An Optimal Transport Approach to Nonlinear Evolution Equations

by

Ehsan Kamalinejad

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Mathematics
University of Toronto

Abstract

An Optimal Transport Approach to Nonlinear Evolution Equations

Ehsan Kamalinejad

Doctor of Philosophy

Graduate Department of Mathematics

University of Toronto

2012

Gradient flows of energy functionals on the space of probability measures with Wasserstein metric has proved to be a strong tool in studying certain mass conserving evolution equations. Such gradient flows provide an alternate formulation for the solutions of the corresponding evolution equations. An important condition, which is known to guarantees existence, uniqueness, and continuous dependence on initial data is that the corresponding energy functional be displacement convex. We introduce a relaxed notion of displacement convexity and we show that it still guarantees short time existence and uniqueness of Wasserstein gradient flows for higher order energy functionals which are not displacement convex in the standard sense. This extends the applicability of the gradient flow approach to larger family of energies. As an application, local and global well-posedness of different higher order non-linear evolution equations are derived. Examples include the thin-film equation and the quantum drift diffusion equation in one spatial variable.

Keywords: Optimal Transport; Wasserstein Gradient Flows; Displacement Convexity; Minimizing Movement; Well-posedness; Non-linear Evolution Equations.

Acknowledgements

I would like to thank my supervisor Almut Burchard for her guidance and support throughout my PhD studies. Her generosity with her energy and time will not be forgotten. I am truly indebted to Robert McCann, Robert Jerrard, Larry Guth, Dejan Slepčev, Catherine Sulem, and Marina Chugunova whose help has been invaluable in my studies. I am grateful to James Colliander, Mary Pugh, Dror Bar-Natan, and Nicola Gigli for all the insightful discussions.

The delightful people at Department of Mathematics have provided me such enjoyable times. I truly appreciate it. In particular, I would like to thank Ida Bulat for providing an infinite source of help at all times. I am very happy to find such lovely friends at University of Toronto. Thank you all.

Last but not least, I wish to express my deep gratitude to the people who supported me with their unconditional love throughout my entire life: my parents Parvaneh and Mohammad, and my brother Ali, who taught me the beauty of Mathematics. And finally I want to thank my angel, Pardis, who gave a meaning to all my efforts and supported me through ups and downs with great patience and love.

Contents

Notations

$T_{\#}\mu$	Push forward of the measure μ via the map T
$C_c^\infty(\mathbb{R}^m)$	Space of compactly supported smooth functions on $\mathbb{R}^m$
$\mathcal{P}(\mathbb{R}^m)$	Space of probability measure on $\mathbb{R}^m$
$\mathcal{P}_2(\mathbb{R}^m)$	Space of probability measure on $\mathbb{R}^m$ with finite second moment
$\mathcal{L}^m$	Lebesgue measure on $\mathbb{R}^m$
$\mathcal{P}_2^a(\mathbb{R}^m)$	Subspace of $\mathcal{P}_2(\mathbb{R}^m)$ containing a.c. measures w.r.t. $\mathcal{L}^m$
Π_i	Projection to the i^{th} coordinate
$L^p(d\mu)$	Space functions with finite p-integration against μ
$\Gamma(\mu,\nu)$	Space of probability measures with marginals μ and ν
$W_2(\mu,\nu)$	Quadratic optimal transport distance between measures μ and ν
$\mathcal{P}(X)$	Probability measures in a separable metric space X
$AC^2([0,t];X)$	Space of absolutely continuous curves with $L^2([0,t])$ integrable derivative
$\partial E(\mu)$	subdifferential of E at μ
$T_\mu(\mathcal{P}_2(\mathbb{R}^m))$	The tangent space of $\mathcal{P}_2(\mathbb{R}^m)$ at μ
$D(E)$	The domain of E, i.e. $\{x \in X \mid E(x) < \infty\}$
E_c	The energy sub level-set given by $\{\mu \in D(E) \mid E(\mu) < c\}$

Chapter 1

Introduction

In the last decade, the theory of Optimal Transport has been a rapidly expanding area of research. With a wide range of applications to differential geometry, mathematical finance, gradient flow solutions of evolution equations and functional inequalities, this theory has received an extensive amount of interest. One of the active applications of optimal transport theory is the application of the quadratic optimal transport in studying evolution equations.

The quadratic optimal transport space $\mathcal{P}_2(\mathbb{R}^m)$, also known as the Wasserstein space, consists of the Borel probability measures on $\mathbb{R}^m$ with finite second moment. **Wasserstein distance** W_2, defines a distance function between pair of measures $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^m)$ given by

$$W_2(\mu,\nu) := \inf \left\{ \int_{\mathbb{R}^m \times \mathbb{R}^m} |x-y|^2 \, d\gamma \; : \; \gamma \in \Gamma(\mu,\nu) \right\}^{\frac{1}{2}}$$

where $\Gamma(\mu,\nu) \subset \mathcal{P}_2(\mathbb{R}^m \times \mathbb{R}^m)$ is the space of probability measures with marginals μ and ν. We will refer to such measures γ as **transport plans**.

It turns out that $\mathcal{P}_2(\mathbb{R}^m)$ has a rich geometric structure and a formal Riemannian calculus can be performed on this space. The first appearance of the Riemannian calculus on $\mathcal{P}_2(\mathbb{R}^m)$ is due to Otto et al in [18] and [25]. It was shown in [25] that the solution of the porous medium equation

$$\partial_t u = \Delta u^m,$$

with $m > 0$ can be reformulated as the Wasserstein gradient flow of the energy

$$E(u) = \int \frac{u^m}{m-1} .$$

Since then, the interaction between the Riemannian space $\mathcal{P}_2(\mathbb{R}^m)$ as a geometric object and evolution equations as analytic objects have attracted a lot of attention. This point of view is commonly called *"Otto calculus"*.

A notion which has been very important in the development of this theory is the notion of *displacement convexity.* McCann in his thesis [23] introduced the notion of displacement convexity of an energy functional on the Wasserstein space. Under the displacement convexity assumption, he proved existence and uniqueness of minimizers of wide classes of energies, commonly referred to as potential, internal, and interactive energies. Displacement convexity had been defined before the development of the Wasserstein gradient flows, but after establishment of the Riemannian structure of the Wasserstein space, it turned out that displacement convexity can be interpreted as the standard convexity along the geodesics of the Wasserstein space. The displacement convexity condition, with its generalization to λ-displacement convexity, has a central role in existence, uniqueness, and long-time behaviour of the gradient flow of an energy functional.

Another important idea in the theory of Wasserstein gradient flows is the notion of *minimizing movement*. The minimizing movement scheme was suggested by De Georgi as a variational approximation of gradient flows in general metric spaces [14]. It was later used by Jordan, Kinderlehrer, Otto [18] and by Ambrosio, Gigli, Savaré [1] to construct a systematic theory of Wasserstein gradient flows. This theory was soon used by many researchers to develop existence, uniqueness, stability, long-time behaviour, and numerical approximation of evolution PDEs such as in [2], [3], [6], [10], [11], [13], [15], and [24].

In recent years, it has become apparent that Otto Calculus [25] also applies to higher-order evolution equations, at least at a formal level. The best-studied example is the thin-film equation

$$\partial_t u = -\nabla \cdot (u \nabla \Delta u),$$

which corresponds to the gradient flow of the Dirichlet energy

$$E(u) = \frac{1}{2} \int |\nabla u|^2 \, dx.$$

The hope is that gradient flow methods might help to resolve long-standing problems concerning well-posedness and long-time behaviour of this PDE. However, taking advantage of the gradient flow method has proved difficult. The main obstruction has been the lack of displacement convexity of the Dirichlet energy. The same problem arises for studying other energy functionals containing derivatives of the density. In [26, open problem 5.17] Villani raised the question whether there is any interesting example of a displacement convex functional that contains derivatives of the density. In [12], Carrillo and Slepčev answered this question by providing a class of displacement convex functionals. Therefore it was proved that there is no fundamental obstruction for existence of such energies. However because of the lack of displacement convexity, the Wasserstein gradient flow method has not been very successful in studying gradient flows of the Dirichlet energy and other interesting energies of higher order.

Our result can be summarized as follows:

- We introduce a relaxed notion of λ-displacement convexity of an energy functional and in Theorem 4.2.2 we prove that, under this relaxed assumption, the general theory of well-posedness of Wasserstein gradient flows holds at least locally.
- In Theorem 5.2.6, we prove that, in spatial dimension, the Dirichlet energy, which is not λ-displacement convex in the standard sense, satisfies the relaxed version of λ-displacement convexity on positive measures. Hence the gradient flow of the Dirichlet energy is locally well-posed and the solution of the thin-film equation with positive initial data exists and is unique as long as positivist is preserved.
- We show that the method developed to study thin-film equation directly applies to a range of PDEs of fourth and higher order.

The thesis is organized as follows:

In Chapter 2 we will gather some basics of the theory of Optimal Transport. The space of probability measures with finite second moments and the structure of optimal transport map is described here. We will focus in particular on quadratic optimal transportation which is the underlying space for our gradient flows.

In chapter 3 the general formulation of Wasserstein gradient flows is presented. We also study the concept of displacement convexity and related displacement convex energy functionals.

We introduce a generalized version of λ-displacement convexity which we call *relaxed λ-convexity* in Chapter 4. Setting technicalities aside, the idea of relaxed λ-convexity can be summarized in two simple principles: Firstly, the modulus of convexity, λ, can vary along the flow. Secondly, one can study λ-convexity *locally on sub-level sets of the energy.* Note that both locallity and energy dissipation have important roles in our analysis. For example, the Dirichlet energy is not even locally λ-convex, because an arbitrarily small neighbourhood of a smooth positive measure contains measures with infinite energy where λ-convexity fails altogether. Instead, by taking advantage of the defining properties of the gradient flow, we study the flow on energy sub-level sets. The key observation is that typically finiteness of the energy implies some regularity on the measure which helps to elevate the formal calculations to a rigorous proof. For example, in one spacial dimension, densities of finite Dirichlet energy are continuous. After introducing relaxed λ-convexity, we state our first result, Theorem 4.2.2. In this theorem we prove that if an energy functional is relaxed λ-convex at a point μ, then the corresponding gradient flow trajectory starting from μ exists and is unique at least for a short time. The proof is based on convergence of the minimizing movement scheme and the subdiffrential properties that are carried over to the limiting curve. The existing proof does not apply directly to a relaxed λ-convex energy because the restricted domain contains only energy dissipating directions and hence having convexity on this restricted set of directions does not imply convexity on all directions even locally. However relaxed λ-convexity is very compatible with the defining properties of the minimizing movement scheme. In particular both of the constraints *locality* and *energy boundedness* are used in the definition of the minimizing movement scheme (3.10). It is interesting to note that since minimizing movement was initially defined to be applied to gradient flows on general metric spaces, Theorem 4.2.2 hence has a wider range of applicability and with minor modifications it can be applied to gradient flows on different metric spaces given that the energy is relaxed λ-convex.

In Chapter 5 we apply the theory developed in the previous chapter to the Dirichlet energy. We prove that the Dirichlet energy on $\mathbb{R}/\mathbb{Z}$, is relaxed λ-convex on the measures with positive density. This theorem re-derives the existing theory [5] of well-posedness of the solutions of the thin-film equation away from degenerate points by a direct geometric proof. To the best of our knowledge, this is the first well-posedness result for the thin-film equation based on Wasserstein gradient flows. Two key ideas are very useful in the proofs of this section: Firstly, the Wasserstein convergence and the uniform convergence are equivalent on energy sub-level sets. Secondly, finiteness of the energy can be used directly in the calculations of the second derivative of the energy along geodesics.

In the final chapter, we show that the method developed in Chapters 4 and 5 can be applied to a wide class of energies of fourth and higher orders. Some important examples have been studied using this method such as equations of higher order of the form $\partial_t u = (-1)^k \partial_x (u \partial_x^{2k+1} u)$, and equations of different forms, for instance the quantum drift diffusion equation $\partial_t u = -\partial_x (u \partial_x \frac{\partial_x^2 \sqrt{u}}{\sqrt{u}})$.

The Wasserstein gradient flow approach to PDEs has some interesting features. For example, it has a unified notion of solution which allows for very weak solutions and it is applicable to equations of higher order even with the lack of maximum principle. Also the minimizing movement scheme is a constructive method. Hence the proofs are constructive and one can derive numerical approximations based on the Wasserstein gradient flows similar to what has been done in [19], [13], and [15].

Chapter 2

The Optimal Transport Problem

In this chapter we start by a simple example motivating the idea of optimal transport and later we state the rigorous definitions and fundamental theorems of the theory which we will use in the next chapters.

2.1 Description of the optimal transport problem

The problem of optimal transport was formalized by the French mathematician Gaspard Monge in 1781. During the world war II the Russian mathematician and economist Leonid Kantorovich made important advancements in the theory. Informally, the problem can be posed as follows. Assume there is a pile of sand that should be transported into a hole to make the surface of the ground even. The question is how to transport the sand with a minimal effort (see figure (2.1)). We can model the pile of sand and the hole by two measures μ and ν on the intervals X and Y. A mapping T which sends every small parcel of the sand to a corresponding location $T(x)$ on the hole represents a particular transportation. Of course, we need the mass of each parcel to be equal to the mass of its

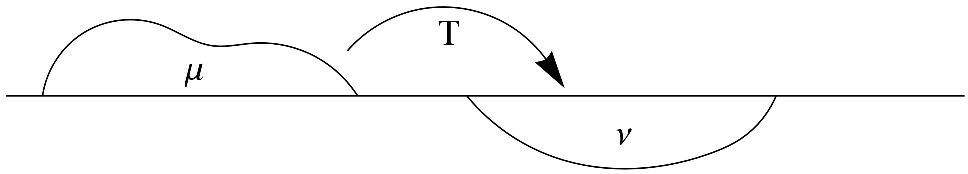

Figure 2.1: Optimal Transport

destination. In other words, we need the *push forward* of μ via T to be equal to ν which is defined by

$$T_{\#}\mu = \nu \quad \Longleftrightarrow \quad \nu(B) = \mu(T^{-1}(B)) \quad \text{for all measurable set } B \subset Y. \tag{2.1}$$

For the displacement of each parcel some energy is needed which is called the *transportation cost*. It is natural to assume that the cost depends on the initial location of the parcel and its destination i.e. cost $= c(x, T(x))$. One natural example of a cost function is the distance of x from $T(x)$. The total cost needed to transport the pile of sand via a particular mapping T is obtain by summing up the cost over all the small parcels of sand

$$I(T) = \int_X c(x, T(x))d\mu.$$

Note that the effect of ν is hidden in this equation via (2.1). The optimal transport problem is to find a map T which gives the minimal cost of transportation constrained to (2.1). This problem is called Monge's optimal transportation problem.

