Weak KAM Theory on the Wasserstein Torus with Multidimensional Underlying Space

WILFRID GANGBO
Georgia Institute of Technology

ADRIAN TUDORASCU
West Virginia University

Abstract

The study of asymptotic behavior of minimizing trajectories on the Wasserstein space $\mathscr{P}(\mathbb{T}^d)$ has so far been limited to the case $d = 1$ as all prior studies heavily relied on the isometric identification of $\mathscr{P}(\mathbb{T})$ with a subset of the Hilbert space $L^2(0,1)$. There is no known analogue isometric identification when $d > 1$. In this article we propose a new approach, intrinsic to the Wasserstein space, which allows us to prove a weak KAM theorem on $\mathscr{P}(\mathbb{T}^d)$, the space of probability measures on the torus, for any $d \geq 1$. This space is analyzed in detail, facilitating the study of the asymptotic behavior/invariant measures associated with minimizing trajectories of a class of Lagrangians of practical importance. © 2013 Wiley Periodicals, Inc.

Contents

Communications on Pure and Applied Mathematics, 0001–0056 (PREPRINT)

1 Introduction

It is now well-known that some systems of PDEs have an underlying Hamiltonian structure in the space of probability measures with finite second moments. It has been shown in [12] that such systems arise as Euler-Lagrange equations for action-minimizing trajectories associated with certain Lagrangians defined on the tangent bundle to the Wasserstein space [2]. However, most of the analysis performed in [12] (for example, on existence of these minimizing trajectories) was carried out in the case of one-dimensional underlying euclidean space and "mechanical" Lagrangians only. We note that the Lagrangian considered determines the choice of the cost function associated to the optimal transport distance used in the analysis (see [15], where the p-Wasserstein space appears). It is not our purpose here to work in full generality (as far as the Lagrangians considered); instead we shall restrict ourselves to the case of "mechanical" Lagrangians, which, at least from the applications point of view, may be the most interesting. It remains an interesting endeavor to extend the results of this paper to more general Lagrangians.

As far as the restriction to the one-dimensional setting is concerned, the main reason was an apparent lack of necessary compactness in general, which, solely in the $d = 1$ case, was offset by the isometric identification of the 2-Wasserstein space with the convex cone of nondecreasing $L^2(0,1)$-functions. Subsequent work [11, 13–15] upheld this restriction for the most part, although some interesting homogenization results were proved in the general case in [15]. The authors have searched for the correct setting and methodology to extend infinite-dimensional weak KAM results from the $\mathscr{P}_2(\mathbb{R})$ case developed in [13, 14] to the $\mathscr{P}_2(\mathbb{R}^d)$ case ever since those works were in progress.

In [13, 14] the manifold considered is an appropriate factorization of the set of monotone nondecreasing $L^2(0,1)$-functions. This is isometric with the space $\mathscr{P}(\mathbb{T})$ (Borel probabilities on the one-dimensional torus), which is compact in the topology induced by the $\mathbb{Z}$-periodic 2-Wasserstein distance (details in [13] and below). Said compactness remains true for the multidimensional torus $\mathbb{T}^d$, and we show that the tangent bundle to $\mathscr{P}(\mathbb{T}^d)$ is relatively compact with respect to an appropriate topology, so at least this important ingredient from classical, finite-dimensional weak KAM is there for us.

A natural definition of gradient in the 2-Wasserstein space, along with the functional indistinguishability of the support points for a probability given by an average of Dirac masses, yields that $\mathbb{R}^d$ fits the role of the de Rham cohomology group for $\mathscr{P}(\mathbb{T}^d)$. However, the similarities with the finite-dimensional case stop there. One of the main differences, which also constitutes an analytical challenge, is the nonuniqueness of the velocity field associated with an absolutely continuous curve in $\mathscr{P}(\mathbb{T}^d)$. Only one such velocity field (of minimal norm) lies in the tangent bundle, but this is not always the relevant one, so an extension of the tangent bundle is necessary. The relevant velocity (for a given absolutely continuous curve in $\mathscr{P}(\mathbb{T}^d)$) is the minimizer of the Lagrangian action. It turns out that for each $\mathbf{c} \in \mathbb{R}^d$ there is a unique $\mathbf{c}$-minimal velocity for an absolutely continuous curve on $\mathscr{P}(\mathbb{T}^d)$, all these velocities being distinguished by the choice of $\mathbf{c}$.

The *rotation vector* of a continuous path $[0, \infty) \ni t \to x(t) \in \mathbb{T}^d$ is the asymptotic limit of $\hat{x}(t)/t$, where $\hat{x}$ is a lift of the path x to the universal cover $\mathbb{R}^d$. Any two such lifts differ by a constant (independent of t) integer, so it is clear that the rotation vector (if it exists) is well-defined (independently of the lift). Things are different in the $\mathscr{P}(\mathbb{T}^d)$ case: first, due to the nonuniqueness of the velocity field along a curve, the notion of lift does not make sense in the context of lifting a continuous path of probabilities in $\mathscr{P}(\mathbb{T}^d)$ to a continuous path in $\mathscr{P}_2(\mathbb{R}^d)$. Instead, we have a well-defined notion of lift for a speed curve $(\sigma, \mathbf{v})$, where $\mathbf{v}$ is a velocity along the path σ. Second, we have not been able to prove that every speed curve can be lifted. However, that has no impact on our study, as we only need to lift pairs $(\sigma, \mathbf{v_c})$, where $\mathbf{v_c}$ is the $\mathbf{c}$-minimal periodic velocity of σ.

The treatment of some first-order nonlinear conservative systems of PDEs as Hamiltonian flows in the context of optimal transport was initiated by Ambrosio and Gangbo [1]. Given a Hamiltonian functional $\mathscr{H}$ defined on the Wasserstein space $\mathscr{P}_2(\mathbb{R}^{2d})$, a natural concept of gradient in $\mathscr{P}_2(\mathbb{R}^{2d})$ is used to make sense of the system

$$\partial_t \nu + \nabla_{(y,p)} \cdot [J \nabla_w \mathscr{H}(\nu)] = 0, \tag{1.1}$$

where J is a constant $2d \times 2d$ matrix so that $J\mathbf{a} \perp \mathbf{a}$ for all $\mathbf{a} \in \mathbb{R}^{2d}$. We have denoted by ∇_w the Wasserstein gradient (to be defined in what follows). When J is the block matrix $J = (0, -\mathrm{id}_d, 0, \mathrm{id}_d)$ we get kinetic systems such as the linear or nonlinear Vlasov, and Vlasov-Monge-Ampère. Under very general conditions on the Hamiltonian $\mathscr{H}$, the theory in [1] ensures existence of a solution to (1.1) for any prescribed initial condition $\nu|_{t=0} = \nu_0$. The paper by Gangbo et al. [9] justifies the terminology "Hamiltonian ODE" for the system (1.1), as the connection with conventional, finite-dimensional Hamiltonian ODEs may initially seem unclear. After all, the Hamiltonian $\mathscr{H}$ is not defined on the cotangent bundle as is customary, but on an extension (in some sense) of it.

When

$$\mathscr{H}(\nu) = \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} |p|^2 \nu(dy, dp) + \mathscr{K}(\mu)$$

where μ is the y-marginal of ν, the restriction of $\mathscr{H}$ to the set of measures of the form $\nu = (\mathrm{id} \times \zeta)_{\#}\mu$ induces the Hamiltonian

$$H(\mu, \zeta) = \frac{1}{2}\|\zeta\|_\mu^2 + \mathscr{K}(\mu). \tag{1.2}$$

It is perhaps not surprising that the Euler-Lagrange equation associated with this Hamiltonian is the monokinetic version of the one associated with the "full" Hamiltonian $\mathscr{H}$. Regarded as a function of the pair (μ, ζ), H appears as a natural generalization of a Hamiltonian defined on the cotangent bundle, a fact supported by the observation that given μ as an average of point masses $\mathbf{x}_1, \dots, \mathbf{x}_n \in \mathbb{R}^d$ and associated values $\mathbf{v}_i := \zeta(\mathbf{x}_i)$, we have

$$H(\mu, \Psi) = \frac{1}{2}|V|_n^2 + F(\mathbf{x}_1, \dots, \mathbf{x}_n),$$

where $V := (\mathbf{v}_1, \dots, \mathbf{v}_n) \in \mathbb{R}^{nd}$, $F(\mathbf{x}_1, \dots, \mathbf{x}_n) := W_2^2(\mu, \chi)/2$ (hereafter W_2 denotes the 2-Wasserstein distance on $\mathscr{P}_2(\mathbb{R}^d)$), and $|V|_n^2 := (|\mathbf{v}_1|^2 + \cdots + |\mathbf{v}_n|^2)/n$. This has the flavor of a finite-dimensional mechanical Hamiltonian. General Hamiltonians such as (1.2) have been considered in [11–15], with applications ranging from action-minimizing solutions for the Euler-Poisson system to globally minimizing trajectories of prescribed asymptotic behavior for the nonlinear Vlasov or Vlasov-Poisson on the torus. The Euler-Lagrange equation for the corresponding (to H) Lagrangian

$$L(\mu, \xi) = \frac{1}{2}\|\xi\|_\mu^2 - \mathscr{K}(\mu) \quad \text{for } \xi \in L^2(\mu)$$

reads

$$\partial_t(\sigma\mathbf{v}) + \nabla_x \cdot (\sigma\mathbf{v} \otimes \mathbf{v}) = -\sigma\nabla_w\mathscr{K}(\sigma) \quad \text{as distributions in } (0, T) \times \mathbb{T}^d. \tag{1.3}$$

This comes coupled with the continuity equation, which expresses the fact that $\mathbf{v}$ is a velocity field along the critical path σ,

$$\partial_t\sigma + \nabla_x \cdot (\sigma\mathbf{v}) = 0 \quad \text{as distributions in } (0, T) \times \mathbb{T}^d. \tag{1.4}$$

If $\mathscr{F} \equiv 0$, we deal with the pressureless Euler system. This is a special case of the monokinetic nonlinear Vlasov system, which is obtained from an even potential $\mathscr{F} \in C^1(\mathbb{T}^d)$ by setting

$$\mathscr{K}(\mu) := \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathscr{F}(x - y)\mu(dx)\mu(dy), \tag{1.5}$$

which yields

$$\nabla_w \mathscr{K}(\mu) = \int_{\mathbb{R}^d} \nabla \mathscr{F}(x-y)\mu(dy).$$

The attractive/repulsive Euler-Poisson system can also be formally brought within this framework by considering a singular potential $\mathscr{F}$ ($\pm\Phi_d$, where Φ_d is the fundamental solution for the Laplace equation in $\mathbb{R}^d$).

We would like to make clear from the beginning of this paper that our main goal here is the complete departure from the one-dimensional setting. To that end we embrace an intrinsic, optimal transport approach, where the analysis is performed directly in the Wasserstein space. The other main achievements are:

(1) we obtain existence of weakly invariant measures of a prescribed rotation vector,
(2) we prove a weak KAM theorem, and
(3) we obtain globally minimizing trajectories of a given rotation vector.

As mentioned early in this introduction, we are concerned with the infinite-dimensional analogue/extension of mechanical Hamiltonians defined over the cotangent bundles to finite-dimensional tori.

Earlier work by Cordero-Erausquin [5] introduced the Borel probability on the torus as the σ-finite measure

$$\widetilde{\mu} := \sum_{\mathbf{k}\in\mathbb{Z}^d} \mu(\cdot + \mathbf{k}),$$

where μ is a Borel probability measure on $\mathbb{R}^d$. The correspondence $\mu \to \widetilde{\mu}$ is not necessarily one-to-one. The set of all Borel probabilities with finite second moments that yield the same measure via the above procedure is identified with said measure. This way, any measure $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ can be considered a measure in $\mathscr{P}(\mathbb{T}^d)$ as well, indistinguishable from all other $\nu \in \mathscr{P}_2(\mathbb{R}^d)$ such that $\widetilde{\mu} = \widetilde{\nu}$. The quadratic Wasserstein distance on $\mathscr{P}_2(\mathbb{R}^d)$ induces canonically a distance on $\mathscr{P}(\mathbb{T}^d)$, called the periodic Wasserstein distance, which renders $\mathscr{P}(\mathbb{T}^d)$ compact. The details are in Section 2. Differentiable paths $t \to x(t)$ on a Riemannian manifold have tangent vectors $\dot{x}(t)$ belonging to the tangent space at $x(t)$, and the pair $\{(x(t), \dot{x}(t))\}_{t\in[a,b]}$ is called the speed curve [7] of the trajectory x. In anticipation of having a need for a similar notion in our context, we define a tangent bundle to $\mathscr{P}(\mathbb{T}^d)$ and an appropriate extension of it (where all velocities along a given curve lie). Section 2 continues with the introduction of differentiability of functionals defined on $\mathscr{P}(\mathbb{T}^d)$; then we prove that there is a natural notion of cohomology on $\mathscr{P}(\mathbb{T}^d)$.

Section 3 is dedicated to issues regarding the absolutely continuous curves on $\mathscr{P}(\mathbb{T}^d)$, such as: periodic velocity, periodic velocity of the $\mathbf{c}$-minimal norm, lifting of curves from $\mathscr{P}(\mathbb{T}^d)$ to $\mathscr{P}_2(\mathbb{R}^d)$, and rotation vectors associated to absolutely continuous speed curves. We show that speed curves with $\mathbf{c}$-minimal periodic velocities can be lifted. We would like to stress the importance of these $\mathbf{c}$-minimal velocities in the $\mathscr{P}(\mathbb{T}^d)$ context. It turns out (Section 4) that the $\mathbf{c}$-calibrated minimizing trajectories given by the weak KAM theorem (Theorem 4.14) have velocities of $\mathbf{c}$-minimal norm. For absolutely continuous curves on $\mathscr{P}_2(\mathbb{R}^d)$, this notion is trivial in the sense that the minimal norm velocity [2] coincides with the $\mathbf{c}$-minimal norm velocity for every $\mathbf{c} \in \mathbb{R}^d$. However, this is no longer valid in the periodic case, where the $\mathbf{c}$-minimal velocity does change with $\mathbf{c}$.

Given an initial condition $(\sigma_0, \mathbf{v}_0)$ existence of a flow (1.3)–1.4 remains a major challenge in fluid mechanics for $d \geq 2$. Only recently has the case $d = 1$ been settled in [3] in great generality. Thus, in Section 4 we define *weakly invariant* measures on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$, whose definition does not rely on the existence of a flow and is a natural and straightforward adaptation of the definition in [6]. For any $V \in \mathbb{R}^d$, we prove that there exists a weakly invariant measure that minimizes the Lagrangian action over $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$ among all weakly invariant measures with rotation vector V.

Mañé's argument [6] proving full-flow invariance for such measures can be easily extended to our case, provided that we have a well-defined flow. This provision is satisfied by the nonlinear Vlasov system for sufficiently smooth potentials (see Section 5). Next we prove a weak KAM theorem in the spirit of [7], and we show that optimizing curves in the variational definition of the Lax-Oleinik semigroup (called $\mathbf{c}$-calibrated curves) exist, and they have a rotation vector equal to $-\nabla \bar{H}(\mathbf{c})$ whenever the effective Hamiltonian $\bar{H}$ is twice differentiable at $\mathbf{c}$ in the sense of Alexandroff.

The Galilean invariant Hamiltonian corresponding to the nonlinear Vlasov system is provided as an example ($\bar{H}$ is twice differentiable everywhere). When $\mathscr{F} \in C^{1,1}(\mathbb{T})$ is even and an initial measure is prescribed, there is a uniquely defined solution to the Vlasov system on $\mathscr{P}(\mathbb{T}^d)$ (cf., e.g., [1] for existence and [18] for uniqueness), and so, we have a well-defined flow.

Section 5 contains an alternate, finite- to infinite-dimensional blowup approach in the case of a potential $\mathscr{F}$ that displays certain symmetry. It turns out that the $\mathbf{c}$-calibrated curves obtained for the monokinetic version satisfy the fully kinetic Vlasov system, which is intrinsically interesting (note that a distributional solution for the monokinetic system does not necessarily give rise to a distributional solution for the fully kinetic system). Imposing that solutions for the monokinetic system satisfy the fully kinetic Vlasov system can be viewed as a criterion for uniquely selecting special solutions for the monokinetic system. The probability measures carried by the ω-limits of these $\mathbf{c}$-calibrated curves are invariant (strongly) with respect to the Vlasov flow and have rotation vector $-\mathbf{c}$. We are able to draw stronger conclusions when $\mathscr{F}$ is even and attains its maximum at the origin. We use $\bar{H}(\mathbf{c}) = |\mathbf{c}|^2/2$ [15] to show that the $\mathbf{c}$-calibrated curves $(\sigma, \mathbf{v})$ satisfy

$$\left\| \frac{\mathrm{id}}{t} + \mathbf{c} \right\|_{L^2(\sigma_t;\mathbb{R}^d)} \leq \frac{C}{\sqrt{t}} + \left\| \frac{\mathrm{id}}{t} + \mathbf{c} \right\|_{L^2(\sigma_0;\mathbb{R}^d)}, \quad \lim_{t\to\infty} \|\mathbf{v}_t + \mathbf{c}\|_{L^2(\sigma_0;\mathbb{R}^d)} = 0.$$

Results from [15] are heavily cited in this section: as the weak KAM solutions are generally not unique (even in finite dimensions), we are not guaranteed that the weak KAM solution constructed in Section 4 is the dimensional blowup limit of finite-dimensional weak KAM solutions corresponding to appropriate discretizations of the Lagrangian. However, such solutions are crucial to our argument. Thus, in Section 5 we show (skipping details, as simply referring to [15] provides the whole picture) how to construct them on the basis of finite-dimensional weak KAM solutions, each of them coming from a standard approximation. Then, some **c**-calibrated curves are obtained as limits of **c**-calibrated curves corresponding to the finite-dimensional weak KAM solutions.

2 Geometry of $\mathscr{P}(\mathbb{T}^d)$

Some of the preliminary results from this section and Section 3 (such as existence of velocities along an absolutely continuous curve) can be found in a recent study by Gigli [16] on $\mathscr{P}_2(M)$, where M is a Riemannian manifold. However, the essence of what we cover here (**c**-minimality, lifting) is not discussed in said reference. Thus, we incorporate our take on the matters, as it also is conceptually and notationally subordinated to the present work.

2.1 A Characterization of the Wasserstein Torus $\mathscr{P}(\mathbb{T}^d)$

We consider the commutative group $\mathbb{Z}^d$ and its action on $\mathbb{R}^d$ given by $(k, x) \to x + k$, and let, as is customary, $\mathbb{T}^d := \mathbb{R}^d/\mathbb{Z}^d$. For each $x \in \mathbb{R}^d$ there is a unique $\overline{x} \in \mathscr{Q} := [0,1)^d$ that is equivalent to x. It is defined by $\overline{x} = (x_1 - \hat{x}_1, \ldots, x_d - \hat{x}_d)$, where $\hat{}\,: \mathbb{R} \to \mathbb{Z}$ is the greatest integer function. We sometimes denote $(\hat{x}_1, \ldots, \hat{x}_d)$ by $\hat{x}$ and identify $[x]$, the class of equivalence of x, with $\overline{x} \in \mathscr{Q} := [0,1)^d$. Recall that $\mathbb{T}^d$ is endowed with the metric $|\cdot|_{\mathbb{T}^d}$ defined by

$$|[x] - [y]|_{\mathbb{T}^d} = \min_{a\in[x],b\in[y]} |a - b|$$

for $x, y \in \mathbb{R}^d$. We also write $|x - y|_{\mathbb{T}^d}$ to mean $|[x] - [y]|_{\mathbb{T}^d}$ (the minus sign should be taken at face value, as $\mathbb{T}^d$ is not a vector space).

By a function $\zeta : \mathbb{T}^d \to \mathbb{R}$ (or periodic function) we mean a function $\zeta : \mathbb{R}^d \to \mathbb{R}$ that satisfies $\zeta(x) = \zeta(y)$ whenever $[x] = [y]$. We write $X : \mathbb{T}^d \to \mathbb{T}^d$ when X is a map of $\mathbb{R}^d$ into $\mathbb{R}^d$ that has the property

$$[x] = [y] \implies [X(x)] = [X(y)].$$

We define an equivalence relation on $\mathscr{P}_2(\mathbb{R}^d)$ (or, more generally, for any vector measures of finite total variation) by

$$\mu \sim \nu \iff \int_{\mathbb{R}^d} \zeta\, d\mu = \int_{\mathbb{R}^d} \zeta\, d\nu \quad \text{for all } \zeta \in C(\mathbb{T}^d).$$

The class of equivalence of $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ is denoted by $[\mu]$, and we denote the collection of all classes of equivalence by $\mathscr{P}(\mathbb{T}^d)$.

If $F : \mathscr{P}_2(\mathbb{R}^d) \to \mathbb{R}$ is invariant under the action of $\sim$ in the sense that $F(\mu) = F(\nu)$ whenever $\mu \sim \nu$, then we write $F : \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$. If $F \in C\big(\mathscr{P}_2(\mathbb{R}^d)\big)$, we write $F \in C\big(\mathscr{P}(\mathbb{T}^d)\big)$.

Given $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ there is a measure $\bar{\mu}$ supported by $\mathscr{Q}$ that lies in $[\mu]$. It is the pushforward of μ by $x \to \bar{x}$ and can be written as

$$\bar{\mu}(A) = \sum_{k \in \mathbb{Z}^d} \mu(A + k)$$

for all Borel sets $A \subset \mathscr{Q}$.

Let $\# : \mathscr{P}_2(\mathbb{R}^d) \to \mathscr{P}_2(\mathbb{R}^d)$ be the pushforward operator. Note that if $X : \mathbb{T}^d \to \mathbb{T}^d$ is a Borel map and $F \in C(\mathbb{T}^d)$, then $F \circ X : \mathbb{T}^d \to \mathbb{R}$. If μ and ν are equivalent in $\mathscr{P}_2(\mathbb{R}^d)$, then so are $X_\#\mu$ and $X_\#\nu$. We define $X_\#[\mu]$ to be $[X_\#\mu]$.

2.2 The Wasserstein Metric on $\mathscr{P}(\mathbb{T}^d)$

Given $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$, we set

$$\mathscr{W}^2(\mu, \nu) = \inf_{\gamma \in \Gamma(\mu,\nu)} \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - y|^2_{\mathbb{T}^d} \gamma(dx, dy),$$

where $\Gamma(\mu, \nu)$ is the set of all joint probability distributions with marginals μ and ν (in this order). According to [10],

$$\sup_{(\zeta,\theta) \in \mathscr{U}_0} J(\zeta, \theta) = \sup_{(\zeta,\theta) \in \mathscr{U}_1} J(\zeta, \theta) = \mathscr{W}^2(\mu, \nu), \tag{2.1}$$

where

$$J(\zeta, \theta) = \int_{\mathbb{R}^d} \zeta \, d\mu + \int_{\mathbb{R}^d} \theta \, d\nu.$$

We have set $\mathscr{U}_0$ to be the set of (ζ, θ) such that $\zeta, \theta : \mathbb{R}^d \to [-\infty, \infty)$ are upper-semicontinuous and satisfy

$$\zeta(x) + \theta(y) \le |x - y|^2_{\mathbb{T}^d}$$

for all $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$. Also, $\mathscr{U}_1$ is the set of (ζ, θ) in $\mathscr{U}_0$ such that

$$\zeta(x) = \inf_{y_* \in \mathbb{R}^d} |x - y_*|^2_{\mathbb{T}^d} - \theta(y_*) \quad \text{and} \quad \theta(y) = \inf_{x_* \in \mathbb{R}^d} |x_* - y|^2_{\mathbb{T}^d} - \zeta(x_*) \tag{2.2}$$

for all $(x, y) \in \mathbb{R}^d \times \mathbb{R}^d$. One readily checks that if (2.2) holds, then $\zeta, \theta \in C(\mathbb{T}^d)$. Hence, if we denote by $\mathscr{U}$ the set of (ζ, θ) in $\mathscr{U}_1$ such that $\zeta, \theta \in C(\mathbb{T}^d)$, then

$$\sup_{(\zeta,\theta) \in \mathscr{U}} J(\zeta, \theta) = \mathscr{W}^2(\mu, \nu). \tag{2.3}$$

The expression on the left-hand side of (2.3) is unchanged if we replace (μ, ν) by (μ_*, ν_*) such that $\mu_* \in [\mu]$ and $\nu_* \in [\nu]$. In other words,

$$\mathscr{W}(\mu, \nu) = \mathscr{W}(\mu_*, \nu_*). \tag{2.4}$$

We define the distance between $[\mu]$ and $[\nu]$ to be $\mathscr{W}(\mu,\nu)$ and exploit Lemma 2.1 below to check that $\mathscr{W}$ is a metric on $\mathscr{P}(\mathbb{T}^d)$.

Given $t \in \mathbb{R}$, we define $m(t) = t - \widehat{t + 1/2}$ and notice that $|m(t)| = |t|_{\mathbb{T}^1}$. Thus, the range of m is $[-1/2, 1/2)$. As a consequence, if for $s \in \mathbb{R}$ we set $m^*(s,t) = s - m(s-t)$, then $|s - m^*(s-t)| = |s-t|_{\mathbb{T}^1}$. This suggests that we introduce the vector-valued Borel map $\mathscr{T} : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d \times \mathbb{R}^d$ defined by

$$\mathscr{T}(x,y) = (x, x - \tilde{m}(x-y)) \quad \text{where } \tilde{m}(z) := (m(z_1), \dots, m(z_d)) \text{ for } z \in \mathbb{R}^d.$$

Since the range of m is $[-1/2, 1/2)$ we conclude that the range of $\mathscr{T}_1 - \mathscr{T}_2$ is $[-1/2, 1/2)^d$, and so

$$|x-y|_{\mathbb{T}^d} = |\mathscr{T}_1(x,y) - \mathscr{T}_2(x,y)|. \tag{2.5}$$

We have denoted by $\mathscr{T}_2$ the second projection of $\mathscr{T}$ and by $\mathscr{T}_1$ the first projection of $\mathscr{T}$. The fact that

$$\mathscr{T}(x+k, y+l) = \mathscr{T}(x,y) + (k,k)$$

holds for all $k, l \in \mathbb{Z}^d$ yields that $\mathscr{T} : \mathbb{T}^d \times \mathbb{T}^d \to \mathbb{T}^d \times \mathbb{T}^d$.

