Available online at www.sciencedirect.com

ScienceDirect

Ann. I. H. Poincaré – AN 26 (2009) 521–545

ANNALES DE L'INSTITUT HENRI POINCARÉ
ANALYSE NON LINÉAIRE

www.elsevier.com/locate/anihpc

Minimizers of Dirichlet functionals on the n-torus and the weak KAM theory

G. Wolansky

Department of Mathematics, Technion – Israel Institute of Technology, Haifa, Israel

Received 2 May 2007; accepted 11 September 2007

Available online 24 January 2008

Abstract

Given a probability measure μ on the n-torus $\mathbb{T}^n$ and a rotation vector $\boldsymbol{P} \in \mathbb{R}^n$, we ask whether there exists a minimizer to the integral $\int_{\mathbb{T}^n} |\nabla\phi + \boldsymbol{P}|^2\,d\mu$. This problem leads, naturally, to a class of elliptic PDE and to an optimal transportation (Monge–Kantorovich) class of problems on the torus. It is also related to higher dimensional Aubry–Mather theory, dealing with invariant sets of periodic Lagrangians, and is known as the "weak-KAM theory".

Résumé

Etant donnés une mesure μ sur le tore n-dimensionnel $\mathbb{T}^n$ et un vecteur de rotation $\boldsymbol{P} \in \mathbb{R}^n$, on étudie la question de l'existence d'un minimiseur pour l'intégrale $\int_{\mathbb{T}^n} |\nabla\phi + \boldsymbol{P}|^2\,d\mu$. Ce problème conduit naturellement à une classe d'équations aux dérivées partielles elliptiques et à une classe de problèmes de transport optimal (Monge–Kantorovich) sur le tore. Il est aussi lié à la théorie d'Aubry–Mather en dimension supérieure, qui traite les ensembles invariants pour des Lagrangiens périodiques, connue sous le nom de théorie KAM faible.

Keywords: Monge–Kantorovich; Optimal mass transport; Periodic Lagrangian; Effective Hamiltonian; Rotation vector; Dirichlet functional; KAM theory

1. Overview

1.1. Motivation

Consider the functional

$$H^{\varepsilon}_{\boldsymbol{P}}(u) := \frac{\varepsilon^2}{2}\int\limits_{\mathbb{T}^n} \left|\nabla u + i\varepsilon^{-1}\boldsymbol{P}u\right|^2 dx - G\left(|u|\right) \tag{1.1}$$

where $\mathbb{T}^n := \mathbb{R}^n \bmod \mathbb{Z}^n$ is the flat n-torus, $\boldsymbol{P} \in \mathbb{R}^n$ a prescribed, constant vector, $u \in \mathbf{W}^{1,2}(\mathbb{T}^n;\mathbb{C})$ is normalized via $\int_{\mathbb{T}^n} |u|^2\,dx = 1$ and G is convex (possibly nonlocal) functional of $|u|$. A critical point u of $H^{\varepsilon}_{\boldsymbol{P}}$ can be considered as a

E-mail address: gershonw@math.technion.ac.il.

0294-1449/$ – see front matter
doi:10.1016/j.anihpc.2007.09.007

periodic function on $\mathbb{R}^n$. The function

$$u_0(x) = e^{i\boldsymbol{P}\cdot x/\varepsilon} u(x) \tag{1.2}$$

is considered as a function on $\mathbb{R}^n$ as well.

Examples.

(i) $G(|u|) = -\int_{\mathbb{T}^n} \Xi(x)|u|^2\,dx$ where Ξ is a smooth potential on $\mathbb{T}^n$. A critical point of (1.1) is an eigenvalues problem of the Schrödinger operator $H_{\boldsymbol{P}} := -\varepsilon^2(\nabla + i\boldsymbol{P})^2 + 2\Xi$ on the torus. The substitution (1.2) leads to a *Bloch state*

$$-\varepsilon^2\Delta u_0 + 2\Xi u_0 + E u_0 = 0 \quad \text{on } \mathbb{R}^n. \tag{1.3}$$

(ii) Self-focusing nonlinear Schrödinger equation. Here $G(|u|) = \int_{\mathbb{T}^n} |u|^\sigma$ where $2 < \sigma < 2(2+n)/n$. An extremum u of $H^\varepsilon_{\boldsymbol{P}}$ with this choice is given by the nonlinear eigenvalue problem for u_0 on $\mathbb{R}^n$:

$$\varepsilon^2\Delta u_0 + \sigma|u_0|^{\sigma-1}u_0 = E u_0. \tag{1.4}$$

(iii) The choice

$$G(|u|) := \sup_{V\in C^1(\mathbb{T}^n)} \left(-\int_{\mathbb{T}^n} \frac{1}{2\gamma}|\nabla V|^2 - V(|u|^2 - 1) \right) dx,$$

leads to the Schrödinger Poisson system (see, e.g. [1,14,26]) for attractive (gravitational) field. Again, u_0 solves

$$-\varepsilon^2\Delta u_0 - V u_0 = E_{(\boldsymbol{P})} u_0$$

where V is a periodic function on $\mathbb{R}^n$ solving $\Delta V + \gamma(|u_0|^2 - 1) = 0$.

In addition to the (rather obvious) spectral asymptotic questions, there are additional motivations for the study of this problem, as described below.

The short wavelength limit of the reduced wave equation in a periodic lattice is described as

$$\Delta u_0 + \varepsilon^{-2}N(x)u_0 = 0,$$

where $N(x)$ is the (ε independent) periodic function representing the *refraction index* of the lattice and $\varepsilon \to 0$ stands for the wavelength. See, e.g. [21]. Suppose one can measure the intensity $|u_0|$ and the direction $\widehat{\boldsymbol{P}} := \boldsymbol{P}/|\boldsymbol{P}|$ of the carrier wavenumber of an electromagnetic wave $u_0(x) = e^{i\boldsymbol{P}\cdot x}u(x)$ in this lattice (here, again, u is periodic). Then N can be recovered from

$$N \equiv \frac{1}{2}|\nabla\phi + \widehat{\boldsymbol{P}}|^2 \tag{1.5}$$

where ϕ is the minimizer of $F(\rho, \widehat{\boldsymbol{P}})$ for a normalized $\rho = |u_0|^2$ (see (1.6) below). Alternatively, suppose we need to *design* a lattice for a prescribed electromagnetic intensity $|u_0|^2$ and wave propagation $\widehat{\boldsymbol{P}}$. Then (1.5) is the solution as well!

1.2. The effective Hamiltonian

Note that, in cases (i)–(iii), I referred to *critical points* of $H^\varepsilon_{\boldsymbol{P}}$. If one looks at minimizers of this functional, or even critical points of finite (ε-independent) Morse index, then one may expect singular limits as $\varepsilon \to 0$. However, there is a formal way to obtain nonsingular limits of these equations as $\varepsilon \to 0$ as follows:

Substitute the WKB anzatz (see [15]) $u_\varepsilon := \sqrt{\rho}\,e^{i\phi/\varepsilon}$ in (1.1), where $\phi \in C^1(\mathbb{T}^n)$ and $\rho \in C^1(\mathbb{T}^n)$ is nonnegative function satisfying $\int_{\mathbb{T}^n} \rho = 1$. Then

$$\lim_{\varepsilon\to 0} H^\varepsilon_{\boldsymbol{P}}(u_\varepsilon) = \frac{1}{2}\int_{\mathbb{T}^n} |\nabla\phi + \boldsymbol{P}|^2\rho(x)\,dx - G(|\rho|^{1/2}).$$

Now, define

$$F(\rho,\boldsymbol{P}):=\frac{1}{2}\inf_{\phi\in C^1(\mathbb{T}^n)}\int_{\mathbb{T}^n}|\nabla\phi+\boldsymbol{P}|^2\rho(x)\,dx \tag{1.6}$$

and

$$H_G(\rho,\boldsymbol{P}):=F(\rho,\boldsymbol{P})-G\big(|\rho|^{1/2}\big). \tag{1.7}$$

Note that F is concave as a function of ρ for fixed $\boldsymbol{P}\in\mathbb{R}^n$. Then we can (at least formally) look for the *maximizer* of $H_G(\cdot,\boldsymbol{P})$ over the set of densities $\rho\in\mathbb{L}^1(\mathbb{T}^n;\mathbb{R}^+)$. It is the *asymptotic energy spectrum H* associated with $H^\varepsilon_{\boldsymbol{P}}$:

$$\widehat{H}_G(\boldsymbol{P}):=\sup_\rho\left\{H_G(\rho,\boldsymbol{P});\ \rho\in\mathbb{L}^1\big(\mathbb{T}^n;\mathbb{R}^+\big),\ \int_{\mathbb{T}^n}\rho=1\right\}. \tag{1.8}$$

The pair (ϕ,ρ) which realizes the minimum (resp. maximum) of F (resp. $\widehat{H}_G(\cdot,\boldsymbol{P})$) corresponds to an asymptotic critical point of $H^\varepsilon_{\boldsymbol{P}}$ and verifies the Euler–Lagrange equations:

$$\nabla\cdot\big[\rho(\nabla\phi+\boldsymbol{P})\big]=0, \tag{1.9}$$

$$\frac{1}{2}|\nabla\phi+\boldsymbol{P}|^2-G_\rho=E \tag{1.10}$$

on $\mathbb{T}^n$, where G_ρ is the Frèchet derivative of $\rho\to G(|\rho|^{1/2})$ and E is a Lagrange multiplier corresponding to the constraint $\int_{\mathbb{T}^n}\rho=1$.

Of particular interest is the linear case (1.3). Here (1.7), (1.8) are reduced into

$$H_\Xi(\rho,\boldsymbol{P}):=F(\rho,\boldsymbol{P})+\int_{\mathbb{T}^n}\Xi\rho, \tag{1.11}$$

$$\widehat{H}_\Xi(\boldsymbol{P}):=\sup_\rho\left\{H_\Xi(\rho,\boldsymbol{P});\ \rho\geqslant 0,\ \int_{\mathbb{T}^n}\rho=1\right\} \tag{1.12}$$

and (1.10) takes the form of the Hamilton–Jacobi equation on the torus:

$$\frac{1}{2}|\nabla\phi+\boldsymbol{P}|^2+\Xi=E \tag{1.13}$$

which is independent of ρ and so is decoupled from (1.9).

Suppose now there exists a maximizer ρ_0 of (1.12). Multiply (1.13) by ρ_0 and integrate over $\mathbb{T}^n$ to obtain

$$E=F(\rho_0,\boldsymbol{P})+\int_{\mathbb{T}^n}\Xi\rho_0:=\widehat{H}_\Xi(\boldsymbol{P}). \tag{1.14}$$

In particular, the Lagrange multiplier E is identical to the asymptotic energy spectrum $\widehat{H}_\Xi(\boldsymbol{P})$. An important point to be noted, at this stage, is that the asymptotic spectrum is in the oscillatory domain of the periodic Schrödinger equation, that is $\widehat{H}_\Xi(\boldsymbol{P})\geqslant\max_{\mathbb{T}^n}\Xi$ necessarily holds for any $\boldsymbol{P}\in\mathbb{R}^n$, and $\widehat{H}_\Xi(\boldsymbol{P})>\max_{\mathbb{T}^n}\Xi$ if $|\boldsymbol{P}|$ is sufficiently large (see Proposition 4.1).

The function $\widehat{H}_\Xi=\widehat{H}_\Xi(\boldsymbol{P})$ defined in (1.14) is considered by Evans and Gomes [5,6] as the *Effective Hamiltonian* corresponding to

$$h_\Xi(p,x)=|p|^2/2+\Xi(x).$$

If $\psi(x,\boldsymbol{P}):=\phi(x)$ is a solution of the Hamilton–Jacobi equation (1.13) corresponding to a given $\boldsymbol{P}$, then a canonical change of variables

$$\boldsymbol{p}=\nabla_x\psi+\boldsymbol{P};\qquad \vec{X}=\nabla_{\boldsymbol{P}}\psi+x$$

reduces the Hamiltonian equation to an integrable system defined by the Hamiltonian $\widehat{H}_\Xi$, that is, $\boldsymbol{P}$ is a cyclic variable and hence a constant of motion.

In general, such a solution does not exist for any $\boldsymbol{P} \in \mathbb{R}^n$. However, (1.12) suggests a way to define the effective Hamiltonian $\widehat{H}_\Xi$ *without* the assumption that (1.13) is solvable. We note, at this stage, that (1.12) seems to be the dual of

$$\widehat{H}_\Xi(\boldsymbol{P}) = \inf_{\phi \in C^\infty(\mathbb{T}^n)} \sup_{x \in \mathbb{T}^n} h_\Xi(\nabla\phi + \boldsymbol{P}, x) \tag{1.15}$$

which was suggested by Gomes and Oberman [9] as a numerical tool for evaluating $\widehat{H}_\Xi$.

1.3. Objectives

As we shall see below, the supremum in (1.12) is not attained in $\mathbb{L}^1(\mathbb{T}^n)$, in general, but in the set $\overline{\mathcal{M}}$ of Borel probability measures on $\mathbb{T}^n$. This, together with (1.11), motivates us to extend the domain of F in (1.6) from the set of nonnegative densities in $\mathbb{L}^1(\mathbb{T}^n)$ to $\overline{\mathcal{M}}$. Similarly, the functional H_Ξ (1.11) is extended to $\overline{\mathcal{M}}$ as well. Our first object is

I. *to define a generalized minimizer ϕ of F.*

The effective Hamiltonian plays a major rule in the *weak KAM theory*. See [7,10,17,19,22] among other references. For the convenience of the reader we review the fundamentals of the weak KAM theory and Mather measures in Section 2. Our second object is

II. *to relate the functional H_Ξ to the weak KAM theory. In particular, to relate the generalized minimizer ϕ of F to the minimal Mather measure.*

An excellent reference to the Monge–Kantorovich theory of optimal transportation is the book of Villani [23]. The relation between M-K theory and the weak KAM theory was suggested in [4] and further elaborated in a series of publications, among which [11,16,3]. Essentially, it relates the minimal (Mather) measures of a given Lagrangian to a measure which minimizes a certain optimal transportation plan. The third object of this paper is

III. *to approximate F and H_Ξ by an optimal transportation functions F_T and $H_{\Xi,T}$, respectively.*

Finally, we use the suggested functionals to establish an alternative to (1.15) for the evaluation of $\widehat{H}_\Xi(\boldsymbol{P})$:

IV. *Establish a combinatorial search algorithm for evaluation of $\widehat{H}_\Xi(\boldsymbol{P})$ to any degree of approximation.*

2. Lagrangian dynamics on the torus

2.1. The Aubrey–Mather theory and minimal orbits

Let

$$L(\boldsymbol{p}, x) := \frac{|\boldsymbol{p}|^2}{2} - \Xi(x), \tag{2.1}$$

a Lagrangian function defined on $\mathbb{R}^n \times \mathbb{R}^n$ where the potential Ξ which is 1-periodic in all the variables $x = (x_1, \ldots, x_n)$. For a given orbit $(x(t), \boldsymbol{p}(t))$ of the associated Euler–Lagrange equation

$$\dot{x} = \boldsymbol{p}; \quad \dot{\boldsymbol{p}} + \nabla_x \Xi = 0, \tag{2.2}$$

a *rotation vector* $\boldsymbol{J} \in \mathbb{R}^n$ is assigned to this orbit provided the limit

$$\boldsymbol{J} = \lim_{|t| \to \infty} t^{-1} x(t) \tag{2.3}$$

exists. As a trivial example, consider the x-independent Lagrangian where $\Xi \equiv 0$. Since $\boldsymbol{p}$ is a constant of motion and $\dot{x} = \boldsymbol{p}$, the rotation vector is defined for each orbit via $\boldsymbol{J} = \boldsymbol{p}$. For general Ξ the rotation vector is *not* defined for any orbit, in general. The object of the *classical* KAM theory (see, e.g. [10]) is the study of small perturbation of an integrable system, e.g. for Lagrangians of the form (2.1) where the potential Ξ is small. In particular, it studies families of solutions of such systems which preserve the rotation vector.