Usually it is too restrictive to assume that each parcel of mass is transported into exactly one location. For example, in the case of sand/hole problem each parcel of sand might be split into smaller parcels which are then transported to different locations. To allow such possibility, one has to consider measures on the product space $X \times Y$. Informally stated, such a measure $\pi(x, y)$ indicates the amount of sand which is transported from location x on the pile to the location y in the hole. We still need the mass at location x to be equal to the total mass transported from x to the hole. This means that π should have marginals μ and ν, i.e.

$$\pi(A \times Y) = \mu(A) \quad \text{and} \quad \pi(X \times B) = \nu(B),$$

for all measurable sets $A \subset X$ and $B \subset Y$. We call such measure on $X \times Y$ a *transportation plan* between μ and ν. The set of all plans between μ and ν is denoted by $\Pi(\mu, \nu)$. The total cost of the transportation $\pi \in \Pi(\mu, \nu)$ is given by

$$I(\pi) = \int_{X \times Y} c(x, y) d\pi.$$

The question now is to find the transportation plan with minimal cost.

The problem in a more general setting when X and Y are measurable spaces equipped with measures μ and ν is called *Kantorovich's optimal transportation* problem. Kantorovich was awarded economics Nobel prize for his work on a related problem in minimization.

In general setting, for an arbitrary cost function, existence of a minimizer is not guaranteed. In the case of the quadratic cost $c(x, y) = |x - y|^2$, and absolutely continuous measures μ and ν, existence of a unique minimizer and characteristic of the optimizers is known. The theorem is due to Brenier [8] with its generalization by McCann [22].

2.2 Definitions and theorems

Let us recall some deinitions and notation. For the moment, we will work on $\mathbb{R}^m$, but we will see later that most of the statements remain valid when $\mathbb{R}^m$ is replaced by a smooth manifold.

The space of probability measures with finite second moment on $\mathbb{R}^m$ is given by

$$\mathcal{P}_2(\mathbb{R}^m) := \left\{ \mu \in \mathcal{B}(\mathbb{R}^m) \;\middle|\; \int_{\mathbb{R}^m} d\mu(x) = 1 \ , \ \int_{\mathbb{R}^m} |x|^2 d\mu(x) < \infty \right\},$$

where $\mathcal{B}(\mathbb{R}^m)$ is the space of Borel measures on $\mathbb{R}^m$. The class of absolutely continuous probability measures $\mathcal{P}_2^a(\mathbb{R}^m)$ are particularly important. Here absolute continuity of a measure refers to absolute continuity with respect to Lebesgue measure, i.e.

$$\mathcal{P}_2^a(\mathbb{R}^m) := \{\mu \in \mathcal{P}_2(\mathbb{R}^m) \;\big|\; \mu = u dx \ , \ u \in L^1(\mathbb{R}^m)\}.$$

For two measures μ, ν in $\mathcal{P}_2(\mathbb{R}^m)$ the space $\Gamma(\mu, \nu)$ is defined to be the space of all joint probability measures on $\mathbb{R}^m \times \mathbb{R}^m$ whose marginals are μ and ν. This means that if $\gamma \in \Gamma(\mu, \nu)$, then for any two bounded measurable functions f and g we have

$$\int_{\mathbb{R}^m \times \mathbb{R}^m} f(x) d\gamma(x, y) = \int_{\mathbb{R}^m} f(x) d\mu(x) \ \text{ and } \ \int_{\mathbb{R}^m \times \mathbb{R}^m} g(y) d\gamma(x, y) = \int_{\mathbb{R}^m} g(y) d\nu(y).$$

The Wasserstein distance of μ and ν, based on Kantorovich's problem, is given by

$$W_2(\mu, \nu) := \inf_\gamma \left\{ \int_{\mathbb{R}^m \times \mathbb{R}^m} |x - y|^2 \, d\gamma \ : \ \gamma \in \Gamma(\mu, \nu) \right\}^{\frac{1}{2}} . \tag{2.2}$$

Brenier's theorem [8] asserts that the minimum is always assumed and the minimal transport plan is concentrated on a graph of a map T_μ^ν, provided that $\mu \in \mathcal{P}_2^a(\mathbb{R}^m)$. This theorem was generalized to smooth manifolds by McCann [22].

Theorem 2.2.1 (Brenier's theorem) *Let $\mu,\ \nu \in \mathcal{P}_2(\mathbb{R}^m)$. Assume that μ is absolutely continuous with respect to Lebesgue measure. Then*

(i) There is a unique minimal plan $\pi \in \Pi(\mu,\nu)$ to the Kantorovich's problem

$$\inf_{\gamma} \left\{ \int_{\mathbb{R}^m \times \mathbb{R}^m} |x-y|^2 \, d\gamma \ : \ \gamma \in \Gamma(\mu,\nu) \right\}^{\frac{1}{2}}.$$

(ii) There exists a convex lower semicontinuous function $\phi : \mathbb{R}^m \to \mathbb{R}$ such that

$$\pi = (Id \times \nabla\phi)_{\#}\mu.$$

(iii) Monge's problem has a unique solution which is given by $T = \nabla\phi$, i.e.

$$\int_{\mathbb{R}^m} |x - \nabla\phi(x)|^2 = \inf_{T} \left\{ \int_{\mathbb{R}^m} |x - T(x)|^2 \, d\mu \ : \ T_{\#}\mu = \nu \right\}.$$

This optimal transport map T is called *Brenier's map.* We usually use T_μ^ν to refer to the optimal transport map between measure μ and ν.

In [23] McCann proved the following theorem which describes Brenier's map explicitly.

Theorem 2.2.2 *Let $\mu,\ \nu \in \mathcal{P}_2^a(\mathbb{R}^m)$ with densities $\mu = udx$ and $\nu = vdx$. Assume that T is the optimal map between μ and ν. Then for almost all $x \in \mathbb{R}^m$ we have*

$$det(DT)v(T(x)) = u(x), \tag{2.3}$$

where DT is the derivative matrix of T.

Equation (2.3) which is the Euler-Lagrange equation for the Monge problem is called *Monge-Ampère* equation.

The following theorem due to McCann [23] describes a time dependent transportation which is given by the push-forward of the linear interpolation between the optimal map $T_{\mu_0}^{\mu_1}$ and the identity map.

Theorem 2.2.3 *Let* $\mu_0, \ \mu_1 \in \mathcal{P}_2^a(\mathbb{R}^m)$. *Assume that* T *is the optimal transport map between* μ_0 *and* μ_1. *Then for any* $s \in [0,1]$ *the map*

$$T_s := (1-s)Id + sT$$

is the optimal map between μ *and*

$$\mu_s = \left((1-s)Id + sT_{\mu_0}^{\mu_1}\right)_{\#} \mu_0. \tag{2.4}$$

Also for the Wasserstein distance of the intermediate points we have

$$W_2(\mu_0, \mu_s) = sW(\mu_0, \mu_1) \quad \forall s \in [0,1].$$

Chapter 3

Wasserstein Gradient Flows

In this chapter, we gather basics of the theory of Wasserstein gradient flows. In the first section we recall the formulation of gradient flows on a Hilbert space. The formulation of the flow on a Hilbert space is very concrete and we will use this transparent structure as a recipe for corresponding notions in the more abstract setting of Wasserstein space. The second section will study some aspects of the Rienmannian structure of the quadratic optimal transport as the underlying space of the flows that we are interested in. Finally, in section three, we study the structure of the Wasserstein gradient flows.

3.1 Gradient Flows on Hilbert Spaces

Informally stated, a gradient flow evolves by the steepest descent of an energy functional. This idea can be formalized in several different ways, some of which carry over to general metric spaces. Let us first recall the notion of a gradient flow on a finite dimensional Riemannian manifold. The ingredients of a gradient flow consist of three parts: a smooth manifold M, a metric g, and an energy E. The gradient flow of the energy E starting from a point x_0 can be formulated as

$$\begin{cases} \lim_{t\downarrow 0} x_t = x_0 & \text{(initial condition)} \\ \partial_t x_t = V_t & \text{(velocity vector)} \\ V_t = -\nabla E(x_t) & \text{(steepest descent).} \end{cases}$$

Note that the role of the metric is to convert the co-vector $dE = \frac{\partial E}{\partial x_j} dx_j$ into the corresponding vector ∇E on the tangent space $T_{x_t}(M)$. In the early 1970's, several groups of mathematicians studied the notion of gradient flows on infinite dimensional Hilbert spaces (see the survey [9]). These articles consider the contraction semi-groups generated by a proper, convex, and lower semicontinuous energy functional E on a Hilbert space. One of the notions that was modified for the infinite dimensional setting was the gradient operator. It turned out that the notion of Fréchet subdifferential (Definition (3.6)) is an appropriate substitution for the gradient on a Hilbert space. Therefore the steepest descent equation $V_t = -\nabla E(x_t)$ on a Hilbert space is reformulated by subdifferential inclusion:

$$V_t \in \partial E(x_t).$$

Later, DeGiorgi and his collaborators studied the gradient flows on general metric spaces through the variational method called *minimizing movement* [14]. To illustrate this method we first consider it on a Hilbert space. The minimizing movement scheme is a time discretization of the steepest descent equation. Assume an energy functional E is given on Hilbert space H. For a fixed time step size τ we consider the following iterative scheme:

$$x_n^\tau := Argmin_{x\in H}\{E(x) + \frac{1}{2\tau}|x - x_{n-1}^\tau|^2\}. \tag{3.1}$$

By taking the time derivative of (3.1) at a formal level, we see that the scheme is approximating the steepest descent equation:

$$\frac{x_n^\tau - x_{n-1}^\tau}{\tau} \in -\partial E(x_n^\tau).$$

It is worth noting that this equation is the reformulation of implicit Euler method in the abstract setting.

Mimicking the structure of gradient flows in setting of a Hilbert space one obtains appropriate formulation for Wasserstein gradient flows. The ingredients of the Wasserstein gradient flows are given by: $\mathcal{P}_2(\mathbb{R}^m)$ as the manifold, the Wasserstein distance (and its infinitesimal version) as the metric, and an energy functional on $\mathcal{P}_2(\mathbb{R}^m)$. More detailed discussion is provided at section 3.3 where we consider the Fréchet subdifferential formulation of Wasserstein gradient flows. Later in Chapter 4 we identify conditions on the energy functional that guarantee short-time existence and uniqueness. The proof is based on a careful analysis of the minimizing movement scheme.

3.2 Geometry of the quadratic optimal transport

In this section we gather some basic elements of the Riemannian calculus on the Wasserstein space $\mathcal{P}_2(\mathbb{R}^m)$. We refer the reader to [1] or [26] for comprehensive discussion of the subject. It is worth mentioning that we consider the Euclidean space $\mathbb{R}^m$ as the underlying space of probability measures but one can replace $\mathbb{R}^m$ with any Hilbert space by slight modifications to the definitions as it is done in [1].

A class of curves in $\mathcal{P}_2(\mathbb{R}^m)$ that supports a natural notion of tangent vectors is given by the absolutely continuous curves with square integrable derivatives $AC^2_{loc}([0,\infty);\mathcal{P}_2(\mathbb{R}^m))$.

A curve μ_t belongs to $AC^2_{loc}([0,\infty);\mathcal{P}_2(\mathbb{R}^m))$ if there exist a locally $L^2(dt)$ integrable function g such that

$$W_2(\mu_a,\mu_b) \leqslant \int_a^b g(t)dt \quad \forall a,b \in [0,\infty).$$

The absolutely continuous curves are given by mass conserving deformations of the measures i.e. they satisfy the *continuity equation*:

$$\partial_t\mu_t + \nabla.(\mu_t V_t) = 0$$

for a *velocity vector field* V_t of deformations of μ_t. This equation is assumed to hold in the distributional sense i.e. for every test function $\psi \in C_c^\infty(\mathbb{R}^m \times (0,s))$ we have

$$\int_0^s \int_{\mathbb{R}^m} \left(\partial_t\psi + \langle V_t, \nabla\psi\rangle\right) d\mu_t dt = 0.$$

The following theorem due to Brenier and Benamou ([4]) shows that $\mathcal{P}_2(\mathbb{R}^m)$ is a *length space* in the sense that the distance of two measures is given by the length of the shortest path between them.

Theorem 3.2.1 *Let $\mu,\ \nu \in \mathcal{P}_2(\mathbb{R}^m)$. We have*

$$W_2(\mu,\nu)^2 = min\left\{\int_0^1\int_{\mathbb{R}^m} |V_t|^2 d\mu_t dt \;\middle|\; \partial_t\mu_t + \nabla.(\mu_t V_t) = 0\right\}, \tag{3.2}$$

where the minimum is taken over all absolutely continuous curves μ_t and velocity vector fields V_t which connect $\mu_0 = \mu$ to $\mu_1 = \nu$.

For a given curve $\mu_t \in AC^2([0,1]);\mathcal{P}_2^a(\mathbb{R}^m))$ there might be many velocity vectors that satisfy the same continuity equation $\partial_t\mu_t + \nabla.(\mu_t V_t) = 0$. For instance vector fields of the form $F_t + V_t$ where $\nabla.(\mu_t F_t) = 0$ all satisfy the same continuity equation. However there is a unique vector field that minimizes (3.2), i.e. the one which defines the distance between μ and ν. This optimal velocity vector field is defined to be the *tangent vector field* to the curve μ_t. This way a tangent vector to a curve can be defined and then one can define the tangent space $Tan_\mu \mathcal{P}_2(\mathbb{R}^m)$ using the tangents to the curves that pass through μ. This norm minimality condition characterizes the tangent space to $\mathcal{P}_2(\mathbb{R}^m)$ by

$$Tan_\mu \mathcal{P}_2(\mathbb{R}^m) = \overline{\{\nabla\phi \;\Big|\; \phi \in C_c^\infty(\mathbb{R}^m)\}}^{L^2(d\mu)}.$$

The tangent vector field of μ_t can also be expressed in term of the optimal maps. If V_t is the tangent vector field of μ_t then

$$V_t = \lim_{\epsilon\to 0} \frac{T_{\mu_t}^{\mu_{t+\epsilon}} - Id}{\epsilon}.$$

The converse is also true, i.e. for a given optimal map T_μ^ν, the vector field $T_\mu^\nu - Id$ is a tangent vector at μ for some curve that passes μ. The tangent vector fields are also useful in calculating the derivative of the Wasserstein metric along curves. Let $\mu_t \in AC^2_{loc}(\mathbb{R}^+; \mathcal{P}_2^a(\mathbb{R}^m))$. By [1, Chapter 8] *the derivative of the Wasserstein metric* along the curve μ_t is given by

$$\frac{d}{dt} W_2(\mu_t, \nu)^2 = 2\int_{\mathbb{R}^m} \langle V_t, Id - T_{\mu_t}^\nu \rangle d\mu_t \quad \forall \nu \in \mathcal{P}_2(\mathbb{R}^m) \tag{3.3}$$

where V_t is the tangent vector field to μ_t and $\langle .,. \rangle$ is the standard inner product on $\mathbb{R}^m$.

