$$

we obtain, using Sobolev's theorem,

$$\left(\int m^{2^*/2}\right)^{2/2^*} \leq \frac{C}{\varepsilon_1}.$$

In particular, for any $1 \le p < \infty$, and any $\alpha > 1$,

$$\int |f|^p m \le C + C \int |\ln m|^p m \le C + \left(\int m^{\frac{2^*}{2}} \right)^{\frac{2}{2^*}\alpha}.$$

Hence

$$\int |f|^p m \le \frac{C_{\alpha,p}}{\varepsilon_1^\alpha}.$$

From (39) we have

$$H^2 \le C\varepsilon_1^2 f^2 + C\varepsilon_2^2 |D^2 u|^2,$$

and thus

$$\int H^2 m \le C,$$

uniformly for small ε. This implies in particular

$$\int |Du|^4 m \le C.$$

□

We will now establish further estimates that allow us to prove the existence of smooth solutions for $d \le 3$. To simplify the notation we will set $\epsilon_1 = \epsilon_2 = 1$.

In what follows we will need to consider the equation

$$\Delta w + G(x, Dw) = 0. \tag{41}$$

We suppose that $G : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfies the following standard hypothesis, namely

$$|D_p G(x,p)|^2 \le C|p|^2 + C, \tag{42}$$

$$|p|^2 \le CG(x,p) + \eta(x), \tag{43}$$

and

$$D_p G(x,p)p - G(x,p) \ge c|p|^2 + \zeta(x), \tag{44}$$

for suitable functions $\eta(x), \zeta(x) \in L^2$. We suppose further that we have the following a priori bounds for any solution w of (41), namely $w \in W^{1,2}$ and

$$|D_x G(x, Dw)| \le A(x)|Dw| + B(x); \ \|w\|_{W^{1,2}}, \|A\|_{L^4}, \|B\|_{L^2} \le C. \tag{45}$$

The estimates that we develop next will be applied to

$$G(x,p) = H(x,p) - \ln m - \bar{H},$$

for $w = u$, and also, after establishing suitable regularity for u, to

$$G(p,x) = |p|^2 - D_p H(x, Du)p - \operatorname{div}(D_p H(x, Du)), \tag{46}$$

for $w = \ln m$.

Observe that any solution w to (41) is also a solution to the time dependent equation

$$w_t + \Delta w + G(x, Dw) = 0. \tag{47}$$

Inspired by the adjoint method by Evans (see [Eva10]) we introduce the adjoint equation:

$$\rho_t + \operatorname{div}(D_p G(x, Dw)\rho) = \Delta \rho, \tag{48}$$

with $\rho(x, 0) = \delta_{x_0}$. Note then that for each fixed t, ρ is a probability measure.

Proposition 7. *Let w be a solution of* (41) *and ρ a solution of* (48)*. Then*

$$-w(x_0) = \int_0^T \int_{\mathbb{T}^d} (D_pG(x, Dw)Dw - G(x, Dw))\rho(x,t)dxdt - \int_{\mathbb{T}^d} w(y)\rho(y,T)dx. \tag{49}$$

Proof. Multiply equation (47) by ρ and add equation (48) multiplied by w. This yields

$$w_t\rho + w\rho_t + G\rho + w\,\mathrm{div}(D_pG\rho) = w\Delta\rho - \rho\Delta w.$$

Note that

$$\int_{\mathbb{T}^d} \rho\Delta w - w\Delta\rho dx = 0.$$

Hence

$$\frac{d}{dt}\int_{\mathbb{T}^d} w\rho dx = \int_{\mathbb{T}^d} (D_pG(x, Dw)Dw - G(x, Dw))\rho dx.$$

Then integrating in time we obtain the result. □

Recall that

$$\|\rho\|_{L^1(L^2(dx),dt)} = \int_0^T \|\rho(\cdot,t)\|_{L^2(\mathbb{T}^d)}dt.$$

The previous estimate yields the following corollary:

Corollary 6.

$$\int_0^T \int_{\mathbb{T}^d} |Dw|^2\rho(x,t)dxdt \le C + C\|\rho\|_{L^1(L^2(dx),dt)} + C\,\mathrm{osc}(w),$$

where $\mathrm{osc}(w)$ *denotes the oscillation of* w.

Proof. Observe that $\int_{\mathbb{T}^d} w(y)\rho(y,T) - w(x_0)$ can be bounded by $\mathrm{osc}(w)$ since, for each time t, ρ is a probability measure. Therefore, using (44) in (49) we get

$$\int_0^T \int_{\mathbb{T}^d} |Dw|^2\rho dxdt \le C + C\,\mathrm{osc}(w) + C\int_0^T \int_{\mathbb{T}^d} |\zeta|\rho dxdt.$$

The estimate follows from

$$\int_0^T \int_{\mathbb{T}^d} |\zeta|\rho dxdt \le \|\zeta\|_{L^2}\|\rho\|_{L^1(L^2(dx),dt)}.$$

□

Proposition 8. *For $0 < \alpha < 1$, and any δ there exists C_δ such that*

$$\int_0^T \int_{\mathbb{T}^d} |D(\rho^{\alpha/2})|^2 dxdt \le C_\delta + \delta\int_0^T \int_{\mathbb{T}^d} |Dw|^2\rho dxdt.$$

Proof. Multiply (48) by $\alpha\rho^{\alpha-1}$. Then

$$\frac{d}{dt}\rho^\alpha + \alpha\rho^{\alpha-1}\,\mathrm{div}(D_pG(x, Dw)\rho) = \alpha\rho^{\alpha-1}\Delta\rho. \tag{50}$$

