JOINT RESEARCH DOCTORAL THESIS
IN MATHEMATICS

Dipartimento di Matematica
*Guido Castelnuovo*
Università di Roma SAPIENZA
Piazzale Aldo Moro 5, 00185 Roma
Italia

Département de Mathématiques
*Centre de Mathématiques Laurent Schwartz*
ÉCOLE POLYTECHNIQUE
91128 Palaiseau cedex
France

# Dynamics of large particle systems

Mikaela Iacobelli

*Defended in Rome, the 21th of December, 2015.*

Doctoral committee:
Emanuele Caglioti, Università di Roma Sapienza, adviser
François Golse, École Polytechnique, adviser
Clément Mouhot, University of Cambridge
Thierry Paul, École Polytechnique
Mario Pulvirenti, Università di Roma Sapienza
Giuseppe Savaré, Università di Pavia

# Abstract

This thesis is devoted to the mathematical study of problems involving the general question of the interaction between continuous and discrete objects. More precisely, one wants to answer to the following questions: how to approximate a probability measure by a finite support measure? How to represent the evolution of a system of interacting particles as the number of particles goes to infinity?

First, we focus on the problem of *quantization of measures*. Given a continuous density, such a problem consists in finding an optimal way to approximate it by a finite distribution of points. The idea is to start from a random distribution and define a dynamic that makes the particles evolve until they reach the optimal position. From a mathematical point of view, finding such a dynamic and showing convergence towards the optimal configuration leads to many difficulties. We solve this problem in the one-dimensional case under suitable regularity hypotheses on the density we want to approximate.

Then, we study the quantization problem on Riemannian manifolds. Under some global assumption on the behaviour of the measure at "infinity" we estimate the quantization error. This generalizes the known results in the flat case. Our new growth assumption depends on the curvature of the manifold and reduces, in the flat case, to a moment condition. We also provided a counterexample showing that such hypothesis is sharp.

In a second time, we are concerned with the mathematical study of some aspects of the Vlasov-Poisson equation, that represents a classical kinetic model in plasma physics. At large spatial and time scales, plasmas have the tendency to be quasineutral, *i.e.* the local charge disappears. On the other hand, at small spatial and time scales, the quasineutrality is not longer verified. The typical degeneracy scales are the oscillation frequency of the electrons and the Debye length. The Debye length is the distance at which electrons screen out electric fields and it is often very small compared to the spatial

observation length.

We study the quasineutral limit, *i.e.* the limit as the Debye length tends to zero, for the Vlasov-Poisson equation with massless electrons. Such a model seems to be more appropriate to describe plasmas created in laboratory, for whom the global neutrality is always verified. The study of these singular limits leads to interesting problems, both in physics and in mathematics. We study the hydrodynamic limit for the Vlasov-Poisson system for ions towards a system of Euler compressible type, in a regime where the Debye length and the temperature are small. Using a family of distances of Wasserstein-type on probability measures, we obtain some convergence and stability estimates that generalise and improve some previous results, both in dimension one and in higher dimensions.

# Contents

**Introduction** **11**

0.1 Quantization of Measures. . . . . . 12

0.1.1 From the Lagrangian to the Eulerian setting. . . . . . 15

0.1.2 What happens on Riemannian manifolds? . . . . . . 16

0.2 Quasineutral limit for the Vlasov-Poisson equation. . . . . . 17

## I Quantization of Measures 21

**1 Transportation theory and its relation to quantization of measures** **23**

1.1 The optimal transport problem . . . . . . 23

1.1.1 Kantorovich's problem . . . . . . 24

1.1.2 Monge problem . . . . . . 26

1.2 Optimal transport metrics . . . . . . 27

1.2.1 Wasserstein distances and $L^p$ norms . . . . . . 29

1.2.2 Kantorovich duality . . . . . . 30

1.2.3 Optimal transport metrics and quantization of measures . . . . . 31

**2 A gradient flow approach to quantization of measures** **35**

2.1 Introduction . . . . . . 35

2.2 Preliminary results . . . . . . 43

2.2.1 The discrete gradient flow . . . . . . 43

2.2.2 The continuous gradient flow . . . . . . 43

2.3 The case $\rho \equiv 1$ . . . . . . 46

2.3.1 The $L^2$ estimate in the continuous case . . . . . . 46

2.3.2 Convergence of the gradient flows . . . . . . 47

2.4 The case $\rho \not\equiv 1$ . . . 55
2.4.1 Convergence of the gradient flows . . . 55
2.4.2 The convergence results . . . 70
2.5 From the discrete to the continuous case . . . 77
2.6 The Hessian of $\mathcal{F}[X]$ . . . 78
2.6.1 Convexity under a smallness assumption on $\rho'$ and $\rho''$ . . . 78
2.6.2 Lack of convexity without a smallness assumption . . . 79

**3 Asymptotic quantization for probability measures on Riemannian manifolds 81**
3.1 Introduction . . . 81
3.2 Proof of Theorem 3.1.4: the compact case . . . 86
3.2.1 Localization argument . . . 87
3.2.2 The local quantization error . . . 88
3.2.3 Upper bound for $V_{N,r}$ . . . 93
3.2.4 Lower bound for $V_{N,r}$ . . . 94
3.2.5 Approximation argument: general compactly supported measures 97
3.3 Proof of Theorem 3.1.4: the non-compact case . . . 98
3.3.1 Proof of Theorem 3.3.1 . . . 98
3.4 Proof of Theorem 3.1.7 . . . 102

**II Quasineutral limit for Vlasov-Poisson 107**

**4 The Vlasov-Poisson equation 109**
4.1 Mean field limit . . . 109
4.1.1 Newton equations . . . 109
4.1.2 Mean field limit of the $N$ particle system . . . 111
4.2 Quantitative stability via Wasserstein type distances . . . 113
4.3 The Vlasov-Poisson equation . . . 118
4.3.1 Basic properties of the Vlasov equation . . . 119
4.4 Quasineutral limit . . . 127
4.4.1 What is a plasma? . . . 127
4.4.2 The Vlasov-Poisson equation for massless electrons . . . 128
4.4.3 The Debye length . . . 130

**5 The quasineutral limit of the Vlasov-Poisson equation in Wasserstein metric 133**

5.1 Introduction . . . . . . . . . . 133

5.2 Weak-strong stability for the VP system with massless electrons: proof of Theorem 5.1.9 . . . . . . . . . . 141

5.2.1 A scaling argument . . . . . . . . . . 142

5.2.2 Weak-strong estimate for the rescaled system . . . . . . . . . . 143

5.2.3 Back to the original system and conclusion of the proof . . . . . . 151

5.3 Proof of Theorem 5.1.1 . . . . . . . . . . 153

5.4 Proofs of Theorems 5.1.3 and 5.1.4 . . . . . . . . . . 156

5.4.1 The fluid point of view and convergence for uniformly analytic initial data . . . . . . . . . . 156

5.4.2 End of the proof of Theorem 5.1.3 and Theorem 5.1.4 . . . . . . . 162

5.5 Proof of Proposition 5.1.6 . . . . . . . . . . 163

5.6 Proof of Proposition 5.1.7 . . . . . . . . . . 165

**6 Quasineutral limit for Vlasov-Poisson via Wasserstein stability estimates in higher dimension 171**

6.1 Introduction . . . . . . . . . . 171

6.2 Overview . . . . . . . . . . 177

6.3 Proofs of Steps 1 and 2 . . . . . . . . . . 178

6.3.1 $W_2$ stability estimate . . . . . . . . . . 178

6.3.2 Control of the growth of the support in velocity in 2D . . . . . . . 182

6.3.3 Control of the growth of the support in velocity in 3D using Batt and Rein's estimates . . . . . . . . . . 188

6.4 Proof of Theorem 6.1.3 . . . . . . . . . . 191

**A Jacobi fields 195**

**Bibliography 197**

# Acknowledgments, remerciements, ringraziamenti

This thesis has been developed in a joint program between Rome and Paris, so I would like to write these lines in three different languages, according to the person the acknowledgments are addressed to.

First and foremost, I would like to thank my advisers, Emanuele Caglioti and François Golse, whose enlightening questions -although not fully answered- drove me toward a better understanding of my subject. I am very grateful to both of you for proposing me a research line that perfectly suits me and which I am extremely passionate about.

Grazie Emanuele per la fiducia che mi hai accordato fin dai tempi della tesi specialistica, lasciandomi scegliere in autonomia come sviluppare la mia ricerca e aiutandomi a realizzare il mio desiderio di una tesi in cotutela. E' stato per me istruttivo e stimolante vedere il mio problema di dottorato attraverso il tuo intuito fisico.

Merci François pour tout ce que tu as fait pour moi, en partant par le soutien constant et les conseils précieux que tu m'as donnés au cours de ces années, jusqu'à l'énorme enjeu d'enseigner à quelqu'un comment faire de la recherche en mathématiques. Tu as été pour moi un vrai maître.

Ces pages sont aussi le résultat des nombreuses discussions que j'ai eu avec Daniel Han-Kwan. C'était un vrai plaisir de travailler avec lui, et je lui suis très reconnaissante pour son encouragement, sa disponibilité et pour avoir partagé avec moi ses connaissances.

I owe my gratitude to Clément Mouhot, Thierry Paul, Mario Pulvirenti, and Giuseppe Savaré for being part of my committee. In addition, I would like to give a special thanks to Clément Mouhot for several useful discussions and for his support.

I had the pleasure and benefit of conversations with many mathematicians about the topics of my thesis. Let me thank Luigi Ambrosio, Matteo Bonforte, José Antonio

Carrillo, Irene Gamba, Jan Maas, Francesco Maggi, Giuseppe Mingione, Pierre Picco, and Filippo Santambrogio.

Ripenso con gratitudine e nostalgia ai corsi magnifici che ho seguito al Dipartimento di Matematica *Guido Castelnuovo*. Vorrei ringraziare quanti rendono questo posto così vivo e scientificamente stimolante, ed anche sottolineare l'incredibile gentilezza di Linda De Riggi e Maura Santoboni.

Un grand merci à toute l'équipe du CMLS, sans laquelle ces années n'auraient pas été aussi agréables. Je voudrais également souligner la gentillesse et la disponibilité des secrétaires du CMLS, Marine Amier, Pascale Fuseau et Carole Juppin.

Grazie Marco Isopi e Andrea Sambusetti per il supporto ed i consigli sinceri. Grazie Alessandro D'Andrea, perché il tempo che mi hai dedicato ha lasciato in me un'impronta forte, scientifica ed umana; e per l'amicizia.

Grazie agli amici di sempre che riescono ad illuminare la mia vita anche da lontano. Penso con particolare intensità ad Anna e Andrea, Serena M., Chiara e Giuseppe, Serena C., Anna e Tania, Cecilia, Carmelo, Leonardo...

Et encore, merci Dorothea pour ton sourire contagieux, ed ad Aldo per avermi scosso quando ne avevo bisogno. Grazie Luca per le innumerevoli pause caffé in cui mi hai strappato un sorriso! Grazie Viv, tu ed il piccolo Luigi siete una fonte di gioia incredibile.

Un ringraziamento di tutto cuore va alla mia famiglia meravigliosa, che mi avvolge come un bosco racchiude un albero, e mi tiene ben salda quando soffia il vento.

Grazie Alessio, compagno di viaggio, perché riponi in me una fiducia illimitata, e non ti stupisci mai.

# Introduction

This thesis is devoted to the study of problems arising from the wide context of statistical mechanics. More precisely, the objective of this thesis is twofold. First, we study the problem of quantization of measures that concerns the approximation in some *optimal* way of a diffuse measure with discrete ones. Second, we focus on the quasineutral limit for the Vlasov-Poisson equation. The common root of these topics is the general task of finding a rigorous justification of the description of a large number of identical objects -typically physical particles as gas molecules or ions and electrons in a plasma- via approximate models that describe the behaviour of the "generic object" of a physical system. As a common feature, in both these problems we make extensive use of some optimal transport metrics. Moreover, this work involves tools in probability, calculus of variations, and Riemannian geometry.

The thesis is split in two parts organized as follows.

In Chapter 1 we review some basic facts on optimal transport theory and we discuss its connection to the quantization problem. Then in Chapters 2 and 3 we collect the results from our two papers [26, 65], dealing respectively with a dynamical approach to the quantization problem in one dimension, and a complete study of the quantization problem on Riemannian manifolds.

In the second part we first explain how to obtain the Vlasov equation as the mean field limit of $N$ particle systems following the Newton's law, then we briefly recall the main results about the Vlasov-Poisson equation, and we motivate the study of its quasineutral limit. Then in Chapters 5 and 6 we present the results from the papers [59, 60].

Regarding the notation, we tried to make it as much unified as possible. However, for the convenience of the reader, the main specific notation and definitions will be briefly reintroduced at the beginning of each chapter.

Before entering into the core of the thesis, in the next Sections we give an overview of all results.

## 0.1 Quantization of Measures.

In our context, the term *quantization* refers to the process of finding the *best* approximation of a $d$-dimensional probability distribution by a convex combination of a finite number $N$ of Dirac masses. This problem arises in several contexts and has applications in information theory (signal compression), numerical integration, mathematical models in economics (optimal location of service centers), and kinetic theory. For a detailed exposition and a complete list of references see [47].

To describe our results, let us introduce the setup of the problem. Fixed $r \geq 1$, consider $\mu$ a probability measure on an open set $\Omega \subset \mathbb{R}^d$. Given $N$ points $x^1, \ldots, x^N \in \Omega$, one wants to find the best approximation of $\mu$, in the Wasserstein distance $W_r$, by a convex combination of Dirac masses centered at $x^1, \ldots, x^N$. Hence one minimizes

$$\inf \left\{ W_r\bigg( \sum_i m_i \delta_{x^i}, \mu \bigg)^r \ : \ m_1, \ldots, m_N \geq 0, \ \sum_i m_i = 1 \right\},$$

with

$$W_r(\nu_1, \nu_2) := \inf \left\{ \bigg( \int_{\Omega\times\Omega} |x-y|^r d\gamma(x,y) \bigg)^{1/r} \ : \ (\pi_1)_\# \gamma = \nu_1, \ (\pi_2)_\# \gamma = \nu_2 \right\},$$

where $\gamma$ varies among all probability measures on $\Omega \times \Omega$, and $\pi_i : \Omega \times \Omega \to \Omega$ $(i = 1, 2)$ denotes the canonical projection onto the $i$-th factor.

**Remark 0.1.1.** We draw the attention on a terminology issue on optimal transport metrics: they can be found equivalently under the name of Monge-Kantorovich distances and Wasserstein distances, and we shall equivalently use either of them. In particular, one can also write the above minimization problem as

$$\inf \left\{ MK_r\bigg( \sum_i m_i \delta_{x^i}, \mu \bigg) \ : \ m_1, \ldots, m_N \geq 0, \ \sum_i m_i = 1 \right\},$$

For more details, we refer to Definition 1.2.1.

The best choice of the masses $m_i$ is explicit and can be expressed in terms of the so-called *Voronoi cells*. Also, the following identity holds (see Lemma 1.2.7):

$$\inf \left\{ W_r\bigg( \sum_i m_i \delta_{x^i}, \mu \bigg)^r \ : \ m_1, \ldots, m_N \geq 0, \ \sum_i m_i = 1 \right\} = F_{N,r}(x^1, \ldots, x^N),$$

where

$$F_{N,r}(x^1, \dots, x^N) := \int_\Omega \min_{1\le i\le N} |x^i - y|^r \, d\mu(y).$$

Hence, the main question becomes:

*Where are the "optimal points" $(x^1, \dots, x^N)$ located?*

The following result describe the asymptotic distribution of the minimizing configurations (see for instance [47]):

**Theorem 0.1.2.** *Let $\mu = h\,dx + \mu^s$ be a probability measure on $\mathbb{R}^d$ satisfying*

$$\int_{\mathbb{R}^d} |x|^{r+\delta} \, d\mu(x) < \infty, \tag{0.1.1}$$

*and let $x^1, \dots, x^N$ minimize the functional $F_{N,r} : (\mathbb{R}^d)^N \to \mathbb{R}^+$. There exists a constant $C_{r,d} > 0$ such that*

$$N^{r/d} F_{N,r}(\mu) \to C_{r,d} \left( \int_{\mathbb{R}^d} h^{d/(d+r)} \, dx \right)^{(d+r)/d}. \tag{0.1.2}$$

*In addition, if $\mu^s \equiv 0$ then*

$$\frac{1}{N} \sum_{i=1}^N \delta_{x^i} \rightharpoonup \frac{h^{d/d+r}}{\int_\Omega h^{d/d+r}(y)dy} \, dx \qquad \text{as } N \to \infty. \tag{0.1.3}$$

These issues are relatively well understood from the point of view of the calculus of variations [47, Chapter 1, Chapter 2]. One of our purposes is to consider instead a *dynamic approach* to this problem, as we shall describe now. Let $\mu = h\,dx$ and given $N$ points $x_0^1, \dots, x_0^N \in \mathbb{R}^d$ consider their evolution under the gradient flow generated by $F_{N,r}$, that is, solve the system of ODEs in $(\mathbb{R}^d)^N$

$$\begin{cases} \big(\dot{x}^1(t), \dots, \dot{x}^N(t)\big) = -\nabla F_{N,r}\big(x^1(t), \dots, x^N(t)\big), \\ \big(x^1(0), \dots, x^N(0)\big) = (x_0^1, \dots, x_0^N) \end{cases}$$

As usual in gradient flow theory, as $t \to \infty$ one expects the points $\big(x^1(t), \dots, x^N(t)\big)$ to converge to a minimizer $(\bar{x}^1, \dots, \bar{x}^N)$ of $F_{N,r}$. Hence, in view of (0.1.3), one expects

$$\frac{1}{N} \sum_{i=1}^N \delta_{\bar{x}^i} \rightharpoonup \frac{h^{d/d+r}}{\int_\Omega h^{d/d+r}(y)dy} \, dx \qquad \text{as } N \to \infty.$$

We want to understand this convergence result also at the level of the ODE. To do that, we need to take a "limit" as $N \to \infty$ in the above ODE system.

In order to give a meaning to this, we need to isometrically embed every $\mathbb{R}^N$ in one space, that we choose to be $L^2(\mathbb{R}^d;\mathbb{R}^d)$. More precisely, we take a set of reference points $(\hat{x}^1,\dots,\hat{x}^N)$ and we parameterize a general family of $N$ points $x^i$ as the image of $\hat{x}^i$ via a map $X:\mathbb{R}^d\to\mathbb{R}^d$, that is

$$x^i = X(\hat{x}^i).$$

In this way, the functional $F_{N,r}(x^1,\dots,x^N)$ can be rewritten in terms of the map $X$ and (a suitable renormalization of it) should converge to a functional $\mathcal{F}[X]$. Hence, we expect the evolution of $x^i(t)$ for $N$ large to be well-approximated by the $L^2$-gradient flow of $\mathcal{F}$.

In Chapter 2 we consider the one-dimensional case $\Omega = (0,1)$ and we show that the gradient flow PDE for the limiting functional $\mathcal{F}$ for the $L^2$-metric is given by the following non-linear parabolic equation

$$\partial_t X = \frac{1}{2^r(r+1)}\Big((r+1)\partial_\theta\big(h(X)|\partial_\theta X|^{r-1}\partial_\theta X\big) - h'(X)|\partial_\theta X|^{r+1}\Big), \qquad (0.1.4)$$

coupled with the Dirichlet boundary condition. Our main result shows that, under the assumptions that $\|h-1\|_{C^2} \ll 1$ and the initial datum is smooth and strictly increasing, the discrete and the continuous gradient flows remain *uniformly* close in $L^2$ for *all* times. In addition, by entropy-dissipation inequalities for the PDE, we can show that the continuous gradient flow converge exponentially fast to the stationary state for the PDE, which corresponds in Eulerian variables (see Section 0.1.1 below) to the measure $\frac{h^{1/3}\,d\theta}{\int h^{1/3}}$, as predicted by the static Theorem 0.1.2.

In particular, under the assumption that $\|h-1\|_{C^2} \ll 1$, we can prove the following quantitative convergence result for the empirical measure associated to the discrete gradient flow:

**Theorem 0.1.3.** *Let $\big(x^1(t),\dots,x^N(t)\big)$ be the gradient flow of $F_{N,2}$, and assume that $h\in C^{3,\alpha}([0,1])$ for some $\alpha\in(0,1)$, with $\|h-1\|_{C^2}\ll 1$. Then there exist two constants $c', C' > 0$ such that*

$$W_1\bigg(\frac{1}{N}\sum_i \delta_{x^i(t)}, \frac{h^{1/3}\,d\theta}{\int h^{1/3}}\bigg) \le C'\,e^{-c't/N^3} + \frac{C'}{N} \qquad \forall\, t\ge 0.$$

*In particular*

$$W_1\bigg(\frac{1}{N}\sum_i \delta_{x^i(t)}, \frac{h^{1/3}\,d\theta}{\int h^{1/3}}\bigg) \le \frac{2C'}{N} \qquad \forall\, t\ge \frac{N^3\log N}{c'}.$$

This theorem formalizes the heuristic argument given before about the fact that the gradient flows of $F_{N,2}$ should converge to a minimizer as $t \to \infty$.

Let us comment on the assumptions. First of all, the $C^{3,\alpha}$ regularity on $h$ allows us to use parabolic regularity theory to control the errors in the discretization of the functional. Secondly, the closeness in $C^2$ to 1 is justified by the fact that, under this assumption, the functional $\mathcal{F}$ is convex. Indeed, in general it is not true that the gradient flow of a functional converge to a global minimizer without some convexity assumptions [4]. In our paper we do not actually show that $F_{N,2}$ is convex (we believe this is false), but by a delicate combination of arguments including the maximum principle and $L^2$-stability estimates for $\mathcal{F}$, we can show that the discrete flow and the continuous one remain close, *uniformly* in time.

This result is completely new in its spirit: it combines tools from nonlinear PDEs and calculus of variations in a problem which, up to now, has always been studied with completely different approaches. As an ongoing project, we are trying to extend the result above to higher dimension. It is worth noticing that the strategy described above is very specific to the one dimensional case, hence completely new ideas and tools are needed.

### 0.1.1 From the Lagrangian to the Eulerian setting.

Equation (0.1.4) provides a Lagrangian description of the evolution of our system of particles in the limit $N \to \infty$. We can also study the Eulerian picture for the gradient flow PDE. If we denote by $f(t,x)$ the image of the Lebesgue measure through the map $X$, i.e.

$$f(t,x)dx = X(t,\theta)_{\#}d\theta,$$

then the PDE satisfied by $f$ takes the form

$$\partial_t f(t,x) = -rC_r\partial_x\bigg(f(t,x)\partial_x\Big(\frac{\rho(x)}{f(t,x)^{r+1}}\Big)\bigg), \tag{0.1.5}$$

with periodic boundary conditions, and we expect the following long time behavior

$$f(t,x) \longrightarrow \frac{\rho^{1/(r+1)}(x)}{\int_0^1 \rho(y)^{1/(r+1)}dy} \qquad \text{as } t \to \infty.$$

Notice that if $\rho \equiv 1$, (0.1.5) becomes

$$\partial_t f = -C_r(r+1)\partial_x^2\big(f^{-r}\big),$$

which is an equation of very fast diffusion type [93]. It is interesting to point out that the above equation set on the whole space $\mathbb{R}$ or with zero Dirichlet boundary conditions has no solutions, since all the mass instantaneously disappear [91, Theorem 3.1]. It is therefore crucial that in our setting the equation has periodic boundary conditions. In particular, in Chapter 2 we proved that our equation satisfies a comparison principle, and this plays a fundamental role in our proof of Theorem 0.1.3.

### 0.1.2 What happens on Riemannian manifolds?

The quantization problem on Riemannian manifolds had been previously studied on compact manifolds. In order to develop a more complete theory for this problem, in Chapter 3 we investigate the quantization problem for probability measures on general Riemannian manifolds.

While on compact manifolds one can prove (0.1.2) and (0.1.3) by using a suitable *localization argument*, the situation is very different when the manifold is *non-compact*. Indeed, some global hypotheses on the behavior of the measure at "infinity" have to be imposed. These new growth assumption that we find depends on the curvature of the manifold and reduces, in the flat case, to a moment condition. We also build an example showing that our hypothesis is sharp.

To state the result we need to introduce some notation: given a point $x_0 \in \mathcal{M}$, we can consider polar coordinates $(\rho, \vartheta)$ on $T_{x_0}\mathcal{M} \simeq \mathbb{R}^d$ induced by the constant metric $g_{x_0}$, where $\vartheta$ denotes a vector on the unit sphere $\mathbb{S}^{d-1}$. Then, we can define the following quantity that measures the size of the differential of the exponential map when restricted to a sphere $\mathbb{S}^{d-1}_\rho$:

$$A_{x_0}(\rho) := \rho \sup_{v \in \mathbb{S}^{d-1}_\rho,\, w \in T_v\mathbb{S}^{d-1}_\rho,\, |w|_{x_0}=1} \Big| d_v \exp_{x_0}(w) \Big|_{\exp_{x_0}(v)}, \tag{0.1.6}$$

The result on non-compact manifolds reads as follow:

**Theorem 0.1.4.** *Let $(\mathcal{M}, g)$ be a complete Riemannian manifold, and let $\mu = h\,d\mathrm{vol} + \mu^s$ be a probability measure on $\mathcal{M}$.*

*Assume that there exist $x_0 \in \mathcal{M}$ and $\delta > 0$ such that*

$$\int_{\mathcal{M}} d(x, x_0)^{r+\delta}\, d\mu(x) + \int_{\mathcal{M}} A_{x_0}\big(d(x, x_0)\big)^r\, d\mu(x) < \infty. \tag{0.1.7}$$

*and let $x^1, \ldots, x^N$ minimize the functional $F_{N,r} : (\mathcal{M})^N \to \mathbb{R}^+$. Then* (0.1.2) *and* (0.1.3) *hold.*

**Remark 0.1.5.** *If $\mathcal{M} = \mathbb{H}^d$ is the hyperbolic space, then $A_{x_0}(\rho) = \sinh \rho$ and* (0.1.7) *reads as*

$$\int_{\mathbb{H}^d} d(x, x_0)^{r+\delta}\, d\mu(x) + \int_{\mathbb{H}^d} \sinh\big(d(x, x_0)\big)^r\, d\mu(x) \approx \int_{\mathbb{H}^d} e^{r\, d(x, x_0)}\, d\mu(x) < \infty.$$

*If $\mathcal{M} = \mathbb{R}^d$ then $A_{x_0}(\rho) = \rho$ and* (0.1.7) *coincides with the finiteness of the* $(r+\delta)$*-moment of $\mu$, as in Theorem 0.1.2.*

We notice that the moment condition (0.1.1) required on $\mathbb{R}^d$ is not sufficient to ensure the validity of the result on $\mathbb{H}^d$. Indeed the following negative result holds:

**Theorem 0.1.6.** *There exists a measure $\mu$ on $\mathbb{H}^2$ such that*

$$\int_{\mathbb{H}^2} d(x, x_0)^p\, d\mu < \infty \qquad \forall\, p > 0,\ \forall\, x_0 \in \mathbb{H}^2,$$

*but*

$$N^{r/d} F_{N,r}(\mu) \to \infty \qquad \text{as } N \to \infty.$$

These results provide a conclusive answer to the static problem in the Riemannian setting.

## 0.2 Quasineutral limit for the Vlasov-Poisson equation.

The two main classes of kinetic equations are the collisional equations of Boltzmann type, and the mean field equations of Vlasov type modeling long-range interactions. Our focus here is on this latter class of equations that, for example, can be used to describe galaxies and plasmas.

To introduce the Vlasov-Poisson equation let us start considering $N$ classical particles of equal masses interacting via Newton's equations in $\mathbb{R}^d$

$$\ddot{x}_i(t) = -\frac{1}{N} \sum_j \nabla W(x_i(t) - x_j(t)),$$

where $x_i(t) \in \mathbb{R}^d$ is the position of particle $i$ at time $t$ and the force is given by an interaction potential $W : \mathbb{R}^d \to \mathbb{R}$ (with a proper scaling).

In applications $N$ may refer to the number of stars in a galaxy or to the electrons in a plasma, so $N$ could be of order of $10^{20}$! This makes the equations untractable in practice. The *mean field* limit $N \to \infty$ transforms this huge system of ODEs to one single PDE. It is not a priori obvious how one can let the dimension of the phase space go to infinity so, to perform the limit, we rewrite the equations in term of the *empirical measure* $\mu_t^N(dxdv) := N^{-1} \sum \delta_{((x_i(t),v_i(t))}$.

This measure belongs to the space of probability measures on the single-particle phase space, which is an infinite-dimensional space, but *independent of the number of particles.* So the Newton's equations become

$$\partial \mu_t^N + v \cdot \nabla_x \mu_t^N + F^N(t,x) \cdot \nabla_v \mu_t^N = 0, \qquad F^N := -\left(\nabla W *_{x,v} \mu_t^N\right);$$

and then, in the limit $N \to \infty$, we obtain an equation for the limiting measure $\mu_t(dxdv)$. Assuming that $\mu_t(dxdv) = f(t,x,v)dxdv$, we obtain the *nonlinear Vlasov equation* for the density function $f(t,x,v)$ with interaction potential $W$:[1]

$$(VE) := \begin{cases} \partial_t f + v \cdot \partial_x f + F(t,x) \cdot \partial_v f = 0, \\ F = -\nabla W *_x \rho \qquad \rho(t,x) = \int f(t,x,v) dv. \end{cases} \tag{0.2.1}$$

Among the possible choices for the interaction potential, a very important case is played by the *Vlasov-Poisson equation* in plasma physics. In this model heavy ions are treated as a fixed background, $f(t,x,v)$ is the density of electrons, and the interaction potential $W$ is the Coulomb potential, i.e., $W$ is the fundamental solution of the Laplacian. Another interesting variant of the Vlasov-Poisson equation is given by the *Vlasov-Poisson system for massless electrons*: this system describes a plasma from the viewpoint of the ions while the electrons move very fast and quasi-instantaneously reach their local thermodynamic equilibrium. Then the density of the electrons follows the classical Maxwell-Boltzmann law and the system reads as follows:

$$(VPME) := \begin{cases} \partial_t f + v \cdot \partial_x f + E \cdot \partial_v f = 0, \\ E = -\nabla U, \\ \Delta U = e^U - \int f \, dv = e^U - \rho, \\ f|_{t=0} = f_0 \geq 0, \quad \int f_0 \, dx \, dv = 1. \end{cases} \tag{0.2.2}$$

All the systems described above have been written in adimensional form. However, an important parameter appearing in plasmas is the so-called *Debye length.* The Debye

[1] Let us just notice that the mean-field limit has been rigorously justified only for smooth interactions.

length is the typical length of electrostatic interaction and in most physical situations is very small compared to the size of the domain. For this reason it is interesting to introduce a small parameter $\varepsilon$ in the system, that now is:

$$(VPME)_\varepsilon := \begin{cases} \partial_t f_\varepsilon + v \cdot \partial_x f_\varepsilon + E_\varepsilon \cdot \partial_v f_\varepsilon = 0, \\ E_\varepsilon = -\nabla U_\varepsilon, \\ \varepsilon^2 \Delta U_\varepsilon = e^{U_\varepsilon} - \int f_\varepsilon \, dv = e^{U_\varepsilon} - \rho_\varepsilon, \\ f_\varepsilon|_{t=0} = f_{0,\varepsilon} \geq 0, \quad \int f_{0,\varepsilon} \, dx \, dv = 1, \end{cases} \tag{0.2.3}$$

(and analogously for the classical Vlasov-Poisson system). The so-called quasineutral limit consists in understanding the behavior of solutions as $\varepsilon \to 0$. At least formally the limit is obtained in a straightforward way by taking $\varepsilon = 0$, and one gets

$$(KIE) := \begin{cases} \partial_t f + v \cdot \partial_x f + E \cdot \partial_v f = 0, \\ E = -\nabla U, \\ U = \log \rho, \\ f_0 \geq 0, \quad \int f_0 \, dx \, dv = 1, \end{cases} \tag{0.2.4}$$

a system we shall call the *kinetic isothermal Euler system.*

This passage to the limit is however extremely delicate to justify: indeed, it is known only in very few cases and it is actually false in some situations. This problem is the focus of the two Chapters 5 and 6.

In Chapter 5 we consider the one-dimensional case. Our main result there is the following quantitative weak-strong stability estimate for $(VPME)_\varepsilon$, in one dimension, with respect to the Wasserstein metric.

**Theorem 0.2.1.** *Let $T > 0$. Let $f^1_\varepsilon, f^2_\varepsilon$ be two measure solutions of* (0.2.3) *on* $[0, T]$, *and assume that $\rho^1_\varepsilon(t, x) := \int f^1_\varepsilon(t, x, v) \, dv$ is bounded in $L^\infty$ on $[0, T] \times \mathbb{T}$. Then, for all $\varepsilon \in (0, 1]$, for all $t \in [0, T]$,*

$$W_1(f^1_\varepsilon(t), f^2_\varepsilon(t)) \lesssim \frac{1}{\varepsilon} e^{\frac{1}{\varepsilon} e^{1/\varepsilon^2}} W_1(f^1_\varepsilon(0), f^2_\varepsilon(0)).$$

The proof of this result is based on a combination of several techniques from calculus of variations, PDEs, and probability theory.

Once this theorem is obtained, we can rely on some previously known convergence results of Grenier [49] where he showed convergence of the Vlasov-Poisson system in the quasineutral limit under analyticity assumptions on the initial data. Thanks to our theorem, we can prove stability for this convergence result under exponentially small perturbations in $W_1$. This implies in particular that even oscillatory behavior of the initial data are allowed, provided the oscillations are sufficiently small.

Although this stability result may look natural, let us mention that the validity of the quasineutral limit is false if the perturbation is only polynomially small (even is the size is measured in a strong Sobolev norm). Indeed, there exist smooth homogeneous equilibria $\mu(v)$ of the limit equation such that the following holds: For any $N > 0$ and $s > 0$, there exists a sequence of non-negative initial data $(f_{0,\varepsilon})$ such that

$$\|f_{\varepsilon,0} - \mu\|_{W^{s,1}_{x,v}} \leq \varepsilon^N,$$

and denoting by $(f_\varepsilon)$ the sequence of solutions to (0.2.3) with initial data $(f_{0,\varepsilon})$, for $\alpha \in [0, 1)$, we have:

$$\liminf_{\varepsilon \to 0} \sup_{t \in [0,\varepsilon^\alpha]} W_1(f_\varepsilon(t), \mu) > 0.$$

We also recall that the above result is (at least for the moment) very specific to the one dimensional case, as it strongly rely on some regularity estimates for the Laplacian that are false in dimension greater or equal than 2.

In Chapter 6 we investigate the 2 and 3-dimensional case for the classical Vlasov-Poisson equation. Although we are unable to prove a weak-strong stability estimate, we can show a strong-strong stability estimate (that is, under the assumption that both $\rho^1_\varepsilon(t, x)$ and $\rho^2_\varepsilon(t, x)$ are bounded in $L^\infty$) with respect to the 2-Wasserstein distance.

The proof of this result is completely different from the previous one, and it is based on some ideas introduced by Loeper in [70]. Since strong-strong stability relies on the $L^\infty$ bound on $\rho$ for both solutions, we need to prove new estimates on the growth of the $L^\infty$-norm of $\rho_\varepsilon(t, x)$ in terms of $t$ and $\varepsilon$. To obtain them we combine regularity estimates for the Laplacian with general results for transport equations.

# Part I

# Quantization of Measures

# Chapter 1

# Transportation theory and its relation to quantization of measures

In this Chapter we introduce the *optimal transport problem* and we explain its relation with the *quantization of measures*. The theory of optimal transportation has deep roots in the past, since it originates with the French geometer Gaspard Monge in 1781. Since then, it has become a classical subject in economics and optimization and, in the last years, it has been widely used in different areas of mathematics, such that partial differential equations, fluid mechanics, probability theory and kinetic theory. Since our purpose is to present the main tools and ideas that have been used in this thesis, this Chapter is definitely not exhaustive and, for a complete discussion in this topic, we refer to the classical monographs [4, 94, 95].

## 1.1 The optimal transport problem

Let $(X, d)$ be a complete, separable metric space. We recall that a *Borel measure* on $(X, d)$ is a probability measure defined on the Borel $\sigma$-algebra of $X$, i.e. the smallest $\sigma$-algebra that contains the open sets of $X$. Let us denote with $\mathcal{P}(X)$ the set of all Borel measures on $X$.

**Definition 1.1.1.** *Let us define $\mathcal{P}_p(X)$ the collection of all probability measures $\mu$ in $\mathcal{P}(X)$ with* finite p-moment*: for some $z$ in $X$*

$$\int_X d(x, z)^p d\mu(x) < \infty.$$

Notice that, if $d$ is bounded, $\mathcal{P}_p(X) = \mathcal{P}(X)$.

**Definition 1.1.2.** *Let $(X, \mu)$ and $(Y, \nu)$ be two probability spaces, let us denote with $\Pi(\mu, \nu)$ the set of all probability measures on $X \times Y$ with* marginals *$\mu$ and $\nu$. More precisely, $\pi \in \Pi(\mu, \nu)$ if and only if $\pi$ is a non negative measure such that*

$$\pi[A \times Y] = \mu[A], \quad \pi[X \times B] = \nu[B,]$$

*for all $A, B$ Borel subset of $X$ and $Y$ respectively.*

Since the tensor product $\mu \otimes \nu$ is in $\Pi(\mu, \nu)$, this set is non empty. Equivalently, $\pi \in \Pi(\mu, \nu)$ if and only if $\pi$ is a non negative measure on $X \times Y$ such that, for all couples of measurable functions $(\varphi, \psi) \in L^1(d\mu) \times L^1(d\nu)$, the following condition is satisfied:

$$\int_{X\times Y} [\varphi(x) + \psi(y)]\, d\pi(x, y) = \int_X \varphi\, d\mu + \int_Y \psi\, d\nu. \tag{1.1.1}$$

### 1.1.1 Kantorovich's problem

Now we can present the problem of *optimal transport* of measures. Assume that we are given a pile of sand, and there is a hole of same volume that we have to completely fill up with the sand. We can normalize the size of the pile to 1 and we can describe the pile of sand and the hole by two probability measures $\mu, \nu$ defined respectively on some measure spaces $X$ and $Y$.

Moving the sand from point $x$ to point $y$ needs some effort, that we can quantify via a measurable cost function $c(x, y)$ defined on $X \times Y$. It remains to clarify what a way of transportation, or a *transference plan* is. We can model transference plans using probability measures $\pi$ on $X \times Y$ with marginals $\mu$ and $\nu$: roughly speaking, $d\pi(x, y)$ measures the quantity of mass moved from point $x$ to location $y$. We do not a priori esclude the possibility that the mass located in point $x$ may be split into several destinations $y$.

For a transference plan to be admissible, we should require that all the mass at point $x$ coincides with $d\mu(x)$ and that all the mass transported at $y$ coincides with $d\nu(y)$. This is exactly the definition of $\pi \in \Pi(\mu, \nu)$.

The problem of transporting all the sand into the hole with the minimal effort can be translated into the problem of minimizing the *Kantorovich functional*:

$$I[\pi] = \int_{X\times Y} c(x, y)\, d\pi(x, y) \qquad \text{for } \pi \in \Pi(\mu, \nu). \tag{1.1.2}$$

For a given transport plan $\pi$, the non negative -and eventually infinite- quantity $I[\pi]$ is the *total transportation cost* associated to $\pi$, while the *optimal transportation cost* between $\mu$ and $\nu$ is

$$\mathcal{T}_c(\mu,\nu) = \inf_{\pi\in\Pi(\mu,\nu)} I[\pi].$$

The *optimal plans* $\pi$, if they exist, are the one that realize $I[\pi] = \mathcal{T}_c(\mu,\nu)$.

Kantorovich's optimal transportation problem has a *probabilistic interpretation.* Let $U$ be a random variable on $X$, i.e. a measurable map defined on a probability space $(\Omega, \mathbf{P})$ with values in $X$ such that the *law* of $U$ is the probability measure $\mu$ on $X$ defined as

$$\mu[A] = \mathbf{P}[U^{-1}(A)].$$

We denote by $\mathbf{E}$ the expectation (the integral with respect to $\mathbf{P}$) and by $l(U)$ the law of $U$. Then, given two probability measures $\mu$ e $\nu$, the goal is to minimize the expectation

$$I(U,V) = \mathbf{E}[c(U,V)], \tag{1.1.3}$$

over all pairs $(U,V)$ of random variables respectively in $X$ and $Y$ such that $l(U) = \mu$ and $l(V) = \nu$. With this reformulation, transference plans $\pi \in \Pi(\mu,\nu)$ are all possible laws associated to the pair $(U,V)$. This is also called a *coupling* of $U$ and $V$. As we shall explain at the end of Section 1.2, this probabilistic formulation will be particularly in Chapters 5 and 6.

Concerning the existence of optimal plans, this holds under very general assumptions on the cost functions. Here we just state the following result which is sufficient for our purposes.

**Theorem 1.1.3.** *Let $X, Y$ be two complete metric spaces, and assume that $c : X \times Y \to [0,\infty)$ is a continuous function. Then Problem* (1.1.2) *admits at least one solution.*

*Proof.* The proof is a simple application of the Direct Methods in the Calculus of Variations. In fact it can be easily checked that the set of transport plans is compact with respect to the weak convergence on $\mathcal{P}(\mathbb{R}^d)$ (see [95, Lemma 4.4]). Consider now a minimizing sequence $\{\pi_k\}_{k\geq 0}$ with $\pi_k \rightharpoonup \pi \in \Pi(\mu,\nu)$. Then, for any $M > 0$ we have that the cost $c_M(x,y) := \min\{c(x,y), M\}$ is bounded and continuous, so by the definition of weak convergence we have

$$\int_{X\times Y} c_M \, d\pi = \lim_{k\to\infty} \int_{X\times Y} c_M \, d\pi_k \leq \liminf_{k\to\infty} \int_{X\times Y} c \, d\pi_k = \mathcal{T}_c(\mu,\nu),$$

where we used that $c_M \leq c$ and that $\pi_k$ is a minimizing sequence. Hence, letting $M \to \infty$ in the left hand side we obtain

$$\int_{X\times Y} c\, d\pi \leq \mathcal{T}_c(\mu,\nu),$$

and since the other inequality is automatic (because $\pi \in \Pi(\mu,\nu)$) this proves that $\pi$ is a minimizer. □

**Remark 1.1.4.** A priori the above result does not guarantee that the infimum in Problem (1.1.2) is finite. However this is the case when $X = Y = \mathbb{R}^d$, $c(x,y) = |x-y|^p$, and $\mu,\nu \in \mathcal{P}_p(\mathbb{R}^d)$. Indeed, by the triangle inequality

$$|x-y|^p \leq (|x|+|y|)^p \leq 2^{p-1}(|x|^p + |y|^p),$$

and integrating the above inequality with respect to $\pi \in \Pi(\mu,\nu)$ we get

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathbb{R}^d} |x-y|^p\, d\pi &\leq 2^{p-1}\left(\int_{\mathbb{R}^d\times\mathbb{R}^d} |x|^p\, d\pi + \int_{\mathbb{R}^d\times\mathbb{R}^d} |y|^p\, d\pi\right)\\ &= 2^{p-1}\left(\int_{\mathbb{R}^d} |x|^p\, d\mu + \int_{\mathbb{R}^d} |y|^p\, d\nu\right) < \infty.\end{aligned}$$

### 1.1.2 Monge problem

Kantorovich's problem is a relaxed version of the original mass transportation problem formulated by Monge in 1781, in the famous paper *Mémoire sur la théorie des déblais et des remblais*. Monge's problem has the additional requirement that no mass is split. Thus, at every point $x$ is associated a unique destination $y$. In terms of random variables, this translates into a dependence of $V$ from $U$ in the equation (1.1.3); if we require this condition to be satisfied by transference plans, this means $\pi$ to have the form

$$d\pi(x,y) = d\pi_T(x,y) := d\mu(x)\,\delta[y = T(x)], \tag{1.1.4}$$

where $T$ is a measurable map from $X$ to $Y$. In this case, the transportation cost associated to $\pi_T$ is

$$I[\pi_T] = \int_X c(x,T(x))\, d\mu(x). \tag{1.1.5}$$

and the condition $\pi_T \in \Pi(\mu,\nu)$ corresponds to the *transport condition* $T_\#\mu = \nu$, that is

$$\nu[B] = \mu[T^{-1}(B)] \qquad \forall\, B \subset Y \text{ measurable set.}$$

In this case, one says that $\nu$ is the *push-forward* of $\mu$ through the map $T$, and $T$ is called a transport map from $\mu$ to $\nu$.

We can now state a strengthened version of Kantorovich's problem: minimize the functional

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

among all measurable maps $T$ such that $T\#\mu = \nu$.

While, as we have seen before, the existence of a solution to Kantorovich's problem is not difficult to prove, solving Monge's problems is highly nontrivial and in general it cannot be done unless one makes some assumptions on the measure $\mu$ and on the cost. We just mention that this problem has been solved for the first time in the case of the quadratic cost $c(x, y) = |x - y|^2$ by Brenier [23], and this result has been particularly interesting for his applications to fluid dynamics. However, since in this thesis we shall never consider this problem, we do not enter into this very interesting theory and we refer to the monographs [4, 94, 95].

## 1.2 Optimal transport metrics

This Section is dedicated to the study of some properties of what we shall call *Monge-Kantorovich distances*. The terminology associated to these distances varies a lot and these distances may be also called Kantorovich-Rubinstein distances and Wasserstein distances. In particular, the name of Wasserstein distance, actually introduced by Dobrushin, is very debatable since these distances were discovered and rediscovered by several authors. Nevertheless, the terminology "Wasserstein distance" has been extremely successful, and most of all recent papers relating optimal transport to partial differential equations, functional inequalities, Riemannian geometry and kinetic theory use this convention. In this thesis we use both the names Monge-Kantorovich distance and Wasserstein distance. In particular, in the Second Part that is focused on kinetic equations, we always use the name Wasserstein distance.

In the following we just require the underlying space $X$ to be a complete, separable metric space. This level of generality allows one to consider several concrete applications that require to use the Monge-Kantorovich distance on spaces like $C([0,1], \mathbb{R}^d)$ and $\mathcal{P}(\mathbb{R}^d)$. Let us now introduce the notions of Monge-Kantorovich distance of order $p$, and some topological properties of such distances (see for instance [94]).

**Definition 1.2.1** (Wasserstein distance)**.** *Let $(X, d)$ be a complete metric space, and recall that $\mathcal{P}_p(X)$ denotes the collection of all probability measures $\mu$ on $X$ with finite p moment. For $p \geq 1$, the p-Wasserstein distance between two probability measures $\mu$ and $\nu$ in $\mathcal{P}_p(X)$ is defined as*

$$W_p(\mu, \nu) := \left( \inf_{\gamma \in \Pi(\mu,\nu)} \int_{X \times X} d(x, y)^p \, \mathrm{d}\gamma(x, y) \right)^{1/p},$$

*where $\Pi(\mu, \nu)$ denotes the collection of all measures on $X \times X$ with marginals $\mu$ and $\nu$ on the first and second factors respectively. We refer equivalently to the Monge-Kantorovich distance:*

$$MK_p(\mu, \nu) = W_p^p(\mu, \nu).$$

*Moreover, recalling the definition of Kantorovich's problem, we can define $W_p = \mathcal{T}_p^{1/p}(\mu, \nu)$ where $\mathcal{T}_p(\mu, \nu)$ is the optimal transportation cost between $\mu$ and $\nu$ with respect to the cost function $c(x, y) = d(x, y)^p$.*

**Proposition 1.2.2** (Weak convergence in $\mathcal{P}_p(X)$)**.** *Let $(X, d)$ be a complete separable metric space. Then the Wasserstein space $\mathcal{P}_p(X)$ endowed with the Wasserstein distance $W_p$ is a complete separable metric space for all $p \in [1, \infty)$. In addition convergence $\mathcal{P}_p(X)$ can be descried as follows.*

*Let $(\mu_k)_{k \in \mathbb{N}}$ be a sequence of probability measures in $\mathcal{P}_p(X)$ and let $\mu$ be another measure in $P(X)$. Then $\mu_k$ converges weakly in $\mathcal{P}_p(X)$ to $\mu$ if any one of the following equivalent properties is satisfied for some (and then any) $x_0 \in X$:*

1. *$\mu_k \rightharpoonup \mu$ and $\int d(x, x_0)^p d\,\mu_k(x) \to \int d(x, x_0)^p d\,\mu(x)$;*
2. *$\mu_k \rightharpoonup \mu$ and $\limsup_{k \to \infty} \int d(x, x_0)^p d\,\mu_k(x) \leq \int d(x, x_0)^p d\,\mu(x)$;*
3. *$\mu_k \rightharpoonup \mu$ and $\lim_{R \to \infty} \limsup_{k \to \infty} \int_{d(x,x_0) \geq R} d(x, x_0)^p d\,\mu_k(x) = 0$;*
4. *For all continuous functions $\varphi$ with $|\varphi(x)| \leq C(1 + d(x, x_0)^p)$, $C \in \mathbb{R}$, one has*
$$\int \varphi(x)\, d\mu_k(x) \to \int \varphi(x) d\,\mu(x).$$

**Remark 1.2.3.**
- *By Hölder's inequality we have that*
$$p \leq q \Rightarrow W_p \leq W_q.$$
*In particular, the Wasserstein distance $W_1$ is the weakest of all, and results in $W_2$ distance are usually stronger than results in $W_1$ distance.*

- *Let* $(X, d)$ *be a compact metric space, and* $p \in [1, \infty)$*; then the Wasserstein distance* $W_p$ *metrizes the weak convergence in* $\mathcal{P}_p(X)$*. In other words, if* $(\mu_k)_{k\in\mathbb{N}}$ *is a sequence of measures in* $\mathcal{P}_p(X)$ *and* $\mu$ *is another measure in* $P(\mathcal{M})$*, then* $\mu_k$ *converges weakly in* $\mathcal{P}_p(X)$ *to* $\mu$ *if and only if*

$$W_p(\mu_k, \mu) \to 0 \qquad \text{as} \quad k \to \infty.$$

As it will also be clear in the Second Part of this thesis, the distances $W_p$ are particularly suited to estimate the distance between solutions to kinetic equations. Indeed, for Vlasov-Poisson equations, it is very natural to consider atomic solutions and $W_p$ is able to control the distance between the supports, while other classical distances, as for instance the total-variation, are too rough for this (recall that the total-variation distance between two Dirac masses is always 2 unless they coincide).

We observe that the most useful exponents in the Wasserstein distances are $p = 1$ and $p = 2$. As a general rule, the $W_1$ distance is more flexible and easier to bound, especially thanks to their duality relation with 1-Lipschitz functions (see Theorem 1.2.6). On the other hand, the quadratic Wasserstein distance $W_2$ encodes better some geometric features and it is widely used for problems with a Riemannian structure. In addition, the exponent 2 makes it very close to an $L^2$ type distance, and this often helps in computations.

### 1.2.1 Wasserstein distances and $L^p$ norms

The probabilistic interpretation of the Kantorovich's problem gives us the following equivalent formulation:

$$W_p(\mu, \nu)^p = \inf \mathbf{E}[|U - V|^p]$$

over all pairs $(U, V)$ of random variables such that $l(U) = \mu$ and $l(V) = \nu$.

Let us notice that, as a consequence of Theorem 1.1.3, the above infimum is actually a minimum. Indeed, if $\pi \in \Pi(\mu, \nu)$ is a minimizer in (1.1.2), it is enough to choose a couple of random variable $(U, V)$ whose joint law is the measure $\pi$.

As we shall see in Chapters 5 and 6, this probabilistic interpretation is very useful when estimating the Wasserstein distance between solutions of some PDE (in our case, they will be solutions of the Vlasov-Poisson system). To explain this point, suppose that we have two families of time-dependent measures $\mu_t$ and $\nu_t$ which solve some equation, and assume we want to estimate $W_p(\mu_t, \nu_t)$. A natural idea would be to differentiate the quantity $W_p(\mu_t, \nu_t)$ with respect to $t$ and try to use the equations for $\mu_t$ and $\nu_t$ to estimate the time derivative. However the derivative of the Wasserstein distance is not so easy to compute and to deal with, so one can use the following alternative strategy.

For $t = 0$ one chooses two random variables $U_0$ and $V_0$ with laws $\mu_0$ and $\nu_0$ respectively, and such that

$$\mathbf{E}[|U_0 - V_0|^p] = W_p(\mu_0, \nu_0)^p.$$

Then one let $U_t$ and $V_t$ evolve in time in such a way that their laws $\mu_t$ and $\nu_t$ are solutions to the equation we are interested in, and then one tries to control the quantity $\mathbf{E}[|U_t - V_t|^p]$ by estimating its time derivative and performing some Gronwall argument. In this way, if for instance one can prove that

$$\mathbf{E}[|U_t - V_t|^p] \le C(t)\, \mathbf{E}[|U_0 - V_0|^p],$$

then it follows by the identities $W_p(\mu_0, \nu_0)^p = \mathbf{E}[|U_0 - V_0|^p]$ and $W_p(\mu_t, \nu_t)^p \le \mathbf{E}[|U_t - V_t|^p]$ that

$$W_p(\mu_t, \nu_t)^p \le C(t)\, W_p(\mu_0, \nu_0)^p.$$

We refer to Chapters 5 and 6 for more details on this argument.

### 1.2.2 Kantorovich duality

Since Kantorovich's problem is a linear minimization problem with convex constraints, it has a *dual formulation* which can be stated in the following general form (see for instance [95, Theorem 5.10]):

**Theorem 1.2.4** (Kantorovich duality)**.** *Let $X$ and $Y$ be two complete metric spaces, and assume that $c : X \times Y \to [0, \infty)$ is a continuous function. Then*

$$\inf_{\pi \in \Pi(\mu,\nu)} I[\pi] = \sup \left\{ \int \varphi\, d\mu + \int \psi\, d\nu : \quad \varphi \in C(X),\, \psi \in C(Y), \quad \varphi(x) + \psi(y) \le c(x, y) \right\}.$$

The above result, which plays a crucial role in optimal transport theory, takes a even more convenient form when the cost function $c(x, y)$ is given by a distance $d(x, y)$. Indeed, in this case it implies that the the 1-Wasserstein metric is equivalent to the 1-Lipschitz distance.

**Definition 1.2.5.** *The* 1-Lipschitz distance $d_{1L}$ *between measure $\nu, \mu \in \mathcal{P}(X)$ is defined as*

$$d_{1L}(\nu, \mu) = \sup_{\varphi \in \mathcal{D}} \left| \int_X \varphi\, d\nu - \int_X \varphi\, d\mu \right|,$$

*where* $\mathcal{D} = \{\varphi : X \to \mathbb{R} \ : \ |\varphi(x) - \varphi(y)| \le |x - y|\}$.

We can now state the following duality result, and we refer to [94, Theorem 1.14] for a proof:

**Theorem 1.2.6** (Kantorovich-Rubinstein)**.** *Let* $(X, d)$ *be a separable complete metric space,* $\mu, \nu \in \mathcal{P}_1(X)$*. Then* $\mathcal{W}_1(\mu, \nu) = d_{1L}(\mu, \nu)$*.*

### 1.2.3 Optimal transport metrics and quantization of measures

We now conclude this introduction with a lemma that allows us to relate the quantization problem to the minimization of a functional depending only on points. As we shall see, this result is the starting point of all the analysis that we shall do in the next chapters.