The Wasserstein metric is closely related to a certain weak topology on $\mathcal{P}_2(\mathbb{R}^m)$, induced by *narrow convergence*:

$$\mu_n \xrightarrow{\text{narrow}} \mu \iff \int_{\mathbb{R}^m} f d\mu_n \to \int_{\mathbb{R}^m} f d\mu \quad \forall f \in C_b^0(\mathbb{R}^m), \tag{3.4}$$

where $C_b^0(\mathbb{R}^m)$ is the set of of continuous bounded real functions on $\mathbb{R}^m$. The topologies induced by the narrow convergence and the Wasserstein distance are equivalent for sequences of measures with uniformly bounded second moments:

$$\lim_{n\to\infty} W_2(\mu_n, \mu) = 0 \quad \Longleftrightarrow \quad \begin{cases} \mu_n \xrightarrow{\text{narrow}} \mu \\ \{\mu_n\} \ \text{ has uniformly bounded 2-moments.} \end{cases} \tag{3.5}$$

3.3 Wasserstein Gradient Flows

Now we describe gradient flows on the Wasserstein space. Consider an energy functional $E : \mathcal{P}_2(\mathbb{R}^m) \to [0, \infty]$ and let its domain, $D(E)$, be the set where E is finite. Let μ lie in $D(E) \cap \mathcal{P}_2^a(\mathbb{R}^m)$. A vector field $\xi \in L^2(d\mu)$ belongs to the **subdifferential** of E at μ, denoted by $\partial E(\mu)$ if

$$\liminf_{\substack{\nu\to\mu \\ \nu\in D(E)}} \frac{E(\nu) - E(\mu) - \int_X \left\langle \xi, T_\mu^\nu - Id \right\rangle \mathrm{d}\mu}{W_2(\mu, \nu)} \geqslant 0. \tag{3.6}$$

We say that a curve $\mu_t \in AC^2_{loc}(\mathbb{R}^+, \mathcal{P}_2^a(\mathbb{R}^m))$ is a trajectory of the **gradient flow** for the energy E, if there exists a velocity field V_t with $|V_t|_{L^2(d\mu_t)} \in L^1_{loc}(\mathbb{R}^+)$ such that

$$\begin{cases} \partial_t \mu_t + \nabla \cdot (\mu_t V_t) = 0 & \text{(continuity equation),} \\ V_t \in -\partial E(\mu_t) & \text{(steepest descent)} \end{cases} \tag{3.7}$$

holds for almost every $t > 0$. We will refer to (3.7) as the **gradient flow equation**. The continuity equation links the curve with its velocity vector field and ensures that the mass is conserved. The steepest descent equation expresses that the gradient flow evolves in the direction of maximal energy dissipation.

Next we describe the link between Wasserstein gradient flows and evolution PDEs. Here, we assume that all of the measures are absolutely with respect to Lebesgue measure and we work with their density functions. More detailed analysis with a discussion of general measures can be found in [1]. Let $\mu = udx$ be in $D(E)$ and let $V \in \partial E(\mu)$ be a tangent vector field at μ. Consider a linear perturbation of μ given by the curve $\mu_\epsilon := (Id + \epsilon W)_{\#}\mu$ for small values of $\epsilon > 0$ where $W \in C_c^\infty(\mathbb{R}^m; \mathbb{R}^m)$. By the subdifferential inequality (3.6) we have

$$\limsup_{\epsilon \uparrow 0} \frac{E(u_\epsilon) - E(u)}{\epsilon} \leqslant \int_{\mathbb{R}^m} \langle V, W \rangle \, udx \leqslant \liminf_{\epsilon \downarrow 0} \frac{E(u_\epsilon) - E(u)}{\epsilon}.$$

On the other hand, assuming C^2 regularity on u_ϵ and using standard first order variations we have

$$\lim_{\epsilon \to 0} \frac{E(u_\epsilon) - E(u)}{\epsilon} = \int_{\mathbb{R}^m} \left(\frac{\delta E(u)}{\delta u} \right) \left(\frac{\partial u_\epsilon}{\partial \epsilon} \right) udx$$

where $\frac{\delta E(u)}{\delta u}$ stands for standard first variations of E. Therefore

$$\int_{\mathbb{R}^m} \langle V, W \rangle \, udx = \int_{\mathbb{R}^m} \frac{\delta E(u)}{\delta u} \partial_\epsilon u_\epsilon dx.$$

The continuity equation for the curve u_ϵ implies that $\partial_\epsilon u_\epsilon = -\nabla.(uW)$. Hence

$$\int_{\mathbb{R}^m} \langle V, W \rangle \, udx = -\int_{\mathbb{R}^m} \{ \frac{\delta E(u)}{\delta u} \nabla \cdot (uW) \} dx.$$

Integrating by parts, we have

$$\int_{\mathbb{R}^m} \langle V, W \rangle udx = \int_{\mathbb{R}^m} \langle \nabla (\frac{\delta E(u)}{\delta u}), W \rangle udx.$$

Since W is arbitrary we have

$$V(x) = \nabla \frac{\delta E(u)}{\delta u}(x) \quad \text{for } \mu\text{-a.e. } x. \tag{3.8}$$

Now assume that a curve $\mu_t = u_t dx$ satisfies the gradient flow equation (3.7). The steepest descent equation and (3.8) imply

$$V_t(x) = -\nabla \frac{\delta E(u)}{\delta u}(x).$$

By plugging V_t into the continuity equation, we have

$$\partial_t u = \nabla \cdot \left(u \nabla \frac{\delta E(u)}{\delta u} \right). \tag{3.9}$$

This is the corresponding PDE for the Wasserstein gradient flow of the energy E.

For example the gradient flow of the energy

$$E(u) = \int \frac{u^{m-1}}{m-1} \ , \ m > 1$$

corresponds to the solution of the porous medium equation

$$\partial_t u = \Delta u^m.$$

Another example which is in particular interesting for us is the Dirichlet energy

$$E(u) = \int_{\mathbb{R}^m} |\nabla u|^2 \, dx.$$

The first variation of the energy is given by

$$\frac{\delta E(u)}{\delta u} = -\Delta u.$$

Therefore the corresponding PDE is the *thin-film equation*:

$$\partial_t u = -\nabla \cdot (u \nabla \Delta u).$$

Our well-posedness proofs are based on the *minimizing movement scheme* as a discrete-time approximation of a gradient flow which is described here. Let $\mu_0 \in D(E)$, and fix a step size $\tau > 0$. Recursively define a sequence $\{M_\tau^n\}_{n=1}^{+\infty}$ by setting $M_0^\tau = \mu_0$, and

$$M_n^\tau = \operatorname*{argmin}_{\mu \in D(E)} \left\{ E(\mu) + \frac{1}{2\tau} W_2^2 \left(M_{n-1}^\tau, \mu \right) \right\} \quad n \geqslant 1. \tag{3.10}$$

To see why this scheme is an approximation of the gradient flow, we formally calculate the Euler-Lagrange equation for this minimization problem. The corresponding Euler-Lagrange is given by

$$U_n^\tau \in -\partial E\left(M_n^\tau\right)$$

where $U_n^\tau = -\frac{T_{M_n^\tau}^{M_{n-1}^\tau} - Id}{\tau}$. Next we define a piecewise constant curve and a corresponding velocity field by

$$\mu_t^\tau := M_n^\tau\,, \qquad V_t^\tau := U_n^\tau\,, \qquad \text{for } (n-1)\tau < t \leq n\tau\,.$$

We have

$$V_t^\tau \in -\partial E(\mu_t^\tau) \quad \forall t > 0. \tag{3.11}$$

Therefore this equation formally suggests that μ_t^τ is an approximation of the gradient flow trajectory of E starting from μ_0.

There is a standard set of hypothesises on the energy functional that we assume throughout this chapter. We gather the hypothesises here:

- E is nonnegative, and its sub-level sets are locally compact in the Wasserstein space.
- E is lower semicontinuous under narrow convergence.
- $D(E) \subseteq \mathcal{P}_2^a(\mathbb{R}^m)$.

The first two conditions guarantee existence of a solution to the minimization problem in the minimizing movement scheme (3.10) and convergence of the scheme to a limiting curve. The third condition ensures that measures of finite energy are absolutely continuous, allowing us to use transport maps rather than more general transportation plans for studying the Wasserstein distance. This simplifies the calculations, and allows us to view the subdifferential as a tangent vector. Note that these conditions can be relaxed as in [1], but to make the presentation more transparent, we prefer to work in this more concrete setting which is sharp enough for the energies we are interested in.

The standard condition in the literature that guarantees existence and uniqueness of the Wasserstein gradient flow of an energy is given by displacement convexity of the energy. This condition asks for convexity of the energy along geodesics of the Wasserstein space. Let $\mu_s : [0,1] \longrightarrow \mathcal{P}_2(\mathbb{R}^m)$ be the geodesic between $\mu_0, \mu_1 \in \mathcal{P}_2(X)$. An energy functional E is called *displacement convex* along μ_s if

$$E(\mu_s) \leqslant (1-s)E(\mu_0) + sE(\mu_1) \qquad s \in [0,1].$$

More generally the energy is called λ-*displacement convex* or in short λ-convex if the convexity is bounded from below by the constant λ, i.e.

$$E(\mu_s) \leqslant (1-s)E(\mu_0) + sE(\mu_1) - \frac{\lambda}{2}s(1-s)W_2^2(\mu_0,\mu_1) \qquad s \in [0,1]. \tag{3.12}$$

Furthermore, assuming that $E(\mu_s)$ is smooth as a function of s, one can write a derivative version of λ-convexity. In this case, E is λ-convex along μ_s if

$$\frac{d^2}{ds^2}E(\mu_s) \geqslant \lambda W_2^2(\mu_0,\mu_1). \tag{3.13}$$

An energy is called λ-convex if it satisfies (3.12) along all geodesics of $\mathcal{P}_2(\mathbb{R}^m)$. It is known that under λ-convexity assumption, the gradient flow of an energy exists and is unique (see [1]). In the next chapter we introduce a relaxed notion of convexity and we prove that, under the relaxed convexity condition of an energy, its gradient flow exists and is unique at least locally in time.

Chapter 4

Relaxed λ-convexity and Local Well-posedness

4.1 Relaxed λ-convexity

In this chapter we start by introducing a generalized notion of convexity which we call *relaxed λ-convexity*. Later in this chapter we will show that if the energy satisfies relaxed convexity, its gradient flow is locally well-posed. Relaxed convexity condition study the convexity of the energy locally on energy sub level-sets, i.e. in $B_\delta(\mu) \cap E_c$, where $B_\delta(\mu) = \left\{\nu \;\middle|\; W_2(\mu, \nu) < \delta\right\}$ and $E_c = \left\{\nu \;\middle|\; E(\nu) < c\right\}$

Definition 4.1.1 (relaxed λ-convexity.) *We say that an energy E is relaxed λ-convex at $\mu \in D(E)$, if there exist $\delta > 0$ and $c > E(\mu)$ such that E is λ-convex along the geodesics connecting any pair of measures $\nu_1, \nu_2 \in B_\delta(\mu) \cap E_c$.*

The following lemma has a key role in the arguments of Theorem 4.2.2. It shows how one can use relaxed λ-convexity assumption to study the subdifferential of an energy.

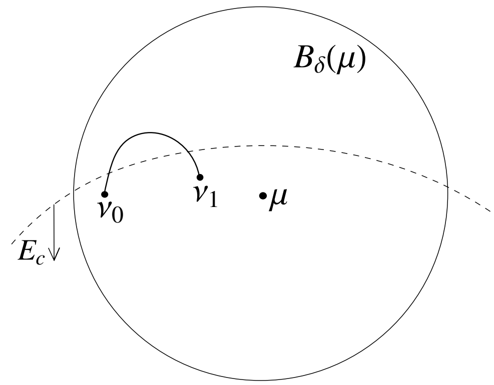

Figure 4.1: relaxed λ-convexity

Lemma 4.1.2 (Subdifferential and relaxed λ-convexity.) *Assume that E is relaxed λ-convex at μ. Then a vector field $\xi \in L^2(d\mu)$ belongs to the subdifferential of E at μ if and only if*

$$E(\nu) - E(\mu) \geqslant \int_X \langle \xi, T_\mu^\nu - Id \rangle \, \mathrm{d}\mu + \frac{\lambda}{2} W_2^2(\mu, \nu) \qquad \forall \nu \in B_\delta(\mu) \cap E_c \tag{4.1}$$

where $B_\delta(\mu) \cap E_c$ is the corresponding relaxed λ-convexity domain at μ.

Proof. First we claim that for studying the subdifferential of the functional, it is enough to consider the restricted domain $B_\delta(\mu) \cap E_c$. Let $\xi \in L^2(d\mu)$, we have to show that

$$\liminf_{\substack{\nu \to \mu \\ \nu \in D(E)}} \frac{E(\nu) - E(\mu) - \int_X \langle \xi, T_\mu^\nu - Id \rangle \, \mathrm{d}\mu}{W_2(\mu, \nu)} \geqslant 0. \tag{4.2}$$

if and only if

$$\liminf_{\substack{\nu \to \mu \\ \nu \in B_\delta(\mu) \cap E_c}} \frac{E(\nu) - E(\mu) - \int_X \langle \xi, T_\mu^\nu - Id \rangle \, \mathrm{d}\mu}{W_2(\mu, \nu)} \geqslant 0. \tag{4.3}$$

$\Downarrow$ is trivial.

For $\Uparrow$ assume that $\{\nu_n\}_1^\infty$ is a minimizing sequence for (4.2).