We now integrate the previous identity on $[0,T]\times\mathbb{T}^d$. First observe that because $\rho(\cdot,t)$ is a probability measure we have

$$\int_{\mathbb{T}^d} \rho^\alpha(x,t)dx \le 1. \tag{51}$$

Thus the integral of the first term of the left hand side of (50) is bounded by (51). We also have

$$\begin{aligned}
&\int_0^T \int_{\mathbb{T}^d} \alpha \rho^{\alpha-1} \operatorname{div}(D_pG(x, Dw)\rho) dxd \\
&= \tilde{c}_\alpha \int_0^T \int_{\mathbb{T}^d} \rho^{\alpha/2} \rho^{\alpha/2-1} D\rho D_pG dxdt \\
&\leq \nu \int_0^T \int_{\mathbb{T}^d} |D(\rho^{\alpha/2})|^2 dxdt + C_\nu \int_0^T \int_{\mathbb{T}^d} |D_pG|^2 \rho^\alpha dxdt,
\end{aligned}$$

for any $\nu > 0$. Now we observe that, since $0 < \alpha < 1$,

$$\rho^\alpha \leq C + \delta\rho.$$

Therefore, from (42), we conclude

$$C \int_0^T \int_{\mathbb{T}^d} |D_pG|^2 \rho^\alpha dxdt \leq C + \delta \int_0^T \int_{\mathbb{T}^d} |Dw|^2 \rho dxdt,$$

since we have $\int_{\mathbb{T}^d} |Dw|^2 dx \leq C$.

The corresponding right hand side term of (50) is

$$\alpha(1-\alpha) \int_0^T \int_{\mathbb{T}^d} |D\rho|^2 \rho^{\alpha-2} dxdt = \frac{4(1-\alpha)}{\alpha} \int_0^T \int_{\mathbb{T}^d} |D(\rho^{\alpha/2})|^2 dxdt.$$

Gathering the previous estimates we get

$$\begin{aligned}
\frac{4(1-\alpha)}{\alpha} \int_0^T \int_{\mathbb{T}^d} |D(\rho^{\alpha/2})|^2 dxdt &\leq C + \nu \int_0^T \int_{\mathbb{T}^d} |D(\rho^{\alpha/2})|^2 dxdt \\
&+ \delta \int_0^T \int_{\mathbb{T}^d} |Dw|^2 \rho dxdt.
\end{aligned}$$

Choosing ν small enough we obtain the result. □

Combining Corollary 6 with the previous estimate we conclude:

Corollary 7.

$$\int_0^T \int_{\mathbb{T}^d} |D(\rho^{\alpha/2})|^2 \leq C + C\delta \|\rho\|_{L^1(L^2(dx),dt)} + C\delta \operatorname{osc}(w). \tag{52}$$

Proposition 9. *If $d \leq 3$ there exists $0 < \mu < 1$ such that*

$$\|\rho\|_{L^1(L^2(dx),dt)} \leq C \left(\int_0^T \int_{\mathbb{T}^d} |D(\rho^{\alpha/2})|^2 \right)^\mu + C. \tag{53}$$

Proof. Recall the following interpolation inequality. For $0 < \theta < 1$, $1 \leq p_0, p_1 \leq \infty$, $g \in L^{p_0} \cap L^{p_1}$, and

$$\frac{1}{p_\theta} = \frac{\theta}{p_1} + \frac{1-\theta}{p_0},$$

we have

$$\|g\|_{L^{p_\theta}} \leq \|g\|_{L^{p_1}}^\theta \|g\|_{L^{p_0}}^{1-\theta}. \tag{54}$$

For $p_1 = 1$ and $p_0 = \frac{p\alpha}{2}$ and $p_\theta = 2$ we have

$$\theta = \frac{\alpha p - 4}{2\alpha p - 4}.$$

Fix $0 < t < T$. From Sobolev's theorem, we have (in dimension 2 for any $1 < p < \infty$, and in higher dimensions for $p = 2^*$, where 2^* is the Sobolev exponent defined by $\frac{1}{2^*} = \frac{1}{2} - \frac{1}{d}$) that

$$\left(\int_{\mathbb{T}^d} \rho^{\alpha p/2}(x,t)dx\right)^{1/p} \leq \left(\int_{\mathbb{T}^d} |D(\rho^{\alpha/2}(x,t))|^2 dx\right)^{1/2} + C.$$

Hence

$$\|\rho(\cdot,t)\|_{L^{\alpha p/2}(\mathbb{T}^d)} \leq \left(\int_{\mathbb{T}^d} |D(\rho^{\alpha/2}(x,t))|^2 dx\right)^{1/\alpha} + C.$$

Using $\rho \geq 0$, $\int \rho(x,t)dx = 1$, for all t and (54), we have

$$\|\rho(\cdot,t)\|_{L^2(\mathbb{T}^d)} \leq C\left(\int_{\mathbb{T}^d} |D(\rho^{\alpha/2})|^2 dx\right)^{\mu} + C,$$

for

$$\mu = \frac{1-\theta}{\alpha} = \frac{p}{2\alpha p - 4}.$$

If $p > 4$ we can choose α close enough to 1 such that $\mu < 1$. Because in dimension 2, p can be chosen arbitrarily large and in dimension 3, $p = 6$, the power $\frac{1-\theta}{\alpha}$ can be chosen < 1 in both cases. In dimension 4 and higher this is no longer possible.

Then, by Jensen's inequality,

$$\begin{aligned} \|\rho\|_{L^1(L^2(dx),dt)} &= \int_0^T \|\rho(\cdot,t)\|_{L^2(\mathbb{T}^d)}dt \\ &\leq C\left(\int_0^T\int_{\mathbb{T}^d} |D\rho^{\alpha/2}|^2 dxdt\right)^{\mu} + C. \end{aligned} \tag{55}$$

□

Corollary 8.

$$\int_0^T\int_{\mathbb{T}^d} |D(\rho^{\alpha/2})|^2 \leq C + C\delta\,\mathrm{osc}(w). \tag{56}$$

Proof. The inequality follows by using (53) in (52). □

Corollary 9.

$$\|\rho\|_{L^1(L^2(dx),dt)} \leq C + C(\mathrm{osc}(w))^{\mu}. \tag{57}$$

Proof. The inequality follows by using (56) in (53). □

Corollary 10.

$$\int_0^T\int_{\mathbb{T}^d} |Dw|^2\rho(x,t)dxdt \leq C + C\,\mathrm{osc}(w). \tag{58}$$

Proof. Using estimate (57) in the estimate of Corollary 6 as well as the elementary inequality $C + C(\mathrm{osc}(u))^{\mu} \leq C + C\,\mathrm{osc}(u)$, we obtain the desired result. □