In the eighties, Aubry [2] and Mather [18] (see also [20]) discovered that Lagrangian flows which induce a monotone, symplectic twist maps on a two-dimensional annulus, possess orbits of any given rotation number (in the twist interval), even if the corresponding Lagrangian is not close to an integrable one. The characterization of these orbits is variational: They are minimizers of the Lagrangian action with respect to any local variation of the orbit. In general, they are embedded in invariant tori of the Lagrangian flow.

There is still another approach to invariant tori of Lagrangian/Hamiltonian systems. An invariant Lagrangian torus can be obtained as a solution of the corresponding Hamilton–Jacobi equation as follows: Suppose there exists, for some $\boldsymbol{P} \in \mathbb{R}^n$, solution $\phi \in C^{1,1}(\mathbb{R}^n)$ which is 1-periodic in each of the coordinates x_j of $x = (x_1, \ldots, x_n)$, for

$$\frac{1}{2}\left|\nabla\phi(x) + \boldsymbol{P}\right|^2 + \Xi(x) = E, \quad E \in \mathbb{R}.$$

Then the graph of the function $(x, \boldsymbol{P} + \nabla\phi(x))$ represents an *invariant torus* of the Lagrangian flow associated with L [17]. The projection on x of any orbit in this invariant set is obtained by a solution of the system

$$\dot{x} = \boldsymbol{P} + \nabla\phi(x). \tag{2.4}$$

In the case $n = 2$ the rotation vector $\boldsymbol{J} \in \mathbb{R}^2$ is defined as in (2.3) for any such orbit, given by (2.4).

2.2. Weak KAM and minimal invariant measures

For dimension higher than 2, there are counter-examples: There exists a Lagrangian system on the 3-dimensional torus, induced by a metric on this torus, for which there are no minimal geodesics, save for a finite number of rotation vectors [12]. Moreover, it is not known that the limit (2.3) exists for any orbit of (2.4), if $n > 2$. Hence, an extension of Aubry–Mather theory to higher dimensions is not a direct one. If, however, we replace the notion of an orbit by an invariant measure, then it is possible to extend the Aubry–Mather theory to higher dimensions. The relaxation of orbits to invariant measures (and the corresponding minimal orbits to minimal invariant measures) leads to the "weak KAM theory".

Let $\mathcal{M}_L$ be the set of all probability measures on the tangent bundle $\mathbb{T}^n \times \mathbb{R}^n$ which are invariant with respect to the flow induced by the Lagrangian L. The *rotation vector* $\boldsymbol{\alpha} : \mathcal{M}_L \to \mathbb{R}^n$ is

$$\boldsymbol{\alpha}(\nu) := \int_{\mathbb{T}^n \times \mathbb{R}^n} \boldsymbol{p}\, d\nu(x, \boldsymbol{p}), \tag{2.5}$$

and, for any $\alpha \in \mathbb{R}^n$, the set of all $\mathcal{M}_L^{\boldsymbol{J}} \subset \mathcal{M}_L$ corresponds to all $\nu \in \mathcal{M}_L$ for which $\boldsymbol{\alpha}(\nu) = \boldsymbol{J}$.

A minimal measure associated with a rotation vector α is defined by

$$\nu_{\boldsymbol{J}} = \arg\min_{\nu \in \mathcal{M}_L^{\boldsymbol{J}}} \int_{\mathbb{T}^n \times \mathbb{R}^n} L\, d\nu \in \mathcal{M}_L^{\boldsymbol{J}}. \tag{2.6}$$

Its dual representation is given by minimizing the Lagrangian $L^{\boldsymbol{P}} := L(\boldsymbol{p}, x) - \boldsymbol{P} \cdot \boldsymbol{p}$ over the *whole set* of invariant measures[1] $\mathcal{M}_L$. The measure $\nu_{\boldsymbol{P}} \in \mathcal{M}_L$ is called a *Mather measure* if

$$\nu_{\boldsymbol{P}} = \arg\min_{\nu \in \mathcal{M}_L} \int_{\mathbb{T}^n \times \mathbb{R}^n} L^{\boldsymbol{P}}\, d\nu \in \mathcal{M}_L. \tag{2.7}$$

These minimal measures are relaxations of minimal invariant orbits of the Aubry–Mather theory. Their properties and the geometry of their supports are the fundamental ingredients of the developing weak KAM theory. For further details, see [19,17,7] or consult [22] for applications and further references.

[1] Note that ν is an invariant measure of L if and only if it is an invariant measure of $L_{\boldsymbol{P}}$.

It should be stressed, however, that the investigation of the functional F (3.1) carried in the present paper is not restricted to minimal (Mather) measures. In fact, Mather measures (and their $\mathbb{T}^n$ projections) are defined only for smooth enough Lagrangian systems which allow the existence of dynamics, e.g. (2.1) where $\Xi \in C^{1,1}(\mathbb{T}^n)$. Since we are motivated, between other things, by quantum dynamics and wave equation, we must assume much less, e.g. the Schrödinger equation (1.3) is well posed if the potential Ξ is only continuous.

3. An overview of the main results

3.1. List of symbols and definitions

(1) $\mathbb{T}^n = \mathbb{R}^n/\mathbb{Z}^n$ the n-dimensional flat torus. It is parameterized by $x = (x_1, \ldots, x_n) \bmod \mathbb{Z}^n$. The Euclidian distance $\|x - y\|_{\mathbb{T}^n}$ on $\mathbb{T}^n$ is defined as $\min_{z\in\mathbb{Z}^n} |x - y - z|$, where $x, y \in \mathbb{R}^n$ and $|\cdot|$ is the Euclidian norm on $\mathbb{R}^n$.
(2) $\mathbb{T}^n \times \mathbb{R}^n$ is the tangent bundle of $\mathbb{T}^n$, that is, $\mathbb{T}^n \times \mathbb{R}^n := \mathbb{R}^n \times \mathbb{T}^n$. A vector in $\mathbb{R}^n$ is denoted by a bold letter, e.g. $\boldsymbol{v}$. The same symbol will also define a vector field, that is, a section in $\mathbb{T}^n \times \mathbb{R}^n$, e.g. $\boldsymbol{v} = \boldsymbol{v}(x)$.
(3) $\mathcal{M}(D)$ stands for the set of all probability normalized Borel measures μ on some metric space D, subjected to the dual topology of $C(D)$: $|\mu| := \sup_{\phi\in C(D),|\phi|_\infty=1} \int_D \phi\, d\mu$. We denote $\overline{\mathcal{M}} := \mathcal{M}(\mathbb{T}^n)$ as the set of all such measures on the torus $\mathbb{T}^n$.
(4) A Borel map $\mathbf{S} : D_1 \to D_2$ induces a map $\mathbf{S}_\# : \mathcal{M}(D_1) \to \mathcal{M}(D_2)$, as follows: $\mathbf{S}_\#\mu(A) := \mu(\mathbf{S}^{-1}(A))$ for any Borel set $A \in D_2$. $\mathbf{S}_\#\mu$ is called the *push-forward* of $\mu \in \mathcal{M}(D_1)$ into $\mathcal{M}(D_2)$.
(5) $\pi : \mathbb{T}^n \times \mathbb{R}^n \to \mathbb{T}^n$ is the projection (natural embedding) of $\mathbb{T}^n$ in $\mathbb{T}^n \times \mathbb{R}^n$, namely $\pi(x, \boldsymbol{p}) = x$ for $(x, \boldsymbol{p}) \in \mathbb{T}^n \times \mathbb{R}^n$. In particular, $\pi_\# : \mathcal{M}(\mathbb{T}^n \times \mathbb{R}^n) \to \overline{\mathcal{M}}$ so $\mu = \pi_\#\nu \in \overline{\mathcal{M}}$ is the $\mathbb{T}^n$ marginal of $\nu \in \mathcal{M}(\mathbb{T}^n \times \mathbb{R}^n)$.
(6) $\overline{\mathcal{M}}_T := C([0, T], \overline{\mathcal{M}})$. An element $\hat{\mu} \in \overline{\mathcal{M}}_T$ is denoted by $\hat{\mu} := \mu_{(t)}$, $0 \leqslant t \leqslant T$. For any $\mu_1, \mu_2 \in \overline{\mathcal{M}}$, the set $\overline{\mathcal{M}}_T(\mu_1, \mu_2) \subset \overline{\mathcal{M}}$ is defined as all $\hat{\mu} \in \overline{\mathcal{M}}_T$ for which $\mu_{(0)} = \mu_1, \mu_{(T)} = \mu_2$. If $\mu_1 = \mu_2 \equiv \mu$ we denote $\overline{\mathcal{M}}_T(\mu) := \overline{\mathcal{M}}_T(\mu, \mu)$.
(7) $\overline{\mathcal{M}}^{(2)} := \mathcal{M}(\mathbb{T}^n \times \mathbb{T}^n)$ and $\Pi_i : \mathbb{T}^n \times \mathbb{T}^n \to \mathbb{T}^n$ the projection on the i factor, $i = 1, 2$. For $\mu_1 \in \overline{\mathcal{M}}$, $\mu_2 \in \overline{\mathcal{M}}$ define $\overline{\mathcal{M}}^{(2)}(\mu_1, \mu_2)$ as the set of all $\sigma \in \overline{\mathcal{M}}^{(2)}$ for which $\Pi_{i,\#}\sigma = \mu_i$, $i = 1, 2$. $\overline{\mathcal{M}}^{(2)}(\mu) := \overline{\mathcal{M}}^{(2)}(\mu, \mu)$.
(8) $C^1(\mathbb{T}^n)$ is the set of all C^1 smooth functions on $\mathbb{T}^n$.
(9) Recall the definition of a subgradient of a function $h : B \to \mathbb{R}$ defined on a Banach space B: For $b \in B$,

$$\partial_b h := \big\{b^* \in B^*,\ h(b') \geqslant h(b) + \langle b' - b, b^*\rangle \text{ for any } b' \in B\big\}.$$

3.2. F and its generalized minimizers

There is a close relation between the minimal (Mather) measures described in Section 2 and the minimizer of the function F defined in (1.6), where ρ is the density of the $\mathbb{T}^n$ marginal of a minimal measure.

In general, however, there are no smooth densities to the marginals of minimal measures on $\mathbb{T}^n$. Motivated by this, we extend the definition of F to the set $\overline{\mathcal{M}}$ of *all* probability Borel measures on $\mathbb{T}^n$:

$$F(\mu, \boldsymbol{P}) := \frac{1}{2} \inf_{\phi\in C^1(\mathbb{T}^n)} \int_{\mathbb{T}^n} |\nabla\phi + \boldsymbol{P}|^2\, d\mu \tag{3.1}$$

and

$$F^*(\mu, \boldsymbol{J}) := \sup_{\boldsymbol{P}\in\mathbb{R}^n} \big[\boldsymbol{P}\cdot\boldsymbol{J} - F(\mu, \boldsymbol{P})\big] \tag{3.2}$$

its convex dual on $\mathbb{R}^n$.

The first question we address is the existence of minimizers of (3.1) for a general measure $\mu \in \overline{\mathcal{M}}$. Evidently, there is no sense of solutions to the elliptic problem (1.9) for such μ. Our first result, given in Theorem 4.1, indicates the existence and uniqueness of a minimizer in a generalized sense (see Definition 4.2 below). In Theorem 4.2 we discuss the relation between these minimizers and the solutions of the elliptic problem (1.9).

3.3. The effective Hamiltonian

The second question concerns the maximizers of (1.14), extended to the entire set $\overline{\mathcal{M}}$. Let

$$\widehat{H}_{\Xi}(\boldsymbol{P}) := \sup_{\mu\in\overline{\mathcal{M}}}\left[F(\mu,\boldsymbol{P}) + \int_{\mathbb{T}^n} \Xi\, d\mu\right], \tag{3.3}$$

where $\Xi \in C(\mathbb{T}^n)$. Theorem 4.3 relates the generalized minimizers of Theorem 4.1 to the minimal (Mather) measure associated with the Lagrangian (2.1) corresponding to $\boldsymbol{P}$ via (2.7). It claims that, if $\Xi \in C^{1,1}(\mathbb{T}^n)$, then the generalized solution of $F(\mu,\boldsymbol{P})$ corresponding to μ which maximizes (3.3) is, indeed, a Mather measure associated with the Lagrangian (2.1). In this sense, the weak minimizers of Theorem 4.1 can be considered as generalized Mather measures for Lagrangians with only continuous potentials.

3.4. On the continuity of F

The third question addressed is the continuity property of F with respect to μ. It is rather easy to observe that F is convex in $\boldsymbol{P}$ on $\mathbb{R}^n$, and concave in μ on $\overline{\mathcal{M}}$. These imply that F is continuous on $\mathbb{R}^n$, but only *upper-semi-continuous* in the natural topology of $\overline{\mathcal{M}}$, which is the weak-$*$ topology induced by $C^*(\mathbb{T}^n)$. That is, if $\mu_j \to \mu$ in $C^*(\mathbb{T}^n)$, then

$$\lim_{j\to\infty} F(\mu_j,\boldsymbol{P}) \leqslant F(\mu,\boldsymbol{P}) \tag{3.4}$$

holds. In general, there is *no continuity* of F over $\overline{\mathcal{M}}$ with the C^* topology.

Examples.

(1) For any atomic measure $\mu = \sum m_i \delta_{x_i} \in \overline{\mathcal{M}}$, we can easily verify that $F(\mu,\boldsymbol{P}) \equiv 0$ for any $\boldsymbol{P} \in \mathbb{R}^n$. In particular, if μ_N is a sequence of *empirical* measures: $\mu_N := N^{-1}\sum_{i=1}^N \delta_{x_i}$ satisfying $\mu_N \to \mu \in \overline{\mathcal{M}}$ in $C^*(\mathbb{T}^n)$, then the inequality in (3.4) is strict whenever $F(\mu,\boldsymbol{P}) > 0$.

(2) Let $n=1$, so $\mathbb{T}^n$ is reduced to the circle $\mathbb{S}^1$. Suppose $\mu \in \overline{\mathcal{M}}(\mathbb{S}^1)$ admits a smooth density $\mu(dx) = \rho(x)\,dx$. If $\rho > 0$ on $\mathbb{S}^1$ then the continuity equation (1.9) reduces to a constant $j = \rho(\phi_x + P)$. This implies

$$F(\mu,P) = \frac{1}{2}\int_{\mathbb{S}^1} \rho|\phi_x + P|^2\,dx = \frac{j^2}{2}\int_{\mathbb{S}^1}\rho^{-1}\,dx \tag{3.5}$$

as well as

$$\int_{\mathbb{S}^1}\rho^{-1}\,dx = \int_{\mathbb{S}^1} j^{-1}(\phi_x + P)\,dx = \frac{P}{j} \quad\Longrightarrow\quad j = P\left(\int_{\mathbb{S}^1}\rho^{-1}\,dx\right)^{-1}.$$

Substitute in (3.5) to obtain

$$F(\mu,P) = \frac{|P|^2}{2}\left(\int_{\mathbb{S}^1}\rho^{-1}\,dx\right)^{-1}.$$

In particular, $\int_{\mathbb{S}^1}\rho^{-1}\,dx = \infty$ iff $F(\mu,P) = 0$ for $P \neq 0$. Any sequence $\mu_j(dx) = \rho_j(x)\,dx$ satisfying $\int_{\mathbb{S}^1}\rho_j^{-1} = \infty$ which converges in $C^*(\mathbb{S}^1)$ to $\mu(dx) = \rho(x)\,dx$ satisfying $\rho \in C^1(\mathbb{S}^1)$, $\rho > 0$ on $\mathbb{S}^1$, is an example of strict inequality in (3.4).