LEMMA 2.1. *For any* $\mu_*, \nu_* \in \mathscr{P}_2(\mathbb{R}^d)$

$$\mathscr{W}(\mu_*, \nu_*) = \min_{\nu}\{W_2(\mu_*, \nu) : \nu \in [\nu_*]\}. \tag{2.6}$$

PROOF. As $|\cdot| \geq |\cdot|_{\mathbb{T}^d}$, we only need to prove that in (2.6) the expression on the left-hand side is greater than or equal to the expression on the right-hand side. Suppose $\nu \in [\nu_*]$ and let $\gamma \in \Gamma(\mu_*, \nu)$. Let γ_* be the pushforward of γ by $\mathscr{T}$, and let ν_0 be the pushforward of γ by $\mathscr{T}_2$. Then $\gamma_* \in \Gamma(\mu_*, \nu_0)$, and so

$$\begin{aligned} W_2^2(\mu_*, \nu_0) &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2 \gamma_*(dx, dy) \\ &= \int_{\mathbb{R}^d \times \mathbb{R}^d} |x - \mathscr{T}_2(x,y)|^2 \gamma(dx, dy) = \int_{\mathbb{R}^d \times \mathbb{R}^d} |x-y|^2_{\mathbb{T}^d} \gamma(dx, dy), \end{aligned}$$

where we have exploited (2.5). As γ is an arbitrary element of $\Gamma(\mu_*, \nu)$, we have then established that

$$W_2(\mu_*, \nu_0) \leq \mathscr{W}(\mu_*, \nu).$$

To conclude the proof, it suffices to show that $\nu_0 \in [\nu_*]$. Take $F \in C(\mathbb{T}^d)$. Since $t - m(t)$ is an integer, $\mathscr{T}_2(x,y) - y \in \mathbb{Z}^d$, and so $F \circ \mathscr{T}_2 = F$. Using these facts and the definition of ν_0, we conclude that $\nu_0 \in [\nu_*]$. □

Remark 2.2. Since $|\tilde{m}(x-y)| < \sqrt{d}/2$, if μ_* is supported by $[0,1]^d$, then ν_* must be supported by the ball of radius $R := 3\sqrt{d}/2$. To see this, one can check that $|x - \tilde{m}(x-y)| < R$ when $x \in [0,1)^d$.

LEMMA 2.3. $(\mathscr{P}(\mathbb{T}^d), \mathscr{W})$ *is compact.*

PROOF. Let $\{[\mu_n]\}_n \subset \mathscr{P}(\mathbb{T}^d)$. Then $\{\overline{\mu}_n\}_n$ is relatively compact in $\mathscr{P}(\mathbb{R}^d)$, and so it admits a subsequence that we still label $\{\overline{\mu}_n\}_n$ converging to some μ in $\mathscr{P}(\mathbb{R}^d)$. By Lemma 2.1 we deduce that $\{[\mu_n]\}_n$ converges to $[\mu]$ in $\mathscr{P}(\mathbb{T}^d)$. $\square$

COROLLARY 2.4. *For $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ we define*

$$\Gamma_{\text{per}}(\mu,\nu) := \Bigg\{\gamma \in \Gamma(\mu,\nu) : \mathscr{W}^2(\mu,\nu) = W_2^2(\mu,\nu) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2_{\mathbb{T}^d}\,\gamma(dx,dy)\Bigg\}.$$

By Lemma 2.1*, for any $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ there exists $\nu^* \in [\nu]$ such that $\Gamma_{\text{per}}(\mu,\nu^*)$ is nonempty.*

2.3 Tangent Bundle and Its Extension

Fix $\mu \in \mathscr{P}_2(\mathbb{R}^d)$. We denote by $L^2(\mu)$ the set of functions $\xi : \mathbb{R}^d \to \mathbb{R}^d$ that are square-integrable with respect to μ. We denote by $\langle\cdot,\cdot\rangle_\mu$ the standard inner product of $L^2(\mu)$ and by $\|\cdot\|_\mu$ its induced norm. We denote by $L^2(\mathbb{T}^d,\mu)$ the closure of $C^\infty(\mathbb{T}^d;\mathbb{R}^d)$ in $L^2(\mu)$, and by $\mathscr{T}_\mu\mathscr{P}(\mathbb{T}^d)$ the closure of $\nabla C^\infty(\mathbb{T}^d)$ in $L^2(\mu)$, where $\nabla C^\infty(\mathbb{T}^d)$ is the set of $\nabla\varphi$ such that $\varphi \in C^\infty(\mathbb{T}^d)$.

Let $\mu_1, \mu_2 \in \mathscr{P}_2(\mathbb{R}^d)$ and let $\xi_i \in L^2(\mu_i)$, $i = 1, 2$. We adopt the following definition:

$$(\mu_1,\xi_2) \sim (\mu_2,\xi_2) \iff \mu_1 \sim \mu_2 \quad \text{and} \quad \mu_1\xi_1 \sim \mu_2\xi_2.$$

We denote by $[\mu,\xi]$ the class of equivalence of (μ,ξ) with respect to $\sim$ and by

$$\mathscr{C}\mathscr{P}(\mathbb{T}^d) = \{[\mu,\xi] : \mu \in \mathscr{P}_2(\mathbb{R}^d),\ \xi \in L^2(\mathbb{T}^d,\mu)\}.$$

Remark 2.5. Note that if $\mu_1 \sim \mu_2$ and $\xi_i \in \mathscr{T}_{\mu_i}\mathscr{P}(\mathbb{T}^d)$, $i = 1, 2$, then

$$(\mu_1,\xi_1) \sim (\mu_2,\xi_2) \iff \langle\nabla F,\xi_1\rangle_{\mu_1} = \langle\nabla F,\xi_2\rangle_{\mu_2} \quad \text{for all } F \in C^\infty(\mathbb{T}^d).$$

PROOF. One implication is obvious. For the other, we start by taking a sequence $\{\phi_n\}_n \subset C^\infty(\mathbb{T}^d)$ such that $\nabla\phi_n \to \xi_1$ in $L^2(\mu_1)$. We have

$$\|\xi_2\|_{\mu_2}\|\nabla\phi_n\|_{\mu_1} = \|\xi_2\|_{\mu_2}\|\nabla\phi_n\|_{\mu_2} \geq \langle\nabla\phi_n,\xi_2\rangle_{\mu_2} = \langle\nabla\phi_n,\xi_1\rangle\mu_1.$$

Passing to the limit yields $\|\xi_2\|_{\mu_2} \geq \|\xi_1\|_{\mu_1}$. Likewise, we get the opposite inequality. Therefore, $\|\xi_2\|_{\mu_2} = \|\xi_1\|_{\mu_1}$, which implies

$$\|\nabla\phi - \xi_1\|_{\mu_1} = \|\nabla\phi - \xi_2\|_{\mu_2} \quad \text{for all } \phi \in C^\infty(\mathbb{T}^d).$$

Consequently, a sequence $\{\phi_n\}_n \subset C^\infty(\mathbb{T}^d)$ satisfies $\nabla\phi_n \to \xi_1$ in $L^2(\mu_1)$ if and only if $\nabla\phi_n \to \xi_2$ in $L^2(\mu_2)$. But $\langle F,\nabla\phi_n\rangle_{\mu_1} = \langle F,\nabla\phi_n\rangle_{\mu_2}$ for any $F \in C(\mathbb{T}^d;\mathbb{R}^d)$ and all n. Pass to the limit in n to finish the argument. $\square$

The *tangent bundle* to $\mathscr{P}(\mathbb{T}^d)$ will then be defined as

$$\mathscr{T}\mathscr{P}(\mathbb{T}^d) := \{[\mu, \xi] : \mu \in \mathscr{P}_2(\mathbb{R}^d),\ \xi \in \mathscr{T}_\mu \mathscr{P}(\mathbb{T}^d)\}.$$

Clearly, $\mathscr{T}\mathscr{P}(\mathbb{T}^d) \subset \mathscr{C}\mathscr{P}(\mathbb{T}^d)$. We point out that the above definitions are anticipatory of the next section, where we shall see that the periodic velocity of minimal norm for an absolutely continuous curve $t \to \mu_t$ on the torus (the counterpart in the periodic case of the object defined in [2]) lies in $\mathscr{T}_{\mu_t}\mathscr{P}(\mathbb{T}^d)$. Nevertheless, this velocity is, generally speaking, not the only periodic velocity associated to a given absolutely continuous curve; however, any periodic velocity $\mathbf{v}$ will satisfy $\mathbf{v}_t \in L^2(\mathbb{T}^d, \mu_t)$. This explains the need of extending $\mathscr{T}\mathscr{P}(\mathbb{T}^d)$ to $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$. What we see here is exactly analogous to the tangent bundle to $\mathscr{P}_2(\mathbb{R}^d)$ being a subset of $\bigcup_{\mu \in \mathscr{P}_2(\mathbb{R}^d)} \{\mu\} \times L^2(\mu)$ in the euclidean (nonperiodic) case [2].

Remark 2.6. Suppose $[\mu, \xi] = [\nu, \zeta] \in \mathscr{C}\mathscr{P}(\mathbb{T}^d)$.

(i) If $\mathbf{c} \in \mathbb{R}^d$, then $\langle \mathbf{c}, \xi\rangle_\mu = \langle \mathbf{c}, \zeta\rangle_\nu$.

(ii) By the definition of $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$ (also, see the proof of Remark 2.5), we have

$$\|\xi\|_\mu = \sup_\phi \langle \xi, \phi\rangle_\mu = \sup_\phi \langle \zeta, \phi\rangle_\nu = \|\zeta\|_\nu, \tag{2.7}$$

where the suprema are performed over the set of $\phi \in C(\mathbb{T}^d; \mathbb{R}^d)$ such that $\|\phi\|_\mu \leq 1$.

Remark 2.7. Let $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ be such that $\mu \sim \nu$.

(i) Recall that $\mathscr{T}_\mu \mathscr{P}_2(\mathbb{R}^d)$ denotes the closure of $\nabla C_c^\infty(\mathbb{R}^d)$ (gradients of C_c^∞-functions) in $L^2(\mu)$ [2]. We have $\mathscr{T}_\mu \mathscr{P}(\mathbb{T}^d) \subset \mathscr{T}_\mu \mathscr{P}_2(\mathbb{R}^d)$.

(ii) Assume $[\mu, \xi] = [\nu, \zeta] \in \mathscr{C}\mathscr{P}(\mathbb{T}^d)$. By Remark 2.6(ii), note that $\{X_n\}_n$ in $C^\infty(\mathbb{T}^d; \mathbb{R}^d)$ tends to ξ in $L^2(\mu)$ if and only if it tends to ζ in $L^2(\nu)$. Furthermore, for any function $F \in C(\mathbb{T}^d \times \mathbb{R}^d)$ with at most quadratic growth, we have

$$\int_{\mathbb{R}^d} F(x, \xi(x))\mu(dx) = \int_{\mathbb{R}^d} F(x, \zeta(x))\nu(dx). \tag{2.8}$$

Proof. We prove (i). Let $\xi \in \mathscr{T}_\mu \mathscr{P}(\mathbb{T}^d)$ and let $\epsilon > 0$. We are to show existence of a φ_ϵ in $C_c^\infty(\mathbb{R}^d)$ such that $\|\xi - \nabla\varphi_\varepsilon\|_\mu < \epsilon$. First, choose $\varphi \in C^\infty(\mathbb{T}^d)$ such that $\|\xi - \nabla\varphi\|_\mu < \epsilon/2$. Second, choose $n > 0$ such that

$$4\|\varphi\|_{C^1(\mathbb{T}^d)}\mu(B_n^c) < \epsilon.$$

Third, let $f_n \in C_c^\infty(\mathbb{R}^d)$ be such that $0 \leq f_n \leq 1$, $|\nabla f_n| \leq 1$, $f_n \equiv 1$ on B_n, and $f_n \equiv 0$ on B_{n+2}^c. Here, $B_n \subset \mathbb{R}^d$ is the open ball of radius n, centered at the origin. We have that $\varphi_\epsilon := \varphi f_n \in C_c^\infty(\mathbb{R}^d)$ and

$$\begin{aligned}\|\xi - \nabla\varphi_\epsilon\|_\mu &\leq \|\xi - \nabla\varphi\|_\mu + \|(f_n - 1)\nabla\varphi\|_\mu + \|\varphi\nabla f_n\|_\mu \\ &\leq \|\xi - \nabla\varphi\|_\mu + \|\varphi\|_{C^1(\mathbb{T}^d)}\mu(B_n^c) < \epsilon.\end{aligned}$$

This proves the remark. $\square$

2.4 A Separable Topology on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$

Let $\{\phi_i\}_{i=1}^{\infty} \subset C^1(\mathbb{T}^d)$ be a dense subset of $C(\mathbb{T}^d;\mathbb{R}^d)$ for the uniform convergence topology. We assume without loss of generality that $\phi_i \not\equiv 0$. For $[\mu_1,\xi_1]$ and $[\mu_2,\xi_2]$ in $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$, we define

$$\mathrm{Dist}([\mu_1,\xi_1],[\mu_2,\xi_2]) = \mathscr{W}(\mu_1,\mu_2) + \sum_{i=1}^{\infty} \frac{1}{2^i} |\langle \xi_1, \hat{\phi}_i\rangle_{\mu_1} - \langle \xi_2, \hat{\phi}_i\rangle_{\mu_2}|,$$

where $\hat{\phi}_i = \phi_i / \|\phi_i\|_{W^{1,\infty}(\mathbb{R}^d;\mathbb{R}^d)}$.

Let $\{[\mu_n,\xi_n]\}_n \subset \mathscr{C}\mathscr{P}(\mathbb{T}^d)$ and let $[\mu,\xi] \in \mathscr{C}\mathscr{P}(\mathbb{T}^d)$. When needed, we will assume that some of the following three possibilities hold:

$$\lim_{n\to\infty} \mathscr{W}(\mu_n,\mu) = 0, \tag{2.9}$$

$$\sup_n \|\xi_n\|_{\mu_n} < \infty, \tag{2.10}$$

$$\lim_{n\to\infty} \langle \psi, \xi_n\rangle_{\mu_n} = \langle \psi, \xi\rangle_{\mu} \quad \text{for all } \psi \in C(\mathbb{T}^d;\mathbb{R}^d). \tag{2.11}$$

Remark 2.8. Note that Dist defines a metric on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$. Furthermore, a sequence $\{[\mu_n,\xi_n]\}_n \subset \mathscr{C}\mathscr{P}(\mathbb{T}^d)$ satisfying (2.10) converges to $[\mu,\xi]$ in the metric Dist if and only if (2.9) and (2.11) hold.

LEMMA 2.9. *If* $C \geq 0$, *then* $\mathbb{B}_C := \{[\mu,\xi] : \|\xi\|_\mu \leq C\}$ *is a compact subset of* $(\mathscr{C}\mathscr{P}(\mathbb{T}^d), \mathrm{Dist})$.

PROOF. Let $\{[\mu_n,\xi_n]\}_n \subset \mathscr{C}\mathscr{P}(\mathbb{T}^d)$. By Lemma 2.3, $\{\mu_n\}_n$ is relatively compact in $\mathscr{P}(\mathbb{T}^d)$, and so, extracting a subsequence if necessary, we may assume existence of a μ such that $\{\mu_n\}_n$ converges to μ in $\mathscr{P}(\mathbb{T}^d)$. Using Lemma 2.1 and Remark 2.2, we may assume without loss of generality that

$$\mathscr{W}(\mu_n,\mu) = W_2(\mu_n,\mu),$$

μ is supported by $[0,1]^d$, and μ_n is supported by B_R, where $R = 3\sqrt{d}/2$. By [2, theorem 5.4.4], there exists $\bar{\xi} \in L^2(\mu)$ such that (up to a subsequence)

$$\lim_{n\to\infty} \langle \Psi, \xi_n\rangle_{\mu_n} = \langle \Psi, \bar{\xi}\rangle_{\mu} \quad \text{for all } \Psi \in C(\mathbb{R}^d;\mathbb{R}^d). \tag{2.12}$$

We conclude that (2.11) holds by setting ξ to be the projection of $\bar{\xi}$ onto $L^2(\mathbb{T}^d,\mu)$. We use Remark 2.8 and the fact that $\|\xi_n\|_{\mu_n} \leq C$ to conclude the proof of the lemma. $\square$

An immediate consequence is the following:

COROLLARY 2.10. *The metric space* $(\mathscr{C}\mathscr{P}(\mathbb{T}^d), \mathrm{Dist})$ *is separable.*

2.5 Optimal Couplings and Their Properties

Throughout the remainder of this paper, $\Gamma_o(\mu,\nu)$ will denote the set of optimal couplings (with respect to W_2) between μ and ν.

LEMMA 2.11. *Suppose that*

$$\mu_0 = \frac{1}{N}\sum_{i=1}^{N}\delta_{x_i}, \quad \bar{\mu}_1 = \frac{1}{N}\sum_{i=1}^{N}\delta_{\bar{y}_i}, \quad \textit{and} \quad \mathscr{W}^2(\mu_0,\bar{\mu}_1) = \frac{1}{N}\sum_{i=1}^{N}|\bar{y}_i - x_i|^2_{\mathbb{T}^d}.$$

Then there exists $y_1,\dots,y_N \in \mathbb{R}^d$ such that

$$\mu_1 =: \frac{1}{N}\sum_{i=1}^{N}\delta_{y_i} \in [\mu_0] \quad \textit{and} \quad W_2(\mu_1,\bar{\mu}_1) \le \mathscr{W}(\mu_0,\bar{\mu}_1).$$

PROOF. We define y_i by setting its k^{th} component y_i^k to be

$$y_i^k = \bar{y}_i^k - \left|\bar{y}_i^k - x_i^k\right|_{\mathbb{T}^1}.$$

We have

$$y_i^k - x_i^k = \left(\bar{y}_i^k - x_i^k\right) - \left|\bar{y}_i^k - x_i^k\right|_{\mathbb{T}^1} \in \mathbb{N},$$

and so $\mu_1 =: 1/N\sum_{i=1}^{N}\delta_{y_i} \in [\mu_0]$. Note that

$$\begin{aligned} W_2^2(\mu_1,\bar{\mu}_1) &\le \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{d}\left|\bar{y}_i^k - y_i^k\right|^2 \\ &= \frac{1}{N}\sum_{i=1}^{N}\sum_{k=1}^{d}\left|\bar{y}_i^k - x_i^k\right|^2_{\mathbb{T}^1} = \mathscr{W}^2(\mu_0,\bar{\mu}_1). \end{aligned}$$ □

PROPOSITION 2.12. *Suppose that $\mu_0,\nu_0,\mu_1 \in \mathscr{P}_2(\mathbb{R}^d)$ are such that $\mu_1 \in [\mu_0]$ and there exists $\gamma \in \Gamma_{\text{per}}(\mu_0,\nu_0)$. Then there exists $\nu_1 \in [\nu_0]$ and $g \in \Gamma_{\text{per}}(\mu_1,\nu_1)$ such that*

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} (x-y)\cdot F(x,y)g(dx,dy) = \int_{\mathbb{R}^d\times\mathbb{R}^d} (x-y)\cdot F(x,y)\gamma(dx,dy) \tag{2.13}$$

for every $F \in C(\mathbb{T}^d\times\mathbb{T}^d;\mathbb{R}^d)$.

PROOF. We choose an increasing sequence of integers $\{N_n\}_n$ and, for each n, a set of points $\{(x_i^n,y_i^n)\}_{i=1}^{N_n} \subset \text{spt}[\gamma]$ such that

$$\lim_{n\to\infty}\int_{\mathbb{R}^d\times\mathbb{R}^d} f(x,y)\gamma^n(dx,dy) = \int_{\mathbb{R}^d\times\mathbb{R}^d} f(x,y)\gamma(dx,dy) \tag{2.14}$$

for any $f \in C(\mathbb{R}^d \times \mathbb{R}^d)$ for which there exists $C > 0$ such that $|f(x,y)| \leq C(1+|x|^2+|y|^2)$ for all $x, y \in \mathbb{R}^d$. Here

$$\gamma^n := \frac{1}{N_n}\sum_{i=1}^{N_n} \delta_{(x_i^n, y_i^n)}.$$

(Note that (2.14) simply expresses the fact that γ^n converges to γ in the Wasserstein space $\mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$.) Let μ_0^n and ν_0^n be the x- and y-marginals of γ^n, respectively. Now (2.14) implies

$$\lim_{n\to\infty} W_2(\mu_0, \mu_0^n) = 0 \quad \text{and} \quad \lim_{n\to\infty} W_2(\nu_0, \nu_0^n) = 0.$$

Since $\gamma \in \Gamma_{\text{per}}(\mu_0, \nu_0)$ and $\{(x_i^n, y_i^n)\} \subset \text{spt}[\gamma]$, the set $\{(x_i^n, y_i^n)\}_{i=1}^{N_n}$ is cyclically monotone (as it belongs to the graph of the subdifferential of a convex function [19]), and so

$$W_2^2(\mu_0^n, \nu_0^n) = \frac{1}{N_n}\sum_{i=1}^{N_n} |x_i^n - y_i^n|^2 = \frac{1}{N_n}\sum_{i=1}^{N_n} |x_i^n - y_i^n|^2_{\mathbb{T}^d} \leq \mathscr{W}^2(\mu_0^n, \nu_0^n).$$

The fact that $\mathscr{W} \leq W_2$ yields

$$\begin{aligned} W_2^2(\mu_0^n, \nu_0^n) &= \frac{1}{N_n}\sum_{i=1}^{N_n} |x_i^n - y_i^n|^2 \\ &= \frac{1}{N_n}\sum_{i=1}^{N_n} |x_i^n - y_i^n|^2_{\mathbb{T}^d} = \mathscr{W}^2(\mu_0^n, \nu_0^n). \end{aligned} \tag{2.15}$$

Next, choose a sequence of points $\{\bar{y}_i^n\} \subset \mathbb{R}^d$ such that

$$\lim_{n\to\infty} W_2(\mu_1, \bar{\mu}_1^n) = 0 \quad \text{where } \bar{\mu}_1^n = \frac{1}{N_n}\sum_{i=1}^{N_n} \delta_{\bar{y}_i^n}. \tag{2.16}$$

By Lemma 2.11 there exist $z_1^n, \ldots, z_{N_n}^n \in \mathbb{R}^d$ such that

$$\mu_1^n =: \frac{1}{N_n}\sum_{i=1}^{N_n} \delta_{z_i^n} \in [\mu_0^n] \quad \text{and} \quad W_2(\mu_1^n, \bar{\mu}_1^n) \leq \mathscr{W}(\mu_0^n, \bar{\mu}_1^n). \tag{2.17}$$

Reordering the points if necessary, we may assume that $z_i^n \in [x_i^n]$. Define $k_i^n := z_i^n - x_i^n \in \mathbb{Z}^d$ and set

$$\nu_1^n := \frac{1}{N_n}\sum_{i=1}^{N_n} \delta_{y_i^n + k_i^n}, \quad g^n =: \frac{1}{N_n}\sum_{i=1}^{N_n} \delta_{(z_i^n, y_i^n + k_i^n)}.$$

We have $\nu_1^n \in [\nu_0^n]$ and

$$W_2^2(\mu_1^n, \nu_1^n) \le \frac{1}{N_n}\sum_{i=1}^{N_n} |z_i^n - (y_i^n + k_i^n)|^2$$
$$= \frac{1}{N_n}\sum_{i=1}^{N_n} |x_i^n - y_i^n|^2_{\mathbb{T}^d} = \mathscr{W}^2(\mu_0^n, \nu_0^n) = \mathscr{W}^2(\mu_1^n, \nu_1^n).$$

We have used (2.15) to obtain the first two inequalities in the previous display. This, together with the fact that $\mathscr{W} \le W_2$, implies $g^n \in \Gamma_{\rm per}(\mu_1^n, \nu_1^n)$ and

$$W_2^2(\mu_1^n, \nu_1^n) = \mathscr{W}^2(\mu_1^n, \nu_1^n) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2_{\mathbb{T}^d}\, g^n(dx, dy). \tag{2.18}$$

If $F \in C(\mathbb{T}^d \times \mathbb{T}^d; \mathbb{R}^d)$, then

$$\frac{1}{N_n}\sum_{i=1}^{N_n}(y_i^n + k_i^n - z_i^n)\cdot F(z_i^n, y_i^n + k_i^n) = \frac{1}{N_n}\sum_{i=1}^{N_n}(y_i^n - x_i^n)\cdot F(x_i^n, y_i^n),$$

or, equivalently,

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} (y-x)\cdot F(x,y) g^n(dx,dy) = \int_{\mathbb{R}^d\times\mathbb{R}^d} (y-x)\cdot F(x,y)\gamma^n(dx,dy). \tag{2.19}$$

We use the inequality in (2.17) and then the triangle inequality to obtain

$$W_2(\mu_1^n, \bar\mu_1^n) \le \mathscr{W}(\mu_0^n, \bar\mu_1^n) \le \mathscr{W}(\mu_0^n, \mu_1) + \mathscr{W}(\mu_1, \bar\mu_1^n)$$
$$= \mathscr{W}(\mu_0^n, \mu_0) + \mathscr{W}(\mu_1, \bar\mu_1^n).$$

We exploit this and the inequality $\mathscr{W} \le W_2$ to obtain

$$W_2(\mu_1^n, \mu_1) \le W_2(\mu_1^n, \bar\mu_1^n) + W_2(\bar\mu_1^n, \mu_1) \le W_2(\mu_0^n, \mu_0) + 2W_2(\mu_1, \bar\mu_1^n). \tag{2.20}$$

The uniform bound on the second moments of the g^n and ν_1^n allows us to find a subsequence of $\{g^n\}_n$ converging narrowly to some $g \in \mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ and a subsequence of $\{\nu_1^n\}_n$ converging narrowly to some $\nu_1 \in \mathscr{P}_2(\mathbb{R}^d)$. Without loss of generality, let us assume that the whole sequences converge. By (2.20), $\{\mu_1^n\}_n$ converges to μ_1 in $\mathscr{P}_2(\mathbb{R}^d)$. Hence, $g \in \Gamma(\mu_1, \nu_1)$. We use the lower semicontinuity property of

$$\mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d) \ni \eta \to \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \eta(dx,dy)$$

with respect to the narrow convergence to obtain

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 g(dx,dy) \le \liminf_{n\to\infty} \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 g^n(dx,dy)$$

$$= \liminf_{n\to\infty} \mathscr{W}^2(\mu_1^n, \nu_1^n) = \mathscr{W}^2(\mu_1, \nu_1).$$

We have used (2.18) to obtain the first equality above, and then we have used the continuity of $\mathscr{W}$ with respect to narrow convergence to obtain the second equality. Since $\mathscr{W} \le W_2$, we conclude that $g \in \Gamma_{\mathrm{per}}(\mu_1, \nu_1)$. Letting n tend to ∞ in (2.19), we conclude that (2.13) holds. $\square$

2.6 Differentials of Functions on $\mathscr{P}(\mathbb{T}^d)$

PROPOSITION 2.13. *Let* $[\mu_0, \xi_0] = [\mu_1, \xi_1] \in \mathscr{C}\mathscr{P}(\mathbb{T}^d)$. *Let* $\nu_0 \in \mathscr{P}_2(\mathbb{R}^d)$ *be such that* $\Gamma_{\mathrm{per}}(\mu_0, \nu_0)$ *is nonempty and let* $\gamma_0 \in \Gamma_{\mathrm{per}}(\mu_0, \nu_0)$. *Then there exist* $\nu_1 \in [\nu_0]$ *and* $\gamma_1 \in \Gamma_{\mathrm{per}}(\mu_1, \nu_1)$ *such that*

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \xi_0(x)\cdot(y-x)\gamma_0(dx,dy) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \xi_1(x)\cdot(y-x)\gamma_1(dx,dy).$$

PROOF. Since $\xi_0 \in L^2(\mathbb{T}^d, \mu_0)$, there exists a sequence $\{\phi_n\}_n \subset C(\mathbb{T}^d; \mathbb{R}^d)$ converging to it in $L^2(\mu_0)$. By Lemma 2.12, we get a measure $\nu_1 \in [\nu_0]$ and an optimal coupling $\gamma_1 \in \Gamma_{\mathrm{per}}(\mu_1, \nu_1)$ such that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \phi_n(x)\cdot(y-x)\gamma_0(dx,dy) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \phi_n(x)\cdot(y-x)\gamma_1(dx,dy) \quad \text{for all } n.$$

The x-marginal of γ_0 is μ_0, which implies ϕ_n converges to ξ_0 in $L^2(\gamma_0)$ as well. Likewise, the limit ξ_1 of $\{\phi_n\}_n$ in $L^2(\mu_1)$ is also the limit in $L^2(\gamma_1)$. Thus, the equality displayed above finishes the proof (by passage to the limit). $\square$

DEFINITION 2.14. Let $U : \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ and $\mu \in \mathscr{P}_2(\mathbb{R}^d)$.