**Lemma 1.2.7.** *Fix* $r \geq 1$ *and let* $\rho$ *be a probability density in* $\mathcal{P}_r(\Omega)$*. Then the following identity holds:*

$$\inf \left\{ MK_r \left( \sum_i m_i \delta_{x^i}, \rho(y)dy \right) \, : \, m_1, \ldots, m_N \geq 0, \, \sum_i m_i = 1 \right\} \\ = F_{N,r}(x^1, \ldots, x^N),$$

*where*

$$F_{N,r}(x^1, \ldots, x^N) := \int_\Omega \min_{1 \leq i \leq N} |x^i - y|^r \rho(y) dy.$$

*Proof.* Let us denote by $m = (m_1, \ldots, m_N)$ a point in $\mathbb{R}^N$ with nonnegative components such that

$$\sum_i m_i = 1.$$

By definition of Monge-Kantorovich distance we have that

$$\begin{aligned} &\inf \left\{ MK_r \left( \sum_i m_i \delta_{x^i}, \rho(y)dy \right) \, : \, m_1, \ldots, m_N \geq 0, \, \sum_i m_i = 1 \right\} \\ &= \inf_m \inf_\gamma \left\{ \int_{\Omega \times \Omega} |x - y|^r d\gamma(x, y) \, : \, (\pi_1)_\# \gamma = \sum_i m_i \delta_{x^i}, \, (\pi_2)_\# \gamma = \rho(y) dy \right\} \end{aligned}$$

where $\gamma$ varies among all probability measures on $\Omega \times \Omega$, and $\pi_i : \Omega \times \Omega \to \Omega$ $(i = 1, 2)$ denotes the canonical projection onto the $i$-th factor.

Let $\gamma_m$ be the optimal plan mapping $\sum_i m_i\delta_{x^i}$ onto $\rho(y)dy$, and let us define $A_{i,m} \subset \Omega$ such that

$$\operatorname{supp}\gamma_m = \bigcup_{i=1}^{N}\{x^i\}\otimes A_{i,m}.$$

Then we have

$$\begin{aligned}
&\inf\left\{MK_r\left(\sum_i m_i\delta_{x^i},\rho(y)dy\right)\ :\ m_1,\dots,m_N\geq 0,\ \sum_i m_i=1\right\}\\
&=\inf_m\inf_\gamma\left\{\int_{\Omega\times\Omega}|x-y|^r d\gamma(x,y)\ :\ (\pi_1)_\#\gamma=\sum_i m_i\delta_{x^i},\ (\pi_2)_\#\gamma=\rho(y)dy\right\}\\
&=\inf_m\int_{\Omega\times\Omega}|x-y|^r d\gamma_m(x,y)=\inf_m\sum_{i=1}^{N}\int_{A_{i,m}}|x^i-y|^r\rho(y)dy\\
&\geq\inf_m\sum_{i=1}^{N}\int_{A_{i,m}}\min_{1\leq j\leq N}|x^j-y|^r\rho(y)dy=\int_\Omega\min_{1\leq i\leq N}|x^i-y|^r\rho(y)dy=F_{N,r}(x^1,\dots,x^N).
\end{aligned}$$

In order to prove equality, we have just to find a plan $\gamma$ for which

$$F_{N,r}(x^1,\dots,x^N)=\int_{\Omega\times\Omega}|x-y|^r d\gamma(x,y).$$

To this aim, let us choose the masses $m_i$ via the definition of Voronoi cell $W(\{x^1,\dots,x^N\}|x^i)$ of the point $x^i$ with respect to the set $\{x^1,\dots,x^N\}$:

$$m_i:=\int_{W(\{x^1,\dots,x^N\}|x^i)}\rho(y)dy,$$

where

$$W(\{x^1,\dots,x^N\}|x^i):=\{y\in\Omega\ :\ |y-x^i|\leq|y-x^j|,\ j\in 1,\dots,N\}.$$

Defining

$$\gamma=\sum_{i=1}^{N}\delta_{x^i}\otimes\rho|_{W(\{x^1,\dots,x^N\}|x^i)},$$

and observing that

$$|x-y|^r=\min_{1\leq j\leq N}|x^j-y|^r\quad\gamma\text{-a.e.},$$

we get

$$\begin{aligned} F_{N,r}(x^1, \dots, x^N) &= \int_\Omega \min_{1\le i\le N} |x^i - y|^r \rho(y) dy \\ &= \int_{\Omega\times\Omega} \min_{1\le i\le N} |x^i - y|^r d\gamma(x,y) \\ &= \int_{\Omega\times\Omega} |x - y|^r d\gamma(x,y). \end{aligned}$$

□

# Chapter 2

# A gradient flow approach to quantization of measures

[1]

## 2.1 Introduction

**The quantization problem in the static case.** The problem of quantization of a $d$-dimension probability distribution by discrete probabilities with a given number of points can be stated as follows: Given a probability density $\rho$, approximate it by a convex combination of a finite number $N$ of Dirac masses. The quality of the approximation is usually measured in terms of the Monge-Kantorovich metric. Much of the early attention in the engineering and statistical literature was concentrated on the one-dimensional quantization problem. This problem arises in several contexts and has applications in information theory (signal compression), cluster analysis (quantization of empirical measures), pattern recognition, speech recognition, numerical integration, stochastic processes (sampling design), mathematical models in economics (optimal location of service centers), and kinetic theory. For a detailed exposition and a complete list of references, we refer to the monograph [47].

We now introduce the setup of the problem. Given $r \geq 1$, consider $\rho$ a probability

[1] This chapter is based on a joint work with Emanuele Caglioti and François Golse [26].

density on an open set $\Omega \subset \mathbb{R}^d$ such that

$$\int_\Omega |y|^r \rho(y)dy < \infty.$$

Given $N$ points $x^1, \dots, x^N \in \Omega$, one wants to find the best approximation of $\rho$, in the sense of Monge-Kantorovich, by a convex combination of Dirac masses centered at $x^1, \dots, x^N$. Hence one minimizes

$$\inf\left\{ MK_r\left(\sum_i m_i \delta_{x^i}, \rho(y)dy\right) \;:\; m_1, \dots, m_N \geq 0,\; \sum_i m_i = 1 \right\},$$

with

$$MK_r(\mu, \nu) := \inf\left\{ \int_{\Omega\times\Omega} |x-y|^r d\gamma(x,y) \;:\; (\pi_1)_\# \gamma = \mu,\; (\pi_2)_\# \gamma = \nu \right\},$$

where $\gamma$ varies among all probability measures on $\Omega \times \Omega$, and $\pi_i : \Omega \times \Omega \to \Omega$ $(i = 1, 2)$ denotes the canonical projection onto the $i$-th factor (see [4, 94] for more details on the Monge-Kantorovitch distance between probability measures).

As shown in the previous Chapter, the following facts hold:

1. The best choice of the masses $m_i$ is given by

$$m_i := \int_{W(\{x^1,\dots,x^N\}|x^i)} \rho(y)dy,$$

   where

$$W(\{x^1, \dots, x^N\}|x^i) := \{y \in \Omega \;:\; |y - x^i| \leq |y - x^j|,\; j \in 1, \dots, N\}$$

   is the so called *Voronoi cell* of $x^i$ in the set $x^1, \dots, x^N$.

2. The following identity holds:

$$\inf\left\{ MK_r\left(\sum_i m_i \delta_{x^i}, \rho(y)dy\right) \;:\; m_1, \dots, m_N \geq 0,\; \sum_i m_i = 1 \right\} = F_{N,r}(x^1, \dots, x^N),$$

   where

$$F_{N,r}(x^1, \dots, x^N) := \int_\Omega \min_{1\leq i\leq N} |x^i - y|^r \rho(y)dy.$$

If one chooses $x^1, \dots, x^N$ in an optimal way by minimizing the functional $F_{N,r} : (\mathbb{R}^d)^N \to \mathbb{R}^+$, in the limit as $N \to \infty$ these points distribute themselves accordingly to a probability density proportional to $\rho^{d/d+r}$. In other words, by [47, Chapter 2, Theorem 7.5] one has

$$\frac{1}{N}\sum_{i=1}^{N} \delta_{x^i} \rightharpoonup \frac{\rho^{d/d+r}}{\int_\Omega \rho^{d/d+r}(y)dy}\, dx. \tag{2.1.1}$$

These issues are relatively well understood from the point of view of the calculus of variations [47, Chapter 1, Chapter 2]. Our goal here is to consider instead a dynamic approach to this problem, as we shall describe now.

**A dynamical approach to the quantization problem.** Given $N$ points $x_0^1, \dots, x_0^N$, we consider their evolution under the gradient flow generated by $F_{N,r}$, that is, we solve the system of ODEs in $(\mathbb{R}^d)^N$

$$\begin{cases} \big(\dot{x}^1(t), \dots, \dot{x}^N(t)\big) = -\nabla F_{N,r}\big(x^1(t), \dots, x^N(t)\big), \\ \big(x^1(0), \dots, x^N(0)\big) = (x_0^1, \dots, x_0^N) \end{cases} \tag{2.1.2}$$

As usual in gradient flow theory, as $t \to \infty$ one expects that the points $\big(x^1(t), \dots, x^N(t)\big)$ converge to a minimizer $(\bar{x}^1, \dots, \bar{x}^N)$ of $F_{N,r}$. Hence (in view of (2.1.1)) the empirical measure

$$\frac{1}{N}\sum_{i=1}^{N} \delta_{\bar{x}^i}$$

is expected to converge to $\frac{\rho^{d/d+r}}{\int_\Omega \rho^{d/d+r}(y)dy} dx$ as $N \to \infty$.

We now want to take the limit in the ODE above as $N \to \infty$. For this, we take a set of reference points $(\hat{x}^1, \dots, \hat{x}^N)$ and we parameterize a general family of $N$ points $x^i$ as the image of $\hat{x}^i$ via a smooth map $X : \mathbb{R}^d \to \mathbb{R}^d$, that is

$$x^i = X(\hat{x}^i).$$

In this way, the functional $F_{N,r}(x^1, \dots, x^N)$ can be rewritten in terms of the map $X$ and (a suitable renormalization of it) should converge to a functional $\mathcal{F}[X]$. Hence, we can expect that the evolution of $x^i(t)$ for $N$ large is well-approximated by the $L^2$-gradient flow of $\mathcal{F}$.

Although this formal argument may look convincing, already the one dimensional case is very delicate. For this reason in this Chapter we shall focus on the one dimensional setting.

**The 1D case.** With no loss of generality we take $\Omega$ to be the open interval $(0,1)$ and we consider $\rho$ a smooth probability density on $\Omega$. In order to obtain a continuous version of the functional

$$F_{N,r}(x^1,\dots,x^N) = \int_0^1 \min_{1\le i\le N} |x^i - y|^r \rho(y)\,dy,$$

with $0 \le x^1 \le \ldots \le x^N \le 1$, assume that

$$x^i = X\left(\frac{i-1/2}{N}\right), \qquad i = 1,\dots,N$$

with $X : [0,1] \to [0,1]$ a smooth non-decreasing map such that $X(0) = 0$ and $X(1) = 1$. Then, as explained in Appendix 2.5,

$$N^r F_{N,r}(x^1,\dots,x^N) \longrightarrow C_r \int_0^1 \rho(X(\theta))|\partial_\theta X(\theta)|^{r+1} d\theta := \mathcal{F}[X]$$

as $N \to \infty$, where $C_r := \frac{1}{2^r(r+1)}$.

By a standard computation [37] we obtain the gradient flow PDE for $\mathcal{F}$ for the $L^2$-metric,

$$\partial_t X(t,\theta) = C_r\Big((r+1)\partial_\theta\big(\rho(X(t,\theta))|\partial_\theta X(t,\theta)|^{r-1}\partial_\theta X(t,\theta)\big) - \rho'(X(t,\theta))|\partial_\theta X(t,\theta)|^{r+1}\Big), \quad (2.1.3)$$

coupled with the Dirichlet boundary condition

$$X(t,0) = 0, \qquad X(t,1) = 1. \quad (2.1.4)$$

Let us notice that, in the particular case $\rho \equiv 1$, (2.1.3) becomes a $p$-Laplacian equation

$$\partial_t X = C_r(r+1)\partial_\theta\big(|\partial_\theta X|^{r-1}\partial_\theta X\big)$$

with $p-1 = r$ (see [32, 92] and references therein for a general treatment of this class of equations).

**From the Lagrangian to the Eulerian setting.** Equation (2.1.3) corresponds a Lagrangian description of the evolution of our system of particles in the limit $N \to \infty$. To consider its Eulerian counterpart, we denote by $f(t,x)$ the image of the Lebesgue measure through the map $X$, i.e.

$$f(t,x)dx = X(t,\theta)_\# d\theta.$$

Then the PDE satisfied by $f$ takes the form [2]

$$\partial_t f(t,x) = -rC_r\partial_x\Big( f(t,x)\partial_x\Big(\frac{\rho(x)}{f(t,x)^{r+1}}\Big)\Big), \tag{2.1.5}$$

with periodic boundary conditions, and we expect the following long time behavior

$$f(t,x) \longrightarrow \frac{\rho^{1/(r+1)}(x)}{\int_0^1 \rho(y)^{1/(r+1)}dy} \qquad \text{as } t \to \infty.$$

We note that if $\rho \equiv 1$, (2.1.5) becomes

$$\partial_t f = -C_r(r+1)\partial_x^2\big(f^{-r}\big),$$

which is an equation of very fast diffusion type [13, 92, 93]. It is interesting to point out that the above equation set on the whole space $\mathbb{R}$ or with zero Dirichlet boundary conditions has no solutions, since all the mass instantaneously disappear [91, Theorem 3.1]. It is therefore crucial that in our setting the equation has periodic boundary conditions. In particular, as we shall see, our equation satisfies a comparison principle (see Lemma 2.2.1).

**Assumptions on $\rho$ and convexity of the functionals.** Notice that our heuristic arguments in the previous section were based on the assumption that both the gradient flows of $F_{N,2}$ and of $\mathcal{F}$ converge to a minimizer as $t \to \infty$. Of course this is true if $F_{N,2}$ and $\mathcal{F}$ are convex [4]. Actually, notice that we are trying to show that the limits as $N \to \infty$ and $t \to \infty$ commute, and for this we need to prove that the discrete and the continuous gradient flows remain close in the $L^2$ sense, uniformly with respect to $t$. Therefore, the convexity of $\mathcal{F}$ and $F_{N,2}$ seems to be a very natural issue for the validity of our gradient flow strategy.

As shown in Appendix 2.6, for the hessian of $\mathcal{F}$ to be nonnegative at "points" $X$ which are Lipschitz and uniformly monotone, one has to assume $\rho$ to be sufficiently close to a constant in $C^2$. We shall therefore adopt this condition on $\rho$.

Whether this condition on $\rho$ ensures that $F_{N,2}$ is also convex is left undecided (actually, we believe this is false). Nevertheless we are able to prove that the discrete flow and

[2]Indeed since $\partial_t X = b(t,X)$ with $b(t,y) := C_r r\left(\frac{\rho(y)}{f(t,y)^{r+1}}\right)$ (this follows by a direct computation starting from (2.1.3)), the function $f \equiv f(t,x)$ solves the continuity equation $\partial_t f(t,x) + \mathrm{div}(b(t,x)f(t,x)) = 0$, as shown for instance in [1].

the continuous one remain close by a combination of arguments including the maximum principle and $L^2$-stability (see Section 2.4).

**Statement of the results.** In order to simplify our presentation, in the whole Chapter we shall focus only on the case $r = 2$. Indeed, this has the main advantage of simplifying some of the computations allowing us to highlight the main ideas. As will be clear from the sequel, this case already incorporates all the main features and difficulties of the problem, and this specific choice does not play any essential role.

As we mentioned in the previous section, the properties of $\rho$ are crucial in the proofs. Notice also that (2.1.3) is of $p$-laplacian type, which is a degenerate parabolic equation. In order to avoid degeneracy, it is necessary for the solution to be an increasing function of $\theta$. For this reason, we assume this on the initial datum and prove that this monotonicity is preserved along the flow.

It is worth noticing that the monotonicity estimate at the discrete level says that if $x^{i+1}(0) - x^i(0) \approx \frac{1}{N}$ for all $i$, this property is preserved in time (up to multiplicative constants). In particular the points $\{x^i(t)\}_{i=1,\dots,N}$ can never collide.

Our main result shows that, under the two above mentioned assumptions (that is, $\rho$ is close to a constant in $C^2$ and the initial datum is smooth and increasing), the discrete and the continuous gradient flows remain uniformly close in $L^2$ for *all* times. Notice however that the proofs of these results in the case $\rho \equiv 1$ and $\rho \not\equiv 1$ are quite different. Indeed, when $\rho \equiv 1$ the equation (2.1.3) depends on $\partial_\theta X$ and $\partial_{\theta\theta} X$, but not on $X$ itself. This fact plays a role in several places, both for showing the monotonicity of solutions (in particular for the discrete case) and in the convergence estimate. In particular, while in the case $\rho \equiv 1$ we obtain convergence of the discrete flow to the continuous one for all initial data, the case $\rho \not\equiv 1$ requires an additional assumption at time 0 (see (2.1.6)).

One further comment concerns the time scaling: notice that, in order to obtain a nontrivial limit of our functional $F_{N,r}$, we needed to rescale them by $1/N^r$. In addition to this, since we want to compare gradient flows, we have to take into account that the Euclidean metric in $\mathbb{R}^N$ has to be rescaled by a factor $1/N$ to be compared with the $L^2$ norm.[3] Hence, to compare the discrete and the continuous gradient flows, we need to

[3]Let $\bar{x} := (x^1, \dots, x^N), \bar{y} := (y^1, \dots, , y^N) \in \mathbb{R}^N$, and embed these points into $L^2([0,1])$ by defining the functions

$$X(\theta) := x^i, \quad Y(\theta) := y^i, \qquad \forall\, \theta \in \left(\frac{i-1}{N}, \frac{i}{N}\right).$$

Then $|\bar{x} - \bar{y}|^2 = \sum_{i=1}^N |x^i - y^i|^2$ while $\|X - Y\|_{L^2}^2 = \frac{1}{N} \sum_{i=1}^N |x^i - y^i|^2$.

rescale the former in time by a factor $N^{r+1}$.

We now state our convergence results, first when $\rho \equiv 1$ and then for the general case. It is worth to point out that the best way to approximate the uniform measure on $[0,1]$ with the sum of $N$ Dirac masses it to put masses of size $1/N$ centered at points $(i-1/2)/N$ and

$$MK_1\left(\frac{1}{N}\sum_{i=1}^{N}\delta_{(i-1/2)/N}, d\theta\right) = \frac{1}{4N}$$

(see the computation in the proof of Theorem 2.3.6). Hence the result in our next theorem shows that the gradient flow approach provides, for $N$ and $t$ large, the best approximation rate.

**Theorem 2.1.1.** *Let $\rho \equiv 1$, $\big(x^1(t),\dots,x^N(t)\big)$ the gradient flow of $F_{N,2}$, and $X(t)$ the gradient flow of $\mathcal{F}$ starting from $X_0$. Assume that $X_0 \in C^{4,\alpha}([0,1])$ and that there exist positive constants $c_0, C_0$ such that*

$$\frac{c_0}{N} \le \bar{x}^i(0) - \bar{x}^{i-1}(0) \le \frac{C_0}{N}, \text{ and } c_0 \le \partial_\theta X_0 \le C_0.$$

*Define $X^i(t) := X\left(t, \frac{i-1/2}{N}\right)$, $\bar{x}^i(t) := x^i(N^3 t)$, and $\mu_t^N := \frac{1}{N}\sum_i \delta_{x^i(t)}$ Then there exist two constants $c', C' > 0$, depending only on $c_0$, $C_0$, and $\|X_0\|_{C^{4,\alpha}([0,1])}$, such that, for all $t \ge 0$,*

$$\frac{1}{N}\sum_{i=1}^{N}\big(\bar{x}^i(t) - X^i(t)\big)^2 \le e^{-c't}\frac{1}{N}\sum_{i=1}^{N}\big(\bar{x}^i(0) - X^i(0)\big)^2 + \frac{C'}{N^4}$$

*and*

$$MK_1(\mu_t^N, d\theta) \le \frac{1}{4N} + C'\, e^{-c't/N^3} + \frac{C'}{N^2}.$$

*In particular*

$$MK_1(\mu_t^N, d\theta) \le \frac{1}{4N} + \frac{2C'}{N^2} \qquad \forall\, t \ge \frac{2N^3 \log N}{c'}.$$

**Theorem 2.1.2.** *Let $\big(x^1(t),\dots,x^N(t)\big)$ be the gradient flow of $F_{N,2}$, and $X(t)$ the gradient flow of $\mathcal{F}$ starting from $X_0$. Assume that $X_0 \in C^{4,\alpha}([0,1])$ for some $\alpha > 0$ and that there exist two positive constants $c_0, C_0$ such that*

$$\frac{c_0}{N} \le x^i(0) - x^{i-1}(0) \le \frac{C_0}{N}, \text{ and } c_0 \le \partial_\theta X_0 \le C_0.$$

*Define* $X^i(t) := X\left(t, \frac{i-1/2}{N}\right)$, $\bar{x}^i(t) := x^i(N^3 t)$, *and* $\mu_t^N := \frac{1}{N}\sum_i \delta_{x^i(t)}$, *and assume that* $\rho : [0,1] \to (0,\infty)$ *is a periodic probability density of class* $C^{3,\alpha}$ *with* $\|\rho'\|_\infty + \|\rho''\|_\infty \le \bar{\varepsilon}$ *and that*

$$|X^i(0) - x^i(0)| \le \frac{\bar{C}}{N^2} \qquad \forall\, i = 1, \ldots, N. \tag{2.1.6}$$

*for some positive constants* $\bar{\varepsilon}, \bar{C}$. *Then there exist two constants* $c', C' > 0$, *depending only on* $c_0$, $C_0$, $\|\rho\|_{C^{3,\alpha}([0,1])}$ *and* $\|X_0\|_{C^{4,\alpha}([0,1])}$, *such that the following holds: if* $\bar{\varepsilon}$ *is small enough (in terms of* $c_0$, $C_0$, *and* $\bar{C}$*) we have*

$$\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}^i(t) - X^i(t)\right)^2 \le \frac{C'}{N^4} \qquad \text{for all } t \ge 0$$

*and*

$$MK_1(\mu_t^N, \gamma\rho^{1/3}\, d\theta) \le C'\, e^{-c't/N^3} + \frac{C'}{N} \qquad \text{for all } t \ge 0,$$

*where*

$$\frac{1}{\gamma} := \int_0^1 \rho(\theta)^{1/3}\, d\theta.$$

*In particular*

$$MK_1(\mu_t^N, \gamma\rho^{1/3}\, d\theta) \le \frac{C'}{N} \qquad \text{for all } t \ge \frac{N^3 \log N}{c'}.$$

As a consequence of our results, under the assumption that $\rho$ is $C^2$ close to 1 we obtain a quantitative version of the results in [47]:

**Corollary 2.1.3.** *There exist two constants* $\bar{\varepsilon} > 0$ *and* $C > 0$ *such that the following holds: assume that* $\|\rho'\|_\infty + \|\rho''\|_\infty \le \bar{\varepsilon}$, *and let* $(x^1, \ldots, x^N)$ *be a minimizer of* $F_{N,2}$. *Then*

$$MK_1(\mu^N, \gamma\rho^{1/3}\, d\theta) \le \frac{C}{N}$$

*where*

$$\mu^N := \frac{1}{N}\sum_i \delta_{x^i}$$

*and*

$$\frac{1}{\gamma} := \int_0^1 \rho(\theta)^{1/3}\, d\theta.$$

This chapter is structured as follows: in the next section we collect several preliminary results both on the discrete and the continuous gradient flow. Then, we prove the convergence result first in the case $\rho \equiv 1$, and finally in the case $\|\rho - 1\|_{C^2([0,1])} \ll 1$.

In the whole Chapter we assume that $0 < \lambda \le \rho \le 1/\lambda$.

## 2.2 Preliminary results

### 2.2.1 The discrete gradient flow

We begin by computing the discrete gradient flow: as shown in the appendix, given points $0 \leq x^1 \leq \ldots \leq x^N \leq 1$, one has

$$F_{N,r}(x^1, \ldots, x^N) = \sum_{i=1}^{N} \int_{x^{i-1/2}}^{x^{i+1/2}} |y - x^i|^r \rho(y)dy$$

where

$$x^{i+1/2} := \frac{x^i + x^{i+1}}{2} \qquad \forall\, i = 2, \ldots, N-1,$$

while we set $x^{1/2} := 0$ and $x^{N+1/2} := 1$. Then, a direct computation gives

$$\frac{\partial F_{N,2}}{\partial x^i}(x^1, \ldots, x^N) = -2 \int_{x^{i-1/2}}^{x^{i+1/2}} (y - x^i)\rho(y)dy. \tag{2.2.1}$$

Moreover, assuming that $\rho$ is at least of class $C^0$ it is easy to check that $\nabla F_{N,2}$ is bounded and continuously differentiable, hence $F_{N,2}$ is of class $C^2$. Thus the gradient flow of $F_{N,2}$ is unique and exists globally for all $t \geq 0$ by the Cauchy-Lipschitz Theorem for ODEs.

### 2.2.2 The continuous gradient flow

In order to construct a solution to the continuous gradient flow (2.2.3) we start from the Eulerian description that we look as a PDE on $[0, 1]$ with periodic boundary conditions.

#### The Eulerian flow

Recall that by assumption $\lambda \leq \rho \leq 1/\lambda$ for some $\lambda > 0$. Given $f(t, x)$ a solution of (2.1.5), we set

$$m(x) := \rho(x)^{1/3}, \qquad u(t, x) := \frac{f(t, x)}{m(x)}.$$

With these new unknowns (2.1.5) becomes

$$\partial_t u = -\frac{1}{4\, m(x)}\, \partial_x \left( m(x) \partial_x \left( \frac{1}{u^2} \right) \right) \qquad \text{on } [0, \infty) \times [0, 1] \tag{2.2.2}$$

with periodic boundary conditions. The advantage of this form is double: first of all, the above PDE enjoys a comparison principle; secondly, constants are solutions. Since for our purposes, only comparison with constants is necessary, we will just show that.

**Lemma 2.2.1.** *Let $u$ be a nonnegative solution of* (2.2.2) *and $c$ be a positive constant. Then both*

$$t \mapsto \int_0^1 (u-c)_-\,dx \quad \text{and} \quad t \mapsto \int_0^1 (u-c)_+\,dx$$

*are nonincreasing functions.*

*Proof.* We show just the first statement (the other being analogous).

Since constants are solutions of (2.2.2), it holds

$$\partial_t(u-c) = -\frac{1}{4\,m}\,\partial_x\biggl(m\,\partial_x\biggl(\frac{1}{u^2}-\frac{1}{c^2}\biggr)\biggr).$$

We now multiply the above equation by $-m\,\phi_\varepsilon\left(\frac{1}{u^2}-\frac{1}{c^2}\right)$, with $\phi_\varepsilon$ a smooth approximation the indicator function of $\mathbb{R}_+$ satisfying $\phi_\varepsilon' \geq 0$. Integrating by parts we get

$$\begin{aligned}\frac{d}{dt}\int_0^1 \Psi_\varepsilon(u-c)\,dx &= -\int_0^1 \phi_\varepsilon\left(\frac{1}{u^2}-\frac{1}{c^2}\right)\,\partial_t(u-c)\,m\,dx\\ &= -\frac{1}{4}\int_0^1\biggl|\partial_x\biggl(\frac{1}{u^2}-\frac{1}{c^2}\biggr)\biggr|^2\phi_\varepsilon'\biggl(\frac{1}{u^2}-\frac{1}{c^2}\biggr)\,m\,dx \leq 0,\end{aligned}$$

where we have set

$$\Psi_\varepsilon(s) := -\int_0^s \phi_\varepsilon\biggl(\frac{1}{(\sigma+c)^2}-\frac{1}{c^2}\biggr)\,d\sigma.$$

Letting $\varepsilon \to 0$ we see that $\Psi_\varepsilon(s) \to s_-$ for $s \geq -c$, hence

$$\frac{d}{dt}\int_0^1 (u-c)_-\,dx \leq 0,$$

proving the result. □

Thus, if $a_0 \leq u(0,x) \leq A_0$, then $a_0 \leq u(t,x) \leq A_0$ for all $t \geq 0$. We now apply this fact to show that if $f$ is bounded away from zero and infinity at the initial time, then so it is for all positive times. More precisely, recalling that by assumption $\lambda \leq \rho \leq \frac{1}{\lambda}$, we have

$$\begin{aligned} a_1 \leq f(0,x) \leq A_1 \quad &\Rightarrow \quad \lambda^{1/3}a_1 \leq u(0,x) \leq \frac{A_1}{\lambda^{1/3}}\\ &\Rightarrow \quad \lambda^{1/3}a_1 \leq u(t,x) \leq \frac{A_1}{\lambda^{1/3}}\\ &\Rightarrow \quad \lambda^{2/3}a_1 \leq f(t,x) \leq \frac{A_1}{\lambda^{2/3}} \qquad \forall\, t \geq 0.\end{aligned}$$

These a priori bounds show that (2.1.5) is a uniformly parabolic equation. In particular, since $f$ is uniformly bounded for all times, by parabolic regularity theory (see for instance [53, Theorem 8.12.1], [41, Chapter 3, Section 3, Theorem 7], and [70, Chapters 5, 6]) we conclude that:

**Proposition 2.2.2.** *Let $\lambda \in (0,1]$, and assume that $\rho : [0,1] \to [\lambda, 1/\lambda]$ is periodic and of class $C^{k,\alpha}$ for some $k \geq 0$ and $\alpha \in (0,1)$. Let $f(0,\cdot) : [0,1] \to \mathbb{R}$ be a periodic function of class $C^{k,\alpha}$ satisfying $0 < a_1 \leq f(0,\cdot) \leq A_1$, and let $f$ solve (2.1.5) with periodic boundary conditions. Then*

$$\lambda^{2/3} a_1 \leq f(t,x) \leq \frac{A_1}{\lambda^{2/3}} \qquad \text{for all } t \geq 0,$$

*$f(0,\cdot)$ is of class $C^{k,\alpha}$ for all $t \geq 0$, and there exists a constant $C$, depending only on $\lambda$, $\|\rho\|_{C^{k,\alpha}}$, $k$, $\alpha$, $a_1$, and $A_1$, such that $\|f(t,\cdot)\|_{C^{k,\alpha}([0,1])} \leq C$ for all $t \geq 0$.*

### The Lagrangian flow

To obtain now existence and uniqueness for the gradient flow of $\mathcal{F}$, we simply define $X(t)$ for any $t \geq 0$ as the solution of the ODE (in $\theta$)

$$\begin{cases} \partial_\theta X(t,\theta) = \frac{1}{f(t,X(t,\theta))} & \text{on } [0,1], \\ X(t,0) = 0, \end{cases} \qquad \forall\, t \geq 0. \tag{2.2.3}$$

Notice that the boundary conditions $X(t,1) = 1$ is automatically satisfied since

$$\int_0^{X(t,1)} f(t,x)\, dx = 1$$

and $f(t) > 0$ is a probability on $[0,1]$. Also, notice that $X(t)$ has exactly one derivative more than $f(t)$. Hence, by Proposition 2.2.2 we obtain:

**Proposition 2.2.3.** *Let $\lambda \in (0,1]$, and assume that $\rho : [0,1] \to [\lambda, 1/\lambda]$ is periodic and of class $C^{k,\alpha}$ for some $k \geq 0$ and $\alpha \in (0,1)$. Let $X(0,\cdot)$ satisfy $0 < a_1 \leq \partial_\theta X(0,\cdot) \leq A_1$, $X(0,0) = 1$, $X(0,1) = 1$, and $\|X(0,\cdot)\|_{C^{k+1,\alpha}([0,1])} < \infty$, and let $X(t,\cdot)$ solve (2.1.3)-(2.1.4). Then*

$$\lambda^{2/3} a_1 \leq \partial_\theta X(t,\theta) \leq \frac{A_1}{\lambda^{2/3}} \qquad \text{for all } t \geq 0,$$

*and there exists a constant $C$, depending only on $\lambda$, $\|\rho\|_{C^{k,\alpha}}$, $k$, $\alpha$, $a_1$, and $A_1$, such that $\|X(t,\cdot)\|_{C^{k+1,\alpha}([0,1])} \leq C$ for all $t \geq 0$.*

# 2.3 The case $\rho \equiv 1$

As we already mentioned we shall focus only on increasing initial data, and as proved in the previous section this monotonicity is preserved in time, hence $\partial_\theta X \geq 0$.

We first observe that, in the case $\rho \equiv 1$, the equation (2.1.3) becomes

$$\partial_t X(t,\theta) = \frac{1}{4}\partial_\theta \left(\partial_\theta X(t,\theta)^2\right) = \frac{1}{2}\partial_\theta X(t,\theta)\partial^2_{\theta\theta}X(t,\theta) \tag{2.3.1}$$

with Dirichlet boundary conditions (2.1.4).

## 2.3.1 The $L^2$ estimate in the continuous case

The following result shows the exponential stability in $L^2$ of the continuous gradient flows.

**Proposition 2.3.1.** *Let $X_1, X_2$ be two solutions of* (2.3.1) *satisfying* (2.1.4) *and*

$$\partial_\theta X_i(0,\theta) \geq c > 0, \qquad i = 1,2. \tag{2.3.2}$$

*Then*

$$\int_0^1 |X_1(t,\theta) - X_2(t,\theta)|^2 \, d\theta \leq \left(\int_0^1 |X_1(0,\theta) - X_2(0,\theta)|^2 \, d\theta\right) e^{-4ct}.$$

*Proof.* We first recall that the monotonicity condition (2.3.2) is preserved in time (apply Proposition 2.2.3 with $\lambda = 1$). Then, since $X_2 - X_1$ vanishes at the boundary, one has

$$\begin{aligned}
\frac{d}{dt}\int_0^1 |X_1 - X_2|^2 \, d\theta &= \int_0^1 (X_1 - X_2)\left(\partial_\theta(\partial_\theta X_1^2) - \partial_\theta(\partial_\theta X_2^2)\right) d\theta \\
&= -\int_0^1 (\partial_\theta X_1 - \partial_\theta X_2)(\partial_\theta X_1^2 - \partial_\theta X_2^2) \, d\theta \\
&= -\int_0^1 (\partial_\theta X_1 - \partial_\theta X_2)^2(\partial_\theta X_1 + \partial_\theta X_2) \, d\theta.
\end{aligned}$$

Using the monotonicity condition $\partial_\theta X_i \geq c$ and the Poincaré inequality on $[0,1]$ (see for instance Lemma 2.3.5 and let $N \to \infty$), we get

$$\begin{aligned}
-\int_0^1 (\partial_\theta X_1 - \partial_\theta X_2)^2(\partial_\theta X_1 + \partial_\theta X_2) \, d\theta &\leq -2c\int_0^1 (\partial_\theta X_1 - \partial_\theta X_2)^2 \, d\theta \\
&\leq -4c\int_0^1 (X_1 - X_2)^2 \, d\theta
\end{aligned}$$

so that

$$\frac{d}{dt}\left(e^{4ct}\int_0^1 |X_1 - X_2|^2(t,\theta)\,d\theta\right) \leq 0.$$

□

This argument shows that, if at time zero $X_1(0,\theta) = X_2(0,\theta)$ for a.e. $\theta \in (0,1)$, in particular $X_1(t,\theta) = X_2(t,\theta)$ for a.e. $\theta \in (0,1)$, for all $t \geq 0$. Moreover if $X_1(0,\theta) - X_2(0,\theta)$ is small in $L^2$ then it remains small in $L^2$ (continuity with respect to the initial datum), and actually converges to zero exponentially fast. In particular, noticing that $X(t,\theta) = \theta$ is a solution (corresponding to $f(t,x) = 1$), we deduce that all solutions converge exponentially to it: indeed, choosing $X_2(t,\theta) = \theta$ and assuming $c \leq 1$ we have

$$\int_0^1 |X(t,\theta) - \theta|^2\,d\theta \leq \left(\int_0^1 |X(0,\theta) - \theta|^2\,d\theta\right) e^{-4ct}.$$

### 2.3.2 Convergence of the gradient flows

The functional $F_{N,2}(x^1,\dots,x^N)$ with $\rho \equiv 1$ is given by

$$F_{N,2}(x^1,\dots,x^N) = \frac{|x^1|^3}{3} + \sum_{i=1}^{N-1}\frac{1}{12}|x^{i+1} - x^i|^3 + \frac{|1 - x^N|^3}{3}, \tag{2.3.3}$$

hence the defining equation for the gradient flow for $F_{N,2}$ is

$$\dot{x}^i = -\frac{\partial F_{N,2}}{\partial x^i} = \frac{1}{4}\Big(\big(x^{i+1} - x^i\big)^2 - \big(x^i - x^{i-1}\big)^2\Big) \qquad \text{for all } i = 1,\dots,N, \tag{2.3.4}$$

where by convention $x^0 := -x^1$ and $x^{N+1} := 2 - x^N$.

The former convention comes from the following observation: in order to avoid problems at the boundary, one could symmetrize the configuration of points $x^1,\dots,x^N$ with respect to 0 to get $N$ points $y^1,\dots,y^N \in [-1,0]$ satisfying $y^i := -x^i$. By identifying $-1$ with 1, we then get a family of $2N$ points on the circle where the dynamics is completely equivalent to ours. This means that, by adding $x^0$ and $x^{N+1}$ defined as above, we can see $x^1$ and $x^N$ as interior points. In the next section we will apply the same observation symmetrizing also the density $\rho$ in the way described above.

In order to prove convergence, we want to find an equation for $X$ evaluated on the grid $(i - 1/2)N$.

**Lemma 2.3.2.** *Let $X(t,\theta)$ be a solution of* (2.3.1)-(2.1.4) *starting from an initial datum $X_0 \in C^{4,\alpha}([0,1])$ with $\partial_\theta X_0 \geq c_0 > 0$. Let $X^i$ be the discretized solution defined at the points $(\frac{i-1/2}{N}, t)$, that is*

$$X^i(t) := X\left(\frac{i-1/2}{N}, t\right) \qquad \text{for all } i = 1, \ldots, N. \tag{2.3.5}$$

*Then*

$$\partial_t X^i - N^3 \frac{\partial F_{N,2}}{\partial x^i}(X^1, \ldots, X^N) = R^i.$$

*with*

$$|R^i(t)| \leq \frac{\hat{C}}{N^2} \qquad \text{for all } t \geq 0, \quad \text{for all } i = 1, \ldots, N,$$

*where $\hat{C}$ depends only on $c_0$ and $\|X_0\|_{C^{4,\alpha}([0,1])}$.*

*Proof.* As we showed in Proposition 2.2.3 we have $\partial_\theta X(t,\theta) \geq c > 0$ for all $t$, so that the equation (2.3.1) remains uniformly parabolic and under our assumptions the solution $X(t)$ remains of class $C^4$ for all times, with

$$\|X(t)\|_{C^4} \leq C \qquad \forall\, t \geq 0.$$

By Taylor's expansion centered at $(\frac{i-1/2}{N}, t)$, one has

$$X^{i+1} = X^i + \frac{1}{N}\partial_\theta X^i + \frac{1}{2N^2}\partial_{\theta\theta} X^i + \frac{1}{6N^3}\partial_{\theta\theta\theta} X^i + O\left(\frac{\|X(t)\|_{C^4}}{N^4}\right),$$

$$X^{i-1} = X^i - \frac{1}{N}\partial_\theta X^i + \frac{1}{2N^2}\partial_{\theta\theta} X^i - \frac{1}{6N^3}\partial_{\theta\theta\theta} X^i + O\left(\frac{\|X(t)\|_{C^4}}{N^4}\right).$$

Thus, with the convention $X^0 := -X^1$ and $X^{N+1} := 2 - X^N$,

$$\begin{aligned}
&\partial_t X^i - \frac{N^3}{4}\Big( \big(X^{i+1} - X^i\big)^2 - \big(X^i - X^{i-1}\big)^2 \Big) = \\
&\qquad \partial_t X^i - \frac{N^3}{4}\left[\frac{1}{N}\partial_\theta X^i + \frac{1}{2N^2}\partial_{\theta\theta} X^i + \frac{1}{6N^3}\partial_{\theta\theta\theta} X^i + O\left(\frac{\|X(t)\|_{C^4}}{N^4}\right)\right]^2 \\
&\qquad\qquad + \frac{N^3}{4}\left[-\frac{1}{N}\partial_\theta X^i + \frac{1}{2N^2}\partial_{\theta\theta} X^i - \frac{1}{6N^3}\partial_{\theta\theta\theta} X^i + O\left(\frac{\|X(t)\|_{C^4}}{N^4}\right)\right]^2,
\end{aligned}$$

hence

$$\partial_t X^i - \frac{N^3}{4}\Big( \big(X^{i+1} - X^i\big)^2 - \big(X^i - X^{i-1}\big)^2 \Big) = \partial_t X^i - \frac{1}{2}\partial_\theta X^i \partial_{\theta\theta} X^i + R^i = R^i,$$

with

$$|R^i(t)| \leq C\,\frac{\|X(t)\|_{C^4}}{N^2} \leq \frac{\hat{C}}{N^2}.$$

with $\hat{C} := C \sup_{t\geq 0} \|X(t)\|_{C^4}$. □

In order to compare $X$ with $x^i$ we need to rescale times. More precisely, let us denote with $\bar{x}^i(t) := x^i(N^3 t)$. Then

$$\dot{\bar{x}}^i = \frac{N^3}{4}\Big( \big(\bar{x}^{i+1} - \bar{x}^i\big)^2 - \big(\bar{x}^i - \bar{x}^{i-1}\big)^2 \Big). \tag{2.3.6}$$

For simplicity of notation we set

$$W_X^i := N\left(X^{i+1} - X^i\right), \quad Y_X^i := \left(W_X^i\right)^2,$$

$$W_{\bar{x}}^i := N\left(\bar{x}^{i+1} - \bar{x}^i\right), \quad Y_{\bar{x}}^i := \left(W_{\bar{x}}^i\right)^2,$$

(recall the convention $X^0 := -X^1$ and $X^{N+1} := 2 - X^N$). The equation for $X^i$ can be written as

$$\partial_t X^i = \frac{N}{4}\big(Y_X^i - Y_X^{i-1}\big) + R^i,$$

while the equation for $W_{\bar{x}}^i$ (which follows easily from (2.3.6)) is given by

$$\partial_t W_{\bar{x}}^i = \frac{N^2}{4}\Big((W_{\bar{x}}^{i+1})^2 - 2(W_{\bar{x}}^i)^2 + (W_{\bar{x}}^{i-1})^2\Big). \tag{2.3.7}$$

We now prove a discrete monotonicity result:

**Lemma 2.3.3.** *Assume that $C \geq \partial_\theta X(0,\theta) \geq c$ and $C \geq W_{\bar{x}}^i(0) \geq c$ for all $i$ and $\theta \in (0,1)$. Then $C \geq W_{\bar{x}}^i(t), W_X^i(t) \geq c$ for all $i = 1,\ldots,N$, and all $t \geq 0$.*

*Proof.* The inequality for $W_X^i$ follows from the fact that the bound $C \geq \partial_\theta X(0) \geq c$ is propagated in time (see Proposition 2.2.3 and recall that here $\lambda = 1$).

To prove that $W_{\bar{x}}^i(t) \geq c > 0$, it suffices to prove that, for any $\varepsilon > 0$ small,

$$W_{\bar{x}}^i(t) \geq c - \varepsilon(2 - e^{-t}) := f(t) \quad \forall i, \forall\, t \geq 0 \tag{2.3.8}$$

(the bound $W_{\bar{x}}^i(t) \leq C$ being obtained in a is completely analogous manner). Notice that, with this choice, $f(0) < \min_i W_{\bar{x}}^i(0)$. Suppose by contradiction that

$$\min_i W_{\bar{x}}^i(t) \not\geq f(t) \text{ in } \mathbb{R}^+$$

Then there exist a first time $t_0$ such that $W_X^{i_0}(t_0) = f(t_0) \geq 0$ for some $i_0$, i.e., $f(t) < W_{\bar{x}}^i(t)$ for all $t \in [0, t_0)$ and all $i = 1, \dots, n$, and $f(t)$ touches $W_{\bar{x}}^i(t)$ from below at $(i_0, t_0)$. From the equation (2.3.7) and the condition (2.3.8) we get a contradiction: indeed, since $t_0$ is the first contact time we get

$$\dot{W}_{\bar{x}}^{i_0}(t_0) \leq \dot{f}(t_0) = -\varepsilon e^{-t_0} < 0,$$

while since $\big(W_{\bar{x}}^{i_0+1}(t_0)\big)^2, \big(W_{\bar{x}}^{i_0-1}(t_0)\big)^2 \geq f(t_0)^2 = \big(W_{\bar{x}}^{i_0}(t_0)\big)^2$ (here we used that $f(t) \geq 0$ provided $\varepsilon$ is sufficiently small to deduce that $W^i \geq f$ implies $(W^i)^2 \geq f^2$)

$$\dot{W}_{\bar{x}}^{i_0}(t_0) = \frac{N^2}{4}\Big(\big(W_{\bar{x}}^{i_0+1}(t_0)\big)^2 - 2\big(W_{\bar{x}}^{i_0}(t_0)\big)^2 + \big(W_{\bar{x}}^{i_0-1}(t_0)\big)^2\Big) \geq 0.$$

This proves that $\min_i W_{\bar{x}}^i(t) \geq f(t)$ for all $t \geq 0$, and letting $\varepsilon \to 0$ we have the desired result. □

We can now prove our convergence theorem.

**Theorem 2.3.4.** *Let $\bar{x}^i$ be a solution of the ODE* (2.3.6), *and let $X^i$ be as in* (2.3.5). *Assume that $X_0 \in C^{4,\alpha}([0,1])$ and that there exist positive constants $c_0, C_0$ such that*

$$\frac{c_0}{N} \leq \bar{x}^i(0) - \bar{x}^{i-1}(0) \leq \frac{C_0}{N}, \quad c_0 \leq \partial_\theta X_0 \leq C_0.$$

*Then there exist two constants $\bar{c}, \bar{C} > 0$, depending only on $c_0$, such that*

$$\frac{1}{N}\sum_{i=1}^{N}\big(\bar{x}^i(t) - X^i(t)\big)^2 \leq e^{-\bar{c}t}\frac{1}{N}\sum_{i=1}^{N}\big(\bar{x}^i(0) - X^i(0)\big)^2 + \bar{C}\left(\frac{\hat{C}}{N^2}\right)^2$$

*for all $t \geq 0$, where $\hat{C}$ is as in Lemma 2.3.2.*

*Proof.* We begin by observing that, because of Lemma 2.3.3,

$$\frac{c_0}{N} \leq \bar{x}^i(t) - \bar{x}^{i-1}(t) \leq \frac{C_0}{N}, \quad \text{and} \quad \frac{c_0}{N} \leq X^i(t) - X^{i-1}(t) \leq \frac{C_0}{N},$$

for all $t \geq 0$. We now estimate the $L^2$ distance between $X^i$ and $\bar{x}^i$: recalling Lemma

2.3.2 we have

$$
\begin{aligned}
&\frac{d}{dt}\frac{1}{N}\sum_{i=1}^{N}|X^i-\bar{x}^i|^2\\
&=\frac{1}{8N}\sum_{i=1}^{N}N\left(X^i-\bar{x}^i\right)\left[Y_X^i-Y_X^{i-1}-(Y_{\bar{x}}^i-Y_{\bar{x}}^{i-1})\right]+\frac{2}{N}\sum_{i=1}^{N}(X^i-\bar{x}^i)R^i\\
&=\frac{1}{8N}\sum_{i=1}^{N}N\left(X^i-\bar{x}^i\right)\left[Y_X^i-Y_{\bar{x}}^i\right]-\frac{1}{8N}\sum_{i=1}^{N}N\left(X^i-\bar{x}^i\right)\left[Y_X^{i-1}-Y_{\bar{x}}^{i-1}\right]\\
&\quad+\frac{2}{N}\sum_{i=1}^{N}(X^i-\bar{x}^i)R^i\\
&=\frac{1}{8N}\sum_{i=1}^{N}N\left(X^i-\bar{x}^i\right)\left[Y_X^i-Y_{\bar{x}}^i\right]-\frac{1}{8N}\sum_{i=0}^{N-1}N\left(X^{i+1}-\bar{x}^{i+1}\right)\left[Y_X^i-Y_{\bar{x}}^i\right]\\
&\quad+\frac{2}{N}\sum_{i=1}^{N}(X^i-\bar{x}^i)R^i\\
&=\frac{1}{8N}\sum_{i=0}^{N-1}N\left(X^i-\bar{x}^i\right)\left[Y_X^i-Y_{\bar{x}}^i\right]-\frac{1}{8N}\sum_{i=0}^{N-1}N\left(X^{i+1}-\bar{x}^{i+1}\right)\left[Y_X^i-Y_{\bar{x}}^i\right]\\
&\quad+\frac{2}{N}\sum_{i=1}^{N}(X^i-\bar{x}^i)R^i\\
&=-\frac{1}{8N}\sum_{i=0}^{N-1}N\left((X^{i+1}-X^i)-(\bar{x}^{i+1}-\bar{x}^i)\right)\left[Y_X^i-Y_{\bar{x}}^i\right]+\frac{2}{N}\sum_{i=1}^{N}(X^i-\bar{x}^i)R^i\\
&=-\frac{1}{8N}\sum_{i=0}^{N-1}(W_X^i-W_{\bar{x}}^i)\left[(W_X^i)^2-(W_{\bar{x}}^i)^2\right]+\frac{2}{N}\sum_{i=1}^{N}(X^i-\bar{x}^i)R^i\\
&\le-\frac{c}{8N}\sum_{i=0}^{N-1}(W_X^i-W_{\bar{x}}^i)^2+\frac{2}{N}\sum_{i=1}^{N}(X^i-\bar{x}^i)R^i,
\end{aligned}
$$

where at the last step we used that $W_{\bar{x}}^i, W_X^i \ge c > 0$. We then apply the following discrete Poincaré inequality (we postpone the proof to the end of the Theorem):

**Lemma 2.3.5.** *Let $(u^0,\dots,u^N)\subset\mathbb{R}^N$ with $u^0=0$. Set*

$$\|u\|_2 := \Big(\frac{1}{N}\sum_{i=0}^{N}(u^i)^2\Big)^{\frac{1}{2}};$$

$$\|u'\|_2 := \Big(\frac{1}{N}\sum_{i=0}^{N-1} N^2(u^{i+1}-u^i)^2\Big)^{\frac{1}{2}}.$$

*Then* $\|u\|_2^2 \le \frac{1}{2}\|u'\|_2^2$.

$$\frac{d}{dt}\frac{1}{N}\sum_{i=1}^{N}|X^i-\bar{x}^i|^2 \le -\bar{c}\frac{1}{N}\sum_{i=1}^{N}|X^i-\bar{x}^i|^2 + \frac{2}{N}\sum_{i=1}^{N}(X^i-\bar{x}^i)R^i.$$

Using that

$$(X^i-\bar{x}^i)R^i \le \epsilon(X^i-\bar{x}^i)^2 + \frac{1}{\epsilon}(R^i)^2,$$

choosing $\epsilon = \bar{c}/4$ we get

$$\frac{d}{dt}\frac{1}{N}\sum_{i=1}^{N}|X^i-\bar{x}^i|^2 \le -\bar{c}\frac{1}{2N}\sum_{i=1}^{N}|X^i-\bar{x}^i|^2 + \frac{2}{N}\sum_{i=1}^{N}(R^i)^2.$$

Recalling that

$$|R^i(t)| \le \frac{\hat{C}}{N^2},$$

(see Lemma 2.3.2), we conclude that

$$\frac{d}{dt}\frac{1}{N}\sum_{i=1}^{N}|X^i-\bar{x}^i|^2 \le -\bar{c}\frac{1}{2N}\sum_{i=1}^{N}|X^i-\bar{x}^i|^2 + \frac{2\,\hat{C}^2}{N^4}.$$

By Gronwall Lemma, this implies

$$\frac{1}{N}\sum_{i=1}^{N}|X^i(t)-\bar{x}^i(t)|^2 \le \frac{1}{N}\sum_{i=1}^{N}|X^i(0)-\bar{x}^i(0)|^2\, e^{-\bar{c}t/2} + \int_0^t e^{-\bar{c}(t-s)/2}\frac{2\,\hat{C}^2}{N^4}\,ds.$$

In particular, using that the third derivatives of $X(t,\cdot)$ are bounded, we get

$$\frac{1}{N}\sum_{i=1}^{N}|X^i(t)-\bar{x}^i(t)|^2 \le \frac{1}{N}\sum_{i=1}^{N}|X^i(0)-\bar{x}^i(0)|^2\, e^{-\bar{c}t/2} + \frac{2\hat{C}^2}{\bar{c}N^4},$$

as desired. □

*Proof of Lemma 2.3.5.* We observe that, since $u^0 = 0$,

$$u^i = \frac{1}{N} \sum_{k=0}^{i-1} N(u^{k+1} - u^k) \quad \text{for} \quad i = 0, \ldots, N,$$

hence

$$\begin{aligned}
\|u\|_2^2 = \frac{1}{N} \sum_{i=0}^{N} (u^i)^2 &= \frac{1}{N} \sum_{i=0}^{N} \Big( \frac{1}{N} \sum_{k=0}^{i-1} N(u^{k+1} - u^k) \Big)^2 \\
&\leq \frac{1}{N} \sum_{i=0}^{N} (i-1) \frac{1}{N^2} \sum_{k=0}^{i-1} N^2 (u^{k+1} - u^k)^2 \\
&= \frac{(N-1)}{2N} \frac{1}{N} \sum_{k=0}^{N-1} N^2 (u^{k+1} - u^k)^2 \leq \frac{1}{2} \|u'\|_2^2.
\end{aligned}$$

□

### The Eulerian counterpart

Let us define $\mu_t^N := \frac{1}{N} \sum_i \delta_{x^i(t)}$. We want to estimate the distance in $MK_1$ between $\mu_t^N$ and the Lebesgue measure on $[0, 1]$.