In the case that $\liminf_{n\to\infty} E(\nu_n) - E(\mu) > 0$ we have

$$\begin{aligned}\frac{E(\nu_n) - E(\mu) - \int_X \langle \xi, T_\mu^{\nu_n} - Id\rangle \, \mathrm{d}\mu}{W_2(\mu,\nu_n)} &\geqslant \frac{E(\nu_n) - E(\mu) - \left(\int_X |\xi|^2 \, \mathrm{d}\mu\right)^{1/2} \left(\int_X \left| T_\mu^{\nu_n} - Id\right|^2 \mathrm{d}\mu\right)^{1/2}}{W_2(\mu,\nu_n)} \\ &= \frac{E(\nu) - E(\mu)}{W_2(\mu,\nu_n)} - \left(\int_X |\xi|^2 \, \mathrm{d}\mu\right)^{1/2} \frac{\left(\int_X \left| T_\mu^{\nu_n} - Id\right|^2 \mathrm{d}\mu\right)^{1/2}}{W_2(\mu,\nu_n)} \\ &= \left[\frac{E(\nu) - E(\mu)}{W_2(\mu,\nu_n)} - |\xi|_{L^2_\mu}\right] \longrightarrow +\infty.\end{aligned}$$

Therefore inequality (4.2) is automatically true if $\liminf_{n\to\infty} E(\nu_n) - E(\mu) > 0$. Hence one only needs to consider sequences $\{\nu_n\}_1^\infty$ such that $\lim_{n\to\infty} E(\nu_n) - E(\mu) \leqslant 0$. Therefore, for large enough n we have $E(\nu_n) < c$. On the other hand, $\nu_n \xrightarrow{W_2} \mu$. Hence (4.3) and (4.2) are equivalent.

It is clear that (4.1) implies (4.3). Conversely, let $\xi \in L^2(d\mu)$ satisfy (4.3). Let $\nu \in B_\delta(\mu) \cap E_c$. Since E is relaxed λ-convex at μ, we have λ-convexity of E along the geodesic μ_s connecting μ to ν. Therefore

$$E(\mu_s) \leqslant (1-s)E(\mu) + sE(\nu) - \frac{\lambda}{2} s(1-s) W_2^2(\mu,\nu) \qquad \forall s \in [0,1].$$

Dividing by s and reordering, we have

$$\frac{E(\mu_s) - E(\mu)}{s} \leqslant E(\nu) - E(\mu) - \frac{\lambda}{2}(1-s) W_2^2(\mu,\nu). \tag{4.4}$$

ξ is in the subdifferential of E at μ. Hence

$$\begin{aligned}\liminf_{s\to 0^+} \frac{E(\mu_s) - E(\mu)}{s} &\geqslant \lim_{s\to 0^+} \frac{1}{s} \int_X \langle \xi, T_\mu^{\mu_s} - Id\rangle \, \mathrm{d}\mu \\ &= \int_X \langle \xi, T_\mu^{\nu} - Id\rangle \, \mathrm{d}\mu\end{aligned} \tag{4.5}$$

where we used linearity of the interpolate map $T_\mu^{\mu_s} = Id + s(T_\mu^\nu - Id)$. Therefore (4.4) and (4.5) imply

$$E(\nu) - E(\mu) \geqslant \int_X \langle \xi, T_\mu^{\nu} - Id\rangle \, \mathrm{d}\mu + \frac{\lambda}{2} W_2^2(\mu,\nu).$$

□

4.2 Local Existence and Uniqueness

The following lemma is used in Theorem 4.2.2 when we study weak convergence of tangent vector fields.

Lemma 4.2.1 *Let* $\mu_t^k, \mu_t \in AC^2([0,\hat{t}]; \mathcal{P}_2^a(\mathbb{R}^m))$ *and let* $V_t^k \in L^2(d\mu_t^k)$, $V_t \in L^2(d\mu_t)$. *Assume that*

- $\mu_t^k \xrightarrow{W_2} \mu_t$ *uniformly on* $[0,\hat{t}]$.
- V_t^k weakly converges *to* V_t *in the sense that* $\forall U \in C_b^0([0,\hat{t}] \times \mathbb{R}^m)$, *we have*

$$\lim_{k\to\infty} \int_0^{\hat{t}} \int_{\mathbb{R}^m} \left\langle V_t^k, U(t,x) \right\rangle d\mu_t^k dt = \int_0^{\hat{t}} \int_{\mathbb{R}^m} \left\langle V_t, U(t,x) \right\rangle d\mu_t dt.$$

Then $\forall \nu \in \mathcal{P}_2(\mathbb{R}^m)$

$$\lim_{k\to\infty} \int_0^{\hat{t}} \int_{\mathbb{R}^m} \left\langle V_t^k, T_{\mu_t^k}^{\nu} - Id \right\rangle d\mu_t^k dt = \int_0^{\hat{t}} \int_{\mathbb{R}^m} \left\langle V_t, T_{\mu_t}^{\nu} - Id \right\rangle d\mu_t dt.$$

Proof. Since V_t^k is weakly convergent by uniform boundedness principle we have $\sup_k \int_0^{\hat{t}} \int_{\mathbb{R}^m} |V_t^k|^2 d\mu_t^k dt < +\infty$. Let

$$M = \sup_k \int_0^{\hat{t}} (\int_{\mathbb{R}^m} \left|V_t^k\right|^2 d\mu_t^k + \int_{\mathbb{R}^m} \left|V_t\right|^2 d\mu_t) dt.$$

Choose $T_t \in C_c^0([0,\hat{t}] \times \mathbb{R}^m)$ such that

$$\int_0^{\hat{t}} \int_{\mathbb{R}^m} |T_{\mu_t}^{\nu} - T_t|^2 d\mu_t dt < \epsilon^2.$$

We have

$$\begin{aligned}
&\left|\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t^k, T^{\nu}_{\mu_t^k}-Id\right\rangle d\mu_t^k dt-\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t, T^{\nu}_{\mu_t}-Id\right\rangle d\mu_t dt\right|\\
&\leqslant \underbrace{\left|\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t^k, Id\right\rangle d\mu_t^k dt-\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t, Id\right\rangle d\mu_t dt\right|}_{A}\\
&+\underbrace{\left|\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t^k, T^{\nu}_{\mu_t^k}-T_t\right\rangle d\mu_t^k dt\right|}_{B}\\
&+\underbrace{\left|\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t^k, T_t\right\rangle d\mu_t^k dt-\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t, T^{\nu}_{\mu_t}\right\rangle d\mu_t dt\right|}_{C}
\end{aligned}$$

We study each of the items separately.

Since μ_t^k is uniformly converging to μ_t, the second moment of μ_t^k is uniformly bounded. In particular there is a compact set $S \subset [0,\hat{t}] \times \mathbb{R}^m$ such that

$$\left(\int_{S^c}|x|^2\mu_t dt+\sup_k\int_{S^c}|x|^2\mu_t^k dt\right)<\epsilon^2. \tag{4.6}$$

We have

$$\begin{aligned}
\lim_{k\to\infty}A &= \lim_{k\to\infty}\left|\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t^k, Id\right\rangle d\mu_t^k dt-\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t, Id\right\rangle d\mu_t dt\right|\\
&\leqslant \lim_{k\to\infty}\left|\int_S\left\langle V_t^k, Id\right\rangle d\mu_t^k dt-\int_S\left\langle V_t, Id\right\rangle d\mu_t dt\right|\\
&+\lim_{k\to\infty}\left|\int_{S^c}\left\langle V_t^k, Id\right\rangle d\mu_t^k dt-\int_{S^c}\left\langle V_t, Id\right\rangle d\mu_t dt\right|.
\end{aligned}$$

Because S is compact one can use weak convergence of V_t^k on S. Hence the limit of the first term vanishes and we have

$$\begin{aligned}\lim_{k\to\infty} A &\leqslant \lim_{k\to\infty}\left|\int_{S^c}\left\langle V_t^k, Id\right\rangle d\mu_t^k dt - \int_{S^c}\left\langle V_t, Id\right\rangle d\mu_t dt\right| \\ &\leqslant \frac{\epsilon}{2}\lim_{k\to\infty}\left(\int_{S^c}|V_t^k|^2 d\mu_t^k dt + \int_{S^c}|V_t|^2 d\mu_t dt\right) \\ &+ \lim_{k\to\infty}\frac{1}{2\epsilon}\left(\int_{S^c}|x|^2 d\mu_t^k dt + \int_{S^c}|x|^2 d\mu_t dt\right)\end{aligned}$$

where we used Young's inequality with the constant ϵ. By (4.6) we have

$$\lim_{k\to\infty} A \leqslant \frac{\epsilon M}{2} + \frac{\epsilon}{2}.$$

Since ϵ is arbitrary we have $\lim_{k\to\infty} A = 0$.

We now study B. Consider the measure γ_t^k on $\mathbb{R}^m \times \mathbb{R}^m$ given by

$$\gamma_t^k = (Id \times T^{\nu}_{\mu_t^k})_{\#}\mu_t^k.$$

Recall that the measure γ_t^k is the optimal plan with marginals μ_t^k and ν. Since $\mu_t^k \to \mu_t$, by the stability of optimal plans [27, Theorem 5.20], the set of optimal plans between μ_t^k and ν is compact in the narrow topology and every limit point is an optimal plan between μ_t and ν. On the other hand, because μ_t is an absolutely continuous measure, Brenier-McCann Theorem ensures that the optimal plan between μ_t and ν is unique. This implies that the sequence γ_t^k converges narrowly for all $t \in [0, \hat{t}]$. Furthermore, the uniform convergence of μ_t^k implies that γ_t^k have uniformly bounded second moment. We have

$$\gamma_t^k = (Id \times T^{\nu}_{\mu_t^k}))_{\#}\mu_t^k \xrightarrow{narrow} \gamma_t = (Id \times T^{\nu}_{\mu_t}))_{\#}\mu_t \quad \forall t \in [0, \hat{t}]. \tag{4.7}$$

Taking the limit of B yields

$$\begin{aligned}\lim_{k\to\infty} B &= \lim_{k\to\infty}\left|\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t^k, T_{\mu_t^k}^{\nu} - T_t\right\rangle d\mu_t^k dt\right| \\ &\leqslant \lim_{k\to\infty}\frac{\epsilon}{2}\int_0^{\hat{t}}\int_{\mathbb{R}^m}|V_t^k|^2 d\mu_t^k dt + \frac{1}{2\epsilon}\int_0^{\hat{t}}\int_{\mathbb{R}^m}|T_{\mu_t^k}^{\nu} - T_t| d\mu_t^k dt \\ &\leqslant \frac{\epsilon M}{2} + \frac{1}{2\epsilon}\lim_{k\to\infty}\int_0^{\hat{t}}\int_{\mathbb{R}^m}|T_{\mu_t^k}^{\nu} - T_t|^2 d\mu_t^k dt.\end{aligned}$$

by lifting to the optimal plans $\gamma_t^k = (Id \times T_{\mu_t^k}^{\nu})_{\#}\mu_t^k$, we have

$$\lim_{k\to\infty} B \leqslant \frac{\epsilon M}{2} + \frac{1}{2\epsilon}\lim_{k\to\infty}\int_0^{\hat{t}}\int_{\mathbb{R}^m\times\mathbb{R}^m}|y - T_t(x)|^2 d\gamma_t^k dt.$$

Since $\gamma_t^k \to \gamma_t$ point-wise, γ_t^k has uniformly bounded second moment, and $|y - T_t(x)|^2$ is dominated by a constant times $|x^2 + y^2 + 1|$, we can use dominated convergence theorem. Therefore

$$\begin{aligned}\lim_{k\to\infty} B &\leqslant \frac{\epsilon M}{2} + \frac{1}{2\epsilon}\lim_{k\to\infty}\int_0^{\hat{t}}\int_{\mathbb{R}^m\times\mathbb{R}^m}|y - T_t(x)|^2 d\gamma_t^k dt \\ &= \frac{\epsilon M}{2} + \frac{1}{2\epsilon}\int_0^{\hat{t}}\int_{\mathbb{R}^m\times\mathbb{R}^m}|y - T_t(x)|^2 d\gamma_t dt \\ &= \frac{\epsilon M}{2} + \frac{1}{2\epsilon}\int_0^{\hat{t}}\int_{\mathbb{R}^m}|T_{\mu_t}^{\nu} - T_t|^2 d\mu_t dt \\ &\leqslant \frac{\epsilon M}{2} + \frac{\epsilon}{2}.\end{aligned}$$

Finally, we study the last term C. We have

$$\begin{aligned}\lim_{k\to\infty} C &= \lim_{k\to\infty}\left|\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t^k, T_t\right\rangle d\mu_t^k dt - \int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t, T_{\mu_t}^{\nu}\right\rangle d\mu_t dt\right| \\ &\leqslant \lim_{k\to\infty}\left|\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t^k, T_t\right\rangle d\mu_t^k dt - \int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t, T_t\right\rangle d\mu_t dt\right| \\ &+ \left|\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t, T_{\mu_t}^{\nu} - T_t\right\rangle d\mu_t dt\right|.\end{aligned}$$

Since $T_t \in C_b^0$ we can use weak convergence of V_t^k for the first term. Hence

$$\begin{aligned}
\lim_{k\to\infty} C &\leqslant \left| \int_0^{\hat{t}} \int_{\mathbb{R}^m} \left\langle V_t, T_{\mu_t}^{\nu} - T_t \right\rangle d\mu_t dt \right| \\
&\leqslant M \int_0^{\hat{t}} \int_{\mathbb{R}^m} |T_{\mu_t}^{\nu} - T_t|^2 d\mu_t dt \\
&\leqslant M\epsilon^2.
\end{aligned}$$

□

Theorem 4.2.2 (Existence and uniqueness of the flow) *Let $E : \mathcal{P}_2(\mathbb{R}^m) \to [0, +\infty]$ be a lower semicontinuous energy functional with locally compact sub-level sets and let $D(E) \subseteq \mathcal{P}_2^a(\mathbb{R}^m)$. Assume $E(\mu) < +\infty$ and that E is relaxed λ-convex at μ. Then there exist $\hat{t} > 0$ and a curve $\mu_t \in AC^2\left([0,\hat{t}]; \mathcal{P}_2^a(\mathbb{R}^m)\right)$ such that μ_t is the unique gradient flow of E starting from μ.*

Proof. Let $\mu_t^k := \mu_t^\tau$ be a piecewise constant solution to the minimizing movement scheme (3.10) with $\tau = \frac{1}{k}$. The minimizing movement sequence is designed in a way that it converges to a limiting curve in a very general setting. In [1, Theorem 11.1.6] it has been proved that, under very weak assumptions which hold here, the minimizing movement scheme converges sub-sequentially to a limiting curve such that (after relabelling) $\forall a > 0$

- $\mu_t^k \xrightarrow{W_2} \mu_t \in AC^2\left([0,a]; \mathcal{P}_2(\mathbb{R}^m)\right)$ uniformly in $[0,a]$.
- The sequence $\{V_t^k\}$ of the velocity tangent vectors to $\{\mu_t^k\}$ converges weakly to $V_t \in L^2(d\mu)$ in $\mathbb{R}^m \times (0,T)$.
- The continuity equation $\partial_t \mu_t + \nabla.\left(\mu_t V_t\right) = 0$ holds for the limiting curve.