(i) We say that $\xi \in L^2(\mathbb{T}^d, \mu)$ is in the differential of U at μ, and we write $\xi \in \partial U(\mu)$, if

$$\sup_{\nu,\gamma} \left| U(\nu) - U(\mu) - \int_{\mathbb{R}^d\times\mathbb{R}^d} \xi(x)\cdot(y-x)\gamma(dx,dy) \right| = o(\mathscr{W}(\mu,\nu)). \tag{2.21}$$

We denote by $\partial U(\mu)$ the differential of U at μ. The supremum is performed over the set of $\nu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\gamma \in \Gamma_{\mathrm{per}}(\mu, \nu)$.

(ii) If $\partial U(\mu)$ is nonempty, then it is a convex set and so it has a unique element of minimal norm, which we denote by $\nabla_w U(\mu)$ and refer to as the Wasserstein gradient of U.

(iii) If $\partial U(\mu)$ is nonempty for all $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\mu \to (\mu, \nabla_w U(\mu))$ is continuous, we say that U is continuously differentiable, and we write $U \in C^1(\mathscr{P}(\mathbb{T}^d))$.

Remark 2.15. Let $[\mu, \xi]$ be an element of $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$.

(i) If $\overline{\xi} \in L^2(\mathbb{T}^d, \mu)$ and $\xi, \overline{\xi} \in \partial U(\mu)$, then $\overline{\xi} - \xi$ is in the orthogonal complement of $\mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d)$ in $L^2(\mathbb{T}^d, \mu)$, denoted by $[\mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d)]^\perp$.

(ii) By Proposition 2.13, if $[\overline{\mu}, \overline{\xi}]$ is equivalent to $[\mu, \xi]$, then $\xi \in \partial U(\mu)$ if and only if $\overline{\xi} \in \partial U(\overline{\mu})$.

(iii) Let $\eta \in [\mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d)]^\perp$, $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, and $\gamma \in \Gamma_o(\mu, \nu)$. According to the proof of Proposition 2.13, one concludes that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \eta(x) \cdot (y - x) \gamma(dx, dy) = 0.$$

(iv) If ξ is in $\partial U(\mu)$, then so is its projection $\overline{\xi}$ onto $\mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d)$. As a consequence, $\overline{\xi} = \nabla_w U(\mu)$, and so either $\partial U(\mu)$ is empty or it is of the form $\nabla_w U(\mu) + [\mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d)]^\perp$.

PROOF. We only prove (i) and (iv). For $\varphi \in C^\infty(\mathbb{T}^d)$, define

$$\gamma_t := [\mathrm{id} \times (\mathrm{id} + t\nabla\varphi)]_\# \mu, \quad \sigma_t := (\mathrm{id} + t\nabla\varphi)_\# \mu.$$

For t small enough, $\gamma_t \in \Gamma_o(\mu, \sigma_t)$. If $\xi, \overline{\xi} \in \partial U(\mu)$, then

$$|\langle \overline{\xi} - \xi, \nabla\varphi \rangle_\mu| = \frac{1}{|t|} o(\mathcal{W}_2(\mu, \sigma_t)) = \frac{1}{|t|} o(t \|\nabla\varphi\|_\mu).$$

Letting t tend to 0 and using that φ is arbitrary in $C^\infty(\mathbb{T}^d)$, we conclude the proof of (i).

If $\xi \in \partial U(\mu)$, $\nu \in \mathcal{P}_2(\mathbb{R}^d)$, and $\gamma \in \Gamma_{\mathrm{per}}(\mu, \nu)$, since $\xi - \overline{\xi}$ is orthogonal to $\mathcal{T}_\mu \mathcal{P}(\mathbb{T}^d)$, we have (by (iii))

$$U(\nu) - U(\mu) \int_{\mathbb{R}^d \times \mathbb{R}^d} \xi(x) \cdot (y - x) \gamma(dx, dy) = \\ U(\nu) - U(\mu) - \int_{\mathbb{R}^d \times \mathbb{R}^d} \overline{\xi}(x) \cdot (y - x) \gamma(dx, dy).$$

This, together with (2.21) implies

$$\sup_{\nu, \gamma} \left| U(\nu) - U(\mu) - \int_{\mathbb{R}^d \times \mathbb{R}^d} \overline{\xi}(x) \cdot (y - x) \gamma(dx, dy) \right| = o\big(\mathcal{W}(\mu, \nu)\big),$$

the supremum being performed over the set of $\nu \in \mathcal{P}_2(\mathbb{R}^d)$ and $\gamma \in \Gamma_{\mathrm{per}}(\mu, \nu)$. This proves (iv). □

LEMMA 2.16. *All functions in $C^1(\mathcal{P}(\mathbb{T}^d))$ are Lipschitz-continuous.*

PROOF. Let $U \in C^1(\mathcal{P}(\mathbb{T}^d))$ and set

$$\Lambda_\mu \phi := \langle \phi, \nabla_w U(\mu) \rangle_\mu \quad \text{for } \mu \in \mathcal{P}(\mathbb{T}^d),\ \phi \in C(\mathbb{T}^d; \mathbb{R}^d).$$

By Definition 2.14(iii), $\Lambda_\mu \phi$ is continuous as a function of μ for fixed ϕ. As μ lies in a compact set, we deduce the family $\{\Lambda_\mu \phi\}_{\mu \in \mathcal{P}(\mathbb{T}^d)}$ is bounded in $\mathbb{R}$ for every $\phi \in C(\mathbb{T}^d;\mathbb{R}^d)$. But, for every $\mu \in \mathscr{P}(\mathbb{T}^d)$, Λ_μ is a bounded, linear operator from $C(\mathbb{T}^d;\mathbb{R}^d)$ (with the sup norm) into $\mathbb{R}$, whose norm we claim equals $\|\nabla_w U(\mu)\|_{L^1(\mu;\mathbb{R}^d)}$. According to the uniform boundedness principle, these norms are uniformly (with respect to μ) bounded, which is the goal of the first part of this proof.

Before moving to the next part, let us justify our claim by taking a sequence $\{\psi_n\}_n \subset C(\mathbb{T}^d;\mathbb{R}^d)$ that converges to $\nabla_w U(\mu)$ in $L^2(\mu)$, chosen such that

$$\sup_{\substack{\phi\in C(\mathbb{T}^d,\mathbb{R}^d)\\ \|\phi\|_\infty\le 1}} \int_{\mathbb{R}^d} |\phi\cdot[\nabla_w U(\mu)-\psi_n]|d\mu \le \frac{1}{n}. \tag{2.22}$$

But it is easy to see that the norm of the linear functional over $C(\mathbb{T}^d;\mathbb{R}^d)$ given by $\phi \to \langle\phi,\psi_n\rangle_\mu$ is $\|\psi_n\|_{L^1(\mu;\mathbb{R}^d)}$. We then use

$$\langle\phi,\psi_n\rangle_\mu = \langle\phi,\nabla_w U(\mu)\rangle_\mu + \langle\phi,\psi_n-\nabla_w U(\mu)\rangle_\mu,$$

take the supremum of both sides over all $\phi \in C(\mathbb{T}^d;\mathbb{R}^d)$ of sup norm at most 1, then use (2.22) to finish the proof of the claim.

Let n be a positive integer. Set $U^n(\mathbf{x}^n) := U(\mu_{\mathbf{x}^n})$, where $\mathbf{x}^n := (x_1^n,\dots,x_n^n) \in \mathbb{T}^{nd}$ and $\mu_{\mathbf{x}^n} := (1/n)\sum_{i=1}^n \delta_{x_i^n}$. One can check that $U^n \in C^1(\mathbb{T}^{nd})$ and

$$\nabla_{x_i^n} U^n(\mathbf{x}^n) = \frac{1}{n}\nabla_w U(\mu_{\mathbf{x}^n})(x_i^n).$$

The uniform (with respect to n and $\mathbf{x}^n$) bound on

$$\|\nabla_w U(\mu_{\mathbf{x}^n})\|_{L^1(\mu_{\mathbf{x}^n};\mathbb{R}^d)} = \frac{1}{n}\sum_{i=1}^n |\nabla_w U(\mu_{\mathbf{x}^n})(x_i^n)|$$

shows that the l^1-norm of the gradient of U^n is bounded; thus U^n is Lipschitz on $\mathbb{T}^{nd}$ and its Lipschitz constant is independent of n. We conclude by using the density of averages of Dirac masses in $\mathscr{P}(\mathbb{T}^d)$ and the continuity of U. □

2.7 Cohomology

The goal of this subsection is to prove that if $\nabla_w U \in C(\mathscr{P}(\mathbb{T}^d))$, then there exists a unique $\tilde{U} \in C^1(\mathscr{P}(\mathbb{T}^d))$ and a unique $\mathbf{c} \in \mathbb{R}^d$ such that

$$U(\mu) = \tilde{U}(\mu) + \mathbf{c}\cdot\int_{\mathbb{R}^d} x\mu(dx).$$

To see this rather quickly, we define the functions $U^n : \mathbb{R}^{nd} \to \mathbb{R}$ by

$$U^n(\mathbf{x}^n) = U(\mu_{\mathbf{x}^n}) \quad \text{where } \mu_{\mathbf{x}^n} := \frac{1}{n}\sum_{i=1}^n \delta_{x_i^n}.$$

As mentioned earlier, one can show that U^n is differentiable and $\nabla_{x_i} U^n(\mathbf{x}^n) = \frac{1}{n}\nabla_w U(\mu_{\mathbf{x}^n})(x_i)$. Thus, ∇U^n is $(\mathbb{Z}^d)^n$-periodic, which implies the existence of a unique periodic differentiable $\widetilde{U}^n : \mathbb{R}^{nd} \to \mathbb{R}$ and of a unique $\mathbf{c}^n \in \mathbb{R}^{nd}$ such that

$$U^n(\mathbf{x}^n) = \widetilde{U}^n(\mathbf{x}^n) + \frac{1}{n}\mathbf{c}^n \cdot \mathbf{x}^n, \tag{2.23}$$

where the $\cdot$ denotes the euclidean inner product in $\mathbb{R}^{nd}$. The permutation invariance of U^n also readily implies that $\mathbf{c}^n$ actually consists of n copies of a vector $\mathbf{c} \in \mathbb{R}^d$, i.e., $\mathbf{c}^n = (\mathbf{c}, \dots, \mathbf{c})$. There is also consistency in the sense that $\mathbf{c}$ is independent of n. Indeed, one can see that by considering $\mathbf{x}^n = (x, \dots, x)$ for some $x \in \mathbb{R}^d$ to discover that $\mathbf{c}$ is, in fact, nothing but $\mathbf{c}^1$ from (2.23) with $n = 1$. With this $\mathbf{c}$ at hand, let

$$\widetilde{U}(\mu) := U(\mu) - \mathbf{c} \cdot \int_{\mathbb{R}^d} x\mu(dx),$$

which is clearly a real-valued, differentiable map on $\mathscr{P}_2(\mathbb{R}^d)$. We also have $\widetilde{U}(\mu_{\mathbf{x}^n}) = \widetilde{U}^n(\mathbf{x}^n)$, so the restriction of $\widetilde{U}$ to the set of averages of Dirac masses is periodic. But this set is dense in $\mathscr{P}_2(\mathbb{R}^d)$, which implies that $\widetilde{U}$ is $\mathscr{P}(\mathbb{T}^d)$-periodic.

PROPOSITION 2.17. *There are closed forms on $\mathscr{P}(\mathbb{T}^d)$ that are not exact.*

PROOF. Let $\mathbf{c} \in \mathbb{R}^d$ and $[\mu] \in \mathscr{P}(\mathbb{T}^d)$. For $\varphi \in C^\infty(\mathbb{T}^d)$, setting $X = \nabla\varphi$ we define

$$\Lambda_\mu(X) = \langle \mathbf{c}, X\rangle_\mu.$$

One can check that $d\Lambda = 0$, and so Λ is a closed form (cf. [9]). Let $\gamma(t) = t\mathbf{a}$ where $\mathbf{a} \in \mathbb{R}^{\mathbf{d}}$ has integer components. Then γ is a closed curve in $\mathscr{P}(\mathbb{T}^d)$. But

$$\int_\gamma \Lambda = \mathbf{c} \cdot \mathbf{a} \neq 0$$

if $\mathbf{a} \in \mathbb{R}^d$ is not perpendicular to $\mathbf{c}$, and so Λ is not exact. $\square$

3 Special Curves on $\mathscr{P}(\mathbb{T}^d)$ and $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$

3.1 Properties of Absolutely Continuous Paths on $\mathscr{P}(\mathbb{T}^d)$

Given $\sigma \in AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$ there exists [2] a Borel map $\mathbf{v} : (0, T) \times \mathbb{R}^d \to \mathbb{R}^d$ such that $t \to \|\mathbf{v}_t\|_{\sigma_t} \in L^2(0, T)$ and

$$\partial_t\sigma + \nabla \cdot (\mathbf{v}\sigma) = 0 \quad \text{in } \mathscr{D}'((0, T) \times \mathbb{R}^d).$$

We say that $\mathbf{v}$ is a velocity associated to σ. One defines $AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$ as the set of all paths $[\sigma] : [0, T] \ni t \to [\sigma_t] \in \mathscr{P}(\mathbb{T}^d)$ for which there exists $\beta \in L^2(0, T)$ such that

$$\mathscr{W}(\sigma_s, \sigma_t) \leq \int_s^t \beta(\tau)d\tau \quad \text{for all } 0 \leq s \leq t \leq T. \tag{3.1}$$

The obvious inequality $\mathscr{W}(\mu, \nu) \le W_2(\mu, \nu)$ (cf. Lemma 2.1) shows that

$$AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d)) \subset AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$$

and

$$|\sigma'|(t) \ge |\sigma'|_{\mathbb{T}^d}(t) \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0, T). \tag{3.2}$$

We use $|\cdot|$ for the metric derivative on $AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$ and $|\cdot|_{\mathbb{T}^d}$ for that on $AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$. A curve $[0, T] \ni t \to \sigma_t \in \mathscr{P}_2(\mathbb{R}^d)$ is said to be $\mathbb{T}^d$-narrowly continuous if

$$[0, T] \ni t \to \int_{\mathbb{R}^d} \varphi \, d\sigma_t \text{ is continuous for all } \varphi \in C(\mathbb{T}^d).$$

It is not difficult to prove that equality holds in (3.2) in the case where σ is the average of n point masses lying simultaneously on n paths in $W^{1,\infty}(0, T; \mathbb{R}^d)$. The discussion below is intended to show that, in general, the inequality is strict.

In the remainder of the section, given $[\sigma], [\tilde{\sigma}] \in AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$, we define the measures Σ and $\tilde{\Sigma}$ on $(0, T) \times \mathbb{R}^d$ by

$$\begin{aligned} \int_0^T \int_{\mathbb{R}^d} F d\Sigma &:= \int_0^T dt \int_{\mathbb{R}^d} F(t, x) \sigma_t(dx), \\ \int_0^T \int_{\mathbb{R}^d} F d\tilde{\Sigma} &:= \int_0^T dt \int_{\mathbb{R}^d} F(t, x) \tilde{\sigma}_t(dx), \end{aligned} \tag{3.3}$$

for every continuous, bounded function $F : (0, T) \times \mathbb{R}^d \to \mathbb{R}$.

DEFINITION 3.1. Let $[0, T] \ni t \to \sigma_t \in \mathscr{P}_2(\mathbb{R}^d)$ and let $\mathbf{v} : (0, T) \times \mathbb{T}^d \to \mathbb{R}^d$ be a Borel velocity field such that $\mathbf{v}_t \in L^2(\mathbb{T}^d, \sigma_t)$ for almost every $t \in (0, T)$ and $t \to \|\mathbf{v}\|_{\sigma_t} \in L^2(0, T)$. If

$$\int_0^T \int_{\mathbb{R}^d} (\partial_t \varphi + \nabla \varphi \cdot \mathbf{v}_t) d\sigma_t \, dt = 0 \tag{3.4}$$

for all $\varphi \in C_c^\infty((0, T); C^\infty(\mathbb{T}^d))$, we say that $\mathbf{v}$ is a *velocity associated to σ in the periodic sense*.

THEOREM 3.2. *Let $[\sigma] \in AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$. Then there exists a velocity $\mathbf{v}$ associated to σ in the periodic sense such that*

$$\|\mathbf{v}_t\|_{\sigma_t} \le |\sigma'|_{\mathbb{T}^d}(t) \quad \textit{for } \mathscr{L}^1\textit{-a.e. } t \in (0, T). \tag{3.5}$$

The proof of theorem 8.3.1 in [2] can be reproduced to prove Theorem 3.2, with obvious modifications required by our periodic setting.

Remark 3.3. Let $\mathbf{v}$ and $\mathbf{w}$ be two periodic velocities for $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$. Clearly, we have that $\mathbf{w}_t - \mathbf{v}_t \in [\mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d)]^\perp$ for a.e. $t \in (0,T)$. Throughout, $[\mathscr{T}_\mu\mathscr{P}(\mathbb{T}^d)]^\perp$ denotes the orthogonal complement of $\mathscr{T}_\mu\mathscr{P}(\mathbb{T}^d)$ in $L^2(\mathbb{T}^d,\mu)$. But the velocity found in Theorem 3.2 lies in $\mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d)$ for a.e. t, which implies, according to the previous observation, that for a.e. t it has the minimal $L^2(\sigma_t)$-norm among all possible periodic velocities. Thus, we call the velocity given by Theorem 3.2 the *velocity of minimal norm in the periodic sense*, or simply the periodic velocity of minimal norm associated to σ.

The concept of minimal norm velocity in the periodic sense can be generalized to the concept of $\mathbf{c}$-minimal norm periodic velocity.

LEMMA 3.4. *Let $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$, and $\Lambda : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ be a Borel map such that $\Lambda(t,\cdot) \in L^2(\mathbb{T}^d,\sigma_t)$ for all $t \in [0,T]$ and $\|\Lambda(t,\cdot)\|_{\sigma_t} \in L^2(0,T)$. Write the decomposition*

$$\Lambda(t,x) = \Phi(t,x) + \Psi(t,x)$$

where

$$\Phi(t,\cdot) \in \mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d) \quad \textit{and} \quad \Psi(t,\cdot) \in [\mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d)]^\perp.$$

Then Φ and Ψ are Borel maps in the sense that there exists a Borel map $X : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ such that for $\mathscr{L}^1$-a.e. $t \in [0,T]$ we have $\Phi(t,\cdot) \equiv X(t,\cdot)$ σ_t-a.e.

PROOF. Define the Borel measure Σ as in (3.3). Let $\mathscr{M}$ be the closure in $L^2(\Sigma)$ of $\{\nabla_x\zeta(t,x) : \zeta \in C([0,T];C^1(\mathbb{T}^d))\}$, and decompose orthogonally $\Lambda = \Lambda_{\mathscr{M}} + \Lambda_{\mathscr{M}^\perp}$. Thus there exists a sequence $\{\zeta^n\}_n \subset C([0,T];C^1(\mathbb{T}^d))$ such that

$$\lim_{n\to\infty}\int_0^T \|\nabla_x\zeta^n(t,\cdot) - \Lambda_{\mathscr{M}}(t,\cdot)\|^2_{\sigma_t}\,dt = 0.$$

Thus, for $\mathscr{L}^1$-a.e. $t \in [0,T]$ we have $\Phi(t,x) \equiv \Lambda_{\mathscr{M}}(t,x)$ for σ_t-a.e. $x \in \mathbb{R}^d$. $\square$

COROLLARY 3.5. *Let $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$. Then any $\mathbf{c} \in \mathbb{R}^d$ decomposes as*

$$\mathbf{c} = \tilde{\mathbf{c}} + \mathbf{c}^\perp \quad \textit{where } \tilde{\mathbf{c}}_t \in \mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d) \textit{ and } \mathbf{c}_t^\perp \in [\mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d)]^\perp. \tag{3.6}$$

Furthermore, $\tilde{\mathbf{c}}$ and $\mathbf{c}^\perp$ are Borel maps as functions of (t,x).

Now we can prove the following:

PROPOSITION 3.6. *Let $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ and $\mathbf{v}$ be its velocity of minimal norm in the periodic sense. Then $\mathbf{v}_\mathbf{c} := \mathbf{v} - \mathbf{c}^\perp$ is the unique velocity in the periodic sense that minimizes $\|\mathbf{w}+\mathbf{c}\|^2_{\sigma_t}$ among all velocities in the periodic sense $\mathbf{w}$.*

PROOF. Every periodic velocity looks like $\mathbf{w} = \mathbf{v} + \psi$, where $\mathbf{v}$ is the minimal norm periodic velocity and $\psi : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ is a Borel map such that $\psi_t \in [\mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d)]^\perp$ for a.e. $t \in (0,T)$. So $\mathbf{w}+\mathbf{c} = \mathbf{v}+\psi+\mathbf{c}$ and

$$\|\mathbf{w}_t+\mathbf{c}\|^2_{\sigma_t} = \|\mathbf{v}_t+\tilde{\mathbf{c}}_t\|^2_{\sigma_t} + \|\psi_t+\mathbf{c}_t^\perp\|^2_{\sigma_t} \quad \text{for a.e. } t \in (0,T).$$

The lowest value is obtained for $\psi = -\mathbf{c}^\perp$. $\square$

DEFINITION 3.7. The velocity for $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ shown in Proposition 3.6 is called the *periodic velocity of* $\mathbf{c}$*-minimal norm associated to* σ.

DEFINITION 3.8. Let $[\sigma] = [\sigma^*] \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$, let $\mathbf{v},\mathbf{v}^*$ be velocities associated to σ,σ^*, respectively, in the periodic sense. We say that $(\sigma,\mathbf{v})$ and $(\sigma^*,\mathbf{v}^*)$ are equivalent and write $(\sigma,\mathbf{v}) \sim (\sigma^*,\mathbf{v}^*)$ if $(\sigma_t,\mathbf{v}_t) \sim (\sigma_t^*,\mathbf{v}_t^*)$ for almost every $t \in (0,T)$. We denote by $[\sigma,\mathbf{v}]$ the class of equivalence of $(\sigma,\mathbf{v})$ and by $\mathscr{C}(0,T;\mathscr{C}\mathscr{P}(\mathbb{T}^d))$ the set of classes of equivalence. Denote by $\mathscr{C}(0,\infty;\mathscr{C}\mathscr{P}(\mathbb{T}^d))$ the intersection (for all $T>0$) of $\mathscr{C}(0,T;\mathscr{C}\mathscr{P}(\mathbb{T}^d))$.

We next show that if $\sigma,\widetilde{\sigma}$ are two curves in $\mathscr{P}_2(\mathbb{R}^d)$ such that $[\sigma] = [\widetilde{\sigma}] \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$, then any periodic velocity $\mathbf{v}$ for σ has, as expected, a counterpart $\widetilde{\mathbf{v}}$ as a periodic velocity for $\widetilde{\sigma}$ in the sense that $(\sigma_t,\mathbf{v}_t) \sim (\widetilde{\sigma}_t,\widetilde{\mathbf{v}}_t)$ for a.e. $t \in [0,T]$ (written as $[\sigma,\mathbf{v}] \equiv [\widetilde{\sigma},\widetilde{\mathbf{v}}]$).