**Theorem 2.3.6.** *Let $\bar{x}^i$ be a solution of the ODE (2.3.6), and let $X^i$ be as in (2.3.5). Assume that $X_0 \in C^{4,\alpha}([0,1])$ and that there exist positive constants $c_0, C_0$ such that*

$$\frac{c_0}{N} \leq \bar{x}^i(0) - \bar{x}^{i-1}(0) \leq \frac{C_0}{N}, \quad c_0 \leq \partial_\theta X_0 \leq C_0.$$

*Then there exist two constants $\bar{\bar{c}}, \bar{\bar{C}} > 0$, depending only on $c_0, C_0, \|X_0\|_{C^{4,\alpha}([0,1])}$, such that*

$$MK_1(\mu_t^N, d\theta) \leq e^{-\bar{\bar{c}}t/N^3} + \frac{\bar{\bar{C}}}{N^2} + \frac{1}{4N} \qquad \forall\, t \geq 0.$$

*In particular*

$$MK_1(\mu_t^N, d\theta) \leq \frac{1}{4N} + \frac{\bar{\bar{C}} + 1}{N^2} \qquad \forall\, t \geq \frac{2N^3 \log N}{\bar{\bar{c}}}.$$

*Proof.* Take $X^0(\theta) = \theta$, so that $X(t, \theta) = \theta$ for all $t$, and apply Theorem 2.3.4: we know that

$$\frac{1}{N} \sum_{i=1}^{N} |X^i(t) - \bar{x}^i(t)|^2 \leq \frac{1}{N} \sum_{i=1}^{N} |X^i(0) - \bar{x}^i(0)|^2 \, e^{-\bar{c}t/2} + \frac{\bar{C}\, \hat{C}^2}{N^4},$$

hence, since $0 \le \bar{x}^i(0) \le 1$, $0 \le X^i(0) \le 1$,

$$\frac{1}{N}\sum_{i=1}^{N}\left|\bar{x}^i(t) - \frac{i-1/2}{N}\right|^2 \le e^{-\bar{c}t} + \frac{\bar{C}\,\hat{C}^2}{N^4}.$$

Recalling that

$$\bar{x}^i(t) := x^i(N^3 t/8),$$

we get

$$\frac{1}{N}\sum_{i=1}^{N}\left|x^i(t) - \frac{i-1/2}{N}\right|^2 \le e^{-\bar{c}t/N^3} + \frac{\bar{C}\,\hat{C}^2}{N^4}.$$

To control $MK_1(\mu_t^N, d\theta)$, we consider a 1-Lipschitz function $\varphi$ and we estimate

$$\begin{aligned}
\int_0^1 \varphi\, d\mu_t^N - \int_0^1 \varphi\, d\theta &= \frac{1}{N}\sum_{i=1}^{N}\varphi(x^i(t)) - \sum_{i=1}^{N}\int_{(i-1)/N}^{i/N}\varphi\, d\theta \\
&= \frac{1}{N}\sum_{i=1}^{N}\left[\varphi(x^i(t)) - \varphi\left(\frac{i-1/2}{N}\right)\right] \\
&+ \sum_{i=1}^{N}\int_{(i-1)/N}^{i/N}\left[\varphi\left(\frac{i-1/2}{N}\right) - \varphi(\theta)\right] d\theta \\
&\le \frac{1}{N}\sum_{i=1}^{N}\left|x^i(t) - \frac{i-1/2}{N}\right| + \sum_{i=1}^{N}\int_{(i-1)/N}^{i/N}\left|\frac{i-1/2}{N} - \theta\right| d\theta \\
&\le \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left|x^i(t) - \frac{i-1/2}{N}\right|^2} + \frac{1}{4N} \\
&\le e^{-\bar{c}t/(2N^3)} + \frac{\bar{C}^{1/2}\,\hat{C}}{N^2} + \frac{1}{4N},
\end{aligned}$$

hence, taking the supremum over all 1-Lipschitz functions we get

$$MK_1(\mu_t^N, d\theta) \le e^{-\bar{c}t/(2N^3)} + \frac{\bar{C}^{1/2}\,\hat{C}}{N^2} + \frac{1}{4N},$$

which proves the result with $\bar{\bar{c}} := \bar{c}/2$ and $\bar{\bar{C}} := \bar{C}^{1/2}\,\hat{C}$. □

## 2.4 The case $\rho \not\equiv 1$

We consider the case $r = 2$ whit $\rho$ a periodic function of class $C^{3,\alpha}$, where

$$\|\rho\|_{C^{3,\alpha}} := \|\rho\|_{C^3} + \sup_{x \neq y} \frac{|\rho'''(x) - \rho'''(y)|}{|x-y|^\alpha}.$$

We recall that

$$\mathcal{F}[X] = \frac{1}{12} \int_0^1 \rho(X(\theta)(\partial_\theta X(\theta))^3 d\theta,$$

and that the gradient flow PDE for $\mathcal{F}$ for the $L^2$-metric is given in (2.1.3).

### 2.4.1 Convergence of the gradient flows

We recall the formula for the gradient of $F_{N,2}$ given in (2.2.1).

**Lemma 2.4.1.** *Let $X(t,\theta)$ be a solution of* (2.1.3)-(2.1.4) *starting from an initial datum $X_0 \in C^{4,\alpha}([0,1])$ for some $\alpha > 0$ with $\partial_\theta X_0 \geq c_0 > 0$, and assume that $0 < \lambda \leq \rho \leq 1/\lambda$. Let $X^i$ be the discrete values of the exact solution at the points $\left(\frac{i-1/2}{N}, t\right)$ as in* (2.3.5). *Then*

$$\partial_t X^i - N^3 \frac{\partial F_{N,2}}{\partial x^i}(X^1, \ldots, X^N) = R^i$$

*with*

$$|R^i(t)| \leq \frac{\hat{C}}{N^2} \qquad \forall\, t \geq 0,\, \forall\, i = 1, \ldots, N, \tag{2.4.1}$$

*where $\hat{C}$ depends only on $c_0$, $\lambda$, $\|\rho\|_{C^{3,\alpha}([0,1])}$, and $\|X_0\|_{C^{4,\alpha}([0,1])}$.*

*Proof.* As we showed in Proposition 2.2.3, under our assumptions $\partial_\theta X(t) \geq c > 0$ for all $t$ and the solution $X(t)$ remains of class $C^4$ for all times, with

$$\|X(t)\|_{C^4} \leq C \qquad \forall\, t \geq 0.$$

A Taylor expansion yields

$$X^{i+1} = X^i + \frac{1}{N}\partial_\theta X^i + \frac{1}{2N^2}\partial_{\theta\theta} X^i + \frac{1}{6N^3}\partial_{\theta\theta\theta} X^i + O\biggl(\frac{\|X(t)\|_{C^4}}{N^4}\biggr);$$

$$X^{i-1} = X^i - \frac{1}{N}\partial_\theta X^i + \frac{1}{2N^2}\partial_{\theta\theta} X^i - \frac{1}{6N^3}\partial_{\theta\theta\theta} X^i + O\biggl(\frac{\|X(t)\|_{C^4}}{N^4}\biggr);$$

$$\rho(y) = \rho(X^i) + \rho'(X^i)\,(y - X^i) + \frac{\rho''(X^i)}{2}\,(y - X^i)^2 + O(|y - X^i|^3),$$

where as before we adopt the convention $X^0 := -X^1$ and $X^{N+1} := 2 - X^N$. In addition, we set

$$\rho(y) := \rho(-y) \quad \text{for } y \in [X^0, 0], \qquad \rho(y) := \rho(2 - y) \quad \text{for } y \in [1, X^{N+1}].$$

Then

$$\begin{aligned}
&-\frac{\partial F_{N,2}}{\partial x^i}(X^1, \ldots, X^N) \\
&= 2\int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} (y - X^i)\left[\rho(X^i) + \rho'(X^i)\,(y - X^i) + \frac{\rho''(X^i)}{2}\,(y - X^i)^2 + O(|y - X^i|^3)\right] dy \\
&= 2\int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} (y - X^i)\rho(X^i) dy \\
&+ 2\int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} (y - X^i)\left[\rho'(X^i)\,(y - X^i) + \frac{\rho''(X^i)}{2}\,(y - X^i)^2 + O(|y - X^i|^3)\right] dy \\
&= \frac{\rho(X^i)}{4}\left[(X^{i+1} - X^i)^2 - (X^i - X^{i-1})^2\right] + 2\rho'(X^i)\frac{1}{24}\left[(X^{i+1} - X^i)^3 - (X^i - X^{i-1})^3\right] \\
&+ \rho''(X^i)\frac{1}{64}\left[(X^{i+1} - X^i)^4 - (X^i - X^{i-1})^4\right] + O(1/N^5).
\end{aligned}$$

Therefore

$$\begin{aligned}
-\frac{\partial F_{N,2}}{\partial x^i}(X^1, \ldots, X^N) = {}& \frac{\rho(X^i)}{4}\left[(X^{i+1} - X^i)^2 - (X^i - X^{i-1})^2\right] \\
&+ \rho'(X^i)\frac{1}{12}\left[(X^{i+1} - X^i)^3 - (X^i - X^{i-1})^3\right] \\
&+ \rho''(X^i)\frac{1}{64}\left[(X^{i+1} - X^i)^4 - (X^i - X^{i-1})^4\right] \\
&+ O(1/N^5).
\end{aligned}$$

We now use the Taylor expansion for $X$ to see that

$$(X^{i+1} - X^i)^2 - (X^i - X^{i-1})^2 = \frac{2\partial_\theta X^i\,\partial_{\theta\theta} X^i}{N^3} + O(1/N^5),$$

$$(X^{i+1}-X^i)^3-(X^i-X^{i-1})^3=\frac{2(\partial_\theta X^i)^3}{N^3}+O(1/N^5),$$
$$(X^{i+1}-X^i)^4-(X^i-X^{i-1})^4=O(1/N^5),$$

thus

$$-\frac{\partial F_{N,2}}{\partial x^i}(X^1,\dots,X^N)$$
$$=\frac{1}{2N^3}\rho(X^i)\partial_\theta X^i\,\partial_{\theta\theta}X^i+\frac{1}{6N^3}\rho'(X^i)(\partial_\theta X^i)^3+O(1/N^5)=O(1/N^5).$$

□

### The $L^\infty$ stability estimate

Let $X$ be a smooth solution of the continuous gradient flow

$$\partial_t X=\frac{1}{2}\rho(X)\partial_\theta X\,\partial_{\theta\theta}X+\frac{1}{6}\rho'(X)(\partial_\theta X)^3 \tag{2.4.2}$$

and define

$$X^i(t):=X\left(t,\frac{i-1/2}{N}\right). \tag{2.4.3}$$

Recall that, according to Lemma 2.4.1, $X^i$ solves the following ODE:

$$\dot{X}^i=2N^3\int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}}(z-X^i)\rho(z)dz+R^i \tag{2.4.4}$$

where $R^i$ satisfies (2.4.1) and we are using the conventions $X^0:=-X^1$, $X^{N+1}:=2-X^N$, and

$$\rho(y):=\rho(-y)\quad\text{for } y\in[X^0,0],\qquad \rho(y):=\rho(2-y)\quad\text{for } y\in[1,X^{N+1}].$$

We also consider the rescaled discrete solution $\big(\bar{x}^i(t)\big)_{1\le i\le N}$

$$\dot{\bar{x}}^i=2N^3\int_{\frac{\bar{x}^i+\bar{x}^{i-1}}{2}}^{\frac{\bar{x}^i+\bar{x}^{i+1}}{2}}(z-\bar{x}^i)\rho(z)dz. \tag{2.4.5}$$

In the following lemma we prove that, over a time scale $\tau>0$, $X^i$ gets at most $\eta\tau$ apart from the exact solution of the ODE, where $\eta$ depends both on $\frac{\hat{C}}{N^2}$ and on the initial distance between the two solutions.

**Lemma 2.4.2.** *Let $\bar{x}^i$ be a solution of the ODE* (2.4.5)*, and let $X^i$ be as in* (2.4.3)*. Set*

$$A_t := \max_{i=1,\dots,N} |\bar{x}^i(t) - X^i(t)|.$$

*There exists a time $T > 0$, depending only on $\sup_{t\geq 0} \|X(t)\|_{C^2}$ and $\|\rho'\|_\infty + \|\rho''\|_\infty$, such that, for any $t^* \geq 0$,*

$$A_{t^*+\tau} \leq A_{t^*} + \eta\,\tau \qquad \forall\,\tau \in [0, T]$$

*with $\eta := \frac{3\hat{C}}{N^2} + \frac{A_{t^*}}{T}$, where $\hat{C}$ is as in* (2.4.1).

*Proof.* Let us define

$$A_t^{\pm} := \max_{i=1,\dots,N} \big(\pm[\bar{x}^i(t) - X^i(t)]\big)_+.$$

Notice that $A_t = \max\{A_t^+, A_t^-\}$, and to prove the result it is enough to prove the following stronger statement:

$$A_{t^*+\tau}^+ \leq A_{t^*}^+ + \eta\,\tau \qquad \forall\,\tau \in [0, T], \tag{2.4.6}$$

$$A_{t^*+\tau}^- \leq A_{t^*}^- - \eta\,\tau \qquad \forall\,\tau \in [0, T].$$

Since the arguments for $A^+$ and $A^-$ are completely analogous , we prove only (2.4.6). Also, without loss of generality we can assume $t^* = 0$. By definition of $A_0^+$, at time 0 the solutions are ordered

$$X^i(0) \leq \bar{x}^i(0) + A_0^+.$$

Let us define $Y^i(t) := X^i(t) - A_0^+ - \eta t$ and assume that there exist $t_0 \in \mathbb{R}^+$ defined as $t_0 := \inf_{t\in\mathbb{R}^+} \big\{Y^i(t) = \bar{x}^i(t)\big\}$. Then,

$$\dot{\bar{x}}^i(t_0) \leq \dot{Y}^i(t_0) \leq \dot{X}^i(t_0) - \eta. \tag{2.4.7}$$

Observing that $\bar{x}^{i+1}(t_0) \geq Y^{i+1}(t_0)$ and $\bar{x}^{i-1}(t_0) \geq Y^{i-1}(t_0)$,

$$\begin{aligned}\dot{\bar{x}}^i(t_0) =& 2N^3 \int_{\frac{\bar{x}^i+\bar{x}^{i-1}}{2}}^{\frac{\bar{x}^i+\bar{x}^{i+1}}{2}} (z - \bar{x}^i)\rho(z)dz \\ &\geq 2N^3 \int_{\frac{Y^i+Y^{i-1}}{2}}^{\frac{Y^i+Y^{i+1}}{2}} (z - Y^i)\rho(z)dz.\end{aligned}$$

Performing a change of variable $\omega = z + A_0^+ + \eta t_0$, we have

$$\begin{aligned}\dot{x}^i(t_0) &\geq 2N^3 \int_{\frac{Y^i+Y^{i-1}}{2}}^{\frac{Y^i+Y^{i+1}}{2}} (z - Y^i)\rho(z)dz \\ &= 2N^3 \int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} (\omega - X^i)\rho(\omega - A_0^+ - \eta t_0)d\omega.\end{aligned}$$

By the fundamental theorem of calculus

$$\begin{aligned}\rho(\omega - A_0^+ - \eta t_0) &= \rho(\omega) - (A_0^+ + \eta t_0)\left(\int_0^1 \rho'(\omega + s(A_0^+ + \eta t_0))ds\right) \\ &:= \rho(\omega) - a(\omega),\end{aligned}$$

so

$$\begin{aligned}\dot{x}^i(t_0) \geq{}& 2N^3 \int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} (\omega - X^i)\rho(\omega)d\omega \\ &- 2N^3(A_0^+ + \eta t_0) \int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} (\omega - X^i)a(\omega)d\omega.\end{aligned}$$

If we recall that $X^i$ solves the ODE (2.4.4) we have

$$\begin{aligned}\dot{x}^i(t_0) &\geq \dot{X}^i - R^i - 2N^3(A_0^+ + \eta t_0) \int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} (\omega - X^i)a(\omega)d\omega \\ &= \dot{X}^i - R^i - 2N^3(A_0^+ + \eta t_0) \int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} (\omega - X^i)\left(a(\omega) - a(X^i)\right)d\omega \\ &+ 2N^3(A_0^+ + \eta t_0) \int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} (\omega - X^i)a(X^i)d\omega \\ &:= \dot{X}^i - R^i - T_1 + T_2.\end{aligned}$$

For $T_1$ we observe that, since $|X^{i+1} - X^i| \leq C/N$ for all $i$,

$$|T_1| \leq CN^3(A_0^+ + \eta t_0)\|a'\|_\infty \int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} |\omega - X^i|^2 d\omega \leq C(A_0^+ + \eta t_0)\|\rho''\|_\infty.$$

For $T_2$ we use the Taylor expansion for $X$:

$$X^{i+1} = X^i + \frac{\partial_\theta X^i}{N} + O\left(\frac{1}{N^2}\right);$$
$$X^{i-1} = X^i - \frac{\partial_\theta X^i}{N} + O\left(\frac{1}{N^2}\right).$$

Thus,

$$\begin{aligned} T_2 &\le CN^3(A_0^+ + \eta t_0)\|\rho'\|_\infty \int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}} (\omega - X^i) d\omega \\ &= CN^3(A_0^+ + \eta t_0)\|\rho'\|_\infty \left[-\frac{1}{2}\left(-\frac{\partial_\theta X^i}{N} + O\left(\frac{1}{N^2}\right)\right)^2 + \frac{1}{2}\left(\frac{\partial_\theta X^i}{N} + O\left(\frac{1}{N^2}\right)\right)^2\right] \\ &\le C(A_0^+ + \eta t_0)\|\rho'\|_\infty. \end{aligned}$$

Then

$$\dot{x}^i(t_0) \ge \dot{X}^i - |R^i| - C(A_0^+ + \eta t_0)\left(\|\rho''\|_\infty + \|\rho'\|_\infty\right),$$

that combined with (2.4.7) and (2.4.1) gives

$$\begin{aligned} \eta &\le C(A_0^+ + \eta t_0)\left(\|\rho''\|_\infty + \|\rho'\|_\infty\right) + |R^i| \\ &\le C(A_0^+ + \eta t_0)\left(\|\rho''\|_\infty + \|\rho'\|_\infty\right) + \frac{\hat{C}}{N^2}. \end{aligned}$$

We now show that there exists a time $T > 0$, depending only on $\sup_{t\ge 0}\|X(t)\|_{C^2}$ and $\|\rho'\|_\infty + \|\rho''\|_\infty$, such that $t_0 > T$. This will prove that (2.4.6) holds on $[0, T]$.

Assume by contradiction that $t_0 \le T$. Then the above estimate gives

$$\eta \le C\left(\frac{A_0^+}{T} + \eta\right)T\left(\|\rho''\|_\infty + \|\rho'\|_\infty\right) + \frac{\hat{C}}{N^2}.$$

Choosing $T$ sufficiently small so that

$$CT\left(\|\rho''\|_\infty + \|\rho'\|_\infty\right) \le \frac{1}{2}$$

we get

$$\eta \le \frac{1}{2}\left(\frac{A_0^+}{T} + \eta\right) + \frac{\hat{C}}{N^2},$$

or equivalently

$$\eta \le \frac{A_0^+}{T} + 2\frac{\hat{C}}{N^2}.$$

This contradicts the definition of $\eta$ and proves the result. □

### The $L^2$ stability estimate

**Lemma 2.4.3.** *Let $\bar{x}^i$ be a solution of the ODE (2.4.5), and let $X^i$ be as in (2.4.3). Let $0 \le T_1 \le T_2 \le \infty$, and assume that there exist two positive constants $c_0, C_0$ such that*

$$\frac{c_0}{N} \le \bar{x}^i(t) - \bar{x}^{i-1}(t) \le \frac{C_0}{N}, \quad \frac{c_0}{N} \le X^i(t) - X^{i-1}(t) \le \frac{C_0}{N}, \qquad \forall\, t \in [T_1, T_2].$$

*Then, there exists $\varepsilon_0 = \varepsilon_0(c_0, C_0) > 0$ such that, if $\|\rho'\|_{L^\infty} + \|\rho''\|_{L^\infty} \le \varepsilon_0$ then one can find two constants $\bar{c}, \bar{C} > 0$, depending only on $c_0$, such that*

$$\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}^i(t) - X^i(t)\right)^2 \le e^{-\bar{c}(t-T_1)}\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}^i(T_1) - X^i(T_1)\right)^2 + \bar{C}\left(\frac{\hat{C}}{N^2}\right)^2$$

*for all $t \in [T_1, T_2]$.*

*Proof.* We compute

$$\begin{aligned}
&\frac{d}{dt}\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2 = \\
&4N^2\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)\left[\int_{\frac{\bar{x}^i+\bar{x}^{i-1}}{2}}^{\frac{\bar{x}^i+\bar{x}^{i+1}}{2}}(z-\bar{x}^i)\rho(z)dz - \int_{\frac{X^i+X^{i-1}}{2}}^{\frac{X^i+X^{i+1}}{2}}(z-X^i)\rho(z)dz\right] \\
&+\frac{2}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)R^i \\
&= 4N^2\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)\left[\int_{\frac{\bar{x}^i+\bar{x}^{i-1}}{2}}^{\bar{x}^i}(z-\bar{x}^i)\rho(z)dz\right. \\
&+\int_{\bar{x}^i}^{\frac{\bar{x}^i+\bar{x}^{i+1}}{2}}(z-\bar{x}^i)\rho(z)dz - \int_{\frac{X^i+X^{i-1}}{2}}^{X^i}(z-X^i)\rho(z)dz \\
&\left.-\int_{X^i}^{\frac{X^i+X^{i+1}}{2}}(z-X^i)\rho(z)dz\right] + \frac{2}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)R^i \\
&:= 4N^2\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)\left[A_{\bar{x}^i} + B_{\bar{x}^i} - A_{X^i} - B_{X^i}\right] + \frac{2}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)R^i.
\end{aligned}$$

For $A_{\bar{x}^i}$ and $B_{\bar{x}^i}$ we have

$$
\begin{aligned}
A_{\bar{x}^i} &= \int_{\frac{\bar{x}^i+\bar{x}^{i-1}}{2}}^{\bar{x}^i} (z-\bar{x}^i)\rho(\bar{x}^{i-1})dz + \int_{\frac{\bar{x}^i+\bar{x}^{i-1}}{2}}^{\bar{x}^i} (z-\bar{x}^i)\big(\rho(z)-\rho(\bar{x}^{i-1})\big)dz \\
&= -\frac{\rho(\bar{x}^{i-1})}{8}(\bar{x}^i-\bar{x}^{i-1})^2 + \int_{\frac{\bar{x}^i+\bar{x}^{i-1}}{2}}^{\bar{x}^i} (z-\bar{x}^i)\big(\rho(z)-\rho(\bar{x}^{i-1})\big)dz \\
&:= D_{\bar{x}}^{i-1} + E_{\bar{x}}^{i,2}.
\end{aligned}
$$

$$
\begin{aligned}
B_{\bar{x}^i} &= \int_{\bar{x}^i}^{\frac{\bar{x}^i+\bar{x}^{i+1}}{2}} (z-\bar{x}^i)\rho(\bar{x}^{i})dz + \int_{\bar{x}^i}^{\frac{\bar{x}^i+\bar{x}^{i+1}}{2}} (z-\bar{x}^i)\big(\rho(z)-\rho(\bar{x}^{i})\big)dz \\
&= \frac{\rho(\bar{x}^{i})}{8}(\bar{x}^{i+1}-\bar{x}^{i})^2 + \int_{\bar{x}^i}^{\frac{\bar{x}^i+\bar{x}^{i+1}}{2}} (z-\bar{x}^i)\big(\rho(z)-\rho(\bar{x}^{i})\big)dz \\
&:= D_{\bar{x}}^{i} + E_{\bar{x}}^{i,1}.
\end{aligned}
$$

Analogously we can set $A_{X^i} := D_X^{i-1} + E_X^{i,2}$ and $B_{X^i} := D_X^i + E_X^{i,1}$. In this way we have

$$
\begin{aligned}
&\frac{d}{dt}\frac{1}{N}\sum_{i=1}^N \left(\bar{x}^i - X^i\right)^2 \\
&= 4N^2 \sum_{i=1}^N \left(\bar{x}^i - X^i\right)\left[A_{\bar{x}^i} + B_{\bar{x}^i} - A_{X^i} - B_{X^i}\right] + \frac{2}{N}\sum_{i=1}^N \left(\bar{x}^i - X^i\right) R^i \\
&= 4N^2 \sum_{i=1}^N \left(\bar{x}^i - X^i\right)\left[D_{\bar{x}}^i - D_{\bar{x}}^{i-1} - D_X^i + D_X^i\right] \\
&+ 4N^2 \sum_{i=1}^N \left(\bar{x}^i - X^i\right)\left[E_{\bar{x}}^{i,1} - E_X^{i,1} + E_{\bar{x}}^{i,2} - E_X^{i,2}\right] + \frac{2}{N}\sum_{i=1}^N \left(\bar{x}^i - X^i\right) R^i \\
&= T_1 + T_2 + \frac{2}{N}\sum_{i=1}^N \left(\bar{x}^i - X^i\right) R^i.
\end{aligned}
$$

Let us estimate $T_1$ and $T_2$ separately. First,

$$
\begin{aligned}
T_1 &= 4N^2 \sum_{i=1}^N \left(\bar{x}^i - X^i\right)\left[D_{\bar{x}}^i - D_{\bar{x}}^{i-1} - D_X^i + D_X^i\right] \\
&= 4N^2 \left( \sum_{i=1}^N \left(\bar{x}^i - X^i\right)\left(D_{\bar{x}}^i - D_X^i\right) - \sum_{i=1}^N \left(\bar{x}^i - X^i\right)\left(D_{\bar{x}}^{i-1} - D_X^{i-1}\right) \right)
\end{aligned}
$$

Using the discrete version of the integration by parts we obtain

$$\begin{aligned}
T_1 &= 4N^2 \left( \sum_{i=1}^{N} \left(\bar{x}^i - X^i\right) \left(D^i_{\bar{x}} - D^i_X\right) - \sum_{i=1}^{N} \left(\bar{x}^i - X^i\right) \left(D^{i-1}_{\bar{x}} - D^{i-1}_X\right) \right) \\
&= 4N^2 \left( \sum_{i=1}^{N} \left(\bar{x}^i - X^i\right) \left(D^i_{\bar{x}} - D^i_X\right) - \sum_{i=1}^{N} \left(\bar{x}^{i+1} - X^{i+1}\right) \left(D^i_{\bar{x}} - D^i_X\right) \right) \\
&= 4N^2 \left( \sum_{i=1}^{N} \Big( \left(\bar{x}^i - \bar{x}^{i+1}\right) - \left(X^i - X^{i+1}\right) \Big) \left(D^i_{\bar{x}} - D^i_X\right) \right)
\end{aligned}$$

Recalling the definitions of $D^i_{\bar{x}}$ and $D^i_X$ we have

$$\begin{aligned}
T_1 &= -\frac{N^2}{4} \left( \sum_{i=1}^{N} \Big( \left(\bar{x}^{i+1} - \bar{x}^i\right) - \left(X^{i+1} - X^i\right) \Big) \Big( \rho(\bar{x}^i)(\bar{x}^{i+1} - \bar{x}^i)^2 - \rho(X^i)(X^{i+1} - X^i)^2 \Big) \right) \\
&= -\frac{N^2}{4} \left( \sum_{i=1}^{N} \Big[ \left(\bar{x}^{i+1} - \bar{x}^i\right) - \left(X^{i+1} - X^i\right) \Big] \Big[ \rho(\bar{x}^i) \Big( (\bar{x}^{i+1} - \bar{x}^i)^2 - (X^{i+1} - X^i)^2 \Big) \Big] \right) \\
&+ \frac{N^2}{4} \left( \sum_{i=1}^{N} \Big[ \left(\bar{x}^{i+1} - \bar{x}^i\right) - \left(X^{i+1} - X^i\right) \Big] \Big[ \left(\rho(\bar{x}^i) - \rho(X^i)\right) \left(X^{i+1} - X^i\right)^2 \Big] \right) \\
&=: T_{1,1} + T_{1,2}.
\end{aligned}$$

Notice that, since $\|\rho'\|_\infty \leq \varepsilon_0$ and $\|\rho\|_{L^1} = 1$, we have $\rho \geq 1/2$ provided $\varepsilon_0$ is small enough. Hence, recalling that $\bar{x}^{i+1} - \bar{x}^i \geq \frac{c_0}{N}$ and $X^{i+1} - X^i \geq \frac{c_0}{N}$, we can estimate the first term

$$\begin{aligned}
T_{1,1} &\leq -\frac{N^2}{8} \sum_{i=1}^{N} \Big[ \left(\bar{x}^{i+1} - \bar{x}^i\right) - \left(X^{i+1} - X^i\right) \Big]^2 \Big[ \Big( (\bar{x}^{i+1} - \bar{x}^i) + (X^{i+1} - X^i) \Big) \Big] \\
&\leq -\frac{c_0}{4} N \sum_{i=1}^{N} \Big[ \left(\bar{x}^{i+1} - \bar{x}^i\right) - \left(X^{i+1} - X^i\right) \Big]^2 \\
&= -\frac{c_0}{4N} \sum_{i=1}^{N} \Big[ N \left(\bar{x}^{i+1} - \bar{x}^i\right) - N \left(X^{i+1} - X^i\right) \Big]^2.
\end{aligned}$$

Hence, recalling that $X^{i+1}-X^i \leq \frac{C_0}{N}$,

$$
\begin{aligned}
|T_{1,2}| &\leq \|\rho'\|_{L^\infty}\frac{N^2}{4}\left(\sum_{i=1}^N \left|\left(\bar{x}^{i+1}-\bar{x}^i\right)-\left(X^{i+1}-X^i\right)\right|\left|\left(\bar{x}^i-X^i\right)\left(X^{i+1}-X^i\right)^2\right|\right)\\
&\leq \frac{C_0^2}{N}\|\rho'\|_{L^\infty}\sum_{i=1}^N \left|N\left(\bar{x}^{i+1}-\bar{x}^i\right)-N\left(X^{i+1}-X^i\right)\right|\left|\bar{x}^i-X^i\right|.
\end{aligned}
$$

Using the inequality $ab \leq a^2+b^2$ we get

$$
\begin{aligned}
|T_{1,2}| \leq &\frac{C_0^2}{N}\|\rho'\|_{L^\infty}\sum_{i=1}^N \left[N\left(\bar{x}^{i+1}-\bar{x}^i\right)-N\left(X^{i+1}-X^i\right)\right]^2\\
&+\frac{C_0^2}{N}\|\rho'\|_{L^\infty}\sum_{i=1}^N \left(\bar{x}^i-X^i\right)^2.
\end{aligned}
$$

Let us now consider $T_2$.

$$
\begin{aligned}
T_2 &= 4N^2\sum_{i=1}^N\left(\bar{x}^i-X^i\right)\left[E_{\bar{x}}^{i,1}-E_X^{i,1}+E_{\bar{x}}^{i,2}-E_X^{i,2}\right]\\
&= 4N^2\sum_{i=1}^N\left(\bar{x}^i-X^i\right)\left[E_{\bar{x}}^{i,1}-E_X^{i,1}\right]+4N^2\sum_{i=1}^N\left(\bar{x}^i-X^i\right)\left[E_{\bar{x}}^{i,2}-E_X^{i,2}\right]\\
&:= T_{2,1}+T_{2,2}.
\end{aligned}
$$

Let us first focus on the differences $E_{\bar{x}}^{i,1}-E_X^{i,1}$ and $E_{\bar{x}}^{i,2}-E_X^{i,2}$. Keeping in mind the definitions of $E_{\bar{x}}^{i,1}$ and $E_X^{i,1}$ we have

$$
\begin{aligned}
&E_{\bar{x}}^{i,1}-E_X^{i,1} =\\
&\int_{\bar{x}^i}^{\frac{\bar{x}^i+\bar{x}^{i+1}}{2}}\left(z-\bar{x}^i\right)\left(\rho(z)-\rho(\bar{x}^i)\right)dz-\int_{X^i}^{\frac{X^i+X^{i+1}}{2}}\left(z-X^i\right)\left(\rho(z)-\rho(X^i)\right)dz.
\end{aligned}
$$

Performing the change of variable $\omega = z-\bar{x}^i$, $\omega = z-X^i$ respectively, we get

$$
\begin{aligned}
E_{\bar{x}}^{i,1}-E_X^{i,1} = &\int_0^{\frac{\bar{x}^{i+1}-\bar{x}^i}{2}}\omega\left(\rho(\omega+\bar{x}^i)-\rho(\bar{x}^i)\right)d\omega\\
&-\int_0^{\frac{X^{i+1}-X^i}{2}}\omega\left(\rho(\omega+X^i)-\rho(X^i)\right)d\omega.
\end{aligned}
$$

Adding and subtracting

$$\int_0^{\frac{\bar{x}^{i+1}-\bar{x}^i}{2}} \omega \left(\rho(\omega + X^i) - \rho(X^i)\right) d\omega$$

we have

$$\begin{aligned} E_{\bar{x}}^{i,1} - E_X^{i,1} = &\int_0^{\frac{\bar{x}^{i+1}-\bar{x}^i}{2}} \omega \Big[\rho(\omega + \bar{x}^i) - \rho(\bar{x}^i) - \rho(\omega + X^i) + \rho(X^i)\Big] d\omega \\ &- \int_{\frac{\bar{x}^{i+1}-\bar{x}^i}{2}}^{\frac{X^{i+1}-X^i}{2}} \omega \left(\rho(\omega + X^i) - \rho(X^i)\right) d\omega. \end{aligned}$$

By the fundamental theorem of calculus and recalling that $(\bar{x}^{i+1} - \bar{x}^i) \leq \frac{C_0}{N}$, $(X^{i+1} - X^i) \leq \frac{C_0}{N}$ we obtain the following estimate

$$\begin{aligned} |E_{\bar{x}}^{i,1} - E_X^{i,1}| = &\Bigg| \int_0^{\frac{\bar{x}^{i+1}-\bar{x}^i}{2}} \omega^2 \left[\int_0^1 \rho'(\bar{x}^i + s\omega) ds - \int_0^1 \rho'(X^i + s\omega) ds\right] d\omega \\ &- \int_{\frac{\bar{x}^{i+1}-\bar{x}^i}{2}}^{\frac{X^{i+1}-X^i}{2}} \omega \left(\rho(\omega + X^i) - \rho(X^i)\right) d\omega \Bigg| \\ &\leq \frac{C_0}{N^3} \|\rho''\|_{L^\infty} |\bar{x}^i - X^i| + \|\rho'\|_{L^\infty} \left| \int_{\frac{\bar{x}^{i+1}-\bar{x}^i}{2}}^{\frac{X^{i+1}-X^i}{2}} \omega^2 d\omega \right| \\ &= \frac{C_0}{N^3} \|\rho''\|_{L^\infty} |\bar{x}^i - X^i| + \frac{\|\rho'\|_{L^\infty}}{8} \left| \left(X^{i+1} - X^i\right)^3 - \left(\bar{x}^{i+1} - \bar{x}^i\right)^3 \right|. \end{aligned}$$

Thus,

$$\begin{aligned} |T_{2,1}| &= 4N^2 \sum_{i=1}^N \left|\bar{x}^i - X^i\right| \left|E_{\bar{x}}^{i,1} - E_X^{i,1}\right| \\ &\leq \frac{C}{N} \|\rho''\|_{L^\infty} \sum_{i=1}^N \left(\bar{x}^i - X^i\right)^2 \\ &+ \frac{N^2}{2} \|\rho'\|_{L^\infty} \sum_{i=1}^N |\bar{x}^i - X^i| \left| \left(X^{i+1} - X^i\right)^3 - \left(\bar{x}^{i+1} - \bar{x}^i\right)^3 \right|. \end{aligned}$$

Recalling that $0 \le (\bar{x}^{i+1} - \bar{x}^i) \le \frac{C_0}{N}$ and $0 \le (X^{i+1} - X^i) \le \frac{C_0}{N}$ we see that

$$\Big| \left(X^{i+1} - X^i\right)^3 - \left(\bar{x}^{i+1} - \bar{x}^i\right)^3 \Big| \le \frac{C}{N^2} \Big| \left(X^{i+1} - X^i\right) - \left(\bar{x}^{i+1} - \bar{x}^i\right) \Big|,$$

therefore

$$\begin{aligned}
|T_{2,1}| &\le \frac{C}{N} \|\rho''\|_{L^\infty} \sum_{i=1}^{N} \left(\bar{x}^i - X^i\right)^2 \\
&\quad + \frac{C}{N} \|\rho'\|_{L^\infty} \sum_{i=1}^{N} |\bar{x}^i - X^i| \big| N\left(\bar{x}^{i+1} - \bar{x}^i\right) - N\left(X^{i+1} - X^i\right) \big| \\
&\le \frac{C}{N} \|\rho''\|_{L^\infty} \sum_{i=1}^{N} \left(\bar{x}^i - X^i\right)^2 \\
&\quad + \frac{C}{N} \|\rho'\|_{L^\infty} \left[ \sum_{i=1}^{N} |\bar{x}^i - X^i|^2 + \sum_{i=1}^{N} \big| N\left(\bar{x}^{i+1} - \bar{x}^i\right) - N\left(X^{i+1} - X^i\right) \big|^2 \right].
\end{aligned}$$

Let us now estimate $E^{i,2}_{\bar{x}} - E^{i,2}_{X}$. By definition we have

$$\begin{aligned}
&E^{i,2}_{\bar{x}} - E^{i,2}_{X} = \\
&\int_{\frac{\bar{x}^i + \bar{x}^{i-1}}{2}}^{\bar{x}^i} (z - \bar{x}^i) \left(\rho(z) - \rho(\bar{x}^{i-1})\right) dz - \int_{\frac{X^i + X^{i-1}}{2}}^{X^i} (z - X^i) \left(\rho(z) - \rho(X^{i-1})\right) dz.
\end{aligned}$$

Performing the change of variable $\omega = z - \bar{x}^{i-1}$, $\omega = z - X^{i-1}$ respectively, we get

$$\begin{aligned}
E^{i,2}_{\bar{x}} - E^{i,2}_{X} &= \int_{\frac{\bar{x}^i - \bar{x}^{i-1}}{2}}^{\bar{x}^i - \bar{x}^{i-1}} \left(\omega + \bar{x}^{i-1} - \bar{x}^i\right) \left(\rho(\omega + \bar{x}^{i-1}) - \rho(\bar{x}^{i-1})\right) d\omega \\
&\quad - \int_{\frac{X^i - X^{i-1}}{2}}^{X^i - X^{i-1}} \left(\omega + X^{i-1} - X^i\right) \left(\rho(\omega + X^{i-1}) - \rho(X^{i-1})\right) d\omega.
\end{aligned}$$

Adding and subtracting

$$- \int_{\frac{X^i - X^{i-1}}{2}}^{X^i - X^{i-1}} \left(\omega + X^{i-1} - X^i\right) \left(\rho(\omega + X^{i-1}) - \rho(X^{i-1})\right) d\omega$$

we get

$$\begin{aligned}
&|E^{i,2}_{\bar{x}} - E^{i,2}_X| \\
&\leq \left| \int_{\frac{\bar{x}^i-\bar{x}^{i-1}}{2}}^{\bar{x}^i-\bar{x}^{i-1}} \left(\omega + \bar{x}^{i-1} - \bar{x}^i\right) \left(\rho(\omega + \bar{x}^{i-1}) - \rho(\bar{x}^{i-1}) - \rho(\omega + X^{i-1}) + \rho(X^{i-1})\right) d\omega \right| \\
&+ \left| - \int_{\frac{X^i-X^{i-1}}{2}}^{X^i-X^{i-1}} \left(\omega + X^{i-1} - X^i\right) \left(\rho(\omega + X^{i-1}) - \rho(X^{i-1})\right) d\omega \right. \\
&+ \left. \int_{\frac{\bar{x}^i-\bar{x}^{i-1}}{2}}^{\bar{x}^i-\bar{x}^{i-1}} \left(\omega + X^{i-1} - X^i\right) \left(\rho(\omega + X^{i-1}) - \rho(X^{i-1})\right) d\omega \right|.
\end{aligned}$$

Arguing as we did for the first term in $E^{i,1}_{\bar{x}} - E^{i,1}_X$, the first term in $E^{i,2}_{\bar{x}} - E^{i,2}_X$ is controlled by

$$\|\rho''\|_\infty \int_{\frac{\bar{x}^i-\bar{x}^{i-1}}{2}}^{\bar{x}^i-\bar{x}^{i-1}} \left|\omega + \bar{x}^{i-1} - \bar{x}^i\right| \left|X^{i-1} - \bar{x}^{i-1}\right| \omega^2 \, d\omega,$$

and recalling that $|\bar{x}^{i-1} - \bar{x}^i| \leq C_0/N$, the above term is bounded by

$$\frac{C}{N^3} \|\rho''\|_\infty \left|X^{i-1} - \bar{x}^{i-1}\right|.$$

Concerning the second term in $E^{i,2}_{\bar{x}} - E^{i,2}_X$, using that

$$\left| \int_{a/2}^{a} - \int_{b/2}^{b} \right| \leq \left| \int_0^a - \int_0^b \right| + \left| \int_0^{a/2} - \int_0^{b/2} \right| = \left| \int_a^b \right| + \left| \int_{a/2}^{b/2} \right|$$

we get

$$
\begin{aligned}
&\Bigg| \int_{\frac{\bar{x}^i-\bar{x}^{i-1}}{2}}^{\bar{x}^i-\bar{x}^{i-1}} \left(\omega + X^{i-1} - X^i\right)\left(\rho(\omega + X^{i-1}) - \rho(X^{i-1})\right) d\omega \\
&- \int_{\frac{X^i-X^{i-1}}{2}}^{X^i-X^{i-1}} \left(\omega + X^{i-1} - X^i\right)\left(\rho(\omega + X^{i-1}) - \rho(X^{i-1})\right) d\omega \Bigg| \\
&\leq \left| \int_{\frac{\bar{x}^i-\bar{x}^{i-1}}{2}}^{\frac{X^i-X^{i-1}}{2}} \left(\omega + X^{i-1} - X^i\right)\left(\rho(\omega + X^{i-1}) - \rho(X^{i-1})\right) d\omega \right| \\
&+ \left| \int_{\bar{x}^i-\bar{x}^{i-1}}^{X^i-X^{i-1}} \left(\omega + X^{i-1} - X^i\right)\left(\rho(\omega + X^{i-1}) - \rho(X^{i-1})\right) d\omega \right|. \\
&\leq \|\rho'\|_{L^\infty} \Bigg[ \Bigg| \int_{\frac{\bar{x}^i-\bar{x}^{i-1}}{2}}^{\frac{X^i-X^{i-1}}{2}} \left|\omega + X^{i-1} - X^i\right| |\omega| \, d\omega \Bigg| \\
&\qquad + \left| \int_{\bar{x}^i-\bar{x}^{i-1}}^{X^i-X^{i-1}} \left|\omega + X^{i-1} - X^i\right| |\omega| \, d\omega \right| \Bigg].
\end{aligned}
$$

We now notice that in the last term the second integral is bounded by the first integral hence we can bound it by

$$
\begin{aligned}
&2\|\rho'\|_{L^\infty} \int_{\bar{x}^i-\bar{x}^{i-1}}^{X^i-X^{i-1}} \omega^2 \, d\omega + 2\|\rho'\|_{L^\infty}(X^i - X^{x-i}) \int_{\bar{x}^i-\bar{x}^{i-1}}^{X^i-X^{i-1}} \omega \, d\omega \\
&\leq C\|\rho'\|_{L^\infty} \left| (X^i - X^{i-1})^3 - (\bar{x}^i - \bar{x}^{i-1})^3 \right| \\
&\qquad + C\|\rho'\|_{L^\infty}(X^{i-1} - X^i) \left| (X^i - X^{i-1})^2 - (\bar{x}^i - \bar{x}^{i-1})^2 \right|.
\end{aligned}
$$

Hence, arguing as for $T_{2,1}$, we obtain

$$
\begin{aligned}
|T_{2,2}| \leq{}& \frac{C}{N}\|\rho''\|_{L^\infty} \sum_{i=1}^{N} \left(\bar{x}^i - X^i\right)^2 \\
&+ \frac{C}{N}\|\rho'\|_{L^\infty} \left[ \sum_{i=1}^{N} |\bar{x}^i - X^i|^2 + \sum_{i=1}^{N} \left| N\left(\bar{x}^{i+1} - \bar{x}^i\right) - N\left(X^{i+1} - X^i\right) \right|^2 \right].
\end{aligned}
$$

Combining all these bounds together, we get

$$\begin{aligned}
&\frac{d}{dt}\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2\\
&= T_1 + T_2 + \frac{2}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)R^i\\
&= T_{1,1} + T_{1,2} + T_{2,1} + T_{2,2} + \frac{2}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)R^i\\
&\leq -\frac{c_0\lambda}{2N}\sum_{i=1}^{N}\Big[N\left(\bar{x}^{i+1} - \bar{x}^i\right) - N\left(X^{i+1} - X^i\right)\Big]^2\\
&\quad + \frac{C}{N}\Big(\|\rho'\|_{L^\infty} + \|\rho''\|_{L^\infty}\Big)\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2\\
&\quad + \frac{C}{N}\Big(\|\rho'\|_{L^\infty} + \|\rho''\|_{L^\infty}\Big)\sum_{i=1}^{N}\Big[N\left(\bar{x}^{i+1} - \bar{x}^i\right) - N\left(X^{i+1} - X^i\right)\Big]^2\\
&\quad + \frac{2}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)R^i.
\end{aligned}$$

Hence, recalling that $\|\rho'\|_{L^\infty} + \|\rho''\|_{L^\infty} \leq \varepsilon_0$, we can choose $\varepsilon_0$ small (the smallness depending only on $c_0, C_0, \lambda$) so that $C\Big(\|\rho'\|_{L^\infty} + \|\rho''\|_{L^\infty}\Big) \leq c_0\lambda/2$ to obtain

$$\begin{aligned}
\frac{d}{dt}\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2 \leq -\frac{c_0}{8N}&\sum_{i=1}^{N}\Big[N\left(\bar{x}^{i+1} - \bar{x}^i\right) - N\left(X^{i+1} - X^i\right)\Big]^2\\
&+ \frac{C}{N}\varepsilon_0\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2 + \frac{2}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)R^i.
\end{aligned}$$

We now use the discrete Poincaré inequality (see Lemma 2.3.5) to get

$$\frac{1}{2}\sum_{i=1}^{N}\Big[N\left(\bar{x}^{i+1} - \bar{x}^i\right) - N\left(X^{i+1} - X^i\right)\Big]^2 \geq \sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2,$$

so that assuming $\varepsilon_0$ small enough we conclude

$$\begin{aligned}\frac{d}{dt}\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2 &\leq -\frac{c_0}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2 + \frac{C}{N}\varepsilon_0\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2 \\ &\quad + \frac{2}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)R^i \\ &\leq -\frac{2c_0}{3N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2 + \frac{2}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)R^i\end{aligned}$$

Finally, using the bound

$$2\left(\bar{x}^i - X^i\right)R^i \leq \epsilon(\bar{x}^i - X^i)^2 + \frac{1}{\epsilon}|R^i|^2$$

with $\epsilon := 2c_0/3$, and recalling that $|R^i| \leq \hat{C}/N^2$ we conclude

$$\frac{d}{dt}\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2 \leq -\frac{c_0}{6N}\sum_{i=1}^{N}\left(\bar{x}^i - X^i\right)^2 + \frac{3}{2c_0}\left(\frac{\hat{C}}{N^2}\right)^2.$$

Integrating this differential inequality over $[T_1, t]$ with $t \leq T_2$, by Gronwall Lemma we obtain

$$\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}^i(t) - X^i(t)\right)^2 \leq e^{-\bar{c}(t-T_1)}\frac{1}{N}\sum_{i=1}^{N}\left(\bar{x}^i(T_1) - X^i(T_1)\right)^2 + \bar{C}\left(\frac{\hat{C}}{N^2}\right)^2$$

for some constants $\bar{c}, \bar{C} > 0$ depending only on $c_0$, as desired. □

### 2.4.2 The convergence results

Combining the results in the previous sections, we can now prove that if a continuous and a discrete solution are close up to $1/N^2$ at time zero, then they remain close for all time. As one can see from the proof, it is crucial that the discrete scheme has a error of order $\frac{1}{N^2}$ (see Lemma 2.4.1).

**Theorem 2.4.4** (Consistency)**.** *Let $\bar{x}^i$ be a solution of the ODE* (2.4.5)*, and let $X^i$ be as in* (2.4.3)*. Assume that $X_0 \in C^{4,\alpha}([0,1])$, $X_0(0) = 0$, $X_0(1) = 1$, and that $a_0 \leq \partial_\theta X_0 \leq A_0$ for some positive constants $a_0, A_0$. Also, suppose that*

$$|X^i(0) - \bar{x}^i(0)| \leq \frac{C'}{N^2} \qquad \forall\, i = 1, \ldots, N. \tag{2.4.8}$$

*for some positive constant $C'$.*

*Then, there exists* $\varepsilon_1 \equiv \varepsilon_1\big(a_0, A_0, \|\rho\|_{C^{3,\alpha}([0,1])}, \|X_0\|_{C^{4,\alpha}([0,1])}\big) > 0$ *such that, if* $\|\rho'\|_\infty + \|\rho''\|_\infty \leq \varepsilon_1$ *we have*

$$\frac{1}{N}\sum_{i=1}^{N}\big(\bar{x}^i(t) - X^i(t)\big)^2 \leq \frac{\bar{\bar{C}}}{N^4} \qquad \forall\, t \in [0,\infty).$$

*Proof.* The idea of the proof is the following: we want to prove the discrete gradient flow and the continuous one are $L^2$ close for all times. This is exactly what is claimed in Lemma 2.4.3 which, on the other hand, is based on the assumption $\frac{c_0}{N} \leq \bar{x}^i(t) - \bar{x}^{i-1}(t) \leq \frac{C_0}{N}$, $c_0, C_0 \in \mathbb{R}^+$. Unfortunately, a priori, these assumptions may not hold for every time. However, by carefully combining Lemmas 2.4.2 and 2.4.3 by an induction argument, we can show that these assumptions actually holds for all times.

*Basis for the induction.* First we observe that, by Proposition 2.2.3, there exist two positive constants $a$ and $A$ such that

$$a \leq \partial_\theta X(t) \leq A \qquad \forall\, t \geq 0. \tag{2.4.9}$$

Recalling the definition of $X^i$ in (2.4.3), we can infer the following inequalities at the discrete level:

$$\frac{a}{N} \leq X^i(t) - X^{i-1}(t) \leq \frac{A}{N} \qquad \forall\, t \geq 0,\ \forall\, i. \tag{2.4.10}$$

Let us now focus on the assumption

$$\frac{c_0}{N} \leq \bar{x}^i(t) - \bar{x}^{i-1}(t) \leq \frac{C_0}{N}, \qquad c_0, C_0 \in \mathbb{R}^+.$$

Using Lemma 2.4.2 we have

$$|\bar{x}^i(t) - X^i(t)| \leq |\bar{x}^i(0) - X^i(0)| + \eta t \qquad \forall\, t \in [0,T].$$

Keeping in mind the definition of $\eta$ and (2.4.8) we have

$$|\bar{x}^i(t) - X^i(t)| \leq \frac{C'}{N^2} + \frac{3\hat{C}}{N^2}t + \frac{C'}{N^2}\frac{t}{T} \qquad \forall\, t \in [0,T],$$

so by the triangular inequality and (2.4.10) we obtain

$$\frac{a}{N} - 2\bigg(\frac{2C'}{N^2} + \frac{3\hat{C}}{N^2}T\bigg) \leq \bar{x}^i(t) - \bar{x}^{i-1}(t) \leq \frac{A}{N} + 2\bigg(\frac{2C'}{N^2} + \frac{3\hat{C}}{N^2}T\bigg)$$

for $t \in [0,T]$. In particular, by choosing $N$ large enough (depending only on $a, A, \hat{C}, C', T$), we can ensure that

$$\frac{a}{2N} \le \bar{x}^i(t) - \bar{x}^{i-1}(t) \le \frac{2A}{N} \qquad \forall\, t \in [0,T]. \tag{2.4.11}$$

*Inductive step.* Our goal is to show that if the above property holds for all $t \in [0, \alpha T]$ then it holds for all $t \in [0, (\alpha+1)T]$. Let us apply Lemma 2.4.3 on $[0, \alpha T]$ and (2.4.8) to get

$$\begin{aligned}\frac{1}{N}\sum_{i=1}^{N}\big(\bar{x}^i(t) - X^i(t)\big)^2 &\le \frac{e^{-\bar{c}t}}{N}\sum_{i=1}^{N}\big(\bar{x}^i(0) - X^i(0)\big)^2 + \bar{C}\bigg(\frac{\hat{C}}{N^2}\bigg)^2 \\ &\le \frac{\bar{\bar{C}}}{N^4} \qquad \forall\, t \in [0, \alpha T]\end{aligned}$$

for some constant $\bar{\bar{C}}$ depending only on $\bar{C}, \hat{C}, C'$. Hence, since

$$|\bar{x}^i(t) - X^i(t)| \le \sqrt{\sum_{i=1}^{N}\big(\bar{x}^i(t) - X^i(t)\big)^2} \qquad \forall\, t \in [0, \alpha T],\ \forall\, i,$$

we obtain in particular,

$$|\bar{x}^i(\alpha T) - X^i(\alpha T)| \le \sqrt{\frac{\bar{\bar{C}}}{N^3}} \qquad \forall\, i = 1, \dots, N.$$

Applying again Lemma 2.4.2 with $\alpha T$ as initial time, we now get

$$\begin{aligned}|\bar{x}^i(\alpha T + t) - X^i(\alpha T + t)| &\le |\bar{x}^i(\alpha T) - X^i(\alpha T)| + \eta \alpha T \\ &\le \sqrt{\frac{\bar{\bar{C}}}{N^3}} + \frac{3\hat{C}}{N^2}\alpha T + \sqrt{\frac{\bar{\bar{C}}}{N^3}}\frac{t}{\alpha T} \qquad \forall\, t \in [0, \alpha T].\end{aligned}$$

Hence, by (2.4.10) and triangle inequality,

$$\begin{aligned}\frac{a}{N} - 2\bigg(2\sqrt{\frac{\bar{\bar{C}}}{N^3}} - \frac{3\hat{C}}{N^2}\alpha T\bigg) &\le \bar{x}^i(t) - \bar{x}^{i-1}(t) \\ &\le \frac{A}{N} + 2\bigg(2\sqrt{\frac{\bar{\bar{C}}}{N^3}} - \frac{3\hat{C}}{N^2}\alpha T\bigg) \qquad \forall\, t \in [\alpha T, (\alpha+1)T].\end{aligned} \tag{2.4.12}$$

Then, if $N$ is big enough so that

$$2\sqrt{\frac{\bar{\bar{C}}}{N^3}} + \frac{3\hat{C}}{N^2}\alpha T \leq \frac{a}{4N} \tag{2.4.13}$$

we have

$$\frac{a}{2N} \leq \bar{x}^i(t) - \bar{x}^{i-1}(t) \leq \frac{2A}{N} \qquad \forall\, t \in [\alpha T, (\alpha+1)T].$$

Recalling the inequality (2.4.11) we get

$$\frac{a}{2N} \leq \bar{x}^i(t) - \bar{x}^{i-1}(t) \leq \frac{2A}{N} \qquad \forall\, t \in [0, (\alpha+1)T].$$

This concludes the inductive step and, in particular, Lemma 2.4.3 applied on $[0,\infty)$ proves the desired estimate for $N \geq N_0$ for some large number $N_0$.