We need to prove that the limiting curve μ_t satisfies the steepest descent equation and that it is unique. Let $B_\delta(\mu) \cap E_c$ be the domain of relaxed convexity at μ. Since μ_t is a continuous curve, we can find $\hat{t}$ such that $\mu_t \in B_{\delta/4}(\mu)$ for all $t \in [0,\hat{t}]$. We have

$$E(\mu_t) \leqslant \liminf_{k\to\infty} E(\mu_t^k) \leqslant E(\mu) < c. \tag{4.8}$$

The first inequality follows from lower semicontinuity of the energy and the second inequality follows from the structure of the minimizing movement scheme (3.10). Hence

$$\mu_t \in B_{\delta/4}(\mu) \cap E_c \quad \forall t \in [0,\hat{t}]. \tag{4.9}$$

Since $\mu_t^k \xrightarrow{W_2} \mu_t$ uniformly in $[0,\hat{t}]$, we can find $K \in \mathbb{N}$ such that $W_2(\mu_t,\mu_t^k) < \delta/4$, $\forall k \geqslant K$ and $\forall t \in [0,\hat{t}]$. Without loss of generality we assume that $K = 1$. Therefore (4.8) and (4.9) imply

$$\mu_t^k \in B_{\delta/2}(\mu) \cap E_c. \tag{4.10}$$

The Euler-Lagrange equation of minimizing movement (3.11) implies that $-V_t^k \in \partial E(\mu_t^k)$. Therefore using (4.10) and the variational formulation of the subdifferential (Lemma 4.1.2) we have

$$E(\nu) - E(\mu_t^k) \geqslant \int_{\mathbb{R}^m} \left\langle -V_t^k, T_{\mu_t^k}^{\nu} - Id \right\rangle d\mu_t^k + \frac{\lambda}{2} W_2^2(\mu_t^k,\nu) \tag{4.11}$$

for all $\nu \in B_{\delta/2}(\mu) \cap E_c$. By construction, $B_{\delta/4}(\mu_t) \cap E_c \subseteq B_{\delta/2}(\mu_t^k) \cap E_c$. Therefore (4.11) holds for all ν in $B_{\delta/4}(\mu_t) \cap E_c$. By integrating (4.11) over t and against a text function $\psi \in C_c^\infty((0,\hat{t});[0,\infty))$ we have

$$\int_0^{\hat{t}} E(\nu)\psi(t)dt - \int_0^{\hat{t}} E(\mu_t^k)\psi(t)dt \geqslant \int_0^{\hat{t}} \int_{\mathbb{R}^m} \left\langle -V_t^k, T_{\mu_t^k}^{\nu} - Id \right\rangle \psi(t) d\mu_t^k dt + \frac{\lambda}{2} \int_0^{\hat{t}} W_2^2(\mu_t^k,\nu)\psi(t)dt \tag{4.12}$$

for all ν in $B_{\delta/4}(\mu_t) \cap E_c$.

We take the limit of (4.12) as $k \to \infty$. By the lower semicontinuity of E

$$\int_0^{\hat{t}} E(\nu)\psi(t)dt - \int_0^{\hat{t}} E(\mu_t)\psi(t)dt \geqslant \int_0^{\hat{t}} E(\nu)\psi(t)dt - \liminf_{k\to\infty} \int_0^{\hat{t}} E(\mu_t^k)\psi(t)dt.$$

Lemma 4.2.1 implies that

$$\lim_{k\to\infty}\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t^k, T^{\nu}_{\mu_t^k}-Id\right\rangle\psi(t)d\mu_t^k dt=\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle V_t, T^{\nu}_{\mu_t}-Id\right\rangle\psi(t)d\mu_t dt.$$

By the triangle inequality

$$W_2(\mu_t,\nu)-W_2(\mu_t^k,\mu_t)\leqslant W_2(\mu_t^k,\nu)\leqslant W_2(\mu_t,\nu)+W_2(\mu_t^k,\mu_t). \tag{4.13}$$

Therefore $\lim_{k\to\infty}W_2(\mu_t^k,\nu)=W_2(\mu_t,\nu)$. Furthermore, since $\mu_t^k\xrightarrow{W_2}\mu_t$ uniformly, inequality (4.13) implies that $W_2(\mu_t^k,\nu)$ is uniformly bounded. Hence, by dominated convergence theorem

$$\lim_{k\to\infty}\int_0^{\hat{t}}W_2^2(\mu_t^k,\nu)\psi(t)dt=\int_0^{\hat{t}}W_2^2(\mu_t,\nu)\psi(t)dt.$$

In conclusion $\forall\nu\in B_{\delta/4}(\mu_t)\cap E_c$ and $\forall\psi\in C_c^\infty((0,\hat{t});[0,\infty))$ we have

$$\int_0^{\hat{t}}E(\nu)\psi(t)dt-\int_0^{\hat{t}}E(\mu_t)\psi(t)dt\geqslant\int_0^{\hat{t}}\int_{\mathbb{R}^m}\left\langle -V_t, T^{\nu}_{\mu_t}-Id\right\rangle\psi(t)d\mu_t dt+\frac{\lambda}{2}\int_0^{\hat{t}}W_2^2(\mu_t,\nu)\psi(t)dt. \tag{4.14}$$

Let t_0 be a Lebesgue point of the map $t\mapsto\int_0^{\hat{t}}E(\mu_t)\psi(t)dt+\int_{\mathbb{R}^m}\left\langle V_t, T^{\nu}_{\mu_t}-Id\right\rangle d\mu_t+\frac{\lambda}{2}\int_0^{\hat{t}}W_2^2(\mu_t,\nu)\psi(t)dt$. By considering a sequence of smooth mollifiers ψ_n converging to the delta function at t_0, the inequality (4.14) is reduced to

$$E(\nu)-E(\mu_{t_0})\geqslant\int_{\mathbb{R}^m}\left\langle -V_{t_0}, T^{\nu}_{\mu_{t_0}}-Id\right\rangle d\mu_{t_0}+\frac{\lambda}{2}W_2^2(\mu_{t_0},\nu). \tag{4.15}$$

Therefore (4.15) holds for almost all t. By Lemma 4.1.2 we have

$$V_t\in-\partial E(\mu_t)\quad\text{for almost all } t\in[0,\hat{t}].$$

We now study uniqueness of the solution. The available uniqueness proofs in the case of λ-convexity can be repeated in the domain of relaxed λ-convexity because as soon as the flow exists clearly it dissipates the energy and the trajectories of the flow starting from μ remain in the domain of relaxed λ-convexity for a short time where one can use λ-convexity. Hence uniqueness arguments can be repeated similar to the available proofs such as in [1, Theorem 11.1.4]. Therefore we provide only the key ideas here.

Assume that we have two gradient flows μ_t^1 and μ_t^2 both starting from μ. One can show that $W_2^2(\mu_t^1, \mu_t^2)$ is absolutely continuous in time and that, by differentiability of Wasserstein metric (3.3), for almost all $t \in [0, \hat{t}]$ we have

$$\frac{d}{dt} W_2^2(\mu_t^1, \mu_t^2) \leqslant \int_{\mathbb{R}^m} \left\langle V_t^2, Id - T_{\mu_t^2}^{\mu_t^1} \right\rangle d\mu_t^2 + \int_{\mathbb{R}^m} \left\langle V_t^1, Id - T_{\mu_t^1}^{\mu_t^2} \right\rangle d\mu_t^1. \tag{4.16}$$

Consider inequality (4.15) along μ_t^1. We have

$$E(\mu_t^2) - E(\mu_t^1) \geqslant \int_{\mathbb{R}^m} \left\langle -V_t, T_{\mu_t^1}^{\mu_t^2} - Id \right\rangle d\mu_t^1 + \frac{\lambda}{2} W_2^2(\mu_t^1, \mu_t^2). \tag{4.17}$$

Rewriting (4.17) again along μ_t^2 and using (4.16) result in

$$\frac{1}{2} \frac{d}{dt} W_2^2(\mu_t^2, \mu_t^1) \leqslant -\lambda W_2^2(\mu_t^2, \mu_t^1).$$

Hence

$$W_2(\mu_t^2, \mu_t^1) \leqslant e^{-\lambda t} W_2(\mu_0^2, \mu_0^1) = 0 \quad \forall t \in [0, \hat{t}]. \tag{4.18}$$

□

Remarks. Since the result of Theorem 4.2.2 holds as long as λ is finite, the flow exists and is unique unless there is a blow up on λ. Also because of the inequality 4.18 the rate of divergence of the flow starting from two close-by points μ_1 and μ_2 is controlled by $min\,\{\lambda_1, \lambda_2\}$ where λ_i is the modulus of relaxed convexity at μ_i. Therefore the theorem implicitly results in continuous dependent on the initial data too.

Chapter 5

Wasserstein Gradient Flow of Dirichlet Energy

In this section we prove a local well-posedness result for the gradient flow of the Dirichlet energy on S^1. In this chapter, we use letter E to refer to the *Dirichlet energy*

$$E(\mu) = \begin{cases} \frac{1}{2}(\partial_x u)^2 & \text{if } \mu = udx, u \in H^1(S^1), \\ +\infty & \text{else.} \end{cases} \tag{5.1}$$

When convenient, we refer to an absolutely continuous measure $\mu = udx$ by its density u. In particular, by a smooth or positive measure we mean a measure with a smooth or positive density.

5.1 Optimal Transport on S^1

The underlying space of the measures that we study in this section is S^1. We identify S^1 with $\mathbb{R}/\mathbb{Z}$. Because S^1 is a manifold, the theory developed in the previous section should be slightly modified. On a Riemannian manifold, there is the issue of existence and regularity of the optimal maps. This question has been an active area of research. Ma-Trudinger-Wang condition in [21] is a famous example that studies this issue. In the case of S^m, the problem has been addressed and positive results are available such as [17] and [20]. The results guarantee that between any pair of smooth positive measures μ, ν on S^m there exists a unique smooth optimal map T_μ^ν. To study the problem on $S^1 \equiv \mathbb{R}/\mathbb{Z}$, because of the periodicity, one has to be careful about the optimal map because $|T(x)-x|$ refers to two different values on S^1. As was suggested in [12], this problem can be solved by relabelling S^1. In short, it was proved in [12] that for any two smooth positive measures μ and ν on S^1 there exist an optimal map $T : [0,1] \to [T(0), T(0)+1]$ between μ and ν. Further more the optimal map T is monotone and the geodesic distance of $S^1 \equiv \mathbb{R}/\mathbb{Z}$ coincides with $|T(x) - x|$.

Let $\mu_0 = u_0 dx$ and $\mu_1 = u_1 dx$ be two smooth measures on S^1 and let T be the optimal map between them. Then by Monge-Ampère equation (2.3) we have

$$u_1(T(X)) = \frac{u_0(x)}{T'(x)}.$$

Since the geodesics are given by the push forward of linear interpolation of the optimal map and the identity map, the explicit form of the geodesic u_s between u_0 and u_1 is given by

$$u_s((1-s)x + sT(x)) = \frac{u_0(x)}{(1-s) + sT'(x)}.$$

In the notation of the previous section, if we think of $f = T - Id$ as the tangent vector field that connects u_0 to u_1, the geodesic equation can be written as

$$u_s(x + sf(x)) = \frac{u_0(x)}{1 + sf'(x)}. \tag{5.2}$$

We will see in Lemma 5.2.3 that for studying relaxed λ-convexity of the energy, it is enough to consider measures with smooth densities. By the derivative formulation of λ-convexity (3.13) the energy is λ-convex along the geodesic μ_s connecting μ_0 to μ_1 if

$$\frac{d^2}{ds^2}E(u_s) \geqslant \lambda W_2^2(u_0, u_1). \tag{5.3}$$

We start by proving that the Dirichlet energy is not λ-convex for any λ. This is known to the community, and in [12] Carrillo and Slepčev proved that the Dirichlet energy is not convex on S^1. We will study a scalable family of functions to prove the lack of convexity for any $\lambda \in \mathbb{R}$. Let u_s be the geodesic connecting $u_0 = u$ to u_1, and let f be the corresponding tangent vector field. We compute the second derivative of the energy along the geodesic

$$\left.\frac{d^2 E(u_s)}{ds^2}\right|_{s=0} = \left.\frac{d^2}{ds^2}\right|_{s=0} \int_{S^1} (\partial_y u_s(y))^2 \, dy \, .$$

By the Monge-Ampère equation (5.2) and the change of variables $y = x + sf(x)$ we have

$$\begin{aligned}
\left.\frac{d^2 E(u_s)}{ds^2}\right|_{s=0} &= \left.\frac{d^2}{ds^2}\right|_{s=0} \int_{S^1} \left(\frac{\partial x}{\partial y} \frac{\partial}{\partial x} \frac{u(x)}{1 + sf'(x)} \right)^2 \frac{\partial y}{\partial x} dx \\
&= \left.\frac{d^2}{ds^2}\right|_{s=0} \int_{S^1} \left(\frac{1}{1 + sf'(x)} \partial_x \frac{u(x)}{1 + sf'(x)} \right)^2 (1 + sf'(x)) \, dx \\
&= \left.\frac{d^2}{ds^2}\right|_{s=0} \int_{S^1} \frac{((1 + sf'(x))u'(x) - su(x)f''(x))^2}{(1 + sf'(x))^3} \, dx \\
&= 2 \int_{S^1} \left\{ (f''u)^2 + 8(f''u)(f'u') + 6(f'u')^2 \right\} dx.
\end{aligned} \tag{5.4}$$

If the energy E was λ-convex, then

$$\underbrace{\int_{S^1} \left\{ (f''u)^2 + 8(f''u)(f'u') + 6(f'u')^2 \right\} dx}_{A} \geqslant \lambda \underbrace{\int_{S^1} f^2 u \, dx}_{B} \, . \tag{5.5}$$

We use (5.5) as a guide to find a counter-example. In the example, u and f are not smooth, and we cannot directly apply this computation but, Lemma 5.2.3 validates the calculations.