3.5. Lagrangian mappings

In Theorem 4.4 we show that F can be approximated, as a function on $\overline{\mathcal{M}}$, by a *weakly continuous* function $F_T(\cdot, \boldsymbol{P})$ which satisfies $F_T(\mu, \boldsymbol{P}) \to F(\mu, \boldsymbol{P})$ as $T \to 0$, for any $\mu \in \overline{\mathcal{M}}$. For this, we represent an extension of F to *orbits* $\hat{\mu} : [0, T] \to \overline{\mathcal{M}}$, $\hat{\mu}|_{(t)} = \mu_{(t)} \in \overline{\mathcal{M}}$, $t \in [0, T]$, given by

$$F(\hat{\mu}, \boldsymbol{P}, T) := \frac{1}{T} \inf_{\phi \in C_0^1(\mathbb{T}^n \times (0,T))} \int_0^T \int_{\mathbb{T}^n} \left[\frac{\partial \phi}{\partial t} + \frac{1}{2} |\nabla_x \phi + \boldsymbol{P}|^2 \right] d\mu_{(t)}\, dt,$$

and set $F_T(\mu, \boldsymbol{P})$ as the *supremum* of $F(\hat{\mu}, \boldsymbol{P}, T)$ over all such orbits satisfying $\mu_{(0)} = \mu_{(T)} = \mu$. It is shown that $F_T(\mu, \boldsymbol{P}) = |\boldsymbol{P}|^2/2 - D_{T\boldsymbol{P}}(\mu)/(2T^2)$ where $D_{\boldsymbol{P}}(\mu)$ is defined by the optimal Monge–Kantorovich transport plant from μ to itself, subjected to the cost function $c(x, y) := \|x - y - \boldsymbol{P}\|^2_{\mathbb{T}^n}$, where $\|\cdot\|_{\mathbb{T}^n}$ is the Euclidian metric on $\mathbb{T}^n$. As an example, consider the case $\mu = \delta_{x_0}$ for some $x_0 \in \mathbb{T}^n$. Then $D_{\boldsymbol{P}}(\delta_{x_0}) = |\{\boldsymbol{P}\}|^2$, where $\{\cdot\}$ stands for the fractional part $\{\boldsymbol{P}\} := \boldsymbol{P} \bmod \mathbb{Z}^n$. So

$$F_T(\delta_{x_0}, \boldsymbol{P}) = |\boldsymbol{P}|^2/2 - \left|\{\boldsymbol{P}T\}\right|^2/\left(2T^2\right).$$

If T is sufficiently small so $\{T\boldsymbol{P}\} = T\boldsymbol{P}$ then $F_T(\delta_{x_0}, \boldsymbol{P}) = 0$.

3.6. Combinatorial search for the minimal measure

Our last object is to suggest an alternative to the numerical algorithm for the calculation of the effective Hamiltonian based on (1.15), introduced in [9]. We take an advantage of the following facts

(i) An optimal transportation functional (such as $F_T(\mu, \boldsymbol{P})$) are continuous in the weak topology of $\overline{\mathcal{M}}$.
(ii) The set of *empirical measures* is dense in the set of all measures $\overline{\mathcal{M}}$.
(iii) On the set of empirical measures of a fixed number of sampling points j, an optimal transportation problem is reduced to a finite combinatorial problem on the set of permutation on $\{1, \ldots, j\}$ (Birkhoff's theorem).

Applying (i)–(iii) to the result of Theorem 4.4, we obtain a discrete, combinatorial algorithm for evaluating the effective Hamiltonian $\widehat{H}_{\Xi}(\boldsymbol{P})$. This is summarized in Theorem 4.6.

4. Detailed description of the main results

4.1. Minimizers of the Dirichlet functional over the n-torus

Let us recall the definition, for $\boldsymbol{P} \in \mathbb{R}^n$, $\boldsymbol{J} \in \mathbb{R}^n$ and $\mu \in \overline{\mathcal{M}}$:

$$F(\mu, \boldsymbol{P}) := \frac{1}{2} \inf_{\phi \in C^1(\mathbb{T}^n)} \int_{\mathbb{T}^n} |\nabla\phi + \boldsymbol{P}|^2\, d\mu, \tag{4.1}$$

$$F^*(\mu, \boldsymbol{J}) := \sup_{\boldsymbol{P} \in \mathbb{R}^n} \left[\boldsymbol{P} \cdot \boldsymbol{J} - F(\mu, \boldsymbol{P})\right]. \tag{4.2}$$

Let also $\mathcal{E}$ a function on $\overline{\mathcal{M}} \times \mathbb{R}^n \times C^1(\mathbb{T}^n)$ defined as

$$\mathcal{E}(\mu, \boldsymbol{J}, \phi) := \frac{1}{2} \left\{ \left| \boldsymbol{J} - \int_{\mathbb{T}^n} \nabla\phi\, d\mu \right|^2 - \int_{\mathbb{T}^n} |\nabla\phi|^2\, d\mu \right\}. \tag{4.3}$$

Next, we consider the notion of *weak solution* of (1.9), corresponding to the minimizer of (4.1):

Definition 4.1.

(1) The set $\Lambda \subset \mathcal{M}(\mathbb{T}^n \times \mathbb{R}^n)$ consists of all probability measures $\nu(dx\,d\boldsymbol{p})$ for which

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \boldsymbol{p} \cdot \nabla\theta(x)\,d\nu = 0 \quad \forall \theta \in C^1(\mathbb{T}^n) \quad \text{and} \quad \int_{\mathbb{T}^n \times \mathbb{R}^n} |\boldsymbol{p}|^2\,d\nu < \infty.$$

(2) Given $\mu \in \overline{\mathcal{M}}$, the set $\Lambda_\mu \subset \Lambda$ of liftings of μ is composed of all $\nu \in \Lambda$ for which $\pi_\#\nu = \mu$.
(3) For each $\boldsymbol{J} \in \mathbb{R}^n$, the set $\Lambda_\mu^{\boldsymbol{J}}$ is defined as all $\nu \in \Lambda_\mu$ which satisfies

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \boldsymbol{p}\,d\nu = \boldsymbol{J}.$$

Remark 4.1. The set Λ_μ is never empty. Indeed, $\nu = \delta_0^{\boldsymbol{p}} \otimes \mu \in \Lambda_\mu$ for any $\mu \in \overline{\mathcal{M}}$. However, the set $\Lambda_\mu^{\boldsymbol{J}}$ can be empty. For example, if $F(\mu, \boldsymbol{P}) = 0$ for all $\boldsymbol{P} \in \mathbb{R}^n$ then $\Lambda_\mu^{\boldsymbol{J}} = \emptyset$ for any $\boldsymbol{J} \neq \boldsymbol{0}$.

Definition 4.2. For given $\boldsymbol{P} \in \mathbb{R}^n$, $\nu \in \Lambda_\mu$ is a weak solution of $F(\mu, \boldsymbol{P})$ provided

$$\int_{\mathbb{T}^n \times \mathbb{R}^n} \left[\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p} \cdot \boldsymbol{P}\right] d\nu \leqslant \int_{\mathbb{T}^n \times \mathbb{R}^n} \left[\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p} \cdot \boldsymbol{P}\right] d\xi \quad \forall \xi \in \Lambda_\mu.$$

The existence and uniqueness of weak solution is described in Theorem 4.1 below.

Theorem 4.1. *For any $\mu \in \overline{\mathcal{M}}$ and $\boldsymbol{P} \in \mathbb{R}^n$, there exists a unique weak solution $\nu \in \Lambda_\mu$ of $F(\mu, \boldsymbol{P})$. Moreover,*

$$F(\mu, \boldsymbol{P}) = -\int_{\mathbb{T}^n \times \mathbb{R}^n} \left[\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p} \cdot \boldsymbol{P}\right] d\nu, \tag{4.4}$$

and

$$F^*(\mu, \boldsymbol{J}) = \frac{1}{2} \inf_{\nu \in \Lambda_\mu^{\boldsymbol{J}}} \int_{\mathbb{T}^n \times \mathbb{R}^n} |\boldsymbol{p}|^2\,d\nu, \tag{4.5}$$

where the RHS of (4.5) *is attained for the weak solution of $F(\mu, \boldsymbol{P})$, provided*[2] $\boldsymbol{J} \in \partial_{\boldsymbol{P}} F(\mu, \boldsymbol{P})$.

If $\mu(dx) = \rho(x)\,dx$ where $\rho \in C^1(\mathbb{T}^n)$ and $\rho > 0$ on $\mathbb{T}^n$, then a weak solution ν of $F(\mu, \boldsymbol{P})$ takes the form $\nu(dx\,d\boldsymbol{p}) = \delta^{\boldsymbol{p}}_{\boldsymbol{P}+\nabla\phi(x)} \otimes \rho(x)\,dx$ where ϕ is the classical solution of the elliptic equation

$$\nabla\big[\rho(\nabla_x\phi + \boldsymbol{P})\big] = 0 \quad \text{on } \mathbb{T}^n. \tag{4.6}$$

Remark 4.2. Eq. (4.6) is strongly elliptic equation if $\rho > 0$, so it has a unique (up to a constant), classical solution. See, e.g. [8].

Remark 4.3. As a by-product we obtain the relation

$$F^*(\mu, \boldsymbol{J}) = \inf_{\phi \in C^1(\mathbb{T}^n)} \mathcal{E}(\mu, \boldsymbol{J}, \phi),$$

see Lemma 5.1(3).

Example 1. If $\mu(dx) = \alpha\rho(x)\,dx + (1-\alpha)\delta_{x_0}$ then the weak solution associated with $\boldsymbol{P} \in \mathbb{R}^n$ is

$$\nu = \delta^{\boldsymbol{p}}_{\boldsymbol{P}+\nabla\phi} \otimes \alpha\rho(x)\,dx + (1-\alpha)\delta_0^{\boldsymbol{p}} \otimes \delta_{x_0},$$

where ϕ is the classical solution of (4.6).

[2] The existence of a subgradient of F with respect to $\boldsymbol{P}$, among other results, is stated and proved in Lemma 5.1, Section 5.2.

We may observe that F is concave on $\overline{\mathcal{M}}$ for fixed $\boldsymbol{P} \in \mathbb{R}^n$. In particular, it is *upper-semi-continuous* in the C^* topology of $\overline{\mathcal{M}}$:

$$\lim_{n\to\infty} F(\mu_n, \boldsymbol{P}) \leqslant F(\mu, \boldsymbol{P}) \tag{4.7}$$

whenever $\mu_n \to \mu$ in $C^*(\mathbb{T}^n)$.

Example 2. If μ_n is an atomic measure then $F(\mu_n, \boldsymbol{P}) = 0$ for any $\boldsymbol{P} \in \mathbb{R}^n$. In particular, if $\mu_n = n^{-1}\sum_{j=1}^n \delta_{x_j^{(n)}}$ is a sequence of empirical measures approximating $\mu \in \overline{\mathcal{M}}$ then the LHS of (4.7) is identically zero.

In Theorem 4.2 and Corollary 4.1 we demonstrate that, in an appropriate sense, any weak solution is a limit of classical ones.

Theorem 4.2. *If* $\lim \mu_j = \mu$ *in the* $C^*(\mathbb{T}^n)$ *and*

$$\lim_{j\to\infty} F(\mu_j, \boldsymbol{P}) = F(\mu, \boldsymbol{P}) \tag{4.8}$$

holds, then there exists a subsequence of weak solutions ν_j *of* $F(\mu_j, \boldsymbol{P})$ *along which*

$$\lim_{j\to\infty} \nu_j = \nu \tag{4.9}$$

holds in $C^*(\mathbb{T}^n \times \mathbb{R}^n)$*, where* ν *is a weak solution of* $F(\mu, \boldsymbol{P})$.

Moreover, there exists a sequence of smooth measures $\mu_j = \rho_j\, dx$ *so that* $\rho_j \in C^\infty(\mathbb{T}^n)$ *and* $\rho_j > 0$ *on* $\mathbb{T}^n$*, for which* (4.8) *holds for any* $\boldsymbol{P} \in \mathbb{R}^n$.

Corollary 4.1. *The weak solution of* $F(\mu, \boldsymbol{P})$ *is the weak limit*

$$\lim_{j\to\infty} \delta_{\boldsymbol{p}-\boldsymbol{P}-\nabla\phi_j}\, d\boldsymbol{p} \otimes \rho_j\, dx = \nu$$

where ϕ_j *are the solutions of*

$$\nabla \cdot \big(\rho_j(\nabla\phi_j + \boldsymbol{P})\big) = 0$$

and $\mu_j = \rho_j dx \to \mu$ *as guaranteed by Theorem* 4.2.

Definition 4.3. Given a continuous function $\Xi \in C(\mathbb{T}^n)$,

$$\widehat{H}^*_\Xi(\boldsymbol{J}) := \sup_{\mu\in\overline{\mathcal{M}}} \int \Xi\, d\mu - F^*(\mu, \boldsymbol{J}). \tag{4.10}$$

Likewise

$$\widehat{H}_\Xi(\boldsymbol{P}) := \sup_{\mu\in\overline{\mathcal{M}}} \int \Xi\, d\mu + F(\mu, \boldsymbol{P}). \tag{4.11}$$

Lemma 4.1. $\widehat{H}^*_\Xi$ *is the negative of the convex dual of* $\widehat{H}_\Xi$ *with respect to* $\mathbb{R}^n$*. Than is*: $\widehat{H}^*_\Xi(\boldsymbol{J}) = -\sup_{\boldsymbol{P}\in\mathbb{R}^n}\{\boldsymbol{P}\cdot\boldsymbol{J} - \widehat{H}_\Xi(\boldsymbol{P})\}$.

Proposition 4.1. $\widehat{H}_\Xi(\boldsymbol{P}) \geqslant \max_{\mathbb{T}^n} \Xi$ *and* $\widehat{H}^*_\Xi(\boldsymbol{J}) \geqslant \max_{\mathbb{T}^n} \Xi$ *hold for any* $\boldsymbol{P}, \boldsymbol{J} \in \mathbb{R}^n$*. If* $|\boldsymbol{P}|$ *(resp.* $|\boldsymbol{J}|$*) is large enough, then the inequality is strong.*

Open problem. Is $\widehat{H}_\Xi(\boldsymbol{P}) > \max_{\mathbb{T}^n} \Xi$ for any $\boldsymbol{P} \neq 0$?

Lemma 4.2. *For any* $\Xi \in C(\mathbb{T}^n)$ $\boldsymbol{P} \in \mathbb{R}^n$ *there exists* $\mu_0 \in \overline{\mathcal{M}}$ *verifying the maximum in* (4.11)*. There exists* $\boldsymbol{J} \in \partial_{\boldsymbol{P}} F(\mu_0, \Xi) \subset \partial_{\boldsymbol{P}} \widehat{H}_\Xi(\cdot)$ *for which* μ_0 *verifies the maximum in* (4.10).

We end this section by stating the connection between maximizers of $\widehat{H}_\Xi$ and $\widehat{H}^*_\Xi$ and the minimal invariant measures of the weak-KAM theory:

Theorem 4.3. *If Ξ is smooth enough (say, $\Xi \in C^2(\mathbb{T}^n)$) and ν is a Mather measure (2.7) of the Lagrangian $L = |\boldsymbol{p}|^2/2 - \Xi(x) - \boldsymbol{p}\cdot\boldsymbol{P}$ on $\mathbb{T}^n \times \mathbb{R}^n$ and $\mu = \pi_\#\nu$ then ν is weak solution of $F(\mu, \boldsymbol{P})$, and is a maximizer of $\widehat{H}_\Xi(\boldsymbol{P})$ in (4.11).*

Moreover, ν verifies (2.6) where $\boldsymbol{J}$ is the rotation number $\boldsymbol{\alpha}(\nu)$ given by (2.5).