Notice that the case $N \leq N_0$ is trivial since (using that $0 \leq \bar{x}^i, X^i \leq 1$)

$$\frac{1}{N}\sum_{i=1}^{N} \left(\bar{x}^i(t) - X^i(t)\right)^2 \leq 1 \leq \frac{N_0^4}{N^4} \qquad \forall\, t \in [0,\infty).$$

□

### The Eulerian description

In order to get a convergence result in Eulerian variable, we will also need a full stability result in $L^2$ in the continuous case. The following result holds:

**Proposition 2.4.5.** *Assume that $\rho : [0,1] \to (0,\infty)$ is a periodic probability density of class $C^2$ and let $X_1, X_2$ be two solutions of the equation* (2.4.2) *satisfying* (2.1.4) *and*

$$0 < c_0 \leq \partial_\theta X_i(0,\theta) \leq C_0, \qquad i = 1,2. \tag{2.4.14}$$

*There exists $\varepsilon_0 \equiv \varepsilon_0(c_0, C_0)$ as in Lemma 2.4.3 such that, if $\|\rho'\|_{L^\infty} + \|\rho''\|_{L^\infty} \leq \varepsilon_0$, then*

$$\int_0^1 |X_1(t,\theta) - X_2(t,\theta)|^2\, d\theta \leq \left(\int_0^1 |X_1(0,\theta) - X_2(0,\theta)|^2\, d\theta\right) e^{-\bar{c}t} \qquad \forall\, t \geq 0$$

*for some $\bar{c} \equiv \bar{c}(c_0)$.*

*Proof.* The proof of this result follows the lines of the proof of Proposition 2.3.1, with the difference that we have to get rid of the extra terms using the smallness of $\|\rho'\|_{L^\infty} + \|\rho''\|_{L^\infty}$. Also, this result could also be obtained as a consequence of Lemma 2.4.3 letting $N \to \infty$. However, since the proof is relatively short, we give it for the convenience of the reader.

We begin by noticing that since $\int_0^1 \rho(x)dx = 1$, if $\|\rho'\|_\infty$ is sufficiently small it follows that $1/2 \le \rho \le 2$, so the monotonicity condition (2.4.14) implies that

$$0 < c_1 \le \partial_\theta X_i(t) \le C_1, \qquad i = 1, 2, \text{ for all}\, t \ge 0 \tag{2.4.15}$$

for some constants $c_1, C_1$ depending only on $c_0, C_0$ (see Proposition 2.2.3). Also, we notice that (2.4.2) can be equivalently rewritten as

$$\partial_t X = \frac{1}{4}\partial_\theta\big(\rho(X)(\partial_\theta X)^2\big) - \frac{1}{12}\rho'(X)(\partial_\theta X)^3.$$

Then, since $X_2 - X_1$ vanishes at the boundary, we compute

$$\begin{aligned}
&\frac{d}{dt}\int_0^1 |X_1 - X_2|^2\, d\theta \\
&= \frac{1}{2}\int_0^1 (X_1 - X_2)\Big(\partial_\theta\big(\rho(X_1)(\partial_\theta X_1)^2\big) - \partial_\theta\big(\rho(X_2)(\partial_\theta X_2)^2\big)\Big)\, d\theta \\
&\qquad - \frac{1}{6}\int_0^1 (X_1 - X_2)\Big(\rho'(X_1)(\partial_\theta X_1)^3 - \rho'(X_2)(\partial_\theta X_2)^3\Big)\, d\theta \\
&= -\frac{1}{2}\int_0^1 \partial_\theta(X_1 - X_2)\Big(\big(\rho(X_1)(\partial_\theta X_1)^2\big) - \big(\rho(X_2)(\partial_\theta X_2)^2\big)\Big)\, d\theta \\
&\qquad - \frac{1}{6}\int_0^1 (X_1 - X_2)\Big(\rho'(X_1)(\partial_\theta X_1)^3 - \rho'(X_2)(\partial_\theta X_2)^3\Big)\, d\theta \\
&= -\frac{1}{2}\int_0^1 \rho(X_2)\partial_\theta(X_1 - X_2)\big((\partial_\theta X_1)^2 - (\partial_\theta X_2)^2\big)\, d\theta \\
&\qquad - \frac{1}{2}\int_0^1 [\rho(X_1) - \rho(X_2)]\,\partial_\theta(X_1 - X_2)\,(\partial_\theta X_1)^2\, d\theta \\
&\qquad - \frac{1}{6}\int_0^1 \rho'(X_2)(X_1 - X_2)\big((\partial_\theta X_1)^3 - (\partial_\theta X_2)^3\big)\, d\theta \\
&\qquad - \frac{1}{6}\int_0^1 [\rho'(X_1) - \rho'(X_2)]\,\rho(X_1 - X_2)\,(\partial_\theta X_1)^3\, d\theta \\
&=: T_{1,1} + T_{1,2} + T_{2,1} + T_{2,2}.
\end{aligned}$$

Recalling that $1/2 \le \rho$, using (2.4.15) we get

$$\begin{aligned} T_{1,1} &\le -\frac{1}{2}\int_0^1 \rho(X_2)\big(\partial_\theta(X_1-X_2)\big)^2\big((\partial_\theta X_1)+(\partial_\theta X_2)\big)\,d\theta \\ &\le -\frac{c_1}{2}\int_0^1 \big(\partial_\theta(X_1-X_2)\big)^2\,d\theta. \end{aligned}$$

Using again (2.4.15) we bound

$$\begin{aligned} |T_{1,2}| &\le \frac{C_1^2}{2}\|\rho'\|_\infty \int_0^1 |X_1-X_2|\,|\partial_\theta X_1-\partial_\theta X_2|\,d\theta \\ &\le \frac{C_1^2}{2}\|\rho'\|_\infty \int_0^1 (X_1-X_2)^2\,d\theta + \frac{C_1^2}{2}\|\rho'\|_\infty \int_0^1 (\partial_\theta X_1-\partial_\theta X_2)^2\,d\theta, \end{aligned}$$

$$\begin{aligned} |T_{2,1}| &\le \frac{C_1^2}{2}\|\rho'\|_\infty \int_0^1 |X_1-X_2|\,|\partial_\theta X_1-\partial_\theta X_2|\,d\theta \\ &\le \frac{C_1^2}{2}\|\rho'\|_\infty \int_0^1 (X_1-X_2)^2\,d\theta + \frac{C_1^2}{2}\|\rho'\|_\infty \int_0^1 (\partial_\theta X_1-\partial_\theta X_2)^2\,d\theta, \end{aligned}$$

$$|T_{2,2}| \le \frac{C_1^3}{6}\|\rho''\|_\infty \int_0^1 (X_1-X_2)^2\,d\theta.$$

Hence, combining all together, if both $\|\rho'\|_\infty$ and $\|\rho''\|_\infty$ are sufficiently small, using Poincaré inequality (see Lemma 2.3.5 and let $N \to \infty$), we obtain

$$\begin{aligned} \frac{d}{dt}\int_0^1 |X_1-X_2|^2\,d\theta &\le -\frac{c_1}{4}\int_0^1 \big(\partial_\theta(X_1-X_2)\big)^2\,d\theta \\ &\quad + C\big(\|\rho'\|_\infty+\|\rho''\|_\infty\big)\int_0^1 (X_1-X_2)^2\,d\theta \\ &\le -\frac{c_1}{2}\int_0^1 (X_1-X_2)^2\,d\theta \\ &\quad + C\big(\|\rho'\|_\infty+\|\rho''\|_\infty\big)\int_0^1 (X_1-X_2)^2\,d\theta \\ &\le -\frac{c_1}{4}\int_0^1 (X_1-X_2)^2\,d\theta, \end{aligned}$$

and the result follows by Gronwall's inequality. □

**Theorem 2.4.6.** *Let $\rho : [0,1] \to (0,\infty)$ be a periodic probability density of class $C^{3,\alpha}$. Let $x^i$ be a solution of the discrete gradient flow starting from an initial datum satisfying*

$$\left| x^i(0) - X_0\left(0, \frac{i-1/2}{N}\right) \right| \leq \frac{C'}{N^2} \qquad \forall\, i = 1, \ldots, N,$$

*where $X_0 \in C^{4,\alpha}([0,1])$, $X_0(0) = 1$, $X_0(1) = 1$, and $0 < c_0 \leq \partial_\theta X_0 \leq C_0$. Then there exist three constants $\varepsilon_1 \equiv \varepsilon_1\big(c_0, C_0, \|\rho\|_{C^{3,\alpha}([0,1])}, \|X_0\|_{C^{4,\alpha}([0,1])}\big)$ as in Theorem 2.4.4; $\bar{\bar{c}} \equiv \bar{\bar{c}}(c_0) > 0, \bar{\bar{C}} \equiv \bar{\bar{C}}(c_0) > 0$, such that,*

$$MK_1(\mu_t^N, \gamma\rho^{1/3}\, d\theta) \leq \bar{\bar{C}}\, e^{-\bar{\bar{c}}t/N^3} + \frac{\bar{\bar{C}}}{N} \qquad \forall\, t \geq 0,$$

*where*

$$\gamma := \frac{1}{\int_0^1 \rho^{1/3}(x)dx}$$

*provided that $\|\rho'\|_\infty + \|\rho''\|_\infty \leq \varepsilon_1$. In particular*

$$MK_1(\mu_t^N, \gamma\rho^{1/3}\, d\theta) \leq \frac{\bar{\bar{C}}}{N} \qquad \textit{for all } t \geq \frac{N^3 \log N}{\bar{\bar{c}}}.$$

*Proof.* Let $\bar{X}$ satisfy

$$\partial_\theta \bar{X} = \frac{1}{\gamma\rho^{1/3} \circ \bar{X}}, \quad \bar{X}(0) = 0.$$

Then $\bar{X}$ is a stationary solution of (2.4.2) satisfying also the boundary condition (2.1.4), hence by Proposition 2.4.5 we deduce that

$$\int_0^1 |X(t) - \bar{X}|^2\, d\theta \leq Ce^{-\bar{c}t},$$

where $X(t)$ is the solution of (2.4.2) starting from $X_0$. We then apply Theorem 2.4.4 to deduce that

$$\frac{1}{N}\sum_{i=1}^N \big(x^i(t) - X^i(t/N^3)\big)^2 \leq \frac{\bar{\bar{C}}}{N^4} \qquad \forall\, t \in [0,\infty),$$

where $X^i(t) := X\left(t, \frac{i-1/2}{N}\right)$. Combining these two estimates and observing that $\bar{X}_\# d\theta = \gamma\, \rho^{1/3}\, d\theta$, the result follows by arguing as in the proof of Theorem 2.3.6. □

## 2.5 From the discrete to the continuous case

In order to obtain a continuous version of the functional

$$F_{N,r}(x^1,\dots,x^N) = \int_0^1 \min_{1\le i\le N} |x^i - y|^r \rho(y)\,dy,$$

with $0 \le x^1 \le \ldots \le x^N \le 1$, we define

$$x^{i+1/2} := \frac{x^i + x^{i+1}}{2},$$

where by convention $x^0 = 0$ and $x^{N+1} = 1$. Then the expression for the minimum becomes

$$\min_{1\le j\le N} |y - x^j|^r = \begin{cases} |y - x^i|^r & \text{for } y \in (x^{i-1/2}, x^{i+1/2}), \\ |y|^r & \text{for } y \in (0, x^{1/2}), \\ |y-1|^r & \text{for } y \in (x^{N+1/2}, 1), \end{cases}$$

and $F_{N,r}$ is given by

$$F_{N,r}(x^1,\dots,x^N) = \sum_{i=1}^N \int_{x^{i-1/2}}^{x^{i+1/2}} |y - x^i|^r \rho(y)dy + \int_0^{x^{1/2}} |y|^r \rho(y)dy + \int_{x^{N+1/2}}^1 |y-1|^r \rho(y)dy.$$

Assume that

$$x^i = X\Big(\frac{i-1/2}{N}\Big), \qquad i = 1,\dots,N$$

with $X : [0,1] \to [0,1]$ a smooth non-decreasing map. Then a Taylor expansion yields

$$F_{N,r}(x^1,\dots,x^N) = \frac{C_r}{N^r}\int_0^1 \rho(X(\theta))|\partial_\theta X(\theta)|^{r+1} d\theta + O\Big(\frac{1}{N^{r+1}}\Big),$$

where $C_r = \frac{1}{2^r(r+1)}$ and $O\left(\frac{1}{N^{r+1}}\right)$ depends on the smoothness of $\rho$ and $X$ (for instance, $\rho \in C^1$ and $X \in C^2$ is enough). Hence

$$N^r F_{N,r}(x^1,\dots,x^N) \longrightarrow C_r \int_0^1 \rho(X(\theta))|\partial_\theta X(\theta)|^{r+1} d\theta := \mathcal{F}[X]$$

as $N \to \infty$.

## 2.6 The Hessian of $\mathcal{F}[X]$

Assume $\lambda \le \rho \le \frac{1}{\lambda}$, and let $X, Y \in L^2([0,1])$ with $0 \le c \le \partial_\theta X \le C$ and $|\partial_\theta Y| \le C$. Also, assume that

$$X(0) = 0, \qquad X(1) = 1, \qquad Y(0) = 0, \qquad Y(1) = 0.$$

Then

$$\begin{aligned} D^2\mathcal{F}[X](Y,Y) &= 6\int_0^1 \rho(X)\,\partial_\theta X\,(\partial_\theta Y)^2\,d\theta \\ &\quad + 6\int_0^1 \rho'(X)\,(\partial_\theta X)^2\,(\partial_\theta Y)\,Y\,d\theta + \int_0^1 \rho''(X)\,(\partial_\theta X)^3\,Y^2\,d\theta. \end{aligned}$$

### 2.6.1 Convexity under a smallness assumption on $\rho'$ and $\rho''$

We want to prove that the Hessian of $\mathcal{F}$ is positive definite provided that

$$\|\rho'\|_\infty + \|\rho''\|_\infty \ll 1.$$

We first observe that

$$\begin{aligned} D^2\mathcal{F}[X](Y,Y) = \frac{d^2}{d\varepsilon^2}\bigg|_{\varepsilon=0} \mathcal{F}(X+\varepsilon Y) &\ge 6\lambda\,c\int_0^1 (\partial_\theta Y)^2\,d\theta \\ &\quad - 6C^2\|\rho'\|_\infty \int_0^1 |\partial_\theta Y|\,|Y|\,d\theta - C^3\|\rho''\|_\infty \int_0^1 Y^2\,d\theta. \end{aligned}$$

Observe that if both $\rho'$ and $\rho''$ are small, we can control both the second and third term by the first one using Young and Poincaré inequalities. In particular one sees that the Hessian is positive at "points" $X$ which are uniformly monotone and Lipschitz [4].

Indeed, using Young's inequality,

$$\begin{aligned} D^2\mathcal{F}[X](Y,Y) &\ge -C^3\|\rho''\|_\infty \int_0^1 Y^2\,d\theta + 6\lambda\,c\int_0^1 (\partial_\theta Y)^2\,d\theta \\ &\quad - 3C^2\|\rho'\|_\infty \left[\int_0^1 Y^2\,d\theta + \int_0^1 (\partial_\theta Y)^2\,d\theta\right]. \end{aligned}$$

[4] Recall that $0 \le c \le \partial_\theta X \le C$

Hence, if $3\|\rho'\|C^2 \leq 3\lambda c$ we have

$$
\begin{aligned}
D^2\mathcal{F}[X](Y,Y) &\geq 3\lambda\, c \int_0^1 (\partial_\theta Y)^2\, d\theta \\
&\qquad - \left[C^3\|\rho''\|_\infty + 3C^2\|\rho'\|_\infty\right] \int_0^1 Y^2\, d\theta \\
&\geq 6\lambda c \int_0^1 Y^2\, d\theta - \left[C^3\|\rho''\|_\infty + 3C^2\|\rho'\|_\infty\right] \int_0^1 Y^2\, d\theta,
\end{aligned}
$$

where for the second inequality we used Poincaré (see for instance Lemma 2.3.5 and let $N \to \infty$). Thus, if $3\lambda c > C^3\|\rho''\|_\infty + 3C^2\|\rho'\|_\infty$ it follows that the Hessian of $\mathcal{F}$ is positive definite.

### 2.6.2 Lack of convexity without a smallness assumption

In this Section it will be convenient to specify the dependence of $\mathcal{F}$ on $\rho$, so we denote

$$
\mathcal{F}_\rho(X) := \int_0^1 \rho(X)\, |\partial_\theta X|^3\, d\theta.
$$

To build a counterexample, we consider $X(t,\theta) = \theta$.

Recalling the formula for the Hessian of $\mathcal{F}$, we see that for any smooth density $\bar\rho$ and for any smooth function $Y$,

$$
D^2\mathcal{F}_{\bar\rho}(X)[Y,Y] = 6\int_0^1 \bar\rho\,(\partial_\theta Y)^2\, d\theta + 6\int_0^1 \bar\rho'\,\partial_\theta Y\, Y\, d\theta + \int_0^1 \bar\rho''\, Y^2\, d\theta.
$$

Integrating by parts we have

$$
\begin{aligned}
D^2\mathcal{F}_{\bar\rho}(X)[Y,Y] &= 6\int_0^1 \bar\rho\,(\partial_\theta Y)^2\, d\theta - 6\int_0^1 \bar\rho\,(\partial_\theta Y)^2 - 6\int_0^1 \bar\rho\,\partial_\theta^2 Y\, Y\, d\theta \\
&\qquad + 2\int_0^1 \bar\rho\left[(\partial_\theta Y)^2 + \partial_\theta^2 Y\, Y\right] d\theta \\
&= 2\int_0^1 \bar\rho\,(\partial_\theta Y)^2\, d\theta - 4\int_0^1 \bar\rho\,\partial_\theta^2 Y\, Y\, d\theta.
\end{aligned}
$$

We now fix $\varepsilon \in (0, 1/8)$ be a small number and define

$$
\bar\rho(\theta) := \begin{cases} 1 & \text{for } \theta \in \left[\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon\right] \\ 0 & \text{for } \theta \in [0,1] \setminus \left[\frac{1}{2}-\varepsilon, \frac{1}{2}+\varepsilon\right]. \end{cases}
$$

Also, let $Y(t,\theta)$ a Lipschitz function, compactly supported in $(0,1)$, that is smooth on $(0,1/2)\cup(1/2,1)$ and coincides with $|\theta-\frac{1}{2}|+1$ in $\left[\frac{1}{2}-\varepsilon,\frac{1}{2}+\varepsilon\right]$.

Since $\bar\rho$ and $Y$ are not smooth, we first extend both of them by periodicity on the whole real line and define $\rho_\delta := \bar\rho * \varphi_\delta$ and $Y_\delta := Y * \varphi_\delta$, with

$$\varphi_\delta(\theta) = \frac{\exp^{-\frac{|\theta|^2}{2\delta}}}{\sqrt{2\pi\delta}}.$$

Then

$$D^2\mathcal{F}_{\rho_\delta}(X)[Y_\delta,Y_\delta] = 2\int_0^1 \rho_\delta\,(\partial_\theta Y_\delta)^2\,d\theta - 4\int_0^1 \rho_\delta\,\partial_\theta^2 Y_\delta\,Y_\delta\,d\theta.$$

Noticing that

$$\rho_\delta \to \bar\rho \quad \text{in } L^1, \qquad \rho_\delta \to 1 \quad \text{uniformly in } [1/2-\varepsilon/2, 1/2+\varepsilon/2],$$

$$Y_\delta \to Y \quad \text{uniformly}, \qquad \partial_\theta Y_\delta \to \partial_\theta Y \quad \text{a.e.}, \qquad \partial_\theta^2 Y_\delta \rightharpoonup 2\delta_{1/2},$$

we see that

$$D^2\mathcal{F}_{\rho_\delta}(X)[Y_\delta,Y_\delta] \to 2\int_{\frac{1}{2}-\varepsilon}^{\frac{1}{2}+\varepsilon} (\partial_\theta Y)^2\,d\theta - 8Y\left(\frac{1}{2}\right) = 4\varepsilon - 8 < 0 \qquad \text{as } \delta\to 0.$$

In particular, by choosing $\delta > 0$ sufficiently small, we have obtained that the Hessian of $\mathcal{F}_{\rho_\delta}$ in the direction $Y_\delta$ is negative when $X(\theta)=\theta$ and $\rho_\delta \in C^\infty([0,1])$ and satisfies $1 \ge \rho_\delta > 0$.

# Chapter 3

# Asymptotic quantization for probability measures on Riemannian manifolds

[1]

## 3.1 Introduction

The problem of quantization of a $d$-dimensional probability distribution deals with constructive methods to find atomic probability measures supported on a finite number of points, which best approximate a given diffuse probability measure. The quality of this approximation is usually measured in terms of the Wasserstein metric, and up to now this problem has been studied in the flat case and on compact manifolds.

The quantization problem arises in several contexts and has applications in signal compression, pattern recognition, speech recognition, stochastic processes, numerical integration, optimal location of service centers, and kinetic theory. For a detailed exposition and a complete list of references, we refer to the monograph [47] and references therein. In this Chapter we study it for probability measures on general Riemannian manifolds. Apart from its own interest, this has several natural applications.

To mention one, in order to find a good approximation of a convex body by polyhedra

[1] This chapter is based on [65].

one may look for the best approximation of the curvature measure of the convex body by discrete measures [52].

To give another natural motivation, let us present the so-called *location problem.* If we want to plan the location of a certain number of grocery stores to meet the demands of the population in a city, we need to chose the optimal location and size of the stores with respect to the distribution of the population. The classical case on $\mathbb{R}^d$ corresponds to the situation of a city on a flat land. Now consider the possibility that the geographical region, instead of being flat, is situated either at the bottom of a valley, or at a pass in the mountains. Then the Wasserstein distance reflects this geography, by depending on the distance $d$ as measured along a spherical cap in the case of the valley or along a piece of a saddle in the case of the mountain pass. It follows by our results that for a city in a valley or on a mountain top the optimal location problem converge as in the flat case, while for a city located on a pass in the mountains the effect of negative curvature badly influences the quality of the approximation. Hence, our results display how geometry and geography can affect the optimal location problem.

We now introduce the setting of the problem. Let $(\mathcal{M}, g)$ be a complete Riemannian manifold, and fixed $r \geq 1$, consider $\mu$ a probability measure on $\mathcal{M}$. Given $N$ points $x^1, \ldots, x^N \in \mathcal{M}$, one wants to find the best approximation of $\mu$, in the Wasserstein distance $W_r$, by a convex combination of Dirac masses centered at $x^1, \ldots, x^N$. Hence one minimizes

$$\inf\left\{W_r\left(\sum_i m_i\delta_{x^i}, \mu\right)^r \,:\, m_1, \ldots, m_N \geq 0,\ \sum_i m_i = 1\right\},$$

with

$$W_r(\nu_1, \nu_2) := \inf\left\{\left(\int_{\mathcal{M}\times\mathcal{M}} d(x, y)^r d\gamma(x, y)\right)^{1/r} \,:\, (\pi_1)_\#\gamma = \nu_1,\ (\pi_2)_\#\gamma = \nu_2\right\},$$

where $\gamma$ varies among all probability measures on $\mathcal{M} \times \mathcal{M}$, $\pi_i : \mathcal{M} \times \mathcal{M} \to \mathcal{M}$ $(i = 1, 2)$ denotes the canonical projection onto the $i$-th factor, and $d(x, y)$ denotes the Riemannian distance. The best choice of the masses $m_i$ is explicit and can be expressed in terms of the so-called *Voronoi cells* [47, Chapter 1.4]. Also, as shown for instance in [47, Chapter 1, Lemmas 3.1 and 3.4], the following identity holds:

$$\inf\left\{W_r\left(\sum_i m_i\delta_{x^i}, \mu\right)^r \,:\, m_1, \ldots, m_N \geq 0,\ \sum_i m_i = 1\right\} = F_{N,r}(x^1, \ldots, x^N),$$

where

$$F_{N,r}(x^1, \dots, x^N) := \int_{\mathcal{M}} \min_{1 \le i \le N} d(x^i, y)^r \, d\mu(y).$$

Hence, the main question becomes: Where are the "optimal points" $(x^1, \dots, x^N)$ located? To answer to this question, at least in the limit as $N \to \infty$, let us first introduce some definitions.

**Definition 3.1.1.** *Let $\mu$ be a probability measure on $\mathcal{M}$, $N \in \mathbb{N}$ and $r \ge 1$. Then, we define the $N$-*th quantization error of order *$r$, $V_{N,r}(\mu)$ as follows:*

$$V_{N,r}(\mu) := \inf_{\alpha \subset \mathcal{M} : |\alpha| \le N} \int_{\mathcal{M}} \min_{a \in \alpha} d(a, y)^r \, d\mu(y), \tag{3.1.1}$$

*where $|\alpha|$ denotes the cardinality of a set $\alpha$.*

*Let us notice that, being the functional $F_{N,r}$ decreasing with respect to the number of points $N$, an equivalent definition of $V_{N,r}$ is:*

$$V_{N,r}(\mu) := \inf_{x^1, \dots, x^N \in \mathcal{M}} F_{N,r}(x^1, \dots, x^N).$$

Let us observe that the above definitions make sense for general positive measures with finite mass. In the sequel we will sometimes consider this class of measures in order to avoid renormalization constants.

A quantity that plays an important role in our result is the following:

**Definition 3.1.2.** *Let $dx$ be the Lebesgue measure and $\chi_{[0,1]^d}$ the characteristic function of the unit cube $[0,1]^d$. We set*

$$Q_r\left([0,1]^d\right) := \inf_{N \ge 1} N^{r/d} V_{N,r}\left(\chi_{[0,1]^d} dx\right).$$

As proved in [47, Theorem 6.2], $Q_r\left([0,1]^d\right)$ is a positive constant. The following result describe the asymptotic distribution of the minimizing configuration in $\mathbb{R}^d$, answering to our question in the flat case (see [25] and [47, Chapter 2, Theorems 6.2 and 7.5]):

**Theorem 3.1.3.** *Let $\mu = h\, dx + \mu^s$ be a probability measure on $\mathbb{R}^d$, where $\mu^s$ denotes the singular part of $\mu$. Assume that $\mu$ satisfies*

$$\int_{\mathbb{R}^d} |x|^{r+\delta} \, d\mu(x) < \infty. \tag{3.1.2}$$

*Then*

$$\lim_{N\to\infty} N^{r/d} V_{N,r}(\mu) = Q_r\left([0,1]^d\right)\left(\int_{\mathbb{R}^d} h^{d/(d+r)}\,dx\right)^{(d+r)/d}. \tag{3.1.3}$$

*In addition, if* $\mu^s \equiv 0$ *and* $x^1,\dots,x^N$ *minimize the functional* $F_{N,r} : (\mathbb{R}^d)^N \to \mathbb{R}^+$*, then*

$$\frac{1}{N}\sum_{i=1}^{N}\delta_{x^i} \rightharpoonup \frac{h^{d/d+r}}{\int_{\mathbb{R}^d} h^{d/d+r}(y)dy}\,dx \qquad \text{as } N\to\infty. \tag{3.1.4}$$

It is worth to mention that the problem of the quantization of measure has been studied also with a Γ-convergence approach in [15], [14], [17], and [80]. It is reasonable that, using the results in [80], the convergence of the empirical measure to a certain power of the measure $h$ in Theorem 3.1.3 holds whenever the measure $\mu$ has an absolutely continuous part. Nevertheless, we do not investigate this question since this Chapter is focused on the extension of the first statement in Theorem 3.1.3 to the case of probability measures on general Riemannian manifolds. Such a statement has been generalized to the case of absolutely continuous probability measures on compact Riemannian manifolds in [52].

Our aim here is twofold: we first give an alternative proof of Theorem 3.1.3 for general probability measures on compact manifolds, and then we extend it to arbitrary measures on non-compact manifolds. As we shall see, passing from the compact to the non-compact setting presents nontrivial difficulties. Indeed, while the compact case relies on a localization argument that allows one to mimic the proof in $\mathbb{R}^d$, the non-compact case requires additional new ideas. In particular one needs to find a suitable analogue of the moment condition (0.1.1) to control the growth at infinity of our given probability measure. We will prove that the needed growth assumption depends on the curvature of the manifold (and more precisely, on the size of the differential of the exponential map).

To state in detail our main result we need to introduce some notation: given a point $x_0 \in \mathcal{M}$, we can consider polar coordinates $(\rho,\vartheta)$ on $T_{x_0}\mathcal{M} \simeq \mathbb{R}^d$ induced by the constant metric $g_{x_0}$, where $\vartheta$ denotes a vector on the unit sphere $\mathbb{S}^{d-1}$ and $\rho$ is the the value of the norm in the metric $g_{x_0}$. Then, we can define the following quantity that measures the size of the differential of the exponential map when restricted to a sphere $\mathbb{S}^{d-1}_\rho \subset T_{x_0}\mathcal{M}$ of radius $\rho$:

$$A_{x_0}(\rho) := \sup_{v\in\mathbb{S}^{d-1}_\rho,\, w\in T_v\mathbb{S}^{d-1}_\rho,\, |w|_{x_0}=\rho} \left| d_v \exp_{x_0}[w]\right|_{\exp_{x_0}(v)}, \tag{3.1.5}$$

To prove asymptotic quantization, we shall impose an analogue of (3.1.2) which involves the above quantity.

**Theorem 3.1.4.** *Let* $(\mathcal{M}, g)$ *be a complete Riemannian manifold without boundary, and let* $\mu = h\,d\mathrm{vol} + \mu^s$ *be a probability measure on* $\mathcal{M}$*. Assume there exist a point* $x_0 \in \mathcal{M}$ *and* $\delta > 0$ *such that*

$$\int_{\mathcal{M}} d(x, x_0)^{r+\delta}\, d\mu(x) + \int_{\mathcal{M}} A_{x_0}\big(d(x, x_0)\big)^r\, d\mu(x) < \infty. \tag{3.1.6}$$

*Then* (3.1.3) *holds.*

Once this theorem is obtained, by the very same argument as in [47, Proof of Theorem 7.5] one gets the following:

**Corollary 3.1.5.** *Let* $(\mathcal{M}, g)$ *be a complete Riemannian manifold without boundary,* $\mu = h\,d\mathrm{vol}$ *an absolutely continuous probability measure on* $\mathcal{M}$ *and let* $x^1, \ldots, x^N$ *minimize the functional* $F_{N,r} : \mathcal{M}^{\otimes N} \to \mathbb{R}^+$*. Assume there exist a point* $x_0 \in \mathcal{M}$ *and* $\delta > 0$ *for which* (3.1.6) *is satisfied. Then* (3.1.4) *holds.*

Notice that the quantity $A_{x_0}$ is related to the curvature of $\mathcal{M}$, being linked to the size of the Jacobi fields (see for instance [73, Chapter 10]). In particular, if $\mathcal{M} = \mathbb{H}^d$ is the hyperbolic space then $A_{x_0}(\rho) = \sinh\rho$, while on $\mathbb{R}^d$ we have $A_{x_0}(\rho) = \rho$. Hence the above condition on $\mathbb{H}^d$ reads as

$$\left(1 + \int_{\mathbb{H}^d} d(x, x_0)^{r+\delta}\, d\mu(x) + \int_{\mathbb{H}^d} \sinh\big(d(x, x_0)\big)^r\, d\mu(x)\right) \approx \int_{\mathbb{H}^d} e^{r\, d(x,x_0)}\, d\mu(x),$$

and on $\mathbb{R}^d$ as

$$\left(1 + \int_{\mathbb{R}^d} d(x, x_0)^{r+\delta}\, d\mu(x) + \int_{\mathbb{R}^d} d(x, x_0)^r\, d\mu(x)\right) \approx \int_{\mathbb{R}^d} d(x, x_0)^{r+\delta}\, d\mu(x).$$

Hence (3.3.2) holds on $\mathbb{H}^d$ for any probability measure $\mu$ satisfying

$$\int_{\mathbb{H}^d} e^{r\, d(x,x_0)}\, d\mu(x) < \infty$$

for some $x_0 \in \mathbb{H}^d$, while on $\mathbb{R}^d$ we only need the finiteness of some $(r+\delta)$-moments of $\mu$, therefore recovering the assumption in Theorem 3.1.3. More in general, thanks to Rauch Comparison Theorem [73, Theorem 11.9], the size of the Jacobi fields on a manifold $\mathcal{M}$

with sectional curvature bounded from below by $-K$ ($K \geq 0$) is controlled by the Jacobi fields on the hyperbolic space with sectional curvature $-K$. Hence in this case

$$A_{x_0}(r) \leq \frac{\sinh(Kr)}{K}.$$

Since $\sinh(Kr) \approx e^{Kr}$ for $r \gg 1$, Theorem 3.1.4 yields the following:

**Corollary 3.1.6.** *Let $(\mathcal{M}, g)$ be a complete Riemannian manifold without boundary, and let $\mu = h\,d\mathrm{vol} + \mu^s$ be a probability measure on $\mathcal{M}$. Assume that the sectional curvature of $\mathcal{M}$ is bounded from below by $-K$ for some $K \geq 0$, and that there exist a point $x_0 \in \mathcal{M}$ and $\delta > 0$ such that*

$$\int_{\mathcal{M}} d(x, x_0)^{r+\delta}\, d\mu(x) + \int_{\mathcal{M}} e^{Kr\, d(x,x_0)}\, d\mu(x) < \infty.$$

*Then* (3.1.3) *holds. In addition, if $\mu^s \equiv 0$ and $x^1, \ldots, x^N$ minimize the functional $F_{N,r} : (\mathbb{R}^d)^N \to \mathbb{R}^+$, then* (3.1.4) *holds.*

Finally, we show that the moment condition (3.1.2) required on $\mathbb{R}^d$ is not sufficient to ensure the validity of the result on $\mathbb{H}^d$. Indeed we can provide the following counter example on $\mathbb{H}^2$.

**Theorem 3.1.7.** *There exists a measure $\mu$ on $\mathbb{H}^2$ such that*

$$\int_{\mathbb{H}^2} d(x, x_0)^p\, d\mu < \infty \qquad \forall\, p > 0,\, \forall\, x_0 \in \mathbb{H}^2,$$

*but*

$$N^{r/2} V_{N,r}(\mu) \to \infty \qquad \text{as } N \to \infty.$$

The chapter is structured as follows: first, in Section 3.2 we prove Theorem 3.1.4 for compactly supported probability measures. Then, in Section 3.3 we deal with the non-compact case concluding the proof of Theorem 3.1.4. Finally, in Section 3.4 we prove Theorem 3.1.7.

## 3.2 Proof of Theorem 3.1.4: the compact case

This section is concerned with the study of asymptotic quantization for probability distributions on compact Riemannian manifolds as the number $N$ of points tends to infinity. Although the problem depends a priori on the global geometry of the manifold (since $V_{N,r}$ involves the Riemannian distance), we shall now show how a localization argument allows us to prove the result.

### 3.2.1 Localization argument

Let $(\mathcal{M}, g)$ be a complete Riemannian manifold without boundary and let $\mu$ be a probability measure on $\mathcal{M}$. We consider $\{\mathcal{U}_i, \varphi_i\}_{i\in I}$ an atlas covering $\mathcal{M}$, and $\varphi_i : \mathcal{W}_i \to \mathbb{R}^d$ smooth charts, where $\mathcal{W}_i \supset\supset \mathcal{U}_i$ for all $i \in I$. As we shall see, in order to be able to split our measure as a sum of measures supported on smaller sets, we want to avoid the mass to concentrate on the boundary of the sets $\mathcal{U}_i$. Hence, up to slightly changing the sets $\mathcal{U}_i$, we may assume that

$$\mu(\partial\mathcal{U}_i) = 0 \qquad \forall\, i \in I. \tag{3.2.1}$$

We want to cover $\mathcal{M}$ with an atlas of disjoint sets, up to sets of $\mu$-measure zero. To do that we define

$$\mathcal{V}_i := \mathcal{U}_i \setminus \left(\bigcup_{j=1}^{i-1} \mathcal{U}_j\right).$$

Notice that we still have $\mathcal{V}_i \subset\subset \mathcal{W}_i$.

Given an open subset of $\mathbb{R}^d$, by [31, Lemma 1.4.2], we can cover it with a countable partition of half-open disjoint cubes such that the maximum length of the edges is a given number $\delta$. We now apply this observation to each open subset $\varphi_i(\mathring{\mathcal{V}}_i) \subset \mathbb{R}^d$ and we cover it with a family $\mathcal{G}_i$ of half-open cubes $\{Q_{i,j}\}_{j\in\mathbb{N}}$ with edges of length $\ell_j \le \delta$.

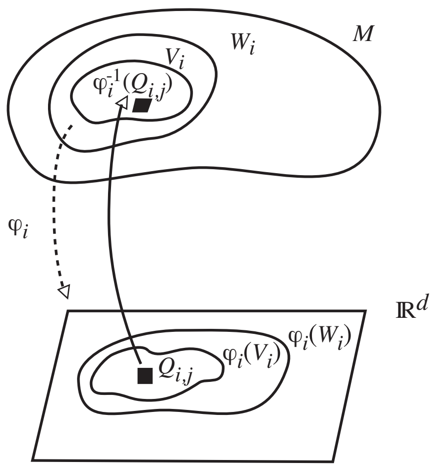


Figure 3.1: We use the map $\varphi_i^{-1} : \varphi_i(\mathcal{W}_i) \subset \mathbb{R}^d \to \mathcal{W}_i \subset \mathcal{M}$ to send the partition in cubes $Q_{i,j}$ of $\varphi_i(\mathcal{V}_i)$ on $\mathcal{M}$.

We notice that the "cubes" $\varphi_i^{-1}(Q_{i,j}) \subset \mathcal{M}$ are disjoint and

$$\bigcup_{i\in I} \bigcup_{Q_{i,j}\in\mathcal{G}_i} Q_{i,j} = \mathcal{M} \setminus \left(\bigcup_i \partial\mathcal{U}_i\right)$$

Since by (3.2.1) the set $\cup_i \partial\mathcal{U}_i$ has zero $\mu$-measure, we can decompose the measure $\mu$ as

$$\mu = \sum_{i\in I} \mu \mathbf{1}_{\mathcal{V}_i} = \sum_{i\in I} \sum_{Q_{i,j}\in\mathcal{G}_i} \mu \mathbf{1}_{\mathcal{V}_i\cap\varphi_i^{-1}(Q_{i,j})}.$$

We now set

$$\alpha_{ij} := \int_{\mathcal{V}_i\cap\varphi_i^{-1}(Q_{i,j})} d\mu, \qquad \mu_{ij} := \frac{\mu\mathbf{1}_{\mathcal{V}_i\cap\varphi_i^{-1}(Q_{i,j})}}{\alpha_{ij}},$$

so that

$$\mu = \sum_{ij} \alpha_{ij}\, \mu_{ij}, \qquad \int_{\mathcal{M}} d\mu_{ij} = 1, \qquad \mathrm{supp}(\mu_{ij}) \subset \mathcal{V}_i \cap \varphi_i^{-1}(Q_{i,j}),$$

where, to simplify the notation, in the above formula the indices $i, j$ implicitly run over $i \in I, Q_{i,j} \in \mathcal{G}_i$. We will keep using this convention also later on.

The idea is now the following: by choosing $\delta$ small enough, each measure $\mu_{ij}$ is supported on a very small set where the metric is essentially constant and allows us to reduce ourselves to the flat case and apply Theorem 3.1.3 to each of these measures. A "gluing argument" then gives the result when $\mu = \sum_{ij} \alpha_{ij}\mu_{ij}$ is compactly supported, $\alpha_{ij} \neq 0$ for at most finitely many indices, and $\mu_{ij}$ has constant density on $\varphi_i^{-1}(Q_{i,j})$. Finally, an approximation argument yields the result for general compactly supported measures.

### 3.2.2 The local quantization error

The goal of this Section is to understand the behavior of $V_{N,r}(\mu)$ when

$$\mu = \lambda \mathbf{1}_{\varphi^{-1}(Q)}\, d\mathrm{vol}, \tag{3.2.2}$$

where $\lambda := \frac{1}{\mathrm{vol}(\varphi^{-1}(Q))}$ (so that $\mu$ has mass 1), $Q$ is a $\delta$-cube in $\mathbb{R}^d$, $\varphi : \mathcal{W} \to \mathbb{R}^d$ is a diffeomorphism defined on a neighborhood $\mathcal{W} \subset \mathcal{M}$ of $\varphi^{-1}(Q)$.

We observe that, in the computation of $V_{N,r}(\mu)$, if the size of the cube is sufficiently small then we can assume that all the points belong to a $K\delta$-neighborhood of $\varphi^{-1}(Q)$, with $K$ a large universal constant, that we denote by $\mathcal{Z}_{K\delta}$. Indeed, if $\mathrm{dist}(b, \varphi^{-1}(Q)) > K\delta$ then

$$\mathrm{dist}(x, b) > \mathrm{dist}(x, y) \qquad \forall\, x, y \in \varphi^{-1}(Q),$$

which implies that, in the definition of $V_{N,r}(\mu)$, it is better to substitute $b$ with an arbitrary point inside $\varphi^{-1}(Q)$. Notice also that, if $\delta$ is small enough, $\mathcal{Z}_{K\delta}$ will be contained in the chart $\mathcal{W}$.

Hence, denoting by $\beta$ a family of $N$ points inside a $\mathcal{Z}_{K\delta}$, and by $\alpha$ a family of $N$ points inside $\varphi(\mathcal{Z}_{K\delta})$, we have

$$\begin{aligned} V_{N,r}(\mu) &= \inf_{\beta} \int_{\varphi^{-1}(Q)} \min_{b\in\beta} d(y, b)^r \, d\mu(y) \\ &= \lambda \inf_{\beta} \int_{\varphi^{-1}(Q)} \min_{b\in\beta} d(y, b)^r \, d\mathrm{vol}(y) \\ &= \lambda \inf_{\beta} \int_{Q} \min_{a\in\alpha} d\big(\varphi^{-1}(x), \varphi^{-1}(a)\big)^r \sqrt{\det\, g_{k\ell}(x)}\, dx. \end{aligned} \tag{3.2.3}$$

We now begin by showing that $d\big(\varphi^{-1}(x), \varphi^{-1}(a)\big)$ can be approximated with a constant metric. Recall that $\delta$ denotes the size of the cube $Q$. Also, we use the notation $g_{k\ell}$ to denote the metric in the chart, that is

$$\sum_{k\ell} g_{k\ell}(x) v^k v^\ell := g_{\varphi^{-1}(x)}\Big(d\varphi^{-1}(x)[v], d\varphi^{-1}(x)[v]\Big), \qquad \forall\, x \in \varphi(\mathcal{W}),\ v \in \mathbb{R}^d. \tag{3.2.4}$$

**Lemma 3.2.1.** *Let $p$ be the center of the cube $Q$ and let $A$ be the matrix with entries $A_{k\ell} := g_{k\ell}(p)$. There exists a universal constant $\hat{C}$ such that, for all $x \in Q$ and $a \in \varphi(\mathcal{Z}_{K\delta})$, it holds*

$$(1 - \hat{C}\delta)\, \langle A(x-a), x-a\rangle \le d\big(\varphi^{-1}(x), \varphi^{-1}(a)\big)^2 \le (1 + \hat{C}\delta)\, \langle A(x-a), x-a\rangle.$$

*Proof.* We begin by recalling that [2]

$$d\big(\varphi^{-1}(x), \varphi^{-1}(a)\big)^2 = \inf_{\substack{\gamma(0)=\varphi^{-1}(x),\\ \gamma(1)=\varphi^{-1}(a)}} \int_0^1 g_{\gamma(t)}\big(\dot\gamma(t), \dot\gamma(t)\big)\, dt.$$

[2]Recall that there are two equivalent definition of the distance between two points:

$$d(x, y) = \inf_{\substack{\gamma(0)=x,\\ \gamma(1)=y}} \int_0^1 \sqrt{g_{\gamma(t)}\big(\dot\gamma(t), \dot\gamma(t)\big)}\, dt = \inf_{\substack{\gamma(0)=x,\\ \gamma(1)=y}} \sqrt{\int_0^1 g_{\gamma(t)}\big(\dot\gamma(t), \dot\gamma(t)\big)\, dt}.$$

Here we will make use of both definitions.

Let $\bar{\gamma} : [0,1] \to \mathcal{M}$ denote a minimizing geodesic.[3] Then the speed of $\bar{\gamma}$ is constant and equal to the distance between the two points, that is

$$\|\dot{\bar{\gamma}}(t)\|_g := \sqrt{g_{\bar{\gamma}(t)}\big(\dot{\bar{\gamma}}(t), \dot{\bar{\gamma}}(t)\big)} = d\big(\varphi^{-1}(x), \varphi^{-1}(a)\big). \tag{3.2.5}$$

We can bound from above $d\big(\varphi^{-1}(x), \varphi^{-1}(a)\big)$ by choosing a curve $\gamma$ obtained by the image via $\varphi^{-1}$ of a segment:

$$d\big(\varphi^{-1}(x), \varphi^{-1}(a)\big)^2 \le \int_0^1 g_{\sigma(t)}\big(\dot{\sigma}(t), \dot{\sigma}(t)\big)\, dt, \qquad \sigma(t) := \varphi^{-1}\big((1-t)x + ta\big).$$

Observe that this formula makes sense since $(1-t)x+ta \in \varphi(\mathcal{W})$ provided $\delta$ is sufficiently small.

Since

$$\sqrt{\int_0^1 g_{\sigma(t)}\big(\dot{\sigma}(t), \dot{\sigma}(t)\big)\, dt} \le C'|x-a| \tag{3.2.6}$$

for some universal constant $C'$, combining (3.2.5) and (3.2.6) we deduce that

$$\|\dot{\bar{\gamma}}(t)\|_g \le C'|x-a| \le C''\delta \qquad \forall\, t \in [0,1].$$

In particular

$$d(\bar{\gamma}(t), x) = d(\bar{\gamma}(t), \bar{\gamma}(0)) \le C''\delta \qquad \text{for all } t \in [0,1],$$

which implies that $\bar{\gamma}$ belongs to the $K\delta$-neighborhood of $\varphi^{-1}(Q)$, that is $\bar{\gamma} \subset \mathcal{Z}_{C''\delta}$.

Thanks to this fact we deduce that in the definition of the distance we can use only curves contained inside $\mathcal{Z}_{C''\delta}$. Since $\mathcal{Z}_{C''\delta} \subset \mathcal{W}$ for $\delta$ sufficiently small, all such curves can be seen as the image through $\varphi^{-1}$ of a curve contained inside $\varphi(\mathcal{W}) \subset \mathbb{R}^d$. Notice that, by (3.2.4), if

$$\sigma(t) := \varphi(\gamma(t)) = \big(\sigma^1(t), \ldots, \sigma^n(t)\big) \in \mathbb{R}^d$$

then

$$g_{\gamma(t)}\big(\dot{\gamma}(t), \dot{\gamma}(t)\big) = \sum_{k\ell} g_{k\ell}(\sigma(t))\dot{\sigma}^k(t)\dot{\sigma}^\ell(t),$$

[3]Notice that the hypothesis of completeness on $\mathcal{M}$ ensures the existence of minimizing geodesics.

therefore

$$
\begin{aligned}
d\big(\varphi^{-1}(x),\varphi^{-1}(a)\big)^2 &= \inf_{\substack{\gamma(0)=\varphi^{-1}(x),\\ \gamma(1)=\varphi^{-1}(a)}} \int_0^1 g_{\gamma(t)}\big(\dot\gamma(t),\dot\gamma(t)\big)\,dt \\
&= \inf_{\substack{\gamma(0)=\varphi^{-1}(x),\\ \gamma(1)=\varphi^{-1}(a),\\ \gamma\subset\mathcal{Z}_{C''\delta}}} \int_0^1 g_{\gamma(t)}\big(\dot\gamma(t),\dot\gamma(t)\big)\,dt \\
&= \inf_{\substack{\sigma(0)=x,\,\sigma(1)=a,\\ \sigma\subset\varphi(\mathcal{Z}_{C''\delta})}} \int_0^1 \sum_{k\ell} g_{k\ell}(\sigma(t))\dot\sigma^k(t)\dot\sigma^\ell(t)\,dt \\
&\le \big(1+\hat C\delta\big) \inf_{\substack{\sigma(0)=x,\,\sigma(1)=a,\\ \sigma\subset\varphi(\mathcal{Z}_{C''\delta})}} \int_0^1 \sum_{k\ell} A_{k\ell}\dot\sigma^k(t)\dot\sigma^\ell(t)\,dt,
\end{aligned}
$$

where in the last inequality we used that, by the Lipschitz regularity of the metric and the fact that $g_{k\ell}$ is positive definite, we have

$$
\sum_{k\ell} g_{k\ell}(z)v^k v^\ell \le (1+\hat C\delta)\sum_{k\ell} A_{k\ell}v^k v^\ell \qquad \forall\, z\in\varphi(\mathcal{Z}_{C''\delta}), \qquad \forall\, v\in\mathbb{R}^d.
$$

Using now that the minimizer for the problem

$$
\inf_{\sigma(0)=x,\,\sigma(1)=a} \int_0^1 \sum_{k\ell} A_{k\ell}\dot\sigma^k(t)\dot\sigma^\ell(t)\,dt
$$

is given by a straight segment, and since this segment is contained inside $\varphi(\mathcal{Z}_{C''\delta})$, we obtain

$$
\inf_{\substack{\sigma(0)=x,\,\sigma(1)=a,\\ \sigma\subset\varphi(\mathcal{Z}_{C''\delta})}} \int_0^1 \sum_{k\ell} A_{k\ell}\dot\sigma^k(t)\dot\sigma^\ell(t)\,dt = \langle A(x-a),x-a\rangle,
$$

which proves

$$
d\big(\varphi^{-1}(x),\varphi^{-1}(a)\big)^2 \le \big(1+\hat C\delta\big)\langle A(x-a),x-a\rangle.
$$

The lower bound is proved analogously using that

$$
\sum_{k\ell} g_{k\ell}(z)v^k v^\ell \ge (1-\hat C\delta)\sum_{k\ell} A_{k\ell}v^k v^\ell \qquad \forall\, z\in\varphi(\mathcal{Z}_{C''\delta}), \qquad \forall\, v\in\mathbb{R}^d,
$$

concluding the proof. □

Applying now this lemma, we can estimate $V_{N,r}(\mu)$ both from above and below. Since the argument in both cases is completely analogous, we just prove the upper bound.

Notice that, by the Lipschitz regularity of the metric and the fact that $\det g_{k\ell}$ is bounded away from zero, we have

$$\sqrt{\det\, g_{k\ell}(x)} \leq (1 + C\delta)\sqrt{\det\, g_{k\ell}(p)} = (1 + C\delta)\sqrt{\det\, A} \qquad \forall\, x \in Q.$$

Combining this estimate with (3.2.3) and Lemma 3.2.1, we get

$$\begin{aligned} V_{N,r}(\mu) &\leq (1 + C'\delta)\, \lambda \inf_{\alpha} \int_Q \min_{a\in\alpha} \langle A(x-a), x-a\rangle^{r/2} \sqrt{\det\, A}\, dx \\ &= (1 + C'\delta)\, \lambda \inf_{\alpha} \int_{A^{1/2}(Q)} \min_{a\in\alpha} |z-a|^r\, dz, \end{aligned}$$

where $|\cdot|$ denotes the Euclidean norm.

We now apply Theorem 3.1.3 to the probability measure $\frac{1}{|A^{1/2}(Q)|}\mathbf{1}_{A^{1/2}(Q)}\, dz$ to get

$$\begin{aligned} &\limsup_{N\to\infty} N^{r/d} V_{N,r}(\mu) \\ &\leq (1 + C'\delta)\, \lambda\, Q_r\big([0,1]^d\big) \left\| \frac{1}{|A^{1/2}(Q)|} \mathbf{1}_{A^{1/2}(Q)}\, dz \right\|_{L^{d/(d+r)}} |A^{1/2}(Q)| \\ &= (1 + C'\delta)\, \lambda\, Q_r\big([0,1]^d\big)\, |A^{1/2}(Q)|^{(d+r)/d}. \end{aligned}$$

Observing that

$$\begin{aligned} |A^{1/2}(Q)| = \int_Q \sqrt{\det A}\, dx \leq (1 + C\delta) \int_Q \sqrt{\det g_{k\ell}(x)}\, dx \\ = (1 + C\delta)\, \mathrm{vol}(\varphi^{-1}(Q)) = (1 + C\delta)\, \frac{1}{\lambda}, \end{aligned}$$

we conclude that

$$\limsup_{N\to\infty} N^{r/d} V_{N,r}(\mu) \leq \big(1 + \bar{C}\,\delta\big)\, Q_r\big([0,1]^d\big)\, \mathrm{vol}(\varphi^{-1}(Q))^{r/d}. \tag{3.2.7}$$

Arguing similarly for the lower bound, we also have

$$\liminf_{N\to\infty} N^{r/d} V_{N,r}(\mu) \geq (1 - \bar{C}\delta)\, Q_r\big([0,1]^d\big)\, \mathrm{vol}(\varphi^{-1}(Q))^{r/d}, \tag{3.2.8}$$

which concludes the local analysis of the quantization error for $\mu$ as in (3.2.2).

In the next two sections we will apply these bounds to study $V_{N,r}(\mu)$ for measures of the form $\mu = \sum_{ij} \alpha_{ij}\mu_{ij}$ where $\alpha_{ij} \neq 0$ for at most finitely many indices, and $\mu_{ij}$ has constant density on $\varphi_i^{-1}(Q_{i,j})$.

### 3.2.3 Upper bound for $V_{N,r}$

We consider a compactly supported measure $\mu = \sum_{ij} \alpha_{ij}\mu_{ij}$ where $\alpha_{ij} \neq 0$ for at most finitely many indices, and $\mu_{ij}$ is of the form $\lambda_{ij}\mathbf{1}_{\varphi_i^{-1}(Q_{i,j})} d\,\mathrm{vol}$ with

$$\varphi_i^{-1}(Q_{i,j}) \cap \varphi_{i'}^{-1}(Q_{j'}) = \emptyset, \qquad \forall\, i, i', \quad \forall\, j \neq j',$$

and $\lambda_{ij} := \frac{1}{\mathrm{vol}(\varphi_i^{-1}(Q_{i,j}))}$ (so that each measure $\mu_{ij}$ has mass 1).

To estimate $V_{N,r}(\mu)$ we first observe that, for any choice of $N_{ij}$ such that $\sum_{ij} N_{ij} \leq N$ the following inequality holds:

$$V_{N,r}(\mu) \leq \sum_{ij} \alpha_{ij}\, V_{N_{ij},r}(\mu_{ij}).$$

Indeed, if for any $i, j$ we consider a family of $N_{ij}$ points $\beta_{ij}$ which is optimal for $V_{N_{ij},r}(\mu_{ij})$, the family $\beta := \cup_{ij}\beta_{ij}$ is an admissible competitor for $V_{N,r}(\mu)$, hence

$$\begin{aligned} V_{N,r}(\mu) &\leq \int_{\mathcal{M}} \min_{b\in\beta} d(x,b)^r\, d\mu = \sum_{ij} \alpha_{ij} \int_{\mathcal{M}} \min_{b\in\beta} d(x,b)^r\, d\mu_{ij} \\ &\leq \sum_{ij} \alpha_{ij} \int_{\mathcal{M}} \min_{b\in\beta_{ij}} d(x,b)^r\, d\mu_{ij} = \sum_{ij} \alpha_{ij}\, V_{N_{ij},r}(\mu_{ij}). \end{aligned}$$

We want to chose the $N_{ij}$ in an optimal way. As it will be clear from the estimates below, the best choice is to set [4]

$$t_{ij} := \frac{\left(\alpha_{ij}\, \mathrm{vol}(\varphi_i^{-1}(Q_{i,j}))^{r/d}\right)^{d/(d+r)}}{\sum_{k\ell} \left(\alpha_{k\ell}\, \mathrm{vol}(\varphi_k^{-1}(Q_\ell))^{r/d}\right)^{d/(d+r)}},$$

and define

$$N_{ij} := [t_{ij} N].$$

Notice that $N_{ij}$ satisfy $\sum_{ij} N_{ij} \leq N$ and

$$\sum_{ij} \frac{N_{ij}}{N} \to 1 \qquad \text{as } N \to \infty.$$

[4]Notice that, if we were on $\mathbb{R}^d$ and $\varphi_i$ are just the identity map, then the formula for $t_{ij}$ simplifies to

$$t_{ij} = \frac{(\alpha_{ij})^{d/(d+r)}}{\sum_{k\ell} (\alpha_{k\ell})^{d/(d+r)}},$$

that is the exact same formula used in [47, Proof of Theorem 6.2, Step 2].