We view S^1 as the interval $[-1/2, 1/2]$ with the endpoints identified. The construction of the example is simple: let $u = 1 - 4|x|$ and $f' = u^{-1}$, this forces the integrand of A to be negative, and the rest follows from a scaling argument. We have to make some modifications to the functions so that the integral converges and the mass is normalized to 1. We define u and f' as follows

$$u(x) = \begin{cases} \frac{81}{16}(1-4|x|)\,, & 0 \leqslant |x| \leqslant \frac{2}{9}, \\ \frac{9}{16}\,, & \frac{2}{9} \leqslant |x| \leqslant \frac{3}{8}, \\ \frac{9}{4}(1-2|x|)\,, & \frac{3}{8} \leqslant |x| \leqslant \frac{1}{2} \end{cases} \quad f'(x) = \begin{cases} \frac{16}{81(1-4|x|)}\,, & 0 \leqslant |x| \leqslant \frac{2}{9}, \\ \frac{16}{11}(3-8|x|)\,, & \frac{2}{9} \leqslant |x| \leqslant \frac{3}{8}, \\ 0\,, & \frac{3}{8} \leqslant |x| \leqslant \frac{1}{2}. \end{cases}$$

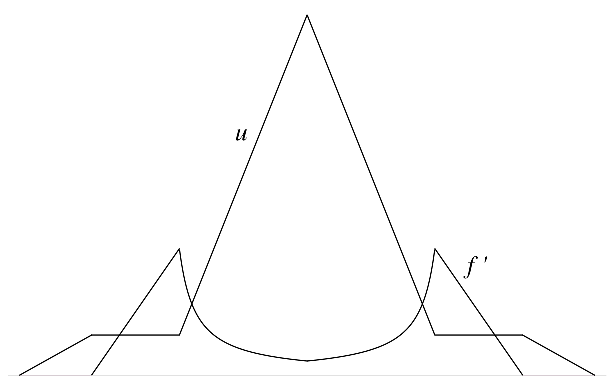

Figure 5.1: Graph of u and f'

By scaling $u_h(x) := hu(hx)$ and $f'_h(x) := \frac{1}{h}f'(hx)$ we have

$$A = -C_1 h^2\,, \quad B = C_2$$

for some positive constants C_1, C_2. If the energy is λ-convex then we must have $A \geqslant \lambda B$ and it should hold uniformly for any such u and f. But for a fixed λ, we can choose h

large enough so that the opposite inequality holds. This means that the Dirichlet energy is not λ-convex on $\mathcal{P}_2(S^1)$.

In the example above, by pushing h to larger numbers the lack of convexity becomes worse. By looking at the equation of $u(x)$, it is clear that the Dirichlet energy of $u(x)$ gets bigger for larger values of h. This example hints that one of the obstructions against the λ-convexity of the Dirichlet energy is the magnitude of the energy which can be controlled on energy sub-level sets.

5.2 Relaxed λ-convexity of Dirichlet Energy

Lemma 5.2.1 (Uniform convergence on energy sub-level sets.) *Uniform convergence and Wasserstein convergence are equivalent on energy sub-level sets of Dirichlet energy on S^1. In particular for two measures $\mu_1 = u_1 dx$ and $\mu_2 = u_2 dx$ with $E(\mu_1), E(\mu_2) < c < +\infty$ we have*

$$W_2^2(\mu_1, \mu_2) \geqslant \alpha |u_1 - u_2|_\infty^\beta \tag{5.6}$$

where $\alpha = \alpha(c)$ and β are constants.

Proof. One side of the equivalence is easy. Assuming $u_n \xrightarrow{uniform} u_0$ we have

$$\int_{S^1} \psi u_n dx \longrightarrow \int_{S^1} \psi u dx \quad \forall \psi \in C^0(S^1)$$

which implies Wasserstein convergence of $\mu_n = u_n dx$ to $\mu = u dx$ by (3.5) and finiteness of the second moments on S^1.

For the converse inequality, we first study the regularity of a measure with finite energy. Let $\nu = vdx \in E_c$. By Poincare's inequality and $\int_{S^1} vdx = 1$ we have

$$\int_{S^1} |v|^2 dx \leqslant \int_{S^1} |v'|^2 dx + 2\int_{S^1} |v| dx + \int_{S^1} dx \leqslant c + 3.$$

Therefore H^1-norm of v is bounded by its energy. The Sobolev embedding theorem implies that v is $C^{0,1/2}$ continuous and we have

$$|v(x) - v(y)| \leqslant \left(\int_{S^1} |v'|^2 dx\right)^{1/2} \left(\int_x^y dx\right)^{1/2} \leqslant \sqrt{c|x-y|} \quad \forall x, y \in S^1. \tag{5.7}$$

Therefore the modulus of continuity is $\sqrt{c}$. Let μ_1 and μ_2 be as in the assumption. Therefore u_1 and u_2 are $C^{0,1/2}$ continuous with constant $\sqrt{c}$. Assume that $|u_1 - u_2|_\infty \geqslant h > 0$. In particular without loss of generality assume that for some point $x_0 \in S^1$ we have $u_1(x_0) - u_2(x_0) \geqslant h$. For every $x \in S^1$, we have

$$\begin{aligned} u_1(x) &\geqslant u_1(x_0) - \sqrt{c|x|} \\ u_2(x) &\leqslant u_2(x_0) + \sqrt{c|x|}. \end{aligned} \tag{5.8}$$

Therefore u_1 lies above and u_2 is below the star-like shape in Figure (5.2). Call the star-like shape by S. Consider a rectangle R in the center of S with height $\frac{h}{2}$ and width $\frac{k}{3}$ where k is the width of S at the height $u_2(x_0) + \frac{h}{4}$. We have $k = \frac{h^2}{8c}$ and the area of R is given by $\frac{h^3}{48c}$. In order to transport the measure μ_1 to μ_2, some mass at least equal to the area of R should be transported outside of S. Therefore

$$\begin{aligned} W_2^2(\mu_1, \mu_2) &\geqslant \{\text{area of R}\}\{\text{distance required to move R outside of S}\}^2 \\ &\geqslant \frac{h^3}{48c}.(\frac{k}{3})^2 \\ &\geqslant \frac{1}{384c^3}|u_1 - u_2|_\infty^7. \end{aligned}$$

$\square$

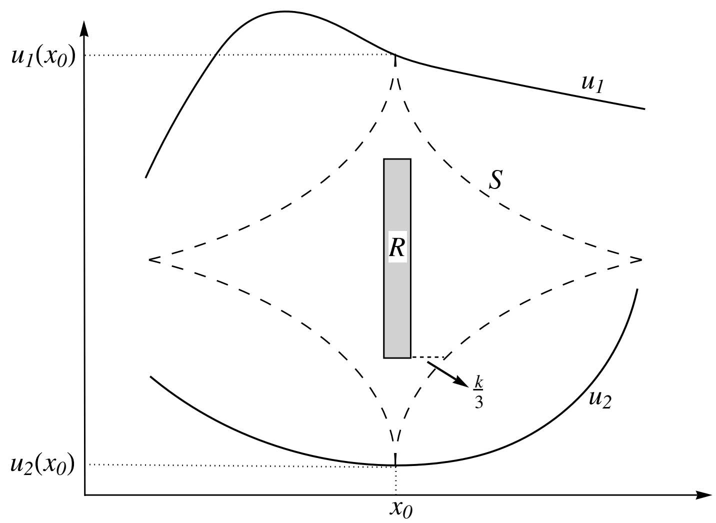

Figure 5.2: Wasserstein $\Longleftrightarrow$ Uniform

Lemma 5.2.2 (Lower Semicontinuity.) *The Dirichlet energy is lower semicontinuous with respect to the Wasserstein metric on S^1.*

Proof. Let $u_n \xrightarrow{W_2} u$. We have to prove $E(u) \leqslant \liminf_{n\to\infty} E(u_n)$. Consider a minimizing subsequnce which (without relabeling) we show it by $\{u_n\}$. Since $\{u_n\}$ is bounded, we can assume that $\{E(u_n)\}$ is bounded. This implies that H^1-norm of the sequence is bounded. By Banach–Alaoglu theorem, u_n has a weak limit point $v \in H^1$ with a subsequence u_{n_k} converges to v strongly in L^2. Since u_{n_k} is also a minimizing sequence, it is enough to show that $E(u) \leqslant \liminf_{n\to\infty} E(u_{n_k})$. We claim that $v = u$. Let $\psi \in C^0(S^1)$ we have

$$\begin{aligned}\left|\int_{s^1} \psi(x)(u-v)(x)dx\right| &\leqslant \left|\int_{s^1} \psi(x)(u-u_{n_k})(x)dx\right| + \left|\int_{s^1} \psi(x)(u_{n_k}-v)(x)dx\right| \\ &\leqslant \left|\int_{s^1} \psi(x)(u-u_{n_k})(x)dx\right| + (\int_{s^1} |\psi|^2 dx)^{1/2}(\int_{s^1} |u_{n_k}-v|^2 dx)^{1/2}\end{aligned}$$

By Lemma 5.2.1 the Wasserstein and uniform convergences are equivalent on energy sublevel sets. Therefore the first term in the last inequality goes to zero. The second term also goes to zero because u_{n_k} converges to v strongly in L^2. Hence $u = v$ almost everywhere. Because u and v are continuous, we have $u = v$. The Dirichlet energy is known to be lower semicontinuous under weak H^1 convergence (for example see [16, Theorem 8.2.1]). Hence we have $E(u) = E(v) \leqslant \liminf_{n\to\infty} E(u_{n_k})$. $\square$

The following lemma validates smooth calculation in the sense that for studying relaxed λ-convexity of the energy, one can study relaxed λ-convexity of the energy only on smooth measures.

Lemma 5.2.3 (Approximation by smooth measures.) *Let $\mu \in D(E)$. Assume that there exist $\delta > 0$ and $c > E(\mu)$ such that the energy is λ-convex along all geodesics with* ***smooth*** *endpoints in $B_\delta \cap E_c$. Then E is relaxed λ-convex at μ.*

Proof. Let $\mu_0 = u_0 dx,\ \mu_1 = u_1 dx \in B_\delta \cap E_c$ and let η_k be a standard smooth mollifier converging to the Dirac delta function. Define $u_{k,i}(x) := \eta_k * u_i(x)$ for $i = 0, 1$ where $*$ is the convolution on S^1. Since $u_{k,i} \xrightarrow{uniformly} u_i$, by Lemma 5.2.1 we have $u_{k,i} \xrightarrow{W_2} u_i$. Therefore for large enough k we have $u_{k,i} \in B_\delta(u)$. The energy also converges, because

$$E(u_{k,i}) = \int_{S^1} (\partial_x(u_i * \eta_k))^2\, dx = \int_{S^1} ((\partial_x u_i) * \eta_k)^2\, dx \xrightarrow[k\to\infty]{} \int_{S^1} (\partial_x u_i)^2\, dx = E(u_i). \tag{5.9}$$

Hence $u_{k,i} \in B_\delta(u) \cap E_c$ for large enough k. By smoothness of $u_{k,i}$ and the assumption of the lemma, we have λ-convexity of the energy along the geodesics $u_{k,s}$ connecting $u_{k,0}$ to $u_{k,1}$

$$E(u_{k,s}) \leqslant (1-s)E(u_{k,0}) + sE(u_{k,1}) - \frac{\lambda}{2} s(1-s) W_2^2(u_{n,0}, u_{k,1}). \tag{5.10}$$

Let γ_k and γ be in order the optimal plan connecting $\mu_{k,0} = u_{k,0}dx$ to $\mu_{k,1} = u_{k,1}dx$ and the optimal plan connecting μ_0 to μ_1. By stability of the optimal plans [27, Theorem

5.20] γ_k converges in narrow topology to γ along a subsequence which after relabelling we assume to be the whole sequence. Equivalence of narrow and Wasserstein convergence (5.6) on $S^1 \times S^1$ implies

$$\mu_{k,s} = \left((1-s)\Pi^1 + s\Pi^2\right)_{\#} \gamma_k \xrightarrow{W_2} \left((1-s)\Pi^1 + s\Pi^2\right)_{\#} \gamma = \mu_s$$

where Π^i is the projection to the i^{th} coordinate and μ_s is the geodesic connecting μ_0 to μ_1. The lower semicontinuity of Dirichlet energy 5.2.2 yields $E(\mu_s) \leqslant \liminf_{k\to+\infty} E(\mu_{k,s})$. Hence by taking the limit of (5.10) we have

$$E(\mu_s) \leqslant (1-s)E(\mu_0) + sE(\mu_1) - \frac{\lambda}{2}s(1-s)W_2^2(\mu_0, \mu_1).$$

□

In the following lemma we prove that the energy is finite along a geodesic, provided that the energies of the endpoints are finite.

Lemma 5.2.4 (Energy of the interpolant.) *Let $\mu_0 = u_0 dx$ and $\mu_1 = u dx$ be two smooth measures with $E(\mu_0), E(\mu_1) < c < +\infty$ and $u_0, u_1 > m > 0$. Then there are constants $\hat{c} < +\infty$ and $\hat{m} > 0$ depending only on c and m such that $E(\mu_s) < \hat{c}$ and $u_s > \hat{m}$ along the geodesic μ_s connecting μ_0 to μ_1.*

Proof. By $C^{0,1/2}$ continuity of the densities, there exists $M = M(c)$ such that $u_0(x), u_1(x) < M$ for all $x \in S^1$. Let $T : S^1 \longrightarrow S^1$ be the the optimal transport map between u_0 and u_1. By Monge–Ampère equation (2.3) we have

$$|T'(x)| \leqslant \frac{M}{m}$$

By taking the derivative of Monge–Ampère equation (2.3) we have

$$\begin{aligned} |T''(x)| &= \left|\frac{u_0'(x)u_1(T(x)) - u_0(x)T'(x)u_1'(T(x))}{u_1(x)^2}\right| \\ &\leqslant \frac{M}{m^2}|u_0'(x)| + \frac{M^2}{m^3}|u_1'(x)|. \end{aligned} \tag{5.11}$$

Now let μ_s be the geodesic connecting μ_0 to μ_1. We have

$$u_s((1-s)x + sT(x)) = \frac{u_0(x)}{(1-s) + sT'(x)}. \tag{5.12}$$

Plugging in bounds on T' yields

$$u_s(x) \geqslant \frac{m^2}{M}.$$

Hence $u_s > \hat{m}$ where $\hat{m} = \hat{m}(c, m)$. Taking derivative of the equation (5.12) we have

$$u_s'((1-s)x + sT(x)) = \frac{u_0'(x)[(1-s) + sT'(x)] - su_0(x)T''(x)}{((1-s) + sT'(x))^3}$$

Using the bounds on T' and (5.11) we have:

$$\begin{aligned} |u_s'((1-s) + sT'(x))| &\leqslant \frac{|u_0'(x)|}{|(1-s) + sT'(x)|^2} + \frac{|u_0(x)||T''(x)|}{|(1-s) + sT'(x)|^3} \\ &\leqslant (\frac{M}{m})^2|u_0'(x)| + (\frac{M}{m})^5|u_0'(x)| + (\frac{M}{m})^6|u_1'(x)|. \end{aligned}$$

Taking integral from both sides yields $E(u_s) < \hat{c}$ where $\hat{c} = \hat{c}(c, m)$. □

The idea of the following lemma was suggested by my supervisor Almut Burchard. This lemma will be used in calculations of the second derivative of the energy in Theorem 5.2.1.