Theorem 6. *Let $G : \mathbb{T}^d \times \mathbb{R}^d \to \mathbb{R}$ satisfy* (42), (43) *and* (44). *Suppose w is a solution of* (41) *satisfying* (45). *Assume $d \leq 3$. Then $w \in W^{1,\infty}$.*

Proof. Differentiate equation (47) with respect to a direction determined by a unit vector $\xi \in \mathbb{R}^d$. Then

$$(w_\xi)_t + D_pGD_xw_\xi + \Delta w_\xi = -G_\xi.$$

Take $\varphi(t)$ smooth with $\varphi(0) = 1$, $\varphi(T) = 0$. Set $v = \varphi(t)w_\xi$. Then

$$v_t + D_pGD_xv + \Delta v = -\varphi G_\xi + \varphi' w_\xi.$$

Integrate the previous identity with respect to ρ. Then

$$-v(x_0, 0) = \int_0^T \int_{\mathbb{T}^d} -\varphi G_\xi \rho + \varphi' w_\xi \rho - \int_{\mathbb{T}^d} v(x, T)\rho(x, T)dx.$$

Note that since $v(x, T) = 0$ we have

$$\int_{\mathbb{T}^d} v(x, T)\rho(x, T)dx = 0.$$

By (43) we have for any small ν

$$|\varphi' w_\xi| \le \nu G(x, Dw) + C_\nu + |\eta| \quad \eta \in L^2.$$

Using this, (45) and Hölder inequality

$$\begin{aligned}
|v(x_0, 0)| &\le \int_0^T \int_{\mathbb{T}^d} C\rho + \nu G\rho + |\eta|\rho + C(A(x)|Dw| + B(x))\rho \\
&\le C + \int_0^T \int_{\mathbb{T}^d} \nu G\rho + C(\|\sqrt{\rho}\|_{L^4(dx)}\|Dw\sqrt{\rho}\|_{L^2(dx)} + \|\rho\|_{L^2(dx)}) \\
&\le C+\nu \int_0^T \int_{\mathbb{T}^d} G\rho + C\big(\|\rho\|_{L^1(L^2(dx),dt)} \int_0^T \int_{\mathbb{T}^d} |Dw|^2\rho\big)^{\frac{1}{2}} + C\|\rho\|_{L^1(L^2(dx),dt)}.
\end{aligned}$$

Then, using (57) and (58)

$$|v(x_0, 0)| \le C + C\nu \operatorname{osc}(w) + C \operatorname{osc}(w)^{\frac{1+\mu}{2}} + C \operatorname{osc}(w)^\mu.$$

Because this inequality is uniform in x_0 and ξ we can take on the left hand side the supremum over x_0 and ξ:

$$\operatorname{osc}(w) \le C \operatorname{Lip}(w) = C \sup_{x_0,\xi} |v(x_0, 0)|.$$

From this we have (choosing ν small enough to absorb the term $C\nu \operatorname{osc}(w)$ in the left-hand side)

$$\operatorname{Lip}(w) \le C + C \operatorname{Lip}(w)^{\frac{1+\mu}{2}}.$$

But then it follows that $\operatorname{Lip}(w)$ is bounded, and so $w \in W^{1,\infty}$. □

From this result we obtain:

Corollary 11. *If $d \le 3$, then $u \in W^{3,2} \cap W^{1,\infty}$.*

Proof. The $W^{1,\infty}$ bound follows from the fact that

$$G(x, p) = H(x, p) - \ln m - \bar{H}$$

satisfies the hypothesis of the previous theorem.

Therefore

$$\Delta u = \bar{H} + \ln m - H(x, Du) \in L^2$$

and so $u \in W^{2,2}$.

We have $D(H(x,Du)) = D_xH + D_pHD^2u \in L^2$, and since $D\ln m \in L^2$ we obtain

$$\Delta Du = D\ln m - D(H(x,Du)) \in L^2$$

which gives $u \in W^{3,2}$. $\square$

To proceed further and obtain higher regularity we need to address the regularity of m.

Proposition 10. $\ln m \in W^{1,\infty}$.

Proof. For this we observe that $w = \ln m$ satisfies

$$\Delta w + |Dw|^2 - \operatorname{div}(D_pH(x,Du)) - D_pH(x,Du)Dw = 0. \tag{59}$$

Therefore $\ln m$ satisfies (41) for

$$G(x,p) = |p|^2 - D_pH(x,Du)p - \operatorname{div}(D_pH(x,Du)).$$

Note first that $D_{x_i}(D_{p_j}H(x,Du)) = D^2_{x_ip_j}H + D^2_{p_jp_l}Hu_{x_lx_i} \in L^4$. Since $u \in W^{3,2} \cap W^{1,\infty}$, from Sobolev's inequality $u \in W^{2,4} \cap W^{1,\infty}$ for $d \leq 3$. So

$$\begin{aligned} D^2_{x_ix_k}(D_{p_l}H(x,Du)) = D^3_{x_ix_kp_j}H + D^3_{x_ip_jp_l}Hu_{x_lx_k} + D^3_{p_jp_lx_k}Hu_{x_lx_i} \\ + D^3_{p_jp_lp_m}Hu_{x_lx_i}u_{x_mx_k} + D^2_{p_jp_l}Hu_{x_lx_ix_k} \in L^2. \end{aligned}$$

From that it is clear that G satisfies all hypothesis of Theorem 6. $\square$

Corollary 12. $m \in W^{2,2}$.

Proof. From the previous proposition we have $m \in W^{1,\infty}$, but then

$$\Delta m = \operatorname{div}(D_pH)m + D_pHDm \in L^2$$

from which we conclude $m \in W^{2,2}$. $\square$

Theorem 7. *If $d \leq 3$, u is smooth.*

Proof. Because $u \in W^{3,2} \cap W^{1,\infty}$ we have $H(x,Du) \in W^{2,2}$. We also have $m \in W^{2,2}$. Thus

$$\Delta u = \ln m + \bar{H} - H(x,Du) \in W^{2,2}.$$

Then by elliptic regularity we have $u \in W^{4,2}$. But then equation (59) satisfied by $w = \ln m$ has $|Dw|^2 \in L^\infty$, $\operatorname{div}(D_pH(x,Du)) \in W^{2,2}$, and $D_pH(x,Du)Dw \in L^\infty$. Hence

$$\Delta w \in L^p,$$

for $1 < p < \infty$ if $d = 2$ or $1 < p \leq 6$ if $d = 2$. This yields $w \in W^{2,p}$ for those values of p. From this we bootstrap $|Dw|^2, D_pHDw \in W^{1,p}$ and so yield $w \in W^{3,2}$, and also $m \in W^{3,2}$.