We observe that each measure $\mu_{ij}$ is a probability measure supported in only one "cube" with constant density. Hence we can apply the local quantization error (3.2.7) to each measure $\mu_{ij}$ to get that

$$\limsup_{N_{ij}\to\infty} N_{ij}^{r/d} V_{N_{ij},r}(\mu_{ij}) \le (1+\bar{C}\delta)\, Q_r\big([0,1]^d\big)\, \mathrm{vol}(\varphi_i^{-1}(Q_{i,j}))^{r/d}.$$

Recalling our choice of $N_{ij}$,

$$\begin{aligned}\limsup_{N\to\infty} N^{r/d} V_{N,r}(\mu) &\le \limsup_{N\to\infty} \sum_{ij} \alpha_{ij} \left(\frac{N}{N_{ij}}\right)^{r/d} N_{ij}^{r/d} V_{N_{ij},r}(\mu_{ij}) \\ &\le (1+\bar{C}\delta)\, Q_r\big([0,1]^d\big) \sum_{ij} \alpha_{ij} t_{ij}^{-r/d}\, \mathrm{vol}(\varphi_i^{-1}(Q_{i,j}))^{r/d},\end{aligned}$$

and observing that

$$\sum_{ij} \alpha_{ij}\, t_{ij}^{-r/d}\, \mathrm{vol}(\varphi_i^{-1}(Q_{i,j}))^{r/d} = \left(\int_{\mathcal{M}} h^{d/(d+r)}\, d\mathrm{vol}\right)^{(d+r)/d},$$

we get

$$\limsup_{N\to\infty} N^{r/d} V_{N,r}(\mu) \le (1+\bar{C}\delta)\, Q_r\big([0,1]^d\big) \left(\int_{\mathcal{M}} h^{d/(d+r)}\, d\mathrm{vol}\right)^{(d+r)/d}.$$

### 3.2.4 Lower bound for $V_{N,r}$

We consider again a compactly supported measure $\mu = \sum_{ij} \alpha_{ij}\mu_{ij}$ where $\alpha_{ij} \neq 0$ for at most finitely many indices, and $\mu_{ij}$ is of the form $\lambda_{ij} \mathbf{1}_{\varphi_i^{-1}(Q_{i,j})} d\,\mathrm{vol}$ with

$$\varphi_i^{-1}(Q_{i,j}) \cap \varphi_{i'}^{-1}(Q_{j'}) = \emptyset, \qquad \forall\, i, i', \quad \forall\, j \neq j',$$

and $\lambda_{ij} := \frac{1}{\mathrm{vol}(\varphi_i^{-1}(Q_{i,j}))}$ (so that $\int_{\mathcal{M}} \mu_{ij} = 1$). Fix $\varepsilon > 0$ with $\varepsilon \ll \delta$, and consider the cubes $Q_{j,\varepsilon}$ given by

$$Q_{j,\varepsilon} := \{y \in Q_{i,j} : \mathrm{dist}(y, \partial Q_{i,j}) > \varepsilon\}.$$

Also, consider a set $\gamma_{ij}$ consisting of $K_{ij}$ points such that

$$\min_{a\in\gamma_{ij}} d(x,a) \le \inf_{z\in\mathcal{M}\setminus\varphi_i^{-1}(Q_{i,j})} d(x,z) \qquad \forall\, x \in \varphi_i^{-1}(Q_{j,\varepsilon}) \text{ s.t. } \varphi_i^{-1}(Q_{i,j}) \cap \mathrm{supp}(\mu) \neq \emptyset.$$

Notice that the property of $\mu$ being compactly supported ensures that

$$K := \max\left\{ K_{ij} : \varphi_i^{-1}(Q_{i,j}) \cap \operatorname{supp}(\mu) \neq \emptyset \right\} < \infty$$

Then, if $\beta$ is a set of $N$ points optimal for $V_{N,r}(\mu)$ and $\beta_{ij} := \beta \cap \varphi_i^{-1}(Q_j)$,

$$\begin{aligned} V_{N,r}(\mu) &= \int_{\mathcal{M}} \min_{b\in\beta} d(x,b)^r \, d\mu \\ &\geq \sum_{ij} \int_{\varphi_i^{-1}(Q_{j,\varepsilon})} \min_{b\in\beta\cup\gamma_{ij}} d(x,b)^r \, d\mu \\ &= \sum_{ij} \int_{\varphi_i^{-1}(Q_{j,\varepsilon})} \min_{b\in\beta_{ij}\cup\gamma_{ij}} d(x,b)^r \, d\mu \\ &= \sum_{ij} \alpha_{ij}^{\varepsilon} \int_{\varphi_i^{-1}(Q_{j,\varepsilon})} \min_{b\in\beta_{ij}\cup\gamma_{ij}} d(x,b)^r \, d\mu_{ij}^{\varepsilon} \\ &\geq \sum_{ij} \alpha_{ij}^{\varepsilon} \, V_{N_{ij}+K_{ij},r}(\mu_{ij}^{\varepsilon}), \end{aligned} \tag{3.2.9}$$

where

$$\alpha_{ij}^{\varepsilon} := \int_{\mathcal{V}_i\cap\varphi_i^{-1}(Q_{j,\varepsilon})} d\mu, \qquad \mu_{ij}^{\varepsilon} := \frac{\mathbf{1}_{\mathcal{V}_i\cap\varphi_i^{-1}(Q_{j,\varepsilon})} \, d\mathrm{vol}}{\mathrm{vol}\big(\varphi_i^{-1}(Q_{j,\varepsilon})\big)}, \qquad N_{ij} := \#\beta_{ij}.$$

We notice that $\alpha_{ij}^{\varepsilon} \to \alpha_{ij}$ as $\varepsilon \to 0$.

Let $L := \liminf_{N\to\infty} N^{r/d} V_{N,r}(\mu)$. Notice that $L < \infty$ by the upper bound proved in the previous step. Choose a subsequence $N(k)$ such that

$$N(k)^{r/d} V_{N(k),r}(\mu) \to L \qquad \text{as } k \to \infty$$

and, for all $i, j$,

$$\frac{N_{ij}(k)}{N(k)} \to v_{ij} \in [0,1] \qquad \text{as } k \to \infty$$

Since $\sum_{ij} N_{ij}(k) = N(k)$ we have $\sum_{ij} v_{ij} = 1$.

Moreover $N_{ij}(k) \to \infty$ for every $i, j$. Indeed, if not, there would exists $\bar{i}, \bar{j}$ such that $N_{\bar{i}\bar{j}}(k) + K_{\bar{i}\bar{j}}(k)$ would be bounded by a number $M$. Hence, since one cannot approximate the absolutely continuous measure $\mu_{ij}^{\varepsilon}$ only with a finite number $M$ of points, it follows that

$$c_0 := V_{M,r}(\mu_{ij}^{\varepsilon}) > 0,$$

that implies in particular

$$V_{N_{\overline{ij}}(k)+K_{\overline{ij}}(k),r}(\mu^{\varepsilon}_{\overline{ij}}) \geq c_0 > 0 \qquad \forall\, k \in \mathbb{N}$$

(since $N_{\overline{ij}}(k) + K_{\overline{ij}}(k) \leq M$). This is impossible as (3.2.9) would give

$$L = \lim_{k\to\infty} N(k)^{r/d} V_{N(k),r}(\mu) \geq \lim_{k\to\infty} N(k)^{r/d} \alpha^{\varepsilon}_{\overline{ij}}\, c_0 = \infty,$$

which contradicts the finiteness of $L$.

Thanks to this fact, we can now apply the local quantization error (3.2.8) to deduce that

$$\begin{aligned} &\liminf_{k\to\infty} N_{ij}(k)^{r/d} V_{N_{ij}(k)+K_{ij}(k),r}(\mu^{\varepsilon}_{ij}) \\ &\qquad = \liminf_{k\to\infty} \big(N_{ij}(k) + K_{ij}(k)\big)^{r/d} V_{N_{ij}(k)+K_{ij}(k),r}(\mu^{\varepsilon}_{ij}) \\ &\qquad\qquad \geq (1-\bar{C}\delta)\, Q_r\big([0,1]^d\big)\, \mathrm{vol}(\varphi_i^{-1}(Q_{j,\varepsilon}))^{r/d}, \end{aligned}$$

which implies that (recalling (3.2.9))

$$L \geq (1-\bar{C}\delta)\, Q_r\big([0,1]^d\big) \sum_{ij} \alpha^{\varepsilon}_{ij}\, v_{ij}^{-r/d} \mathrm{vol}(\varphi_i^{-1}(Q_{j,\varepsilon}))^{r/d}.$$

Letting $\varepsilon \to 0$ this gives

$$L \geq (1-\bar{C}\delta)\, Q_r\big([0,1]^d\big) \sum_{ij} \alpha_{ij}\, v_{ij}^{-r/d} \mathrm{vol}(\varphi_i^{-1}(Q_j))^{r/d},$$

and applying [47, Lemma 6.8] we finally obtain

$$\begin{aligned} L &\geq (1-\bar{C}\delta)\, Q_r\big([0,1]^d\big) \sum_{ij} \Big(\alpha_{ij} \mathrm{vol}(\varphi_i^{-1}(Q_j))^{r/d}\Big)\, v_{ij}^{-r/d} \\ &\geq (1-\bar{C}\delta)\, Q_r\big([0,1]^d\big) \left(\sum_{ij} \Big(\alpha_{ij} \mathrm{vol}(\varphi_i^{-1}(Q_j))^{r/d}\Big)^{d/(d+r)}\right)^{(d+r)/d} \\ &= (1-\bar{C}\delta)\, Q_r\big([0,1]^d\big) \left(\int_{\mathcal{M}} h^{d/(d+r)}\, d\mathrm{vol}\right)^{(d+r)/d}. \end{aligned}$$

### 3.2.5 Approximation argument: general compactly supported measures

In the previous two sections we proved that if $\mu$ is compactly supported and it is of the form

$$\mu = \sum_{ij} \alpha_{ij} \frac{\mathbf{1}_{\varphi_i^{-1}(Q_{i,j})}}{\operatorname{vol}\big(\varphi_i^{-1}(Q_{i,j})\big)} \, d\operatorname{vol}$$

where $Q_{i,j}$ is a family of cubes in $\mathbb{R}^d$ of size at most $\delta$ and $\alpha_{ij} \neq 0$ for finitely many indices, then

$$\begin{aligned}(1-\bar{C}\delta)\, Q_r\big([0,1]^d\big) \left(\int_{\mathcal{M}} h^{d/(d+r)}\, d\mathrm{vol}\right)^{(d+r)/d} &\leq \liminf_{N\to\infty} N^{r/d} V_{N,r}(\mu) \\ \leq \limsup_{N\to\infty} N^{r/d} V_{N,r}(\mu) \leq (1+\bar{C}\delta)\, Q_r\big([0,1]^d\big) &\left(\int_{\mathcal{M}} h^{d/(d+r)}\, d\mathrm{vol}\right)^{(d+r)/d}. \end{aligned} \tag{3.2.10}$$

To prove the quantization result for general measures with compact support, we need three approximation steps.

First, given a compactly supported measure $\mu = h\, d\mathrm{vol}$, we can approximate it with a sequence $\{\mu_k\}_{k\in\mathbb{N}}$ of measures as above where the size of the cubes $\delta_k \to 0$, and this allows us to prove that

$$N^{r/d} V_{N,r}(\mu) \to Q_r\big([0,1]^d\big) \left(\int_{\mathcal{M}} h^{d/(d+r)}\, d\mathrm{vol}\right)^{(d+r)/d} \tag{3.2.11}$$

for any compactly supported measure of the form $h\, d\mathrm{vol}$. Then, given a singular measure with compact support $\mu = \mu^s$, we show that

$$N^{r/d} V_{N,r}(\mu) \to 0.$$

Finally, given an arbitrary measure with compact support $\mu = h\, d\mathrm{vol} + \mu^s$, we show that (3.2.11) still holds true.

The proofs of these three steps is performed in detail in [47, Theorem 6.2, Step 3, Step 4, Step 5] for the case of $\mathbb{R}^d$. As it can be easily checked, such a proof applies immediately also in our case, so we will not repeat here for the sake of conciseness.

This concludes the proof of Theorem 3.1.4 when $\mu$ is compactly supported (in particular, whenever $\mathcal{M}$ is compact).

## 3.3 Proof of Theorem 3.1.4: the non-compact case

The aim of this Section is to study the case of non-compactly supported measures. As we shall see, this situation is very different with respect to the flat case as we need to deal with the growth at infinity of $\mu$.

To state our result, let us recall the notation we already presented in the introduction: given a point $x_0 \in \mathcal{M}$, we can consider polar coordinates $(\rho, \vartheta)$ on $T_{x_0}\mathcal{M} \simeq \mathbb{R}^d$ induced by the constant metric $g_{x_0}$, where $\vartheta$ denotes a vector on the unit sphere $\mathbb{S}^{d-1}$. Then we define the quantity $A_{x_0}(\rho)$ as in (3.1.5). Our goal is to prove the following result which implies Theorem 3.1.4.

**Theorem 3.3.1.** *Let $(\mathcal{M}, g)$ be a complete Riemannian manifold, and let $\mu = h\,\mathrm{dvol} + \mu^s$ be a probability measure on $\mathcal{M}$. Then, for any $x_0 \in \mathcal{M}$ and $\delta > 0$, there exists a constant $C = C(\delta) > 0$ such that*

$$N^r V_{N^d,r}(\mu) \le C\biggl(1 + \int_{\mathcal{M}} d(x,x_0)^{r+\delta}\,d\mu(x) + \int_{\mathcal{M}} A_{x_0}\bigl(d(x,x_0)\bigr)^r\,d\mu(x)\biggr). \tag{3.3.1}$$

*In particular, if there exists a point $x_0 \in \mathcal{M}$ and $\delta > 0$ for which the right hand side is finite, we have*

$$N^{r/d} V_{N,r}(\mu) \to Q_r\bigl([0,1]^d\bigr)\left(\int_{\mathcal{M}} h^{d/(d+r)}\,\mathrm{dvol}\right)^{(d+r)/d}. \tag{3.3.2}$$

### 3.3.1 Proof of Theorem 3.3.1

We begin by the proof of (3.3.1). For this we will need the following result, whose proof is contained in [47, Lemma 6.6].

**Lemma 3.3.2.** *Let $\nu$ be a probability measure on $\mathbb{R}$. Then*

$$N^r V_{N,r}(\nu) \le C\biggl(1 + \int_{\mathbb{R}} |t|^{r+\delta}\,d\nu(t)\biggr). \tag{3.3.3}$$

To simplify the notation, given $v \in T_{x_0}\mathcal{M}$ we use $|v|_{x_0}$ to denote $\sqrt{g_{x_0}(v,v)}$.

In order to construct a family of $N^d$ points on $\mathcal{M}$, we argue as follows: first of all we consider polar coordinates $(\rho, \vartheta)$ on $T_{x_0}\mathcal{M} \simeq \mathbb{R}^d$ induced by the constant metric $g_{x_0}$, where $\vartheta$ denotes a vector on the unit sphere $\mathbb{S}^{d-1}$, and then we consider a family of "radii"

$0 < \rho_1 < \ldots < \rho_N < \infty$ and a set of $N^{d-1}$ points $\{\vartheta_1, \ldots, \vartheta_{N^{d-1}}\} \subset \mathbb{S}^{d-1}$ distributed in a "uniform" way on the sphere so that

$$\min_k d_\theta(\vartheta, \vartheta_k) \leq \frac{C}{N} \qquad \forall\, \vartheta \in \mathbb{S}^{d-1}, \tag{3.3.4}$$

where $d_\theta(\vartheta, \vartheta_k)$ denotes the distance on the sphere induced by $g_{x_0}$.

We then define the family of points $p_{i,k}$ on the tangent space $T_{x_0}\mathcal{M}$ that, in polar coordinates, are given by $p_{i,k} := (\rho_i, \vartheta_k)$, and we take the family of points on $\mathcal{M}$ given by

$$x_{i,k} := \exp_{x_0}(p_{i,k}) \qquad i = 1, \ldots, N; \quad k = 1, \ldots, N^{d-1}.$$

We notice the following estimate: given a point $x \in \mathcal{M}$, we consider the vector $p = (\rho, \vartheta) \in T_{x_0}\mathcal{M}$ defined as $p := \dot\gamma(0)$ where $\gamma : [0,1] \to \mathcal{M}$ is a constant speed minimizing geodesic. By the definition of the exponential map we notice that $x = \exp_{x_0}(p)$ and $\rho = |p|_{x_0} = d(x, x_0)$. Then, we can estimate the distance between $x := \exp_{x_0}(p)$ and $x_{i,k}$ as follows: first we consider $\sigma : [0,1] \to \mathbb{S}^{d-1} \subset T_{x_0}\mathcal{M}$ a geodesic (on the unit sphere) connecting $\vartheta$ to $\vartheta_k$ and we define $\eta := \exp_{x_0}(\rho\, \sigma)$, and then we connect $\exp_{x_0}\big((\rho, \vartheta_k)\big)$ to $x_{i,k}$ considering $\gamma|_{[\rho,\rho_i]}$, where $\gamma(s) := \exp_{x_0}\big((s, \vartheta)\big)$ is a unit speed geodesic (see Figure 3.2).

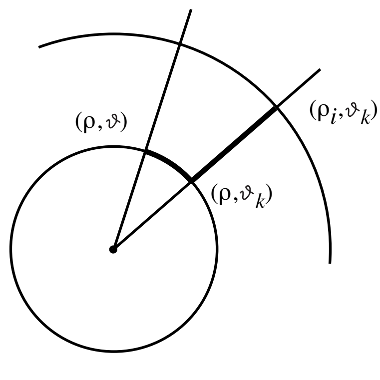


Figure 3.2: The bold curve joining $(\rho, \vartheta)$ and $(\rho_i, \vartheta_k)$ provides an upper bound for the distance between the two points.

Setting $\eta := \exp_{x_0}(\rho\,\sigma)$, this gives the bound

$$\begin{aligned} d(x, x_{i,k}) &\leq \int_0^1 \big|\dot\eta(t)\big|_{\eta(t)}\,dt + \left|\int_\rho^{\rho_i} \big|\dot\gamma(s)\big|_{\gamma(s)}\,ds\right| \\ &= \rho \int_0^1 \big|d_{\rho\sigma(t)}\exp_{x_0}[\dot\sigma(t)]\big|_{\eta(t)}\,dt + |\rho - \rho_i| \\ &\leq A_{x_0}(\rho)\int_0^1 \big|\dot\sigma(t)\big|_{x_0}\,dt + |d(x,x_0) - \rho_i| \\ &= A_{x_0}\big(d(x,x_0)\big)\,d_\theta(\vartheta_k,\vartheta) + |d(x,x_0) - \rho_i|, \end{aligned}$$

where $A_{x_0}(\rho)$ is defined in (3.1.5), and we used that $\sigma(t)$ is a geodesic (on the sphere) from $\vartheta_k$ to $\vartheta$ and that $\rho = d(x, x_0)$.

Notice that, thanks to the estimate above and by (3.3.4),

$$\begin{aligned} \min_{i,k} d(x, x_{i,k})^r &\leq \min_{i,k}\Big[A_{x_0}\big(d(x,x_0)\big)\,d_\theta(\vartheta,\vartheta_k) + |d(x,x_0) - \rho_i|\Big]^r \\ &\leq \min_i \Big[A_{x_0}\big(d(x,x_0)\big)\,\frac{C}{N} + |d(x,x_0) - \rho_i|\Big]^r. \end{aligned}$$

We can now estimate the quantization error:

$$\begin{aligned} N^r V_{N^d,r}(\mu) &\leq N^r \int_{\mathcal{M}} \min_{i,k} d(x, x_{i,k})^r\,d\mu(x) \\ &\leq N^r \int_{\mathcal{M}} \min_{i,k}\Big[A_{x_0}\big(d(x,x_0)\big)\,\frac{C}{N} + |d(x,x_0) - \rho_i|\Big]^r\,d\mu(x). \end{aligned}$$

Using that $(a+b)^r \leq 2^{r-1}(a^r + b^r)$ for $a, b > 0$ we get

$$\begin{aligned} N^r V_{N^d,r}(\mu) &\leq N^r 2^{r-1}\left[\int_{\mathcal{M}} \min_i |d(x,x_0) - \rho_i|^r\,d\mu(x) + \int_{\mathcal{M}} A_{x_0}\big(d(x,x_0)\big)^r \left(\frac{C}{N}\right)^r d\mu(x)\right] \\ &= N^r 2^{r-1}\int_{\mathcal{M}} \min_i |d(x,x_0) - \rho_i|^r\,d\mu(x) + C^r 2^{r-1}\int_{\mathcal{M}} A_{x_0}\big(d(x,x_0)\big)^r\,d\mu(x). \end{aligned}$$

Let us now consider the map $d_{x_0} : \mathcal{M} \to \mathbb{R}$ defined as $d_{x_0}(x) := d(x, x_0)$, and define the probability measure on $\mathbb{R}$ given by $\mu_1 := (d_{x_0})_\# \mu$. In this way

$$\int_{\mathcal{M}} \min_i |d(x,x_0) - \rho_i|^r\,d\mu(x) = \int_{\mathbb{R}} \min_i |s - \rho_i|^r\,d\mu_1(s).$$

We now choose the radii $\rho_i$ to be optimal for the quantization problem in one dimension for $\mu_1$. Then the above estimate and Lemma 3.3.2 yield

$$\begin{aligned}
N^r V_{N^d,r}(\mu) &\le N^r 2^{r-1} V_{N,r}(\mu_1) + C^r 2^{r-1} \int_{\mathcal{M}} A_{x_0}\big(d(x,x_0)\big)^r \, d\mu(x), \\
&\le C'\bigg(1 + \int_0^\infty s^{r+\delta}\, d\mu_1(s) + \int_{\mathcal{M}} A_{x_0}\big(d(x,x_0)\big)^r \, d\mu(x)\bigg) \\
&= C'\bigg(1 + \int_{\mathcal{M}} d(x,x_0)^{r+\delta}\, d\mu(x) + \int_{\mathcal{M}} A_{x_0}\big(d(x,x_0)\big)^r \, d\mu(x)\bigg),
\end{aligned}$$

that concludes the proof of (3.3.1).

To show why this bound implies (3.3.2) (and hence Theorem 3.1.4 in the general non-compact case), we first notice that by (3.3.1) it follows that, for any $M \ge 1$,

$$\begin{aligned}
&M^{r/d} V_{M,r}(\mu) \\
&\qquad \le C\bigg(1 + \int_{\mathcal{M}} d(x,x_0)^{r+\delta}\, d\mu(x) + \int_{\mathcal{M}} A_{x_0}\big(d(x,x_0)\big)^r \, d\mu(x)\bigg). \qquad (3.3.5)
\end{aligned}$$

Indeed, for any $M \ge 1$ there exists $N \ge 1$ such that $N^d \le M < (N+1)^d$, hence (since $V_{M,r}$ is decreasing in $M$)

$$\begin{aligned}
M^{r/d} V_{M,r}(\mu) &\le (N+1)^r V_{N^d,r}(\mu) = \left(1 + \frac{1}{N}\right)^r N^r V_{N^d,r}(\mu) \\
&\le C\bigg(1 + \int_{\mathcal{M}} d(x,x_0)^{r+\delta}\, d\mu(x) + \int_{\mathcal{M}} A_{x_0}\big(d(x,x_0)\big)^r \, d\mu(x)\bigg),
\end{aligned}$$

which proves (3.3.5).

We now prove (3.3.2). Observe that, as shown in [47, Proof of Theorem 6.2, Step 5], once the asymptotic quantization is proved for compactly supported probability measures, by the monotone convergence theorem one always has

$$\liminf_{N\to\infty} N^{r/d} V_{N,r}(\mu) \ge Q_r\big([0,1]^d\big) \left(\int_{\mathcal{M}} h^{d/(d+r)}\, d\mathrm{vol}\right)^{(d+r)/d},$$

hence one only have to prove the limsup inequality.

For that, one splits the measure $\mu$ as the sum of $\mu_R^1 := \chi_{B_R(x_0)}\mu$ and $\mu_R^2 := \chi_{\mathcal{M}\setminus B_R(x_0)}\mu$, where $R \gg 1$. Then one applies [47, Lemma 6.5(a)] to bound from above $N^{r/d} V_{N,r}(\mu)$

in terms of $N^{r/d}V_{N,r}(\mu_R^1)$ and $N^{r/d}V_{N,r}(\mu_R^2)$, and uses the result in the compact case for $N^{r/d}V_{N,r}(\mu_R^1)$, to obtain that, for any $\varepsilon \in (0,1)$

$$\limsup_{N\to\infty} N^{r/d}V_{N,r}(\mu) \le (1-\varepsilon)^{-r/d}\, Q_r\big([0,1]^d\big) \left(\int_{B_R(x_0)} h^{d/(d+r)}\, d\mathrm{vol}\right)^{(d+r)/d} + \mu(\mathcal{M}\setminus B_R(x_0))\, \varepsilon^{-r/d} \limsup_{N\to\infty} N^{r/d}V_{N,r}\left(\frac{1}{\mu(\mathcal{M}\setminus B_R(x_0))}\mu_R^2\right).$$

Thanks to (3.3.5), we can bound the limsup in the right hand side by

$$\varepsilon^{-r/d}\left(\mu(\mathcal{M}\setminus B_R(x_0)) + \int_{\mathcal{M}} d(x,x_0)^{r+\delta}\, d\mu_R^2(x) + \int_{\mathcal{M}} A_{x_0}\big(d(x,x_0)\big)^r\, d\mu_R^2(x)\right),$$

that tends to 0 as $R\to\infty$ by dominated convergence. Hence, letting $R\to\infty$ we deduce that

$$\begin{aligned}\limsup_{N\to\infty} N^{r/d}V_{N,r}(\mu) &\le (1-\varepsilon)^{-r/d}\, Q_r\big([0,1]^d\big) \lim_{R\to\infty}\left(\int_{B_R(x_0)} h^{d/(d+r)}\, d\mathrm{vol}\right)^{(d+r)/d}\\ &= (1-\varepsilon)^{-r/d}\, Q_r\big([0,1]^d\big)\left(\int_{\mathcal{M}} h^{d/(d+r)}\, d\mathrm{vol}\right)^{(d+r)/d},\end{aligned}$$

and the result follows letting $\varepsilon\to 0$.

## 3.4 Proof of Theorem 3.1.7

We begin by noticing that if

$$\int_{\mathbb{H}^2} d(x,x_0)^p\, d\mu < \infty$$

for some $x_0\in\mathbb{H}^2$, then this holds for any other point: indeed, given $x_1\in\mathbb{H}^2$,

$$\int_{\mathbb{H}^2} d(x,x_1)^p\, d\mu \le 2^{p-1}\int_{\mathbb{H}^2}\big[d(x,x_0)^p + d(x_0,x_1)^p\big]\, d\mu < \infty.$$

In particular, it suffices to check the moment condition at only one point.

We fix a point $x_0\in\mathbb{H}^2$ and we use the exponential map at $x_0$ to identify $\mathbb{H}^2$ with $(\mathbb{R}^2, d^2\rho + \sinh\rho\, d^2\vartheta)$. Then, we define the measure

$$\mu := \sum_{k\in\mathbb{N}} e^{-(1+\varepsilon)k}\mathcal{H}^1 \llcorner \mathbb{S}^1_k,$$

where $\mathcal{H}^1 \llcorner \mathbb{S}^1_R$ denotes the 1-dimensional Haudorff measure restricted to the circle around the origin of radius $R$, and $\varepsilon > 0$ is a constant to be fixed.

We begin by noticing that

$$\int_{\mathbb{H}^2} d(x, x_0)^p \, d\mu = \sum_{k\in\mathbb{N}} e^{-(1+\varepsilon)k} \int_{\mathbb{S}^1_k} \rho^p \, d\mathcal{H}^1$$
$$= \sum_{k\in\mathbb{N}} e^{-(1+\varepsilon)k} k^p \, 2\pi \sinh(k) \approx \sum_{k\in\mathbb{N}} e^{-\varepsilon k} k^p < \infty$$

for all $p > 0$.

An important ingredient of the proof will be the following estimate on the quantization error for the uniform measure on a circle around the origin.

**Lemma 3.4.1.** *For any $R \geq 1$ and $M \in \mathbb{N}$ we have*

$$V_{M,r}\big(\mathcal{H}^1 \llcorner \mathbb{S}^1_R\big) \gtrsim \left(\frac{e^R}{2R} - M\right)_+ R.$$

*Proof.* To prove the above estimate, we built a good competitor for the minimization problem. Let us denote with $[\cdot]$ the integer part, and define

$$L := \left[\frac{e^R}{2R}\right].$$

We split $\mathbb{S}^1_R$ in $2L$ arcs $\Sigma_{i,R}$ of equal length. Notice that the following estimate holds: there exists a positive constant $c$, independent of $R$, such that

$$d(\Sigma_{2j,R}, \Sigma_{2j',R}) > c \qquad \forall\ j \neq j' \in \{1, \ldots, L\}. \tag{3.4.1}$$

To show this fact, one argues as follows: consider a geodesic connecting a point $x_1 \in \Sigma_{2j,R}$ to $x_2 \in \Sigma_{2j',R}$. Because $j \neq j'$ any curve connecting them has to rotate by an angle of order at least $R/e^R$. Now, two cases arise: either the geodesic $\gamma : [0, 1] \to \mathbb{H}^2$ is always contained inside $\mathbb{R}^d \setminus B_{R-1}(0)$, or not. In the first case we exploit that the metric is always larger than $\sinh^2(R-1)d^2\vartheta$. More precisely, if we denote by $(e_\rho, e_\theta)$ a basis of tangent vectors in polar coordinates

$$\begin{aligned} d(x, y) &= \int_0^1 \sqrt{g_{\gamma(t)}\big(\dot\gamma(t), \dot\gamma(t)\big)} \, dt \\ &= \int_0^1 \sqrt{\big(\dot\gamma(t)\cdot e_\rho\big)^2 + \sinh^2(\rho)\big(\dot\gamma(t)\cdot e_\theta\big)} \, dt \\ &\geq \sinh(R-1) \int_0^1 |\dot\gamma(t)\cdot e_\theta| \, dt \gtrsim e^{R-1}\frac{R}{e^R} \approx R \geq 1, \end{aligned}$$

where for the last inequality we used that $\gamma$ has to rotate by an angle of order at least $R/e^R$. In the second case, to enter inside the ball $B_{R-1}(0)$ the geodesic has to travel a distance at least 1, so its length is greater that 1. This proves the validity of (3.4.1).

We pick now a family of $M$ points $\{x_\ell\}_{\ell=1}^M$. Then, by (3.4.1) and triangle inequality, we have that for every index $\ell$ there exists at most one index $j(\ell)$ such that

$$d(x_\ell, \Sigma_{2j,R}) > \frac{c}{2} \qquad \forall\, j \neq j(\ell).$$

Therefore there exists a family of indices $J \in \{1, \dots, L\}$ of cardinality at least $(L-M)_+$ such that

$$d(x_\ell, \Sigma_{2j,R}) > \frac{c}{2} \qquad \forall j \in J, \quad \forall \ell = 1, \dots, M.$$

We can now estimate the quantization error:

$$\begin{aligned} V_{M,r}\big(\mathcal{H}^1 \llcorner \mathbb{S}^1_R\big) &= \min_{\alpha \subset \mathbb{H}^2 : |\alpha| = M} \int_{\mathbb{S}^1_R} \min_{x_\ell \in \alpha} d(x, x_\ell)^r \, d\mathcal{H}^1 \\ &\geq \min_{\alpha \subset \mathbb{H}^2 : |\alpha| = M} \sum_{j=1}^{L} \int_{\Sigma_{2j,R}} \min_{x_\ell \in \alpha} d(x, x_\ell)^r \, d\mathcal{H}^1 \\ &\geq \sum_{j \in J} \int_{\Sigma_{2j,R}} \left(\frac{c}{2}\right)^r \, d\mathcal{H}^1 \gtrsim (L-M)_+ R, \end{aligned}$$

where at the last step we used that $\mathcal{H}^1(\Sigma_{2j,R}) \approx R$.

□

We can now conclude the proof. Indeed, given a set of points $\{x_\ell\}_{1 \leq \ell \leq N^2}$ optimal for $\mu$, these points are admissible for the quantization problem of each measure $\mathcal{H}^1 \llcorner \mathbb{S}^1_k$, therefore

$$\begin{aligned} V_{N^2,r}(\mu) &= \sum_{k \in \mathbb{N}} e^{-(1+\varepsilon)k} \int_{\mathbb{S}^1_k} \min_\ell d(x, x_\ell)^r \, d\mathcal{H}^1(x) \\ &\geq \sum_{k \in \mathbb{N}} e^{-(1+\varepsilon)k} V_{N^2,r}\big(\mathcal{H}^1 \llcorner \mathbb{S}^1_k\big) \\ &\gtrsim \sum_{k \in \mathbb{N}} e^{-(1+\varepsilon)k} \left(\frac{e^k}{2k} - N^2\right)_+ k, \end{aligned}$$

where at the last step we used Lemma 3.4.1. Noticing that, for $N$ large,

$$\frac{e^k}{2k} - N^2 \geq \frac{1}{4}\frac{e^k}{k} \qquad \text{for } k \geq \log(N^4),$$

we conclude that

$$\begin{aligned} N^r V_{N^2,r}(\mu) &\gtrsim \frac{N^r}{4} \sum_{k \geq \log(N^4)} e^{-(1+\varepsilon)k} \frac{e^k}{k}\, k \\ &= \frac{N^r}{4} \sum_{k \geq \log(N^4)} e^{-\varepsilon k} \\ &\gtrsim N^r \int_{\log(N^4)}^{\infty} e^{-\varepsilon t}\, dt \approx \frac{N^r N^{-4\varepsilon}}{\varepsilon} \to \infty \end{aligned}$$

as $N \to \infty$ provided we choose $\varepsilon < r/4$.

# Part II

# Quasineutral limit for Vlasov-Poisson

# Chapter 4

# The Vlasov-Poisson equation

The general purpose of *kinetic theory* is to describe the time evolution of a system consisting of a very large number of identical particles. The mathematical study of such systems leads to a class of partial differential equations called *kinetic equations*.

## 4.1 Mean field limit

In this Section we introduce the Vlasov-Poisson equation as a typical example of mean-field equation. We shall explain how to obtain the Vlasov-Poisson equation as an approximation of the system of Newton equations of motion in the limit when the number of particles tends to infinity.

### 4.1.1 Newton equations

The standard model describing the collective interaction of a $N$ particle system in $\mathbb{R}^d$ is the system of Newton equations:

$$m_i \ddot{x}_i(t) = \sum_j F_{j \to i}(t), \tag{4.1.1}$$

where $m_i$ is the mass of the $i$-th particle, $x_i(t) \in \mathbb{R}^d$ and $\ddot{x}_i(t)$ are respectively its position and its acceleration at time $t$, and $F_{j\to i}(t)$ denote the force exerted by particle $j$ on particle $i$. In the following we consider the case where all masses are equal, that is, $m_i = m$ for all $i$. Since we will not deal with magnetic forces, we can focus on the case of forces generated by an *interaction potential*, i.e., there exists $W : \mathbb{R}^d \to \mathbb{R}$ such that $F(x-y) = -\nabla W(x-y)$ is the force exerted from a particle located in $y$ on position $x$.

A classical example of interaction kernel is the Poisson kernel,

$$F = c\frac{x}{|x|^d} \quad \text{in } \mathbb{R}^d.$$

This corresponds to consider particles under gravitational interaction for $c < 0$ or electrostatic interactions (ions in a plasma) for $c > 0$.

Describing a system of $N$ interacting particles via Newton's equations leads to a Hamiltonian system which has as many equations as the dimension of the phase space of the system. More precisely, the single particle phase space is the set $\mathbb{R}^d \times \mathbb{R}^d$ of pairs of all possible positions and momenta of an unconstrained single point particle in the $d$-dimensional Euclidean space $\mathbb{R}^d$. Thus, for a system of $N$ identical point particles in $\mathbb{R}^d$, the number of degrees of freedom is $dN$ and the $N$-particle phase space is the space $(\mathbb{R}^d \times \mathbb{R}^d)^N$ of $2N$-tuples of all possible positions and velocities of the $N$ point particles. This corresponds to a *microscopic* description of the system. However, in most physical cases the number $N$ is very large, usually of the order of $10^{23}$ (the Avogadro number), as one can see by considering as an example a rarefied gas or a plasma. This means that, even from the point of view of numerical analysis, it is unreasonable to exactly determine the dynamic of each individual particle.

An alternative approach to circumvent this difficulty caused the the presence of so many particles consists in applying Newton's second law of motion to each infinitesimal volume element of a fluid. In this way one ends up with the *hydrodynamic equations*, such as the Euler or Navier-Stokes equations, which describe the *macroscopic* behavior of the system.

Kinetic equations, instead, describe the evolution of a $N$ particle system at a *mesoscopic* level, which is an intermediate viewpoint between the Newtonian dynamic of the microscopic particles and a macroscopic hydrodynamical model. In this context, the evolution of the same many body system is described by a partial differential equation set on the single-particle phase space. In the following we consider the *mean-field limit* as a reduction of the $N$-particle phase space to the single particle phase space. From the theoretical point of view, the mean-field approximation is extremely important because it establishes the basic limit equation, and in addition it shows that the qualitative behaviour of the system does not depend on the exact value of the number of particles. This latter observation is very useful in numerical simulations.

### 4.1.2 Mean field limit of the $N$ particle system

Roughly speaking, a mean-field model describes the evolution of a typical particle subject to the collective interaction created by a large number $N$ of other particles, that we suppose identical. The state of the typical particle is given by its phase space density; the force field exerted by the $N$ other particles on this typical particle is approximated by the average with respect to the phase space density of the force field exerted on that particle from each point in the phase space. This approach is suitable for interaction potentials that are not too "sensitive" to the precise position of each particle. Often this assumption is called *long range interaction*.

The goal now is to understand how one can let the dimension of the phase space (that is $2dN$, $d$ being the dimension of the physical space) go to infinity in a rigorous way. First, let us convert the second-order Newton equations into a first-order system introducing for each position variable $x_i$ the velocity variable $v_i := \dot{x}_i$. Then, the whole state of the system at time $t$ is described by $(x_1, v_1), \dots, (x_N, v_N)$, where the positions $x_i$ belong to the $d$-dimensional space of positions $X^d$ (which may be the Euclidean space $\mathbb{R}^d$, or a subset $\Lambda$ of $\mathbb{R}^d$, or also the $d$-dimensional torus $\mathbb{T}^d$) and the velocities $v_i$ belong is $\mathbb{R}^d$. In order to balance the kinetic and the potential energy of the system and to obtain a nontrivial limit as $N \to \infty$, we shall assume that the particles have mass $1/N$. Let $H_N$ be the Hamiltonian of the $N$-particles system interacting via $W$, that is,

$$H_N(x_1, \dots, x_N, v_1, \dots, v_N) = \sum_i \frac{1}{2}|v_i|^2 + \frac{1}{N}\sum_j W(x_i - x_j). \tag{4.1.2}$$

Then, Newton laws (4.1.1) are equivalent to the Hamiltonian system

$$\dot{x}_i = \nabla_{v_i} H_N, \quad \dot{v}_i = -\nabla_{x_i} H_N, \tag{4.1.3}$$

and, by Liouville' s theorem we obtain the following equation:

$$\frac{df_N}{dt} = \partial_t f_N + \nabla_x f_N \cdot \dot{x} + \nabla_v f_N \cdot \dot{v} = \partial_t f_N + [H_N, f_N] = 0 \tag{4.1.4}$$

where $[,]$ is the canonical Poisson bracket and $f_N = f_N(t, x_1, \dots, x_N, v_1, \dots, v_N)$ denotes the joint *distribution function* in the $N$-particle phase space, i.e.,

$$f_N(t, x_1, \dots, x_N, v_1, \dots, v_N)\, dx_1 \dots dx_N\, dv_1 \dots dv_N \tag{4.1.5}$$

is the probability of finding, at time $t$, the $i$-th particle in a volume $dx_i\, dv_i$ around the point $(x_i, v_i)$ in the $N$-particle phase space $(X^d \times \mathbb{R}^d)^N$. In this context, Liouville's equation translates into *incompressibility* of the flow in phase space. This property is central to classical mechanics, and can be thought of as information conservation, where trajectories in phase space cannot merge or diverge. Let us notice that the $N$-particle Liouville equation, although it allows for considering superpositions of all trajectories at the same time, still contains exactly the same amount of information as the original Newton equations.

Since all particles are identical, it is enough to know the state of the system up to permutation of particles. Thus, let us consider $C = \{(x_1, v_1), \ldots, (x_N, v_N)\}$ to be the quotient of the phase space $(X^d \times \mathbb{R}^d)^N$ by the permutation group $S_N$. Then there is a one-to-one correspondence between such a cloud of undistinguishable points $C$ and the associated *empirical measure*

$$\mu_C^{(N)} := \frac{1}{N} \sum_i \delta_{(x_i, v_i)},$$

where $\delta_{(x,v)}$ is the Dirac mass in phase space at $(x, v)$. The empirical measure belongs to the space of probability measures on the one particle phase space $\mathcal{P}(X^d \times \mathbb{R}^d)$, even if it depends on all the configuration $C$. This probability measure should not be confused with the joint distribution function $f_N$ defined in (4.1.5), the latter being a probability measure on the $N$-particle phase space. The main advantage of introducing $\mu_C^{(N)}$ is that it belongs to an infinite-dimensional space, but independent of the number of particles. It is now clear that the plan is to rewrite the Newton equations in terms of the empirical measure, and then pass to the limit as $N \to \infty$.

Integrating $\mu_t^{(N)} := \frac{1}{N} \sum_i \delta_{(x_i(t), v_i(t))}$ against a test function $\varphi \in C_c^\infty(X^d \times \mathbb{R}^d)$, we find out that the $N$ body problem (4.1.3) associated to the rescaled Hamiltonian (4.1.2) yields the following equation for $\mu_t^{(N)}$:

$$\partial_t \int \varphi \, d\mu_t^{(N)} = \int v \cdot \nabla_x \varphi \, d\mu_t^{(N)} - \iint \nabla W(x-y) \cdot \nabla_v \varphi(x, v) \, d\mu_t^{(N)}(x, v) \, d\mu_t^{(N)}(y, w). \tag{4.1.6}$$

The equation above is the *Vlasov equation* in distributional form.

Recalling that weak convergence for probability measures is equivalent to convergence in the sense of distributions, one can prove that, if $W \in C^1(X^d)$, then as $N \to \infty$ the empirical measure $\mu_t^{(N)}$ will converge to some limit measure $f(t, x, v)$ solving the Vlasov equation, that is:

$$\partial_t f + v \cdot \nabla_x f + F[f](t, x) \cdot \nabla_v f = 0 \tag{4.1.7}$$

where

$$F[f](x) = -\iint \nabla W(x-y)\, f(dy, dw) = -\nabla W *_{x,v} f.$$

The function $f(t, x, v)$ represents the distribution function of the system at time $t$, that is the number density of particles that are located at the position $x$ and have instantaneous velocity $v$ at time $t$. This formal discussion shows a possible way to replace a very large number of simple ordinary differential equations by just one partial differential equation in the limit as $N$ goes to infinity. In this system there is no direct interaction between the particles, and the dynamic of each particular particle is affected by a field which is generated collectively by all particles together.

## 4.2 Quantitative stability via Wasserstein type distances

In the previous Section we assumed that $W$ is continuously differentiable. Although it was enough to pass to the limit in the Vlasov equation for the empirical measure, this assumption does not implies that the solution of the Vlasov equation is unique, nor it provides any information on the rate of convergence.

If $W$ is smoother then one can prove more precise results of quantitative convergence, involving distances on probability measures, for instance Wasserstein distances. To prove the following stability result, it would be sufficient to use the 1-Wasserstein distance, assuming finiteness of first moments. However, we can avoid this restriction introducing another distance on probability measures which does not need any moment assumption.

**Definition 4.2.1.** *The* bounded Lipschitz distance $d_{bL}$ *between measure* $\nu, \mu \in \mathcal{P}(X^d \times \mathbb{R}^d)$ *is defined as*

$$d_{bL}(\nu, \mu) = \sup_{f \in \mathcal{D}} \left| \int_{X^d \times \mathbb{R}^d} f(x, v)\, d\nu(x, v) - \int_{X^d \times \mathbb{R}^d} f(x, v)\, d\mu(x, v) \right|,$$

*where* $\mathcal{D} = \left\{ f \,:\, f : X^d \times \mathbb{R}^d \to [0, 1] \text{ and } |f(x, v) - f(y, w)| \leq |(x, v) - (y, w)| \right\}$.

We recall that this distance induces the weak topology on $\mathcal{P}(X^d \times \mathbb{R}^d)$. Our goal here is to present an existence and uniqueness result for solutions on the Vlasov equation when the potential $W \in C^2$ [36].

**Theorem 4.2.2.** *Let* $W \in C^2_b(X^d)$. *Then the following holds:*

**i)** *Equation (4.1.7) has a unique solution in* $\mathcal{P}(X^d \times \mathbb{R}^d)$ *for any initial datum* $\bar{\mu} \in \mathcal{P}(X^d \times \mathbb{R}^d)$*;*

**ii)** *if* $\mu_t^{(1)}$ *and* $\mu_t^{(2)}$ *solve (4.1.7), then there exists a constant* $c > 0$ *depending only on* $W$ *such that*

$$d_{bL}(\mu_t^{(1)}, \mu_t^{(2)}) \leq e^{ct} d_{bL}(\mu_0^{(1)}, \mu_0^{(2)}) \qquad \forall\, t \geq 0. \tag{4.2.1}$$

Notice that, as a corollary of this result, one gets a quantitative information on the rate of convergence of the empirical measure to its limit: on any finite time interval the closeness (measured with respect to $d_{bL}$) of $\mu_0^{(N)}$ to its limit is propagated linearly in time.

*Proof.* By assumption there exists $B, L \in \mathbb{R}$ such that

$$|\nabla^2 W| \leq B; \quad |W(x) - W(y)| \leq L|x - y|. \tag{4.2.2}$$

Given a family of measures $\{\mu_t\} \subset \mathcal{P}(X^d \times \mathbb{R}^d)$ weakly continuous with respect to $t$, this family defines a time dependent force field

$$F_t^{\mu}(x) = -\int \nabla W(x - y)\, d\mu_t(y, w) \tag{4.2.3}$$

Thanks to our assumptions on $W$, by Cauchy-Lipschitz Theorem, the ODEs

$$\dot{x}(t) = v(t), \qquad \dot{v}(t) = F_t^{\mu}(x(t)), \tag{4.2.4}$$

have a unique global solution which induces a two-parameter family of maps $T_{\{\mu\}}^{s,t} : X^d \times \mathbb{R}^d \to X^d \times \mathbb{R}^d$ which represent the evolution of a point $(x, v)$ from time $s$ to time $t$. These maps induce by duality a family of maps on $\mathcal{P}(X^d \times \mathbb{R}^d)$:

$$\lambda_t = (T_{\{\mu\}}^{0,t})_{\#}\lambda_0 \qquad \forall\, \lambda_0 \in \mathcal{P}(X^d \times \mathbb{R}^d). \tag{4.2.5}$$

In particular, (4.1.7) is equivalent to the fixed point equation

$$\mu_t = (T_{\{\mu\}}^{0,t})_{\#}\mu_0, \tag{4.2.6}$$

so we shall prove existence and uniqueness via the contraction approach.

First, we shall prove (4.2.1).

From (4.2.6) and by triangle inequality we have

$$\begin{aligned} d_{bL}(\mu_t, \lambda_t) &= d_{bL}((T^{0,t}_{\{\mu\}})_\# \mu_0, (T^{0,t}_{\{\lambda\}})_\# \lambda_0) \\ &\leq d_{bL}((T^{0,t}_{\{\mu\}})_\# \mu_0, (T^{0,t}_{\{\lambda\}})_\# \mu_0) + d_{bL}((T^{0,t}_{\{\lambda\}})_\# \mu_0, (T^{0,t}_{\{\lambda\}})_\# \lambda_0). \end{aligned} \tag{4.2.7}$$

Call $w(t) = (x(t), v(t))$. Then (4.2.4) can be written as

$$\dot{w}(t) = G^\mu_t(w), \tag{4.2.8}$$

where $G^\mu_t(w) = \left(\dot{x}(t), -\int \nabla W(x - x')\, d\mu_t(x', v')\right)$.

Notice that

$$\begin{aligned} d_{bL}((T^{0,t}_{\{\lambda\}})_\# \mu_0, (T^{0,t}_{\{\lambda\}})_\# \lambda_0) &= \sup_{f \in \mathcal{D}} \left| \int \Big( (T^{0,t}_{\{\lambda\}})_\# \mu_0(dw) - (T^{0,t}_{\{\lambda\}})_\# \lambda_0(dw) \Big) f(w) \right| \\ &= \sup_{f \in \mathcal{D}} \left| \int \big(\mu_0(dw) - \lambda_0(dw)\big) f(T^{0,t}_{\{\lambda\}} w) \right|. \end{aligned}$$

We claim that the Lipschitz constant of $f(T^{0,t}_{\{\lambda\}} w)$ is bounded by $e^{L't}$ with $L' = B + 1$. To this aim, since $f \in \mathcal{D}$ it is enough to bound

$$\Delta_t := \left| T^{0,t}_{\{\lambda\}}(w) - T^{0,t}_{\{\lambda\}}(w') \right|.$$

By the definition of flux

$$T^{0,t}_{\{\lambda\}}(w) = w + \int_0^t G^\lambda_\tau(w)\, d\tau, \tag{4.2.9}$$

since

$$\begin{aligned} \left| G^\lambda_t(w) - G^\lambda_t(w) \right|^2 &= |v - v'|^2 + \left| \int \left(\nabla W(x - \bar{x}) - \nabla W(x' - \bar{x})\right)\, d\mu_t(\bar{x}, \bar{v}) \right|^2 \\ &\leq |v - v'|^2 + B^2 |x - x'|^2 \leq \big(L'\, |w - w'|\big)^2, \end{aligned} \tag{4.2.10}$$

hence $\frac{d}{dt}|\Delta_t| \leq L'|\Delta_t|$ which implies that

$$|\Delta_t| \leq e^{L't} |\Delta_0|,$$

proving the claim.

Thanks to the claim, going back to (4.2.9) we get

$$\begin{aligned} d_{bL}((T^{0,t}_{\{\lambda\}})_{\#}\mu_0,(T^{0,t}_{\{\lambda\}})_{\#}\lambda_0) &= \sup_{f\in\mathcal{D}}\left|\int(\mu_0(dw)-\lambda_0(dw))\,f(T^{0,t}_{\{\lambda\}}w)\right| \\ &\le e^{L't}d_{dL}(\mu_0,\lambda_0). \end{aligned} \tag{4.2.11}$$

We now esitante the first term in the right hand side of (4.2.7). Using again Liouville's Theorem, we have

$$d_{bL}\left((T^{0,t}_{\{\mu\}})_{\#}\mu_0,(T^{0,t}_{\{\lambda\}})_{\#}\mu_0\right)=\sup_{f\in\mathcal{D}}\left|\int\mu_0(dw)\left(f(T^{0,t}_{\{\mu\}}w)-f(T^{0,t}_{\{\lambda\}}w)\right)\right|,$$

and using (4.2.9) and that $f$ is 1-Lipschitz we get

$$\begin{aligned} d_{bL}\left((T^{0,t}_{\{\mu\}})_{\#}\mu_0,(T^{0,t}_{\{\lambda\}})_{\#}\mu_0\right) &\le \int\mu_0(dw)\left|T^{0,t}_{\{\mu\}}w-T^{0,t}_{\{\lambda\}}w\right| \\ &= \int\mu_0(dw)\left|\int_0^t d\tau\left[G^{\mu}_{\tau}(T^{0,\tau}_{\{\mu\}}w)-G^{\lambda}_{\tau}(T^{0,\tau}_{\{\lambda\}}w)\right]\right| \\ &:= R(t). \end{aligned} \tag{4.2.12}$$

Using (4.2.10) and the triangle inequality, we have

$$\begin{aligned} R(t) &\le \int\mu_0(dw)\left|\int_0^t d\tau\left[G^{\mu}_{\tau}(T^{0,\tau}_{\{\mu\}}w)-G^{\lambda}_{\tau}(T^{0,\tau}_{\{\mu\}}w)\right]\right| \\ &\quad+\int\mu_0(dw)\left|\int_0^t d\tau\left[G^{\lambda}_{\tau}(T^{0,\tau}_{\{\mu\}}w)-G^{\lambda}_{\tau}(T^{0,\tau}_{\{\lambda\}}w)\right]\right| \\ &\le \int_0^t d\tau\left[\int\mu_\tau(dw)\left|G^{\mu}_{\tau}(w)-G^{\lambda}_{\tau}(w)\right|+L'\int\mu_0(dw)\int_0^t d\tau\left|T^{0,\tau}_{\{\mu\}}w-T^{0,\tau}_{\{\lambda\}}w\right|\right]. \end{aligned}$$

To estimate the integrand in the first term of the right hand side we use that the map $(x',v')\mapsto\nabla W(x-x')$ is $B$-Lipschitz and bounded by $L$ for any $x\in X^d$ to deduce that

$$\begin{aligned} \left|G^{\mu}_{\tau}(w)-G^{\lambda}_{\tau}(w)\right| &= \left|\int(\mu_\tau(dw')-\lambda_\tau(dw'))\nabla W(x-x')\right| \\ &\le \max\{B,L\}\,d_{bL}(\mu_\tau,\lambda_\tau), \end{aligned}$$

Hence, recalling (4.2.9), we obtain

$$R(t) \leq \max\{B, L\} \int_0^\tau d_{bL}(\mu_\tau, \lambda_\tau)\, d\tau + \sqrt{L'} \int_0^\tau R(\tau)\, d\tau,$$

and Gronwall Lemma gives

$$R(t) \leq \max\{B, L\} \int_0^\tau e^{L'(t-\tau)} d_{bL}(\mu_\tau, \lambda_\tau)\, d\tau.$$

Hence, combining this bound with (4.2.12), we have shown that

$$d_{bL}(\mu_\tau, \lambda_\tau) \leq e^{L't} d_{bL}(\mu_0, \lambda_0) + \max\{B, L\} \int_0^\tau e^{L'(t-\tau)} d_{bL}(\mu_\tau, \lambda_\tau)\, d\tau, \tag{4.2.13}$$

and a further application of Gronwall Lemma yields the estimate

$$d_{bL}(\mu_\tau, \lambda_\tau) \leq e^{ct} d_{bL}(\mu_0, \lambda_0) \qquad \forall\, t \geq 0, \tag{4.2.14}$$

where $c$ is a constant depending only on $W$. This proves (4.2.1).