PROPOSITION 3.9. *Let* $\mu \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ *and* $\sigma,\widetilde{\sigma}$ *be two curves such that* $\sigma_t,\widetilde{\sigma}_t \in \mathscr{P}_2(\mathbb{R}^d)$ *and* $\mu_t = [\sigma_t] = [\widetilde{\sigma}_t]$ *for all* $t \in [0,T]$. *Then, for every periodic velocity* $\mathbf{v}$ *for* σ *there exists a periodic velocity* $\widetilde{\mathbf{v}}$ *for* $\widetilde{\sigma}$ *such that* $[\sigma,\mathbf{v}] \equiv [\widetilde{\sigma},\widetilde{\mathbf{v}}]$.

PROOF. Consider the Borel measures $\Sigma,\widetilde{\Sigma}$ defined on $[0,T]\times\mathbb{R}^d$ whose disintegrations are $d\sigma_t\,dt$ and $d\widetilde{\sigma}_t\,dt$, respectively (see (3.3)). Set

$$-\int_0^T\int_{\mathbb{R}^d}\partial_t\varphi\,d\Sigma = L(\nabla_x\varphi) = \widetilde{L}(\nabla_x\varphi) = -\int_0^T\int_{\mathbb{R}^d}\partial_t\varphi\,d\widetilde{\Sigma}, \tag{3.7}$$

which is a linear functional defined on the vector space of spatial gradients of functions $\varphi \in C_c^1(0,T;C^1(\mathbb{T}^d))$. Since $\mathbf{v}$ is a periodic velocity for σ, we know

$$L(\nabla_x\varphi) = \int_0^T\int_{\mathbb{R}^d}\mathbf{v}\cdot\nabla_x\varphi\,d\Sigma,$$

which yields that L can be extended to the closure of $\nabla_x C_c^1(0,T;C^1(\mathbb{T}^d))$ into $L^2(\Sigma)$ as a linear, continuous functional whose norm is at most $\|\mathbf{v}\|_{L^2(\Sigma)}$. In particular, due to (3.7), we have

$$|\widetilde{L}(\nabla_x\varphi)| \le \|\mathbf{v}\|_\Sigma\|\nabla_x\varphi\|_\Sigma = \|\mathbf{v}\|_{L^2(\Sigma;\mathbb{R}^d)}\|\nabla_x\varphi\|_{\widetilde{\Sigma}}$$
$$\text{for all } \varphi \in C_c^1(0,T;C^1(\mathbb{T}^d)),$$

which means $\widetilde{L}$ can also be extended into a linear, continuous functional defined on the closure of $\nabla_x C_c^1(0,T;C^1(\mathbb{T}^d))$ into $L^2(\widetilde{\Sigma})$ and whose norm is at most $\|\mathbf{v}\|_\Sigma$. By Riesz representation, there exists a Borel map $\widetilde{\mathbf{v}}:[0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ such that $\widetilde{\mathbf{v}} \in L^2(\widetilde{\Sigma})$ and

$$\widetilde{L}(\nabla_x\varphi) = \int_0^T\int_{\mathbb{R}^d}\widetilde{\mathbf{v}}\cdot\nabla_x\varphi\,d\widetilde{\Sigma} \quad \text{for every } \varphi \in C_c^1(0,T;C^1(\mathbb{T}^d)),$$

which, in light of (3.7) and upon using $\varphi(t,x) := f(t)\xi(x)$ with $f \in C_c^1(0,T)$ and $\xi \in \mathscr{C}^1(\mathbb{T}^d)$, yields the conclusion thanks to Remark 2.5. $\square$

We now turn to the relationship between "full" velocities (in the [2] sense) and periodic velocities.

PROPOSITION 3.10. *Let $\sigma \in AC^2(0,T;\mathscr{P}_2(\mathbb{R}^d))$ and let $\mathbf{v}$ be a velocity associated to σ. We define the measure Σ as in* (3.3). *Let $\mathscr{E}$ be the vector space obtained as the closure in $L^2(\Sigma)$ of the set $C([0,T];C^\infty(\mathbb{T}^d))$, and denote its orthogonal complement by $\mathscr{E}^\perp$. Write the orthogonal decomposition*

$$\mathbf{v} = \mathbf{v}^{\mathrm{per}} + \mathbf{w} \quad \textit{where } \mathbf{v}^{\mathrm{per}} \in \mathscr{E} \textit{ and } \mathbf{w} \in \mathscr{E}^\perp.$$

We have

$$\|\mathbf{v}^{\mathrm{per}} + \mathbf{c}\|_\Sigma \le \|\mathbf{v} + \mathbf{c}\|_\Sigma \quad \textit{for every } \mathbf{c} \in \mathbb{R}^d \tag{3.8}$$

and

$$\mathbf{v}_t^{\mathrm{per}} \in L^2(\mathbb{T}^d, \sigma_t) \quad \textit{for almost every } t \in (0,T). \tag{3.9}$$

Also, $\mathbf{v}^{\mathrm{per}}$ is a velocity associated to σ in the periodic sense.

PROOF. Since

$$\mathbf{v} + \mathbf{c} = (\mathbf{v}^{\mathrm{per}} + \mathbf{c}) + \mathbf{w}, \quad \mathbf{v}^{\mathrm{per}} + \mathbf{c} \in \mathscr{E}, \quad \text{and} \quad \mathbf{w} \in \mathscr{E}^\perp,$$

we obtain (3.8).

Next, choose a sequence $\{G^n\} \subset C^1([0,T];C^\infty(\mathbb{T}^d))$ such that

$$\lim_{n\to\infty} \int_0^T \int_{\mathbb{R}^d} \left|\mathbf{v}_t^{\mathrm{per}}(x) + \mathbf{c} - G_t^n(x)\right|^2 \Sigma(dt,dx) = 0.$$

Then up to a subsequence of $\{G^n\}$ we have

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} \left|\mathbf{v}_t^{\mathrm{per}}(x) + \mathbf{c} - G_t^n(x)\right|^2 \sigma_t(dx) = 0$$

for almost every t, which proves (3.9).

Let $\varphi \in C_c^\infty((0,T);C_c^\infty(\mathbb{T}^d))$. Since $\mathbf{v} - \mathbf{v}^{\mathrm{per}} \in \mathscr{E}^\perp$, we have

$$\int_0^T dt \int_{\mathbb{R}^d} (\partial_t\varphi + \mathbf{v}^{\mathrm{per}} \cdot \nabla\varphi)\sigma_t(dx) = \int_0^T dt \int_{\mathbb{R}^d} (\partial_t\varphi + \mathbf{v} \cdot \nabla\varphi)\sigma_t(dx) = 0.$$

This, together with (3.9), proves that $\mathbf{v}^{\mathrm{per}}$ is a velocity associated to σ in the periodic sense. $\square$

As a direct consequence of Proposition 3.10, we obtain the following proposition:

COROLLARY 3.11. *Let* $\sigma \in AC^2(0,T;\mathscr{P}_2(\mathbb{R}^d))$ *and let* $\mathbf{v} : (0,T)\times\mathbb{R}^d \to \mathbb{R}^d$ *be a velocity associated to* σ. *For* $\mathscr{L}^1$*-a.e.* $t \in (0,T)$ *let* $\mathbf{w}_t$ *be the projection of* $\mathbf{v}_t$ *onto* $L^2(\mathbb{T}^d,\sigma_t)$. *Then* $\mathbf{w}$ *is a velocity associated to* σ *in the periodic sense and*

$$\|\mathbf{w}_t\|_{\sigma_t} < \|\mathbf{v}_t\|_{\sigma_t} \quad \text{for } \mathscr{L}^1\text{-a.e. } t \in (0,T) \tag{3.10}$$

unless $\mathbf{v}_t \in L^2(\mathbb{T}^d,\sigma_t)$. *In particular, the vector field* $\mathbf{w}$ *coincides with a Borel map up to a set of zero measure.*

THEOREM 3.12. *Let* $\sigma \in AC^2(0,T;\mathscr{P}_2(\mathbb{R}^d))$ *and let* $\mathbf{v}$ *be a velocity associated to* σ. *Let* $\mathbf{w} : (0,T)\times\mathbb{R}^d \to \mathbb{R}^d$ *be a velocity associated to* σ *in the periodic sense. Then, for almost every* $t \in (0,T)$, $\mathbf{w}_t$ *and* $\mathbf{v}_t$ *have the same projection on* $\mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d)$.

PROOF. Let $\varphi \in C^1(\mathbb{T}^d)$. First we use that $\mathbf{w}$ is a velocity associated to σ in the periodic sense and then the fact that φ and $\nabla\varphi$ are bounded to deduce

$$\langle \mathbf{w}_t, \nabla\varphi\rangle_{\sigma_t} = \frac{d}{dt}\int_{\mathbb{R}^d} \varphi\, d\sigma_t = \langle \mathbf{v}_t, \nabla\varphi\rangle_{\sigma_t} \tag{3.11}$$

in the sense of distributions. We conclude that there exists a set of zero measure $\mathscr{N} \subset (0,T)$ such that (3.11) holds for all $t \in (0,T)\setminus\mathscr{N}$ and all $\varphi \in \Xi$, where $\Xi \subset C^1(\mathbb{T}^d)$ is countably dense in $C(\mathbb{T}^d)$ for the uniform topology. Hence (3.11) holds for all $\varphi \in C^1(\mathbb{T}^d)$. $\square$

We below recall a proposition whose proof is standard (see, e.g., [2] for the nonperiodic case).

LEMMA 3.13. *Let*

$$\mathbf{v} \in L^2(0,T;W^{1,\infty}(\mathbb{T}^d;\mathbb{R}^d)). \tag{3.12}$$

For any $\sigma_0 \in \mathscr{P}_2(\mathbb{R}^d)$ *there exists exactly one solution* $[\sigma] \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ *to the continuity equation*

$$\partial_t\sigma + \nabla_x\cdot(\mathbf{v}\sigma) = 0 \quad \text{in } \mathscr{D}'((0,T)\times\mathbb{T}^d)$$

with the property that $[\sigma(0,\cdot)] \equiv [\sigma_0]$. *This is given by* $\sigma_t \sim X_{t\#}\sigma_0$ *for all* $t \in (0,T)$, *where* X *is the classical flow of* $\mathbf{v}$.

3.2 c-Optimality

For $\mathbf{c} \in \mathbb{R}^d$ and $\mu,\nu \in \mathscr{P}_2(\mathbb{R}^d)$ we define

$$\mathscr{W}_{\mathbf{c}}(\mu,\nu) := \mathscr{W}(\mu,(\mathrm{id}+\mathbf{c})_\#\nu) = \mathscr{W}((\mathrm{id}-\mathbf{c})_\#\mu,\nu)$$

and

$$W_{\mathbf{c}}(\mu,\nu) := W_2(\mu,(\mathrm{id}+\mathbf{c})_\#\nu) = W_2((\mathrm{id}-\mathbf{c})_\#\mu,\nu).$$

Remark 3.14. Let μ and ν be two Borel probability measures with bounded second moments.

(i) Observe that if γ has μ and ν as its marginals, then

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y-\mathbf{c}|^2\gamma(dx,dy) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\gamma(dx,dy) + |\mathbf{c}|^2 + 2\mathbf{c}\cdot\int_{\mathbb{R}^d} y\nu(dy) - 2\mathbf{c}\cdot\int_{\mathbb{R}^d} x\mu(dx),$$

and so γ minimizes the expression on the left-hand side of the previous identity if and only if it minimizes the first expression on its right-hand side. In particular,

$$W_{\mathbf{c}}^2(\mu,\nu) = W_2^2(\mu,\nu) + |\mathbf{c}|^2 + 2\mathbf{c}\cdot\int_{\mathbb{R}^d} y\nu(dy) - 2\mathbf{c}\cdot\int_{\mathbb{R}^d} x\mu(dx).$$

(ii) Denote $\mu_{\mathbf{c}} := (\mathrm{id}-\mathbf{c})_{\#}\mu$. By Lemma 2.1, there exists $\nu_* \in [\nu]$ and $\gamma_{\mathbf{c}} \in \Gamma_o(\mu_{\mathbf{c}}, \nu_*)$ (the set of optimal transport plans between $\mu_{\mathbf{c}}$ and ν_*) such that

$$\begin{aligned}\mathscr{W}^2(\mu_{\mathbf{c}},\nu) = W^2(\mu_{\mathbf{c}},\nu_*) &= \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\gamma_{\mathbf{c}}(dx,dy)\\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2_{\mathbb{T}^d}\gamma_{\mathbf{c}}(dx,dy).\end{aligned}$$

Let $\gamma := [(\pi_1+\mathbf{c})\times\pi_2]_{\#}\gamma_{\mathbf{c}}$ and notice that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y-\mathbf{c}|^2_{\mathbb{T}^d}\gamma(dx,dy) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2_{\mathbb{T}^d}\gamma_{\mathbf{c}}(dx,dy) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2\gamma_{\mathbf{c}}(dx,dy) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y-\mathbf{c}|^2\gamma(dx,dy).$$

According to (i), we necessarily have $\gamma \in \Gamma_o(\mu,\nu_*)$. Furthermore, by the definitions of $W_{\mathbf{c}}$ and $\mathscr{W}_{\mathbf{c}}$, we get

$$\mathscr{W}_{\mathbf{c}}^2(\mu,\nu) = \min_{\bar\nu\sim\nu} W_{\mathbf{c}}^2(\mu,\bar\nu) = W_{\mathbf{c}}^2(\mu,\nu_*) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y-\mathbf{c}|^2_{\mathbb{T}^d}\gamma(dx,dy).$$

3.3 Approximation by Measures with Smooth Densities

For $\varepsilon > 0$ take

$$\eta^\varepsilon(x) := (2\pi\varepsilon)^{-d/2}\exp(-|x|^2/(2\varepsilon))$$

and introduce the periodic mollifying kernel

$$\sum_{\mathbf{k}\in\mathbb{Z}^d} \eta^\varepsilon(\cdot+\mathbf{k}) =: \mathscr{E}^\varepsilon \in C^\infty(\mathbb{T}^d), \quad \mathscr{E}^\varepsilon > 0 \text{ in } \mathbb{R}^d. \tag{3.13}$$

It is easy to see that the convergence of the series is absolute, so the properties listed above follow.

Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $\zeta \in L^2(\mu)$, and set

$$\widetilde{\rho}^\varepsilon = \mu * \mathscr{E}^\varepsilon, \quad Z^\varepsilon = (\zeta\mu) * \mathscr{E}^\varepsilon, \quad \zeta^\varepsilon = \frac{Z^\varepsilon}{\widetilde{\rho}^\varepsilon}, \quad \mu^\varepsilon = \widetilde{\rho}^\varepsilon\big|_{\mathscr{Q}}, \tag{3.14}$$

where the convolution is standard, so that the integrals involved are performed over $\mathbb{R}^d$. We shall not make a notational distinction between measures absolutely continuous with respect to $\mathscr{L}^d$ and their densities. Note that the first three functions defined at (3.14) are smooth and $\mathbb{Z}^d$-periodic. Since $\widetilde{\rho}^\varepsilon$ is also everywhere positive, it has infinite total mass. One checks that if $\varphi \in C(\mathbb{T}^d)$, then

$$\mathscr{E}^\varepsilon * (\varphi\chi_{\mathscr{Q}}) = \eta^\varepsilon * \varphi. \tag{3.15}$$

Setting $\varphi \equiv 1$ in (3.15) we obtain that the total mass of μ^ε is 1.

3.4 The Metric Derivative in Terms of the Velocity of Minimal Norm

We consider $\mathbb{T}^d$-narrowly continuous curve $\sigma : [0,T] \to \mathscr{P}_2(\mathbb{R}^d)$ for which there exists an associated velocity $\mathbf{v}$ in the periodic sense. Set

$$\widetilde{\rho}_t^\varepsilon = \sigma_t * \mathscr{E}^\varepsilon, \quad E_t^\varepsilon = (\mathbf{v}_t\sigma_t) * \mathscr{E}^\varepsilon, \quad \mathbf{v}_t^\varepsilon = \frac{E_t^\varepsilon}{\widetilde{\rho}_t^\varepsilon}, \quad \overline{\sigma}_t^\varepsilon = \widetilde{\rho}_t^\varepsilon\big|_{\mathscr{Q}}, \tag{3.16}$$

where the periodic mollifying kernel is defined in (3.13). From the fact that

$$\mathscr{W}(\overline{\sigma}_t^\varepsilon, \sigma_t) \le \varepsilon^2 \int\limits_{\mathbb{R}^d} \eta(z)|z|^2\,dz \tag{3.17}$$

we see that the following holds:

LEMMA 3.15. *The measures $[\overline{\sigma}_t^\varepsilon]$ converge to $[\overline{\sigma}_t]$ in $\mathscr{P}(\mathbb{T}^d)$ as $\varepsilon \to 0^+$, uniformly with respect to $t \in [0,T]$.*

We shall next state a lemma that carries some statements of [2, lemma 8.1.9] to our context.

LEMMA 3.16. *If $\mathbf{v}$ is a velocity associated to σ in the periodic sense, $\mathbf{v}^\varepsilon$ is also a velocity associated to $\overline{\sigma}^\varepsilon$ in the periodic sense. Moreover, for $t \in [0,T]$,*

$$\|\mathbf{v}_t^\varepsilon\|_{\overline{\sigma}_t^\varepsilon} \le \|\mathbf{v}_t\|_{\sigma_t}. \tag{3.18}$$

PROOF. Let $\varphi \in C_c^1(0,T;C^1(\mathbb{T}^d))$. We first use the fact that the convolution operator is self-adjoint and then (3.15) to conclude that

$$\begin{aligned}\int_0^T dt \int\limits_{\mathbb{R}^d} \mathbf{v}^\varepsilon \cdot \nabla\varphi_t\,d\overline{\rho}_t^\varepsilon &= \int_0^T dt \int\limits_{\mathbb{R}^d} \mathbf{v}_t \cdot \mathscr{E}^\varepsilon * (\chi_{\mathscr{Q}}\nabla\varphi_t)\sigma_t(dx) \\ &= \int_0^T dt \int\limits_{\mathbb{R}^d} \mathbf{v}_t \cdot \eta^\varepsilon * \nabla\varphi_t\,d\sigma_t.\end{aligned}$$

Hence

$$\int_0^T dt \int_{\mathbb{R}^d} \mathbf{v}_t^\varepsilon \cdot \nabla\varphi_t \, d\overline{\rho}_t^\varepsilon = \int_0^T dt \int_{\mathbb{R}^d} \mathbf{v}_t \cdot \nabla(\eta^\varepsilon * \varphi_t) d\sigma_t. \tag{3.19}$$

Similarly,

$$\int_0^T dt \int_{\mathbb{R}^d} \partial_t\varphi_t \, d\overline{\rho}_t^\varepsilon = \int_0^T dt \int_{\mathbb{R}^d} \partial_t(\eta^\varepsilon * \varphi) d\sigma_t. \tag{3.20}$$

Combining (3.19)–(3.20) we see that $(\overline{\sigma}^\varepsilon, \mathbf{v}^\varepsilon)$ satisfies (3.4). We have (cf. [2, p. 178])

$$|\mathbf{v}_t^\varepsilon(x)|^2 \rho_t^\varepsilon(x) \le \int_{\mathbb{R}^d} |\mathbf{v}_t(y)|^2 \mathscr{E}_t^\varepsilon(x-y)\sigma_t(dy). \tag{3.21}$$

Since

$$\int_{\mathscr{Q}} \mathscr{E}_t^\varepsilon(x-y)dx = 1,$$

integrating both sides of (3.21) over $\mathscr{Q}$ we conclude the proof of the lemma. □

PROPOSITION 3.17.

(i) *One has* $\mathbf{v}^\varepsilon \in L^2(0,T; W^{1,\infty}(\mathbb{T}^d;\mathbb{R}^d))$.

(ii) *There exists a unique (flow)* $X^\varepsilon : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ *such that*

$$\dot{X}_t^\varepsilon(x) = \mathbf{v}_t^\varepsilon\big(X_t^\varepsilon(x)\big), \quad X_0^\varepsilon(x) = x, \quad X_t^\varepsilon(x+\mathbf{k}) = X_t^\varepsilon(x) + \mathbf{k} \tag{3.22}$$

for all $x \in \mathbb{R}^d$ *and* $\mathbf{k} \in \mathbb{Z}^d$.

(iii) *We have*

$$X_{t\#}^\varepsilon[\overline{\sigma}_0^\varepsilon] = [\overline{\sigma}_t^\varepsilon] \quad \textit{in } \mathscr{P}(\mathbb{T}^d) \tag{3.23}$$

for all $t \in [0,T]$ *and all* $\varepsilon > 0$.

PROOF.

(i) The time distributional derivative of $\widetilde{\rho}^\varepsilon$ is

$$\partial_t \widetilde{\rho}_t^\varepsilon(x) = -\int_{\mathbb{R}^d} \mathbf{v}_t(y)\cdot\nabla\mathscr{E}^\varepsilon(x-y)\sigma_t(dy),$$

and so

$$|\partial_t\widetilde{\rho}_t^\varepsilon(x)| \le \|E_t^\varepsilon\|_{W1,\infty}\|\mathbf{v}_t\|_{\sigma_t}.$$

This proves that for each x, $t \to \widetilde{\rho}_t^\varepsilon(x)$ is 2-absolutely continuous, uniformly in x, and so it is Hölder-continuous. Hence $\widetilde{\rho}^\varepsilon$ is continuous in both variables (t,x). As it is pointwise positive and periodic in x, it is bounded below by a constant depending only on ε. But

$$\|\nabla\widetilde{\rho}_t^\varepsilon\|_\infty \le \|\nabla\mathscr{E}^\varepsilon\|_\infty, \quad \|E_t^\varepsilon\|_{W1,\infty} \le \|\mathscr{E}^\varepsilon\|_{W^{1,\infty}}\|\overline{\mathbf{v}}_t\|_{\overline{\sigma}_t}.$$

These prove that $\mathbf{v}^\varepsilon \in L^2(0,T;W^{1,\infty}(\mathbb{T}^d;\mathbb{R}^d))$.

(ii) Existence and uniqueness of a flow $X^\varepsilon : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$ satisfying (3.22) follows from (i). By uniqueness and due to the periodicity of $\mathbf{v}_t^\varepsilon$ we deduce that

$$X_t^\varepsilon(x+\mathbf{k}) = X_t^\varepsilon(x) + \mathbf{k}$$

for all $\mathbf{k} \in \mathbb{Z}^d$.

(iii) By Lemma 3.16 $(\bar\sigma^\varepsilon, \mathbf{v}^\varepsilon)$ satisfies (3.4). We use Lemma 3.13 to obtain (iii). □

The Eulerian characterization of curves in $AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ follows.

THEOREM 3.18. *Let $[\sigma] \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ and $\mathbf{v}$ be a velocity in the periodic sense. If $0 \le s < t \le T$ and we set $h = t - s$, then*

$$\mathscr{W}_{h\mathbf{c}}(\sigma_s,\sigma_t) \le \int_s^t \|\mathbf{v}_\tau + \mathbf{c}\|_{\sigma_\tau}\, d\tau \quad \textit{for every } \mathbf{c} \in \mathbb{R}^d. \tag{3.24}$$

As a consequence, by choosing $\mathbf{c} = 0$ we get

$$|\sigma'|_{\mathbb{T}^d}(t) \le \|\mathbf{v}_t\|_{\sigma_t} \quad \textit{for } \mathscr{L}^1\textit{-a.e. } t \in (0,T). \tag{3.25}$$

PROOF. Define

$$\sigma_t^\varepsilon = X_\#^\varepsilon \bar\sigma_0^\varepsilon.$$

Note that for any $0 \le s < t \le T$, we have that $\gamma_{s,t}^\varepsilon := (X_s^\varepsilon \times X_t^\varepsilon)_\# \bar\sigma_0^\varepsilon$ belongs to $\Gamma(\sigma_s^\varepsilon, \sigma_t^\varepsilon)$. Thus,

$$\begin{aligned}\mathscr{W}_{h\mathbf{c}}^2(\sigma_s^\varepsilon,\sigma_t^\varepsilon) &\le \int\limits_{\mathbb{R}^d} |X_t^\varepsilon(x) + h\mathbf{c} - X_s^\varepsilon(x)|_{\mathbb{T}^d}^2\, \bar\sigma_0^\varepsilon(dx)\\ &= \int\limits_{\mathbb{R}^d} \left|\int_s^t (\dot X_\tau^\varepsilon(x) + \mathbf{c})\, d\tau\right|_{\mathbb{T}^d}^2 \bar\sigma_0^\varepsilon(dx).\end{aligned}$$

We then use Minkowski's inequality to conclude that

$$\begin{aligned}\mathscr{W}_{h\mathbf{c}}(\sigma_s^\varepsilon,\sigma_t^\varepsilon) &\le \int_s^t \left\|\dot X_\tau^\varepsilon(x) + \mathbf{c}\right\|_{\bar\sigma_0^\varepsilon} d\tau\\ &= \int_s^t \|\mathbf{v}_\tau^\varepsilon(x) + \mathbf{c}\|_{\bar\sigma_\tau^\varepsilon}\, d\tau \le \int_s^t \|\mathbf{v}_\tau + \mathbf{c}\|_{\sigma_\tau}\, d\tau.\end{aligned} \tag{3.26}$$

We have used Lemma 3.16 to obtain the last inequality in (3.26). Thanks to Lemma 3.15, letting $\varepsilon \to 0^+$ in (3.26) we establish (3.24).

Setting $\mathbf{c} = 0$ in (3.24) we conclude that (3.25) holds whenever t is a Lebesgue point of $\tau \to \|\mathbf{v}_\tau + \mathbf{c}\|_{\sigma_\tau}$ and $|\sigma'|_{\mathbb{T}^d}(t)$ exists. □

Proposition 3.10 and Theorem 3.18 yield the following lemma:

LEMMA 3.19. *Let $\sigma \in AC^2(0,1;\mathscr{P}_2(\mathbb{R}^d))$ and let $\mathbf{w}$ be a velocity associated to σ. Then*

$$\int_0^1 \|\mathbf{w}_t + \mathbf{c}\|_{\sigma_t}^2 \, dt \geq W_{\mathbf{c}}^2(\sigma_0,\sigma_1) \geq \mathscr{W}_{\mathbf{c}}^2(\sigma_0,\sigma_1) \quad \textit{for all } \mathbf{c} \in \mathbb{R}^d. \tag{3.27}$$

Next, we shall show that there are cases where the inequalities in Lemma 3.19 can be reversed. We introduce the projections $\pi_1(x,y) = x$ and $\pi_2(x,y) = y$.