Then $\Delta u + H(x,Du) \in W^{3,2}$ and we iterate to get $u \in W^{5,2}$. Then we can repeat the iteration to get increased regularity on m and u to any degree of smoothness. $\square$

6. Convexity and uniqueness

In this section we study the convexity of $\bar{H}_\varepsilon$ and $\bar{L}_\varepsilon$ and the uniqueness of minimizers.

Lemma 1. *The effective Lagrangian $\bar{L}_\varepsilon$ is convex.*

Proof. Let μ_1, μ_2 be such that $A_L(\mu_i) = \bar{L}_\varepsilon(h_i)$ and $\rho(\mu_i) = h_i$. Note that, for $0 \leq \lambda \leq 1$,

$$\rho(\lambda\mu_1 + (1-\lambda)\mu_2) = \lambda\rho(\mu_1) + (1-\lambda)\rho(\mu_2) = \lambda h_1 + (1-\lambda)h_2$$

and

$$m_{\lambda\mu_1+(1-\lambda)\mu_2} = \lambda m_{\mu_1} + (1-\lambda)m_{\mu_2}.$$

Then, since the entropy $\bar{S}$ is convex, we have that

$$\begin{aligned}\bar{L}_\varepsilon(\lambda\mu_1 + (1-\lambda)\mu_2) &\leq A_L(\lambda\mu_1 + (1-\lambda)\mu_2)\\ &\leq \lambda A_L(\mu_1) + (1-\lambda)A_L(\mu_2)\\ &= \lambda\bar{L}_\varepsilon(h_1) + (1-\lambda)\bar{L}_\varepsilon(h_2).\end{aligned}$$

□

Lemma 2. *Either for quadratic Hamiltonians, or for $H(p,x)$ strictly convex Hamiltonians in p and in dimension 2, $\bar{H}_\varepsilon(P)$ is strictly convex. Furthermore for each P* (1) *admits at most one minimizer, up to the addition of constants.*

Proof. In fact, let $P_0, P_1 \in \mathbb{R}^d$ and $0 < \lambda < 1$ be such that

$$\bar{H}_\varepsilon(\lambda P_0 + (1-\lambda)P_1) = \lambda\bar{H}_\varepsilon(P_0) + (1-\lambda)\bar{H}_\varepsilon(P_1).$$

Let $f, g \in C^2(\mathbb{T}^d)$ such that

$$\begin{aligned}\bar{H}_\varepsilon(P_0) &= \varepsilon_1 \log \int e^{\frac{\varepsilon_2\Delta f + H(x,P_0+Df)}{\varepsilon_1}}dx,\\ \bar{H}_\varepsilon(P_1) &= \varepsilon_1 \log \int e^{\frac{\varepsilon_2\Delta g + H(x,P_1+Dg)}{\varepsilon_1}}dx.\end{aligned}$$

Set $\varphi = \lambda f + (1-\lambda)g$. Then

$$\begin{aligned}\Delta\varphi &= \lambda\Delta f + (1-\lambda)\Delta g,\\ D\varphi + \lambda P_0 + (1-\lambda)P_1 &= \lambda(Df + P_0) + (1-\lambda)(Dg + P_1),\end{aligned}$$

and, by convexity,

$$H(x, \lambda P_0 + (1-\lambda)P_1 + D\varphi) \leq \lambda H(x, Df + P_0) + (1-\lambda)(H(x, Dg + P_1)). \tag{60}$$

By convexity of the exponential function and Hölder inequality we get

$$\begin{aligned}e^{\frac{\bar{H}(\lambda P_0+(1-\lambda)P_1)}{\varepsilon_1}} &\leq \int e^{\frac{\varepsilon_2\Delta\varphi + H(x,\lambda P_0+(1-\lambda)P_1+D\varphi)}{\varepsilon_1}}dx\\ &\leq \int e^{\frac{\lambda}{\varepsilon_1}(\varepsilon_2\Delta f + H(x,Df+P_1)) + \frac{(1-\lambda)}{\varepsilon_1}(\varepsilon_2\Delta g + H(x,Dg+P_2))}dx\\ &\leq \left(\int e^{\frac{\varepsilon_2\Delta f + H(x,Df+P_1)}{\varepsilon_1}}dx\right)^\lambda \cdot \left(\int e^{\frac{\varepsilon_2\Delta g + H(x,Dg+P_2)}{\varepsilon_1}}dx\right)^{(1-\lambda)}\\ &= e^{\frac{\lambda\bar{H}_\varepsilon(P_0)}{\varepsilon_1}} e^{\frac{(1-\lambda)\bar{H}_\varepsilon(P_1)}{\varepsilon_1}} = e^{\frac{\bar{H}_\varepsilon(\lambda P_0+(1-\lambda)P_1)}{\varepsilon_1}}. \qquad (61)\end{aligned}$$

Therefore all inequalities in (61) are equalities, and so is (60). Since H is strictly convex, $Df + P_0 = Dg + P_1$ at all points. Hence $P_1 - P_0 = D(f-g)$ is an exact differential and then $P_0 = P_1$ and $f = g$, up to a constant. $\square$

7. Smoothness of the effective Lagrangian and Hamiltonian functions

In this section we fix ε and study the differentiability of $\bar{H}_\varepsilon$ and $\bar{L}_\varepsilon$, i.e. Theorem 3, both for mechanical Hamiltonians and for the two dimensional case.

Proof. As in [I-SM], to prove the differentiability of $\bar{H}_\varepsilon$ we use the implicit function theorem (see for example [D1], Chapter X).