4.2. Extension to time dependent measures

We now extend the definition of F and F^* to the set of $\overline{\mathcal{M}}$-valued orbits on the interval $[0,T]$.

Define, for[3] $\hat{\mu} \in \overline{\mathcal{M}}_T$

$$F(\hat{\mu}, \boldsymbol{P}, T) := \frac{1}{T} \inf_{\phi \in C^1_0(\mathbb{T}^n \times (0,T))} \int_0^T \int_{\mathbb{T}^n} \left[\frac{\partial \phi}{\partial t} + \frac{1}{2}|\nabla_x \phi + \boldsymbol{P}|^2\right] d\mu_{(t)}\, dt. \tag{4.12}$$

Let also $\mathcal{E}$ a function on $\overline{\mathcal{M}}_T \times \mathbb{R}^n \times C^1_0(\mathbb{T}^n \times (0,T))$ defined as:

$$\mathcal{E}(\hat{\mu}, \boldsymbol{J}, \phi, T) := \left\{\frac{1}{2}\left|\boldsymbol{J} - \frac{1}{T}\int_0^T \int_{\mathbb{T}^n} \nabla\phi\, d\mu_{(t)}\, dt\right|^2 - \frac{1}{T}\int_0^T \int_{\mathbb{T}^n} \left(\phi_t + \frac{1}{2}|\nabla\phi|^2\right) d\mu_{(t)}\, dt\right\}. \tag{4.13}$$

The analog of Definition 4.1 is

Definition 4.4.

(1) The set $\hat{\Lambda}_T \subset \mathcal{M}_T$ consists of all orbits of probability measures $\hat{\nu} : [0,T] \to \mathcal{M}(\mathbb{T}^n \times \mathbb{R}^n)$, $\hat{\nu}|_{(t)} := \nu_{(t)} \in \mathcal{M}(\mathbb{T}^n \times \mathbb{R}^n)$, for which

$$\int_0^T \int_{\mathbb{T}^n \times \mathbb{R}^n} \big(\theta_t + \boldsymbol{p}\cdot \nabla_x \theta(x,t)\big)\, d\nu_{(t)}\, dt = 0 \quad \forall \theta \in C^1_0\big(\mathbb{T}^n \times (0,T)\big) \quad \text{and} \quad \int_0^T \int_{\mathbb{T}^n \times \mathbb{R}} |\boldsymbol{p}|^2\, d\nu_{(t)}\, dt < \infty.$$

(2) Given $\hat{\mu} \in \overline{\mathcal{M}}_T$, the set $\widehat{\Lambda}_{T,\hat{\mu}} \subset \Lambda$ of liftings of $\hat{\mu}$ is composed of all $\hat{\nu} \in \hat{\Lambda}_T$ for which $\pi_\#\hat{\nu} = \hat{\mu}$, that is, $\pi_\#\nu_{(t)} = \mu_{(t)}$ for any $t \in [0,T]$.

(3) For each $\boldsymbol{J} \in \mathbb{R}^n$, the set $\widehat{\Lambda}^{\boldsymbol{J}}_{T,\hat{\mu}}$ is defined as all $\hat{\nu} \in \hat{\Lambda}_{T,\hat{\mu}}$ which satisfies

$$T^{-1} \int_0^T \int_{\mathbb{T}^n \times \mathbb{R}^n} \boldsymbol{p}\, d\nu_{(t)}\, dt = \boldsymbol{J}.$$

Definition 4.5. For given $\boldsymbol{P} \in \mathbb{R}^n$, $\hat{\nu} \in \hat{\Lambda}_{T,\hat{\mu}}$ is a weak solution of $F(\hat{\mu}, \boldsymbol{P}, T)$ provided

$$\int_0^T \int_{\mathbb{T}^n \times \mathbb{R}^n} \left[\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p}\cdot\boldsymbol{P}\right] d\nu_{(t)}\, dt \leqslant \int_0^T \int_{\mathbb{T}^n \times \mathbb{R}^n} \left[\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p}\cdot\boldsymbol{P}\right] d\xi_{(t)} \quad \forall \hat{\xi} \in \hat{\Lambda}_{\hat{\mu}}.$$

The T-orbit analogue of Theorem 4.1 is

[3] See point (6) in Section 3.1.

Proposition 4.2. *For any $\hat\mu \in \overline{\mathcal{M}}_T$ and $\boldsymbol{P} \in \mathbb{R}^n$, there exists a unique weak solution $\hat\nu^{(0)} \in \hat\Lambda_{T,\hat\mu}$ of $F(\hat\mu, \boldsymbol{P}, T)$. Moreover,*

$$F(\hat\mu, \boldsymbol{P}, T) = -\frac{1}{T}\int_0^T \int_{\mathbb{T}^n\times\mathbb{R}^n} \left[\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p}\cdot\boldsymbol{P}\right] d\nu_{(t)}^{(0)}\, dt = -\inf_{\hat\nu\in\hat\Lambda_{T,\hat\mu}} \frac{1}{T}\int_0^T \int_{\mathbb{T}^n\times\mathbb{R}^n} \left[\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p}\cdot\boldsymbol{P}\right] d\nu_{(t)}\, dt. \tag{4.14}$$

The Legendre transform of $F(\hat\mu, \boldsymbol{P}, T)$ with respect to $\boldsymbol{P}$ is

$$F^*(\hat\mu, \boldsymbol{J}, T) = \sup_{\phi\in C_0^1(\mathbb{T}^n\times(0,T))} \mathcal{E}(\hat\mu, \boldsymbol{J}, \phi) = \frac{1}{T}\inf_{\hat\nu\in\hat\Lambda^{\boldsymbol{J}}_{T,\hat\mu}} \int_0^T \int_{\mathbb{T}^n\times\mathbb{R}^n} \left[\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p}\cdot\boldsymbol{P}\right] d\nu_{(t)}\, dt, \tag{4.15}$$

where the RHS of (4.15) *is attained for the weak solution of $F(\hat\mu, \boldsymbol{P}, T)$, provided $\boldsymbol{J} \in \partial_{\boldsymbol{P}} F(\hat\mu, \boldsymbol{P}, T)$.*

If $\mu_{(t)}(dx) = \rho(x,t)\,dx$ where $\rho \in C_0^1(\mathbb{T}^n \times (0,T))$ and $\rho > 0$ on $\mathbb{T}^n \times [0,T]$ then a weak solution $\hat\nu$ of $F(\hat\mu, \boldsymbol{P}, T)$ takes the form $\nu_{(t)} = \delta_{\boldsymbol{p}-\boldsymbol{P}-\nabla\phi(x,t)}\, d\boldsymbol{p} \otimes \rho(x,t)\,dx$ where ϕ is the classical solution of the elliptic equation

$$\nabla_x\big[\rho(\nabla_x\phi + \boldsymbol{P})\big] = -\rho_t. \tag{4.16}$$

Definition 4.6. Given $\mu_1, \mu_2 \in \overline{\mathcal{M}}$, $\boldsymbol{P} \in \mathbb{R}^n$ and $\Xi \in C(\mathbb{T}^n)$, define[4]

$$H_{\Xi,T}(\mu_1, \mu_2, \boldsymbol{P}) := \sup_{\hat\mu\in\overline{\mathcal{M}}_T(\mu_1,\mu_2)} \left[\frac{1}{T}\int_0^T\int_{\mathbb{T}^n} \Xi\, d\mu_{(t)}\, dt + F(\hat\mu, \boldsymbol{P}, T)\right]. \tag{4.17}$$

Likewise, for any $\boldsymbol{J} \in \mathbb{R}^n$:

$$H^*_{\Xi,T}(\mu_1, \mu_2, \boldsymbol{J}) := \sup_{\hat\mu\in\overline{\mathcal{M}}_T(\mu_1,\mu_2)} \left[\frac{1}{T}\int_0^T\int_{\mathbb{T}^n} \Xi\, d\mu_{(t)}\, dt - F^*(\hat\mu, \boldsymbol{J}, T)\right]. \tag{4.18}$$

If $\mu_1 = \mu_2 \equiv \mu \in \overline{\mathcal{M}}$ then $H_{\Xi,T}(\mu, \boldsymbol{P}) := H_{\Xi,T}(\mu, \mu, \boldsymbol{P})$ and $H^*_{\Xi,T}(\mu, \boldsymbol{J}) := H^*_{\Xi,T}(\mu, \mu, \boldsymbol{J})$. If $\Xi \equiv 0$, set $F_T(\mu, \boldsymbol{P}) := H_{0,T}(\mu, \boldsymbol{P})$ and $F^*_T(\mu, \boldsymbol{J}) := -H^*_{0,T}(\mu, \boldsymbol{J})$.

Proposition 4.3. *For any $\Xi \in C(\mathbb{T}^n)$, $\boldsymbol{P} \in \mathbb{R}^n$ and $\mu_1, \mu_2 \in \overline{\mathcal{M}}$ there exists an orbit $\hat\mu \in \overline{\mathcal{M}}(\mu_1, \mu_2)$ realizing* (4.17). *Likewise, for any $\boldsymbol{J} \in \mathbb{R}^n$ there exists an orbit $\hat\mu \in \overline{\mathcal{M}}(\mu_1, \mu_2)$ realizing* (4.18).

Proposition 4.4. *For any $T > 0$, $\Xi \in C(\mathbb{T}^n)$, $\boldsymbol{P} \in \mathbb{R}^n$ and $\mu \in \overline{\mathcal{M}}$,*

$$H_{\Xi,T}(\mu, \boldsymbol{P}) \geqslant \int_{\mathbb{T}^n} \Xi\, d\mu + F(\mu, \boldsymbol{P}), \tag{4.19}$$

but

$$\sup_{\mu\in\overline{\mathcal{M}}} H_{\Xi,T}(\mu, \boldsymbol{P}) = \widehat{H}_\Xi(\boldsymbol{P}) := \sup_{\mu\in\overline{\mathcal{M}}} \int_{\mathbb{T}^n} \Xi\, d\mu + F(\mu, \boldsymbol{P}) \quad \forall T > 0, \tag{4.20}$$

and the maximizer of $\widehat{H}_\Xi(\boldsymbol{P})$ (4.11) *is the same as the maximizer of $H_{\Xi,T}(\mu, \boldsymbol{P})$ for any $T > 0$.*

4.3. Optimal transportation

Recall the definition of the action associated with the Lagrangian (2.1): $A^\Xi_{\boldsymbol{P}} : \mathbb{T}^n \times \mathbb{T}^n \times \mathbb{R} \to \mathbb{R}$:

$$A^\Xi_{\boldsymbol{P}}(y, x, T) = \inf_{x(\cdot)} \left\{ \frac{1}{T}\int_0^T \left(\frac{|\dot{\boldsymbol{x}} - \boldsymbol{P}|^2}{2} - \Xi\big(x(s)\big)\right) ds;\ x(0) = y,\ x(t) = x \right\}. \tag{4.21}$$

[4] Recall Section 3.1(6).

Definition 4.7. For $\boldsymbol{P} \in \mathbb{R}^n$, $\mu_1, \mu_2 \in \overline{\mathcal{M}}$ define[5] the Monge–Kantorovich distance with respect to the action A^{Ξ}:

$$D_{\boldsymbol{P}}^T(\mu_1, \mu_2, \Xi) := \min_{\sigma \in \overline{\mathcal{M}}^{(2)}(\mu_1, \mu_2)} \int_{\mathbb{T}^n} A_{\boldsymbol{P}}^{\Xi}(x, y, T)\, d\sigma(x, y). \tag{4.22}$$

If $\mu_1 = \mu_2 \equiv \mu \in \overline{\mathcal{M}}$ we denote

$$D_{\boldsymbol{P}}^T(\mu, \Xi) := \min_{\sigma \in \overline{\mathcal{M}}^{(2)}(\mu)} \int_{\mathbb{T}^n} A_{\boldsymbol{P}}^{\Xi}(x, y, T)\, d\sigma(x, y). \tag{4.23}$$

Example 4.1. If $\Xi \equiv 0$ then $A_{\boldsymbol{P}}^0(y, x, T) = \|x - y - T\boldsymbol{P}\|^2/(2T^2)$, where $\|\cdot\|$ is the Euclidian metric on $\mathbb{T}^n$. In particular

$$D_{\boldsymbol{P}}^T(\mu) := D_{\boldsymbol{P}}^T(\mu, 0) = \min_{\sigma \in \overline{\mathcal{M}}^{(2)}(\mu)} \frac{1}{2T^2} \int_{\mathbb{T}^n} \|x - y - T\boldsymbol{P}\|^2\, d\sigma(x, y). \tag{4.24}$$

Proposition 4.5. *For any $\mu_1, \mu_2 \in \overline{\mathcal{M}}$, $\Xi \in C(\mathbb{T}^n)$, $\boldsymbol{P} \in \mathbb{R}$ and $T > 0$*

$$H_{\Xi,T}(\mu_1, \mu_2, \boldsymbol{P}) = \frac{|\boldsymbol{P}|^2}{2} - D_{\boldsymbol{P}}^T(\mu_1, \mu_2, \Xi).$$

In particular

$$H_{\Xi,T}(\mu, \boldsymbol{P}) = \frac{|\boldsymbol{P}|^2}{2} - D_{\boldsymbol{P}}^T(\mu, \Xi). \tag{4.25}$$

Lemma 4.3. $D_{\boldsymbol{P}}^T : \overline{\mathcal{M}} \times C(\mathbb{T}^n) \to \mathbb{R}$ *is continuous in both the weak C^* topology of $\overline{\mathcal{M}}$ and in the* sup *topology of $C(\mathbb{T}^n)$.*

Proposition 4.6. *For any $\mu \in \overline{\mathcal{M}}$, $\Xi \in C(\mathbb{T}^n)$, $\boldsymbol{P} \in \mathbb{R}$*

$$\lim_{T \to 0} H_{\Xi,T}(\mu, \boldsymbol{P}) = \int_{\mathbb{T}^n} \Xi\, d\mu + F(\mu, \boldsymbol{P}). \tag{4.26}$$

As a corollary to (4.26) and Lemma 4.3, evaluated for $\Xi = 0$, we obtain

Theorem 4.4. *The functional $H_{\Xi,T}$ (Definition 4.6) is continuous on $\overline{\mathcal{M}}$ with respect to the C^* topology. In addition*

$$\lim_{T \to 0} H_{\Xi,T}(\mu, \boldsymbol{P}) = H_{\Xi}(\mu, \boldsymbol{P})$$

for any $\mu \in \overline{\mathcal{M}}$, $\boldsymbol{P} \in \mathbb{R}^n$.

By Definition 4.6 and Theorem 4.4

Corollary 4.2.

$$\lim_{T \to 0} F_T(\mu, \boldsymbol{P}) = F(\mu, \boldsymbol{P})$$

for any $\mu \in \overline{\mathcal{M}}$, $\boldsymbol{P} \in \mathbb{R}^n$.

[5] Recall Section 3.1(7).