LEMMA 3.20. *Let $\sigma_0, \sigma_1 \in \mathscr{P}_2(\mathbb{R}^d)$ and let $\gamma \in \Gamma(\sigma_0,\sigma_1)$ be such that*

$$\mathscr{W}_{\mathbf{c}}^2(\sigma_0,\sigma_1) = W_{\mathbf{c}}^2(\sigma_0,\sigma_1) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y-\mathbf{c}|_{\mathbb{T}^d}^2 \gamma(dx,dy).$$

For $t \in [0,1]$ we set

$$\sigma_t := [(1-t)\pi_1 + t\pi_2]_{\#}\gamma$$

and define Σ as in (3.3).

(i) *Whenever $0 \leq s < t \leq 1$ we have*

$$W_2(\sigma_t,\sigma_s) = (t-s)W_2(\sigma_0,\sigma_1),$$

and so $\sigma \in AC^2(0,1;\mathscr{P}_2(\mathbb{R}^d))$.

(ii) *Let $\mathbf{v} : (0,1)\times\mathbb{R}^d \to \mathbb{R}^d$ be implicitly defined by*

$$\int_0^1 \int_{\mathbb{R}^d} \mathbf{v}_t(x)\cdot G_t(x)\Sigma(dt,dx) := \int_0^1 dt \int_{\mathbb{R}^d\times\mathbb{R}^d} (y-x)\cdot G_t\big((1-t)x+ty\big)\gamma(dx,dy)$$

for all continuous and bounded test functions G. Then $\mathbf{v}$ is the velocity of the minimal norm associated to σ and

$$\|\mathbf{v}_t + \mathbf{c}\|_{\sigma_t} = W_{\mathbf{c}}(\sigma_0,\sigma_1) = \mathscr{W}_{\mathbf{c}}(\sigma_0,\sigma_1) \quad \textit{for almost every } t \in (0,1). \tag{3.28}$$

PROOF. The fact that $\mathbf{v}$ is a velocity associated to σ is readily checked. We refer to Section 7.2 [2] for the proof of (i) and for the fact that $\mathbf{v}$ is of minimal norm.

It remains to study its properties. Let $\phi \in C(\mathbb{R}^d,\mathbb{R}^d)$ be a bounded function. We have

$$\begin{aligned}\langle \phi, \mathbf{v}_t + \mathbf{c}\rangle_{\sigma_t} &= \int_{\mathbb{R}^d\times\mathbb{R}^d} \phi((1-t)x+ty)\cdot(y-x+\mathbf{c})\gamma(dx,dy)\\ &\leq \|\phi\|_{\sigma_t}\mathscr{W}_{\mathbf{c}}(\sigma_0,\sigma_1).\end{aligned}$$

Since ϕ is arbitrary, we conclude that (3.28) holds. This, together with Lemma 3.19, proves that (3.28) holds for almost every $t \in (0,1)$. □

3.5 Lifting with Curves of c-Minimal Velocity

Here our goal is to show that for any $\mathbf{c} \in \mathbb{R}^d$, any curve in $AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ admits a lift "of $\mathbf{c}$-minimal velocity" onto $AC^2(0,T;\mathscr{P}_2(\mathbb{R}^d))$. We begin with a well-known proposition which is formulated so that it can also be adapted to the periodic case. Let us fix $T > 0$.

PROPOSITION 3.21. *Assume $\{\sigma^n\}_n \subset AC^2(0,T;\mathscr{P}_2(\mathbb{R}^d))$ and $\mathbf{v}^n$ is a velocity associated to σ^n (respectively, $\mathbf{v}^n$ is a velocity associated to σ^n in the periodic sense). Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $C_0 \geq \sqrt{d}$ be such that $W_2(\sigma_t^n,\mu) \leq C_0$ (respectively, $\mathscr{W}(\sigma_t^n,\mu) \leq C_0$) for all $t \in [0,T]$ and all integers $n \geq 1$, and*

$$\sup_{n\in\mathbb{N}} \int_0^T \|\mathbf{v}_t^n\|^2_{\sigma_t^n}\, dt \leq C_0. \tag{3.29}$$

Then there exists an increasing sequence $\{n_k\}_k \subset \mathbb{N}$, a path $\sigma \in AC^2(0,T;\mathscr{P}_2(\mathbb{R}^d))$ (respectively, $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$), and $\mathbf{v}$ a velocity associated to σ ($\mathbf{v}$ a velocity associated to σ in the periodic sense) such that

(i) *$\{\sigma_t^{n_k}\}_k$ converges narrowly (respectively, in $\mathscr{P}(\mathbb{T}^d)$) to σ_t for $t \in [0,T]$, and*

(ii) *it holds that*

$$\liminf_{k\to\infty} \int_0^T \|\mathbf{v}^{n_k} + \mathbf{c}\|^2_{\sigma_t^{n_k}}\, dt \geq \int_0^T \|\mathbf{v} + \mathbf{c}\|^2_{\sigma_t}\, dt \tag{3.30}$$

for any $\mathbf{c} \in \mathbb{R}^d$.

THEOREM 3.22. *Let $\mathbf{c} \in \mathbb{R}^d$, $\sigma^* \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$, and $\mathbf{v}^*$ be a velocity associated to σ^* in the periodic sense. Then there exist $\sigma \in AC^2(0,T;\mathscr{P}_2(\mathbb{R}^d))$ and a velocity $\mathbf{v}$ associated to σ such that*

$$\int_0^T \|\mathbf{v}_t + \mathbf{c}\|^2_{\sigma_t}\, dt \leq \int_0^T \|\mathbf{v}_t^* + \mathbf{c}\|^2_{\sigma_t^*}\, dt, \tag{3.31}$$

$\sigma_0 = \sigma_0^*$, and $[\sigma_t] = [\sigma_t^*]$ for all $t \in [0,T]$. Furthermore, $\mathbf{v}$ is associated to σ in the periodic sense.

PROOF. Let $n \geq 1$ be an integer and set $h = T/n$ and $t_k = kh$ for $k = 0,\dots,n$. We use Remark 3.14(ii) to inductively find $\sigma_{t_k}^h$ such that $\sigma_0^h = \sigma_0^*$,

$$W_{h\mathbf{c}}\big(\sigma_{t_{k-1}}^h,\sigma_{t_k}^h\big) = \mathscr{W}_{h\mathbf{c}}\big(\sigma_{t_{k-1}}^h,\sigma_{t_k}^h\big) = \mathscr{W}_{h\mathbf{c}}\big(\sigma_{t_{k-1}}^*,\sigma_{t_k}^*\big),$$

and $[\sigma_{t_k}^h] = [\sigma_{t_k}^*]$ for all $k = 1,\dots,n$. Choose $\gamma_k \in \Gamma(\sigma_{t_{k-1}}^h,\sigma_{t_k}^h)$ such that

$$\begin{aligned}\mathscr{W}_{h\mathbf{c}}^2(\sigma_{t_{k-1}}^*,\sigma_{t_k}^*) &= \int\limits_{\mathbb{R}^d\times\mathbb{R}^d} |x-y-h\mathbf{c}|^2_{\mathbb{T}^d}\gamma_k(dx,dy)\\ &= \int\limits_{\mathbb{R}^d\times\mathbb{R}^d} |x-y-h\mathbf{c}|^2\gamma_k(dx,dy).\end{aligned}$$

(1) Remark 3.14(i) ensures that

$$W_2^2\big(\sigma^h_{t_{k-1}}, \sigma^h_{t_k}\big) = \int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^2 \gamma_k(dx, dy).$$

For $t \in [t_{k-1}, t_k]$, we set

$$\sigma^h_t := \Big[\Big(1 - \frac{t-t_{k-1}}{h}\Big)\pi_1 + \frac{t-t_{k-1}}{h}\pi_2\Big]_{\#}\gamma$$

and define a Borel measure Σ^h on $(0,T)\times\mathbb{R}^d$ by

$$\int_0^T\int_{\mathbb{R}^d} G\, d\Sigma^h = \int_0^T dt \int_{\mathbb{R}^d} G_t(x)\sigma^h_t(dx)$$

for all continuous and bounded functions $G : [0,T]\times\mathbb{R}^d \to \mathbb{R}$. We define the vector field $\mathbf{v}^h : (0,1)\times\mathbb{R}^d \to \mathbb{R}^d$ by

$$\int_0^T\int_{\mathbb{R}^d} G\cdot \mathbf{v}^h\, d\Sigma^h := \sum_{k=1}^n \int_{t_{k-1}}^{t_k} dt \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{y-x}{h}\cdot G_t\Big(x + (y-x)\frac{t-t_{k-1}}{h}\Big)\gamma_k(dx, dy)$$

for all continuous and bounded functions $G : [0,T]\times\mathbb{R}^d \to \mathbb{R}^d$. We have

$$\int_{\mathbb{R}^d} \mathbf{v}^h_t(x)\cdot G_t(x)\sigma^h_t(dx) = \int_{\mathbb{R}^d\times\mathbb{R}^d} \frac{y-x}{h}\cdot G_t\Big(x + (y-x)\frac{t-t_{k-1}}{h}\Big)\gamma_k(dx, dy).$$

By Lemma 3.20, $\sigma^h \in AC^2(0,T;\mathscr{P}_2(\mathbb{R}^d))$ and $\mathbf{v}^h$ is the velocity of minimal norm associated to σ^h. We have

$$\begin{aligned}\int_{\mathbb{R}^d} |\mathbf{v}^h_t|^2\, d\sigma^h_t &= \int_{\mathbb{R}^d\times\mathbb{R}^d} \Big|\frac{y-x}{h}\Big|^2 \gamma_k(dx, dy)\\ &= \frac{1}{h^2} W_2^2\big(\sigma^h_{t_{k-1}}, \sigma^h_{t_k}\big) \le \frac{1}{h^2}\Big(\int_{t_{k-1}}^{t_k} \|\mathbf{v}^*_s\|_{\sigma^*_s}\, ds\Big)^2\end{aligned}$$

for every $t \in (t_{k-1}, t_k)$. We sum up over $k = 1,\dots,n$ and use Jensen's inequality to obtain

$$\int_0^T \|\mathbf{v}^h_t(x)\|^2_{\sigma^h_t}\, dt \le \int_0^T \|\mathbf{v}^*_t\|^2_{\sigma^*_t}\, dt, \tag{3.32}$$

and so, using (3.1) we obtain

$$W_2(\sigma_t^h, \sigma_0^*) \le \int_0^t \|\mathbf{v}_s^h\|_{\sigma_s^h}\, ds \le \sqrt{T}\sqrt{\int_0^T \|\mathbf{v}_t^*\|^2_{\sigma_t^*}\, dt}. \tag{3.33}$$

Similarly,

$$\int_{\mathbb{R}^d} |\mathbf{v}_t^h + \mathbf{c}|^2\, d\sigma_t^h = \int_{\mathbb{R}^d\times\mathbb{R}^d} \left|\frac{y-x}{h} + \mathbf{c}\right|^2 \gamma_k(dx, dy)$$
$$= \frac{1}{h^2} W^2_{h\mathbf{c}}\big(\sigma^h_{t_{k-1}}, \sigma^h_{t_k}\big) = \frac{1}{h^2} \mathscr{W}^2_{h\mathbf{c}}\big(\sigma^*_{t_{k-1}}, \sigma^*_{t_k}\big)$$

for every $t \in (t_{k-1}, t_k)$, and so

$$\int_0^T dt \int_{\mathbb{R}^d} |\mathbf{v}_t^h + \mathbf{c}|^2\, d\sigma_t^h = \frac{1}{h}\sum_{k=1}^n \mathscr{W}^2_{h\mathbf{c}}(\sigma^*_{t_{k-1}}, \sigma^*_{t_k}). \tag{3.34}$$

Using (3.34) and combining Lemma 3.18 with Jensen's inequality, we obtain

$$\begin{aligned} \int_0^T dt \int_{\mathbb{R}^d} |\mathbf{v}_t^h + \mathbf{c}|^2\, d\sigma_t^h &\le \sum_{k=1}^n \int_{t_{k-1}}^{t_k} \|\mathbf{v}_t^* + \mathbf{c}\|^2_{\sigma_t^*}\, dt \\ &= \int_0^T dt \int_{\mathbb{R}^d} |\mathbf{v}_t^* + \mathbf{c}|^2\, d\sigma_t^*. \end{aligned} \tag{3.35}$$

Thanks to Proposition 3.21, (3.32) combined with (3.33) provides us with a sequence $\{\sigma^{h_k}, \mathbf{v}^{h_k}\}_k$, $\sigma \in AC^2(0,T;\mathscr{P}_2(\mathbb{R}^d))$, and a velocity $\mathbf{v}$ associated to σ such that $\{\sigma_t^{h_k}\}_k$ converges to σ_t in $\mathscr{P}_2(\mathbb{R}^d)$ for every $t \in [0,T]$ and

$$\int_0^T \|\mathbf{v}_t + \mathbf{c}\|^2_{\sigma_t}\, dt \le \liminf_{k\to\infty} \int_0^T \|\mathbf{v}_t^{h_k} + \mathbf{c}\|^2_{\sigma_t^{h_k}}\, dt. \tag{3.36}$$

We combine (3.35) and (3.36) to obtain (3.31).

(2) We claim that $[\sigma_t] = [\sigma_t^*]$ for all $t \in [0,T]$. Indeed, if $t \in [t_{k-1}, t_k]$, then

$$\begin{aligned} \mathscr{W}\big(\sigma_t^h, \sigma_t^*\big) &\le \mathscr{W}\big(\sigma_t^h, \sigma^h_{t_{k-1}}\big) + \mathscr{W}\big(\sigma^h_{t_{k-1}}, \sigma_t^*\big) \\ &= \mathscr{W}\big(\sigma_t^h, \sigma^h_{t_{k-1}}\big) + \mathscr{W}\big(\sigma^*_{t_{k-1}}, \sigma_t^*\big). \end{aligned}$$

We use Lemma 3.20 to conclude that

$$\mathscr{W}\big(\sigma_t^h, \sigma_t^*\big) \le \frac{t - t_{k-1}}{h}\mathscr{W}\big(\sigma^*_{t_{k-1}}, \sigma^*_{t_k}\big) + \mathscr{W}\big(\sigma^*_{t_{k-1}}, \sigma_t^*\big) \le 2\int_{t_{k-1}}^{t_k} \|\mathbf{v}_\tau^*\|_{\sigma_\tau^*}\, d\tau.$$

Replacing h by h_k in the previous inequality and letting k tend to ∞ yields that $\{\sigma_t^{h_k}\}_k$ converges to σ_t^* in $\mathscr{P}(\mathbb{T}^d)$ for every $t \in [0,T]$, which proves the claim.

(3) Define $\mathscr{E}$ as in Proposition 3.10 and let $\mathbf{v}^{\text{per}}$ be the projection of $\mathbf{v}$ onto $\mathscr{E}$. We combine claim (2) and (3.34) and then combine Lemma 3.18 with Jensen's inequality to obtain

$$\int_0^T \left\|\mathbf{v}_t^h + \mathbf{c}\right\|_{\sigma_t^h}^2 dt = \frac{1}{h}\sum_{k=1}^n \mathscr{W}_{h\mathbf{c}}^2(\sigma_{t_{k-1}}, \sigma_{t_k}) \le \int_0^T \left\|\mathbf{v}_t^{\text{per}} + \mathbf{c}\right\|_{\sigma_t}^2 dt.$$

This, together with (3.36), yields

$$\int_0^T \|\mathbf{v}_t + \mathbf{c}\|_{\sigma_t}^2\, dt \le \int_0^T \left\|\mathbf{v}_t^{\text{per}} + \mathbf{c}\right\|_{\sigma_t}^2 dt,$$

and so by (3.8), we have $\mathbf{v} = \mathbf{v}^{\text{per}}$. □

This theorem has an interesting application when $\mathbf{v}^*$ is a special periodic velocity.

COROLLARY 3.23. *Let $\mathbf{c} \in \mathbb{R}^d$ and $\sigma^* \in AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$ with periodic velocity of $\mathbf{c}$-minimal norm denoted by $\mathbf{v}_{\mathbf{c}}^*$. There exists $\sigma \in AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$ such that*:

(i) $\sigma_0 = \sigma_0^*$;
(ii) *If $\mathbf{v}_{\mathbf{c}}$ is the $\mathbf{c}$-minimal norm periodic velocity for σ, then $[\sigma, \mathbf{v}_{\mathbf{c}}] \equiv [\sigma^*, \mathbf{v}_{\mathbf{c}}^*]$*;
(iii) *$\mathbf{v}_{\mathbf{c}}$ is also a full (in the $AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$-sense) velocity for σ*;
(iv) *If $\mathbf{c} = \mathbf{0}$, then $\mathbf{v}_{\mathbf{0}}$ is the full velocity of minimal norm of σ. Thus,*

$$|(\sigma^*)'|_{\mathbb{T}^d}(t) = \|\mathbf{v}_{\mathbf{0},t}^*\|_{\sigma_t^*} = \|\mathbf{v}_{\mathbf{0},t}\|_{\sigma_t} = |\sigma'|(t) \quad \text{for a.e. } t \in [0, T]. \tag{3.37}$$

PROOF. Let us consider the curve $\sigma \in AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$ given by Theorem 3.22. By Proposition 3.9, we produce a periodic velocity $\mathbf{w}$ for σ such that $[\sigma, \mathbf{w}] \equiv [\sigma^*, \mathbf{v}_{\mathbf{c}}^*]$. Clearly, $\mathbf{w}$ must be the $\mathbf{c}$-minimal norm periodic velocity for σ, otherwise we would take $\tilde{\mathbf{w}}$ to be said velocity and Proposition 3.9 would give a counterpart velocity $\tilde{\mathbf{w}}^*$ for σ^*, which, by Remark 2.6, would satisfy

$$\|\tilde{\mathbf{w}}_t^* + \mathbf{c}\|_{\sigma_t^*} = \|\tilde{\mathbf{w}} + \mathbf{c}\|_{\sigma_t} < \|\mathbf{w} + \mathbf{c}\|_{\sigma_t} = \|\mathbf{v}_{\mathbf{c}}^* + \mathbf{c}\|_{\sigma_{\mathbf{c}}^*}$$

for all t in a nonnegligible subset of $(0, T)$, which contradicts the $\mathbf{c}$-minimality of $\mathbf{v}_{\mathbf{c}}^*$.

Finally, consider the dual role (full and periodic) velocity given by Theorem 3.22 and denote it by $\mathbf{v}_{\mathbf{c}}$. As $\mathbf{w}$ is the $\mathbf{c}$-minimal norm periodic velocity for σ, we necessarily have

$$\|\mathbf{w}_t + \mathbf{c}\|_{\sigma_t} \le \|\mathbf{v}_{\mathbf{c},t} + \mathbf{c}\|_{\sigma_t} \quad \text{for a.e. } t \in (0, T). \tag{3.38}$$

According to (3.31) and due to $[\sigma, \mathbf{w}] \equiv [\sigma^*, \mathbf{v}_{\mathbf{c}}^*]$, we get

$$\int_0^T \|\mathbf{v}_{\mathbf{c},t} + \mathbf{c}\|_{\sigma_t}^2\, dt \le \int_0^T \|\mathbf{v}_{\mathbf{c},t}^* + \mathbf{c}\|_{\sigma_t^*}^2\, dt = \int_0^T \|\mathbf{w}_t + \mathbf{c}\|_{\sigma_t}^2\, dt,$$

which, in view of (3.38), implies

$$\|\mathbf{v}_{\mathbf{c},t} + \mathbf{c}\|_{\sigma_t} = \|\mathbf{v}_{\mathbf{c},t}^* + \mathbf{c}\|_{\sigma_t^*} = \|\mathbf{w}_t + \mathbf{c}\|_{\sigma_t} \quad \text{for a.e. } t \in (0, T).$$

Thus, for a.e. t we have $\mathbf{v}_{\mathbf{c},t} \equiv \mathbf{w}_t$, σ_t-a.e., which means (i)–(iii) are proved. To prove (iv), note that for a.e. $t \in [0, T]$ we have

$$|\sigma'|(t) \leq \|\mathbf{v}_{\mathbf{0},t}\|_{\sigma_t} = \|\mathbf{v}^*_{\mathbf{0},t}\|_{\sigma^*_t} = |(\sigma^*)'|_{\mathbb{T}^d}(t) = |\sigma'|_{\mathbb{T}^d}(t).$$

But the metric inequality $\mathscr{W} \leq W_2$ implies the obvious $|\sigma'| \geq |\sigma'|_{\mathbb{T}^d}$; thus the display above consists of equalities only. In particular, $|\sigma'|(t) = \|\mathbf{v}_{\mathbf{0},t}\|_{\sigma_t}$ for a.e. $t \in [0, T]$; i.e., $\mathbf{v}_\mathbf{0}$ is the minimal norm full velocity of σ. □

DEFINITION 3.24. Let $\sigma \in AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$ and $\mathbf{v}$ be a periodic velocity. We say that the pair $(\tilde{\sigma}, \tilde{\mathbf{v}})$ is a *lift* of $(\sigma, \mathbf{v})$ if:

(i) $\tilde{\sigma} \in AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$;
(ii) $\tilde{\sigma}_0 = \sigma_0$;
(iii) $[\sigma, \mathbf{v}] \equiv [\tilde{\sigma}, \tilde{\mathbf{v}}]$;
(iv) $\tilde{\mathbf{v}}$ is a dual-role velocity for $\tilde{\sigma}$ (both full and periodic).

Note that Corollary 3.23 proves that any pair $[\sigma, \mathbf{v}] \in \mathscr{C}(0, T; \mathscr{C}\mathscr{P}(\mathbb{T}^d))$ such that $\mathbf{v}$ is a $\mathbf{c}$-minimal periodic velocity for σ (for some $\mathbf{c} \in \mathbb{R}^d$) admits at least one lift. It is not clear if this statement is true for general periodic velocities $\mathbf{v}$.

PROPOSITION 3.25. *Let* $[\sigma, \mathbf{v}] \in \mathscr{C}(0, T; \mathscr{C}\mathscr{P}(\mathbb{T}^d))$ *and let* $(\tilde{\sigma}, \tilde{\mathbf{v}})$ *be a lift of* $(\sigma, \mathbf{v})$. *Then*

$$\int_{\mathbb{R}^d} x\tilde{\sigma}(t, dx) = \int_{\mathbb{R}^d} x\sigma_0(dx) + \int_0^t \int_{\mathbb{R}^d} \mathbf{v}_\tau \, d\sigma_\tau \, d\tau \quad \text{for all } t \in [0, T]. \tag{3.39}$$

In other words, the mean (*center of mass*) *of* $\tilde{\sigma}$ *is lift invariant.*

PROOF. The equality $[\sigma, \mathbf{v}] \equiv [\tilde{\sigma}, \tilde{\mathbf{v}}]$ implies

$$\int_{\mathbb{R}^d} \tilde{\mathbf{v}}_t \, d\tilde{\sigma}_t = \int_{\mathbb{R}^d} \mathbf{v}_t \, d\sigma_t \quad \text{for a.e. } t \in [0, T]. \tag{3.40}$$

But $\tilde{\mathbf{v}}$ is also a full velocity for $\tilde{\sigma}$. By the definition of a full velocity [2], we have that

$$[0, T] \ni t \to \int_{\mathbb{R}^d} \phi(x)\sigma(t, dx)$$

is absolutely continuous for any continuous function ϕ of at most quadratic growth. Furthermore,

$$\frac{d}{dt}\int_{\mathbb{R}^d} \phi(x)\tilde{\sigma}(t, dx) = \int_{\mathbb{R}^d} \nabla\phi(x) \cdot \tilde{\mathbf{v}}_\mathbf{c}(t, x)\tilde{\sigma}(t, dx) \quad \text{for a.e. } t \in [0, T].$$

If we apply this to $\phi_i(x) = x_i$, $i = 1, \ldots, d$, we get

$$\frac{d}{dt}\int_{\mathbb{R}^d} x\tilde{\sigma}(t, dx) = \int_{\mathbb{R}^d} \tilde{\mathbf{v}}(t, x)\tilde{\sigma}(t, dx) \quad \text{for a.e. } t \in [0, T],$$

which, by the absolute continuity recalled above, is equivalent to

$$\int_{\mathbb{R}^d} x\tilde{\sigma}(t,dx) = \int_{\mathbb{R}^d} x\sigma(0,dx) + \int_0^t \int_{\mathbb{R}^d} \tilde{\mathbf{v}}(\tau,x)\tilde{\sigma}(\tau,dx)d\tau \quad \text{for all } t \in [0,T].$$

We use (3.40) to finish the proof. □

This invariance allows us to make the following definition:

DEFINITION 3.26. Let $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ and $\mathbf{v}$ be a periodic velocity for σ. We say that the speed curve $(\sigma,\mathbf{v})$ has *rotation vector* $V \in \mathbb{R}^d$ if

$$\lim_{t\to\infty} \frac{1}{t} \int_{\mathbb{R}^d} x\tilde{\sigma}(t,dx) = V$$

for a lift $(\tilde{\sigma},\tilde{\mathbf{v}})$ of $(\sigma,\mathbf{v})$.

Note that no lift may exist, or the limit above may not exist. In each case, the rotation vector is undefined. It should be noted that if $\sigma_t = \delta_{x(t)}$, where $x(t) \in \mathbb{T}^d$, then $(\sigma,\mathbf{v})$ has rotation vector V if and only if $(x(t),\dot{x}(t))$ has rotation vector V.