The minimizer $\phi(\cdot, P)$ in (1) satisfies (22) with $\varphi(x) = Px + \phi(x,P)$. Let

$$H_0^k(\mathbb{T}^d, \mathbb{R}) = \left\{ f \in H^k(\mathbb{T}^d, \mathbb{R}) : \int f = 0 \right\}.$$

Define

$$F : \mathbb{R}^d \times H_0^k(\mathbb{T}^d, \mathbb{R}) \to H_0^{k-4}(\mathbb{T}^d, \mathbb{R})$$

by

$$F(P,\phi) := \varepsilon_2 \Delta m_\varphi - \operatorname{div}(m_\varphi D_p H(x, P + D\phi)). \tag{62}$$

For k large enough, the map F is C^∞. Indeed the map F can be obtained by composing derivatives of ϕ, which is a C^∞ operation from H^k to H^{k-1}, with smooth functions. But for k large enough if $G : \mathbb{R}^n \to \mathbb{R}^m$ is C^∞, then the map

$$\bar{G} : H^k(\mathbb{T}^d, \mathbb{R}^n) \to H^k(\mathbb{T}^d, \mathbb{R}^m), \quad \bar{G}(\Psi) = G \circ \Psi$$

is C^∞. The partial derivative of F, the Euler-Lagrange equation, with respect to the variable ϕ can be regarded as a linear map

$$\mathcal{L} := D_2 F(P,\phi) : H_0^k(\mathbb{T}^d, \mathbb{R}) \to H_0^{k-4}(\mathbb{T}^d, \mathbb{R})$$

which is given by

$$\begin{aligned}
\mathcal{L}(\psi) \quad &= \quad \frac{\varepsilon_2}{\varepsilon_1} \Delta \left[m_\varphi (\varepsilon_2 \Delta\psi + D_p H(x, P + D\phi) D\psi) \right] \\
&- \quad \operatorname{div} \left[\frac{m_\varphi}{\varepsilon_1} (\varepsilon_2 \Delta \psi + D_p H(x, P + D\phi) D\psi) D_p H(x, P + D\phi) \right] \\
&- \quad \operatorname{div} \left[m_\varphi D_{pp}^2 H(x, P + D\phi) D\psi \right].
\end{aligned}$$

$\mathcal{L}$ is a fourth-order elliptic Fredholm PDO of index 0.

Lemma 3. *$\mathcal{L}$ is an isomorphism.*

Proof. We have

$$\begin{aligned}
\int \mathcal{L}(\psi)\psi \quad &= \quad \int \frac{m_\varphi}{\varepsilon_1} (\varepsilon_2 \Delta\psi + D_p H(x, P + D\phi) D\psi)^2 \\
&+ \quad \int m_\varphi \langle D_{pp}^2 H(x, P + D\phi) D\psi, D\psi \rangle.
\end{aligned} \tag{63}$$

We claim that there is a constant C such that $\int L(\psi)\psi \geq C\|\psi\|_{H_0^2}$. Otherwise there is a sequence $\psi_n \in H_0^2$ with $\|\psi\|_{H_0^2} = 1$ such that $\int L(\psi_n)\psi_n$ converges to zero. In such a case (63) implies that $\|D\psi_n\|_{L^2}$ converges to zero. Since

$$2\varepsilon_2 \Delta\psi D_p H(x, P + D\phi) D\psi \geq -\frac{1}{2} (\varepsilon_2 \Delta \psi)^2 - 2(D_p H_p(P + D\phi) D\psi)^2,$$

there is a constant $\delta > 0$ such that

$$\int \frac{m_\varphi}{\varepsilon_1}[(\varepsilon_2\Delta\psi + H_p(I,P+D\phi)D\psi)^2 + (H_p(I,P+D\phi)D\psi)^2] \geq \int \frac{m_\varphi}{2\varepsilon_1}(\varepsilon_2\Delta\psi)^2 \geq \delta \int (\Delta\psi)^2. \tag{64}$$

Therefore $\|D^2\psi_n\|_{L^2} = \|\Delta\psi_n\|_{L^2}$ converges to zero. Thus $\mathcal{L}$ is one to one and then an isomorphism. □

The previous lemma makes clear the use of H_0^k as the domain and target for $\mathcal{L}$, that is, to have $\ker(\mathcal{L})$ and $\operatorname{coker}(\mathcal{L})$ equal to zero.

We can now apply the implicit function theorem to conclude that $\phi(,P)$ is smooth in P.

Consequently, $\bar{H}_\varepsilon$ is smooth, and since it is also strictly convex, the map $D\bar{H}_\varepsilon$ has a smooth inverse γ_ε and therefore $\bar{L}_\varepsilon(Q) = h\gamma_\varepsilon(Q) - \bar{H}_\varepsilon(\gamma_\varepsilon(Q))$ is also smooth. □

8. Limiting behavior

In this section we fix $P \in \mathbb{R}^d$ and prove Theorem 4. We assume we are either in dimension $d = 2$ or in the mechanical case. Let u_ε be the minimizer in (1). By (37), u_ε is a priori bounded in $W^{1,2}$. Thus, through some subsequence there exists $u \in W^{1,2}$ such that $u_\varepsilon \rightharpoonup u$, in $W^{1,2}$, as $\varepsilon \to 0$. Also, since $\bar{H}_\varepsilon$ is bounded it converges through a subsequence to a limit. We have the following two results:

Proposition 11.

$$\limsup_{\varepsilon\to(0,0)} \bar{H}_\varepsilon(P) \leq \bar{H}(P).$$

Proof. Let ϕ_{ε_2} be a solution to (5). Then

$$e^{\bar{H}_\varepsilon(P)/\varepsilon_1} \leq \int e^{\varepsilon_2\Delta\phi_{\varepsilon_2}(x)+H(x,P+D\phi_{\varepsilon_2})/\varepsilon_1}dx = e^{\bar{H}_\varepsilon/\varepsilon_1} \leq e^{\check{H}_{\varepsilon_2}(P)/\varepsilon_1}$$

and so

$$\limsup_{\varepsilon\to(0,0)} \bar{H}_\varepsilon(P) \leq \lim_{\varepsilon_2\to 0} \check{H}_{\varepsilon_2}(P) = \bar{H}(P).$$

□

We now address the opposite inequality:

Proposition 12. *Suppose $\varepsilon \to 0$ with $\frac{\varepsilon_1}{\varepsilon_2}$ bounded. Then*

$$\liminf \bar{H}_\varepsilon(P) \geq \bar{H}(P). \tag{65}$$

Proof. Suppose that (65) is not true. Then there is $\delta > 0$ such that

$$\liminf_{\varepsilon\to(0,0)} \bar{H}_\varepsilon(P) < \bar{H}(P) - 2\delta.$$

Let

$$B_\varepsilon = \{x \in \mathbb{T}^d : \varepsilon_2\Delta u_\varepsilon(x) + H(x,P+Du_\varepsilon) > \bar{H}(P) - \delta\}.$$

Then for some sequence $\varepsilon^k \to (0,0)$ we have

$$\varepsilon_1^k \log|B_{\varepsilon^k}| + \bar{H}(P) - \delta \leq \bar{H}(P) - 2\delta.$$

Therefore,

$$|B_{\varepsilon^k}| \leq \exp(-\frac{\delta}{\varepsilon_1^k}).$$

Let $f_\varepsilon = \varepsilon_2 \Delta u_\varepsilon + H(P + Du_\varepsilon, x)$ and $g \in C^\infty(\mathbb{T}^d)$ be a non-negative function. By lower semicontinuity

$$\begin{aligned}
\int_{\mathbb{T}^d} H(x, P+Du)g &\leq \liminf_{k\to\infty} \int_{\mathbb{T}^d} -\varepsilon_2^k Du_{\varepsilon^k} Dg + H(x, P + Du_{\varepsilon^k})g \\
&\leq \liminf_{k\to\infty} \int_{\mathbb{T}^d} f_{\varepsilon^k} g \\
&= \liminf_{k\to\infty} (\int_{B_\varepsilon} f_{\varepsilon^k} g + \int_{B_\varepsilon^c} f_{\varepsilon^k} g) \\
&\leq (\bar{H}(P) - \delta) \int_{\mathbb{T}^d} g + \liminf_{k\to\infty} \int_{B_\varepsilon} f_{\varepsilon^k} g \\
&\leq (\bar{H}(P) - \delta) \int_{\mathbb{T}^d} g + \liminf_{k\to\infty} \|f_{\varepsilon^k}\|_2 \|\chi_{B_\varepsilon} g\|_2 \\
&\leq (\bar{H}(P) - \delta) \int_{\mathbb{T}^d} g.
\end{aligned}$$

Since, assuming $\frac{\varepsilon_1}{\varepsilon_2}$ is bounded, we have by Corollary 5 that $\|f_\varepsilon\|_2 \leq C$. This in particular shows that

$$H(x, P + Du) \leq \bar{H}(P) - \delta$$

almost everywhere, and so u is Lipschitz. Then there is $\rho > 0$ such that

$$H(x, P + Du(y)) < H(y, P + Du(y)) + \frac{\delta}{2} \text{ a.e.,} \quad d(x, y) < \frac{\rho}{2}.$$

Set $\psi_\rho = \eta_\rho * u$, where η_ρ is the standard mollifier supported in $B(0, \rho)$. From the convexity of H and Jensen's inequality

$$\begin{aligned}
H(x, P + D\psi_\rho(x)) &\leq \int H(x, P + Du(y))\eta_\rho(y - x)dy \\
&\leq \int (H(y, P + Du(y)) + \frac{\delta}{2})\eta_\rho(y - x)dy \\
&\leq \bar{H}(P) - \frac{\delta}{2} \qquad (66)
\end{aligned}$$

for all $x \in \mathbb{T}^d$, which contradicts (3). □

The two previous propositions combined yield the following corollary.

Corollary 13. *Let u_ε be a sequence of minimizers of* (1). *Suppose as $\varepsilon \to (0, 0)$, with $\frac{\varepsilon_1}{\varepsilon_2}$ bounded, $u_\varepsilon \rightharpoonup u$ in $W^{1,2}$. Then*

$$H(x, P + Du) \leq \bar{H}(P)$$

almost everywhere and in viscosity sense.

Suppose H is a mechanical Hamiltonian. Let $u = \lim_{\varepsilon^k \to (0,0)} u_\varepsilon$, where the limit is taken through an appropriate sequence such that $\frac{\varepsilon_1^k}{\varepsilon_2^k}$ is bounded. Because u_ε is uniformly Lipschitz, we can assume that the convergence of u_{ε^k} to u is uniform.

Proposition 13. *Suppose H is a mechanical Hamiltonian. Let $u = \lim_{\varepsilon^k \to (0,0)} u_\epsilon$, where the limit is taken through an appropriate sequence such that $\frac{\varepsilon_1^k}{\varepsilon_2^k} \to 0$. Then u is a solution to*

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

Proof. Let v_{ε^k} be defined through (28). Because $u_{\varepsilon^k} - v_{\varepsilon^k}$ is uniformly Lipschitz, by Theorem 5 we have $\frac{\varepsilon_1^k}{\varepsilon_2^k}(u_{\varepsilon^k} - v_{\varepsilon^k})$, which is uniformly bounded by (31) and whose Lipschitz constant converges to zero, must converge, through a suitable subsequence, to a constant, which then has to be $\bar{H}(P)$. □

Proposition 14. *Assuming $\frac{\varepsilon_1}{\varepsilon_2}$ bounded we have $\mu^\varepsilon \to \mu$ with μ Mather.*

Proof. For some sequence $\varepsilon^k \to (0,0)$ the measures μ_{ε^k} converge weakly to a measure μ. For any $\psi \in C^2(\mathbb{T}^d)$

$$\begin{aligned}\int_{\mathbb{T}^d\times R^d} D\psi(v)d\mu &= \lim_{k\to\infty}\int D\psi\cdot H_p(x,P+Du_{\varepsilon^k}(x))m_{\varepsilon^k}(x)dx\\ &= -\lim_{k\to\infty}\varepsilon_2^k\int \Delta\psi(x)m_{\varepsilon^k}(x)dx\\ &= 0.\end{aligned}$$