Lemma 5.2.5 (Interpolation inequality.) *For every $\alpha \in \mathbb{R}$ there exists a constant $\lambda < 0$ such that*

$$|f''|_{L^2} - \alpha|f'^2|_\infty - \lambda|f|_{L^2} \geqslant 0 \quad \forall f \in C^\infty(S^1).$$

Proof. Consider the Fourier expansion $f(x) = \sum_{k\in\mathbb{Z}} a_k e^{i2k\pi x}$. We have

$$|f'|_\infty = \sup_{x\in S^1} |\sum_{k\in\mathbb{Z}} i2k\pi a_k e^{i2k\pi x}| \leqslant 2\pi \sum_{k\in\mathbb{Z}} |ka_k| = 2\pi \sum_{k\in\mathbb{Z}} |k^4 a_k^2|^{\frac{4}{10}} |a_k^2|^{\frac{1}{10}} |\frac{1}{k}|^{\frac{6}{10}}.$$

Hölder's inequality with exponents $\frac{4}{10}$, $\frac{1}{10}$, and $\frac{5}{10}$ yields

$$|f'|_\infty \leqslant 2\pi \left(\sum_{k\in\mathbb{Z}} (|k^4 a_k^2|\right)^{\frac{4}{10}} \left(\sum_{k\in\mathbb{Z}} |a_k^2|\right)^{\frac{1}{10}} \left(\sum_{k\in\mathbb{Z}} |\frac{1}{k}|^{\frac{6}{5}}\right)^{\frac{1}{2}}.$$

The term $2\pi(\sum_{k\in\mathbb{Z}} |\frac{1}{k}|^{\frac{6}{5}})^{1/2} = d$ is a constant independent of a_k. Therefore

$$|f'|_\infty \leqslant d|f''|_{L^2}^{4/5} |f|_{L^2}^{1/5}.$$

By the arithmetic-geometric inequality for a constant β we have

$$\begin{aligned} |f'|_\infty &\leqslant d|f''|_{L^2}^{4/5} |f|_{L^2}^{1/5} \\ &= d(\beta^{5/4}|f''|_{L^2})^{4/5} (\beta^{-5}|f|_{L^2})^{1/5} \\ &\leqslant \frac{4d}{5}\beta^{5/4}|f''|_{L^2} + \frac{d}{5}\beta^{-5}|f|_{L^2}. \end{aligned}$$

Putting $\beta = (\frac{5}{4}\alpha d)^{\frac{-4}{5}}$ and $\lambda = -\dfrac{d\alpha\beta^{-5}}{5}$ yields

$$\alpha|f'^2|_\infty \leqslant |f''|_{L^2} - \lambda|f|_{L^2}.$$

□

We are now ready to prove the main theorem of this section which shows that the Dirichlet energy is relaxed λ-convex at positive measures.

Theorem 5.2.6 (relaxed λ-convexity of the Dirichlet energy.) *Let $\mu = udx$ be a measure with $E(u) < c < +\infty$ and $u > m > 0$. Then $\exists\lambda = \lambda_{c,m}$ such that E is relaxed λ-convex at μ.*

Proof. We first claim that the second derivative of the energy at a positive measure ν is uniformly bounded from below along any smooth vector field. Let $\nu = vdx$ be a measure with $E(v) < c$ and $v > m$. Let ν_1 be another smooth measure and let f be the vector field defining the geodesic $\nu_s = (Id + sf)_{\#}\nu$ that connects ν to ν_1. By (5.4) we have

$$\left.\frac{d^2E(\nu_s)}{ds^2}\right|_{s=0} = 2\int_{S^1}(f''v)^2 + 8(vf'')(v'f') + 6(v'f')^2dx.$$

Recall that $W_2^2(\nu,\nu_1) = \int_{S^1} f(x)^2v(x)dx$. By (5.3) the energy is λ-convex at v, if for all such vector fields

$$2\int_{S^1}(vf'')^2 + 8(vf'')(v'f') + 6(v'f')^2dx - \lambda\int_{S^1} vf^2dx \geqslant 0.$$

By completing the squares we have

$$2\int_{S^1}\{(vf'')^2+8(vf'')(v'f')+6(v'f')^2\}dx-\lambda\int_{S^1} vf^2dx \geqslant \int_{S^1}\{f''^2v^2-52f'^2v'^2\}dx-\lambda\int_{S^1} vf^2dx.$$

The lower bound on the density $v > m$ yields

$$\int_{S^1}\{f''^2v^2 - 52f'^2v'^2\}dx - \lambda\int_{S^1} vf^2dx \geqslant m^2\int_{s^1} f''^2dx - 52\int_{s^1} f'^2u'^2dx - m\lambda\int_{s^1} f^2dx.$$

Hölder's inequality and energy bound $E(u) < c$ imply

$$m^2\int_{s^1} f''^2dx - 52\int_{s^1} f'^2u'^2dx - m\lambda\int_{s^1} f^2dx \geqslant m^2\int_{s^1} f''^2dx - 52c|f'^2|_\infty - m\lambda\int_{s^1} f^2dx.$$

By reordering and absorbing the constants in λ, the energy is λ-convex along ν_s at ν if $\forall f \in C^\infty(S^1)$ we have

$$|f''|_{L^2}^2 - \alpha|f'|_\infty^2 - \lambda|f|_{L^2}^2 \geqslant 0 \tag{5.13}$$

where $\alpha = \frac{52c}{m^2}$. By Lemma 5.2.5 the claim has been proved.

Now consider the energy sub-level set E_c. By Theorem 5.2.1 Wasserstein convergence implies uniform convergence on E_c. Therefore there exists a $\delta = \delta_c$ such that we have

$v > m$ for all $\nu = vdx \in E_c \cap B_\delta(\mu)$. Assume that $\nu_0, \nu_1 \in E_c \cap B_\delta(\nu)$. Let ν_s be the geodesic connecting ν_0 to ν_1. By Lemma 5.2.4 there exist $\hat{m}$ and $\hat{c}$ depending only on c and m such that $E(\nu_s) < \hat{c} < +\infty$ and $v_s > \hat{m} > 0$. By the argument at the beginning of the proof there exists a $\hat{\lambda} = \hat{\lambda}_{m,c}$ such that E is $\hat{\lambda}$-convex along the geodesic ν_s. The constant $\hat{\lambda}$ is uniform for all pairs of smooth measures inside $E_c \cap B_\delta(\mu)$. Therefore, by Lemma 5.2.3 E is relaxed $\hat{\lambda}$-convex at μ. □

Corollary 5.2.7 *The gradient flow trajectory of the Dirichlet energy on S^1 with a positive initial data exists and is unique at least for a short period of time.*

Corollary 5.2.8 *The positive periodic solutions of the thin-film equation $\partial_t u = -\partial_x(u\partial_x^3 u)$ are locally well-posed.*

Chapter 6

Other Equations of Fourth and Higher Order

In this chapter, we show that the theory developed in the last two chapters can be applied to a wide class of energy functionals and evolution equations. Note that the result of Theorem 4.2.2 is general and it can be applied to any energy functional, provided that it is relaxed λ-convex. The corresponding lemmas from Chapter 5 for the energies studied here can be derived in a similar fashion with minor modifications. Hence, we discuss the proofs only briefly.

6.1 Higher Order Equations

The family that we study here is of the form $E(u) = \frac{1}{2}\int_{S^1} |u^{(k)}|^2 dx$ for $k \in \mathbb{N}$. The flow of this family of energies corresponds to the solution of the higher order non-linear equations of the form $\partial_t u = (-1)^k \partial_x (u \partial_x^{2k+1} u)$.

Consider $u \in D(E)$. Finiteness of $|u|_{H^k}$ in particular implies that the H^1-norm of u is bounded. Since we only used the H^1-norm bounds in Lemmas 5.2.1, 5.2.2, and 5.2.3, they automatically follow for this class of energies. Therefore, Wasserstein and uniform convergence are equivalent on energy sub-level sets, E is lower semicontinuous, and one can use approximation by smooth functions to study convexity.

In Lemma 5.2.4 we derived bounds on T'' by taking derivatives of the explicit formula of T' given by the Monge–Ampère equation. In the same fashion, one can find bounds on higher derivatives of the optimal map by taking more derivatives of the Monge–Ampère equation. For generalization of Lemma 5.2.5, we will have to show that $\forall\alpha\ \exists\lambda$ such that

$$|f^{(m+1)}|_{L^2}^2 - \alpha|f^{(m)}|_\infty^2 - \lambda|f|_{L^2}^2 \geqslant 0$$

for every smooth vector field f. We will now use induction. Assume that for any $\alpha_m > 0$ there exists $\lambda_m \leqslant 0$ such that

$$|f^{(m)}|_\infty^2 \leqslant \frac{1}{\alpha_m}|f^{(m+1)}|_{L^2}^2 - \lambda_m|f|_{L^2}^2. \tag{6.1}$$

Let $\alpha_{m+1} > 0$ be given. By applying Lemma 5.2.5 to $f^{(m+1)}$, there exists a $\hat{\lambda} \leqslant 0$ such that

$$|f^{(m+1)}|_\infty^2 \leqslant \frac{1}{2\alpha_{m+1}}|f^{(m+2)}|_{L^2}^2 - \hat{\lambda}|f^{(m)}|_{L^2}^2. \tag{6.2}$$

Put $\alpha_m = -2\hat{\lambda}$, by (6.1) there exists $\lambda_m \leqslant 0$ such that

$$|f^{(m)}|_\infty^2 \leqslant \frac{-1}{2\hat{\lambda}}|f^{(m+1)}|_{L^2}^2 - \lambda_m|f|_{L^2}^2.$$

$|f^{(m)}|_{L^2}^2 \leqslant |f^{(m)}|_\infty^2$ on $\mathbb{R}/\mathbb{Z}$. Therefore

$$-\hat{\lambda}|f^{(m)}|_{L^2}^2 \leqslant \frac{1}{2}|f^{(m+1)}|_{L^2}^2 + \hat{\lambda}\lambda_m|f|_{L^2}^2.$$

Plugging into (6.2) yields

$$|f^{(m+1)}|_\infty^2 \leqslant \frac{1}{2\alpha_{m+1}}|f^{(m+2)}|_{L^2}^2 + \frac{1}{2}|f^{(m+1)}|_{L^2}^2 + \hat{\lambda}\lambda_m|f|_{L^2}^2.$$

Therefore

$$|f^{(m+1)}|_\infty^2 - \frac{1}{2}|f^{(m+1)}|_{L^2}^2 \leqslant \frac{1}{2\alpha_{m+1}}|f^{(m+2)}|_{L^2}^2 + \hat{\lambda}\lambda_m|f|_{L^2}^2.$$

By $|f^{(m+1)}|_{L^2}^2 \leqslant |f^{(m+1)}|_\infty^2$ and by setting $\lambda_{m+1} = \frac{-1}{2}\hat{\lambda}\lambda_m$, we have

$$\forall\alpha_{m+1} > 0 \;\; \exists\lambda_{m+1} \leqslant 0 \quad s.t. \quad |f^{(m+1)}|_\infty^2 \geqslant \frac{1}{\alpha_{m+1}}|f^{(m+2)}|_{L^2}^2 - \lambda_{m+1}|f|_{L^2}^2. \tag{6.3}$$

In conclusion, all the lemmas in the previous section can be applied to higher order energies. We now study convexity of the energies along smooth vector fields on a measure $\mu = udx$ with positive density $u > m$ and finite energy $E(u) < c < \infty$.

$$\begin{aligned}\frac{d^2}{ds^2}|_{s=0}E(u_s) &= \frac{d^2}{ds^2}|_{s=0}\int_{S^1}(\partial_y^k u_s(y))^2 dy \\ &= \frac{d^2}{ds^2}|_{s=0}\int_{S^1}\{(\frac{\partial x}{\partial y}\frac{\partial}{\partial x})^k\frac{u(x)}{1+sf'(x)}\}^2\frac{\partial y}{\partial x}dx.\end{aligned}$$

Since we study all the different orders at the same time, we consider the general form given by a polynomial P which is determined by the order of the energy. We have

$$\frac{d^2}{ds^2}|_{s=0}E(u_s) = \int_{S^1}|uf^{(k)}|^2 + P(u, u^{(1)}, ..., u^{(k)}; f, f^{(1)}, ..., f^{(k-1)})dx$$

where P is of order at most 2 with respect to each of its entries, and the order of the derivative of each term in P is at most k. At a measure with positive density and finite energy, we have $u > m$ and $|u^{(i)}|_\infty < M$ for all $i < k$ where M depends only on $E(u)$. Also $|f^{(i)}|_{L^2} \leqslant |f^{(k-1)}|_\infty$ for all $i < k-1$. Therefore similar to the calculation of Theorem 4.2.2, for positive constants β_1, β_2 we have

$$\frac{d^2}{ds^2}|_{s=0}E(u_s) \geqslant \beta_1|f^{(k)}|_{L^2}^2 - \beta_2|f^{(k-1)}|_\infty^2.$$

Therefore E is relaxed convex at u if we can find λ such that

$$\beta_1|f^{(k)}|_{L^2}^2 - \beta_2|f^{(k-1)}|_\infty^2 \geqslant \lambda \int_{S^1} uf^2 dx.$$

This implies that the energy is relaxed λ-convex at u because by (6.3) for $\alpha = \alpha_{m,c}$ there exists λ such that

$$|f^{(k)}|_{L^2}^2 - \alpha|f^{(k-1)}|_\infty^2 - \lambda|f|_{L^2}^2 \geqslant 0 \quad \forall f \in C^\infty(S^1)$$

Hence we have proved the following theorem.