We can now use a fixed point argument to show that (4.2.6) has a unique solution in $\mathcal{P}(X^d \times \mathbb{R}^d)$ for any initial datum $\bar{\mu} \in \mathcal{P}(X^d)$. To this aim we consider the space $C_{\mathcal{P}}$ of all continuous curves $\mu_t : [0, \infty) \to \mathcal{P}(X^d \times \mathbb{R}^d)$ (the continuity being with respect to the metric $d_{bL}$), endowed with the metric

$$d_\alpha(\{\mu\}, \{\lambda\}) = \sup_{t \in [0,\infty)} d_{bL}(\mu_t, \lambda_t) e^{-\alpha t},$$

with $\alpha > 0$ to be chosen. Since $(\mathcal{P}(X^d \times \mathbb{R}^d), d_{bL})$ is a complete metric space, so is $(C_{\mathcal{P}}, d_\alpha)$.

Fix $\bar{\mu} \in \mathcal{P}(X^d)$, and let $[0, \infty) \ni t \to \mu_t$ be a curve in $C_{\mathcal{P}}$ such that $\mu_0 = \bar{\mu}$. The flux $T^{0,t}_{\{\mu\}}$ associated to the curve $\{\mu\}$ induces a new element of $\mathcal{P}(X^d)$ via the formula $(T^{0,t}_{\{\mu\}})_\# \bar{\mu}$, and we call $\mathcal{F} : C_{\mathcal{P}} \to C_{\mathcal{P}}$ the map defined in this way. We claim that $\mathcal{F}$ is a contraction. Indeed (4.2.13) gives

$$\begin{aligned} d_{bL}(\mathcal{F}(\{\mu\})(t), \mathcal{F}(\{\lambda\})(t)) &= d_{bL}((T^{0,t}_{\{\mu\}})_\# \bar{\mu}, (T^{0,t}_{\{\lambda\}})_\# \bar{\mu}) \\ &\leq \max\{L, B\} \int_0^t e^{L'(t-\tau)} d_{bL}(\mu_\tau, \lambda_\tau)\, d\tau, \end{aligned}$$

and by choosing $\alpha$ large enough we get

$$\begin{aligned}
d_\alpha(\mathcal{F}\{\mu\},\mathcal{F}\{\lambda\}) &= \sup_{t\in[0,T]} d_{bL}(\mathcal{F}(\{\mu\})(t),\mathcal{F}(\{\lambda\})(t))e^{-\alpha t}\\
&\leq \max\{L,B\}\sup_{t\in[0,T]}\left[e^{-\alpha t}\int_0^t e^{L'(t-\tau)}e^{\alpha\tau}d_\alpha(\{\mu\},\{\lambda\})\,d\tau\right]\\
&\leq \frac{\max\{L,B\}}{\alpha-L'}d_\alpha(\{\mu\},\{\lambda\})\\
&\leq \gamma\, d_\alpha(\{\mu\},\{\lambda\}),
\end{aligned}$$

with $\gamma\in(0,1)$.

This proves that $\mathcal{F}: C_\mathcal{P}\to C_\mathcal{P}$ is a contraction, so it has a unique fixed point. □

In this Chapter we have discussed a possible way to reduce a system of $N$ particles, in the limit $N\to\infty$, to a single partial differential equation in the phase space. To properly justify this passage to the limit we have needed to assume strong regularity assumptions on the potential ($W\in C^1$ to pass to the limit, $W\in C^2$ to ensure uniqueness and quantitative stability).

While the mean-field limit for smooth potential has been well understood for more than three decades, in the singular case the rigorous justification of the mean-field limit is still an open problem, and trying to lower the regularity of $W$ is a great challenge which has produced a lot of work in the latest years. We refer the interested reader to [18, 82, 12] for a detailed discussion of several classical results on the topic, and to the lecture notes [45, 67] and the references therein for several recent developments in the area.

## 4.3 The Vlasov-Poisson equation

In the previous sections we explained a possible way to recover a kinetic equation from the microscopic, Newtonian description of an $N$-body system. In particular, we were interested in the derivation and well-posedness of the Vlasov equation, and we rigorously justified the mean-field limit for smooth potentials.

Unfortunately, in the most relevant physical situations, the interaction potential is not smooth at all. Such is the case in particular for one of the most important nonlinear

Vlasov equations, namely the *Vlasov-Poisson equations*, where $W$ is the fundamental solution of $\pm\Delta$ :

$$(VPE)\ \begin{cases} \partial_t f + v \cdot \nabla_x f + F(t,x) \cdot \nabla_v f = 0 \\ F = \pm\nabla\Delta^{-1}\rho \quad\ \rho(t,x) = \int f(t,x,v)dv. \end{cases}$$

Notice that, up to a change of sign in the interaction, the same equation describes both plasma systems, where the particles are ions and electrons, and galaxies, in which each star counts as one particle.

At the beginning of the last century the astrophysicist Sir J. Jeans used this system to model stellar clusters and galaxies [65] and to study their stability properties. In this context it appears in many textbooks on astrophysics such as [10, 42]. In the repulsive case, this system was introduced by A. A. Vlasov around 1937 [96, 97]

Because of the considerable importance in plasma physics and in astrophysics, there is a huge literature on the Vlasov-Poisson system.

The global existence and uniqueness of classical solutions of the Cauchy problem for the Vlasov-Poisson system was obtained by Iordanskii [66] in dimension 1, Ukai-Okabe [90] in the 2-dimensional case, and independently by Lions-Perthame [75] and Pfaffelmoser [86] in the 3-dimensional case. To our knowledge, there are currently no results about existence and uniqueness of classical solutions in dimension greater than 3.

It is important to mention that, parallel to the existence of classical solutions, there have been a considerable amount of work on the existence of weak solutions, in particular under very low assumptions on the initial data. Since in our case we shall never deal with any of the issues related to dealing with weak solutions, we just mention the classical result by Arsen'ev [7], who proved global existence of weak solutions under the hypothesis that $f(0)$ is bounded and has finite kinetic energy, and the result of Horst and Hunze [63], where the authors relax the integrability assumption on $f(0)$.

If one wishes to relax even more then integrability assumptions on the initial data then one enters into the framework of the so called renormalized solutions introduced by Di Perna and Lions [33, 34, 35]. The interested reader is referred to the recent papers [2, 11] for more details and references.

### 4.3.1 Basic properties of the Vlasov equation

In the following Sections we will be focused on the study of the Vlasov-Poisson system in the context of plasma physics. Before recalling some basic properties of plasmas and the main features of the Vlasov-Poisson system, we do some general considerations on the qualitative behaviour of the Vlasov equation.

$$(VE)\begin{cases} \partial_t f + v\cdot\nabla_x f + F(t,x)\cdot\nabla_v f = 0, \\ F = -\nabla_x W *_x \rho, \quad \rho(t,x) = \int f(t,x,v)\,dv \\ f|_{t=0} = f_0 \geq 0, \quad \int_{X^d\times\mathbb{R}^d} f_0\,dx\,dv = 1, \end{cases} \tag{4.3.1}$$

**Boundary conditions**

Boundary conditions deeply affect the qualitative behaviour of the Vlasov equation and a thorough discussion on the topic is beyond the aims of this work. Thus, we assume the position space to be either $X^d = \mathbb{R}^d$, or the $d$-dimensional torus $X^d = \mathbb{T}^d$. The latter case is widely used in physics to model confined plasma. Moreover, this is still by far the simplest way to describe a confined geometry, avoiding effects such as dispersion at infinity which completely change the qualitative behaviour of the nonlinear Vlasov equation.

In the case of the Poisson coupling, the total potential $W*\rho$ is defined as $W = \pm\Delta^{-1}\rho$, which makes sense on $\mathbb{T}^d$ only if $\rho$ has zero mean. The natural solution consists in removing the mean of $\rho$ and consider $W*(\rho - \langle\rho\rangle)$, where $\langle\rho\rangle = \int \rho\,dx$. An informal justification is the following: in the plasma model, one may argue that the density of ions should be taken into account. Thus, to preserve the global neutrality of the plasma, we can assume the density of ions to be equal to the mean density of electrons. Such a reasoning is based on the existence of two different species of particles, but even if there is just one species of particles, as in the case for gravitational interaction, it is still possible to remove the mean of $\rho$. This procedure is known as *Jeans swindle* in galactic dynamics. Recently Kiessling [51] rigorously justified the Jeans swindle using a limit procedure.

If $W$ is a given potential in $\mathbb{R}^d$, then, in a periodic setting, it should be replaced by its periodic version

$$W^{\mathrm{per}}(x) = \sum_{k\in\mathbb{Z}^d} W(x-k).$$

Thus, if $W$ decays fast enough at infinity and $\rho$ is periodic,

$$\nabla W * \rho = \nabla W^{\mathrm{per}} * \rho = \nabla W^{\mathrm{per}} * (\rho - \langle\rho\rangle).$$

In our case, since the Coulomb potential does not decay fast enough at infinity, we may approximate $W$ by some potential $W_\epsilon$ exhibiting a "cutoff" at large distances. Then, if

$\nabla W_\varepsilon \in L^1(\mathbb{R}^d)$, the convolution $\nabla W_\varepsilon * \rho$ makes sense for a periodic $\rho$. Moreover

$$\nabla W_\varepsilon * \rho = \nabla W_\varepsilon^{\text{per}} * (\rho - \langle \rho \rangle).$$

Passing to the limit for $\varepsilon \to 0$, one gets

$$\int_{\mathbb{T}^d} \nabla W^{\text{per}}(x-y)\rho(y)dy = \int_{\mathbb{T}^d} \nabla W^{\text{per}}(x-y)\left(\rho - \langle \rho \rangle\right)(y)dy = \pm \nabla \Delta_{\text{per}}^{-1}\left(\rho - \langle \rho \rangle\right),$$

where $\Delta_{\text{per}}^{-1}\left(\rho - \langle \rho \rangle\right)$ is the inverse Laplace operator on $\mathbb{T}^d$. We refer to [51] for a detailed explanation of the physics meaning of this procedure.

### Structure of the Vlasov-Poisson equation and invariants

In contrast to models incorporating collisions, the Vlasov equation is *time reversible*, i.e. if $f = f(t,x,v)$ is a solution, then $g(t,x,v) := f(-t,x,-v)$ is also a solution where time and velocity have been reversed. This means that the Vlasov-Poisson equation inherited the reversibility feature of the Newton equations in the mean-field limit. As a consequence, the nonlinear Vlasov equation does not have any regularizing effect.

The nonlinear Vlasov equation is a *transport equation* and can be described by the *method of characteristics*: if $f(t,x,v)$ solves the equation, then the measure $f(t,x,v)dx\,dv$ is the *push-forward* of the initial measure $f_0(x,v)dx\,dv$ by the flow $\Phi_t = (X_t, V_t)$ in phase space, solving the characteristic equations:

$$\begin{cases} \dot{X}_t = V_t \\ \dot{V}_t = F(X_t,t) \quad \text{with} \quad F(X_t,t) = -\nabla W * \rho \\ (X_0, V_0) = (x,v). \end{cases}$$

The Vlasov equation (4.3.1) is a Hamiltonian system with the so-called Lie-Poisson bracket structure, and (as we already mentioned in the previous Chapters) can be expressed as Liouville equation

$$\partial_t f + \left[H_{f(t,\cdot,\cdot)}, f\right] = 0$$

where the Hamiltonian is given by

$$H_{f(t,\cdot,\cdot)}(x,v) = \frac{1}{2}|v|^2 + \iint_{X^d \times \mathbb{R}^d} W(x-y) f(t,y,w)dydw.$$

A consequence of this property is that the flow $\Phi_t$ induced by $F(t,x)$ has Jacobian determinant equal to 1, thus the flow preserves the Liouville measure $dx\,dv$ and the push-forward equation

$$f(t,x,v)dx\,dv = \Phi_t \# f_0(x,v)dx\,dv$$

becomes a transport condition for densities:

$$f(t,x,v) = f(0,(\Phi_t)^{-1}(x,v)).$$

Let us remark that the Proof of Theorem 4.2.2 we already relied on the structure of the Vlasov equation as a transport equation to prove both existence and uniqueness of solutions.

We can also consider the Vlasov equation in terms of the infinite-dimensional space of distributions, and the physical observables $\mathcal{F}[f]$, which are functionals on this space. In this framework, the Vlasov equation is equivalent to the Hamiltonian equation (see for instance [79])

$$\frac{d\mathcal{F}}{dt} + \{\mathcal{H}, \mathcal{F}\} = 0$$

where $\mathcal{H}$ is the Hamiltonian whose functional derivative is $H_{f(t,\cdot,\cdot)}$, i.e., $\delta\mathcal{H}/\delta f = H_{f(t,\cdot,\cdot)}$, and $\{\,;\}$ is a Lie-Poisson bracket, defined by

$$\{\mathcal{F}_1, \mathcal{F}_2\} = \int f \left[\frac{\delta\mathcal{F}_1}{\delta f}, \frac{\delta\mathcal{F}_2}{\delta f}\right] dx\,dv. \tag{4.3.2}$$

Regarding the distribution of particles itself to be a functional

$$f(t,x,v) = \iint f(t,x',v')\delta((x,v)-(x',v'))\,dx'dv'$$

we recover the Vlasov equation from (4.3.2) in the following way:

$$\partial_t f = \{f, \mathcal{H}\} = -\left[f, H\right].$$

For a derivation of this Lie-Poisson bracket from the canonical Hamiltonian formalism for the particle motion, see [69, 98]. An important property of the Lie-Poisson bracket is the existence of an infinite number of observables of the form

$$\mathcal{C}[f] = \iint C(f(x,v))\,dx\,dv$$

where $C(f)$ is an arbitrary smooth function, which commutes with any Hamiltonian $\mathcal{H}$ and thus constitutes an infinite number of invariants of motion, one for each choice of $C(f)$. These invariants are known as the *Casimirs* of the equation; their level sets foliate the space of distributions into invariant subspaces on which the dynamics is constrained.

It is a fundamental question whether a certain bracket actually correspond to a valid Hamiltonian structure if one wants to go further than a formal rewriting of the equation. A milestone on the topic is certainly the work of Arnold who was able to formulate a stability theorem for plane flows using a method now known as the *Energy-Casimir method* [5, 6], see also [83].

Coming back to the main features of Vlasov equations, we observe that an immediate consequence of the Hamiltonian nature of the Vlasov equation is the conservation of all $L^p$ norms for $p \in [1, \infty]$, as well as the *entropy*

$$S = -\iint f \log f \, dx \, dv.$$

Entropy conservation is in contrast with what happen for collisional equations, as for instance the Boltzmann equation, for which the entropy can only increase in time, unless it is at equilibrium. This property can be also thought ad *preservation of information*: whatever information we have about the distribution of particles at initial time, is preserved in time.

In addition, the Vlasov equation preserves the *total energy*:

$$H(x,v,t) := \iint f(t,x,v)\frac{|v|^2}{2} dx\, dv + \frac{1}{2}\iint W(x-y)\rho(x)\rho(y) dx dy = T + U.$$

The total energy is the sum of the kinetic energy $T$ and the potential energy $U$ and it is a kind of "mesoscopic Hamiltonian" corresponding to the average of the microscopic Hamiltonian against the particle distribution. The factor $1/2$ in the definition of the potential energy $U$ comes from the fact that we count unordered pairs of particles.

Before moving on, we underline that some of these features are shared by other partial differential equations, in particular the two-dimensional incompressible Euler equation, for which a classical reference is [77]. The similarity between the Vlasov equation and the two-dimensional Euler equation with nonnegative vorticity is well-known; and mathematicians try systematically to adapt tools and results from one equation to the other.

### General considerations about uniqueness for Vlasov-Poisson

We now make some basic considerations about the analytical difficulties arising in the study of the Vlasov-Poisson equation.

Notice that in $\mathbb{R}^d$ (that is, $W = -\Delta^{-1}\rho$) the interaction potential is given by

$$W(t,x) = \frac{c}{|x|^{d-2}} * \rho(t).$$

In particular we see that $W \geq 0$, so the conservation of energy implies that the kinetic energy is bounded uniformly in time.

Hence, if we start from a nice initial condition $f_0$ (say smooth and compactly supported), the basic informations that are propagated in time are

$$\|f(t)\|_{L^\infty(\mathbb{R}^d\times\mathbb{R}^d)} \leq C, \qquad \int_{\mathbb{R}^d\times\mathbb{R}^d} |v|^2 f(t,x,v)\,dx\,dv \leq C \qquad \forall\, t \geq 0.$$

While in this case one can prove existence of weak solutions [7], the existence and uniqueness of classical solutions is a very delicate issue. Indeed, assuming that $f(0)$ is smooth, to propagate some regularity of $f$ in time one needs to have some smoothness of the force field $F(t) = \nabla\Delta^{-1}\rho(t)$, whose regularity is strictly related to the one of $\rho(t)$.

To explain this point better, let us just focus on the uniqueness issue. As we have seen in the proof of Theorem 4.2.2, to prove uniqueness one would like to know that the force field $F(t) = \nabla\Delta^{-1}\rho(t)$ is Lipschitz. Since, by elliptic regularity, $F(t)$ has one derivative more than $\rho(t)$, for $F(t)$ to be Lipschitz we would like to know that $\rho(t)$ is bounded (this is not completely true since elliptic regularity fails in $L^\infty$, but this heuristic argument is correct).

Hence, the main issue behind uniqueness consists proving that $\rho(t) \in L^\infty$ for all times, and this indeed at the basis of the uniqueness argument in [75] (see also [76] for a different proof based on stability in the Wasserstein metric, and also our proof of Theorem 6.3.1). However, in order to deduce boundedness of $\rho(t)$ from the one of $f(t)$, one would like to know that $f$ is compactly supported in the velocity space: indeed, if $\mathrm{supp} f(t,x,\cdot) \subset B_{R(t)}$ for some radius $R(t)$ then

$$\rho(t,x) = \int_{B_{R(t)}} f(t,x,v)\,dv \leq \|f(t)\|_{L^\infty(\mathbb{R}^d\times\mathbb{R}^d)}\,|B_{R(t)}| \qquad \forall\, x,$$

which proves that $\rho(t)$ is bounded. This shows that the main difficulty behind uniqueness of solutions is to show that if $f(0)$ is compactly supported, then it remains so for all times.

While this has been proved up to dimension 3, the existence and uniqueness of smooth solutions is currently unknown in higher dimension.

This brief discussion shows just one of the main mathematical difficulty behind the PDE theory of Vlasov-Poisson. For a general presentation of the mathematical analysis of this system, the interested reader is referred to [44, 16, 88].

### Collisionless relaxation: Landau damping

We conclude this Chapter with a discussion about the large-time behavior in the Vlasov-Poisson equation.

An archetypal continuous model for $N$ body systems is the Boltzmann equation, when interactions among particles occur on a scale that is negligible compared to the spatial scale. In its simplest form, the Boltzmann equation is obtained from Newtonian dynamics in the Boltzmann-Grad limit, that is, when the number of particles becomes extremely large while the total cross-section remains of order unity. It reads

$$\partial_t f + v \cdot \nabla_x f = Q(f, f),$$

where $Q$ is the Boltzmann collision operator, which is localised in space-time and takes into account the effect of collisions of particles; these collisions are classically assumed to be elastic, that is, energy-preserving. The well known $H$ theorem shows that for the classical elastic Boltzmann equation, under the action of uncorrelated collisions, the Boltzmann entropy cannot decrease in time. This means that, even if Newton equations are mechanically reversible, the Boltzmann equation is an irreversible model where the density $f$ evolves towards a maximal entropy state. The latter observation provides a guideline for the long-time behaviour of the Boltzmann equation, see for instance [29, 43] for a complete discussion.

Up to now we stressed the reversible nature of the Vlasov-Poisson equation and the presence of infinitely many conservation laws, including the entropy and the energy of the system. This behaviour is in sharp contrast with what happen for collisional models. Another difference between the collisional and the mean field models concerns equilibria. Indeed, while the Boltzmann equation only has Gaussian equilibria, the Vlasov-Poisson equation has infinitely many shapes of equilibria: in the absence of an external field or boundaries, this includes in particular homogeneous distributions $f_0(v)$, and also many inhomogeneous periodic stationary solutions. This seems to oppose the idea that solutions of the Vlasov-Poisson equation would have a definite large time behaviour. It thus was surprising when, in 1946, Landau [71, 72] stunned the physical community by predicting an irreversible behaviour for the Vlasov-Poisson equation. Landau's proof relied on the solution of the Cauchy problem for the *linearized* Vlasov-Poisson equation around a spatially homogeneous Maxwellian equilibrium. In fact, when one linearizes the equation around a homogeneous equilibrium, the resulting linear equation can be turned into a completely integrable system. Then, once one formally solves the equation by Fourier and Laplace transform, an analysis of singularities in the complex plane leads to the conclusion that the electric field decays exponentially fast. Since Landau's analysis concerns the linearized case, his result does not guarantee that the asymptotic behaviour of the linear Vlasov equation is an approximation of the asymptotic behaviour of the nonlinear Vlasov-Poisson equation. Nevertheless, it suggested that a homogenization mechanism may occur in a collisionless system, even if the entropy does not increase. This entropy-preserving relaxation means that mixing trajectories generate fast kinetic oscillations which globally compensate each other in the velocity averaging procedure leading from the distribution function to the force field.

This discovery has been extremely influential, and led to considerable speculation about its driving mechanism. For a historical background on the Landau damping, we refer to [81, Section 1].

The first results in a nonlinear setting was proved by Caglioti-Maffei [27]. They consider the one dimensional torus and use fixed-point theorems and perturbative arguments to prove that there exists a class of solutions of the nonlinear Vlasov-Poisson equation that converge weakly, as $t$ tends to infinity, to a stationary homogeneous equilibrium. Since solutions of free transport weakly converge to spatially homogeneous distributions, the solutions constructed by this "scattering" approach are indeed damped. They also noted that this implies, by time-reversibility, the instability in the weak topology. Let us mention that another construction to damped solutions in the one dimensional setting was also performed by Hwang and Velazquez [64].

The gap between the linear and nonlinear theory of Landau damping was only bridged recently by Mouhot and Villani [81] who showed that the Landau damping survives nonlinearity, and proved the first rigorous result that establish an exponential decay to equilibrium in confined, collisionless and time-reversible dynamics. In that paper, the damping phenomenon is reinterpreted in terms of exchanges of regularity between spatial and kinetic modes, rather than energy, showing that it is driven by the phase mixing mechanism associated with the trajectories of particles. Some of their new tools are the introduction of families of analytic norms measuring regularity by comparison to solutions of the free transport equation, distinctive functional inequalities, a control of nonlinear echoes, sharp scattering estimates in analytic regularity, and a Newton approximation scheme, whose extremely fast convergence is fully exploited.

As already mentioned before, there are strong similarities between the Vlasov-Poisson equation and the 2d Euler equation for incompressible fluids, although the latter model is more singular. In particular, the 2d Euler equation should exhibit the same kind of relaxation to equilibrium, and this has been rigorously proved by Bedrossian and Masmoudi for the Couette flows: in [19] the authors establish asymptotic stability of shear flows close to the planar Couette flow. Their proof requires several new ideas and tools, in particular a delicate paraproduct decompositions and controlled regularity loss. We also mention that a combination of the techniques from [81] and [19] has recently allowed Bedrossian, Masmoudi, and Mouhot to give a new simpler proof of nonlinear Landau damping in Gevrey regularity under less restrictive assumptions on the size of the perturbations.

## 4.4 Quasineutral limit

In this Section we recall some physical properties of plasmas and we introduce the Debye length, one of the characteristic parameter to describe a plasma. Then, we introduce the kinetic models investigated in the thesis and the *quasineutral limit*.

### 4.4.1 What is a plasma?

When the temperature of a material grows, its state changes from solid to liquid and then to gas. If the temperature further increases, the atoms are ionized and the gas reach a new state of the matter in which the charge numbers of ions and electrons are almost the same and charge neutrality is globally achieved.

In 1927, the American Nobel prize Irving Langmuir first used the term *plasma* to describe a ionized gas, since the way blood plasma carries red and white corpuscles reminded him the way a ionized gas carries electrons and ions. Langmuir, in collaboration with his colleague Lewi Tonks, was investigating the physics and chemistry of tungsten-filament light-bulbs, with the purpose to find a way to extend the lifetime of the filament, a goal which he eventually achieved. After Langmuir, research on plasma gradually spread in other directions, and many kind of plasmas have been created for industrial purposes. Moreover, research in plasma physics has been driven by the aim to create and confine hot plasmas in fusion research.

Plasma is often called the fourth state of matter and it is the most abundant form of ordinary matter in the Universe, most of which is in the rarefied intergalactic regions, and in stars. Although it is closely related to the gas phase, since it also has no definite form or volume, it has some peculiarities that distinguish it and determine a completely different behaviour. Indeed, when electrons move, they interact with the long range Coulomb force and these interactions create electric and magnetic fields following Maxwell equations. Thus, various collective movements occur in the plasma, along with many kinds of instabilities and wave phenomena.

As we already mentioned, the complete model of collisionless plasma describes the behaviour of two different species of particles: ions and electrons. As the ratio of the masses of the ion and the electron is of several orders of magnitude, we are allowed to assume that the ions are at rest, and uniformly distributed. The latter observation was already mentioned when we discussed boundary conditions for the Vlasov equation in Section 4.3.1. The ratio of the masses of the ion and the electron is just one of the problems related to the presence in plasmas of typical parameters with dramatic differences at level of magnitude. One may also think to electric permittivity in the vacuum and the Debye length compared to the magnetic and the electric fields and the observation length. In the following we will introduce the Debye length, which has an important role in the study of plasmas, and the quasineutral limit.

### 4.4.2 The Vlasov-Poisson equation for massless electrons

In the kinetic description of a plasma, we usually consider the point of view of electrons, from which ions are very slow, motionless at equilibrium. This assumption leads to the classical Vlasov-Poisson system:

$$(VPE)\begin{cases} \partial_t f + v\cdot\nabla_x f + E(t,x)\cdot\nabla_v f = 0, \\ E = -\nabla_x W \quad \rho(t,x) = \int f(t,x,v)\,dv \\ -\Delta W = \rho - 1 \\ f|_{t=0} = f_0 \geq 0, \quad \int_{X^d\times\mathbb{R}^d} f_0\,dx\,dv = 1, \end{cases}$$

On the other hand, we can consider the viewpoint of ions, assuming that the electrons have 0 mass. In this case, the electrons move very fast and they reach their local thermodynamic equilibrium quasi-instantaneously. Notice that, since the mass of the electrons is negligible with respect to the mass of the ions, the typical collision frequency for the electrons is much bigger than for the ions, so collisions for the electrons may be not negligible and they can reach their local thermodynamic equilibrium. Then, their density $n_e$ follows the Maxwell-Boltzmann law [74]

$$n_e = \int f_e dv = g(x) \exp\left(\frac{eW}{k_B T_e}\right),$$

where $W$ denotes the Coulomb potential, $k_B$ is the Boltzmann constant, $T_e$ the average temperature of the electrons, and $g \in L^1(\mathbb{R}^n)$ is a term due to an external potential preventing the particles from going to infinity (we refer to [21] and references therein for more details).

More precisely, we have:

$$g(x) = n_0 \exp\left(-\frac{H(x)}{k_B T_e}\right),$$

where $H$ denote the external confining potential and $n_0 \in \mathbb{R}$ is a normalizing constant. Thus, the Poisson equation becomes:

$$-\Delta_x W = \int f\, dv - g \exp\left(\frac{eW}{k_B T_e}\right).$$

Let us remark that, in this situation, the total number of electrons is not a priori fixed. This means that the global neutrality of the plasma is not anymore satisfied:

$$\int \left(\int f\, dv - g \exp\left(\frac{eW}{k_B T_e}\right)\right) dx \neq 0.$$

However, one may also focus on the case when the total charge of the electrons is fixed, and in this case the Poisson equation reads:

$$-\Delta_x W = \int f\, dv - \frac{g \exp\left(\frac{eW}{k_B T_e}\right)}{\int_{\mathbb{R}^d} g \exp\left(\frac{eW}{k_B T_e}\right)}.$$

The existence of global weak solutions to these two systems in dimension three has been investigated by Bouchut [21]. A natural approximation of the latter equation comes from the linearization of the exponential law. This approximation is valid from the physical point of view as long as the electric energy is small compared to the kinetic energy:

$$\frac{eW}{k_B T_e} \ll 1$$

In the following we will focus on models with such Maxwell-Boltzmann laws on the torus $\mathbb{T}^d$ with $d = 1, 2$ or 3, thus we do not need a confining potential, and we take $g = 1$.

### 4.4.3 The Debye length

The Debye length is one of the fundamental length scales in plasma physics. It describes a screening distance, beyond which charges are unaware of other charges. The Debye sphere is a sphere whose radius is the Debye length, outside of which charges are electrically screened.

It is defined as $\lambda_D^{(\alpha)}$

$$\lambda_D^{(\alpha)} = \sqrt{\frac{\varepsilon_0 k_B T_\alpha}{N_\alpha e^2}},$$

where $k_B$ is the Boltzmann constant, $T_\alpha$ and $N_\alpha$ are respectively the average temperature and density of electrons (for $\alpha = e$) or ions (for $\alpha = i$). This parameter is of tremendous importance in plasmas. It can be interpreted as the typical length below which charge separation occurs. In plasmas, this length may vary by many orders of magnitude (typical values go from $10^{-3}$ m to $10^{-8}$ m). In practical situations, for terrestrial plasmas, it is always small compared to the other characteristic lengths in consideration, in particular the characteristic observation length, denoted by $L$. Actually, the condition $\lambda_D \ll L$ is sometimes required in the definition itself of a plasma. Therefore, if we set

$$\lambda_D = \varepsilon \ll 1,$$

then in many regimes, after considering dimensionless variables, it is relevant to observe that the Poisson equation formally reads

$$-\varepsilon^2 \Delta W_\varepsilon = \pm(n_i - n_e).$$

The *quasineutral limit* precisely consists in considering the limit $\varepsilon \to 0$.

#### Why quasineutral limits?

From the numerical point of view, kinetic equations are harder to handle than fluid equations. Indeed the main difficulty is that kinetic equations live on a phase space of dimension $2d$ (for $x, v \in \mathbb{R}^d$). Actually, another problem for simulating plasmas is the following: there are characteristic lengths and times of completely different magnitude (think of the Debye length and the observation length) that make numerics really delicate.

In our works our purpose is to get, in the quasineutral limit, simplified hydrodynamic systems. In the following we give an idea of the reasons that make fluid descriptions more convenient:

- First, using a fluid description, we deal with a lower dimensional phase space. Moreover, after taking the limit, we obtain a system with only one characteristic time and length. Both these reasons make numerical simulations easier to perform. Of course it is well-known that the fluid approximation is not always accurate for simulations of plasmas, but it is nevertheless valid in some regimes that we may describe in the analysis. So it is important to be aware of the physical assumptions we make when we derive the equations.

- Macroscopic quantities, such as charge or current density, are easier to measure in experiments.

- A simplified fluid description allows one to give a better qualitative description of the behaviour of the plasma.

# Chapter 5

# The quasineutral limit of the Vlasov-Poisson equation in Wasserstein metric

[1]

## 5.1 Introduction

In this Chapter we study the Vlasov-Poisson system in the presence of massless thermalized electrons. We shall focus on the one-dimensional case and consider that the equations are set on the phase space $\mathbb{T} \times \mathbb{R}$ (we will sometimes identify $\mathbb{T}$ to $[-1/2, 1/2)$ with periodic boundary conditions). The system, which we shall refer to as the Vlasov-Poisson system with *massless electrons*, encodes the fact that electrons move very fast and quasi-instantaneously reach their local thermodynamic equilibrium. It reads as follows:

$$(VPME) := \begin{cases} \partial_t f + v \cdot \partial_x f + E \cdot \partial_v f = 0, \\ E = -U', \\ U'' = e^U - \int_{\mathbb{R}} f \, dv =: e^U - \rho, \\ f|_{t=0} = f_0 \geq 0, \quad \int_{\mathbb{T}\times\mathbb{R}} f_0 \, dx \, dv = 1. \end{cases} \tag{5.1.1}$$

[1]This chapter is based on a joint work with Danial Han-Kwan [59].

Here, as usual, $f(t, x, v)$ stands for the distribution function of the ions in the phase space $\mathbb{T} \times \mathbb{R}$ at time $t \in \mathbb{R}^+$, while $U(t, x)$ and $E(t, x)$ represent the electric potential and field respectively, and $U'$ (resp. $U''$) denotes the first (resp. second) spatial derivative of $U$. In the Poisson equation, the semi-linear term $e^U$ stands for the density of electrons, which therefore are assumed to follow a Maxwell-Boltzmann law.

We are interested in the behavior of solutions to the (VPME) system in the so-called *quasineutral limit*, i.e., when the Debye length of the plasma vanishes. Loosely speaking, the Debye length can be interpreted as the typical scale of variations of the electric potential. It turns out that it is always very small compared to the typical observation length, so that the quasineutral limit is relevant from the physical point of view. As a result, the approximation which consists in considering a Debye length equal to zero is widely used in plasma physics, see for instance [30]. This leads to the study of the limit as $\varepsilon \to 0$ of the scaled system:

$$(VPME)_\varepsilon := \begin{cases} \partial_t f_\varepsilon + v \cdot \partial_x f_\varepsilon + E_\varepsilon \cdot \partial_v f_\varepsilon = 0, \\ E_\varepsilon = -U_\varepsilon', \\ \varepsilon^2 U_\varepsilon'' = e^{U_\varepsilon} - \int_{\mathbb{R}} f_\varepsilon \, dv =: e^{U_\varepsilon} - \rho_\varepsilon, \\ f_\varepsilon|_{t=0} = f_{0,\varepsilon} \geq 0, \quad \int_{\mathbb{T}\times\mathbb{R}} f_{0,\varepsilon} \, dx \, dv = 1. \end{cases} \tag{5.1.2}$$

The formal limit is obtained in a straightforward way by taking $\varepsilon = 0$ (which corresponds to a Debye length equal to 0):

$$(KIE) := \begin{cases} \partial_t f + v \cdot \partial_x f + E \cdot \partial_v f = 0, \\ E = -U', \\ U = \log \rho, \\ f_0 \geq 0, \quad \int_{\mathbb{T}\times\mathbb{R}} f_0 \, dx \, dv = 1, \end{cases} \tag{5.1.3}$$

a system we shall call the *kinetic isothermal Euler system*.

We point out that there are variants of the (VPME) system which are also worth studying, such as the linearized (VPME), in which semi-linear term in the Poisson equation is linearized (this turns out to be a standard approximation in plasma physics, see also [54, 55, 57, 58]),

$$U'' = U + 1 - \rho,$$

and the Vlasov-Poisson system for electrons with fixed ions (the most studied model in the mathematical literature), in which the Poisson equation reads as follows

$$U'' = 1 - \rho,$$

which we shall refer to as the *classical* Vlasov-Poisson system. As a matter of fact, our results concerning the (VPME) system have analogous statements for the linearized (VPME) or the classical Vlasov-Poisson system. We have made the choice to study the (VPME) system since the semi-linear term in the Poisson equation creates additional interesting difficulties. As we shall mention in Remark 5.1.5, our analysis applies as well, *mutatis mutandis*, to these models, and actually provides a stronger result in terms of the class of data that we are allowed to consider.

The justification of the limit $\varepsilon \to 0$ from (5.1.2) to (5.1.3) is far from trivial. Indeed, this is known to be true only in few cases (see also [22, 48, 58] for further insights): when the sequence of initial data $f_{0,\varepsilon}$ enjoys uniform analytic regularity with respect to the space variable (as we shall describe later in Section 5.4.1, this is just an adaptation of a work of Grenier [49] on the classical Vlasov-Poisson system); when $f_{0,\varepsilon}$ converge to a Dirac measure in velocity $f_0(x, v) = \rho_0(x)\delta_{v=v_0(x)}$ (see [55] and [24, 78, 46]); and, following [58], when $f_{0,\varepsilon}$ converge to a homogeneous initial condition $\mu(v)$ which is symmetric with respect to some $\overline{v} \in \mathbb{R}$ and which is first increasing then decreasing. Also, it is conjectured (see [50]) that this result should hold when the sequence of initial data $f_{0,\varepsilon}$ converges to some $f_0$ such that, for all $x \in \mathbb{T}$, $v \mapsto f_0(x, v)$ satisfies a stability condition *a la Penrose* [85] (typically when $v \mapsto f_0(x, v)$ is increasing then decreasing). On the other hand, the limit is known to be false in general, as we will explain later.

In this work, we shall study this convergence issue in a Wasserstein metric. More precisely, we consider the distance between finite (possibly signed) measures given by

$$W_1(\mu, \nu) := \sup_{\|\varphi\|_{\mathrm{Lip}} \leq 1} \left( \int \varphi \, d\mu - \int \varphi \, d\nu \right),$$

where $\|\cdot\|_{\mathrm{Lip}}$ stand for the usual Lipschitz semi-norm and which is referred to as the 1-Wasserstein distance (see Section 1.2 for more details and references). We recall that $W_1$ induces the weak topology on the space of Borel probability measures with finite first moment, that we denote by $\mathcal{P}_1(\mathbb{T} \times \mathbb{R})$. Notice that, since $\mathbb{T}$ is compact, this corresponds to measures $\mu$ with finite first moment in velocity.

As it will also be clear from our arguments, $W_1$ is particularly suited to estimate the distance between solutions to kinetic equations. Indeed, for Vlasov-Poisson equations, it is very natural to consider atomic solutions (that it, measures concentrated on finitely many points) and $W_1$ is able to control the distance between the supports, while other classical distances (as for instance the total-variation) are too rough for this.

Before stating our convergence results, we first deal with the existence of global weak solutions in $\mathcal{P}_1(\mathbb{T} \times \mathbb{R})$.

**Theorem 5.1.1.** *Let $f_0$ be a probability measure in $\mathcal{P}_1(\mathbb{T} \times \mathbb{R})$, that is,*

$$\int |v| \, df_0(x, v) < \infty. \tag{5.1.4}$$

*Then there exists a global weak solution to* (5.1.2) *with initial datum $f_0$.*

The analogous result for the classical Vlasov-Poisson equation was proved by Zheng and Majda [99], and more recently by Hauray [61] with a new proof.

We shall prove Theorem 5.1.1 by combining the method introduced by Hauray (see [61]) with new stability estimates for the massless electron system.

Roughly speaking, the main results of this Chapter are the following: if we consider initial data for (5.1.2) of the form $f_{0,\varepsilon} = g_{0,\varepsilon} + h_{0,\varepsilon}$ with $g_{0,\varepsilon}$ analytic (or equal to a finite sum of Dirac masses in velocity, with analytic moments) and $h_{0,\varepsilon}$ converging very fast to 0 in the $W_1$ distance, then the solution starting from $f_{0,\varepsilon}$ converges to the solution of (5.1.3) with initial condition $g_0 := \lim_{\varepsilon \to 0} g_{0,\varepsilon}$. This means that small perturbations in the $W_1$ distance do not affect the quasineutral limit. Notice that the fact that the size of the perturbation has to be small only in $W_1$ means that $h_{0,\varepsilon}$ could be arbitrarily large in any $L^p$ norm.

To state our main results, we first introduce some notation. The following analytic norm has been used by Grenier [49] to show convergence results for the quasineutral limit in the context of the classical Vlasov-Poisson system.

Such a norm is useful to study the quasineutral limit since the formal limit is false in general in Sobolev regularity (see Proposition 5.1.6 and the discussion below); one can also see that the formal limit equation exhibits a loss of derivative (in the force term), which can be overcome with analytic regularity.

**Definition 5.1.2.** *Given $\delta > 0$ and a function $g : \mathbb{T} \to \mathbb{R}$, we define*

$$\|g\|_{B_\delta} := \sum_{k \in \mathbb{Z}} |\widehat{g}(k)| \delta^{|k|},$$

*where $\widehat{g}(k)$ is the $k$-th Fourier coefficient of $g$. We also define $B_\delta$ as the space of functions $g$ such that $\|g\|_{B_\delta} < +\infty$.*

**Theorem 5.1.3.** *Consider a sequence of non-negative initial data in $\mathcal{P}_1(\mathbb{T} \times \mathbb{R})$ for (5.1.2) of the form*

$$f_{0,\varepsilon} = g_{0,\varepsilon} + h_{0,\varepsilon},$$

*where $(g_{0,\varepsilon})$ is a sequence of continuous functions satisfying*

$$\sup_{\varepsilon\in(0,1)} \sup_{v\in\mathbb{R}} (1+v^2)\|g_{0,\varepsilon}(\cdot,v)\|_{B_{\delta_0}} \leq C,$$

$$\sup_{\varepsilon\in(0,1)} \left\| \int_{\mathbb{R}} g_{0,\varepsilon}(\cdot,v)\,dv - 1 \right\|_{B_{\delta_0}} < \eta,$$

*for some $\delta_0, C, \eta > 0$, with $\eta$ small enough, and admitting a limit $g_0$ in the sense of distributions.*

*There exists a function $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$, with $\lim_{\varepsilon\to 0^+} \varphi(\varepsilon) = 0$, such that the following holds.*

*Assume that $(h_{0,\varepsilon})$ is a sequence of measures with finite first moment, satisfying*

$$\forall \varepsilon > 0, \quad W_1(h_{0,\varepsilon}, 0) \leq \varphi(\varepsilon).$$

*Then there exist $T > 0$ and $g(t)$ a weak solution on $[0,T]$ of (5.1.3) with initial condition $g_0 = \lim_{\varepsilon\to 0} g_{0,\varepsilon}$, such that, for any global weak solution $f_\varepsilon(t)$ of (5.1.2) with initial condition $f_{0,\varepsilon}$,*

$$\sup_{t\in[0,T]} W_1(f_\varepsilon(t), g(t)) \to_{\varepsilon\to 0} 0.$$

*We can explicitly take $\varphi(\varepsilon) = \frac{1}{\varepsilon}\exp\left(\frac{\lambda}{\varepsilon^3}\exp\frac{15}{2\varepsilon^2}\right)$ for some $\lambda < 0$.*

We now state an analogous result for initial data consisting of a finite sum of Dirac masses in velocity:

**Theorem 5.1.4.** *Let $N \geq 1$ and consider a sequence of non-negative initial data in $\mathcal{P}_1(\mathbb{T} \times \mathbb{R})$ for (5.1.2) of the form*

$$f_{0,\varepsilon} = g_{0,\varepsilon} + h_{0,\varepsilon},$$

$$g_{0,\varepsilon}(x,v) = \sum_{i=1}^{N} \rho^i_{0,\varepsilon}(x)\delta_{v=v^i_{0,\varepsilon}(x)},$$

*where the $(\rho^i_{0,\varepsilon},\, v^i_{0,\varepsilon})$ is a sequence of analytic functions satisfying*

$$\sup_{\varepsilon\in(0,1)} \sup_{i\in\{1,\cdots,N\}} \|\rho^i_{0,\varepsilon}\|_{B_{\delta_0}} + \|v^i_{0,\varepsilon}\|_{B_{\delta_0}} \leq C,$$

$$\sup_{\varepsilon\in(0,1)} \left\| \sum_{i=1}^{N} \rho^i_{0,\varepsilon} - 1 \right\|_{B_{\delta_0}} < \eta$$

*for some $\delta_0, C, \eta > 0$, with $\eta$ small enough, and admitting limits $(\rho^i_0, v^i_0)$ in the sense of distributions.*

*There exists a function $\varphi : \mathbb{R}^+ \to \mathbb{R}^+$, with $\lim_{\varepsilon\to 0^+} \varphi(\varepsilon) = 0$, such that the following holds.*

*Assume that $(h_{0,\varepsilon})$ is a sequence of measures with finite first moment, satisfying*

$$\forall \varepsilon > 0, \quad W_1(h_{0,\varepsilon}, 0) \leq \varphi(\varepsilon).$$

*Then there exist $T > 0$, such that, for any global weak solution $f_\varepsilon(t)$ of (5.1.2) with initial condition $f_{0,\varepsilon}$,*

$$\sup_{t\in[0,T]} W_1(f_\varepsilon(t), g(t)) \to_{\varepsilon\to 0} 0,$$

*where*

$$g(t,x,v) = \sum_{i=1}^{N} \rho^i(t,x)\delta_{v=v^i(t,x)},$$

*and $(\rho^i, v^i)$ satisfy the multi fluid isothermal system on $[0,T]$*

$$\begin{cases} \partial_t \rho^i + \partial_x(\rho^i v^i) = 0, \\ \partial_t v^i + v^i \partial_x v^i = E, \\ E = -U', \\ U = \log\left(\sum_{i=1}^{N} \rho^i\right), \\ \rho^i|_{t=0} = \rho^i_0, v^i|_{t=0} = v^i_0. \end{cases} \tag{5.1.5}$$

*We can explicitly take $\varphi(\varepsilon) = \frac{1}{\varepsilon}\exp\left(\frac{\lambda}{\varepsilon^3}\exp\frac{15}{2\varepsilon^2}\right)$ for some $\lambda < 0$.*

**Remark 5.1.5.** *It is worth mentioning that the previous convergence results can be slightly improved when dealing with the classical Vlasov-Poisson equation. Indeed, thanks to Remark 5.2.3, the analogue of Theorems 5.1.3 and 5.1.4 holds for a larger class of initial data. In fact, it is possible to take*

$$\varphi(\varepsilon) = \frac{1}{\varepsilon}\exp\left(\frac{\lambda}{\varepsilon}\right)$$

*for some $\lambda < 0$.*

In the following, we shall say that a function $\varphi$ is admissible if it can be chosen in the statements of Theorems 5.1.3 and 5.1.4.

The interest of these results is the following: they prove that it is possible to justify the quasineutral limit without making analytic regularity or stability assumption. The price to pay is that the constants involved in the explicit functions $\varphi$ above are extremely small, so that we are very close to the analytic regime.

On the other hand one should have in mind the following negative result, which roughly means that the functions $\varphi(\varepsilon) = \varepsilon^s$, for any $s > 0$ are not admissible (this therefore yields a lower bound on admissible functions); this is the consequence of *instability* mechanisms described in [50] and [58].

**Proposition 5.1.6.** *There exist smooth homogeneous equilibria $\mu(v)$ such that the following holds. For any $N > 0$ and $s > 0$, there exists a sequence of non-negative initial data $(f_{0,\varepsilon})$ such that*

$$\|f_{\varepsilon,0} - \mu\|_{W^{s,1}_{x,v}} \leq \varepsilon^N,$$

*and denoting by $(f_\varepsilon)$ the sequence of solutions to (5.1.2) with initial data $(f_{0,\varepsilon})$, for $\alpha \in [0, 1)$, the following holds:*

$$\liminf_{\varepsilon \to 0} \sup_{t \in [0, \varepsilon^\alpha]} W_1(f_\varepsilon(t), \mu) > 0.$$

We can make the following observations.

- In Proposition 5.1.6, one can take some equilibrium $\mu$ satisfying the same regularity as in Theorem 5.1.3. However, there is no contradiction with our convergence results since in Theorem 5.1.3 we assume that $g_{0,\varepsilon}$ approximates in an analytic way $g_0$ and that $h_{0,\varepsilon}$ converges faster than any polynomial in $\varepsilon$. Therefore, the quantification of the "fast" convergence in Theorem 5.1.3 is important.

- Note that we can have $W_1(h_{0,\varepsilon}, 0) = o_{\varepsilon \to 0}\left(\frac{1}{\varepsilon} \exp\left(\frac{\lambda}{\varepsilon^3} \exp \frac{15}{2\varepsilon^2}\right)\right)$, but

  $$\|h_{0,\varepsilon}\|_{L^p} \sim 1 \qquad \text{for any } p \in [1, \infty],$$

  as fast convergence to 0 in the $W_1$ distance can be achieved for sequences exhibiting fast oscillations.

Theorem 5.1.4 can also be compared to the following result in the *stable* case, that corresponds to initial data consisting of one single Dirac mass (see [55]). In this case, the analogue of Theorem 5.1.4 can be proved with weak assumptions on the initial data.

**Proposition 5.1.7.** *Consider*

$$g_0(x,v) = \rho_0(x)\delta_{v=u_0(x)}.$$

*where* $\rho_0 > 0$ *and* $\rho_0, u_0 \in H^s(\mathbb{T})$, *for* $s \geq 2$. *Consider a sequence* $(f_{0,\varepsilon})$ *of non-negative initial data in* $L^1 \cap L^\infty$ *for* (5.1.2) *such that, for all* $\varepsilon > 0$,

$$\frac{1}{2}\int f_{0,\varepsilon}|v|^2 dv dx + \int \left(e^{U_{0,\varepsilon}} \log e^{U_{0,\varepsilon}} - e^{U_{0,\varepsilon}} + 1\right) dx + \frac{\varepsilon^2}{2}\int |U_{0,\varepsilon}'|^2 dx \leq C$$

*for some* $C > 0$, *and* $U_{0,\varepsilon}$ *is the solution to the Poisson equation*

$$\varepsilon^2 U_{0,\varepsilon}'' = e^{U_{0,\varepsilon}} - \int f_{0,\varepsilon}\, dv.$$

*Also, assume that*

$$\frac{1}{2}\int g_{0,\varepsilon}|v-u_0|^2 dv dx + \int \left(e^{U_{0,\varepsilon}} \log\left(e^{U_{0,\varepsilon}}/\rho_0\right) - e^{U_{0,\varepsilon}} + \rho_0\right) dx + \frac{\varepsilon^2}{2}\int |U_{0,\varepsilon}'|^2 dx \to_{\varepsilon\to 0} 0,$$

*Then there exists* $T > 0$ *such that for any global weak solution* $f_\varepsilon(t)$ *of* (5.1.2) *with initial condition* $f_{0,\varepsilon}$,

$$\sup_{t\in[0,T]} W_1(f_\varepsilon(t), g(t)) \to_{\varepsilon\to 0} 0,$$

*where*

$$g(t,x,v) = \rho(t,x)\delta_{v=u(t,x)},$$

*and* $(\rho, u)$ *satisfy the isothermal Euler system on* $[0,T]$

$$\left\{\begin{array}{c} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t u + u\partial_x u = E, \\ E = -U', \\ U = \log \rho, \\ \rho|_{t=0} = \rho_0, u|_{t=0} = u_0. \end{array}\right. \tag{5.1.6}$$

**Remark 5.1.8.** *We could also have stated another similar result using the estimates around* stable symmetric homogeneous equilibria *of [58], but will not do so for the sake of conciseness.*

In what follows, we study the quasineutral limit by using Wasserstein stability estimates for the Vlasov-Poisson system. Such stability estimates were proved for the classical Vlasov-Poisson system by Loeper [76], in dimension three. In the one-dimensional case, they can be improved, as recently shown by Hauray in the note [61].

The key estimate is a weak-strong stability result for the $(\mathrm{VPME})_\varepsilon$ system, which basically consists in an adaptation of Hauray's proof, and which we believe is of independent interest.

**Theorem 5.1.9.** *Let $T > 0$. Let $f^1_\varepsilon, f^2_\varepsilon$ be two measure solutions of (5.1.2) on $[0,T]$, and assume that $\rho^1_\varepsilon(t,x) := \int f^1_\varepsilon(t,x,v)\,dv$ is bounded in $L^\infty$ on $[0,T]\times\mathbb{T}$. Then, for all $\varepsilon \in (0,1]$, for all $t \in [0,T]$,*

$$W_1(f^1_\varepsilon(t), f^2_\varepsilon(t)) \le \frac{1}{\varepsilon} e^{\frac{1}{\varepsilon}\left[(1+3/\varepsilon^2)t + (8+\frac{1}{\varepsilon^2}e^{15/(2\varepsilon^2)})\int_0^t \|\rho^1_\varepsilon(\tau)\|_\infty \, d\tau\right]} W_1(f^1_\varepsilon(0), f^2_\varepsilon(0)).$$

The proofs of Theorems 5.1.3 and 5.1.4 rely on this stability estimate and on a method introduced by Grenier [49] to justify the quasineutral limit for initial data with uniform analytic regularity.

This Chapter is organized as follows. In Section 5.2, we start by proving Theorem 5.1.9. We then turn to the global weak existence theory in $\mathcal{P}_1(\mathbb{T}\times\mathbb{R})$: in Section 5.3, we prove Theorem 5.1.1, using some estimates exhibited in the previous section. Section 5.4 is then dedicated to the proof of the main Theorems 5.1.3 and 5.1.4. Then, we conclude with the study of auxiliary results: in Section 5.5 we prove Proposition 5.1.6, while in Section 5.6 we prove Proposition 5.1.7.

## 5.2 Weak-strong stability for the VP system with massless electrons: proof of Theorem 5.1.9

In this Section we prove Theorem 5.1.9, i.e., the weak-strong stability estimate for solutions of the $(\mathrm{VPME})_\varepsilon$ system. Notice that our weak-strong stability estimate encloses in particular the case (5.1.1) by taking $\varepsilon = 1$.

Let us introduce the setup of the problem. We follow the same notations as in [61]. In particular, we will use a Lagrangian formulation of the problem.

As a preliminary step, it will be convenient to split the electric field in a singular part behaving as the electric field in Vlasov-Poisson and a regular term. More precisely, let us decompose $E_\varepsilon$ as $\bar{E}_\varepsilon + \widehat{E}_\varepsilon$ where

$$\bar{E}_\varepsilon = -\bar{U}_\varepsilon', \qquad \widehat{E}_\varepsilon = -\widehat{U}_\varepsilon',$$

and $\bar{U}_\varepsilon$ and $\widehat{U}_\varepsilon$ solve respectively

$$\varepsilon^2 \bar{U}_\varepsilon'' = 1 - \rho_\varepsilon, \qquad \varepsilon^2 \widehat{U}_\varepsilon'' = e^{\bar{U}_\varepsilon + \widehat{U}_\varepsilon} - 1.$$

Notice that in this way $U_\varepsilon := \bar{U}_\varepsilon + \widehat{U}_\varepsilon$ solves

$$\varepsilon^2 U_\varepsilon'' = e^{U_\varepsilon} - \rho_\varepsilon.$$

Then we can rewrite (5.1.2) as

$$(VPME)_\varepsilon := \begin{cases} \partial_t f_\varepsilon + v \cdot \partial_x f_\varepsilon + (\bar{E}_\varepsilon + \widehat{E}_\varepsilon) \cdot \partial_v f_\varepsilon = 0, \\ \bar{E}_\varepsilon = -\bar{U}_\varepsilon', \qquad \widehat{E}_\varepsilon = -\widehat{U}_\varepsilon', \\ \varepsilon^2 \bar{U}_\varepsilon'' = 1 - \rho_\varepsilon, \\ \varepsilon^2 \widehat{U}_\varepsilon'' = e^{\bar{U}_\varepsilon + \widehat{U}_\varepsilon} - 1, \\ f_\varepsilon(x, v, 0) \geq 0, \quad \int f_\varepsilon(x, v, 0)\, dx\, dv = 1. \end{cases}$$

To prove Theorem 5.1.9, we shall first show a weak-strong stability estimate for a rescaled system (see $(\mathrm{VPME})_{\varepsilon,2}$ below), and then obtain our result by a further scaling argument.