3.6 The Chain Rule

The proof of proposition 8.4.6 in [2] can be reproduced to obtain the following:

PROPOSITION 3.27. *Let $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$ and let $\mathbf{v}$ be the velocity of minimal norm associated to σ in the periodic sense and characterized by $\mathbf{v}_t \in \mathscr{T}_{\sigma_t}\mathscr{P}(\mathbb{T}^d)$. Then, for almost every $t \in (0,T)$ the following properties hold: for any $\gamma_h \in \Gamma(\sigma_t,\sigma^*_{t+h})$ such that*

$$\sigma^*_{t+h} \in [\sigma_{t+h}] \quad \text{and} \quad \gamma \in \Gamma_{\text{per}}(\sigma_t,\sigma^*_{t+h}),$$

we have

$$\lim_{h\to 0} \left(\pi_1, \frac{\pi_2-\pi_1}{h}\right)_{\#} \gamma_h = (\mathrm{id}\times\mathbf{v}_t)_{\#}\sigma_t \tag{3.41}$$

and

$$\lim_{h\to 0} \frac{\mathscr{W}\big(\sigma_{t+h},(\mathrm{id}+h\mathbf{v}_t)_{\#}\sigma_t\big)}{|h|} = 0. \tag{3.42}$$

LEMMA 3.28. *Let $U \in C^1(\mathscr{P}(\mathbb{T}^d))$, $\sigma \in AC_2(0,T;\mathscr{P}(\mathbb{T}^d))$, and $\mathbf{v}$ be the velocity of minimal norm associated to σ in the periodic sense. Then $t \to U(\sigma_t)$ is in $W^{1,2}(0,T)$ and, in the sense of distributions, we have*

$$\frac{d}{dt}U(\sigma_t) = \langle \nabla_w U(\sigma_t), \mathbf{v}_t\rangle_{\sigma_t}. \tag{3.43}$$

PROOF. Since U is Lipschitz, $t \to U(\sigma_t)$ is in $W^{1,2}(0,T)$. To prove the theorem, it remains to show that (3.43) holds almost everywhere. Let t be such that (3.41) and (3.42) hold and

$$\lim_{h\to 0} \frac{\mathscr{W}(\sigma_t, \sigma_{t+h})}{|h|} = |\sigma'|_{\mathbb{T}^d}(t). \tag{3.44}$$

Set $\xi := \nabla_w U(\sigma_t)$. For $h \in \mathbb{R}$ small enough, we first use Lemma 2.1 to choose $\sigma^*_{t+h} \in [\sigma_{t+h}]$, and, second, choose $\gamma_h \in \Gamma_{\text{per}}(\sigma_t, \sigma^*_{t+h})$. By the definition of ξ, there exists $\epsilon : \mathbb{R} \to \mathbb{R}$ continuous at 0 such that $\epsilon(0) = 0$ and

$$\begin{aligned} &\left| \frac{U(\sigma_{t+h}) - U(\sigma_t)}{h} - \langle \xi, \mathbf{v}_t \rangle_{\sigma_t} \right| \\ &\quad \le \epsilon\Big(\mathscr{W}(\sigma_t, \sigma_{t+h})\Big) \frac{\mathscr{W}(\sigma_t, \sigma_{t+h})}{|h|} \\ &\qquad + \left| \langle \xi, \mathbf{v}_t \rangle_{\sigma_t} - \int_{\mathbb{R}^d \times \mathbb{R}^d} \xi(x) \cdot \frac{y-x}{h} \gamma_h(dx, dy) \right| \end{aligned} \tag{3.45}$$

Letting h tend to 0 in (3.45) and exploiting (3.41) and (3.44), we conclude that (3.43) holds at t. □

4 Elements of Weak KAM theory

We assume we are given a function

$$\mathscr{K} \in C^1(\mathscr{P}(\mathbb{T}^d)). \tag{4.1}$$

We write $\mathscr{K}$ as the difference of its positive and negative parts: $\mathscr{K} = \mathscr{K}^+ - \mathscr{K}^-$. As $\mathscr{K}$ is continuous and $\mathscr{P}(\mathbb{T}^d)$ is compact (cf. Lemma 2.3), $\mathscr{K}^\pm$ attains its maximum, which we denote by $\mathscr{K}^\pm_\infty$.

We define the Lagrangian

$$\begin{gathered} L(\mu, \mathbf{w}) = \frac{1}{2}\Psi(\mu, \mathbf{w}) - \mathscr{K}(\mu), \quad \Psi(\mu, \mathbf{w}) := \|\mathbf{w}\|^2_\mu, \\ \mu \in \mathscr{P}_2(\mathbb{R}^d),\ \mathbf{w} \in L^2(\mu). \end{gathered}$$

We fix $\mathbf{c} \in \mathbb{R}^d$ and define

$$L_{\mathbf{c}}(\mu, \mathbf{w}) = L(\mu, -\mathbf{w}) + \mathbf{c} \cdot \int_{\mathbb{R}^d} \mathbf{w}\, d\mu.$$

Remark 4.1. By Remark 2.6, Ψ, L, and $L_{\mathbf{c}}$ are defined on the quotient space $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$. One exploits the variational formulation (2.7) to obtain that Ψ is lower-semicontinuous on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$. Since $\mathscr{F}$ is continuous, we conclude that L is lower-semicontinuous on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$.

Let $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$, and let $\mathbf{v}$ be a velocity associated to σ in the periodic sense. We exploit again the variational formulation (2.7) to show that $t \to \|\mathbf{v}_t\|_{\sigma_t}$ is lower-semicontinuous at almost every point $t \in (0,T)$. Hence it is a measurable function. Similarly, $t \to \langle \mathbf{c}, \mathbf{v}_t\rangle_{\sigma_t}$ is measurable. The function $t \to \mathscr{K}(\sigma_t)$ is continuous as the composition of two continuous functions. We conclude that $t \to L(\sigma_t, \mathbf{v}_t), L_{\mathbf{c}}(\sigma_t, \mathbf{v}_t)$ are measurable.

4.1 Definition of Weakly Invariant Measures

Let m be a Borel probability measure on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$. We say that m is weakly invariant and of rotation number $V \in \mathbb{R}^d$ if

$$\int_{\mathscr{C}\mathscr{P}(\mathbb{T}^d)} \left(\int_{\mathbb{R}^d} \mathbf{v}\, d\mu \right) m(d\mu, d\mathbf{v}) = V \tag{4.2}$$

and

$$\int_{\mathscr{C}\mathscr{P}(\mathbb{T}^d)} \delta U \, dm = 0 \tag{4.3}$$

for all $U \in C^1(\mathscr{P}(\mathbb{T}^d))$ such that

$$(\mu, \mathbf{v}) \to \delta U(\mu, \mathbf{v}) =: \langle \nabla_w U(\mu), \mathbf{v}\rangle_\mu$$

is continuous.

Remark 4.2. Note that $\mu \to \delta U(\mu, \nabla_w U(\mu)) = \|\nabla_w U(\mu)\|^2_\mu$ is continuous as a composition of two continuous functions. Since $\mathscr{P}(\mathbb{T}^d)$ is compact, we conclude that $\mu \to \|\nabla_w U(\mu)\|_\mu$ is bounded.

LEMMA 4.3. *Let $\sigma \in AC^2(0,T;\mathscr{P}(\mathbb{T}^d))$, let $\mathbf{v}$ be a velocity associated to σ in the periodic sense, and let C be such that $\|\mathbf{v}_t\|_{\sigma_t} \le C$ for all $t \in [0,T]$.*

(i) *If F is continuous on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$, then $t \to F(\sigma_t, \mathbf{v}_t)$ is Lebesgue measurable.*

(ii) *It makes sense then to define the measures m_T on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$ by setting*

$$\int_{\mathscr{C}\mathscr{P}(\mathbb{T}^d)} F \, dm_T = \frac{1}{T}\int_0^T F(\sigma_t, \mathbf{v}_t) dt \tag{4.4}$$

for any F continuous on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$. Here $\mathbb{B}_C$ is the compact set defined in Lemma 2.9.

PROOF.

(i) Due to the continuity of F, it suffices to show that $(0,T) \ni t \to (\sigma_t, \mathbf{v}_t) \in \mathscr{C}\mathscr{P}(\mathbb{T}^d)$ is Lebesgue-measurable. By the definition of the topology of $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$, since $t \to [\sigma_t]$ is continuous in the $\mathscr{W}$-metric, it is enough to prove that the function $t \to \langle \xi, \mathbf{v}_t\rangle_{\sigma_t}$ is Lebesgue-measurable for all $\xi \in C(\mathbb{T}^d;\mathbb{R}^d)$. We now refer to

Corollary 3.23 in order to replace $\langle \xi, \mathbf{v}_t \rangle_{\sigma_t}$ by $\langle \xi, \tilde{\mathbf{v}}_t \rangle_{\tilde{\sigma}_t}$, where $(\tilde{\sigma}, \tilde{\mathbf{v}})$ is the lift indicated there for $\mathbf{c} = \mathbf{0}$ (so that $\tilde{\mathbf{v}}$ is its minimal norm velocity).

Consider now the optimal maps $\nabla\Phi(t, \cdot)$ pushing the Lebesgue measure restricted to $\mathcal{Q}$ forward to $\tilde{\sigma}_t$ for all $t \in [0, T]$. By lemma 4.2 and remark 4.4 in [8], the map $(0, T) \times \mathcal{Q} \ni (t, x) \to \nabla\Phi(t, x)$ is $\mathcal{L}^{1+d}$-a.e. equal to a Borel map, say Ψ. Thus, the map

$$(0, T) \times \mathcal{Q} \ni (t, x) \to \xi(\Psi(t, x)) \cdot \tilde{\mathbf{v}}_t(\Psi(t, x))$$

is Borel as a composition of Borel maps. Since $\tilde{\mathbf{v}}_t \in L^2(\tilde{\sigma}_t)$ for a.e. $t \in (0, T)$ and $t \to \|\tilde{\mathbf{v}}_t\|_{\tilde{\sigma}_t}$ lies in $L^2(0, T)$, we deduce that $|\tilde{\mathbf{v}} \circ \Lambda|^2$ is integrable over $(0, T) \times \mathcal{Q}$, where $\Lambda(t, x) := (t, \Psi(t, x))$. But ξ is bounded, so the map in the display above is integrable as well over $(0, T) \times \mathcal{Q}$. We apply Fubini's theorem to deduce that $t \to \langle \xi, \tilde{\mathbf{v}}_t \rangle_{\tilde{\sigma}_t}$ is integrable over $(0, T)$, which, in particular, yields the conclusion (since $\xi \in C(\mathbb{T}^d; \mathbb{R}^d)$ is arbitrary).

(ii) By Lemma 2.9, $\mathbb{B}_C$ is a compact Hausdorff space that contains the range of $t \to (\sigma_t, \mathbf{v}_t)$. By the Riesz representation theorem, the linear functional on $C(\mathbb{B}_C,)$ $F \to (1/T)\int_0^T F(\sigma_t, \mathbf{v}_t)dt$, defines uniquely a Borel measure m_T on $\mathbb{B}_C$. We extend that measure to $\mathcal{C}\mathcal{P}(\mathbb{T}^d)$ by requiring that it be null on the complement of $\mathbb{B}_C$. Observe that m_T satisfies (4.4). □

LEMMA 4.4. *Suppose that* σ *is as in Lemma* 4.3 *and the* L^∞*-norm of* $t \to \|\mathbf{v}_t\|_{\sigma_t} \in L^\infty(0, T)$ *is bounded by a constant* C *independent of* T. *Then*:

(i) $\{m_T\}_{T>0}$ *has at least one cluster point* m *for the narrow convergence topology.*

(ii) *All cluster points* m *satisfy* (4.3).

PROOF.

(i) Recall that m_T is supported by the compact set $\mathbb{B}_C$. Since $\mathbb{B}_C$ is a Polish space, in light of remark 5.1.5 in [2], we conclude that $\{m_T\}_{T>0}$ is precompact for the narrow convergence. Thus, there exists an increasing sequence $\{T_n\}_n$ converging to ∞ such that $\{m_{T_n}\}_n$ converges to a Borel probability m on $\mathbb{B}_C$.

(ii) Let m be a point of accumulation as in (i), and let $U \in C^1(\mathcal{P}(\mathbb{T}^d))$ such that δU is continuous. Since δU is continuous on $\mathbb{B}_C$, we use Lemma 3.28 to conclude that

$$\int_{\mathbb{B}_C} \delta U \, dm = \lim_{n\to\infty} \int_{\mathbb{B}_C} \delta U(\mu, \mathbf{v}) m_{T_n}(d\mu, d\mathbf{v}) = \lim_{n\to\infty} \frac{U(\sigma_{T_n}) - U(\sigma_0)}{T_n} = 0.$$

This concludes the proof of the lemma. □

Example 4.5. Let $V \in \mathbb{R}^d$, and define $\sigma_t := \delta_{tV}$ and $\mathbf{v}_t := V$.

(i) The map $t \to F(\sigma_t, \mathbf{v}_t)$ is continuous as a composition of continuous functions.

(ii) By Lemma 2.9, the set $\mathbb{B}_C$ is a compact Hausdorff space that contains the range of $t \to (\sigma_t, V)$. By the Riesz representation theorem there exists a unique Borel measure m_T on $\mathbb{B}_C$ such that

$$\int_{\mathbb{B}_C} F\, dm_T = \frac{1}{T}\int_0^T F(\sigma_t, \mathbf{v}_t)dt \tag{4.5}$$

for $F \in C(\mathbb{B}_C)$. We extend this measure to the entire $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$ by setting it equal to 0 on the complement of $\mathbb{B}_C$. Observe that (4.5) still holds for $F \in C(\mathscr{C}\mathscr{P}(\mathbb{T}^d))$.

(iii) We have

$$\int_{\mathscr{C}\mathscr{P}(\mathbb{T}^d)} \left(\int_{\mathbb{R}^d} \mathbf{v}\, d\mu\right) m_T(d\mu, d\mathbf{v}) = V. \tag{4.6}$$

We may choose an unbounded increasing sequence $\{T_n\}_n$ such that $\{m_{T_n}\}_n$ converges narrowly to some Borel m on $\mathbb{B}_C$. Since $(\mu, \mathbf{v}) \to E(\mu, \mathbf{v}) := \int_{\mathbb{R}^d} \mathbf{v}\, d\mu$ is continuous, we use (4.6) to obtain

$$\int_{\mathbb{B}_C} \left(\int_{\mathbb{R}^d} \mathbf{v}\, d\mu\right) m(d\mu, d\mathbf{v}) = \lim_{n\to\infty} \int_{\mathbb{B}_C} \left(\int_{\mathbb{R}^d} \mathbf{v}\, d\mu\right) m_{T_n}(d\mu, d\mathbf{v}) = V.$$

In conclusion, for each $V \in \mathbb{R}^d$ there exists a weakly invariant measure μ of rotation vector V.

4.2 Existence of Minimal Weakly Invariant Measures

LEMMA 4.6. *Let F be a continuous function on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$ such that there exists $C_1 > 0$ such that $|F(\mu, \mathbf{v})| \le C_1\|\mathbf{v}\|_\mu$ on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$. Let $C_2 > 0$ and define $\mathscr{S}$ to be the set of Borel probability measures m on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$ such that*

$$\int_{\mathscr{C}\mathscr{P}(\mathbb{T}^d)} F dm = V \quad \textit{and} \quad \int_{\mathscr{C}\mathscr{P}(\mathbb{T}^d)} \Psi^2\, dm \le C_2.$$

Then every sequence in $\mathscr{S}$ admits a subsequence that converges narrowly in $\mathscr{S}$.

PROOF. Let $\{m_n\}_n \subset \mathscr{S}$. For each $\lambda \in \mathbb{R}$, Lemma 2.9 asserts that $\{\Psi \le \lambda\}$ is a compact set. Since $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$ is a Polish space and

$$\sup_n \int_{\mathscr{C}\mathscr{P}(\mathbb{T}^d)} \Psi^2\, dm_n \le C_2 < \infty,$$

we use remark 5.1.5 in [2] to conclude that $\{m_n\}_n$ admits a subsequence—which we still label $\{m_n\}_n$—that converges narrowly to some Borel probability m on

$\mathscr{CP}(\mathbb{T}^d)$. Remark 4.1 asserts that Ψ is lower-semicontinuous. Thus (cf., e.g., [2, sec. 5.1.1])

$$\int_{\mathscr{CP}(\mathbb{T}^d)} \Psi^2 \, dm \le \liminf_{n\to\infty} \int_{\mathscr{CP}(\mathbb{T}^d)} \Psi^2 \, dm_n \le C_2. \tag{4.7}$$

Let $f_r \in C(\mathbb{R})$ be defined by

$$f_r(t) = \begin{cases} r & \text{if } t \ge r, \\ t & \text{if } |t| \le r, \\ -r & \text{if } t \le -r. \end{cases}$$

We have

$$\int_{\mathscr{CP}(\mathbb{T}^d)} |F - f_r(F)| dm_n \le 2 \int_{\{|F|>r\}} |F| dm_n \le \frac{2}{r} \int_{\{|F|>r\}} |F|^2 \, dm_n \le \frac{2C_2}{r}.$$

Since m satisfies (4.7), the previous inequality holds if we substitute m_n by m. Consequently,

$$\left| \int F \, dm_n - \int F \, dm \right| \le \frac{4C_2}{r} + \left| \int f_r(F) dm_n - \int f_r(F) dm \right|,$$

where all the integrals above are over $\mathscr{CP}(\mathbb{T}^d)$. Since $f_r(F)$ is continuous and bounded, and $\{m_n\}_n$ converges narrowly to m, we have

$$\limsup_{n\to\infty} \left| \int_{\mathscr{CP}(\mathbb{T}^d)} F \, dm_n - \int_{\mathscr{CP}(\mathbb{T}^d)} F \, dm \right| \le \frac{4C_2}{r}. \tag{4.8}$$

We then let r tend to ∞ in (4.8) and use that $m_n \in \mathscr{S}$ to conclude that

$$V = \int_{\mathscr{CP}(\mathbb{T}^d)} F \, dm.$$

This, together with (4.7), yields that $m \in \mathscr{S}$. □

THEOREM 4.7. *Let $V \in \mathbb{R}^d$. Then there exists a Borel probability measure m that minimizes*

$$m \to J(m) := \int_{\mathscr{CP}(\mathbb{T}^d)} L \, dm$$

over the set of weakly invariant Borel probability measures of rotation vector V.

PROOF. In light of Example 4.5, the set of weakly invariant Borel probability measures of rotation vector V is nonempty. Pick an element m_*. There exists a minimizing sequence $\{m_n\}_n$ for J that we may assume satisfies

$$J(m_n) \le J(m_*).$$

Hence

$$\int_{\mathscr{C}\mathscr{P}(\mathbb{T}^d)} \Psi^2 \, dm_n \leq 2J(m_*) + 2\max \mathscr{F} =: C_2^2.$$

Let $U \in C^1(\mathscr{P}(\mathbb{T}^d))$ be such that δU is continuous. By Remark 4.2 there exists a constant C such that $\|\nabla_w U(\mu)\|_\mu \leq C$ on $\mathscr{P}(\mathbb{T}^d)$. Thus, $|\delta U(\mu, \mathbf{v})| \leq C\|\mathbf{v}\|_\mu$. If we set $a(\mu, \mathbf{v}) := \int_{\mathbb{T}^d} \mathbf{v}\, d\mu$, Hölder's inequality yields $|a(\mu, \mathbf{v})| \leq \|\mathbf{v}\|_\mu$. We have that $\{m_n\}_n$ is contained in $\mathscr{S}$, the set m of Borel probability measures on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$ such that (4.2), (4.3), and

$$\int_{\mathscr{C}\mathscr{P}(\mathbb{T}^d)} |\Psi|^2 \, dm \leq C_2^2$$

hold. By Lemma 4.6, up to a subsequence, $\{m_n\}_n$ converges narrowly to some m in $\mathscr{S}$. This proves that m is a weakly invariant measure of rotation vector V. By Remark 4.1, L is lower-semicontinuous on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$. Thus, (cf., e.g., [2, sec. 5.1.1])

$$J(m) \leq \liminf_{n\to\infty} J(m_n),$$

which implies that m minimizes J over the set of weakly invariant Borel probability measures of rotation vector V. □

4.3 Actions

If $[\sigma] \in AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$ and $\mathbf{v}$ is a velocity associated to σ in the periodic sense, we define the actions

$$\mathscr{A}_T^\varepsilon(\sigma, \mathbf{v}) := \int_0^T e^{-\varepsilon t} L_{\mathbf{c}}(\sigma_t, \mathbf{v}_t)dt, \quad \mathscr{A}_T(\sigma, \mathbf{v}) := \int_0^T L_{\mathbf{c}}(\sigma_t, \mathbf{v}_t)dt.$$

If, in addition,

$$t \to e^{-\varepsilon t}\|\mathbf{v}\|_{\sigma_t}^2 \in L^1(0, \infty), \tag{4.9}$$

then we define

$$\mathscr{A}^\varepsilon(\sigma, \mathbf{v}) := \int_0^\infty e^{-\varepsilon t} L_{\mathbf{c}}(\sigma_t, \mathbf{v}_t)dt.$$

CONVENTION. Let $[\sigma^*, \mathbf{v}^*] \in \mathscr{C}(0, T; \mathscr{C}\mathscr{P}(\mathbb{T}^d))$. By Corollary 3.23 and Proposition 3.9, we may choose $\sigma \in AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$ such that $\sigma \in [\sigma^*]$, and $\mathbf{v}$ a velocity associated to σ in the periodic sense such that $[\sigma, \mathbf{v}] \equiv [\sigma^*, \mathbf{v}^*]$. Thus, unless explicitly stated otherwise, we tacitly assume that whenever we write $[\sigma, \mathbf{v}]$, we have $\sigma \in AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$.

Remark 4.8.

(i) By Remark 2.6, if $[\mu, \xi_1] = [\nu, \xi_2] \in \mathscr{C}\mathscr{P}(\mathbb{T}^d)$, then $L_{\mathbf{c}}(\mu, \xi_1) = L_{\mathbf{c}}(\nu, \xi_2)$, and so $L_{\mathbf{c}}$ is well-defined on the quotient space $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$. As a consequence, if $(\sigma, \mathbf{v})$ and $(\sigma^*, \mathbf{v}^*)$ are as in Definition 3.8, then for almost every

$t \in (0,T)$, we have $L_{\mathbf{c}}(\sigma_t, \mathbf{v}_t) = L_{\mathbf{c}}(\sigma_t^*, \mathbf{v}_t^*)$. Thus, for each $\varepsilon \in (0,1)$, the actions defined below are well-defined on $\mathscr{C}(0,T;\mathscr{C}\mathscr{P}(\mathbb{T}^d))$.

(ii) Let $\sigma \in AC^2(0,T;\mathscr{P}_2(\mathbb{R}^d))$ and let $\mathbf{v}$ be a velocity associated to σ. If $\mathbf{w}_t$ is the projection of $\mathbf{v}_t$ onto $L^2(\mathbb{T}^d, \sigma_t)$, then

$$\mathbf{c} \cdot \int_{\mathbb{R}^d} \mathbf{v}_t \, d\sigma_t = \mathbf{c} \cdot \int_{\mathbb{R}^d} \mathbf{w}_t \, d\sigma_t \quad \text{for all } \mathbf{c} \in \mathbb{R}^d.$$

This, together with Corollary 3.11, proves that $\mathbf{w}$ is a velocity associated to σ in the periodic sense and $\mathscr{A}^\varepsilon(\sigma, \mathbf{w}) \le \mathscr{A}^\varepsilon(\sigma, \mathbf{v})$.

For $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$ we consider the costs for transporting μ onto ν within time $T > 0$:

$$\mathscr{C}_T^\varepsilon(\mu,\nu) = \inf_{[\sigma,\mathbf{v}]} \{\mathscr{A}_T^\varepsilon(\sigma,\mathbf{v}) : \sigma_0 = \mu, \sigma_T \in [\nu],\ [\sigma,\mathbf{v}] \in \mathscr{C}(0,T;\mathscr{C}\mathscr{P}(\mathbb{T}^d))\},$$

$$\mathscr{C}_T(\mu,\nu) = \inf_{[\sigma,\mathbf{v}]} \{\mathscr{A}_T(\sigma,\mathbf{v}) : \sigma_0 = \mu, \sigma_T \in [\nu],\ [\sigma,\mathbf{v}] \in \mathscr{C}(0,T;\mathscr{C}\mathscr{P}(\mathbb{T}^d))\}.$$

4.4 Value Functions Depending on a Parameter

For $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ we set

$$V^\varepsilon(\mu) := \inf_{(\sigma,\mathbf{v})} \{\mathscr{A}^\varepsilon(\sigma,\mathbf{v}) : (\sigma,\mathbf{v}) \in \mathscr{C}_\varepsilon(0,\infty;\mathscr{C}\mathscr{P}(\mathbb{T}^d)),\ [\sigma_0] = [\mu]\}.$$

By Remark 4.8

$$\begin{aligned} V^\varepsilon(\mu) &= \inf_{(\sigma,\mathbf{v})} \{\mathscr{A}^\varepsilon(\sigma,\mathbf{v}) : (\sigma,\mathbf{v}) \in \mathscr{C}_\varepsilon(0,\infty;\mathscr{T}\mathscr{P}_2(\mathbb{R}^d)), \sigma_0 = \mu\} \\ &= V^\varepsilon(\overline{\mu}) \end{aligned} \tag{4.10}$$

if $[\mu] = [\overline{\mu}]$.