4.4. A combinatorial search algorithm

Next, Birkhoff's theorem implies

Lemma 4.4. *Given $x_1, \dots, x_j \in \mathbb{T}^n$. For $\mu_j := j^{-1}\sum_1^j \delta_{x_i}$,*

$$D_{\boldsymbol{P}}^T(\mu_j, \Xi) = \min_{\sigma\in\Pi_j}\sum_{i=1}^{j} A_{\boldsymbol{P}}^{\Xi}(x_i, x_{\sigma(i)}, T)$$

where Π_j is the set of all permutations of $\{1, \dots, j\}$. In particular

$$D_{\boldsymbol{P}}^T(\mu_j) = \frac{1}{2T^2}\min_{\sigma\in\Pi_j}\sum_{i=1}^{j}\|x_i - x_{\sigma(i)} - T\boldsymbol{P}\|^2.$$

By Proposition 4.5, Theorem 4.4, Corollary 4.2 and Lemma 4.4 we obtain the following algorithm for evaluation of $H_\Xi(\mu, \boldsymbol{P})$ and $F(\mu, \boldsymbol{P})$:

Theorem 4.5. *Let $\mu_j := j^{-1}\sum_{i=1}^j \delta_{x_i^{(j)}}$ is a sequence of measures converging C^* to $\mu \in \overline{\mathcal{M}}$. Then*

$$\lim_{T\to 0}\lim_{j\to\infty} D_{\boldsymbol{P}}^T(\mu_j, \Xi) = \frac{|\boldsymbol{P}|^2}{2} - H_\Xi(\mu, \boldsymbol{P}). \tag{4.27}$$

In particular

$$\lim_{T\to 0}\lim_{j\to\infty}\frac{1}{2T^2}\min_{\sigma\in\Pi_j}\sum_{i=1}^{j}\left\|x_i^{(j)} - x_{\sigma(i)}^{(j)} - T\boldsymbol{P}\right\|^2 = \frac{|\boldsymbol{P}|^2}{2} - F(\mu, \boldsymbol{P}).$$

We may use now Theorem 4.5 to evaluate the effective Hamiltonian $\widehat{H}_\Xi(\boldsymbol{P})$. In fact, we do not need to take the limit $T \to 0$, as shown below:

Definition 4.8. Given $j \in \mathbb{N}$, let

$$D_{\boldsymbol{P}}^T(j, \Xi) := \min_{x_1,\dots,x_j\in\mathbb{T}^n}\min_{\sigma\in\Pi_j}\sum_{i=1}^{j} A_{\boldsymbol{P}}^{\Xi}(x_i, x_{\sigma(i)}, T)$$

where Π_j as defined in Lemma 4.4.

Theorem 4.6. *For any $\Xi \in C(\mathbb{T}^n)$, $\boldsymbol{P} \in \mathbb{R}^n$ and $T > 0$,*

$$\widehat{H}_\Xi(\boldsymbol{P}) = \frac{|\boldsymbol{P}|^2}{2} - \lim_{j\to\infty} D_{\boldsymbol{P}}^T(j, \Xi). \tag{4.28}$$

5. Proof of the main results

5.1. Duality

The key duality argument for minimizing convex functionals under affine constraints is summarized in the following proposition. This is a slight generalization of Proposition 4.1 in [24]. The proof is sketched in the appendix of this paper. See also [25]).

Proposition 5.1. *Let $\mathbf{C}$ a real Banach space and $\mathbf{C}^*$ its dual. Let $\mathbf{Z}$ a subspaces of $\mathbf{C}$. Let $h \in \mathbf{C}^*$. Let $\mathbf{Z}^* \subset \mathbf{C}^*$ given by the condition $z^* \in \mathbf{Z}^*$ iff $\langle z^* - h, z\rangle = 0$ for any $z \in \mathbf{Z}$. Let $\mathcal{F} : \mathbf{C}^* \to \mathbb{R} \cup \{\infty\}$ a convex function and*

$$E := \inf_{c^*\in\mathbf{Z}^*}\mathcal{F}(c^*). \tag{5.1}$$

Then

$$\sup_{z\in\mathbf{Z}} \inf_{c^*\in\mathbf{C}^*}\big[\mathcal{F}(c^*) - \langle c^* - h, z\rangle\big] \leqslant E, \tag{5.2}$$

and if $\overline{A}_0 := \{c^* \in \mathbf{C}^*;\ \mathcal{F}(c^*) \leqslant E\}$ *is compact* (*in the* $*$*-topology of* $\mathbf{C}^*$)*, then there is an equality in* (5.2).

In particular, $E < \infty$ *if and only if* $\mathbf{Z}^* \neq \emptyset$*. In this case there exists* $z^* \in \mathbf{Z}^*$ *for which* $E = \mathcal{F}(z^*)$.

Remark 5.1. The case $E < \infty$ does not implies, in general, the existence of $z \in \mathbf{Z}$ realizing (5.2).

5.2. An auxiliary result

Lemma 5.1.

1. F *is convex on* $\mathbb{R}^n$ *as function of* $\boldsymbol{P}$ *and concave on* $\overline{\mathcal{M}}$ *as function of* μ.
2. F^* *is convex on both* $\mathbb{R}^n$ (*as a function of* $\boldsymbol{J}$) *and* $\overline{\mathcal{M}}$.
3. $F^*(\mu,\boldsymbol{J}) = \inf_{\phi\in C^1(\mathbb{T}^n)} \mathcal{E}(\mu,\boldsymbol{J},\phi)$.
4. *The sub-gradients* $\partial_{\boldsymbol{P}} F(\mu,\cdot)$ *and* $\partial_{\boldsymbol{J}} F^*(\mu,\cdot)$ *are nonempty convex cones in* $\mathbb{R}^n$ *for any* $\mu \in \overline{\mathcal{M}}$ *and* $\boldsymbol{P} \in \mathbb{R}^n$ (*resp.* $\boldsymbol{J} \in \mathbb{R}^n$) *and satisfies* $\boldsymbol{P} \in \partial_{\boldsymbol{J}} F^*(\mu,\cdot)$ *iff* $\boldsymbol{J} \in \partial_{\boldsymbol{P}} F(\mu,\cdot)$.
5. F *is upper-semi-continuous in the* C^* *topology of* $\overline{\mathcal{M}}$ *for any* $\boldsymbol{P} \in \mathbb{R}^n$*, and* F^* *is lower-semi-continuous for the same topology for nay* $\boldsymbol{J} \in \mathbb{R}^n$.

In particular, from point (5) of this lemma:

Corollary 5.1. *If* $x_0 \in \mathbb{T}^n$ *and* $\mu_n \to \delta_{x_0}$ *then* $\lim_{n\to\infty} F(\mu_n,\boldsymbol{P}) = F(\delta_{x_0},\boldsymbol{P}) = 0$.

Proof of Lemma 5.1. Concavity of F on $\overline{\mathcal{M}}$ is a result of its definition as a n infimum of functionals over this convex set. The strict convexity on $\mathbb{R}^n$ follows from its quadratic dependence:

$$\frac{|\boldsymbol{P}_1 + \nabla\phi_1|^2}{2} + \frac{|\boldsymbol{P}_2 + \nabla\phi_2|^2}{2} \geqslant \left|\frac{\boldsymbol{P}_1 + \boldsymbol{P}_2}{2} + \frac{\nabla\phi_1 + \nabla\phi_2}{2}\right|^2,$$

holds for any $\boldsymbol{P}_1, \boldsymbol{P}_2 \in \mathbb{R}^n$ and any $\phi_1, \phi_2 \in C^1(\mathbb{T}^n)$. If ϕ_1 approximates the maximizer of $F(\mu,\boldsymbol{P}_1)$ (resp. ϕ_2 approximates the maximizer of $F(\mu,\boldsymbol{P}_2)$), then integrating the above inequality with respect to μ yields

$$F(\mu,\boldsymbol{P}_1) + F(\mu,\boldsymbol{P}_2) \geqslant \int_{\mathbb{T}^n} \left|\frac{\boldsymbol{P}_1 + \boldsymbol{P}_2}{2} + \frac{\nabla\phi_1 + \nabla\phi_2}{2}\right|^2 d\mu \geqslant 2F\left(\mu, \frac{\boldsymbol{P}_1 + \boldsymbol{P}_2}{2}\right).$$

The same arguments apply to F^*. In addition, from (4.1), (4.2)

$$F^*(\mu,\boldsymbol{J}) = \sup_{\boldsymbol{P}\in\mathbb{R}^n}\big[\boldsymbol{P}\cdot\boldsymbol{J} - F(\mu,\boldsymbol{P})\big] = \sup_{\phi\in C^1(\mathbb{T}^n)}\left\{\sup_{\boldsymbol{P}\in\mathbb{R}^n}\left[\boldsymbol{P}\cdot\boldsymbol{J} - \frac{1}{2}\int_{\mathbb{T}^n} |\boldsymbol{P} + \nabla\phi|^2\,d\mu\right]\right\},$$

and from

$$\mathcal{E}(\mu,\boldsymbol{J},\phi) = \sup_{\boldsymbol{P}\in\mathbb{R}^n}\left[\boldsymbol{P}\cdot\boldsymbol{J} - \frac{1}{2}\int_{\mathbb{T}^n} |\boldsymbol{P} + \nabla\phi|^2\,d\mu\right]$$

we obtain (3). (4)–(6) follow from (1)–(2). □

5.3. Proof of Theorem 4.1

We now apply Proposition 5.1 as follows:

Let $\mathbf{C}$ the space of all continuous functions on $\mathbb{T}^n \times \mathbb{R}^n$, equipped with the norm

$$\|q\| := \sup_{(x,\boldsymbol{p})\in\mathbb{T}^n\times\mathbb{R}^n}\left\{\frac{|q(x,\boldsymbol{p})|}{1 + |\boldsymbol{p}|}\right\}.$$

Define

$$\mathbf{Z} := \{\boldsymbol{p}\cdot\nabla_x\phi,\ \phi\in C^1(\mathbb{T}^n)\}. \tag{5.3}$$

The dual space $\mathbf{C}^*$ *contains* the set $\mathcal{M}(\mathbb{T}^n\times\mathbb{R}^n)$ of finite Borel measures on $\mathbb{T}^n\times\mathbb{R}^n$ which admit a finite first moment. If $\nu\in\mathbf{C}^*$ is such a measure then the duality relation is given by

$$\langle\nu,q\rangle = \int_{\mathbb{T}^n\times\mathbb{R}^n} q\,d\nu \quad \forall q\in\mathbf{C}.$$

Given a probability measure $\mu\in\overline{\mathcal{M}}$, define

$$\mathcal{F}_\mu(\nu) = \begin{cases} \int_{\mathbb{T}^n\times\mathbb{R}^n}(|\boldsymbol{p}|^2/2-\boldsymbol{P}\cdot\boldsymbol{p})\,d\nu & \text{if } \nu\in\mathcal{M}(\mathbb{T}^n\times\mathbb{R}^n)\cap\mathbf{C}^* \text{ and satisfies } \nu(dx,\mathbb{R}^n)=\mu(dx),\\ \infty & \text{otherwise}\end{cases}$$

(recall Definition 4.1). Evidently, $\mathcal{F}_\mu$ is a convex function on $\mathbf{C}^*$. Note also that the set $\overline{A}_0 := \{c^*;\ \mathcal{F}_\mu(c^*)<E\}\subset\mathbf{C}^*$ is compact for any $E<\infty$ by Prokhorov theorem.

Lemma 5.2. *$\nu\in\mathbf{Z}^*$ and $\mathcal{F}_\mu(\nu)<\infty$ if and only if $\nu\in\Lambda_\mu$.*

Substitute this $\mathcal{F}_\mu$ for $\mathcal{F}$ in (5.1) where $h\equiv 0$ it follows that

$$E = \inf_{\nu\in\Lambda_\mu}\int_{\mathbb{T}^n\times\mathbb{R}^n}\big(|\boldsymbol{p}|^2/2-\boldsymbol{P}\cdot\boldsymbol{p}\big)\nu(dx). \tag{5.4}$$

On the other hand

$$\begin{aligned}\inf_{c^*\in\mathbf{C}^*}\mathcal{F}(c^*)-\langle c^*,z\rangle &= \inf_{\nu\in\mathbf{C}^*}\int_{\mathbb{T}^n\times\mathbb{R}^n}\left(\frac{1}{2}|\boldsymbol{p}|^2-\boldsymbol{p}\cdot(\nabla_x\phi+\boldsymbol{P})\right)d\nu\\ &= \inf_{\nu\in\Lambda_\mu}\frac{1}{2}\int_{\mathbb{T}^n\times\mathbb{R}^n}|\boldsymbol{p}-\boldsymbol{P}-\nabla_x\phi|^2\,d\nu-\frac{1}{2}\int_{\mathbb{T}^n}|\nabla_x\phi+\boldsymbol{P}|^2\,d\mu. \end{aligned} \tag{5.5}$$

We choose $\nu=\delta^p_{(\boldsymbol{P}+\nabla_x\phi)}\otimes\mu$, so the first term in (5.5) is zero. Then

$$\sup_{z\in\mathbf{Z}}\inf_{c^*\in\mathbf{C}^*}\mathcal{F}(c^*)-\langle c^*,z\rangle = \sup_{\phi\in C^1(\mathbb{T}^n)}\left[-\frac{1}{2}\int_{\mathbb{T}^n}|\nabla_x\phi+\boldsymbol{P}|^2\,d\mu\right] := -F(\mu,\boldsymbol{P}) \tag{5.6}$$

where F as defined in (4.1). The last part of Proposition 5.1 implies the existence of a weak solution $\nu\in\Lambda_\mu$ of $F(\mu,\boldsymbol{P})$.

To show the uniqueness of the weak solution, note that any Borel measure ν on $\mathbb{T}^n\times\mathbb{R}^n$ whose $\mathbb{T}^n$ marginal is μ can be written as $\nu(dx\,d\boldsymbol{p})=\mu(dx)Q_x(d\boldsymbol{p})$ where Q_x is a Borel probability measure on R^n defined for μ-a.e. $x\in\mathbb{T}^n$. If ν satisfies Definition 4.2, then $Q_x=\delta^p_{\boldsymbol{v}}$ where $\boldsymbol{v}(x):=\int_{\mathbb{R}^n}\boldsymbol{p}\,Q_x(d\boldsymbol{p})$ is a Borel vector filed, defined μ-a.e. If there are $\nu_1\neq\nu_2$ which realize the minimum in Definition 4.2 and $\boldsymbol{v}_1,\boldsymbol{v}_2$ the corresponding vector fields, then

$$F(\mu,\boldsymbol{P}) = \frac{1}{2}\int_{\mathbb{T}^n}|\boldsymbol{v}_1|^2\,d\mu = \frac{1}{2}\int_{\mathbb{T}^n}|\boldsymbol{v}_2|^2\,d\mu$$

implies

$$\frac{1}{2}\int_{\mathbb{T}^n}\left|\frac{\boldsymbol{v}_1+\boldsymbol{v}_2}{2}\right|^2 d\mu < F(\mu,\boldsymbol{P}),$$

unless $\boldsymbol{v}_1=\boldsymbol{v}_2$ μ-a.e., which contradicts the minimality of ν_1 and ν_2.

From (5.6) it follows that the *Legendre transform* of the function $F(\mu,\cdot)$ is

$$F^*(\mu, \boldsymbol{J}) = \sup_{\boldsymbol{P}\in\mathbb{R}^n} \sup_{\phi\in C^1(\mathbb{T}^n)} \left[\boldsymbol{P}\cdot\boldsymbol{J} - \frac{1}{2}\int_{\mathbb{T}^n} |\nabla_x\phi + \boldsymbol{P}|^2\, d\mu \right]$$
$$= \sup_{\phi\in C^1(\mathbb{T}^n)} \sup_{\boldsymbol{P}\in\mathbb{R}^n} \left[\boldsymbol{P}\cdot\boldsymbol{J} - \frac{1}{2}\int_{\mathbb{T}^n} |\nabla_x\phi + \boldsymbol{P}|^2\, d\mu \right] = \sup_{\phi\in C^1(\mathbb{T}^n)} \mathcal{E}(\mu, \boldsymbol{J}, \phi) \tag{5.7}$$

where $\mathcal{E}$ as defined in (4.3).