### 5.2.1 A scaling argument

Let us define

$$\mathcal{F}_\varepsilon(t, x, v) := \frac{1}{\varepsilon} f_\varepsilon\left(\varepsilon t, x, \frac{v}{\varepsilon}\right).$$

Then a direct computation gives

$$(VPME)_{\varepsilon,2} := \begin{cases} \partial_t \mathcal{F}_\varepsilon + v \cdot \partial_x \mathcal{F}_\varepsilon + (\bar{\mathcal{E}}_\varepsilon + \widehat{\mathcal{E}}_\varepsilon) \cdot \partial_v \mathcal{F}_\varepsilon = 0, \\ \bar{\mathcal{E}}_\varepsilon = -\bar{\mathcal{U}}_\varepsilon', \qquad \widehat{\mathcal{E}}_\varepsilon = -\widehat{\mathcal{U}}_\varepsilon', \\ \bar{\mathcal{U}}_\varepsilon'' = 1 - \varrho_\varepsilon, \\ \widehat{\mathcal{U}}_\varepsilon'' = e^{(\bar{\mathcal{U}}_\varepsilon + \widehat{\mathcal{U}}_\varepsilon)/\varepsilon^2} - 1, \\ \mathcal{F}_\varepsilon(x, v, 0) \geq 0, \quad \int \mathcal{F}_\varepsilon(x, v, 0)\, dx\, dv = 1, \end{cases}$$

where

$$\varrho_\varepsilon(t, x) := \int \mathcal{F}_\varepsilon(t, x, v)\, dv.$$

We remark that $\bar{\mathcal{U}}_\varepsilon$ is just the classical Vlasov-Poisson potential so, as in [61],

$$\bar{\mathcal{E}}_\varepsilon(t,x) = -\int_{\mathbb{T}} W'(x-y)\varrho_\varepsilon(t,y)\,dy,$$

where

$$W(x) := \frac{x^2-|x|}{2}$$

(recall that we are identifying $\mathbb{T}$ with $[-1/2,1/2)$ with periodic boundary conditions). In addition, since $W$ is 1-Lipschitz and $|W|\le 1$, recalling that

$$\bar{\mathcal{U}}_\varepsilon(t,x) = \int_{\mathbb{T}} W(x-y)\varrho_\varepsilon(t,y)\,dy \tag{5.2.1}$$

we see that $\bar{\mathcal{U}}_\varepsilon$ is 1-Lipschitz and $|\bar{\mathcal{U}}_\varepsilon|\le 1$.

### 5.2.2 Weak-strong estimate for the rescaled system

The goal of this Section is to prove a quantitative weak-strong convergence for the rescaled system $(\text{VPME})_{\varepsilon,2}$. In order to simplify the notation, we omit the subscript $\varepsilon$. In the sequel we will need the following elementary result:

**Lemma 5.2.1.** *Let $h:[-1/2,1/2]\to\mathbb{R}$ be a continuous function such that $\int_{-1/2}^{1/2} h = 0$. Then*

$$\|h\|_\infty \le \int_{-1/2}^{1/2} |h'|.$$

*Proof.* Since $\int_{-1/2}^{1/2} h = 0$ there exists a point $\bar{x}$ such that $h(\bar{x})=0$. Then, by the Fundamental Theorem of Calculus,

$$|h(x)| = \Big|\int_{\bar{x}}^{x} h'\Big| \le \int_{-1/2}^{1/2}|h'| \qquad \forall\, x\in[-1/2,1/2].$$

□

We can now prove existence of solutions to the equation for $\widehat{\mathcal{U}}$.

**Lemma 5.2.2.** *There exists a unique solution of*

$$\widehat{\mathcal{U}}'' = e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2} - 1 \qquad on\ \mathbb{T} \tag{5.2.2}$$

*and this solution satisfies*

$$\|\widehat{\mathcal{U}}\|_\infty \leq 3, \quad \|\widehat{\mathcal{U}}'\|_\infty \leq 2, \quad \|\widehat{\mathcal{U}}''\|_\infty \leq \frac{3}{\varepsilon^2}.$$

*Proof.* We prove existence of $\widehat{U}$ by finding a minimizer for

$$h \mapsto E[h] := \int_{\mathbb{T}} \left( \frac{1}{2}(h')^2 + \varepsilon^2 e^{(\bar{\mathcal{U}}+h)/\varepsilon^2} - h \right) dx$$

among all periodic functions $h : [-1/2, 1/2] \to \mathbb{R}$. Indeed, as we shall see later, the Poisson equation we intend to solve is nothing but the Euler-Lagrange equation of the above functional.

Notice that since $E[h]$ is a strictly convex functional, solutions of the Euler-Lagrange equation are minimizers and the minimizer is unique. Let us now prove the existence of such a minimizer.

Take $h_k$ a minimizing sequence, that is

$$E[h_k] \to \inf_h E[h] =: \alpha.$$

Notice that by choosing $h = -\bar{\mathcal{U}}$ we get (recall that $|\bar{\mathcal{U}}|, |\bar{\mathcal{U}}'| \leq 1$, see (5.2.1))

$$\alpha \leq E[-\bar{\mathcal{U}}] = \int_{\mathbb{T}} \left( \frac{1}{2}(\bar{\mathcal{U}}')^2 + \bar{\mathcal{U}} \right) dx \leq 2,$$

hence

$$E[h_k] \leq 3 \qquad \text{for } k \text{ large enough.}$$

We first want to prove that $h_k$ is uniformly bounded in $H^1$.

We observe that, since $\bar{\mathcal{U}} \geq -1$, for any $s \in \mathbb{R}$

$$\varepsilon^2 e^{(\bar{\mathcal{U}}(x)+s)/\varepsilon^2} - s \geq \varepsilon^2 e^{(s-1)/\varepsilon^2} - s.$$

Now, for $s \geq 2$ (and $\varepsilon \in (0, 1]$) we have

$$\varepsilon^2 e^{(s-1)/\varepsilon^2} - s \geq e^{s-1} - s \geq s - 2\log 2 \geq s - 3,$$

while for $s \leq 2$ we have

$$\varepsilon^2 e^{(s-1)/\varepsilon^2} - s \geq -s \geq |s| - 4,$$

thus

$$e^{(s-1)/\varepsilon^2} - s \geq |s| - 4 \qquad \forall\, s \in \mathbb{R}.$$

Therefore

$$3 \geq E[h_k] \geq \int_{\mathbb{T}} \frac{1}{2}(h_k')^2 + |h_k| - 4, \tag{5.2.3}$$

which gives

$$\int_{\mathbb{T}} \frac{1}{2}(h_k')^2 \leq 8.$$

In particular, by the Cauchy-Schwarz inequality this implies

$$\begin{aligned} |h_k(x) - h_k(z)| &\leq \left| \int_z^x |h_k'(y)|\, dy \right| \leq \sqrt{|x-z|} \sqrt{\int_{\mathbb{T}} |h_k'(y)|^2\, dy} \\ &\leq 4\sqrt{|x-z|}. \end{aligned} \tag{5.2.4}$$

Up to now we have proved that $h_k'$ are uniformly bounded in $L^2$. We now want to control $h_k$ in $L^\infty$.

Let $M_k$ denote the maximum of $|h_k|$ over $\mathbb{T}$. Then by (5.2.4) we deduce that

$$h_k(x) \geq M_k - 4 \qquad \forall\, x \in \mathbb{T},$$

hence, recalling (5.2.3),

$$3 \geq E[h_k] \geq \int_{\mathbb{T}} (|h_k(x)| - 4)\, dx \geq M_k - 8,$$

which implies $M_k \leq 11$. Thus, we proved that $|h_k| \leq 11$ for all $k$ large enough, which implies in particular that $h_k$ are uniformly bounded in $L^2$.

In conclusion, we have proved that $h_k$ are uniformly bounded in $H^1$ (both $h_k$ and $h_k'$ are uniformly bounded in $L^2$) and in addition they are uniformly bounded and uniformly continuous (as a consequence of (5.2.4)). Hence, up to a subsequence, they converge weakly in $H^1$ (by weak compactness of balls in $H^1$) and uniformly (by the Ascoli-Arzelà theorem) to a function $\widehat{\mathcal{U}}$. We claim that $\widehat{\mathcal{U}}$ is a minimizer. Indeed, by the lower semicontinuity of the $L^2$ norm under weak convergence,

$$\int_{\mathbb{T}} |\widehat{\mathcal{U}}'(x)|^2\, dx \leq \liminf_k \int_{\mathbb{T}} |h_k'(x)|^2\, dx.$$

On the other hand, by uniform convergence,

$$\int_{\mathbb{T}} \left( \varepsilon^2 e^{(\bar{\mathcal{U}}(x) + h_k(x))/\varepsilon^2} - h_k(x) \right) dx \to \int_{\mathbb{T}} \left( \varepsilon^2 e^{(\bar{\mathcal{U}}(x) + \widehat{\mathcal{U}}(x))/\varepsilon^2} - \widehat{\mathcal{U}}(x) \right) dx.$$

In conclusion

$$E[\widehat{\mathcal{U}}] \leq \liminf_k E[h_k] = \alpha,$$

which proves that $\widehat{\mathcal{U}}$ is a minimizer.

By the minimality,

$$0 = \frac{d}{d\eta}\bigg|_{\eta=0} E[\widehat{\mathcal{U}} + \eta h] = \int_{\mathbb{T}} \left(\widehat{\mathcal{U}}'\, h' + e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2} h - h\right)\, dx = \int_{\mathbb{T}} [-\widehat{\mathcal{U}}'' + e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2} - 1] h\, dx,$$

which proves that $\widehat{\mathcal{U}}$ solves (5.2.2) by the arbitrariness of $h$.

We now prove the desired estimates on $\widehat{\mathcal{U}}$. Since $\widehat{\mathcal{U}}'$ is a periodic continuous function we have

$$0 = \int_{\mathbb{T}} \widehat{\mathcal{U}}''\, dx = \int_{\mathbb{T}} \left(e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2} - 1\right)\, dx.$$

Thus we get

$$\int_{\mathbb{T}} |\widehat{\mathcal{U}}''|\, dx \leq \int_{\mathbb{T}} \left| \left(e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2} - 1\right) \right| dx \leq \int_{\mathbb{T}} e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2}\, dx + 1 = 2,$$

and so, by Lemma 5.2.1, we deduce

$$\|\widehat{\mathcal{U}}'\|_\infty \leq \int_{\mathbb{T}} |\widehat{\mathcal{U}}''|\, dx \leq 2.$$

Since $\|\widehat{\mathcal{U}}'\|_\infty \leq 2$, $\|\bar{\mathcal{U}}\|_\infty \leq 1$, and $\int_{\mathbb{T}} e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2}\, dx = 1$ we claim that $\|\widehat{\mathcal{U}}\|_\infty \leq 3$. Indeed, suppose that there exists $\bar{x}$ such that $\widehat{\mathcal{U}}(\bar{x}) \geq M$. Then, recalling that $\|\widehat{\mathcal{U}}'\|_\infty \leq 2$, we have $\widehat{\mathcal{U}}(x) \geq M - 2$ for all $x$. Using that $\|\bar{\mathcal{U}}\|_\infty \leq 1$ we get $\widehat{\mathcal{U}}(x) + \bar{\mathcal{U}}(x) \geq M - 3$. Then,

$$1 = \int_{\mathbb{T}} e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2}\, dx \geq \int_{\mathbb{T}} e^{(M-3)/\varepsilon^2}\, dx = e^{(M-3)/\varepsilon^2} \Rightarrow M \leq 3.$$

On the other hand, if there exists $\bar{x}$ such that $\widehat{\mathcal{U}}(\bar{x}) \leq -M$, then an analogous argument gives

$$1 = \int_{\mathbb{T}} e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2}\, dx \leq \int_{\mathbb{T}} e^{-(M-3)/\varepsilon^2}\, dx = e^{-(M-3)/\varepsilon^2} \Rightarrow M \leq 3.$$

Hence we have that $\|\widehat{\mathcal{U}}\|_\infty \leq 3$. Finally, to estimate $\widehat{\mathcal{U}}''$ we differentiate the equation

$$\widehat{\mathcal{U}}'' = \left(e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2} - 1\right),$$

recall that $\|\bar{\mathcal{U}}'\|_\infty \le 1$, $\|\widehat{\mathcal{U}}'\|_\infty \le 2$, and $\int_{\mathbb{T}} e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2} = 1$, to obtain

$$\int_{\mathbb{T}} |\widehat{\mathcal{U}}'''| = \int_{\mathbb{T}} \left| e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2} \left( \frac{\bar{\mathcal{U}}' + \widehat{\mathcal{U}}'}{\varepsilon^2} \right) \right| \le \frac{\|\bar{\mathcal{U}}'\|_\infty + \|\widehat{\mathcal{U}}'\|_\infty}{\varepsilon^2} \int_{\mathbb{T}} e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2} \le \frac{3}{\varepsilon^2},$$

so, by Lemma 5.2.1 again, we get

$$\|\widehat{\mathcal{U}}''\|_\infty \le \int_{\mathbb{T}} |\widehat{\mathcal{U}}'''| \le \frac{3}{\varepsilon^2},$$

as desired. □

To prove the weak-strong stability result for $(\mathrm{VPME})_{\varepsilon,2}$, following the strategy used in [61] for the classical Vlasol-Poisson system, we will represent solutions in Lagrangian variables instead of using the Eulerian formulation. In this setting, the phase space is $\mathbb{T} \times \mathbb{R}$ and particles in the phase-space are represented by $Z = (X, V)$, where the random variables $X : [0,1] \to \mathbb{T}$ and $V : [0,1] \to \mathbb{R}$ are maps from the probability space $([0,1], ds)$ to the physical space. The idea is that elements in $[0,1]$ do not have any physical meaning but they just label the particles $\{(X(s), V(s))\}_{s\in[0,1]} \subset \mathbb{T} \times \mathbb{R}$.

We mention that this "probabilistic" point of view was already introduced for ODEs by Ambrosio in his study of linear transport equations [1] and generalized later by Figalli to the case of SDEs [39].

To any random variable as above, one associates the mass distribution of particles in the phase space as follows:[2]

$$\mathcal{F}(x,v)\,dx\,dv = (X,V)_\# ds,$$

that is $\mathcal{F}$ is the law of $(X,V)$. So, instead of looking for the evolution of $\mathcal{F}$, we rather let $Z_t := (X_t, V_t)$ evolve accordingly to the following Lagrangian system (recall that, to simplify the notation, we are omitting the subscript $\varepsilon$)

$$(VPME)_{L,2} := \begin{cases} \dot{X}_t = V_t, \\ \dot{V}_t = \bar{\mathcal{E}}(X_t) + \widehat{\mathcal{E}}(X_t), \\ \bar{\mathcal{E}} = -\bar{\mathcal{U}}', \qquad \widehat{\mathcal{E}} = -\widehat{\mathcal{U}}', \\ \bar{\mathcal{U}}'' = 1 - \rho, \\ \widehat{\mathcal{U}}'' = e^{(\bar{\mathcal{U}}+\widehat{\mathcal{U}})/\varepsilon^2} - 1, \\ \rho(t) = (X_t)_\# ds. \end{cases} \tag{5.2.5}$$

[2] Note that the law of $(X,V)$ may not be absolutely continuous, we just wrote the formula to explain the heuristic.

(Such a formulation is rather intuitive if one thinks of the evolution of finitely many particles.) Notice that the fact that $\rho(t)$ is the law of $X_t$ is a consequence of the fact that $\mathcal{F}$ is the law of $Z_t$.

As initial condition we impose that at time zero $Z_t$ is distributed accordingly to $\mathcal{F}_0$, that is

$$(Z_0)_\# ds = \mathcal{F}_0(x,v)\,dx\,dv.$$

We recall the following characterization of the 1-Wasserstein distance, used also by Hauray in [61]:

$$W_1(\mu,\nu) = \min_{X_\# ds=\mu,\, Y_\# ds=\nu} \int_0^1 |X(s)-Y(s)|\,ds.$$

Hence, if $\mathcal{F}_1, \mathcal{F}_2$ are two solutions of $(\mathrm{VPME})_{\varepsilon,2}$, in order to control $W_1(\mathcal{F}_1(t), \mathcal{F}_2(t))$, it is sufficient to do the following: choose $Z_0^1$ and $Z_0^2$ such that

$$(Z_0^i)_\# ds = df^i(0,x,v), \qquad i=1,2$$

and

$$W_1(f^1(0), f^2(0)) = \int_0^1 |Z_0^1(s)-Z_0^2(s)|\,ds,$$

and prove a bound on $\int_0^1 |Z_t^1(s)-Z_t^2(s)|\,ds$ for $t\ge 0$. In this way we automatically get a control on

$$W_1(\mathcal{F}^1(t), \mathcal{F}^2(t)) \le \int_0^1 |Z_t^1(s)-Z_t^2(s)|\,ds.$$

So, our goal is to estimate $\int_0^1 |Z_t^1(s)-Z_t^2(s)|\,ds$. For this, as in [61] we consider

$$\frac{d}{dt}\int_0^1 |Z_t^1(s)-Z_t^2(s)|\,ds.$$

Using (5.2.5), this is bounded by

$$\begin{aligned}
&\int_0^1 |V_t^1(s)-V_t^2(s)|\,ds + \int_0^1 |\mathcal{E}_t^1(X_t^1)-\mathcal{E}_t^2(X_t^2)|\,ds \\
&\le \int_0^1 |Z_t^1(s)-Z_t^2(s)|\,ds + \int_0^1 |\bar{\mathcal{E}}_t^1(X_t^1)-\bar{\mathcal{E}}_t^2(X_t^2)|\,ds + \int_0^1 |\widehat{\mathcal{E}}_t^1(X_t^1)-\widehat{\mathcal{E}}_t^2(X_t^2)|\,ds \\
&\le \int_0^1 |Z_t^1(s)-Z_t^2(s)|\,ds + 8\|\varrho^1(t)\|_\infty \int_0^1 |Z_t^1(s)-Z_t^2(s)|\,ds \\
&+ \int_0^1 |\widehat{\mathcal{E}}_t^1(X_t^1)-\widehat{\mathcal{E}}_t^2(X_t^1)|\,ds + \int_0^1 |\widehat{\mathcal{E}}_t^2(X_t^1)-\widehat{\mathcal{E}}_t^2(X_t^2)|\,ds,
\end{aligned}$$

where we split $\mathcal{E}_t^1$ and $\mathcal{E}_t^2$ as a sum of $\bar{\mathcal{E}}_t^1 + \widehat{\mathcal{E}}_t^1$ and $\bar{\mathcal{E}}_t^2 + \widehat{\mathcal{E}}_t^2$, and we applied the estimate in [61, Proof of Theorem 1.8] to control

$$\int_0^1 |\bar{\mathcal{E}}_t^1(X_t^1) - \bar{\mathcal{E}}_t^2(X_t^2)|\, ds$$

by

$$8\|\varrho^1(t)\|_\infty \int_0^1 |Z_t^1(s) - Z_t^2(s)|\, ds.$$

To estimate the last two terms, we argue as follows: for the second one we recall that $\widehat{\mathcal{E}}_t^2$ is $(3/\varepsilon^2)$-Lipschitz (see Lemma 5.2.2), hence

$$\int_0^1 |\widehat{\mathcal{E}}_t^2(X_t^1) - \widehat{\mathcal{E}}_t^2(X_t^2)|\, ds \leq \frac{3}{\varepsilon^2}\int_0^1 |X_t^1 - X_t^2|\, ds \leq \frac{3}{\varepsilon^2}\int_0^1 |Z_t^1 - Z_t^2|\, ds.$$

For the first term, we first observe the following fact: recalling (5.2.1) and that $W$ is 1-Lipschitz, we have

$$\begin{aligned} |\bar{\mathcal{U}}_t^1 - \bar{\mathcal{U}}_t^2|(x) &= \left| \int_0^1 W(x - X_t^1) - W(x - X_t^2)\, ds \right| \\ &\leq \int_0^1 |X_t^1 - X_t^2|\, ds \leq \int_0^1 |Z_t^1 - Z_t^2|\, ds \end{aligned} \tag{5.2.6}$$

for all $x$. Now we want to estimate $\widehat{\mathcal{E}}_t^1 - \widehat{\mathcal{E}}_t^2$ in $L^2$: for this we start from the equation

$$(\widehat{\mathcal{U}}_t^1)'' - (\widehat{\mathcal{U}}_t^2)'' = e^{(\bar{\mathcal{U}}_t^1 + \widehat{\mathcal{U}}_t^1)/\varepsilon^2} - e^{(\bar{\mathcal{U}}_t^2 + \widehat{\mathcal{U}}_t^2)/\varepsilon^2}.$$

Multiplying by $\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2$ and integrating by parts, we get

$$\begin{aligned} 0 &= \int_{\mathbb{T}} \Big( (\widehat{\mathcal{U}}_t^1)' - (\widehat{\mathcal{U}}_t^2)' \Big)^2 dx + \int_{\mathbb{T}} \Big[ e^{(\bar{\mathcal{U}}_t^1 + \widehat{\mathcal{U}}_t^1)/\varepsilon^2} - e^{(\bar{\mathcal{U}}_t^2 + \widehat{\mathcal{U}}_t^2)/\varepsilon^2} \Big] [\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2]\, dx \\ &= \int_{\mathbb{T}} \Big( (\widehat{\mathcal{U}}_t^1)' - (\widehat{\mathcal{U}}_t^2)' \Big)^2 dx + \int_{\mathbb{T}} \Big[ e^{(\bar{\mathcal{U}}_t^1 + \widehat{\mathcal{U}}_t^1)/\varepsilon^2} - e^{(\bar{\mathcal{U}}_t^1 + \widehat{\mathcal{U}}_t^2)/\varepsilon^2} \Big] [\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2]\, dx \\ &\quad + \int_{\mathbb{T}} \Big[ e^{(\bar{\mathcal{U}}_t^1 + \widehat{\mathcal{U}}_t^2)/\varepsilon^2} - e^{(\bar{\mathcal{U}}_t^2 + \widehat{\mathcal{U}}_t^2)/\varepsilon^2} \Big] [\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2]\, dx. \end{aligned}$$

For the second term we observe that, by the Fundamental Theorem of Calculus,

$$e^{(\bar{\mathcal{U}}_t^1 + \widehat{\mathcal{U}}_t^1)/\varepsilon^2} - e^{(\bar{\mathcal{U}}_t^1 + \widehat{\mathcal{U}}_t^2)/\varepsilon^2} = \frac{1}{\varepsilon^2}\left( \int_0^1 e^{[\bar{\mathcal{U}}_t^1 + \lambda \widehat{\mathcal{U}}_t^1 + (1-\lambda)\widehat{\mathcal{U}}_t^2]/\varepsilon^2}\, d\lambda \right) [\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2].$$

Hence,

$$\begin{aligned}
&\int_{\mathbb{T}} \left[ e^{(\bar{\mathcal{U}}_t^1+\widehat{\mathcal{U}}_t^1)/\varepsilon^2} - e^{(\bar{\mathcal{U}}_t^1+\widehat{\mathcal{U}}_t^2)/\varepsilon^2} \right] [\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2]\, dx \\
&= \int_{\mathbb{T}} \frac{1}{\varepsilon^2} \left( \int_0^1 e^{[\bar{\mathcal{U}}_t^1+\lambda\widehat{\mathcal{U}}_t^1+(1-\lambda)\widehat{\mathcal{U}}_t^2]/\varepsilon^2}\, d\lambda \right) (\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2)^2 dx \\
&\geq \frac{1}{\varepsilon^2} e^{-5/\varepsilon^2} \int_{\mathbb{T}} (\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2)^2 dx
\end{aligned}$$

where we used that $\bar{\mathcal{U}}$ and $\widehat{\mathcal{U}}$ are bounded by 1 and 4, respectively.

For the third term, we simply estimate

$$|e^{(\bar{\mathcal{U}}_t^1+\widehat{\mathcal{U}}_t^2)/\varepsilon^2} - e^{(\bar{\mathcal{U}}_t^2+\widehat{\mathcal{U}}_t^2)/\varepsilon^2}| \leq \frac{1}{\varepsilon^2} e^{5/\varepsilon^2} |\bar{\mathcal{U}}_t^1 - \bar{\mathcal{U}}_t^2|,$$

hence, combining all together,

$$\begin{aligned}
&\int_{\mathbb{T}} \Big( (\widehat{\mathcal{U}}_t^1)' - (\widehat{\mathcal{U}}_t^2)' \Big)^2 dx + e^{-5/\varepsilon^2} \int_{\mathbb{T}} (\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2)^2 dx \\
&\leq \frac{1}{\varepsilon^2} e^{5/\varepsilon^2} \int_{\mathbb{T}} |\bar{\mathcal{U}}_t^1 - \bar{\mathcal{U}}_t^2|\, |\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2|\, dx \\
&\leq \frac{1}{\varepsilon^2} e^{5/\varepsilon^2} \delta \int_{\mathbb{T}} |\widehat{\mathcal{U}}_t^1 - \widehat{\mathcal{U}}_t^2|^2\, dx + \frac{1}{\varepsilon^2} \frac{e^{5/\varepsilon^2}}{\delta} \int_{\mathbb{T}} |\bar{\mathcal{U}}_t^1 - \bar{\mathcal{U}}_t^2|^2\, dx.
\end{aligned}$$

Thus, choosing $\delta := \varepsilon^2 e^{-10/\varepsilon^2}$ , we finally obtain

$$\int_{\mathbb{T}} \Big( (\widehat{\mathcal{U}}_t^1)' - (\widehat{\mathcal{U}}_t^2)' \Big)^2 dx \leq \frac{1}{\varepsilon^4} e^{15/\varepsilon^2} \int_{\mathbb{T}} |\bar{\mathcal{U}}_t^1 - \bar{\mathcal{U}}_t^2|^2\, dx$$

Observing now that $(\widehat{\mathcal{U}}_t^i)' = -\widehat{\mathcal{E}}_t^1$ and recalling (5.2.6), we obtain

$$\sqrt{\int_{\mathbb{T}} \Big( \widehat{\mathcal{E}}_t^1 - \widehat{\mathcal{E}}_t^2 \Big)^2 dx} \leq \frac{1}{\varepsilon^2} e^{15/(2\varepsilon^2)} \int_0^1 |Z_t^1 - Z_t^2|\, ds. \tag{5.2.7}$$

Thanks to this estimate, we conclude that

$$\begin{aligned}\int_0^1 |\widehat{\mathcal{E}}_t^1(X_t^1) - \widehat{\mathcal{E}}_t^2(X_t^1)|\, ds &= \int_{\mathbb{T}} |\widehat{\mathcal{E}}_t^1(x) - \widehat{\mathcal{E}}_t^2(x)|\, \varrho^1(t,x)\, dx \\ &\le \|\varrho^1(t)\|_\infty \int_{\mathbb{T}} |\widehat{\mathcal{E}}_t^1 - \widehat{\mathcal{E}}_t^2|\, dx \\ &\le \|\varrho^1(t)\|_\infty \sqrt{\int_{\mathbb{T}} |\widehat{\mathcal{E}}_t^1 - \widehat{\mathcal{E}}_t^2|^2\, dx} \\ &\le \frac{1}{\varepsilon^2} e^{15/(2\varepsilon^2)} \|\varrho^1(t)\|_\infty \int_0^1 |Z_t^1 - Z_t^2|\, ds.\end{aligned}$$

Hence, combining all together we proved that

$$\frac{d}{dt}\int_0^1 |Z_t^1 - Z_t^2|\, ds \le \Big(1 + 8\|\varrho^1(t)\|_\infty + \frac{3}{\varepsilon^2} + \frac{1}{\varepsilon^2} e^{15/(2\varepsilon^2)} \|\varrho^1(t)\|_\infty\Big) \int_0^1 |Z_t^1 - Z_t^2|\, ds,$$

so that, by Gronwall inequality,

$$\int_0^1 |Z_t^1 - Z_t^2|\, ds \le e^{(1+3/\varepsilon^2)t + (8+\frac{1}{\varepsilon^2} e^{15/(2\varepsilon^2)}) \int_0^t \|\varrho^1(\tau)\|_\infty\, d\tau} \int_0^1 |Z_0^1 - Z_0^2|\, ds,$$

which implies (recalling the discussion at the beginning of this computation)

$$W_1(\mathcal{F}^1(t), \mathcal{F}^2(t)) \le e^{(1+3/\varepsilon^2)t + (8+\frac{1}{\varepsilon^2} e^{15/(2\varepsilon^2)}) \int_0^t \|\varrho^1(\tau)\|_\infty\, d\tau} W_1(\mathcal{F}^1(0), \mathcal{F}^2(0)). \tag{5.2.8}$$

This proves the desired weak-strong stability for the rescaled system.

### 5.2.3 Back to the original system and conclusion of the proof

To obtain the weak-strong stability estimate for our original system, we simply use (5.2.8) together with the definition of $W_1$. More precisely, given two densities $f_1(x,v)$ and $f_2(x,v)$, consider

$$\mathcal{F}_i(x,v) := \frac{1}{\varepsilon} f_i(x, v/\varepsilon), \qquad i = 1, 2.$$

Then

$$\begin{aligned}W_1(\mathcal{F}_1, \mathcal{F}_2) &= \sup_{\|\varphi\|_{\mathrm{Lip}} \le 1} \int \varphi(x,v)[\mathcal{F}_1(x,v) - \mathcal{F}_2(x,v)]\, dx\, dv \\ &= \sup_{\|\varphi\|_{\mathrm{Lip}} \le 1} \int \varphi(x,v) \frac{1}{\varepsilon}[f_1(x, v/\varepsilon) - f_2(x, v/\varepsilon)]\, dx\, dv \\ &= \sup_{\|\varphi\|_{\mathrm{Lip}} \le 1} \int \varphi(x, \varepsilon w)[f_1(x,w) - f_2(x,w)]\, dx\, dw.\end{aligned}$$

We now observe that if $\varphi$ is 1-Lipschitz so is $\varphi(x, \varepsilon w)$ for $\varepsilon \le 1$, hence

$$\sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \int \varphi(x,\varepsilon w)[f_1(x,w) - f_2(x,w)]\,dx\,dw \le \sup_{\|\psi\|_{\mathrm{Lip}}\le 1} \int \psi(x,w)[f_1(x,w) - f_2(x,w)]\,dx\,dw = W_1(f_1,f_2).$$

Reciprocally, given any 1-Lipschitz function $\psi(x,w)$, the function $\varphi(x,w) := \varepsilon\psi(x, w/\varepsilon)$ is still 1-Lipschitz, hence

$$W_1(f_1,f_2) = \sup_{\|\psi\|_{\mathrm{Lip}}\le 1} \int \psi(x,w)[f_1(x,w) - f_2(x,w)]\,dx\,dw \le \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \int \frac{1}{\varepsilon}\varphi(x,\varepsilon w)[f_1(x,w) - f_2(x,w)]\,dx\,dw = \frac{1}{\varepsilon} W_1(\mathcal{F}_1,\mathcal{F}_2).$$

Hence, in conclusion, we have

$$\varepsilon W_1(f_1,f_2) \le W_1(\mathcal{F}_1,\mathcal{F}_2) \le W_1(f_1,f_2).$$

In particular, when applied to solutions of (VPME), we deduce that

$$\varepsilon W_1(f_1(\varepsilon t), f_2(\varepsilon t)) \le W_1(\mathcal{F}_1(t),\mathcal{F}_2(t)) \le W_1(f_1(\varepsilon t), f_2(\varepsilon t)). \tag{5.2.9}$$

Observing that

$$\int_0^t \|\varrho^1(\tau)\|_\infty\,d\tau = \int_0^t \|\rho^1(\varepsilon\tau)\|_\infty\,d\tau = \frac{1}{\varepsilon}\int_0^{\varepsilon t} \|\rho^1(\tau)\|_\infty\,d\tau,$$

(5.2.9) together with (5.2.8) gives

$$W_1(f^1(t), f^2(t)) \le \frac{1}{\varepsilon} e^{\frac{1}{\varepsilon}\left[(1+3/\varepsilon^2)t + (8+\frac{1}{\varepsilon^2}e^{15/(2\varepsilon^2)})\int_0^t \|\rho^1(\tau)\|_\infty\,d\tau\right]} W_1(f^1(0), f^2(0)),$$

which concludes the proof of Theorem 5.1.9.

**Remark 5.2.3.** *Notice that, if we were working with the classical Vlasov-Poisson system, the stability estimate would have simply been*

$$W_1(\mathcal{F}^1(t),\mathcal{F}^2(t)) \le e^{t+8\int_0^t \|\varrho^1(\tau)\|_\infty\,d\tau} W_1(\mathcal{F}^1(0),\mathcal{F}^2(0)),$$

*(compare with [61]), so in terms of f*

$$W_1(f^1(t), f^2(t)) \le \frac{1}{\varepsilon} e^{\frac{1}{\varepsilon}\left[t+8\int_0^t \|\rho^1(\tau)\|_\infty\,d\tau\right]} W_1(f^1(0), f^2(0)),$$

## 5.3 Proof of Theorem 5.1.1

In this Section, we prove the existence of global weak solutions in $\mathcal{P}_1(\mathbb{T} \times \mathbb{R})$ for the (VPME) system. Without loss of generality we prove the existence of solutions when $\varepsilon = 1$ (that is, for (5.1.1)).

To prove existence of weak solutions we follow [61, Proposition 1.2 and Theorem 1.7]. For this, we take a random variable $(X_0, V_0) : [0,1] \to \mathbb{T} \times \mathbb{R}$ whose law is $f_0$, that is $(X_0, V_0)_\# ds = f_0$, and we solve

$$(VPME)_{L,2} := \begin{cases} \dot{X}_t = V_t, \\ \dot{V}_t = \bar{\mathcal{E}}(X_t) + \widehat{\mathcal{E}}(X_t), \\ \bar{\mathcal{E}} = -\bar{\mathcal{U}}', \qquad \widehat{\mathcal{E}} = -\widehat{\mathcal{U}}', \\ \bar{\mathcal{U}}'' = 1 - \rho, \\ \widehat{\mathcal{U}}'' = e^{(\bar{\mathcal{U}} + \widehat{\mathcal{U}})} - 1, \\ \rho(t) = (X_t)_\# ds. \end{cases} \tag{5.3.1}$$

Indeed, once we have $(X_t, V_t)$, $f_t := (X_t, V_t)_\# ds$ will be a solution to (5.1.1). We split the argument in several steps.

*Step 1: Solution of the $N$ particle system (compare with [61, Proof of Proposition 1.2]).* We start from the $N$ particle systems of ODEs, $i = 1, \ldots, N$,

$$\begin{cases} \dot{X}_t^i = V_t^i, \\ \dot{V}_t^i = \bar{E}(X_t^i) + \widehat{E}(X_t^i), \\ \bar{E} = -\bar{U}', \qquad \widehat{E} = -\widehat{U}', \\ \bar{U}'' = \frac{1}{N}\sum_{i=1}^N \delta_{X^i} - 1, \\ \widehat{U}'' = e^{\bar{U} + \widehat{U}} - 1, \end{cases}$$

Because the electric field $\bar{E}(X^i) = -\frac{1}{N}\sum_{j \neq i} W'(X^i - X^j)$ is singular when $X^i = X^j$ for some $i \neq j$, to prove existence we want to rewrite the above system as a differential inclusion

$$\dot{\mathcal{Z}}^N(t) \in \mathcal{B}^N(\mathcal{Z}^N(t)),$$

where $\mathcal{B}^N$ is a multivalued map from $\mathbb{R}^{2N}$ into the set of parts of $\mathbb{R}^{2N}$. For this, we write

$$\dot{\mathcal{Z}}^N(t) \in \mathcal{B}^N(\mathcal{Z}^N(t)) \quad \Leftrightarrow \quad \dot{X}^i = V^i, \quad \dot{V}^i \in \frac{1}{N}\sum_{j=1}^N F_{ij}$$

where

$$F_{ij} = -F_{ji} = -W'(X^i - X^j) + \widehat{E}(X^i) \quad \text{when } X^i \neq X^j,$$

$$F_{ij} = -F_{ji} \in [\widehat{E}(X^i) - 1/2, \widehat{E}(X^i) + 1/2] \quad \text{when } X^i = X^j.$$

As in [61], this equation is solved by Filippov's Theorem (see [40]) which provides existence of a solution, and as shown in [61, Step 2 of proof of Proposition 1.2] the solution of the differential inclusion is a solution to the $N$ particle problem.

*Step 2: Approximation argument.* To solve $(VPME)_{L,2}$ we approximate $f_0$ with a family of empirical measures

$$f_0^N := \frac{1}{N} \sum_{i=1}^{N} \delta_{(x^i,v^i)},$$

that we can assume to satisfy (thanks to (5.1.4))

$$\int |v| \, df_0^N(x,v) \leq C \qquad \forall\, N, \tag{5.3.2}$$

and we apply Step 1 to solve the ODE system and find solutions $(X_t^N, V_t^N) \in \mathbb{T} \times \mathbb{R}$ of $(VPME)_{L,2}$ starting from an initial condition $(X_0^N, V_0^N)$ whose law is $f_0^N$.

Next, we notice that in [61, Step 2, Proof of Theorem 1.7] the only property on the vector field used in the proof is the fact that $F_{ij}$ are bounded by $1/2$, and it is used to show that

$$\sup_{u,s\in[0,t]} \frac{|Z^N(u) - Z^N(u)|}{|s-u|} \leq |V_0^N| + \frac{1}{2}(1+t),$$

which, combined with (5.3.2), is enough to ensure tightness (see [61, Step 2, Proof of Theorem 1.7] for more details). Since in our case the $F_{ij}$ are also bounded (as we are simply adding a bounded term $\widehat{E}$), we deduce that for some $C > 0$,

$$\sup_{u,s\in[0,t]} \frac{|Z^N(u) - Z^N(u)|}{|s-u|} \leq |V_0^N| + C(1+t),$$

so the sequence $Z^N := (X^N, V^N)$ is still tight and (up to a subsequence) converge to a process $Z = (X, V)$: it holds

$$\int_0^1 \sup_{t\in[0,T]} \left| Z_t^N(s) - Z_t(s) \right| ds \to 0 \qquad \text{as } N \to \infty \tag{5.3.3}$$

for any $T > 0$.

*Step 3: Characterization of the limit process.* We now want to prove that the limit process $Z_t = (X_t, V_t)$ is a solution of $(VPME)_{L,2}$.

Let us denote by $\bar{E}^N$ and $\widehat{E}^N$ the electric fields associated to the solution $(X^N, V^N)$. Recall that $(X^N, V^N)$ solve

$$\dot{X}^N = V^N, \qquad \dot{V}^N = \bar{E}^N(X^N) + \widehat{E}^N(X^N),$$

or equivalently

$$X_t^N = \int_0^t V_\tau^N \, d\tau, \qquad V_t^N = \int_0^t \bar{E}_\tau^N(X_\tau^N) + \widehat{E}_\tau^N(X_\tau^N) \, d\tau.$$

In [61, Step 3, Proof of Theorem 1.7], using (5.3.3), it is proved that

$$\int_0^t \bar{E}_\tau^N(X_\tau^N) \to \int_0^t \bar{E}_\tau(X_\tau) \quad \text{in } L^1([0,1], \, ds),$$

so, to ensure that $(X, V)$ solves

$$X_t = \int_0^t V_\tau \, d\tau, \qquad V_t = \int_0^t \bar{E}_\tau(X_\tau) + \widehat{E}_\tau(X_\tau) \, d\tau,$$

it suffices to show that, for any $\tau \geq 0$,

$$\int_0^1 |\widehat{E}_\tau^N(X_\tau^N(s)) - \widehat{E}_\tau(X_\tau(s))| \, ds \to 0 \qquad \text{as } N \to \infty.$$

To show this we see that

$$\begin{aligned}\int_0^1 |\widehat{E}_\tau^N(X_\tau^N(s)) - \widehat{E}_\tau(X_\tau(s))| \, ds &\leq \int_0^1 |\widehat{E}_\tau^N(X_\tau^N(s)) - \widehat{E}_\tau(X_\tau^N(s))| \, ds \\ &\quad + \int_0^1 |\widehat{E}_\tau(X_\tau^N(s)) - \widehat{E}_\tau(X_\tau(s))| \, ds \\ &=: I_1 + I_2\end{aligned}$$

For $I_2$ we use that $\widehat{E}_\tau$ is $M$-Lipschitz (recall Lemma 5.2.2) to estimate

$$I_2 \leq M \int_0^1 |X_\tau^N(s) - X_\tau(s)| \, ds$$

that goes to 0 thanks to (5.3.3).

For $I_1$, we notice that the Cauchy-Schwarz inequality, (5.2.7), and (5.3.3) imply that, as $N \to \infty$,

$$\int_{\mathbb{T}} |\widehat{E}_\tau^N(x) - \widehat{E}_\tau(x)|\, dx \leq \sqrt{\int_{\mathbb{T}} |\widehat{E}_\tau^N(x) - \widehat{E}_\tau(x)|^2\, dx} \\ \leq \bar{C} \int_0^1 |Z_\tau^N(s) - Z_\tau(s)|\, ds \to 0.$$

Hence, we know that $\widehat{E}_\tau^N$ converge to $\widehat{E}_\tau$ in $L^1(\mathbb{T})$. We now recall that $\{\widehat{E}_\tau^N\}_{N\geq 1}$ are $M$-Lipschitz, which implies by Ascoli-Arzelà that, up to subsequences, they converge uniformly to some limit, but by uniqueness of the limit they have to converge uniformly to $\widehat{E}_\tau$. Thanks to this fact we finally obtain

$$I_1 \leq \sup_{x\in\mathbb{T}} |\widehat{E}_\tau^N(x) - \widehat{E}_\tau(x)| \to 0,$$

which concludes the proof.

# 5.4 Proofs of Theorems 5.1.3 and 5.1.4

Our aim is now to prove Theorems 5.1.3 and 5.1.4. The principle is first to adapt some results from [49] for the (VPME) system in terms of the $W_1$ distance, which allows us to settle the case where $h_{\varepsilon,0} = 0$. In a second time, we apply the stability estimate of Theorem 5.1.9.

## 5.4.1 The fluid point of view and convergence for uniformly analytic initial data

We describe in this Section the approach introduced by Grenier in [49] for the study of the quasineutral limit for the classical Vlasov-Poisson system. As we shall see, this can be adapted without difficulty to (VPME)$_\varepsilon$.

We assume that, for all $\varepsilon \in (0,1)$, $g_{0,\varepsilon}(x,v)$ is a *continuous* function; following Grenier [49], we write each initial condition as a “superposition of Dirac masses in velocity”:

$$g_{0,\varepsilon}(x,v) = \int_{\mathcal{M}} \rho^\theta_{0,\varepsilon}(x)\delta_{v=v^\theta_{0,\varepsilon}(x)}\, d\mu(\theta)$$

with $\mathcal{M} := \mathbb{R}$, $d\mu(\theta) = \frac{1}{\pi}\frac{d\theta}{1+\theta^2}$,

$$\rho^\theta_{0,\varepsilon} = \pi(1+\theta^2) g_{0,\varepsilon}(x,\theta), \quad v^\theta_{0,\varepsilon} = \theta.$$

This leads to the study of the behavior as $\varepsilon \to 0$ for solutions to the multi-fluid pressureless Euler-Poisson system

$$\begin{cases} \partial_t \rho^\theta_\varepsilon + \partial_x(\rho^\theta_\varepsilon v^\theta_\varepsilon) = 0, \\ \partial_t v^\theta_\varepsilon + v^\theta_\varepsilon \partial_x v^\theta_\varepsilon = E_\varepsilon, \\ E_\varepsilon = -U'_\varepsilon, \\ \varepsilon^2 U''_\varepsilon = e^{U_\varepsilon} - \int_{\mathcal{M}} \rho^\theta_\varepsilon \, d\mu(\theta), \\ \rho^\theta_\varepsilon|_{t=0} = \rho^\theta_{0,\varepsilon}, v^\theta_\varepsilon|_{t=0} = v^\theta_{0,\varepsilon}. \end{cases} \tag{5.4.1}$$

One then checks that defining

$$g_\varepsilon(t,x,v) = \int_{\mathcal{M}} \rho^\theta_\varepsilon(t,x) \delta_{v = v^\theta_\varepsilon(t,x)} \, d\mu(\theta)$$

gives a weak solution to (5.1.2) (as an application of Theorem 5.1.9 and of the subsequent estimates, this is actually *the unique* weak solution to (5.1.2) with initial datum $g_{0,\varepsilon}$).

The formal limit system, which is associated to the kinetic isothermal Euler system, is the following multi fluid isothermal Euler system:

$$\begin{cases} \partial_t \rho^\theta + \partial_x(\rho^\theta v^\theta) = 0, \\ \partial_t v^\theta + v^\theta \partial_x v^\theta = E, \\ E = -U', \\ U = \log\left(\int_{\mathcal{M}} \rho^\theta \, d\mu(\theta)\right), \\ \rho^\theta|_{t=0} = \rho^\theta_0, v^\theta|_{t=0} = v^\theta_0, \end{cases} \tag{5.4.2}$$

where the $\rho^\theta_0$ are defined as the limits of $\rho^\theta_{0,\varepsilon}$ (which are thus supposed to exist) and $v^\theta_0 = \theta$.

As before, one checks that defining

$$g(t,x,v) = \int_{\mathcal{M}} \rho^\theta(t,x) \delta_{v = v^\theta(t,x)} \, d\mu(\theta)$$

gives a weak solution to the kinetic Euler isothermal system.

Recalling the analytic norms used by Grenier in [49] (see Definition 5.1.2), we can adapt the results of [49, Theorems 1.1.2, 1.1.3 and Remark 1 p. 369] to get the following proposition.

**Proposition 5.4.1.** *Assume that there exist $\delta_0, C, \eta > 0$, with $\eta$ small enough, such that*

$$\sup_{\varepsilon \in (0,1)} \sup_{v \in \mathbb{R}} (1+v^2) \|g_{0,\varepsilon}(\cdot, v)\|_{B_{\delta_0}} \leq C,$$

*and that*

$$\sup_{\varepsilon \in (0,1)} \left\| \int_{\mathbb{R}} g_{0,\varepsilon}(\cdot, v)\, dv - 1 \right\|_{B_{\delta_0}} < \eta.$$

*Denote for all $\theta \in \mathbb{R}$,*

$$\rho^\theta_{0,\varepsilon} = \pi(1+\theta^2) g_{0,\varepsilon}(x, \theta), \quad v^\theta_{0,\varepsilon} = v^\theta = \theta.$$

*Assume that for all $\theta \in \mathbb{R}$, $\rho^\theta_{0,\varepsilon}$ has a limit in the sense of distributions and denote*

$$\rho^\theta_0 = \lim_{\varepsilon \to 0} \rho^\theta_{0,\varepsilon}.$$

*Then there exist $\delta_1 > 0$ and $T > 0$ such that:*

- *for all $\varepsilon \in (0,1)$, there is a unique solution $(\rho^\theta_\varepsilon, v^\theta_\varepsilon)_{\theta \in M}$ of (5.4.1) with initial data $(\rho^\theta_{0,\varepsilon}, v^\theta_{0,\varepsilon})_{\theta \in M}$, such that $\rho^\theta_\varepsilon, v^\theta_\varepsilon \in C([0,T]; B_{\delta_1})$ for all $\theta \in M$ and $\varepsilon \in (0,1)$, with bounds that are uniform in $\varepsilon$;*
- *there is a unique solution $(\rho^\theta, v^\theta)_{\theta \in M}$ of (5.4.2) with initial data $(\rho^\theta_0, v^\theta_0)_{\theta \in M}$, such that $\rho^\theta, v^\theta \in C([0,T]; B_{\delta_1})$ for all $\theta \in M$;*
- *for all $s \in \mathbb{N}$, we have*

$$\sup_{\theta \in M} \sup_{t \in [0,T]} \left[ \|\rho^\theta_\varepsilon - \rho^\theta\|_{H^s(\mathbb{T})} + \|v^\theta_\varepsilon - v^\theta\|_{H^s(\mathbb{T})} \right] \to_{\varepsilon \to 0} 0. \tag{5.4.3}$$

Remark that analyticity is actually needed only in the position variable, and not in the velocity variable. This allows us, for instance, to consider initial data which are compactly supported in velocity.

We shall not give a complete proof of this result (which is of Cauchy-Kovalevski type), since it is very close to the one given by Grenier in [49] for the classical Vlasov-Poisson system, but we just emphasize the main differences.

First of all we begin by noticing that one difficulty in the classical case comes from the fact the one can not directly use the Poisson equation

$$-\varepsilon^2 U''_\varepsilon = \rho_\varepsilon - 1$$

if one wants some useful uniform analytic estimates for the electric field. Because of this issue, a combination of the Vlasov and Poisson equation is used in [49], which allows one to get a kind of wave equation solved by $U_\varepsilon$. This shows in particular that the electric field has a highly oscillatory behavior in time (the fast oscillations in time correspond to the so-called plasma waves) which have to be filtered in order to obtain convergence. For this reason, Grenier needs to introduce some correctors in order to get convergence of the velocity fields (these oscillations and correctors vanish only if the initial conditions are well-prepared, *i.e.* verify some compatibility conditions).

For the (VPME) system, that is when one adds the exponential term in the Poisson equation, such a problem does not occur. To explain this, consider first the linearized Poisson equation

$$-\varepsilon^2 U_\varepsilon'' + U_\varepsilon = \int_{\mathcal{M}} \rho_\varepsilon^\theta \, d\mu(\theta) - 1$$

and observe that this equation is appropriate to get uniform analytic estimates. Indeed, writing

$$U_\varepsilon = (Id - \varepsilon^2 \partial_{xx})^{-1} \left( \int_{\mathcal{M}} \rho_\varepsilon^\theta \, d\mu(\theta) - 1 \right),$$

this shows that if $\rho_\varepsilon$ is analytic then also $U_\varepsilon$ (and so $E_\varepsilon$) is analytic, which implies that there are no fast oscillations in time, contrary to the classical case. In particular, our convergence result holds without the need of adding any correctors.

A second difference concerns the existence of analytic solutions on an interval of time $[0, T]$ independent of $\varepsilon$: the construction of Grenier of analytic solutions is based on a Cauchy-Kovalevski type proof based on an iteration procedure in a scale of Banach spaces (see [49, Section 2.1]). Most of the estimates used to prove that such iteration converge use the Fourier transform, that is unavailable in our case since the Poisson equation

$$-\varepsilon^2 U_\varepsilon'' + e^{U_\varepsilon} = \int_{\mathcal{M}} \rho_\varepsilon^\theta \, d\mu(\theta)$$

is nonlinear. However, since we deal with analytic functions, we can express everything in power series to use the Fourier transform and obtain some a priori estimates in the analytic norm. Furthermore, one can write the Poisson equation as

$$-\varepsilon^2 U_\varepsilon'' + U_\varepsilon = \int_{\mathcal{M}} \rho_\varepsilon^\theta \, d\mu(\theta) - 1 - (e^{U_\varepsilon} - U_\varepsilon - 1)$$

and rely on the fact that the "error term" $(e^{U_\varepsilon} - U_\varepsilon - 1)$ is quadratic in $U_\varepsilon$ (which is expected to be small in the regime where $\sup_{\varepsilon\in(0,1)} \left\| \int_{\mathcal{M}} \rho_\varepsilon^\theta \, d\mu(\theta) - 1 \right\|_{B_{\delta_0}} \ll 1$), and thus can be handled in the approximation scheme used in [49].

We deduce the next corollary.

**Corollary 5.4.2.** *With the same assumptions and notation as in Proposition 5.4.1, we have*

$$\sup_{t\in[0,T]} W_1(g_\varepsilon(t), g(t)) \to_{\varepsilon\to 0} 0, \tag{5.4.4}$$

*where*

$$g_\varepsilon(t,x,v) = \int_{\mathcal{M}} \rho_\varepsilon^\theta(t,x)\delta_{v=v_\varepsilon^\theta(t,x)} \, d\mu(\theta), \qquad g(t,x,v) = \int_{\mathcal{M}} \rho^\theta(t,x)\delta_{v=v^\theta(t,x)} \, d\mu(\theta). \tag{5.4.5}$$

*Proof.* The convergence (5.4.4) follows from (5.4.3), and the Sobolev embedding theorem. We have indeed for all $t \in [0,T]$:

$$\begin{aligned}
&W_1(g_\varepsilon(t), g(t)) = \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \langle g_\varepsilon - g, \, \varphi\rangle \\
&= \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \left\{ \int_{\mathbb{T}} \int_{\mathcal{M}} (\rho_\varepsilon^\theta(t,x)\varphi(x, v_\varepsilon^\theta(t,x)) - \rho^\theta(t,x)\varphi(x, v^\theta(t,x))) \, d\mu(\theta) \, dx \right\} \\
&= \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \left\{ \int_{\mathbb{T}} \int_{\mathcal{M}} \rho_\varepsilon^\theta(t,x)(\varphi(x, v_\varepsilon^\theta(t,x)) - \varphi(x, v^\theta(t,x)) \, d\mu(\theta) \, dx \right\} \\
&+ \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \left\{ \int_{\mathbb{T}} \int_{\mathcal{M}} (\rho_\varepsilon^\theta(t,x) - \rho^\theta(t,x))\varphi(x, v^\theta(t,x)) \, d\mu(\theta) \, dx \right\}.
\end{aligned}$$

Thus, we deduce the estimate

$$\begin{aligned}
W_1(g_\varepsilon(t), g(t)) \le &\sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \sup_{\varepsilon\in(0,1),\, \theta\in M} \|\rho_\varepsilon^\theta\|_\infty \|\varphi\|_{\mathrm{Lip}} \int_{\mathcal{M}} \|v_\varepsilon^\theta(t,x) - v^\theta(t,x)\|_\infty \, d\mu(\theta) \\
&+ \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \int_{\mathcal{M}} \|\rho_\varepsilon^\theta - \rho_\theta\|_\infty \, d\mu(\theta) \|\varphi\|_{\mathrm{Lip}} \left( 1/2 + \sup_{\theta\in M} \|v^\theta(t,x)\|_\infty \right) \\
&+ \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \left\{ \int_{\mathbb{T}\times\mathbb{R}} \int_{\mathcal{M}} (\rho_\varepsilon^\theta(t,x) - \rho^\theta(t,x))\varphi(0,0) \, d\mu(\theta) \, dx \right\}.
\end{aligned}$$

We notice that the last term is equal to 0 since for all $t \ge 0$,

$$\int_{\mathbb{T}} \int_{\mathcal{M}} \rho_\varepsilon^\theta(t,x) \, d\mu(\theta) \, dx = \int_{\mathbb{T}} \int_{\mathcal{M}} \rho_\varepsilon^\theta(t,x) \, d\mu(\theta) \, dx = 1,$$

by conservation of the total mass. After taking the supremum in time, we also see that the other two terms converge to 0, using the $L^\infty$ convergence of $(\rho_\varepsilon^\theta, v_\varepsilon^\theta)$ to $(\rho_0^\theta, v_0^\theta)$. This concludes the proof. □

This approach is also relevant for singular initial data such as the sum of Dirac masses in velocity:

$$g_{0,\varepsilon}(x,v) = \sum_{i=1}^N \rho_{0,\varepsilon}^i(x)\delta_{v=v_{0,\varepsilon}^i(x)}.$$

and we have a similar theorem assuming that $(\rho_{0,\varepsilon}^i, v_{0,\varepsilon}^i)$ is uniformly analytic.