Remark 4.9. If $\mu \in \mathscr{P}_2(\mathbb{R}^d)$, then

$$-\frac{|\mathbf{c}|^2/2 + \mathscr{K}_\infty^+}{\varepsilon} \le V^\varepsilon(\mu) \le \frac{2 + 2|\mathbf{c}| + \mathscr{K}_\infty^-}{\varepsilon}.$$

PROOF. The lower bound results from the fact that $L_{\mathbf{c}} \ge -(|\mathbf{c}|^2/2 + \mathscr{K}_\infty^+)$. Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and denote by ν_0 the Lebesgue measure restricted to $[0,1]^d$. We define σ to be the geodesic of constant speed (cf. [2]) starting at $\overline{\mu}$ at $t = 0$ and ending at ν_0 at $t = 1$. Let $\mathbf{v}$ be the associated velocity of minimal norm. If we set

$$\sigma_t = \nu_0, \quad \mathbf{v}_t = -\mathbf{c} \quad \text{for } t \in (1,\infty),$$

then

$$L_{\mathbf{c}}(\sigma_t, \mathbf{v}_t) = -\frac{1}{2}|\mathbf{c}|^2 + \mathscr{K}_\infty^- \le \mathscr{K}_\infty^-.$$

Using that $\|\mathbf{v}_t\|_{\sigma_t}^2 = W_2(\overline{\mu}, \nu_0) \le 2\sqrt{d}$, we conclude that

$$L_{\mathbf{c}}(\sigma_t, \mathbf{v}_t) \le \frac{1}{2}\|\mathbf{v}_t\|_{\sigma_t}^2 + |\mathbf{c}|\|\mathbf{v}_t\|_{\sigma_t} + \mathscr{K}_\infty^- \le 2d(1 + |\mathbf{c}|) + \mathscr{K}_\infty^-.$$

These upper bounds on $L_{\mathbf{c}}(\sigma_t, \mathbf{v}_t)$ for the above particular pair $(\sigma, \mathbf{v})$ yield the required upper bound on $V^\varepsilon(\mu)$. □

The following quantity will be useful:

$$m(\mathbf{c}) := \max\left\{2d(1+|\mathbf{c}|) + \mathscr{K}_\infty^-, \frac{|\mathbf{c}|^2}{2} + \mathscr{K}_\infty^+\right\}. \tag{4.11}$$

THEOREM 4.10. *Let* $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$. *Then*

$$V^\varepsilon(\mu) = \inf_{\nu^* \in \mathscr{P}_2(\mathbb{R}^d)} \mathscr{C}_T^\varepsilon(\mu, \nu^*) + e^{-\varepsilon T} V^\varepsilon(\nu^*) \tag{4.12}$$

and

$$|V^\varepsilon(\mu) - V^\varepsilon(\nu)| \le 2m(\mathbf{c})\mathscr{W}(\mu, \nu).$$

PROOF. The proof of (4.12) is a consequence of $[\sigma, \mathbf{v}] \in \mathscr{C}(0, \infty; \mathscr{C}\mathscr{P}(\mathbb{T}^d))$ implying

$$\mathscr{A}^\varepsilon(\sigma, \mathbf{v}) = \mathscr{A}_T^\varepsilon(\sigma, \mathbf{v}) + e^{-\varepsilon T}\mathscr{A}^\varepsilon(\sigma^*, \mathbf{v}^*),$$

where

$$\sigma_t^* = \sigma_{T+t}, \quad \mathbf{v}_t^* := \mathbf{v}_{T+t} \quad \text{for } t \ge 0.$$

Next, assume without loss of generality that $V^\varepsilon(\mu) \ge V^\varepsilon(\nu)$. Let σ be a geodesic of constant speed connecting μ to ν within time T (cf. [2]) and let $\mathbf{v}$ be its velocity of minimal norm. We have

$$\mathscr{C}_T^\varepsilon(\mu, \nu) \le \mathscr{A}^\varepsilon(\sigma, \mathbf{v}) = \int_0^T e^{-\varepsilon t}\left(\frac{1}{2}\|\mathbf{v}_t\|_{\sigma_t}^2 + \mathbf{c} \cdot \int_{\mathbb{R}^d} \mathbf{v}_t \, d\sigma_t - \mathscr{K}(\sigma_t)\right) dt.$$

We use Hölder's inequality and the fact that $T\|\mathbf{v}_t\|_{\sigma_t} = W_2(\mu, \nu)$ to conclude that

$$\mathscr{C}_T^\varepsilon(\mu, \nu) \le T\left(\frac{1}{2}\frac{W_2^2(\mu, \nu)}{T^2} + |\mathbf{c}|\frac{W_2(\mu, \nu)}{T} + \mathscr{K}_\infty^+\right) dt.$$

If $\mu \ne \nu$ and $T = W_2(\mu, \nu)$, we conclude that

$$\mathscr{C}_T^\varepsilon(\mu, \nu) \le W_2(\mu, \nu)\left(\frac{1}{2} + |\mathbf{c}| + \mathscr{K}_\infty^+\right). \tag{4.13}$$

This, together with (4.12), implies

$$V^\varepsilon(\mu) - V^\varepsilon(\nu) \le V^\varepsilon(\nu)(e^{-\varepsilon W_2(\mu,\nu)} - 1) + W_2(\mu, \nu)\left(\frac{1}{2} + |\mathbf{c}| + \mathscr{K}_\infty^+\right),$$

which, by Remark 4.9, yields

$$\begin{aligned} |V^\varepsilon(\mu) - V^\varepsilon(\nu)| &\le V^\varepsilon(\nu)\varepsilon W_2(\mu, \nu) + W_2(\mu, \nu)\left(\frac{1}{2} + |\mathbf{c}| + \mathscr{K}_\infty^+\right) \\ &\le 2m(\mathbf{c})W_2(\mu, \nu). \end{aligned}$$

We use (cf. equation (4.10)) the continuity of $V^\varepsilon : \mathscr{P}(\mathbb{T}^d) \to \mathbb{R}$ and exploit Lemma 2.1 to get

$$|V^\varepsilon(\mu) - V^\varepsilon(\nu)| = \min_{\mu^* \in [\mu], \nu^* \in [\nu]} |V^\varepsilon(\mu^*) - V^\varepsilon(\nu^*)| \leq 2m(\mathbf{c})\mathscr{W}(\mu, \nu).$$

□

4.5 Lower-Semicontinuity Properties of Actions

COROLLARY 4.11. *Let $T > 0$ and $\mu_0, \mu_T \in \mathscr{P}_2(\mathbb{R}^d)$ be arbitrary. Then, there exists a path $[\sigma, \mathbf{v}] \in \mathscr{C}(0, T; \mathscr{C}\mathscr{P}(\mathbb{T}^d))$ such that $\sigma_0 = \mu_0$, $\sigma_T \in [\mu_T]$, $\sigma \in AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$, and*

$$\mathscr{C}_T(\mu_0, \mu_T) = \int_0^T L_{\mathbf{c}}(\sigma_t, \mathbf{v}_t)dt.$$

PROOF. Take a sequence

$$\{[\sigma^n, \mathbf{v}^n]\}_n \subset AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$$

such that $\{\mathscr{A}_T(\sigma^n, \mathbf{v}^n)\}_n$ converges to $\mathscr{C}_T(\mu_0, \mu_T)$. By Lemma 3.22, we can assume without loss of generality that $\sigma_0^n = \mu_0$, $[\sigma_T^n] = [\mu_T]$ and $\{\sigma^n\}_n \subset AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$. Thanks to Proposition 3.21 we may assume existence of a path $\sigma \in AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d))$ and a velocity $\mathbf{v}$ associated to σ in the periodic sense such that $\{\sigma_t^n\}_n$ converges to σ_t in $\mathscr{P}(\mathbb{T}^d)$ for all $t \in [0, T]$ and (3.30) holds. Since $\mathscr{K}$ is continuous, we have that $\mathscr{C}_T(\mu_0, \mu_T) \geq \int_0^T L_{\mathbf{c}}(\sigma_t, \mathbf{v}_t)dt$. □

Remark 4.12. By the definition of $L_{\mathbf{c}}$, we infer that $\mathbf{v}$ is the $\mathbf{c}$-minimal norm periodic velocity associated to σ. By Corollary 3.23, the speed curve $(\sigma, \mathbf{v})$ admits a lift in the sense of Definition 3.26.

4.6 Compactness Properties of the Value Function and a Fixed Point Theorem

Let V^ε be the value function defined in the previous subsection. As V^ε is a Lipschitz map on $\mathscr{P}(\mathbb{T}^d)$ it attains its minimum there. We set

$$U^\varepsilon := V^\varepsilon - \min V^\varepsilon.$$

LEMMA 4.13. *Up to a subsequence, $\{U^\varepsilon\}_\varepsilon$ converges uniformly to a function $U \in C(\mathscr{P}_2(\mathbb{T}^d))$ such that $|U(\mu) - U(\nu)| \leq 2m(\mathbf{c})\mathscr{W}(\mu, \nu)$ for $\mu, \nu \in \mathscr{P}_2(\mathbb{R}^d)$. The subsequence can be chosen so that $\{\varepsilon V^\varepsilon\}_\varepsilon$ converges uniformly to a constant function, which we denote by $\bar{H}(\mathbf{c})$.*

PROOF. It suffices to show the uniform convergence on $\mathscr{P}(\mathbb{T}^d)$. By Remark 4.9, $\{\varepsilon V^\varepsilon\}_\varepsilon$ is bounded on $\mathscr{P}(\mathbb{T}^d)$. Viewed as a function on that set, Theorem 4.10 ensures that the Lipschitz constant of $\varepsilon V^\varepsilon$ is no greater than $2\varepsilon m(\mathbf{c})$. The Ascoli-Arzelà theorem yields uniform convergence of a subsequence of $\{\varepsilon V^\varepsilon\}_\varepsilon$ to a function whose Lipschitz constant is null. That function must be a constant on $\mathscr{P}(\mathbb{T}^d)$.

Since the minimum value of U^ε is 0 and U^ε is a $2m(\mathbf{c})$-Lipschitz as a function on $\mathscr{P}(\mathbb{T}^d)$ and a set whose diameter is smaller than or equal to $2\sqrt{d}$, we conclude that $\{U^\varepsilon\}_\varepsilon$ is bounded. The Ascoli-Arzelà theorems yield uniform convergence of a subsequence of the above subsequence to a function $U \in C(\mathscr{P}(\mathbb{T}^d))$ whose Lipschitz constant does not exceed $2m(\mathbf{c})$. □

THEOREM 4.14. *Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and let $T > 0$.*

(i) *We have*

$$U(\mu) = \inf_{\nu \in \mathscr{P}_2(\mathbb{R}^d)} \mathscr{C}_T(\mu, \nu) + U(\nu) + T\bar{H}(\mathbf{c}). \tag{4.14}$$

(ii) *Then there exists $(\sigma, \mathbf{v})$ (independent of T) such that $\sigma_0 = \mu$,*

$$\sigma \in AC^2(0, T; \mathscr{P}_2(\mathbb{R}^d)), \quad [\sigma, \mathbf{v}] \in \mathscr{C}(0, T; \mathscr{C}\mathscr{P}(\mathbb{T}^d)),$$

and

$$U(\mu) - U(\sigma_T) = \mathscr{C}_T(\mu, \sigma_T) + T\bar{H}(\mathbf{c}). \tag{4.15}$$

PROOF.

(i) For all $t \in [0, T]$ we have $1 - \varepsilon T \leq e^{-\varepsilon T} \leq e^{-\varepsilon t} \leq 1$, so

$$(1 - \varepsilon T)\mathscr{C}_T \leq \mathscr{C}_T^\varepsilon \leq \mathscr{C}_T. \tag{4.16}$$

By Theorem 4.10(i) we have

$$U^\varepsilon(\mu) = \min_\nu \mathscr{C}_T(\mu, \nu) + e^{-\varepsilon T} U^\varepsilon(\nu) + \min V^\varepsilon (e^{-\varepsilon T} - 1).$$

The uniform convergence of $\{U^\varepsilon\}_\varepsilon$ and $\{\varepsilon V^\varepsilon\}_\varepsilon$ provided by Lemma 4.13 yields the proof of (i).

(ii) We use Corollary 4.11 to inductively find $(\sigma^i, \mathbf{v}^i)$ such that

$$\sigma^i \in AC^2(i, (i+1); \mathscr{P}_2(\mathbb{R}^d)), \quad [\sigma^i, \mathbf{v}^i] \in \mathscr{C}(i, (i+1); \mathscr{C}\mathscr{P}(\mathbb{T}^d)),$$
$$\sigma^i_{i+1} = \sigma^{i+1}_{i+1}$$

and

$$U(\sigma^i_i) - U(\sigma^i_{i+1}) = \mathscr{C}_1(\sigma^i_i, \sigma^i_{i+1}) + \bar{H}(\mathbf{c}) \tag{4.17}$$

for $i = 0, 1, \ldots$. Set

$$\sigma_t = \sigma^i_t \ \text{ if } t \in [i, (i+1)] \quad \text{and} \quad [\mathbf{v}_t] = [\mathbf{v}^i_t] \ \text{ if } t \in (i, (i+1)).$$

Choose an integer $m \geq 0$ such that $m < T \leq m + 1$. As

$$\mathscr{C}_1(\sigma_m, \sigma_{m+1}) + \bar{H}(\mathbf{c}) \geq \big(\mathscr{C}_{T-m}(\sigma_m, \sigma_T) + (T-m)\bar{H}(\mathbf{c})\big) + \big(\mathscr{C}_{m+1-T}(\sigma_T, \sigma_{m+1}) + (m+1-T)\bar{H}(\mathbf{c})\big),$$

(i) and (4.17) yield

$$U(\sigma_m) - U(\sigma_T) = \mathscr{C}_{T-m}(\sigma_m, \sigma_T) + (T-m)\bar{H}(\mathbf{c}). \tag{4.18}$$

If $m = 0$, this completes the proof of the theorem. Assume in what follows that $m > 0$. By (4.17) and (4.18),

$$\begin{aligned} U(\sigma_0) - U(\sigma_T) &= \sum_{i=0}^{m-1} U(\sigma_i) - U(\sigma_{i+1}) + U(\sigma_m) - U(\sigma_T) \\ &= \sum_{i=0}^{m-1} \mathscr{C}_1(\sigma_i^i, \sigma_{i+1}^i) + \bar{H}(\mathbf{c}) + \mathscr{C}_{T-N}(\sigma_m, \sigma_T) + (T-m)\bar{H}(\mathbf{c}) \\ &\geq \mathscr{C}_T(\sigma_0, \sigma_T) + T\bar{H}(\mathbf{c}). \end{aligned}$$

This, together with (i), completes the proof. □

Remark 4.15. Equations (4.14) and (4.15) imply uniqueness for $\bar{H}(\mathbf{c})$. Let us outline the argument. Indeed, if there exists another weak KAM solution (U satisfying the properties listed in the above theorem is called a *weak KAM solution*) with its own constant α replacing $\bar{H}(\mathbf{c})$, then, according to (4.14), the optimal path for one solution (which gives the equation (4.15) for said solution) yields an inequality in (4.15) for the other solution. By subtraction and taking into account that the weak KAM solutions are bounded, we obtain (by letting $T \to \infty$) both $\alpha \geq \bar{H}(\mathbf{c})$ and $\alpha \leq \bar{H}(\mathbf{c})$. Thus, we may call $\bar{H}(\mathbf{c})$ the *effective Hamiltonian*.

THEOREM 4.16. *Let $\bar{H}$ be the function defined in Lemma* 4.13. *Then $\bar{H}$ is convex. Consequently, the set where $\bar{H}$ fails to be differentiable in the sense of Alexandroff is a set of null measure in $\mathbb{R}^d$.*

PROOF. For $\mathbf{c} \in \mathbb{R}^d$ arbitrary, we display the dependence in $\mathbf{c}$ in the value function $U(\cdot\,;\mathbf{c})$ and the curve $\sigma(\cdot\,;\mathbf{c})$ for which (4.15) holds. We are to show that if $\mathbf{c} = (\mathbf{c}_1 + \mathbf{c}_2)/2$, then

$$0 \leq \lambda =: \frac{\bar{H}(\mathbf{c}_1) + \bar{H}(\mathbf{c}_2)}{2} - \bar{H}(\mathbf{c}).$$

As

$$U(\mu; \mathbf{c}_i) - U(\sigma_T(\mathbf{c}); \mathbf{c}_i) \leq \mathscr{C}_T(\mu, \sigma_T(\mathbf{c})) + T\bar{H}(\mathbf{c}_i) \quad \text{for } i = 1, 2,$$

(4.15) implies

$$\begin{aligned} f(T) :=& \left(\frac{U(\mu;\mathbf{c}_1) + U(\mu;\mathbf{c}_2)}{2} - U(\mu;\mathbf{c})\right) \\ &- \left(\frac{U(\sigma_T(\mathbf{c});\mathbf{c}_1) + U(\sigma_T(\mathbf{c});\mathbf{c}_2)}{2} - U(\mu;\mathbf{c})\right) \leq T\lambda. \end{aligned} \tag{4.19}$$

By Lemma 4.13, the functions $U(\cdot\,;\mathbf{c}_1)$, $U(\cdot\,;\mathbf{c}_2)$, and $U(\cdot\,;\mathbf{c})$ are Lipschitz, and so f is bounded below. Hence (4.19) implies that $\lambda \geq 0$. □

THEOREM 4.17. *Let $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ and $(\sigma, \mathbf{v})$ be as in Theorem* 4.14. *If $\bar{H}$ is twice differentiable at* $\mathbf{c}$ *in the sense of Alexandroff, then*

$$\left| \nabla \bar{H}(\mathbf{c}) + \frac{1}{T} \int_0^T \left(\int_{\mathbb{R}^d} \mathbf{v}_t \, d\sigma_t \right) dt \right| \leq \frac{\bar{C}}{\sqrt{T}}.$$

Here $\bar{C}$ is a constant depending only on $\mathbf{c}$.

PROOF. The proof we give here is an adaptation of what is done in the finite-dimensional setting (cf., e.g., [17]). We use the notation $U(\cdot\,;\mathbf{c})$ and $\sigma(\cdot\,;\mathbf{c})$ as in the proof of Theorem 4.16. Assume that $\bar{H}$ is twice differentiable at $\mathbf{c}$. We denote by $\mathbf{v}$ a velocity associated to $\sigma(\cdot\,;\mathbf{c})$ in the periodic sense and set $\eta := \nabla \bar{H}(\mathbf{c})$.

Choose $\delta, C > 0$ depending on $\mathbf{c}$ such that

$$|\bar{H}(\mathbf{c}') - \bar{H}(\mathbf{c}) - \eta \cdot (\mathbf{c}' - \mathbf{c})| \leq C|\mathbf{c}' - \mathbf{c}|^2 \quad \forall |\mathbf{c} - \mathbf{c}'| \leq \delta. \tag{4.20}$$

We only consider in the following $\mathbf{c}' \in \mathbb{R}^d$ such that $|\mathbf{c} - \mathbf{c}'| \leq \delta$ and T such that the diameter of $\mathscr{P}(\mathbb{T}^d)$ is smaller than $T\delta^2$.

We use that

$$U(\mu; \mathbf{c}') - U(\sigma_T(\mathbf{c}); \mathbf{c}') \leq \int_0^T \big(L_{\mathbf{c}}(\sigma_t(\mathbf{c}), \mathbf{v}_t) + \bar{H}(\mathbf{c}')\big) dt \tag{4.21}$$

and

$$U(\mu; \mathbf{c}) - U(\sigma_T(\mathbf{c}); \mathbf{c}) = \int_0^T \big(L_{\mathbf{c}}(\sigma_t(\mathbf{c}), \mathbf{v}_t) + \bar{H}(\mathbf{c})\big) dt \tag{4.22}$$

to conclude that

$$\begin{aligned} U(\mu; \mathbf{c}) - U(\mu; \mathbf{c}') &\geq U(\sigma_T(\mathbf{c}); \mathbf{c}) - U(\sigma_T(\mathbf{c}); \mathbf{c}') \\ &\quad + \int_0^T (\langle \mathbf{c} - \mathbf{c}', \mathbf{v}_t \rangle_{\sigma_t} + \bar{H}(\mathbf{c}) - \bar{H}(\mathbf{c}')) dt. \end{aligned} \tag{4.23}$$

Hence, by (4.20),

$$\begin{aligned} U(\mu; \mathbf{c}) - U(\mu; \mathbf{c}') &\geq U(\sigma_T(\mathbf{c}); \mathbf{c}) - U(\sigma_T(\mathbf{c}); \mathbf{c}') \\ &\quad + \int_0^T (\langle \mathbf{c} - \mathbf{c}', \mathbf{v}_t + \eta \rangle_{\sigma_t} - C|\mathbf{c}' - \mathbf{c}|^2) dt. \end{aligned} \tag{4.24}$$

By Lemma 4.13 we may choose a constant $M > 0$ depending only on $\mathbf{c}$ such that the Lipschitz constant of $U(\cdot\,;\mathbf{c})$ and $U(\cdot\,;\mathbf{c}')$ is less than M. Thus,

$$(U(\mu; \mathbf{c}) - U(\sigma_T(\mathbf{c}); \mathbf{c})) + (U(\sigma_T(\mathbf{c}); \mathbf{c}') - U(\mu; \mathbf{c}')) \leq 2M\mathscr{W}(\mu, \sigma_T(\mathbf{c})). \tag{4.25}$$

We use (4.24) and (4.25) to get

$$\int_0^T (\langle \mathbf{c} - \mathbf{c}', \mathbf{v}_t + \eta \rangle_{d\sigma_t} - C|\mathbf{c}' - \mathbf{c}|^2) dt \leq 2M\mathscr{W}(\mu, \sigma_T(\mathbf{c})). \tag{4.26}$$

Now, choose

$$\mathbf{c}' := \mathbf{c} - \sqrt{\frac{\mathscr{W}(\mu, \sigma_T(\mathbf{c}))}{T}} \frac{\int_0^T dt \int_{\mathbb{R}^d} (\mathbf{v}_t + \eta) d\sigma_t}{|\int_0^T dt \int_{\mathbb{R}^d} (\mathbf{v}_t(\mathbf{c}) + \eta) d\sigma_t|}$$

and substitute the expression of $\mathbf{c}'$ in (4.26) to conclude that

$$\left| \int_0^T dt \int_{\mathbb{R}^d} (\mathbf{v}_t + \eta) d\sigma_t \right| \leq (2M + C) \sqrt{T \mathscr{W}(\mu, \sigma_T(\mathbf{c}))}. \tag{4.27}$$

$\square$

By Remark 4.12, we have the following:

COROLLARY 4.18. *Let* $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ *and* $(\sigma, \mathbf{v})$ *be as in Theorem* 4.14. *If* $\bar{H}$ *is twice differentiable at* $\mathbf{c}$ *in the sense of Alexandroff, then* $(\sigma, \mathbf{v})$ *has rotation vector* $-\nabla \bar{H}(\mathbf{c})$.

Remark 4.19. Note that for the Galilean invariant Lagrangian defined in [15], i.e., where

$$\mathscr{K}(\mu) = \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d} \mathscr{F}(x - y) \mu(dx) \mu(dy) \tag{4.28}$$

for some "nice" potential $\mathscr{F}$, we proved that $\bar{H}(\mathbf{c}) = |\mathbf{c}|^2/2$ for all $\mathbf{c} \in \mathbb{R}^d$. Thus, the conclusion of the above theorem is satisfied at every $\mathbf{c} \in \mathbb{R}^d$.

Remark 4.20. To each $(\sigma, \mathbf{v})$ as in the statement of Theorem 4.17 we may associate the probability measure γ_T defined on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$ as follows: if F is a continuous function on $\mathscr{C}\mathscr{P}(\mathbb{T}^d)$, then

$$\int_{\mathscr{C}\mathscr{P}(\mathbb{T}^d)} F \, d\gamma_T = \frac{1}{T} \int_0^T F(\sigma_t, \mathbf{v}_t) dt.$$

Corollary 4.18 yields that every point of accumulation (such points always exist!) of $\{\gamma_T\}_{T \geq 1}$ has rotation vector $-\nabla \bar{H}(\mathbf{c})$.

5 The Nonlinear Vlasov System

The goal of this section is to show that more can be said in an important particular case: even if the Euler-Lagrange equation associated with the Lagrangian L might fail to define a flow, its fully kinetic version does. We would like to show that the $\mathbf{c}$-calibrated curves given by Theorem 4.14 give rise to solutions for the fully kinetic PDE. Some measures coming directly from the measures in Remark 4.20 are invariant with respect to the Vlasov flow and have rotation vector $\mathbf{c}$ (the sign variation comes from considering the forward Lax-Oleinik semigroup in this section, rather than the backward semigroup). Theorem 5.6 and Corollary 5.7 provide interesting (and new, to our knowledge) asymptotic results on monokinetic solutions starting at any initial measure in $\mathscr{P}_2(\mathbb{R}^d \times \mathbb{R}^d)$ supported on d-dimensional graphs, which also have a prescribed asymptotic velocity.

If $\mathscr{K}$ is the one considered in (4.28), the requirement

$$\mathscr{F} \in C^{1,1}(\mathbb{T}^d) \tag{5.1}$$

is sufficient to ensure existence and uniqueness of solution for the nonlinear Vlasov system

$$\partial_t \mu + v \cdot \nabla_x \mu = \nabla_v \cdot [\mu(\nabla \mathscr{F} * \sigma)] \quad \text{where } \sigma = \operatorname{proj}_x \mu, \tag{5.2}$$

when μ_0 is prescribed. We point the reader to [1, 18] for issues like existence, uniqueness, and stability properties for the nonlinear Vlasov equation. We denote by Φ the Vlasov flow: $\mu_t = \Phi(t, \mu_0)$. Further assumptions on $\mathscr{F}$ will be made below.

5.1 Discrete Approximations of Paths

Here we indicate how to approximate any curve $\mu \in AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$ by curves of the form

$$\sigma_n(t) = \frac{1}{n^d} \sum_{i=1}^{n^d} \delta_{\mathbf{x}_n^i(t)},$$

where the curves $t \to \mathbf{x}_n^i(t) \in \mathbb{R}^d$ lie in $H^1(0, T; \mathbb{R}^d)$. The approximation is to be understood in the sense presented by Proposition 5.1 below.