To prove the last part, note that $\nu_0 := \rho(x)\,dx \otimes \delta_{(\boldsymbol{p}-\nabla\psi-\boldsymbol{P})}\,d\boldsymbol{p} \in \Lambda_\mu$ whenever ψ is the solution of (4.6). Indeed, by (4.6) and integration by parts

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \nabla\phi\cdot\boldsymbol{p}\,d\nu_0 = \int_{\mathbb{T}^n} \nabla\phi\cdot(\boldsymbol{P}+\nabla\psi)\rho(x)\,dx = \int_{\mathbb{T}^n} \phi\nabla\cdot\big[(\boldsymbol{P}+\nabla\psi)\rho(x)\big]\,dx = 0 \tag{5.8}$$

for any $\phi \in C^1(\mathbb{T}^n)$. Hence, (5.4) implies

$$E \leqslant \int_{\mathbb{T}^n\times\mathbb{R}^n} \big(|\boldsymbol{p}|^2/2 - \boldsymbol{P}\cdot\boldsymbol{p}\big)\,d\nu_0 = \frac{1}{2}\int_{\mathbb{T}^n} |\nabla\psi+\boldsymbol{P}|^2\rho(x)\,dx - \boldsymbol{P}\cdot\int_{\mathbb{T}^n}(\nabla\psi+\boldsymbol{P})\rho(x)\,dx. \tag{5.9}$$

However, ψ realizes the infimum in (4.1), so (5.4), (5.6) and Proposition 5.1 imply

$$E = -\frac{1}{2}\int_{\mathbb{T}^n} |\nabla\psi+\boldsymbol{P}|^2\rho(x)\,dx$$

which, together with (5.9), imply

$$\int_{\mathbb{T}^n} |\nabla\psi+\boldsymbol{P}|^2\rho(x)\,dx - \boldsymbol{P}\cdot\int_{\mathbb{T}^n}(\nabla\psi+\boldsymbol{P})\rho(x)\,dx \geqslant 0.$$

However, (5.8) with $\phi=\psi$ implies the equality above, hence the equality in (5.9) as well. In particular, ν_0 minimizes (5.4).

5.4. Proof of Theorem 4.2

First, the sequence $\{\nu_j\}$ is tight in $\mathcal{M}(\mathbb{T}^n\times\mathbb{R}^n)$ since $\mathbb{T}^n$ is compact and $\int_{\mathbb{T}^n\times\mathbb{R}^n}|\boldsymbol{p}|^2\,d\nu_j$ are uniformly bounded. By Prokhorov theorem it follows that the weak limit $\nu_j \to \nu \in \mathcal{M}(\mathbb{T}^n\times\mathbb{R}^n)$ exists (for a subsequence). Also, $\nu\in\Lambda_\mu$ since the condition given in Definition 4.1 is preserved under the weak-$*$ convergence.

Next

$$-\lim_{j\to\infty} F(\mu_j, \boldsymbol{P}) = \lim_{j\to\infty}\int_{\mathbb{T}^n\times\mathbb{R}^n}\big(|\boldsymbol{p}|^2/2 - \boldsymbol{P}\cdot\boldsymbol{p}\big)\,d\nu_j \geqslant \int_{\mathbb{T}^n\times\mathbb{R}^n}\big(|\boldsymbol{p}|^2/2 - \boldsymbol{P}\cdot\boldsymbol{p}\big)\,d\nu$$
$$\geqslant \inf_{\xi\in\Lambda_\mu}\int_{\mathbb{T}^n\times\mathbb{R}^n}\big(|\boldsymbol{p}|^2/2 - \boldsymbol{P}\cdot\boldsymbol{p}\big)\,d\xi = -F(\mu, \boldsymbol{P}). \tag{5.10}$$

By assumption (4.8) it follows that the equality folds in (5.10). In particular, ν is the weak solution of $F(\mu, \boldsymbol{P})$.

To prove the second part, let $\eta_\varepsilon \in C^\infty(\mathbb{T}^n)$ a sequence of *positive* mollifiers on $\mathbb{T}^n$ satisfying $\lim_{\varepsilon\to 0}\eta_\varepsilon = \delta_{(\cdot)}$, and $\mu_\varepsilon = \eta_\varepsilon * \mu$. Then $\mu_\varepsilon(dx) = \rho_\varepsilon(x)\,dx$ where $\rho_\varepsilon \in C^\infty(\mathbb{T}^n)$ are strictly positive on $\mathbb{T}^n$ and

$$\lim_{\varepsilon\to 0}\mu_\varepsilon = \mu.$$

Next, let ν be a weak solution of $F(\mu, \boldsymbol{P})$ and $\nu_\varepsilon = \eta_\varepsilon * \nu$. If $\nu = \mu(dx)\nu_x(d\boldsymbol{p})$ and $q = q(p)$ any ν measurable function, then $\tilde{q}(x) := \int_{\mathbb{R}^n} q(\boldsymbol{p})\nu_x(d\boldsymbol{p})$ is μ measurable and

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} q(\boldsymbol{p})\,d\nu = \int_{\mathbb{T}^n}\tilde{q}(x)\mu(dx)$$

while

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} q(\boldsymbol{p})\,d\nu_\varepsilon = \int_{\mathbb{T}^n\times\mathbb{R}^n}\int_{\mathbb{T}^n} dx\,\eta_\varepsilon\big(|y-x|\big)\mu(dy)\nu_y(d\boldsymbol{p})q(\boldsymbol{p})$$

$$= \int_{\mathbb{T}^n}\int_{\mathbb{T}^n} \eta_\varepsilon\big(|x-y|\big)\,dx\,\mu(dy)\tilde{q}(y) = \int_{\mathbb{T}^n} dx\,\mu(dy)\tilde{q}(y) = \int_{\mathbb{T}^n\times\mathbb{R}^n} q(\boldsymbol{p})\,d\nu \tag{5.11}$$

for any ν-measurable function q on $\mathbb{R}^n$. Then

$$\int_{\mathbb{T}^n\times\mathbb{R}^n} \nabla\theta\cdot\boldsymbol{p}\,d\nu_\varepsilon(x) = \int_{\mathbb{T}^n\times\mathbb{R}^n} \nabla(\eta_\varepsilon * \theta)\cdot\boldsymbol{p}\,d\nu(x) = 0$$

for any $\theta\in C^\infty(\mathbb{T}^n)$ hence $\nu_\varepsilon\in\Lambda_{\mu_\varepsilon}$.

Define

$$\boldsymbol{v}_\varepsilon(x) = \rho_\varepsilon^{-1}(x)\int_{\mathbb{R}^n}\boldsymbol{p}\,d\nu_\varepsilon(x,d\boldsymbol{p}),\quad x\in\mathbb{T}^n.$$

Then $\boldsymbol{v}_\varepsilon\in C^\infty(\mathbb{T}^n)$ and $\nabla\cdot(\rho_\varepsilon\boldsymbol{v}_\varepsilon)=0$. Let ϕ_ε be the unique solution of the elliptic equation

$$\nabla\cdot\big(\rho_\varepsilon[\nabla\phi_\varepsilon+\boldsymbol{P}]\big)=0. \tag{5.12}$$

Define

$$\hat{\nu}_\varepsilon(dx\,d\boldsymbol{p}) := \rho_\varepsilon(x)\,dx\otimes\delta_{(\boldsymbol{p}-\boldsymbol{P}-\nabla\phi_\varepsilon)}\,d\boldsymbol{p}.$$

Then (5.12) implies, as in (5.8), that $\hat{\nu}_\varepsilon\in\Lambda_{\mu_\varepsilon}$. By (4.4) and (5.11) and the second part of Theorem 4.1

$$-F(\mu,\boldsymbol{P}) := \int_{\mathbb{T}^n\times\mathbb{R}^n}\big(|\boldsymbol{p}|^2/2-\boldsymbol{P}\cdot\boldsymbol{p}\big)\,d\nu = \int_{\mathbb{T}^n\times\mathbb{R}^n}\big(|\boldsymbol{p}|^2/2-\boldsymbol{P}\cdot\boldsymbol{p}\big)\,d\nu_\varepsilon$$

$$\geqslant \int_{\mathbb{T}^n\times\mathbb{R}^n}\big(|\boldsymbol{p}|^2/2-\boldsymbol{P}\cdot\boldsymbol{p}\big)\,d\hat{\nu}_\varepsilon = -F(\mu_\varepsilon,\boldsymbol{P}), \tag{5.13}$$

hence

$$\lim_{\varepsilon\to 0} F(\mu_\varepsilon,\boldsymbol{P}) \geqslant F(\mu,\boldsymbol{P}) \tag{5.14}$$

for any $\boldsymbol{P}\in\mathbb{R}^n$. But, since F is concave in μ via Lemma 5.1 it follows that there is an equality in (5.14).

5.5. Proof of Lemma 4.1

Since F is convex on $\mathbb{R}^n$ for fixed μ and concave on $\overline{\mathcal{M}}$ for fixed $\boldsymbol{P}$ we may use the Min–Max theorem [13] to obtain

$$\widehat{H}^*_\Xi(\boldsymbol{J}) = \sup_{\mu\in\overline{\mathcal{M}}}\inf_{\boldsymbol{P}\in\mathbb{R}^n}\left[\int_{\mathbb{T}^n}\Xi\,d\mu + F(\mu,\boldsymbol{P}) - \boldsymbol{P}\cdot\boldsymbol{J}\right]$$

$$= \inf_{\boldsymbol{P}\in\mathbb{R}^n}\left\{\sup_{\mu\in\overline{\mathcal{M}}}\left[\int_{\mathbb{T}^n}\Xi\,d\mu + F(\mu,\boldsymbol{P})\right] - \boldsymbol{P}\cdot\boldsymbol{J}\right\} = -\sup_{\boldsymbol{P}\in\mathbb{R}^n}\big\{\boldsymbol{P}\cdot\boldsymbol{J} - \widehat{H}_\Xi(\boldsymbol{P})\big\}. \tag{5.15}$$

5.6. Proof of Proposition 4.1

Let $\Xi(x_0) = \max_{\mathbb{T}^n}\Xi(x)$. Let $\mu_n\to\delta_{x_0}$. By Lemma 5.1(5), we obtain the weak inequality for the r.h.s. of (4.11) defining $\widehat{H}_\Xi(\boldsymbol{P})$. To obtain the strong inequality use, e.g., the uniform Lebesgue measure $\mu = dx$ on $\mathbb{T}^n$. Then the minimizer ϕ of F (4.1) verifies $\nabla\cdot(\nabla\phi+\boldsymbol{P}) = 0$, that is, $\Delta\phi = 0$ on $\mathbb{T}^n$ which implies $\nabla\phi\equiv 0$. Hence $F(dx,\boldsymbol{P}) = |\boldsymbol{P}|^2/2$. We obtain the strong inequality if $|\boldsymbol{P}|^2/2 + \int_{\mathbb{T}^n}\Xi > \max_{\mathbb{T}^n}\Xi$. The proof for $\widehat{H}^*$ is analogous.

5.7. Proof of Lemma 4.2

The existence of a maximizer μ of (4.10) follows from the lower-semi-continuity of F^* (hence of H^*_Ξ) with respect to the C^* topology of $\overline{\mathcal{M}}$, as claimed in Lemma 5.1(5).

Let

$$H_\Xi(\mu, \boldsymbol{P}) := \int \Xi \, d\mu + F(\mu, \boldsymbol{P}), \qquad H^*_\Xi(\mu, \boldsymbol{J}) := \int \Xi \, d\mu - F^*(\mu, \boldsymbol{J}). \tag{5.16}$$

Then, by Definition 4.3, $\widehat{H}_\Xi(\boldsymbol{P}) = H_\Xi(\mu_0, \boldsymbol{P}) \geqslant H_\Xi(\mu, \boldsymbol{P})$ for any $\mu \in \overline{\mathcal{M}}$, and $\widehat{H}^*_\Xi(\boldsymbol{J}) = \sup_{\mu \in \overline{\mathcal{M}}} H^*_\Xi(\mu, \boldsymbol{J})$. By Lemma 4.1

$$\widehat{H}_\Xi(\boldsymbol{P}) - \widehat{H}^*_\Xi(\boldsymbol{J}') \geqslant \boldsymbol{P} \cdot \boldsymbol{J}'$$

holds for any $\boldsymbol{J}' \in \mathbb{R}^n$, and the equality above takes place if and only if $\boldsymbol{J}' = \boldsymbol{J} \in \partial_{\boldsymbol{P}} \widehat{H}_\Xi(\cdot)$. Hence

$$H_\Xi(\mu_0, \boldsymbol{P}) - H^*_\Xi(\mu, \boldsymbol{J}') \geqslant \boldsymbol{P} \cdot \boldsymbol{J}' \tag{5.17}$$

holds for any $\mu \in \overline{\mathcal{M}}$ and any $\boldsymbol{J}' \in \mathbb{R}^n$. The equality holds if and only if $\boldsymbol{J}' = \boldsymbol{J} \in \partial_{\boldsymbol{P}} \widehat{H}_\Xi(\cdot)$ and μ which verifies the maximum of $H^*_\Xi(\cdot, \boldsymbol{J})$. From (5.16), (5.17)

$$F(\mu_0, \boldsymbol{P}) + F^*(\mu, \boldsymbol{J}') + \int_{\mathbb{T}^n} \Xi (d\mu_0 - d\mu) \geqslant \boldsymbol{P} \cdot \boldsymbol{J}'. \tag{5.18}$$

Let now $\mu = \mu_0$. Then (5.18) implies

$$F(\mu_0, \boldsymbol{P}) + F^*(\mu_0, \boldsymbol{J}') \geqslant \boldsymbol{P} \cdot \boldsymbol{J}', \tag{5.19}$$

and, *if* there is an equality in (5.19) for some $\boldsymbol{J}'$, then $\boldsymbol{J}' \in \partial_{\boldsymbol{P}} \widehat{H}_\Xi(\cdot)$. However, we know, by definition of F^* as the Legendre transform of F with respect to $\boldsymbol{P}$, that there is, indeed, an equality in (5.19) provided $\boldsymbol{J}' \in \partial_{\boldsymbol{P}} F(\mu_0, \cdot)$. This verifies $\partial_{\boldsymbol{P}} F(\mu_0, \cdot) \subset \partial_{\boldsymbol{P}} \widehat{H}_\Xi$.

5.8. Proof of Theorem 4.3

Assume $\nu \in \mathcal{M}_L$ is a Mather measure on $\mathbb{T}^n \times \mathbb{R}^n$. Theorem 5.1.2 of [7] implies the existence of a conjugate pair $\phi_+ \geqslant \phi_- \in Lip(\mathbb{T}^n)$ where the domain $\phi_+ = \phi_-$ contains the projected Mother set which, in turn, contains the support of the projection μ of ν on $\mathbb{T}^n$. In addition, Corollary 4.2.20 of [7] implies that either functions satisfies

$$\frac{1}{2}|\nabla\phi + \boldsymbol{P}|^2 + \Xi \leqslant E, \tag{5.20}$$

and for some $E \in \mathbb{R}$, with an equality on the projected Mather set (in particular, on the support of μ).

We show that μ is also a maximizer of $\widehat{H}_\Xi(\boldsymbol{P})$ (4.11).