Thus μ is holonomic. Recall that $f_\varepsilon = \varepsilon_2\Delta u_\varepsilon + H(P+Du_\varepsilon, x)$ is uniformly bounded from below for the mechanical Hamiltonian or in L_2 assuming $\frac{\varepsilon_1}{\varepsilon_2}$ is bounded for a general Hamiltonian. For $\lambda > 0$

$$\begin{aligned}\bar{H}_\varepsilon(P) &= \int_{\{f_\varepsilon\geq \bar{H}_\varepsilon(P)-\lambda\}} \bar{H}_\varepsilon(P)m_\varepsilon + \int_{\{f_\varepsilon< \bar{H}_\varepsilon(P)-\lambda\}} \bar{H}_\varepsilon(P)m_\varepsilon\\ &\leq \int_{\{f_\varepsilon\geq \bar{H}_\varepsilon(P)-\lambda\}} (f_\varepsilon+\lambda)m_\varepsilon + \int_{\{f_\varepsilon< \bar{H}_\varepsilon(P)-\lambda\}} (\bar{H}_\varepsilon(P)+f_\varepsilon-f_\varepsilon)m_\varepsilon\\ &\leq \int f_\varepsilon m_\varepsilon + \lambda + Ce^{-\varepsilon_1\lambda}.\end{aligned}$$

Thus

$$\bar{H}(P) = \lim_{\varepsilon\to 0}\bar{H}_\varepsilon(P) \leq \liminf_{\varepsilon\to 0}\int f_\varepsilon m_\varepsilon + \lambda.$$

Since $\lambda > 0$ is arbitrary

$$\bar{H}(P) \leq \liminf_{\varepsilon\to 0}\int f_\varepsilon m_\varepsilon.$$

Thus

$$\begin{aligned}\int L(x,v)d\mu &= \lim_{k\to\infty}\int L(x, D_pH(P+Du_{\varepsilon^k},x))m_{\varepsilon^k}(x)dx\\ &= \lim_{k\to\infty}\int ((P+Du_{\varepsilon^k})D_pH(P+Du_{\varepsilon^k},x) - H(P+Du_{\varepsilon^k},x))m_{\varepsilon^k}(x)dx\\ &= \lim_{k\to\infty}\int (PD_pH(P+Du_{\varepsilon^k},x) - f_{\varepsilon^k})\, m_{\varepsilon^k}(x)dx\\ &\leq P\int v d\mu - \bar{H}(P) \leq \bar{L}(\int v d\mu) \leq \int L(x,v)d\mu\end{aligned}$$

so all inequalities are equalities. □

Proof of Theorem 4. The first point of the theorem, the convergence of $\bar{H}_\varepsilon$, follows from Propositions 11 and 12. The second and third points follow from Corollary 13 and the last one is simply Proposition 14. □

9. Connection with mean-field games

Recently, Caines and his co-workers [HMC06], [HMC07], and independently Lions and Lasry [LL06a], [LL06b], [LL07a] and [LL07b] (see also [LLG10a], [LLG10b]), introduced a new class of problems called mean-field games, some of which have a surprising connection with the Evans-Aronsson problems.

In the stationary setting, one of the main examples of mean-field games is described by systems of equations of the form

$$\begin{cases} \varepsilon_2 \Delta u + F(x, Du, m) = \bar{F}, \\ \varepsilon_2 \Delta m - \operatorname{div}(D_p F m) = 0, \end{cases}$$

where $\bar{F}$ is a suitable constant. The variational problem studied in this paper is included in this class, and if we set

$$F(x, p, m) = H(x, p) - \varepsilon_1 \ln m$$

we have

$$\begin{cases} \varepsilon_2 \Delta u + H(x, Du) = \bar{H}_\varepsilon + \varepsilon_1 \ln m, \\ \varepsilon_2 \Delta m - \operatorname{div}(D_p H m) = 0. \end{cases} \tag{67}$$

The results in this paper imply the existence of stationary solutions of this class of mean-field games both for quadratic Hamiltonians and for two dimensional problems (when H satisfies (9)-(12)). For quadratic Hamiltonians Hopf-Cole transformations are widely used; see for example [Ge1] where time dependent mean-field games are studied. However, the regularity estimates in section 4 are new. The convergence results of the previous section establish the existence of a vanishing coupling ($\varepsilon_1 \to 0$) and viscosity ($\varepsilon_2 \to 0$) limit for the mean-field games. Estimates similar to the ones for $\int m^{2^*/2}$ and $\int |D^2 u|^2 m + |D \ln m|^2$ were obtained in the deterministic case by Evans [E2]. We end the paper with a new estimate for third derivatives of the solution:

Proposition 15. *Let u be a solution of* (67)*. Then, if ε_2 is small enough*

$$\int_{\mathbb{T}^d} |\varepsilon_2 D\Delta u|^2 m \leq \frac{C}{\varepsilon_2}.$$

Proof. Differentiating the first equation in (67) and multiplying by $2Du$, we get

$$\varepsilon_2 \Delta |Du|^2 - 2\varepsilon_2 |D^2 u|^2 + D_p H D|Du|^2 + 2D_x H Du = 2\varepsilon_1 Df Du,$$

where $f = \ln m$, as before. Letting $v = |Du|^2$ and multiplying by v^k we obtain

$$(\varepsilon_2 \Delta v) v^k - 2\varepsilon_2 |D^2 u|^2 v^k + (D_p H Dv) v^k + 2(D_x H Du) v^k = 2\varepsilon_1 (Df Du) v^k.$$

Multiplying by m we have

$$\frac{m}{k+1}(\varepsilon_2 \Delta v^{k+1} + D_p H D v^{k+1}) - 2\varepsilon_2 |D^2 u|^2 v^k m - k\varepsilon_2 |Dv|^2 v^{k-1} m \\ + 2(D_x H Du) v^k m = 2\varepsilon_1 (Df Du) v^k m.$$

Integrating,

$$\varepsilon_2 \int 2|D^2 u|^2 v^k m + k|Dv|^2 v^{k-1} m = 2\int (D_x H Du - \varepsilon_1 Df Du) v^k m. \tag{68}$$