Let us denote by

$$\mathscr{P}^n(\mathbb{T}^d) = \left\{ \frac{1}{n^d} \sum_{i=1}^{n^d} \delta_{z^i} : z^i \in \mathbb{T}^d,\ i = 1, \dots, n^d \right\}.$$

PROPOSITION 5.1. *Let $\mu \in AC^2(0, T; \mathscr{P}(\mathbb{T}^d))$ and $\varepsilon > 0$ be given. Then there exist a positive integer n and a curve $\mu^n \in AC^2(0, T; \mathscr{P}^n(\mathbb{T}^d))$ such that*

$$\begin{gathered} \sup_{t \in [0,T]} \mathscr{W}(\mu_t^n, \mu_t) \le \varepsilon \quad \text{and} \\ \int_0^T \left|(\mu_t^n)'\right|_{\mathbb{T}^d}^2(t)\,dt \le \varepsilon + \int_0^T \left|\mu'\right|_{\mathbb{T}^d}^2(t)\,dt. \end{gathered} \tag{5.3}$$

PROOF. We first periodically mollify μ and its periodic minimal norm velocity field $\mathbf{v}$ to $(\mu^\delta, \mathbf{v}^\delta)$ for $\delta > 0$ such that, according to Lemmata 3.15 and 3.16, we have

$$\begin{gathered} \mathbf{v}^\delta \in L^2(0, T; W^{1,2}(\mathbb{T}^d, \mathbb{R}^d)), \quad \mu^\delta = \rho^\delta \mathscr{L}^d, \quad \rho^\delta \in W^{1,1}\big((0, T) \times \mathbb{T}^d\big), \\ \sup_{t \in [0,T]} \mathscr{W}(\mu_t^\delta, \mu_t) \le \varepsilon/2, \end{gathered}$$

and

$$|(\mu^\delta)'|_{\mathbb{T}^d}(t) \le \|\mathbf{v}_t^\delta\|_{\mu_t^\delta} \le \|\mathbf{v}_t\|_{\mu_t} = |\mu'|_{\mathbb{T}^d}(t) \quad \text{for a.e. } t \in (0, T).$$

Since $\mathbf{v}^\delta$ and ρ^δ are smooth, it is then easy to show that for n sufficiently large we may choose $\mu^{\delta,n} \in AC^2(0,T;\mathscr{P}^n(\mathbb{T}^d))$ such that

$$\sup_{t\in[0,T]} \mathscr{W}(\mu_t^\delta,\mu_t^{\delta,n}) \le \varepsilon/2 \quad \text{and}$$
$$\int_0^T |(\mu^{\delta,n})'|^2_{\mathbb{T}^d}(t)dt \le \varepsilon/2 + \int_0^T |(\mu^\delta)'|^2_{\mathbb{T}^d}(t)dt$$

to conclude the proof. □

5.2 From Finite-Dimensional to Infinite-Dimensional Weak KAM

Here we add three assumptions on $\mathscr{F}$, i.e.,

$$\mathscr{F}(z) = \mathscr{F}(-z) \le \mathscr{F}(\mathbf{0}) = 0 \quad \text{for all } z \in \mathbb{R}^d. \tag{5.4}$$

The fact that $\mathscr{F}$ is even is not restrictive in any way, as we pointed out in the Introduction. One can easily check that $\mathscr{F}(\mathbf{0}) = 0$ can also be done away with, even though the conclusions should be modified accordingly (the effective Hamiltonian and the rotation vectors involved will take into account the maximum value $\mathscr{F}(\mathbf{0})$). This leaves $\mathbf{0}$ being a maximum point for $\mathscr{F}$ as the only serious restriction.

Remark 5.2. We have imposed the condition $\mathscr{F}(0) = \max \mathscr{F}$ only to argue that the function $U^{kn}(\cdot;\mathbf{c})$ defined in (5.10) when restricted to $\mathscr{P}^n(\mathbb{T}^d)$ coincides with $U^n(\cdot;\mathbf{c})$ (cf. [15]) (note that $\mathscr{P}^n(\mathbb{T}^d)$ is a subset of $\mathscr{P}^{kn}(\mathbb{T}^d)$ for all integers $k,n \ge 1$). This is called the *consistency* property. This can still be obtained under fewer restrictions on $\mathscr{F}$.

Let $n \ge 1$ be an integer and set $m := n^d$. The finite-dimensional weak KAM theory [4, 7] (cf. also [15]) ensures the following:

PROPOSITION 5.3.

(i) *For every* $\mathbf{c} \in \mathbb{R}^d$ *there exists a Lipschitz function* $\tilde{U}^n(\cdot;\mathbf{c}) : \mathbb{T}^{md} \to \mathbb{R}$ *such that for every* $T > 0$ *and* $x \in AC^2(0,\infty;\mathbb{T}^{md})$ *we have*

$$\tilde{U}^n(x(0);\mathbf{c}) - \tilde{U}^n(x(T);\mathbf{c}) \le \int_0^T \left\{ \frac{1}{2m}\sum_{i=1}^m |\dot{x}_i + \mathbf{c}|^2 - \frac{1}{2m^2}\mathscr{F}(x_i - x_j) \right\} dt. \tag{5.5}$$

(ii) *Given* $X \in \mathbb{T}^{md}$ *there exists* $x^n \in AC^2(0,\infty;\mathbb{T}^{md})$*, an optimizer in* (5.5)*, i.e., such that*

$$\tilde{U}^n(x^n(0);\mathbf{c}) - \tilde{U}^n(x^n(T);\mathbf{c}) = \int_0^T \left\{ \frac{1}{2m}\sum_{i=1}^m |\dot{x}_i^n + \mathbf{c}|^2 - \frac{1}{2m^2}\mathscr{F}(x_i^n - x_j^n) \right\} dt \tag{5.6}$$

for all $T > 0$. Here $x^n(0) = X$. The following Euler-Lagrange equations are satisfied:

$$\ddot{x}_i^n(t) = -\frac{1}{m}\sum_{j=1}^{m} \nabla \mathscr{F}(x_i^n(t) - x_j^n(t)), \;\; i = 1, \dots, m. \tag{5.7}$$

We set

$$\sigma^n = \frac{1}{m}\sum_{i=1}^{m} \delta_{x_i^n} \in AC^2(0, T; \mathscr{P}^n(\mathbb{T}^d)). \tag{5.8}$$

Note that the unique velocity associated to σ^n in the periodic sense is given by

$$\mathbf{v}_t^n(x_i^n(t)) = \dot{x}_i^n(t). \tag{5.9}$$

Due to the symmetric property of the actions, we minimize in (5.5); $\widetilde{U}^n(\,\cdot\,;\mathbf{c})$ is invariant under any permutation of m letters. Hence, it is meaningful to define

$$\begin{gathered} U^n(\,\cdot\,;\mathbf{c}) : \mathscr{P}^n(\mathbb{T}^d) \to \mathbb{R} \quad \text{given by } U^n(\mu;\mathbf{c}) := \widetilde{U}^n(X;\mathbf{c}), \\ \text{where } \mu := \frac{1}{m}\sum_{i=1}^{m} \delta_{X_i}. \end{gathered} \tag{5.10}$$

As in [15], one can show that property (5.4) implies that the Lipschitz constant of $U^n(\,\cdot\,;\mathbf{c})$ is bounded by a constant independent of n. By this and the consistency property (Remark 5.2), we see that there exists a Lipschitz functional $U(\,\cdot\,;\mathbf{c})$ defined on $\bigcup_{n\ge 1} \mathscr{P}^n(\mathbb{T}^d)$ (which is dense in $\mathscr{P}(\mathbb{T}^d)$ with the $\mathscr{W}$-induced topology) such that $U(\,\cdot\,;\mathbf{c}) \equiv U^n(\,\cdot\,;\mathbf{c})$ on $\mathscr{P}^n(\mathbb{T}^d)$. Its extension by continuity to $\mathscr{P}(\mathbb{T}^d)$ is therefore Lipschitz-continuous and we still denote it by $U(\,\cdot\,;\mathbf{c})$. The following result is immediate:

PROPOSITION 5.4.

(i) *One has*

$$U(\sigma(0);\mathbf{c}) \le U(\sigma(T);\mathbf{c}) + \int_0^T L_{\mathbf{c}}(\sigma(t), \mathbf{v}(t))dt + \frac{1}{2}T|\mathbf{c}|^2 \tag{5.11}$$

for all $\sigma \in AC^2(0, \infty; \mathscr{P}^n(\mathbb{T}^d))$ such that $\sigma(0) = \mu^n$ and all $T > 0$.

(ii) *Given $\mu^n \in \mathscr{P}^n(\mathbb{T}^d)$, there exists a path $\sigma^n \in AC^2(0, \infty; \mathscr{P}^n(\mathbb{T}^d))$ with $\sigma^n(0) = \mu^n$ such that equality is achieved in* (5.11) *for all $T > 0$.*

This leads to the announced infinite-dimensional weak KAM solution obtained by our dimensional blowup approach. Indeed, we have the following:

THEOREM 5.5.

(i) *Let $\mu \in \mathscr{P}(\mathbb{T}^d)$. One has*

$$U(\mu;\mathbf{c}) \le U(\sigma(T);\mathbf{c}) + \int_0^T L_{\mathbf{c}}(\sigma(t), \mathbf{v}(t))dt + \frac{1}{2}T|\mathbf{c}|^2 \tag{5.12}$$

for all $\sigma \in AC^2(0,\infty;\mathscr{P}(\mathbb{T}^d))$ *such that* $\sigma(0)=\mu$ *and all* $T>0$.

(ii) *There exists a path* $\sigma^* \in AC^2(0,\infty;\mathscr{P}(\mathbb{T}^d))$ *with* $\sigma^*(0)=\mu$ *such that equality is achieved in* (5.12) *for all* $T>0$.

PROOF. To prove (i) it suffices to approximate the path σ by paths consisting of averages of Dirac masses as in Proposition 5.1, then use (5.11).

For (ii) we take the optimal paths σ^n given by Proposition 5.4(ii), then use Proposition 3.21 to pass to the limit as $n\to\infty$. This yields a path σ satisfying the opposite (compared to (i)) inequality. We use (i) to finish the proof. □

The next statement brings up the fact that some $\mathbf{c}$-calibrated speed curves $(\sigma,\mathbf{v})$ from Proposition 5.5(ii) have the interesting property that the measures $\mu_t \in \mathscr{P}_2(\mathbb{T}^d\times\mathbb{R}^d)$ that disintegrate as $\mu_t(dx,dv)=\sigma_t(dx)\delta_{\mathbf{v}_t(x)}(v)$ yield distributional solutions for the fully kinetic Vlasov equation in $(0,\infty)\times\mathbb{T}^d\times\mathbb{R}^d$. These curves are obtained by a limiting procedure from the curves in (5.6).

THEOREM 5.6. *For any* $\mathbf{c}\in\mathbb{R}^d$ *and any* $\overline{\sigma}\in\mathscr{P}_2(\mathbb{R}^d)$ *there exists a* $\mathbf{c}$*-calibrated speed curve* $(\sigma,\mathbf{v})\in\mathscr{C}(0,\infty;\mathscr{C}\mathscr{P}(\mathbb{T}^d))$ *such that* $\sigma_0=\overline{\sigma}$, $\mathbf{v}$ *is a dual-role velocity for* σ, *and the curve* μ *defined by* $\mu_t(dx,dv)=\sigma_t(dx)\delta_{\mathbf{v}_t(x)}(dv)$ *solves the Vlasov equation* (5.2) *in* $\mathscr{D}'((0,\infty)\times\mathbb{T}^d\times\mathbb{R}^d)$.

PROOF. We assume that $\overline{\sigma}$ is supported by $[0,1)^d$. Approximate $\overline{\sigma}$ in the W_2-metric by $\sigma_0^n=\frac{1}{m}\sum_{i=1}^m \delta_{x_i^n(0)}$ where $\{x_i^n(0)\}_{i=1}^m\subset[0,1)^d$. Let σ^n be as in (5.8), $\mathbf{v}^n$ be as in (5.9), where $\{x_i^n\}_{i=1}^m$ is as in Proposition 5.3. Recall that Φ is the Vlasov flow and observe that by (5.7)

$$\mu_t^n=\Phi(t,\mu_0^n)\quad\text{where } \mu_t^n=\frac{1}{m}\sum_{j=1}^m \delta_{(x_i^n(t),\dot{x}_i^n(t))}. \tag{5.13}$$

In terms of μ^n, (5.6) can be written as

$$U(\sigma_0^n;\mathbf{c})=U(\sigma_T^n;\mathbf{c})+\int_0^T\left\{\int_{\mathbb{R}^d\times\mathbb{R}^d}\frac{1}{2}|v+\mathbf{c}|^2\mu_t^n(dx,dv)-\mathscr{F}(\sigma_t^n)\right\}dt. \tag{5.14}$$

Since $\mathscr{K}$ is bounded and the Hamiltonian $\frac{1}{2}\|\mathbf{v}_t^n\|_{\sigma_t^n}^2+\mathscr{K}(\sigma_t^n)$ is time independent, we exploit (5.14) and the fact that $U(\cdot\,;\mathbf{c})$ is bounded to conclude that $\{\mu_t^n\}_{t,n}$ is bounded in $\mathscr{P}_2(\mathbb{R}^d\times\mathbb{R}^d)$. Here we have used that the initial measure σ_0^n is supported by $[0,1)^d$ and used (5.7) to control the growth in time of the diameter of the support of σ_t^n. Extracting a subsequence if necessary, let us assume without loss of generality that $\{\mu_0^n\}_n$ converges narrowly to some μ_0. Since $\mathscr{F}$ satisfies (5.1), Φ is continuous in the sense that $\{\mu_t^n\}_n$ converges narrowly to $\mu_t=\Phi(t,\mu_0)$ for every t. Hence $\{\sigma_t^n\}_n$ converges narrowly to σ_t, the x-marginal of μ_t.

Letting n tend to ∞ in (5.14), we obtain

$$
\begin{aligned}
U(\sigma_0;\mathbf{c}) &\geq U(\sigma_T;\mathbf{c}) + \int_0^T \Bigg\{ \int\limits_{\mathbb{R}^d\times\mathbb{R}^d} \frac{1}{2}|v+\mathbf{c}|^2 \mu_t(dx,dv) - \mathscr{K}(\sigma_t)\Bigg\}\,dt \\
&\geq U(\sigma_T;\mathbf{c}) + \int_0^T \left\{ \frac{1}{2}\|\tilde{\mathbf{v}}_t + \mathbf{c}\|_{\sigma_t}^2 - \mathscr{K}(\sigma_t)\right\} dt \qquad (5.15)\\
&\geq U(\sigma_T;\mathbf{c}) + \int_0^T \left\{ \frac{1}{2}\|\mathbf{v}_t + \mathbf{c}\|_{\sigma_t}^2 - \mathscr{K}(\sigma_t)\right\} dt.
\end{aligned}
$$

Here $\tilde{\mathbf{v}}$ is obtained by disintegrating μ_t with respect to its x-marginal σ_t as follows:

$$
\int\limits_{\mathbb{R}^d\times\mathbb{R}^d} \varphi(x,v)\mu_t(dx,dv) = \int\limits_{\mathbb{R}^d} \sigma_t(dx) \int\limits_{\mathbb{R}^d} \varphi(x,v)\mu_t^x(dv),
$$

and then setting

$$
\tilde{\mathbf{v}}_t(x) = \int\limits_{\mathbb{R}^d} v\mu_t^x(dv).
$$

The periodic velocity $\mathbf{v}_t$ is defined as the projection of $\tilde{\mathbf{v}}_t$ onto $L^2(\mathbb{T}^d,\sigma_t)$. Since μ solves (5.2), it suffices to pick a test function $\xi \in C^1(\mathbb{T}^d)$ and compute

$$
\begin{aligned}
\frac{d}{dt}\int\limits_{\mathbb{R}^d} \xi\, d\sigma_t &= \frac{d}{dt} \int\limits_{\mathbb{R}^d\times\mathbb{R}^d} \xi(x)\mu_t(dx,dv) \\
&= \int\limits_{\mathbb{R}^d\times\mathbb{R}^d} v\cdot\nabla\xi(x)\mu_t(dx,dv) \\
&= \int\limits_{\mathbb{R}^d} \nabla\xi(x)\cdot\Bigg(\int\limits_{\mathbb{R}^d} v\mu_t^x(dv)\Bigg)\sigma_t(dx) = \int\limits_{\mathbb{R}^d} \nabla\xi\cdot\tilde{\mathbf{v}}_t\, d\sigma_t
\end{aligned}
$$

to see that $\mathbf{v}$ is a velocity for σ in the periodic sense. It follows that

$$
U(\sigma_0;\mathbf{c}) \leq U(\sigma_T;\mathbf{c}) + \int_0^T \left\{ \frac{1}{2}\|\mathbf{v}_t + \mathbf{c}\|_{\sigma_t}^2 - \mathscr{F}(\sigma_t)\right\} dt,
$$

which implies that all inequalities in (5.15) are, in fact, equalities. In particular,

$$
\tilde{\mathbf{v}}_t \equiv \mathbf{v}_t \quad \text{in } L^2(\sigma_t) \text{ for a.e. } t>0 \tag{5.16}
$$

and

$$
\int_0^T \int\limits_{\mathbb{R}^d\times\mathbb{R}^d} |v+\mathbf{c}|^2\mu_t(dx,dv)dt = \int_0^T \|\mathbf{v}_t+\mathbf{c}\|_{\sigma_t}^2 dt \quad \text{for all } T\geq 0. \tag{5.17}
$$

On the other hand, by Jensen's inequality,

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |v+\mathbf{c}|^2\mu_t(dx,dv) = \int_{\mathbb{R}^d} \sigma_t(dx)\int_{\mathbb{R}^d} |v+\mathbf{c}|^2\mu_t^x(dv)$$
$$\geq \int_{\mathbb{R}^d} \sigma_t(dx)\left|\int_{\mathbb{R}^d} (v+\mathbf{c})\mu_t^x(dv)\right|^2.$$

Hence

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |v+\mathbf{c}|^2\mu(dx,dv) \geq \|\tilde{\mathbf{v}}_t+\mathbf{c}\|_{\sigma_t}^2 \geq \|\mathbf{v}_t+\mathbf{c}\|_{\sigma_t}^2,$$

and equality holds if and only if $\mu_t^x = \delta_{\mathbf{v}_t(x)}$. Hence, according to (5.17), for almost every t we have

$$\mu_t^x = \delta_{\mathbf{v}_t(x)}.$$

By the optimality condition in (5.15) (recall that all inequalities are equalities), $(\sigma,\mathbf{v})$ satisfies the monokinetic system (since it minimizes the Lagrangian action over $[0,T]$ for all $T > 0$). By (5.16), $\mathbf{v}$ is a dual-role (both full and periodic) velocity field for σ. □

COROLLARY 5.7. *Let $\sigma_0 \in \mathscr{P}_2(\mathbb{R}^d)$ and $\mathbf{c}\in\mathbb{R}^d$ be given. There exists a curve $\sigma \in AC^2(0,\infty;\mathscr{P}_2(\mathbb{R}^d))$ originating at σ_0, along with a dual-role velocity $\mathbf{v}$ associated to it, such that*:

(i) *The measures $\mu_t(dx,dv) := \sigma_t(dx)\delta_{\mathbf{v}_t(x)}(v)$ solve the nonlinear Vlasov equation* (5.2).

(ii) *We have*

$$\left\|\frac{\mathrm{id}}{t}+\mathbf{c}\right\|_{\sigma_t} \leq \frac{C}{\sqrt{t}} + \left\|\frac{\mathrm{id}}{t}+\mathbf{c}\right\|_{\sigma_0} \quad \textit{for a.e. } t>0, \qquad \lim_{t\to\infty}\|\mathbf{v}_t+\mathbf{c}\|_{\sigma_t} = 0.$$

PROOF. The first part follows directly from Theorem 5.6. To prove (ii), in order to concentrate on the main ideas, let us only deal with the case $\mathbf{c}=\mathbf{0}$. As $\mathscr{F}\leq 0$, so is $\mathscr{K}$. Thus, the optimal curve from Theorem 5.5 satisfies

$$\int_0^\infty \|\mathbf{v}_t\|_{\sigma_t}^2\,dt < \infty. \tag{5.18}$$

The Hamiltonian energy is conserved along this curve, i.e.,

$$\frac{1}{2}\|\mathbf{v}_t\|_{\sigma_t}^2 + \mathscr{K}(\sigma_t) = \text{const} \quad \text{for all } t\geq 0.$$

Indeed, one can either prove this directly as in [11] or use the same property of the discrete approximations σ^n and then pass to the limit. Thus,

$$\beta(t) := \int_0^\infty \|\mathbf{v}_t\|_{\sigma_t}^2\,dt = \text{const} - 2\mathscr{K}(\sigma_t). \tag{5.19}$$

The AC^2-regularity of σ shows that β is absolutely continuous, and one can compute its a.e. derivative to get

$$\dot{\beta}(t) = -2 \int\limits_{\mathbb{R}^d \times \mathbb{R}^d} \nabla \mathscr{F}(x-y) \cdot \mathbf{v}_t(y) \sigma_t(dy) \sigma_t(dx),$$

which implies $|\dot{\beta}(t)| \le \|\nabla \mathscr{F}\|_\infty \sqrt{\beta(t)}$. But $\mathscr{F} \in C(\mathbb{T}^d)$ and (5.19) imply $\beta \in L^\infty(0,\infty)$, and the previous inequality and (5.18) lead to $\beta \in W^{1,\infty}(0,\infty) \cap L^1(0,\infty)$. It follows that

$$\lim_{t\to\infty} \beta(t) = 0.$$

Next, we use the fact that $\mathbf{v}$ is a full velocity for σ to write

$$\frac{d}{dt} \|\mathrm{id}\|_{\sigma_t}^2 = \int\limits_{\mathbb{R}^d} x \cdot \mathbf{v}_t(x) \sigma_t(dx) \le \|\mathrm{id}\|_{\sigma_t} \|\mathbf{v}_t\|_{\sigma_t} \quad \text{for a.e. } t > 0,$$

which implies

$$\|\mathrm{id}\|_{\sigma_t} - \|\mathrm{id}\|_{\sigma_0} \le \int_0^t \|\mathbf{v}_s\|_{\sigma_s} ds \le \sqrt{t} \left(\int_0^t \|\mathbf{v}_s\|_{\sigma_s}^2 \, ds \right)^{1/2},$$

which, in view of (5.18), finishes the proof. □

Acknowledgment. WG gratefully acknowledges the support provided by National Science Foundation grants DMS-0901070 and DMS-1160939. AT gratefully acknowledges the support provided by Simons Foundation Grant #246063.

Bibliography

[1] Ambrosio, L.; Gangbo, W. Hamiltonian ODEs in the Wasserstein space of probability measures. *Comm. Pure Appl. Math.* **61** (2008), no. 1, 18–53. doi:10.1002/cpa.20188

[2] Ambrosio, L.; Gigli, N.; Savaré, G. *Gradient flows in metric spaces and in the Wasserstein spaces of probability measures*. Lectures in Mathematics ETH Zurich. Birkhäuser, Basel, 2005.

[3] Brenier, Y.; Gangbo, W.; Savaré, G.; Westdickenberg, M. Sticky particle dynamics with interactions. *J. Math. Pures Appl. (9)* **99** (2013), no. 5, 577–617. doi:10.1016/j.matpur.2012.09.013

[4] Concordel, M. C. Periodic homogenization of Hamilton-Jacobi equations: additive eigenvalues and variational formula. *Indiana Univ. Math. J.* **45** (1996), no. 4, 1095–1117. doi:10.1512/iumj.1996.45.1145

[5] Cordero-Erausquin, D. Sur le transport de mesures périodiques. *C. R. Acad. Sci. Paris Sér. I Math* **329** (1999), no. 3, 199–202. doi:10.1016/S0764-4442(00)88593-6

[6] Evans, L. C. A survey of partial differential equations methods in weak KAM theory. *Comm. Pure Appl. Math.* **57** (2004), no. 4, 445–480. doi:10.1002/cpa.20009

[7] Fathi, A. Weak KAM theory in Lagrangian dynamics. Lecture notes, 2003.

[8] Feldman, M.; Tudorascu, A. On Lagrangian solutions for the semi-geostrophic system with singular data. *SIAM J. Math. Anal.* **45** (2013), no. 3, 1616–1640. doi:10.1137/120870116

[9] Gangbo, W.; Kim, H. K.; Pacini, T. Differential forms on Wasserstein space and infinite-dimensional Hamiltonian systems. *Mem. Amer. Math. Soc.* **211** (2011), no. 993, vi+77 pp. doi:10.1090/S0065-9266-2010-00610-0

[10] Gangbo, W.; McCann, R. J. The geometry of optimal transportation. *Acta Math.* **177** (1996), no. 2, 113–161. doi:10.1007/BF02392620

[11] Gangbo, W.; Nguyen, T.; Tudorascu, A. Hamilton-Jacobi equations in the Wasserstein space. *Methods Appl. Anal.* **15** (2008), no. 2, 155–183.

[12] Gangbo, W.; Nguyen, T.; Tudorascu, A. Euler-Poisson systems as action-minimizing paths in the Wasserstein space. *Arch. Ration. Mech. Anal.* **192** (2009), no. 3, 419–452. doi:10.1007/s00205-008-0148-y

[13] Gangbo W.; Tudorascu, A. Lagrangian dynamics on an infinite-dimensional torus; a weak KAM theorem. *Adv. Math.* **224** (2010), no. 1, 260–292. doi:10.1016/j.aim.2009.11.005

[14] Gangbo, W.; Tudorascu, A. A weak KAM theorem; from finite to infinite dimension. *Optimal transportation, geometry and functional inequalities*, 45–72. CRM Series, 11. Edizioni Della Normale, Pisa, 2010.

[15] Gangbo, W.; Tudorascu, A. Homogenization for a class of integral functionals in spaces of probability measures. *Adv. Math.* **230** (2012), no. 3, 1124–1173. doi:10.1016/j.aim.2012.03.005

[16] Gigli, N. Second order analysis on $(\mathscr{P}_2(M), W_2)$. *Mem. Amer. Math. Soc.* **216** (2012), no. 1018, xii+154 pp. doi:10.1090/S0065-9266-2011-00619-2

[17] Gomes, D. A. Viscosity solutions of Hamilton-Jacobi equations, and asymptotics for Hamiltonian systems. *Calc. Var. Partial Differential Equations* **14** (2002), no. 3, 345–357. doi:10.1007/s005260100106

[18] Kim, H. K. Hamiltonian systems and the calculus of differential forms on the Wasserstein space. Doctoral dissertation. Georgia Institute of Technology, Atlanta, 2008.

[19] Villani, C. *Topics in optimal transportation*. Graduate Studies in Mathematics, 58. American Mathematical Society, Providence, R.I., 2003. doi:10.1007/b12016

W. GANGBO
Georgia Institute of Technology
Atlanta, GA 30332-0160
USA
E-mail: `gangbo@math.gatech.edu`

A. TUDORASCU
West Virginia University
Morgantown, WV 26506-6310
USA
E-mail: `adriant@math.wvu.edu`

Received January 2012.