Let ϕ be either ϕ_+ or ϕ_-. Let $\eta_\varepsilon \in C^\infty(\mathbb{T}^n)$ nonnegative mollifier function on $\mathbb{T}^n$, supported in the ball $|x| < \varepsilon$ and satisfying $\int_{\mathbb{T}^n} \eta_\varepsilon = 1$. Let $\phi^\varepsilon := \phi * \eta_\varepsilon$. Then $\phi^\varepsilon \in C^\infty(\mathbb{T}^n)$ and $\nabla\phi^\varepsilon = \eta_\varepsilon * \nabla\phi$. The Jensen's inequality implies that

$$\eta_\varepsilon * |\nabla\phi + \boldsymbol{P}|^2 \geqslant \left|\nabla\phi^\varepsilon + \boldsymbol{P}\right|^2, \tag{5.21}$$

so, by (5.20), (5.21)

$$\frac{1}{2}\left|\nabla\phi^\varepsilon + \boldsymbol{P}\right|^2 + \Xi \leqslant E + \Xi - \eta_\varepsilon * \Xi.$$

Given $\delta > 0$, there exists $\varepsilon > 0$ for which $|\Xi - \eta_\varepsilon * \Xi| < \delta$ on $\mathbb{T}^n$. Hence

$$\frac{1}{2}\left|\nabla\phi^\varepsilon + \boldsymbol{P}\right|^2 + \Xi \leqslant E + \delta.$$

So, for any $\tilde{\mu} \in \overline{\mathcal{M}}$:

$$E + \delta \geqslant \int_{\mathbb{T}^n} \left(\frac{1}{2}\left|\nabla\phi^\varepsilon + \boldsymbol{P}\right|^2 + \Xi\right) d\tilde{\mu} \geqslant F(\tilde{\mu}, \boldsymbol{P}) + \int_{\mathbb{T}^n} \Xi \, d\tilde{\mu} = \widehat{H}_\Xi(\boldsymbol{P}),$$

where the second inequality follows from the definition (4.1) of $F(\mu, \boldsymbol{P})$ and from $\phi^\varepsilon \in C^1(\mathbb{T}^n)$. Since $\delta > 0$ can be arbitrarily small it follows that $\widehat{H}_\Xi(\boldsymbol{P}) \leqslant E$. Hence $\mu = \pi_{\#}\nu$ is, indeed, a maximizer of $\widehat{H}_\Xi$ (4.11).

Finally, if $\boldsymbol{J} = \boldsymbol{\alpha}(\nu)$ then $\boldsymbol{J} \in \partial_{\boldsymbol{P}} \widehat{H}_\Xi$ and the last part of the theorem follows from Lemma 4.2.

5.9. *Proof of Proposition 4.2*

The proof is analogous to this of Theorem 4.1, utilizing Proposition 5.1. We only sketch the new definitions involved, generalizing those given in the proof of Theorem 4.1 to the time periodic case.

Let $\mathbf{C}$ to be the space of all continuous functions on $\mathbb{T}^n \times \mathbb{R}^n \times [0, T]$, equipped with the norm

$$\|q\| := \sup_{(x,\boldsymbol{p},t)\in\mathbb{T}^n\times\mathbb{R}^n\times[0,T]} \left\{ \frac{|q(x, \boldsymbol{p}, t)|}{1 + |\boldsymbol{p}|} \right\}.$$

Define

$$\mathbf{Z} := \left\{ \phi_t + \boldsymbol{p} \cdot \nabla_x \phi,\ \phi \in C_0^1\left(\mathbb{T}^n \times (0, T)\right) \right\}. \tag{5.22}$$

The dual space $\mathbf{C}^*$ *contains* the set $\overline{\mathcal{M}}_T$ of $\overline{\mathcal{M}}$-valued orbits on $[0, T]$ of bounded first moment. If $\hat{\nu} \in \mathbf{C}^*$ is such an orbit then the duality relation is given by

$$\langle \hat{\nu}, q \rangle := \int_0^T \int_{\mathbb{T}^n\times\mathbb{R}^n} q\, d\nu_{(t)}\, dt \quad \forall q \in \mathbf{C}.$$

Given an orbit $\hat{\mu} \in \overline{\mathcal{M}}_T$, define

$$\begin{aligned}
&\mathcal{F}_{\hat{\mu}}(\hat{\nu}, T) := \int_{\mathbb{T}^n\times\mathbb{R}^n\times[0,T]} \left(|\boldsymbol{p}|^2/2 - \boldsymbol{P} \cdot \boldsymbol{p}\right) d\nu_{(t)}\, dt \\
&\qquad \text{if } \hat{\nu} \in \mathcal{M}_T\left(\mathbb{T}^n \times \mathbb{R}^n \times [0, T]\right) \cap \mathbf{C}^* \quad \text{and} \quad \nu_{(t)}\left(dx, \mathbb{R}^n\right) = \mu_{(t)}(dx) \quad \text{a.e.}, \\
&\mathcal{F}_{\hat{\mu}}(\hat{\nu}, T) = \infty \quad \text{otherwise.}
\end{aligned} \tag{5.23}$$

The rest of the proof is equivalent to this of Theorem 4.1.

5.10. *Proof of Proposition 4.3*

First we note that $F_T(\mu, \boldsymbol{P}) \geqslant 0$ for any $\mu \in \overline{\mathcal{M}}$ and $\boldsymbol{P} \in \mathbb{R}^n$. Indeed, by Theorem 4.2:

$$F_T(\mu, \boldsymbol{P}) \geqslant F(\mu, \boldsymbol{P}) \geqslant 0. \tag{5.24}$$

Let $\hat{\mu}^{(n)}$ be a maximizing sequence of (4.17). By (4.14) and (5.24) it follows that there exists $C > 0$ for which

$$\frac{1}{T} \int_0^T \int_{\mathbb{T}^n\times\mathbb{R}^n} \left(\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p} \cdot \boldsymbol{P} \right) d\nu_{(t)}^{(n)} dt \leqslant C \tag{5.25}$$

where $\hat{\nu}^{(n)}$ are the weak solutions corresponding to $\hat{\mu}^{(n)}$.

Let

$$\|\hat{\mu}\|_T^2 := - \inf_{\phi\in C^1(\mathbb{T}^n\times[0,T])} \int_0^T \left(\phi_t + \frac{1}{2} |\nabla_x \phi|^2 \right) d\mu_{(t)}\, dt.$$

We recall form [24] that

$$\|\hat{\mu}\|_T^2 = \frac{1}{2} \inf_{\hat{\nu}\in\hat{\Lambda}_{\hat{\mu}}} \int_0^T \int_{\mathbb{T}^n\times\mathbb{R}^n} |\boldsymbol{p}|^2\, d\nu_{(t)}\, dt. \tag{5.26}$$

Moreover, Lemma 2.2 in [24] implies that the set $\{\hat{\mu} \in \hat{\Lambda}_{\hat{\mu}};\ \|\hat{\mu}\|_T < C\}$ is uniformly bounded in the 1/2-Hölder norm with respect to the Wasserstein metric:

$$W_1(\mu_1, \mu_2) := \sup_{\phi \in C^1(\mathbb{T}^n), |\nabla\phi| \leqslant 1} \int_{\mathbb{T}^n} \phi(x)(d\mu_1 - d\mu_2) = \inf_{\sigma \in \overline{\mathcal{M}}^{(2)}(\mu)} \int_{\mathbb{T}^n}\int_{\mathbb{T}^n} |x-y|_T \, d\sigma,$$

(recall points (i) and (vi) in list of symbols). Since The W_1-Wasserstein metric is a metrization of the weak topology of measures on compact domains, it follows (see Corollary 2.1 in [24]) that the set $\{\hat{\mu} \in \hat{\Lambda}_{\hat{\mu}};\ \|\hat{\mu}\|_T < C\}$ is pre-compact in the topology of $C([0,T]; C^*(\mathbb{T}^n))$.

Next, since $|\boldsymbol{p}|^2/2 - \boldsymbol{P}\cdot\boldsymbol{p} \geqslant |\boldsymbol{p}|^2/4 - |\boldsymbol{P}|^2$ it follows from (5.25) and (5.26) that

$$C \geqslant \frac{1}{T} \inf_{\hat{\nu} \in \hat{\Lambda}_{\hat{\mu}^{(n)}}} \int_0^T \int_{\mathbb{T}^n} \left(\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{P}\cdot\boldsymbol{p}\right) d\nu_{(t)}\, dt \geqslant \frac{1}{2T} \big\|\hat{\mu}^{(n)}\big\|_T^2 - |\boldsymbol{P}|^2.$$

Hence, the limit

$$\hat{\hat{\mu}} = \lim_{n\to\infty} \hat{\mu}^{(n)} \in \hat{\Lambda}_{\hat{\mu}}$$

exists in the weak topology of $C^*([0,T]; C^*(\mathbb{T}^n))$, along a subsequence of the maximizing sequence. Moreover, $\hat{\hat{\mu}}$ is a maximizer of (4.17) by concavity of $F(\hat{\mu}, \boldsymbol{P}, T)$ with respect to $\hat{\mu}$.

5.11. *Proof of Proposition 4.4*

The inequality (4.19) follows directly from (4.17), upon substituting the *constant* orbit $\hat{\mu} \equiv \mu_{(t)} \equiv \mu$ for $t \in [0,T]$. To verify (4.20) we use (4.14) to write

$$\sup_{\mu \in \overline{\mathcal{M}}} H_{\Xi,T}(\mu, \boldsymbol{P}) = \sup_{\hat{\nu} \in \hat{\Lambda}_T} \frac{1}{T} \int_0^T \int_{\mathbb{T}^n \times \mathbb{R}^n} \left\{ \Xi - \left[\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p}\cdot\boldsymbol{P}\right]\right\} d\nu_{(t)}\, dt. \tag{5.27}$$

But, the RHS of (5.27) is unchanged if we replace $\hat{\nu}$ by $\nu := T^{-1}\int_0^T d\nu_{(t)}$. Moreover, $\nu \in \Lambda$ since Definition 4.4 is reduced to Definition 4.1 for $\theta(x,t) \to \theta(x) := T^{-1}\int_0^T \theta(x,t)\,dt$. Hence (5.27) is reduced into

$$\sup_{\mu \in \overline{\mathcal{M}}} H_{\Xi,T}(\mu, \boldsymbol{P}) = \sup_{\nu \in \Lambda} \int_{\mathbb{T}^n \times \mathbb{R}^n} \left\{ \Xi - \left[\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{p}\cdot\boldsymbol{P}\right]\right\} d\nu,$$

which yields (4.20) via (4.14).

5.12. *Proof of Lemma 4.3*

The lower-semi-continuity of $D^T_{\boldsymbol{P}}$ with respect to $C^*(\overline{\mathcal{M}})$ follows from the dual formulation

$$D^T_{\boldsymbol{P}}(\mu) = \sup_{\psi_1,\psi_2} \left(\int_{\mathbb{T}^n} \psi_1\, d\mu + \int_{\mathbb{T}^n} \psi_2\, d\mu \right)$$

where the supremum above is taken over all pairs $\psi_1, \psi_2 \in C(\mathbb{T}^n)$ verifying $\psi_2(y) + \psi_1(x) \leqslant A^{\Xi}_{\boldsymbol{P}}(x,y,T)$ for any $x, y \in \mathbb{T}^n$. For details, see [23, Chapter 1].

The upper-semi-continuity follows directly from definition (4.23). Indeed, let $\sigma_j \in \overline{\mathcal{M}}^{(2)}(\mu)$ verifies (4.23) for $\mu_i \in \overline{\mathcal{M}}$ and $\mu_j \to \mu$ in $C^*(\overline{\mathcal{M}})$, then the sequence σ_j is compact in the set $\mathcal{M}(\mathbb{T}^n \times \mathbb{T}^n)$ in the weak topology. Let σ be a limit of this sequence. Then $\sigma \in \overline{\mathcal{M}}^{(2)}(\mu)$, and

$$D^T_{\boldsymbol{P}}(\mu) \leqslant \int_{\mathbb{T}^n} A^{\Xi}_{\boldsymbol{P}}(x,y,T)\, d\sigma = \lim_{j\to\infty} \int_{\mathbb{T}^n} A^{\Xi}_{\boldsymbol{P}}(x,y,T)\, d\sigma_j = \lim_{j\to\infty} D^{\Xi}_{\boldsymbol{P}}(\mu_j).$$

The continuity of D^{Ξ}_T with respect to Ξ in the $C^0(\mathbb{T}^n)$ topology is verified by the continuous dependence of A^{Ξ}_T on Ξ.

5.13. Proof of Proposition 4.5

Definition 4.6 of $\widehat{H}_T$ corresponds to Definition 3.1 of $\mathcal{L}$ in [24]. In addition, Definition 3.3 [24] of $\mathcal{K}$ corresponds to (4.22). Then, Proposition 4.6 is a result of the identity $\mathcal{L}(\mu_1, \mu_2) = \mathcal{K}(\mu_1, \mu_2)$, which follows from the Main Theorem of [24].

In fact, the extended Lagrangian $\mathcal{L}$ is defined, in [24], for a Lagrangian $L = |\boldsymbol{p}|^2/2 - \Xi(x)$, i.e. for $\boldsymbol{P} = 0$, but the proof of the Main Theorem in [24] can be extended to $L = |\boldsymbol{P} - \boldsymbol{p}|^2/2 - \Xi$ in a direct way.

5.14. Proof of Proposition 4.6

Since $\widehat{H}_T(\mu, \boldsymbol{P}, \Xi) \geqslant \int \Xi\, d\mu + F(\mu, \boldsymbol{P})$ by Definition 4.6, it follows from Proposition 4.5 that

$$\frac{|\boldsymbol{P}|^2}{2} - D^T_{\boldsymbol{P}}(\mu, \Xi) \geqslant \int_{\mathbb{T}^n} \Xi\, d\mu + F(\mu, \boldsymbol{P}).$$

Thus, we only need to show that

$$\lim_{T\to 0} D^T_{\boldsymbol{P}}(\mu, \Xi) \geqslant \frac{|\boldsymbol{P}|^2}{2} - \int_{\mathbb{T}^n} \Xi\, d\mu - F(\mu, \boldsymbol{P}).$$

We may reduce to the case $\Xi \equiv 0$, hence we need to verify

$$\lim_{T\to 0} \frac{|\boldsymbol{P}|^2}{2} - D^T_{\boldsymbol{P}}(\mu) \leqslant F(\mu, \boldsymbol{P}). \tag{5.28}$$

Indeed, from (4.21), (4.22) we observe

$$\lim_{T\to 0} D^T_{\boldsymbol{P}}(\mu, \Xi) = -\int_{\mathbb{T}^n} \Xi\, d\mu + \lim_{T\to 0} D^T_{\boldsymbol{P}}(\mu).$$

If σ_T verifies the minimum in (4.22), then

$$D^T_{\boldsymbol{P}}(\mu) = \frac{1}{2T^2} \int_{\mathbb{T}^n}\int_{\mathbb{T}^n} \|y - x + \boldsymbol{P}T\|^2\, d\sigma_T(x, y) \leqslant \frac{|\boldsymbol{P}|^2}{2}. \tag{5.29}$$

Let

$$G := \{(x, y) \in \mathbb{T}^n \times \mathbb{T}^n;\ \|y - x\|_\infty \geqslant 1/3\},$$

where $\|x\|_\infty$ is the metric on $\mathbb{T}^n$ defined as $\min_{z\in\mathbb{Z}} |x - z|_\infty$. Then, for sufficiently small T, $\|y - x + \boldsymbol{P}T\|^2 \geqslant 1/16$ for $(x, y) \in G$ so from (5.29)

$$\iint_G d\sigma_T(x, y) \leqslant 16T^2|\boldsymbol{P}|^2. \tag{5.30}$$

Let $B^n \subset \mathbb{R}^n$ be the unite box $-1/2 \leqslant x_i \leqslant 1/2$, $i = 1, \dots, n$. Let $\phi \in C_0(\mathbb{T}^n \times T^{-1}B^n/3)$. Extend ϕ to a function in $C_0(\mathbb{T}^n \times T^{-1}B^n)$ by $\phi(x, \boldsymbol{p}) = 0$ if $\boldsymbol{p} \in T^{-1}B^n - T^{-1}B^n/3$. Further, extend ϕ into a function on $\mathbb{T}^n \times \mathbb{R}^n$ as a T^{-1} periodic function in $\boldsymbol{p}$, that is, ϕ is a function on $\mathbb{T}^n \times (\mathbb{T}^n/T)$. Set

$$y = x + \boldsymbol{p}T, \qquad \hat{\phi}(x, y) := \phi\left(x, \frac{y - x}{T}\right).$$