Thus, for $k=1$ we have (using $|D_xH| \le C + Cv$)

$$2\varepsilon_2 \int |D^2u|^2 vm + \varepsilon_2 \int |Dv|^2 m$$
$$\le C \int (v^{3/2} + v^{5/2} + \varepsilon_1 |Df|^2) m + \varepsilon_1 \int v^3 m.$$

Then, because

$$v^{5/2} \le Cv + Cv^3 \le v(C + C\varepsilon_2^2 |D^2u|^2 + C\varepsilon_1^2 f^2)$$
$$\varepsilon_1 v^3 \le \varepsilon_1 v(C + C\varepsilon_2^2 |D^2u|^2 + C\varepsilon_1^2 f^2)$$

we have

$$2\varepsilon_2 \int |D^2u|^2 vm + \varepsilon_2 \int |Dv|^2 m$$
$$\le C \int (v^{3/2} + \varepsilon_1 v + \varepsilon_1 v^2 + \varepsilon_1^5 f^4 + \varepsilon_1 |Df|^2) m + \varepsilon_2^2 C \int |D^2u|^2 vm.$$

Then $\varepsilon_2 \int |D^2u|^2 vm$, $\varepsilon_2 \int |Dv|^2 m$ are uniformly bounded, if ε_2 is small enough.

Now $D(H(x,Du)) = D_xH + D_pHD^2u$, so

$$|D(H(x,Du)|^2 \le Cv^2 + C + C(v+1)|D^2u|^2$$

and then

$$\int |D(H(x,Du)|^2 m \le \frac{C}{\varepsilon_2}.$$

So

$$\int |\varepsilon_2 D\Delta u|^2 m \le C\varepsilon_1^2 \int |Df|^2 m + C \int |D(H)|^2 m \le C + \frac{C}{\varepsilon_2}.$$

□

For low dimensions we could use (68) for $k > 1$ to get additional estimates.

As a final remark we should observe that variational methods in mean-field games is, as far as the authors are aware, a largely unexplored direction which may yield new estimates and results in this very challenging set of problems.

References

[HMC06] M.Y. Huang, R.P. Malhame and P.E. Caines. Large Population Stochastic Dynamic Games: Closed Loop McKean-Vlasov Systems and the Nash Certainty Equivalence Principle. *Special issue in honor of the 65th birthday of Tyrone Duncan, Communications in Information and Systems.* Vol. 6, Number 3, 2006, pp 221 - 252. MR2346927 (2009f:91008)

[HMC07] M.Y. Huang, R.P. Malhame and P.E. Caines. Population Cost-Coupled LQG Problems with Non-uniform Agents: Individual-Mass Behavior and Decentralized - Nash Equilibria. *IEEE Trans. on Automatic Control*, Vol. 52. No. 9, September, 2007, pp 1560 - 1571. MR2352434 (2008g:91047)

[D1] J. Dieudonné. *Foundations of Modern Analysis.* Vol. I, Academic Press, New York, 1969. MR0349288 (50:1782)

[E1] L.C. Evans. *Some new PDE methods for weak KAM theory.* Calc. Var. Partial Differential Equations **17** (2003), no. 2, 159 – 177. MR1986317 (2004e:37097)

[E2] L.C. Evans. *Further PDE methods for weak KAM theory.* Calc. Var. Partial Differential Equations **35** (2009), no. 4, 435 – 462. MR2496651 (2011f:37111)

[Eva10] L. C. Evans, *Adjoint and compensated compactness methods for Hamilton-Jacobi PDE*, Arch. Ration. Mech. Anal., **197** (2010), 1053–1088. MR2679366 (2011h:35042)

[G] D. Gomes. *A stochastic analog of Aubry-Mather theory.* Nonlinearity **15** (2002), 3, 581–603. MR1901094 (2003b:37096)

[GISMY] D. Gomes, Iturriaga R., Sánchez-Morgado H., Y. Yu. *Mather measures selected by an approximation scheme.* Proc. Amer. Math. Soc. **138** (2010), no. 10, 3591–3601. MR2661558 (2011h:37093)

[Ge1] O. Guéant. *Mean field games equations with quadratic hamiltonian: a specific approach.* working paper, 2011.

[I-SM] R. Iturriaga and H. Sánchez-Morgado. *On the stochastic Aubry Mather theory.* Bol. Soc. Mat. Mex. (3) vol 11 no. 1 (2005) 91–99. MR2198592 (2006k:37166)

[LL06a] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. I. Le cas stationnaire. *C. R. Math. Acad. Sci. Paris*, 343(9):619–625, 2006. MR2269875 (2007m:91021)

[LL06b] Jean-Michel Lasry and Pierre-Louis Lions. Jeux à champ moyen. II. Horizon fini et contrôle optimal. *C. R. Math. Acad. Sci. Paris*, 343(10):679–684, 2006. MR2271747 (2007m:91022)

[LL07a] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Jpn. J. Math.*, 2(1):229–260, 2007. MR2295621 (2008k:91034)

[LL07b] Jean-Michel Lasry and Pierre-Louis Lions. Mean field games. *Cahiers de la Chaire Finance et Développement Durable*, 2007. MR2295621 (2008k:91034)

[LLG10a] Jean-Michel Lasry, Pierre-Louis Lions, and O. Guéant. Application of mean field games to growth theory. *preprint*, 2010.

[LLG10b] Jean-Michel Lasry, Pierre-Louis Lions, and O. Guéant. Mean field games and applications. *Paris-Princeton lectures on Mathematical Finance*, 2010. MR2762362

[M] R. Mañé. *Generic properties and problems of minimizing measures of Lagrangian systems.* Nonlinearity **9** (1996), 273-310. MR1384478 (97d:58118)

[Ma] J. Mather. *Action minimizing invariant measure for positive definite Lagrangian systems.* Math. Z. **207** (1991), 169-207. MR1109661 (92m:58048)

[W] K. Wang. *Action minimizing stochastic invariant measures for a class of Lagrangian systems.* Comm. Pure App. Analysis. **7** no. 5 (2008). 1211-1223 MR2410876 (2010a:37105)

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

E-mail address: dgomes@math.ist.utl.pt

Instituto de Matemáticas, Universidad Nacional Autónoma de México, México DF 04510, México

E-mail address: hector@matem.unam.mx