Theorem 6.1.1 *The energies of the form*

$$E(u) = \begin{cases} \int_{S^1} |\partial_x^k u(x)|^2 dx & \mu = udx, \; u \in H^k(S^1), \\ +\infty & \textit{else.} \end{cases}$$

are relaxed λ-convex on the positive measures with finite energy. In particular, periodic gradient flow solutions of

$$\partial_t u = (-1)^k \partial_x (u \partial_x^{2k+1} u)$$

with positive initial data exist and are unique for a short time.

6.2 Other Equations of Fourth Order

Consider the energies of the form $E(u) = \int_{S^1} g(u, \partial_x u) dx$. We start by calculating the second derivative of the energy along a geodesic induced by a vector field $f \in C^\infty(S^1)$.

$$\begin{aligned} \frac{d^2}{ds^2}|_{s=0} E(u_s) &= \frac{d^2}{ds^2}|_{s=0} \int_{S^1} g(u_s(y), \partial_y u_s(y)) dy \\ &= \frac{d^2}{ds^2}|_{s=0} \int_{S^1} g\left(\frac{u(x)}{1+sf'(x)}, \frac{\partial x}{\partial y}\frac{\partial}{\partial x}(\frac{u(x)}{1+sf'(x)})\right)\frac{\partial y}{\partial x} dx \\ &= \frac{d^2}{ds^2}|_{s=0} \int_{S^1} g\left(\frac{u(x)}{1+sf'(x)}, \frac{1}{1+sf'(x)}\partial_x(\frac{u(x)}{1+sf'(x)})\right)(1+sf'(x)) dx \end{aligned}$$

where we used the change of variable $y = x + sf(x)$. Therefore we have

$$\frac{d^2}{ds^2}|_{s=0}E(\mu_s) = \int_{s^1} \begin{bmatrix} f' & f'' \end{bmatrix} A \begin{bmatrix} f' \\ f'' \end{bmatrix} dx \tag{6.4}$$

where the matrix A is given by

$$A = \begin{bmatrix} 2u'g^{(0,1)} + 4u'^2g^{(0,2)} + 4uu'g^{(1,1)} + u^2g^{(2,0)} & 2ug^{(0,1)} + 2uu'g^{(0,2)} + u^2g^{(1,1)} \\ 2ug^{(0,1)} + 2uu'g^{(0,2)} + u^2g^{(1,1)} & u^2g^{(0,2)} \end{bmatrix}$$

Note that if A is positive definite, then the energy is convex. We study the class of the form

$$g(u, \partial_x u) = |\partial_x(u^a)|^2 \quad a > 0.$$

Finiteness of the energy implies that u^a is $C^{0,\frac{1}{2}}$ continuous with modulus of continuity smaller than the energy. Because $a > 0$, u is continuous and since $\int_{S^1} u dx = 1$, there exists a point x_0 with $u(x_0) = 1$. Without loss of generality we assume $x_0 = 0$. We have

$$|u(x)^a - 1| \leqslant \sqrt{c|x|} \;\Rightarrow\; u(x) \leqslant (1 + \sqrt{c|x|})^{\frac{1}{a}}.$$

Therefore there exists a uniform $M < \infty$ such that $u < M$ for all $u \in E_c$. We now briefly discuss the corresponding lemmas from Chapter 5.

Equivalence of Wasserstein and uniform convergence on energy sub-level sets. Let $u_2(x_0) - u_1(x_0) > h$. Then we have $u_2^a(x) \geqslant u_2^a(x_0) - \sqrt{c|x - x_0|}$ and $u_1^a(x) \leqslant u_1^a(x_0) + \sqrt{c|x - x_0|}$. Therefore the star-like shape in Lemma 5.2.1 should be replaced by a modified version, given by $(u_2^a(x_0) - \sqrt{c|x - x_0|})^{\frac{1}{a}}$ and $(u_1^a(x_0) + \sqrt{c|x - x_0|})^{\frac{1}{a}}$, and the rest of the proof goes similarly. Hence, we have equivalence of the Wasserstein and uniform convergence on the energy sub-level sets.

Lower semicontinuity and smooth approximation. Having Lemma 5.2.1 for this class of energies, the proof of Lemmas 5.2.2 and 5.2.3 can be repeated by replacing u with u^a. Hence the energy $E(u) = \int_{S^1} |\partial_x u^a| dx$ is lower semicontinuous and one can use approximation by smooth functions.

Energy of the interpolant. Let u be bounded away from zero $u > m > 0$. When $a \geqslant 1$

$$m^{2(a-1)} \int_{S^1} |\partial_x u|^2 dx \;\leqslant\; E(u) = \int_{S^1} u^{2(a-1)} |\partial_x u|^2 dx \;\leqslant\; M^{2(a-1)} \int_{S^1} |\partial_x u|^2 dx \qquad (6.5)$$

and when $0 < a < 1$

$$M^{2(a-1)} \int_{S^1} |\partial_x u|^2 dx \;\leqslant\; E(u) = \int_{S^1} u^{2(a-1)} |\partial_x u|^2 dx \;\leqslant\; m^{2(a-1)} \int_{S^1} |\partial_x u|^2 dx. \qquad (6.6)$$

By equivalence of Wasserstein and uniform convergence, there exists δ such that $v > m$ for all $v \in B_\delta(u) \cap E_c$. Also we have proved that $v < M$ for all $v \in E_c$. Therefore we can refer to the calculation for the Dirichlet energy and just compare the energy of the geodesic with the corresponding Dirichlet energy using (6.5) and (6.6) to find a bound on the energy of interpolate points along a geodesic.

In conclusion, all of the required lemmas are true. By (6.4), along a geodesic induced by a smooth vector field f we have

$$\begin{aligned} \frac{d^2}{ds^2}|_{s=0} E(u_s) &= \int_{S^1} 2a^2 u^{2(a-1)} \left((uf'')^2 + 4(1+a)(uf'')(u'f') + (1+a)(1+2a)(u'f')^2 \right) dx \\ &\geqslant \int_{S^1} \alpha_1 (u^a f'')^2 - \alpha_2 ((u^a)' f')^2 dx \end{aligned}$$

for some constants α_1, α_2. Similar to (5.13), we have

$$\frac{d^2}{ds^2}|_{s=0} E(u_s) - m\lambda \int_{S^1} u f^2 dx \;\geqslant\; |f''|^2_{L^2} - \alpha |f'|^2_\infty - \lambda |f|^2_{L^2} \;\geqslant\; 0$$

where the last inequality follows from Lemma 5.2.5. We have proved the following theorem.

Theorem 6.2.1 *For every $a > 0$*

$$E(u) = \begin{cases} \int_{S^1} |\partial_x u(x)^a|^2 dx & \mu = udx,\ u \in H^1(S^1) \\ +\infty & else. \end{cases}$$

is relaxed λ-convex on positive measures with finite energy. In particular, periodic gradient flow solutions of

$$\partial_t u = -2a\partial_x(u\partial_x(u^{a-1}\partial_x^2 u^a))$$

with positive initial data exist and are unique for a short time.

An interesting example is the *Fisher Information*

$$E(u) = \frac{1}{2}\int_{S^1} |\partial_x u(x)^{\frac{1}{2}}|^2 dx$$

which corresponds to the *quantum drift diffusion equation* [7]

$$\partial_t u = -\partial_x(u\partial_x \frac{\partial_x^2\sqrt{u}}{\sqrt{u}}).$$

Therefore we have local well-posedness of periodic solutions of the quantum drift diffusion equation with positive initial data.

Another interesting case is the limiting case $a = 0$. The corresponding energy can be written as $E(u) = \frac{1}{2}\int_{S^1} |\partial_x \log u|^2 dx$. Finiteness of the energy result in $C^{0,\frac{1}{2}}$ continuity of $\log u$. All of the lemmas can be repeated in a similar fashion for this energy. Furthermore, finiteness of the energy implies a lower bound for the measure because

$$|\log u(x) - \log(1)| \leqslant \sqrt{c|x|} \implies e^{-\sqrt{c}} \leqslant u(x) \leqslant e^{\sqrt{c}}. \tag{6.7}$$

Therefore positivity is preserved along the flow. By (6.4) we have

$$\frac{d^2}{ds^2}|_{s=0}E(\mu_s) = \int_{S^1} 2(f'')^2 + 4(f'')(\frac{u'}{u}f') - 2(\frac{u'}{u}f')^2 dx.$$

By 5.2.5 there exists λ such that

$$\frac{d^2}{ds^2}|_{s=0}E(\mu_s) \geqslant |f''|_{L^2}^2 - 6ce^{\sqrt{c}}|f'|_\infty^2 - \lambda|f|_{L^2} \geqslant 0.$$

Hence E is relaxed λ convex at $u \in E_c$. Furthermore, since there is a uniform lower bound $e^{-\sqrt{c}}$ for all $v \in E_c$, the constant λ is uniformly bounded along the flow. Therefore, the gradient flow is globally well-posed and we have the following theorem.

Theorem 6.2.2 *Wasserstein gradient flow of the energy*

$$E(u) = \begin{cases} \frac{1}{2}\int_{S^1} |\partial_x \log u|^2 dx & \mu = udx,\ u \in H^1(S^1) \\ +\infty & \textit{else.} \end{cases}$$

is globally well-posed. Hence the equation

$$\partial_t u = -\partial_x(u\partial_x^2 \frac{\partial_x u}{u^2})$$

with periodic boundary condition is well-posed.

Remarks. There are several directions to extend the developed method to other classes of energies and equations. As a simple application, one can construct other classes of relaxed λ-convex functionals by combining the ones already studied. For example, the gradient flow of the energy $E(u) = \int_{S^1}\{|\partial_x u|^2 + \epsilon\frac{1}{u^2}\}dx$, which is the Dirichlet energy with a perturbation, is globally well-posed. The reason is that the second term forces the solution to remain positive. One more challenging problem is the analysis of equations in higher dimensions. Our method is utilizing Sobolev embedding theorem on energy sub-level sets which is getting weaker on higher dimensions. An interesting question is whether it is possible to solve this problem with studying higher order energies. Another interesting question is to study wellposedness of gradient flows on the Wasserstein space with mobility. The difficulty here is that one needs to find the form of the geodesics in the Wasserstein space with mobility at least for close-by points.

Bibliography

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[2] Luigi Ambrosio, Giuseppe Savaré, and Lorenzo Zambotti. Existence and stability for Fokker-Planck equations with log-concave reference measure. *Probab. Theory Related Fields*, 145(3-4):517–564, 2009.

[3] Sebastian Andres and Max-K. von Renesse. Particle approximation of the Wasserstein diffusion. *J. Funct. Anal.*, 258(11):3879–3905, 2010.

[4] Jean-David Benamou and Yann Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numer. Math.*, 84(3):375–393, 2000.

[5] A. L. Bertozzi and M. Pugh. The lubrication approximation for thin viscous films: regularity and long-time behavior of weak solutions. *Comm. Pure Appl. Math.*, 49(2):85–123, 1996.

[6] Adrien Blanchet, Vincent Calvez, and José A. Carrillo. Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model. *SIAM J. Numer. Anal.*, 46(2):691–721, 2008.

[7] L. L. Bonilla and R. Escobedo. Wigner-Poisson and nonlocal drift-diffusion model equations for semiconductor superlattices. *eprint arXiv:cond-mat/0503109*, March 2005.

[8] Yann Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44(4):375–417, 1991.

[9] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Publishing Co., Amsterdam, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).

[10] E. A. Carlen and W. Gangbo. Constrained steepest descent in the 2-Wasserstein metric. *Ann. of Math. (2)*, 157(3):807–846, 2003.

[11] José A. Carrillo, Robert J. McCann, and Cédric Villani. Contractions in the 2-Wasserstein length space and thermalization of granular media. *Arch. Ration. Mech. Anal.*, 179(2):217–263, 2006.

[12] José A. Carrillo and Dejan Slepčev. Example of a displacement convex functional of first order. *Calc. Var. Partial Differential Equations*, 36(4):547–564, 2009.

[13] Fausto Cavalli and Giovanni Naldi. A Wasserstein approach to the numerical solution of the one-dimensional Cahn-Hilliard equation. *Kinet. Relat. Models*, 3(1):123–142, 2010.

[14] Ennio De Giorgi. New problems on minimizing movements. In *Boundary value problems for partial differential equations and applications*, volume 29 of *RMA Res. Notes Appl. Math.*, pages 81–98. Masson, Paris, 1993.

[15] Bertram Düring, Daniel Matthes, and Josipa Pina Milišić. A gradient flow scheme for nonlinear fourth order equations. *Discrete Contin. Dyn. Syst. Ser. B*, 14(3):935–959, 2010.

[16] Lawrence C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[17] Alessio Figalli, Young-Heon Kim, and Robert McCann. Regularity of optimal transport maps on multiple products of spheres. *To appear in J. Eur. Math. Soc. (JEMS).*

[18] Richard Jordan, David Kinderlehrer, and Felix Otto. The variational formulation of the Fokker-Planck equation. *SIAM J. Math. Anal.*, 29(1), 1998.

[19] David Kinderlehrer and Noel J. Walkington. Approximation of parabolic equations using the Wasserstein metric. *M2AN Math. Model. Numer. Anal.*, 33(4):837–852, 1999.

[20] Grégoire Loeper. Regularity of optimal maps on the sphere: the quadratic cost and the reflector antenna. *Arch. Ration. Mech. Anal.*, 199(1):269–289, 2011.

[21] Xi-Nan Ma, Neil S. Trudinger, and Xu-Jia Wang. Regularity of potential functions of the optimal transportation problem. *Arch. Ration. Mech. Anal.*, 177(2):151–183, 2005.

[22] Robert J. McCann. Polar factorization of maps on Riemannian manifolds. *Geom. Funct. Anal.*, 11(3):589–608, 2001.

[23] Robert John McCann. *A convexity theory for interacting gases and equilibrium crystals*. ProQuest LLC, Ann Arbor, MI, 1994. Thesis (Ph.D.)–Princeton University.

[24] Luca Natile and Giuseppe Savaré. A Wasserstein approach to the one-dimensional sticky particle system. *SIAM J. Math. Anal.*, 41(4):1340–1365, 2009.

[25] Felix Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.

[26] Cédric Villani. *Topics in optimal transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[27] Cédric Villani. *Optimal transport, Old and New*, volume 338 of *Grundlehren der Mathematischen Wissenschaften*. Springer-Verlag, Berlin, 2009.