Then $\hat{\phi} \in C(\mathbb{T}^n \times \mathbb{T}^n)$. Given σ_T which verifies the minimum in (4.22), we define a corresponding measure ν_T on $\mathbb{T}^n \times \mathbb{R}^n$, supported in $\mathbb{T}^n \times (B^n/(3T))$, as follows: For any $\phi \in C_0(\mathbb{T}^n \times B^n/(3T))$,

$$\int_{\mathbb{T}^n}\int_{B^n/(3T)} \phi(x, \boldsymbol{p})\, d\nu_T(x, \boldsymbol{p}) = \int_{\mathbb{T}^n}\int_{\mathbb{T}^n} \hat{\phi}(x, y)\, d\sigma_T(x, y). \tag{5.31}$$

By (5.30) we obtain

$$1 \geqslant \int_{\mathbb{T}^n}\int_{\mathbb{R}^n} d\nu_T \geqslant 1 - 16T^2|\boldsymbol{P}|^2. \tag{5.32}$$

We now verify that

$$\lim_{T\to 0} \nu_T = \nu_0 \in \Lambda_\mu \tag{5.33}$$

(see Definition 4.1(2)). First, we show that the sequence of measures on $\mathbb{T}^n \times \mathbb{R}^n$ is tight. For this, we use (5.31) with $\phi(x,\boldsymbol{p}) = |\boldsymbol{p}-\boldsymbol{P}|^2 \cdot 1_{B^n/(3T)}(\boldsymbol{p})$ and (5.29) to obtain

$$\frac{1}{2}\int_{\mathbb{T}^n}\int_{\mathbb{R}^n} |\boldsymbol{p}-\boldsymbol{P}|^2\, d\nu_T < \frac{|\boldsymbol{P}|^2}{2}.$$

This, and (5.32), imply that the limit (5.33) is a probability measure on $\mathbb{T}^n \times \mathbb{R}^n$, that is $\nu_0 \in \mathcal{M}(\mathbb{T}^n \times \mathbb{R}^n)$. Moreover, $\nu_0(dx, \mathbb{R}^n) = \mu(dx)$. To show that $\nu_0 \in \Lambda_\mu$, we proceed as follows: Let $q \in C^\infty(\mathbb{R}^+)$ satisfies:

(i) $q \in C^\infty(\mathbb{R}^+)$.
(ii) $q(s) = 1$ for $0 \leqslant s \leqslant 1/3$.
(iii) $q(s) = 0$ for $2/5 \leqslant s \leqslant 1/2$.
(iv) $q(1/2 - s) = q(1/2 + s)$ for any $s \in [0, 1/2]$.
(v) $q(s) \leqslant 1$ for $s \in [0, 1]$.
(vi) $q(s+1) = q(s)$ for all $s \in \mathbb{R}^+$.

Let

$$Q(\boldsymbol{p}) = \Pi_1^N q\big(|p_i|\big) \quad \text{for } \boldsymbol{p} = (p_1, \dots, p_n) \in \mathbb{R}^n, \qquad Q_T(\boldsymbol{p}) = Q(T\boldsymbol{p}).$$

Given $\psi \in C^1(\mathbb{T}^n)$, set $\phi_T^{(1)}(x,p) = \psi(x + pT)Q_T(\boldsymbol{p})$ and $\phi_T^{(2)}(x,p) = \psi(x)Q_T(\boldsymbol{p})$ if $\boldsymbol{p} \in B^n/(3T)$, $\phi_T = 0$ otherwise. Then, by (5.31)

$$\begin{aligned}\int_{\mathbb{T}^n}\int_{\mathbb{R}^n} \phi_T^{(1)}(x,p)\, d\nu_T(x,p) &= \int_{\mathbb{T}^n}\int_{\mathbb{T}^n} Q(y-x)\psi(y)\, d\sigma_T(x,y) \\ &= \int_{\mathbb{T}^n}\int_{\mathbb{T}^n} \psi(y)\, d\sigma_T(x,y) + \int_{\mathbb{T}^n}\int_{\mathbb{T}^n} \big(Q(y-x)-1\big)\psi(y)\, d\sigma_T(x,y) \\ &= \int_{\mathbb{T}^n} \psi(y)\, d\mu(y) + \int_{\mathbb{T}^n}\int_{\mathbb{T}^n} \big(Q(y-x)-1\big)\psi(y)\, d\sigma_T(x,y).\end{aligned} \tag{5.34}$$

The same argument applies also to $\phi^{(2)}$ and yields

$$\int_{\mathbb{T}^n}\int_{\mathbb{R}^n} \phi^{(2)}(x,p)\, d\nu_T(x,\boldsymbol{p}) = \int_{\mathbb{T}^n} \psi(x)\, d\mu(x) + \int_{\mathbb{T}^n}\int_{\mathbb{T}^n} \big(Q(y-x)-1\big)\psi(x)\, d\sigma_T(x,y). \tag{5.35}$$

However, $Q(y-x) - 1 = 0$ on the set G so, by (5.30)

$$\int_{\mathbb{T}^n}\int_{\mathbb{T}^n} \big(Q(y-x)-1\big)\psi(x)\, d\sigma_T(x,y) = \int_{\mathbb{T}^n}\int_{\mathbb{T}^n} \big(Q(y-x)-1\big)\psi(y)\, d\sigma_T(x,y) \leqslant 16T^2|\boldsymbol{P}|^2|\psi|_\infty. \tag{5.36}$$

Subtract (5.35) from (5.34), divide by T and let $T \to 0$ and use (5.36) to obtain

$$\begin{aligned}0 &= \lim_{T\to 0}\int_{\mathbb{T}^n}\int_{\mathbb{R}^n} \frac{\phi^{(2)}(x,\boldsymbol{p}) - \phi^{(1)}(x,\boldsymbol{p})}{T}\, d\nu_T(x,\boldsymbol{p}) \\ &= \lim_{T\to 0}\int_{\mathbb{T}^n}\int_{\mathbb{R}^n} Q_T(\boldsymbol{p})\frac{\psi(x+T\boldsymbol{p}) - \psi(x)}{T}\, d\nu_T(x,\boldsymbol{p}) = \int_{\mathbb{T}^n}\int_{\mathbb{R}^n} \nabla\psi \cdot \boldsymbol{p}\, d\nu_0(x,\boldsymbol{p}),\end{aligned} \tag{5.37}$$

which implies for any $\psi \in C^1(\mathbb{T}^n)$, hence $\nu_0 \in \Lambda_\mu$ as claimed.

Let now consider $\int_{\mathbb{T}^n} Q_T(\boldsymbol{p})|\boldsymbol{p}-\boldsymbol{P}|^2\,d\nu_T$. By (5.31)

$$\frac{1}{2}\int\limits_{\mathbb{T}^n}\int\limits_{B^n/(3T)} Q_T(\boldsymbol{p})|\boldsymbol{p}-\boldsymbol{P}|^2\,d\nu_T(x,\boldsymbol{p}) = \frac{1}{2}T^{-2}\int\limits_{\mathbb{T}^n}\int\limits_{\mathbb{T}^n} Q(x-y)\|x-y-T\boldsymbol{P}\|^2\,d\sigma_T(x,y). \tag{5.38}$$

Since σ_T is a minimizer of (4.24) and $Q_T \leqslant 1$, it follows that

$$\frac{1}{2}T^{-2}\int\limits_{\mathbb{T}^n}\int\limits_{\mathbb{T}^n} Q(x-y)\|x-y-T\boldsymbol{P}\|^2\,d\sigma_T(x,y) \leqslant D^T_{\boldsymbol{P}}(\mu). \tag{5.39}$$

On the other hand, ν_T is supported on $B^n/(3T)$ by definition, so

$$\frac{1}{2}\int\limits_{\mathbb{T}^n}\int\limits_{B^n/(3T)} Q_T(\boldsymbol{p})|\boldsymbol{p}-\boldsymbol{P}|^2\,d\nu_T(x,\boldsymbol{p}) = \frac{1}{2}\int\limits_{\mathbb{T}^n}\int\limits_{\mathbb{R}^n} |\boldsymbol{p}-\boldsymbol{P}|^2\,d\nu_T(x,\boldsymbol{p}). \tag{5.40}$$

On the other hand

$$\lim_{T\to 0}\int\limits_{\mathbb{T}^n}\int\limits_{\mathbb{R}^n} |\boldsymbol{p}-\boldsymbol{P}|^2\,d\nu_T(x,\boldsymbol{p}) \geqslant \int\limits_{\mathbb{T}^n}\int\limits_{\mathbb{R}^n} |\boldsymbol{p}-\boldsymbol{P}|^2\,d\nu_0(x,\boldsymbol{p}). \tag{5.41}$$

From(5.38)–(5.41) we obtain

$$\frac{1}{2}\int\limits_{\mathbb{T}^n}\int\limits_{\mathbb{R}^n} |\boldsymbol{p}-\boldsymbol{P}|^2\,d\nu_0(x,\boldsymbol{p}) \leqslant \lim_{T\to 0} D^T_{\boldsymbol{P}}(\mu),$$

hence

$$\lim_{T\to 0}\left[\frac{|\boldsymbol{P}|^2}{2} - D^T_{\boldsymbol{P}}(\mu)\right] \leqslant -\int\limits_{\mathbb{T}^n}\int\limits_{\mathbb{R}^n}\left(\frac{|\boldsymbol{p}|^2}{2} - \boldsymbol{P}\cdot\boldsymbol{p}\right)d\nu_0(x,\boldsymbol{p}). \tag{5.42}$$

Since we already proved that $\nu_0 \in \Lambda_\mu$, then Theorem 4.1 and (5.42) verify (5.28).

5.15. Proof of Theorem 4.6

To prove (4.28) we use (4.20) and (4.25) together with

$$\lim_{j\to\infty} D^T_{\boldsymbol{P}}(j,\Xi) = \min_{\mu\in\overline{\mathcal{M}}} D^T_{\boldsymbol{P}}(\mu,\Xi). \tag{5.43}$$

To establish (5.43) note, first, that $D^T_{\boldsymbol{P}}(j,\Xi) \geqslant \inf_{\mu\in\overline{\mathcal{M}}} D^T_{\boldsymbol{P}}(\mu,\Xi)$ for any j by definition, so it is enough to establish the inequality

$$\limsup_{j\to\infty} D^T_{\boldsymbol{P}}(j,\Xi) \leqslant \inf_{\mu\in\overline{\mathcal{M}}} D^T_{\boldsymbol{P}}(\mu,\Xi).$$

Let now $\{\mu_j\}$ be a sequence of empirical measures, where, for each j, μ_j contains exactly j atoms, and so that $\mu = \lim_{j\to\infty}\mu_j$ in C^*. Then $D^T_{\boldsymbol{P}}(\mu_j,\Xi) \geqslant D^T_{\boldsymbol{P}}(j,\Xi)$ by definition, while $\lim_{j\to\infty} D^T_{\boldsymbol{P}}(\mu_j,\Xi) = D^T_{\boldsymbol{P}}(\mu,\Xi)$ by Lemma 4.3.

References

[1] E.A. Arriola, J. Soler, A variational approach to the Schrödinger–Poisson system: Asymptotic behaviour, Breathers, and Stability, J. Stat. Phys. 103 (5–6) (2001) 1069–1106.
[2] S. Aubry, The twist map, the extended Frenkel–Kontrovna model and the devil's staircase, Physica D 7 (1983) 240–258.
[3] P. Bernard, B. Buffoni, Optimal mass transportation and mather theory, Preprint, http://arxiv.org/abs/math.DS/0412299.
[4] L.C. Evans, A survey of partial differential equations methods in weak KAM theory, Comm. Pure Appl. Math. 57 (4) (2004) 445–480.
[5] L.C. Evans, D. Gomes, Effective Hamiltonians and averaging for Hamiltonian dynamics. I, Arch. Ration. Mech. Anal. 157 (1) (2001) 1–33.

[6] L.C. Evans, D. Gomes, Effective Hamiltonians and averaging for Hamiltonian dynamics. II, Arch. Ration. Mech. Anal. 161 (4) (2002) 271–305.
[7] A. Fathi, The weak KAM Theorem in Lagrangian Dynamics, Cambridge Studies in Advanced Mathematics Series, vol. 88, Cambridge University Press, 2003.
[8] D. Gilbarg, N.S. Trudinger, Elliptic Partial Differential Equations of Second Order, Springer-Verlag, Berlin, 1977.
[9] D.A. Gomes, A.M. Oberman, Computing the effective Hamiltonian using a variational approach, SIAM J. Control Optim. 43 (3) (2004) 792–812.
[10] M. Govin, C. Chandre, H.R. Jauslin, Kolmogorov–Arnold–Moser-Renormalization-Group analysis of stability in Hamiltonian flows, Phys. Rev. Lett. 79 (1997) 3881–3884.
[11] L. Granieri, On action minimizing measures for the Monge–Kantorovich problem, Preprint, July 2004.
[12] G.A. Hedlund, Geodesics on a 3 dimensional Riemannian manifolds with periodic coefficients, Ann. of Math. 33 (1932) 719–739.
[13] J.B. Hiriart-Urruty, C. Lemarechal, Convex Analysis and Minimization Algorithms II, Grundlehren der Mathematischen Wissenschaften, vol. 306, Springer-Verlag, 1993 (Chapter 10).
[14] R. Illner, P.F. Zweifel, H. Lange, Global existence, uniqueness and asymptotic behaviour of solutions of the Wigner–Poisson and Schrödinger–Poisson systems, Math. Meth. Appl. Sci. 17 (1994) 349–376.
[15] J.B. Keller, Semiclassical mechanics, SIAM Rev. 27 (4) (1985) 485–504.
[16] D.P. Luigi, G.M. Stella, L. Granieri, Minimal measures, one-dimensional currents and the Monge–Kantorovich problem, Calc. Var. Partial Differential Equations 27 (1) (2006) 1–23.
[17] R. Mañè, On the minimizing measures of Lagrangian dynamical systems, Nonlinearity 5 (1992) 623–638.
[18] J.N. Mather, Existence of quasi-periodic orbits for twist homeomorphisms on the annulus, Topology 21 (1982) 457–467.
[19] J.N. Mather, Minimal measures, Comment. Math. Helv. 64 (1989) 375–394.
[20] J. Moser, Monotone twist mappings and the calculus of variations, Ergodic Theory Dynam. Systems 6 (1986) 401–413.
[21] J. Rubinstein, G. Wolansky, Eikonal functions: Old and new, in: D. Givoli, M.J. Grote, G. Papanicolaou (Eds.), A Celebration of Mathematical Modeling: The Joseph B. Keller Anniversary Volume, Kluwer, 2004.
[22] K.F. Siburg, The Principle of Least Action in Geometry and Dynamics, Lecture Notes in Mathematics, vol. 1844, Springer, 2004.
[23] C. Villani, Topics in Optimal Transportation, Graduate Studies in Math., vol. 58, Amer. Math. Soc., 2003.
[24] G. Wolansky, Optimal transportation in the presence of a prescribed pressure field, Preprint, arXiv: math-ph/0306070 v5, 12 Jan 2006.
[25] G. Wolansky, On time reversible description of the process of coagulation and fragmentation, Arch. Rat. Mech., submitted for publication.
[26] P. Markowich, G. Rein, G. Wolansky, Existence and nonlinear stability of stationary states of the Schrödinger–Poisson system, J. Stat. Phys. 106 (2002) 1221–1239.