In this case $\mathcal{M} = \{1, \cdots, N\}$ and $d\mu$ is the counting measure. This leads to the study of the behavior as $\varepsilon \to 0$ of the system (for $i \in \{1, \cdots, N\}$)

$$\begin{cases} \partial_t \rho_\varepsilon^i + \partial_x(\rho_\varepsilon^i v_\varepsilon^i) = 0, \\ \partial_t v_\varepsilon^i + v_\varepsilon^i \partial_x v_\varepsilon^i = E_\varepsilon, \\ E_\varepsilon = -U_\varepsilon', \\ \varepsilon^2 U_\varepsilon'' = e^{U_\varepsilon} - \left(\sum_{i=1}^N \rho_\varepsilon^i\right), \\ \rho_\varepsilon^i|_{t=0} = \rho_{0,\varepsilon}^i, v_\varepsilon^i|_{t=0} = v_{0,\varepsilon}^i. \end{cases} \tag{5.4.6}$$

and the formal limit is the following multi fluid isothermal system

$$\begin{cases} \partial_t \rho^i + \partial_x(\rho^i v^i) = 0, \\ \partial_t v^i + v^i \partial_x v^i = E, \\ E = -U', \\ U = \log\left(\sum_{i=1}^N \rho^i\right), \\ \rho^i|_{t=0} = \rho_0^i, v^i|_{t=0} = v_0^i. \end{cases} \tag{5.4.7}$$

As before, adapting the arguments in [49], we obtain the following proposition and its corollary.

**Proposition 5.4.3.** *Assume that there exist $\delta_0, C, \eta > 0$, with $\eta$ small enough, such that*

$$\sup_{\varepsilon\in(0,1)} \sup_{i\in\{1,\cdots,N\}} \|\rho_{0,\varepsilon}^i\|_{B_{\delta_0}} + \|v_{0,\varepsilon}^i\|_{B_{\delta_0}} \leq C,$$

*and that*

$$\sup_{\varepsilon\in(0,1)} \left\| \sum_{i=1}^N \rho_{0,\varepsilon}^i - 1 \right\|_{B_{\delta_0}} < \eta.$$

*Assume that for all* $i = 1, \cdots, N$, $\rho^i_{0,\varepsilon}, v^i_{0,\varepsilon}$ *admit a limit in the sense of distributions and denote*

$$\rho^i_0 = \lim_{\varepsilon\to 0} \rho^i_{0,\varepsilon}, \quad v^i_0 = \lim_{\varepsilon\to 0} v^i_{0,\varepsilon}.$$

*Then there exist* $\delta_1 > 0$ *and* $T > 0$ *such that:*

- *for all* $\varepsilon \in (0,1)$, *there is a unique solution* $(\rho^i_\varepsilon, v^i_\varepsilon)_{i\in\{1,\cdots,N\}}$ *of* (5.4.6) *with initial data* $(\rho^i_{0,\varepsilon}, v^i_{0,\varepsilon})_{i\in\{1,\cdots,N\}}$, *such that* $\rho^i_\varepsilon, v^i_\varepsilon \in C([0,T]; B_{\delta_1})$ *for all* $i \in \{1, \cdots, N\}$ *and* $\varepsilon \in (0,1)$, *with bounds that are uniform in* $\varepsilon$;

- *there is a unique solution* $(\rho^i, v^i)_{i\in\{1,\cdots,N\}}$ *of* (5.4.7) *with initial data* $(\rho^i_0, v^i_0)_{i\in\{1,\cdots,N\}}$, *such that* $\rho^i, v^i \in C([0,T]; B_{\delta_1})$ *for all* $i \in \{1, \cdots, N\}$;

- *for all* $s \in \mathbb{N}$, *we have*

$$\sup_{i\in\{1,\cdots,N\}} \sup_{t\in[0,T]} \left[\|\rho^i_\varepsilon - \rho^i\|_{H^s(\mathbb{T})} + \|v^i_\varepsilon - v^i\|_{H^s(\mathbb{T})}\right] \to_{\varepsilon\to 0} 0.$$

**Corollary 5.4.4.** *With the same assumptions and notation as in the Proposition 5.4.3, for all* $t \in [0,T]$ *we have*

$$W_1(g_\varepsilon(t), g(t)) \to_{\varepsilon\to 0} 0, \tag{5.4.8}$$

*where*

$$g_\varepsilon(t,x,v) = \sum_{i\in\{1,\cdots,N\}} \rho^i_\varepsilon(t,x)\delta_{v=v^i_\varepsilon(t,x)}, \quad g(t,x,v) = \sum_{i\in\{1,\cdots,N\}} \rho^i(t,x)\delta_{v=v^i(t,x)}. \tag{5.4.9}$$

### 5.4.2 End of the proof of Theorem 5.1.3 and Theorem 5.1.4

We are now in position to conclude. Let $(f_\varepsilon)$ a sequence of global weak solutions to (5.1.2) with initial conditions $(f_{0,\varepsilon})$ (obtained thanks to Theorem 5.1.1).

We denote by $(g_\varepsilon)$ the sequence of weak solutions to (5.1.2) with initial conditions $(g_{0,\varepsilon})$, defined by (5.4.5) for the case of Theorem 5.1.3 and (5.4.9) for the case of Theorem 5.1.4. Using the triangle inequality, we have

$$W_1(f_\varepsilon(t), g(t)) \le W_1(f_\varepsilon(t), g_\varepsilon(t)) + W_1(g_\varepsilon(t), g(t)),$$

where $g$ is defined by (5.4.5) for the case of Theorem 5.1.3 and (5.4.9) for the case of Theorem 5.1.4.

For the first term, we use Theorem 5.1.9 to get

$$\begin{aligned} W_1(f_\varepsilon(t), g_\varepsilon(t)) &\le W_1(g_{0,\varepsilon} + h_{0,\varepsilon}, g_{0,\varepsilon}) \frac{1}{\varepsilon} e^{\frac{1}{\varepsilon}\left[(1+3/\varepsilon^2)t + (8+\frac{1}{\varepsilon^2}e^{15/(2\varepsilon^2)})\int_0^t \|\rho_\varepsilon(\tau)\|_\infty \, d\tau\right]} \\ &= W_1(h_{0,\varepsilon}, 0) \frac{1}{\varepsilon} e^{\frac{1}{\varepsilon}\left[(1+3/\varepsilon^2)t + (8+\frac{1}{\varepsilon^2}e^{15/(2\varepsilon^2)})\int_0^t \|\rho_\varepsilon(\tau)\|_\infty \, d\tau\right]}, \end{aligned}$$

where $\rho_\varepsilon$ is here the local density associated to $g_\varepsilon$. By Proposition 5.4.1 (for the case of Theorem 5.1.3) and Proposition 5.4.3 (for the case of Theorem 5.1.4), there exists $C_0 > 0$ such that for all $\varepsilon \in (0, 1)$,

$$\sup_{\tau \in [0,T]} \|\rho_\varepsilon(\tau)\|_\infty \le C_0.$$

Consequently, we observe that taking

$$\varphi(\varepsilon) = \frac{1}{\varepsilon} \exp\left(\frac{\lambda}{\varepsilon^3} \exp \frac{15}{2\varepsilon^2}\right),$$

with $\lambda < 0$, we have, by assumption on $h_{0,\varepsilon}$ (take a smaller $T$ if necessary) that

$$\sup_{t \in [0,T]} W_1(f_\varepsilon(t), g_\varepsilon(t)) \to_{\varepsilon \to 0} 0.$$

We also get that $W_1(g_\varepsilon(t), g(t))$ converges to 0, applying Corollary 5.4.2 for the case of Theorem 5.1.3, and Corollary 5.4.4 for the case of Theorem 5.1.4.

This concludes the proofs of Theorems 5.1.3 and 5.1.4.

## 5.5 Proof of Proposition 5.1.6

We now discuss a non-derivation result which was first stated by Grenier in the note [50], and then studied in more details by the first author and Hauray in [58] (the latter was stated for a general class of homogeneous data, i.e., independent of the position). In [58] such results are given either for the classical Vlasov-Poisson system or for the linearized (VPME) system, but the proofs can be adapted to (VPME).

We first recall two definitions from [58].

**Definition 5.5.1.** *We say that a homogeneous profile $\mu(v)$ with $\int \mu \, dv = 1$ satisfies the Penrose instability criterion if there exists a local minimum point $\bar{v}$ of $\mu$ such that the following inequality holds:*

$$\int_{\mathbb{R}} \frac{\mu(v) - \mu(\bar{v})}{(v - \bar{v})^2} \, dv > 1. \tag{5.5.1}$$

*If the local minimum is flat, i.e., is reached on an interval* $[\bar{v}_1, \bar{v}_2]$, *then* (5.5.1) *has to be satisfied for all* $\bar{v} \in [\bar{v}_1, \bar{v}_2]$.

**Definition 5.5.2.** *We say that a positive and* $C^1$ *profile* $\mu(v)$ *satisfies the* $\delta$*-condition*[3] *if*

$$\sup_{v\in\mathbb{R}} \frac{|\mu'(v)|}{(1+|v|)\mu(v)} < +\infty. \tag{5.5.2}$$

We can now state the theorem taken from [58].

**Theorem 5.5.3.** *Let* $\mu(v)$ *be a smooth profile satisfying the Penrose instability criterion. Assume that* $\mu$ *is positive and satisfies the* $\delta$*-condition*[4]. *For any* $N > 0$ *and* $s > 0$, *there exists a sequence of non-negative initial data* $(f_{0,\varepsilon})$ *such that*

$$\|f_{\varepsilon,0} - \mu\|_{W^{s,1}_{x,v}} \leq \varepsilon^N,$$

*and denoting by* $(f_\varepsilon)$ *the sequence of solutions to* (5.1.2) *with initial data* $(f_{0,\varepsilon})$, *the following holds:*

1. $L^1$ **instability for the macroscopic observables:** *consider the density* $\rho_\varepsilon := \int f_\varepsilon \, dv$ *and the electric field* $E_\varepsilon = -\partial_x U_\varepsilon$. *For all* $\alpha \in [0, 1)$, *we have*

$$\liminf_{\varepsilon\to 0} \sup_{t\in[0,\varepsilon^\alpha]} \|\rho_\varepsilon(t) - 1\|_{L^1_x} > 0, \qquad \liminf_{\varepsilon\to 0} \sup_{t\in[0,\varepsilon^\alpha]} \varepsilon \|E_\varepsilon\|_{L^1_x} > 0. \tag{5.5.3}$$

2. **Full instability for the distribution function:** *for any* $r \in \mathbb{Z}$, *we have*

$$\liminf_{\varepsilon\to 0} \sup_{t\in[0,\varepsilon^\alpha]} \|f_\varepsilon(t) - \mu\|_{W^{r,1}_{x,v}} > 0. \tag{5.5.4}$$

We deduce a proof of Proposition 5.1.6 from this result. Indeed, take a smooth $\mu$ satisfying the assumptions of Theorem 5.5.3, and consider the sequence of initial conditions $(f_{0,\varepsilon})$ given by this theorem.

By the Sobolev imbedding theorem in dimension 1, the space $W^{2,1}(\mathbb{T}\times\mathbb{R})$ is continuously imbedded in the space $W^{1,\infty}(\mathbb{T}\times\mathbb{R})$ (i.e., bounded Lipschitz functions), hence

[3] The appellation is taken from [58].

[4] It is also possible to consider a non-negative $\mu$ but the relevant condition is rather involved, we refer to [58] for details.

there exists a constant $C > 0$ such that, for all $\varepsilon \in (0,1)$,

$$\begin{aligned} W_1(f_\varepsilon, \mu) &= \sup_{\|\varphi\|_{\mathrm{Lip}} \le 1} \langle f_\varepsilon - \mu, \, \varphi \rangle \\ &\ge \sup_{\|\varphi\|_{W^{2,1}} \le C} \langle f_\varepsilon - \mu, \, \varphi \rangle \\ &= C \| f_\varepsilon - \mu \|_{W^{-2,1}}. \end{aligned}$$

Therefore, by (5.5.4) with $r = -2$, we deduce that

$$\liminf_{\varepsilon \to 0} \sup_{t \in [0, \varepsilon^\alpha]} W_1(f_\varepsilon, \mu) > 0,$$

which proves the claimed result.

## 5.6 Proof of Proposition 5.1.7

As we already mentioned in the introduction, in the case of one single Dirac mass in velocity, the situation is much more favorable. This was first shown by Brenier in [24] for the quasineutral limit of the classical Vlasov-Poisson system, using the so-called relative entropy (or modulated energy) method. It was then adapted by the first author in [55] for the quasineutral limit of (VPME).

In this case, the expected limit is the Dirac mass in velocity

$$f(t,x,v) = \rho(t,x,v) \delta_{v = u(t,x)}$$

which is a weak solution of (5.1.3) whenever $(\rho, u)$ is a strong solution to the isothermal Euler system

$$\begin{cases} \partial_t \rho + \partial_x(\rho u) = 0, \\ \partial_t u + u \partial_x u + \frac{\partial_x \rho}{\rho} = 0, \\ \rho|_{t=0} = \rho_0, v|_{t=0} = u_0. \end{cases} \tag{5.6.1}$$

This is a hyperbolic and symmetric system, that admits local smooth solutions for smooth initial data (in this Section, smooth means $H^s$ with $s$ larger than 2). From [55] we deduce the following stability result[5].

[5] In [55], computations are done for the model posed on $\mathbb{R}^3$, but the same holds for the model set on $\mathbb{T}$.

**Theorem 5.6.1.** *Let* $\rho_0 > 0, u_0$ *be some smooth initial conditions for* (5.6.1)*, and* $\rho, u$ *the associated strong solutions of* (5.6.1) *defined on some interval of time* $[0,T]$*, where* $T > 0$*. Let* $f_\varepsilon$ *be a non-negative global weak solution of* (5.1.2) *such that* $f_\varepsilon \in L^1 \cap L^\infty$*,* $\int f_\varepsilon \, dxdv = 1$*, and with uniformly bounded energy, i.e., there exists* $A > 0$*, such that for all* $\varepsilon \in (0,1)$*,*

$$\mathcal{E}_\varepsilon(t) := \frac{1}{2}\int f_\varepsilon |v|^2 dv\, dx + \int \left(e^{U_\varepsilon}\log e^{U_\varepsilon} - e^{U_\varepsilon} + 1\right) dx + \frac{\varepsilon^2}{2}\int |U_\varepsilon'|^2 dx \leq A.$$

*For all* $\varepsilon \in (0,1)$*, define the relative entropy*

$$\mathcal{H}_\varepsilon(t) := \frac{1}{2}\int f_\varepsilon |v-u|^2 dv\, dx + \int \left(e^{U_\varepsilon}\log\left(e^{U_\varepsilon}/\rho\right) - e^{U_\varepsilon} + \rho\right) dx + \frac{\varepsilon^2}{2}\int |U_\varepsilon'|^2 dx.$$

*Then there exists* $C > 0$ *and a function* $G_\varepsilon(t)$ *satisfying* $\|G_\varepsilon\|_{L^\infty([0,T])} \leq C\varepsilon$ *such that, for all* $t \in [0,T]$*,*

$$\mathcal{H}_\varepsilon(t) \leq \mathcal{H}_\varepsilon(0) + G_\varepsilon(t) + C\int_0^t \|\partial_x u\|_{L^\infty}\mathcal{H}_\varepsilon(s) ds.$$

*In particular, if* $\mathcal{H}_\varepsilon(0) \to_{\varepsilon\to 0} 0$*, then* $\mathcal{H}_\varepsilon(t) \to_{\varepsilon\to 0} 0$ *for all* $t \in [0,T]$*.*

*In addition, if there is* $C_0 > 0$ *such that* $\mathcal{H}_\varepsilon(0) \leq C_0\varepsilon$*, then there is* $C_T > 0$ *such that* $\mathcal{H}_\varepsilon(t) \leq C_T\varepsilon$ *for all* $t \in [0,T]$ *and* $\varepsilon \in (0,1)$*.*

Notice that, by a convexity argument, one also deduces that $\rho_\varepsilon = \int f_\varepsilon \, dv \rightharpoonup \rho$ (and $e^{U_\varepsilon} \rightharpoonup \rho$ as well) and $j_\varepsilon = \int f_\varepsilon v \, dv \rightharpoonup \rho u$ in a weak-$\star$ sense (see [55]).

We can actually deduce the following corollary, which is a precise version of Proposition 5.1.7.

**Corollary 5.6.2.** *With the same assumptions and notation as in the previous theorem, the following convergence results hold:*

1. *If* $\mathcal{H}_\varepsilon(0) \to_{\varepsilon\to 0} 0$*, then*
$$\sup_{t\in[0,T]} W_1(f_\varepsilon, \rho\,\delta_{v=u}) \to_{\varepsilon\to 0} 0.$$

2. *If* $\mathcal{H}_\varepsilon(0) \leq C_0\varepsilon$*, then there is* $C_T' > 0$ *such that, for all* $\varepsilon \in (0,1)$*,*
$$\sup_{t\in[0,T]} W_1(f_\varepsilon, \rho\,\delta_{v=u}) \leq C_T'\sqrt{\varepsilon}.$$

*Proof.* Recall that we denote $\rho_\varepsilon = \int f_\varepsilon \, dv$. Let $\varphi$ such that $\|\varphi\|_{\mathrm{Lip}} \leq 1$ and compute

$$\begin{aligned}\langle f_\varepsilon - \rho \, \delta_{v=u}, \, \varphi \rangle &= \langle f_\varepsilon - \rho_\varepsilon \, \delta_{v=u}, \, \varphi \rangle + \langle (\rho_\varepsilon - \rho) \, \delta_{v=u}, \, \varphi \rangle \\ &=: A_1 + A_2.\end{aligned}$$

Using the bound $\|\varphi\|_{\mathrm{Lip}} \leq 1$, the Cauchy-Schwarz inequality, the fact that $f_\varepsilon$ is non-negative and of total mass 1, and the definition of $\mathcal{H}_\varepsilon(t)$, we have

$$\begin{aligned}|A_1| &= \left| \int_{\mathbb{T}\times\mathbb{R}} f_\varepsilon(t,x,v)(\varphi(t,x,v) - \varphi(t,x,u(t,x)) \, dvdx \right| \\ &\leq \int_{\mathbb{T}\times\mathbb{R}} f_\varepsilon(t,x,v)|\varphi(t,x,v) - \varphi(t,x,u(t,x))| \, dvdx \\ &\leq \int_{\mathbb{T}\times\mathbb{R}} f_\varepsilon(t,x,v)|v - u(t,x)| \, dvdx \\ &\leq \left( \int_{\mathbb{T}\times\mathbb{R}} f_\varepsilon(t,x,v)|v - u(t,x)|^2 \, dvdx \right)^{1/2} \left( \int_{\mathbb{T}\times\mathbb{R}} f_\varepsilon(t,x,v) \, dvdx \right)^{1/2} \\ &\leq \sqrt{2}\sqrt{\mathcal{H}_\varepsilon(t)}.\end{aligned}$$

Considering $A_2$, we first have

$$\begin{aligned}A_2 &= \int_{\mathbb{T}} (\rho_\varepsilon(t,x) - \rho(t,x))\varphi(x,u(t,x)) \, dx \\ &= \int_{\mathbb{T}} (\rho_\varepsilon(t,x) - \rho(t,x))\big(\varphi(x,u(t,x)) - \varphi(0,0)\big) \, dx\end{aligned}$$

since the total mass is preserved (and equal to 1). Furthermore, we use the Poisson equation

$$\rho_\varepsilon = e^{U_\varepsilon} - \varepsilon^2 U_\varepsilon'',$$

to rewrite $A_2$ as

$$\begin{aligned}A_2 = &\int_{\mathbb{T}} (e^{U_\varepsilon} - \rho(t,x)) \left(\varphi(x,u(t,x)) - \varphi(0,0)\right) \, dx \\ &- \varepsilon^2 \int U_\varepsilon'' \left(\varphi(x,u(t,x)) - \varphi(0,0)\right) \, dx \\ =: &\; A_2^1 + A_2^2.\end{aligned}$$

Let us start with $A_2^2$. By integration by parts, the Cauchy-Schwarz inequality, and the bound $\|\varphi\|_{\mathrm{Lip}} \leq 1$, we have

$$\begin{aligned}
|A_2^2| &= \varepsilon^2 \left| \int_{\mathbb{T}} U_\varepsilon' [\partial_x \varphi(x, u(t,x)) + \partial_x u(t,x) \partial_v \varphi(x, u(t,x))] \, dx \right| \\
&\leq \varepsilon [1 + \|\partial_x u\|_\infty] \left( \varepsilon^2 \int_{\mathbb{T}} |U_\varepsilon'|^2 \, dx \right)^{1/2} \\
&\leq \varepsilon [1 + \|\partial_x u\|_\infty] \sqrt{\mathcal{E}_\varepsilon(t)} \\
&\leq \sqrt{A} [1 + \|\partial_x u\|_\infty] \varepsilon.
\end{aligned}$$

For $A_2^1$, we shall use the classical inequality

$$(\sqrt{y} - \sqrt{x})^2 \leq x \log(x/y) - x + y,$$

for $x, y > 0$, and proceed as follows:

$$\begin{aligned}
|A_2^1| &= \left| \int_{\mathbb{T}} \left( e^{U_\varepsilon} - \rho(t,x) \right) \left( \varphi(x, u(t,x)) - \varphi(0,0) \right) \, dx \right| \\
&= \left| \int_{\mathbb{T}} \left( e^{\frac{1}{2} U_\varepsilon} - \sqrt{\rho(t,x)} \right) \left( e^{\frac{1}{2} U_\varepsilon} + \sqrt{\rho(t,x)} \right) \left( \varphi(x, u(t,x)) - \varphi(0,0) \right) \, dx \right| \\
&\leq \left( \int_{\mathbb{T}} \left( e^{\frac{1}{2} U_\varepsilon} - \sqrt{\rho(t,x)} \right)^2 \left| \varphi(x, u(t,x)) - \varphi(0,0) \right| \, dx \right)^{1/2} \\
&\quad \times \left( \int_{\mathbb{T}} \left( e^{\frac{1}{2} U_\varepsilon} + \sqrt{\rho(t,x)} \right)^2 \left| \varphi(x, u(t,x)) - \varphi(0,0) \right| \, dx \right)^{1/2}.
\end{aligned}$$

We have

$$\begin{aligned}
&\left( \int_{\mathbb{T}} \left( e^{\frac{1}{2} U_\varepsilon} - \sqrt{\rho(t,x)} \right)^2 \left| \varphi(x, u(t,x)) - \varphi(0,0) \right| \, dx \right) \\
&\qquad \leq (1 + \|u\|_\infty) \int \left( e^{U_\varepsilon} \log \left( e^{U_\varepsilon} / \rho \right) - e^{U_\varepsilon} + \rho \right) \, dx \\
&\qquad \leq (1 + \|u\|_\infty) \mathcal{H}_\varepsilon(t),
\end{aligned}$$

and likewise we obtain the rough bound

$$\begin{aligned}
&\left(\int_{\mathbb{T}} \left(e^{\frac{1}{2}U_\varepsilon} + \sqrt{\rho(t,x)}\right)^2 |\varphi(x,u(t,x)) - \varphi(0,0)|\, dx\right) \\
&\quad\le 2\int_{\mathbb{T}} \left(e^{\frac{1}{2}U_\varepsilon} - 1\right)^2 |\varphi(x,u(t,x)) - \varphi(0,0)|\, dx \\
&\quad + 2\int_{\mathbb{T}} \left(\sqrt{\rho(t,x)} + 1\right)^2 |\varphi(x,u(t,x)) - \varphi(0,0)|\, dx \\
&\quad\le 2(1+\|u\|_\infty)\left(\mathcal{E}_\varepsilon(t) + \int_{\mathbb{T}} \left(\sqrt{\rho(t,x)} + 1\right)^2 dx\right) \\
&\quad\le 2(1+\|u\|_\infty)\left(A + \int_{\mathbb{T}} \left(\sqrt{\rho(t,x)} + 1\right)^2 dx\right).
\end{aligned}$$

As a consequence, we get

$$|A_2^2| \le \sqrt{2}(1+\|u\|_\infty)\left(A + \int_{\mathbb{T}} \left(\sqrt{\rho(t,x)} - 1\right)^2 dx\right)^{1/2} \sqrt{\mathcal{H}_\varepsilon(t)}.$$

Gathering all pieces together, we have shown

$$\begin{aligned}
&\langle f_\varepsilon - \rho\delta_{v=u}, \varphi\rangle \\
&\le \sqrt{2}\left[1 + (1+\|u\|_\infty)\left(A + \int_{\mathbb{T}} \left(\sqrt{\rho(t,x)} + 1\right)^2 dx\right)^{1/2}\right] \sqrt{\mathcal{H}_\varepsilon(t)} + \sqrt{A}[1+\|\partial_x u\|_\infty]\varepsilon,
\end{aligned}$$

which allows us to conclude the proof applying Theorem 5.6.1. □

# Chapter 6

# Quasineutral limit for Vlasov-Poisson via Wasserstein stability estimates in higher dimension

[1]

## 6.1 Introduction

In a non relativistic setting the dynamics of electrons in a plasma with heavy ions uniformly distributed in space is described by the Vlasov-Poisson system. Throughout this Chapter, we will focus on the 2 and 3 dimensional periodic (in space) case. We introduce the distribution function of the electrons $f(t, x, v)$, for $t \in \mathbb{R}^+$, $(x, v) \in \mathbb{T}^d \times \mathbb{R}^d$ where $\mathbb{T}^d$ is the $d$-dimensional torus and $d = 2, 3$. As usual, $f(t, x, v)dx\,dv$ can be interpreted as the probability of finding particles with position and velocity close to the point $(x, v)$ in the phase space at time $t$. We also define the electric potential $U(t, x)$ and the associated electric field $E(t, x)$.

We introduce the positive parameter $\varepsilon$ defined as the ratio of the *Debye length* of the plasma to the size of the domain. Adding a subscript in order to emphasize on the

[1] This chapter is based on a joint work with Danial Han-Kwan [60].

dependance on $\varepsilon$, we end up with the rescaled Vlasov-Poisson system:

$$\begin{cases} \partial_t f_\varepsilon + v \cdot \nabla_x f_\varepsilon + E_\varepsilon \cdot \nabla_v f_\varepsilon = 0, \\ E_\varepsilon = -\nabla_x U_\varepsilon, \\ -\varepsilon^2 \Delta_x U_\varepsilon = \int_{\mathbb{R}^d} f_\varepsilon \, dv - \int_{\mathbb{T}^d \times \mathbb{R}^d} f_\varepsilon \, dv \, dx, \\ f_\varepsilon|_{t=0} = f_{0,\varepsilon} \geq 0, \quad \int_{\mathbb{T}^d \times \mathbb{R}^d} f_{0,\varepsilon} \, dx \, dv = 1, \end{cases} \tag{6.1.1}$$

and the energy of this system is

$$\mathcal{E}(f_\varepsilon(t)) := \frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} f_\varepsilon |v|^2 \, dv dx + \frac{\varepsilon^2}{2} \int_{\mathbb{T}^d} |\nabla_x U_\varepsilon|^2 \, dx. \tag{6.1.2}$$

Our goal here is to study the behavior of solutions to the system (6.1.1) as $\varepsilon$ goes to 0. Let us observe that, if $f_\varepsilon \to f$ and $U_\varepsilon \to U$ in some sense as $\varepsilon \to 0$, the formal limit of our system is

$$\begin{cases} \partial_t f + v \cdot \nabla_x f + E \cdot \nabla_v f = 0, \\ E = -\nabla_x U, \\ \int_{\mathbb{R}^d} f \, dv = 1, \\ f|_{t=0} = f_0 \geq 0, \quad \int_{\mathbb{T}^d \times \mathbb{R}^d} f_0 \, dx \, dv = 1, \end{cases} \tag{6.1.3}$$

and the total energy of the system reduces to the kinetic part of (6.1.2)

$$\mathcal{E}(f(t)) := \frac{1}{2} \int_{\mathbb{T}^d \times \mathbb{R}^d} f |v|^2 \, dv dx.$$

In this system, the force $E$ is a Lagrange multiplier, or a pressure, associated to the constraint $\int_{\mathbb{R}^d} f \, dv = 1$.

The justification of the quasineutral limit from the rescaled Vlasov-Poisson system (6.1.1) to (6.1.3) is subtle and has a long history. Up to now, this limit is known to be true only in few cases and we refer to [22, 48, 49, 24, 78, 58] for a deeper understanding of this problem.

One of the first mathematical works on the quasineutral limit of the Vlasov-Poisson system was performed by Grenier in [49]. He introduces an interpretation of the plasma as a superposition of a -possibly uncountable- collection of fluids and he shows that the quasineutral limit holds when the sequence of initial data $f_{0,\varepsilon}$ enjoys uniform analytic regularity with respect to the space variable. This convergence result has been improved by Brenier [24], who gives a rigorous justification of the quasineutral limit in the so called "cold electron" case, i.e. when the initial distribution $f_{0,\varepsilon}$ converges to a monokinetic profile

$$f_0(x, v) = \rho_0(x) \delta_{v = v_0(x)}$$

where $\delta_v$ denotes the Dirac measure in velocity. For further insight on this direction see also [24, 78, 46].

A different approach, more focused on the question of stability, or eventually instability, around homogeneous equilibria in the quasineutral limit is developed in [58]. They show that the limit is true for homogeneous profiles that satisfy some monotonicity condition, together with a symmetry condition, i.e. when the initial distribution $f_{0,\varepsilon}$ converges to an homogeneous initial condition $\mu(v)$ which is symmetric with respect to some $\overline{v} \in \mathbb{R}$ and which is first increasing then decreasing.

In the previous Chapter, we considered the quasineutral limit of the one-dimensional Vlasov-Poisson equation for ions with massless thermalized electrons (considering that electrons move very fast and quasi-instantaneously reach their local thermodynamic equilibrium), and we proved that the limit holds for very small but rough perturbations of analytic data. In this context, small means small in the Wasserstein distance $W_1$, which implies that highly oscillatory perturbations are for instance allowed. Our aim here is to show that an analogue of this result holds in higher dimension.

In this Chapter we shall always deal with the *Wasserstein space* $\mathcal{P}_2(\mathcal{M})$, that is, the space of probability measures which have a finite moment of order 2 equipped with the quadratic Wasserstein distance $W_2$ (see Section 1.2 for more details).

In order to state our main result, let us introduce the fluid point of view and the convergence result for uniformly analytic initial data introduced by Grenier in [49].

Below we shall use again the higher dimensional analogue of the analytic norm $\|\cdot\|_{B_\delta}$ introduced in Definition 5.1.2.

We assume that, for all $\varepsilon \in (0,1)$, $g_{0,\varepsilon}(x,v)$ is a *continuous* function; following Grenier [49], we write each initial condition as a "superposition of Dirac masses in velocity":

$$g_{0,\varepsilon}(x,v) = \int_{\mathcal{M}} \rho^\theta_{0,\varepsilon}(x)\delta_{v=v^\theta_{0,\varepsilon}(x)}\, d\mu(\theta)$$

with $\mathcal{M} := \mathbb{R}^d$, $d\mu(\theta) = c_d \frac{d\theta}{1+|\theta|^{d+1}}$, where $c_d$ is a normalizing constant (depending only on the dimension $d$),

$$\rho^\theta_{0,\varepsilon} = \frac{1}{c_d}(1+|\theta|^{d+1})g_{0,\varepsilon}(x,\theta), \quad v^\theta_{0,\varepsilon} = \theta.$$

This leads to the study of the behavior as $\varepsilon \to 0$ for solutions to the *multi-fluid pres-*

*sureless Euler-Poisson system*

$$\begin{cases} \partial_t \rho_\varepsilon^\theta + \nabla_x \cdot (\rho_\varepsilon^\theta v_\varepsilon^\theta) = 0, \\ \partial_t v_\varepsilon^\theta + v_\varepsilon^\theta \cdot \nabla_x v_\varepsilon^\theta = E_\varepsilon, \\ E_\varepsilon = -\nabla_x U_\varepsilon, \\ -\varepsilon^2 \Delta_x U_\varepsilon = \int_{\mathcal{M}} \rho_\varepsilon^\theta \, d\mu(\theta) - 1, \\ \rho_\varepsilon^\theta|_{t=0} = \rho_{0,\varepsilon}^\theta, v_\varepsilon^\theta|_{t=0} = v_{0,\varepsilon}^\theta. \end{cases} \tag{6.1.4}$$

One then checks that defining

$$g_\varepsilon(t,x,v) = \int_{\mathcal{M}} \rho_\varepsilon^\theta(t,x) \delta_{v=v_\varepsilon^\theta(t,x)} \, d\mu(\theta)$$

provides a weak solution to (6.1.1).

The formal limit system, which is associated to the kinetic incompressible Euler system (6.1.3), is the following *multi fluid incompressible Euler system*:

$$\begin{cases} \partial_t \rho^\theta + \nabla_x \cdot (\rho^\theta v^\theta) = 0, \\ \partial_t v^\theta + v^\theta \cdot \nabla_x v^\theta = E, \\ \operatorname{curl} E = 0, \int_{\mathbb{T}^d} E \, dx = 0, \\ \int_{\mathcal{M}} \rho^\theta \, d\mu(\theta) = 1, \\ \rho^\theta|_{t=0} = \rho_0^\theta, v^\theta|_{t=0} = v_0^\theta, \end{cases} \tag{6.1.5}$$

where the $\rho_0^\theta$ are defined as the limits of $\rho_{0,\varepsilon}^\theta$ (which are thus supposed to exist) and $v_0^\theta = \theta$.

As before, one checks that defining

$$g(t,x,v) = \int_{\mathcal{M}} \rho^\theta(t,x) \delta_{v=v^\theta(t,x)} \, d\mu(\theta)$$

gives a weak solution to the kinetic Euler incompressible system (6.1.3).

We are now in position to state the results of [49, Theorems 1.1.2, 1.1.3 and Remark 1 p. 369].

**Proposition 6.1.1.** *Assume that there exist $\delta_0, C, \eta > 0$, with $\eta$ small enough, such that*

$$\sup_{\varepsilon \in (0,1)} \sup_{v \in \mathbb{R}} (1+v^2) \|g_{0,\varepsilon}(\cdot, v)\|_{B_{\delta_0}} \le C,$$

*and that*

$$\sup_{\varepsilon \in (0,1)} \left\| \int_{\mathbb{R}} g_{0,\varepsilon}(\cdot, v) \, dv - 1 \right\|_{B_{\delta_0}} < \eta.$$

*Denote for all $\theta \in \mathbb{R}$,*

$$\rho^\theta_{0,\varepsilon} = \pi(1+\theta^2) g_{0,\varepsilon}(x,\theta), \quad v^\theta_{0,\varepsilon} = v^\theta = \theta.$$

*Assume that for all $\theta \in \mathbb{R}$, $\rho^\theta_{0,\varepsilon}$ has a limit in the sense of distributions and denote*

$$\rho^\theta_0 = \lim_{\varepsilon \to 0} \rho^\theta_{0,\varepsilon}.$$

*Then there exist $\delta_1 > 0$ and $T > 0$ such that:*

- *for all $\varepsilon \in (0,1)$, there is a unique solution $(\rho^\theta_\varepsilon, v^\theta_\varepsilon)_{\theta \in M}$ of (6.1.4) with initial data $(\rho^\theta_{0,\varepsilon}, v^\theta_{0,\varepsilon})_{\theta \in M}$, such that $\rho^\theta_\varepsilon, v^\theta_\varepsilon \in C([0,T]; B_{\delta_1})$ for all $\theta \in M$ and $\varepsilon \in (0,1)$, with bounds that are uniform in $\varepsilon$;*
- *there is a unique solution $(\rho^\theta, v^\theta)_{\theta \in M}$ of (6.1.5) with initial data $(\rho^\theta_0, v^\theta_0)_{\theta \in M}$, such that $\rho^\theta, v^\theta \in C([0,T]; B_{\delta_1})$ for all $\theta \in M$;*
- *for all $s \in \mathbb{N}$, we have*

$$\sup_{\theta \in M} \sup_{t \in [0,T]} \left[ \|\rho^\theta_\varepsilon - \rho^\theta\|_{H^s(\mathbb{T})} + \|v^\theta_\varepsilon - \frac{1}{i}(d_+(t,x) e^{\frac{it}{\sqrt{\varepsilon}}} - d_-(t,x) e^{-\frac{it}{\sqrt{\varepsilon}}}) - v^\theta\|_{H^s(\mathbb{T})} \right] \to_{\varepsilon \to 0} 0 \tag{6.1.6}$$

*where $d_\pm(t,x)$ are the correctors introduced to avoid the so called "plasma oscillations". They are defined as the solution of*

$$\text{curl } d_\pm = 0, \qquad \text{div} \left( \partial_t d_\pm + \left( \int \rho_\theta v_\theta \mu(d\theta) \cdot \nabla \right) d_\pm \right) = 0, \tag{6.1.7}$$

$$\text{div}\, d_\pm(0) = \lim_{\varepsilon \to 0} \text{div} \frac{\sqrt{\varepsilon} E^\varepsilon(0) \pm i j^\varepsilon(0)}{2}, \tag{6.1.8}$$

*where $j^\varepsilon := \int \rho^\varepsilon_\theta v^\varepsilon_\theta \mu(d\theta)$.*

**Remark 6.1.2.** *If in (6.1.8), $\text{div}\, d_\pm(0) = 0$, then the initial data are said to be well-prepared and there are no plasma oscillations in the limit $\varepsilon \to 0$.*

The main result is the following:

**Theorem 6.1.3.** *Let $\gamma$, $\delta_0$, and $C_0$ be positive constants. Consider a sequence $(f_{0,\varepsilon})$ of non-negative initial data in $L^1$ for (6.1.1) such that for all $\varepsilon \in (0,1)$, and all $x \in \mathbb{T}^d$,*

- *(uniform estimates)*

$$\|f_{0,\varepsilon}\|_\infty \le C_0, \quad \mathcal{E}(f_{0,\varepsilon}) \le C_0,$$

- *(compact support in velocity)*

$$f_{0,\varepsilon}(x,v) = 0 \quad \text{if } |v| > \frac{1}{\varepsilon^\gamma},$$

- *(analytic + perturbation) There exists a function* $\varphi : (0,1] \to \mathbb{R}^+$, *with* $\lim_{\varepsilon\to 0}\varphi(\varepsilon) = 0$ *such that the following hold. Assume the following decomposition:*

$$f_{0,\varepsilon} = g_{0,\varepsilon} + h_{0,\varepsilon},$$

*where* $(g_{0,\varepsilon})$ *is a sequence of continuous functions satisfying*

$$\sup_{\varepsilon\in(0,1)} \sup_{v\in\mathbb{R}^d} (1+|v|^2)\|g_{0,\varepsilon}(\cdot,v)\|_{B_{\delta_0}} \le C,$$

*admitting a limit* $g_0$ *in the sense of distributions. Furthemore,* $(h_{0,\varepsilon})$ *is a sequence of functions satisfying for all* $\varepsilon > 0$

$$W_2(f_{0,\varepsilon}, g_{0,\varepsilon}) = \varphi(\varepsilon).$$

*For all* $\varepsilon \in (0,1)$, *consider* $f_\varepsilon(t)$ *a global weak solution of* (6.1.1) *with initial condition* $f_{0,\varepsilon}$, *in the sense of Arsenev [7]. Define the filtered distribution function*

$$\widetilde{f_\varepsilon}(t,x,v) := f_\varepsilon\Big(t,x,v - \frac{1}{i}(d_+(t,x)e^{\frac{it}{\sqrt{\varepsilon}}} - d_-(t,x)e^{-\frac{it}{\sqrt{\varepsilon}}})\Big) \tag{6.1.9}$$

*where* $(d_\pm)$ *are defined in* (6.1.7).

*There exist* $T > 0$ *and* $g(t)$ *a weak solution on* $[0,T]$ *of* (6.1.3) *with initial condition* $g_0$ *such that*

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} W_1(\widetilde{f_\varepsilon}(t), g(t)) = 0.$$

*Explicitly, we can take*

- *in two dimensions,* $\varphi(\varepsilon) = \exp\left[\exp\left(-\frac{K}{\varepsilon^{2(1+\max(\beta,\gamma))}}\right)\right]$, *for some constant* $K > 0$, $\beta > 2$;
- *in three dimensions,* $\varphi(\varepsilon) = \exp\left[\exp\left(-\frac{K}{\varepsilon^{2+\max(38,3\gamma))}}\right)\right]$, *for some constant* $K > 0$.

**Remark 6.1.4.** *Let us notice that in Theorem 6.1.3 we consider sequences of initial conditions with compact support in velocity (yet, we allow the support to grow polynomially as* $\varepsilon \to 0$*). The reason is that, in the spirit of [59], we rely on a Wasserstein stability estimate to control the difference between the unperturbed analytic solution and the perturbed one. In dimensions* 2 *and* 3*, as we shall explain below, we need* $L^\infty$ *bounds on the densities of* both *solutions. In order to have such a bound on the* $L^\infty$ *norm of the densities we need to control the support in velocity. Such a condition was not required in the previous Chapter since, in the 1D case, we could use a "weak-strong" Wasserstein stability estimate and only a* $L^\infty$ *bound on the unperturbed solution was needed.*

**Remark 6.1.5.** *In the opposite direction, we recall that in the one dimensional case there is a negative result stating that an initial rate of convergence of the form* $\varphi(\varepsilon) = \varepsilon^s$ *for any* $s > 0$ *is not sufficient to ensure the convergence for positive times. This is the consequence of* instability *mechanisms described in [50] and [58]. As a matter of fact, we expect that an analogue of this result holds also in higher dimension.*

## 6.2 Overview

The following is entirely devoted to the proof of Theorem 6.1.3. Let us describe the main steps that are needed to achieve this convergence result.

1. We first revisit Loeper's Wasserstein stability estimates [76] on the torus $\mathbb{T}^d$ and with quasineutral scaling, which allows us to control $W_2(f_1, f_2)$, where $f_1$ and $f_2$ are two given solutions of (6.1.1) in terms of the initial distance $W_2(f_1(0), f_2(0))$ and of the $L^\infty$ norm of the densities $\rho_1 = \int_{\mathbb{R}^d} f_1\, dv$ and $\rho_2 = \int_{\mathbb{R}^d} f_2\, dv$. This first step is performed in Section 6.3.1.

2. In the one dimensional case studied in the previous Chapter we had a "weak-strong" type stability estimate; as a consequence a control of the $L^\infty$ norm of the density of the perturbed solution $f_\varepsilon$ (following the notations of Theorem 6.1.3) was not required.

   In the higher dimensional case under study, such an estimate is needed. To achieve this, we give quantitative estimates of the growth of the support in velocity for solutions of (6.1.1). We separate the 2 and the 3-dimensional case since different tools are involved.

While in the two dimensional case studied in Section 6.3.2 only elementary considerations are needed, in the three dimensional case, we shall use a more involved bootstrap argument due to Batt and Rein [8], see Section 6.3.3.

3. We finally conclude in Section 6.4 by combining the results of the two previous steps and Grenier's convergence result stated in Proposition 6.1.1.

# 6.3 Proofs of Steps 1 and 2

## 6.3.1 $W_2$ stability estimate

We start by giving the relevant $W_2$ stability estimate, adapting from the work of Loeper [76].

**Theorem 6.3.1.** *Let $f_1, f_2$ be two weak solutions of the Vlasov-Poisson system* (6.1.1)*, and set*

$$\rho_1 := \int_{\mathbb{R}^d} f_1\, dv, \quad \rho_2 = \int_{\mathbb{R}^d} f_2\, dv.$$

*Define the function*

$$A(t) := \left[1+\frac{1}{\varepsilon^2}\sqrt{\|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}}\Big[\max\big\{\|\rho_1(t)\|_{L^\infty(\mathbb{T}^d)}, \|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}\big\}\Big]^{1/2}\right]+\frac{\|\rho_1(t)-1\|_{L^\infty(\mathbb{T}^d)}}{\varepsilon^2}, \tag{6.3.1}$$

*and assume that $A(t) \in L^1([0,T])$ for some $T > 0$. Also, set*

$$F_t[z] := 16\, d\, e^{\log\left(\frac{z}{16\, d}\right) \exp\left[C_0 \int_0^t A(s)\, ds\right]} \qquad \forall\, z \in [0,d],\ t \in [0,T]. \tag{6.3.2}$$

*Then there exists a dimensional constant $C_0 > 1$ such that, if $W_2(f_1(0), f_2(0)) \leq d$, then for all $t \in [0,T]$,*

$$W_2(f_1(t), f_2(t)) \leq \begin{cases} F_t[W_2(f_1(0), f_2(0))] & \text{if } F_T[W_2(f_1(0), f_2(0))] \leq d, \\ d\, e^{C_0 \int_0^t A(s)\, ds} & \text{if } F_T[W_2(f_1(0), f_2(0))] > d. \end{cases} \tag{6.3.3}$$

*Proof of Theorem 6.3.1.* Before starting the proof we recall two important estimates that follow immediately from [76, Theorem 2.7] and the analogue of [76, Lemma 3.1] on the torus (notice that $|x-y| \leq \sqrt{d}$ for all $x, y \in \mathbb{T}^d$):

**Lemma 6.3.2.** *Let $\Psi_i : \mathbb{T}^d \to \mathbb{R}$ solve*

$$-\varepsilon^2 \Delta \Psi_i = \rho_i - 1, \qquad i = 1, 2.$$

*Then*

$$\varepsilon^2 \|\nabla \Psi_1 - \nabla \Psi_2\|_{L^2(\mathbb{T}^d)} \leq \Big[\max\big\{\|\rho_1\|_{L^\infty(\mathbb{T}^d)}, \|\rho_2\|_{L^\infty(\mathbb{T}^d)}\big\}\Big]^{1/2} W_2(\rho_1, \rho_2),$$

$$\varepsilon^2 |\nabla \Psi_i(x) - \nabla \Psi_i(y)| \leq C\,|x-y|\,\log\left(\frac{4\sqrt{d}}{|x-y|}\right) \|\rho_i - 1\|_{L^\infty(\mathbb{T}^d)} \qquad \forall\, x, y \in \mathbb{T}^d,\ i = 1, 2.$$

To prove Theorem 6.3.1, we define the quantity

$$Q(t) := \int_0^1 |Y_1(t, S_1(s)) - Y_2(t, S_2(s))|^2\, ds$$

where

$$S_1, S_2 : [0,1] \to \mathbb{T}^d \times \mathbb{R}^d$$

are measurable maps such that $(S_i)_\# ds = f_i(0)$ and

$$W_2(f_1(0), f_2(0))^2 = \int_0^1 |S_1(s) - S_2(s)|^2\, ds,$$

while $Y_i = (X_i, V_i)$ solve the ODE

$$\begin{aligned} \dot{X}_i &= V_i, \\ \dot{V}_i &= -\nabla \Psi_i(t, X_i) \end{aligned}$$

with the initial condition $Y_i(0, x, v) = (x, v)$.

Thus, thanks to [76, Corollary 3.3] it follows that $f_i(t) = Y_i(t)_\# f_i(0) = [Y_i(t, S_i)]_\# ds$.

Then we compute

$$\begin{aligned} \frac{1}{2}\frac{d}{dt} Q(t) = &\int_0^1 [X_1(t, S_1) - X_2(t, S_2)]\,[V_1(t, S_1) - V_2(t, S_2)]\, ds \\ &+ \int_0^1 [V_1(t, S_1) - V_2(t, S_2)]\,\big[\nabla \Psi_1\big(t, X_1(t, S_1)\big) - \nabla \Psi_2\big(t, X_2(t, S_2)\big)\big]\, ds. \end{aligned}$$

By Cauchy-Schwarz inequality, we have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}Q(t) \leq &\sqrt{\int_0^1 |X_1(t,S_1) - X_2(t,S_2)|^2\,ds}\sqrt{\int_0^1 |V_1(t,S_1) - V_2(t,S_2)|^2\,ds}\\ &+ \sqrt{\int_0^1 |V_1(t,S_1) - V_2(t,S_2)|^2\,ds}\sqrt{\int_0^1 \big|\nabla\Psi_1\big(t,X_1(t,S_1)\big) - \nabla\Psi_2\big(t,X_2(t,S_2)\big)\big|^2\,ds}\\ &\leq Q(t) + \sqrt{Q(t)}\sqrt{\int_0^1 \big|\nabla\Psi_1\big(t,X_1(t,S_1)\big) - \nabla\Psi_1\big(t,X_2(t,S_2)\big)\big|^2\,ds}\\ &+ \sqrt{Q(t)}\sqrt{\int_0^1 \big|\nabla\Psi_1\big(t,X_2(t,S_2)\big) - \nabla\Psi_2\big(t,X_2(t,S_2)\big)\big|^2\,ds}.\end{aligned}$$

Using the definition of the push-forward, we finally get

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}Q(t) \leq\ &Q(t) + \sqrt{Q(t)}\sqrt{\int_0^1 \big|\nabla\Psi_1\big(t,X_1(t,S_1)\big) - \nabla\Psi_1\big(t,X_2(t,S_2)\big)\big|^2\,ds}\\ &+ \sqrt{Q(t)}\sqrt{\int_{\mathbb{T}^d\times\mathbb{R}^d} |\nabla\Psi_1(t,x) - \nabla\Psi_2(t,x)|^2 f_2(t,x,v)\,dx\,dv}\\ \leq\ &Q(t) + \sqrt{Q(t)}\sqrt{\int_0^1 \big|\nabla\Psi_1\big(t,X_1(t,S_1)\big) - \nabla\Psi_1\big(t,X_2(t,S_2)\big)\big|^2\,ds}\\ &+ \sqrt{\|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}}\sqrt{Q(t)}\sqrt{\int_{\mathbb{T}^d} |\nabla\Psi_1(t,x) - \nabla\Psi_2(t,x)|^2\,dx}.\end{aligned}$$

We now apply Lemma 6.3.2 to the last two terms and we bound them respectively by

$$C\,\frac{\|\rho_1(t) - 1\|_{L^\infty(\mathbb{T}^d)}}{\varepsilon^2}\sqrt{Q(t)}\sqrt{\int_0^1 \big|X_1(t,S_1) - X_2(t,S_2)\big|^2 \log^2\left(\frac{4\sqrt{d}}{\big|X_1(t,S_1) - X_2(t,S_2)\big|}\right)ds}$$

and

$$\frac{1}{\varepsilon^2}\sqrt{\|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}}\Big[\max\big\{\|\rho_1(t)\|_{L^\infty(\mathbb{T}^d)}, \|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}\big\}\Big]^{1/2} W_2(\rho_1(t),\rho_2(t)).$$

Since $W_2(\rho_1(t),\rho_2(t)) \le \sqrt{Q(t)}$ (see for instance [76, Lemma 3.6]) we conclude that

$$\frac{1}{2}\frac{d}{dt}Q(t) \le \left[1+\frac{1}{\varepsilon^2}\sqrt{\|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}}\Big[\max\big\{\|\rho_1(t)\|_{L^\infty(\mathbb{T}^d)},\|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}\big\}\Big]^{1/2}\right]Q(t)$$
$$+C\,\frac{\|\rho_1(t)-1\|_{L^\infty(\mathbb{T}^d)}}{\varepsilon^2}\sqrt{Q(t)}\sqrt{\int_0^1 \big|X_1(t,S_1)-X_2(t,S_2)\big|^2\log^2\left(\frac{4\sqrt{d}}{\big|X_1(t,S_1)-X_2(t,S_2)\big|}\right)ds}.$$

Noticing that $\big|X_1(t,S_1)-X_2(t,S_2)\big| \le \sqrt{d}$ (since $X_1$ and $X_2$ are points on the torus) and

$$\log\left(\frac{4\sqrt{d}}{z}\right)=\frac{1}{2}\log\left(\frac{16\,d}{z^2}\right)\qquad \forall\, z>0,$$

we get

$$\int_0^1 \big|X_1(t,S_1)-X_2(t,S_2)\big|^2\log^2\left(\frac{4\sqrt{d}}{\big|X_1(t,S_1)-X_2(t,S_2)\big|}\right)ds$$
$$=\frac{1}{4}\int_0^1 \big|X_1(t,S_1)-X_2(t,S_2)\big|^2\log^2\left(\frac{16\,d}{\big|X_1(t,S_1)-X_2(t,S_2)\big|^2}\right)ds$$
$$=\frac{1}{4}\int_0^1 g(s)\log^2\left(\frac{16\,d}{g(s)}\right)ds,$$

where we set $g(s) := \big|X_1(t,S_1)-X_2(t,S_2)\big|^2$.

Hence, since the function

$$z\mapsto H(z) := \begin{cases} z\log^2\left(\frac{16\,d}{z}\right) & \text{for } 0\le z\le d,\\ d\log^2(16) & \text{for } z\ge d,\end{cases} \tag{6.3.4}$$

is concave and increasing, recalling that $g\le d$ and applying Jensen's inequality to $H$ we get

$$\frac{1}{2}\frac{d}{dt}Q(t) \le \left[1+\frac{1}{\varepsilon^2}\sqrt{\|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}}\Big[\max\big\{\|\rho_1(t)\|_{L^\infty(\mathbb{T}^d)},\|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}\big\}\Big]^{1/2}\right]Q(t)$$
$$+\,C\,\frac{\|\rho_1(t)-1\|_{L^\infty(\mathbb{T}^d)}}{\varepsilon^2}\sqrt{Q(t)}\sqrt{H\left(\int_0^1 g(s)\,ds\right)}$$
$$\le \left(1+\frac{1}{\varepsilon^2}\sqrt{\|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}}\Big[\max\big\{\|\rho_1(t)\|_{L^\infty(\mathbb{T}^d)},\|\rho_2(t)\|_{L^\infty(\mathbb{T}^d)}\big\}\Big]^{1/2}\right)Q(t)$$
$$+\,C\,\frac{\|\rho_1(t)-1\|_{L^\infty(\mathbb{T}^d)}}{\varepsilon^2}\sqrt{Q(t)}\sqrt{H\big(Q(t)\big)},$$

where for the last inequality we used that $\int_0^1 g(s)\,ds \leq Q(t)$.

In particular, recalling the definition of $A(t)$ in (6.3.1), by (6.3.4) we deduce that there exists a dimensional constant $C_0 > 0$ such that

$$\frac{d}{dt}Q(t) \leq C_0\,A(t)\,Q(t)\log\left(\frac{16\,d}{Q(t)}\right) \qquad \text{as long as } Q(t) \leq d, \tag{6.3.5}$$

while

$$\frac{d}{dt}Q(t) \leq C_0\,A(t)\,Q(t) \qquad \text{when } Q(t) \geq d. \tag{6.3.6}$$

In particular, assuming $Q(0) \leq d$, by (6.3.5) we get

$$Q(t) \leq 16\,d\,e^{\log\left(\frac{Q(0)}{16\,d}\right)\exp\left[C_0\int_0^t A(s)\,ds\right]} =: F_t[Q(0)] \tag{6.3.7}$$

as long as $Q(t) \leq d$, which is the case in particular if $F_t[Q(0)] \leq d$. On the other hand, if there is some time $t_0$ such that $F_{t_0}[Q(0)] = d$, since $Q(t_0) \leq F_{t_0}[Q(0)]$ by (6.3.6) we get

$$Q(t) \leq d\,e^{C_0\int_{t_0}^t A(s)\,ds} \leq d\,e^{C_0\int_0^t A(s)\,ds} \qquad \text{for } t \geq t_0. \tag{6.3.8}$$

Noticing that $F_t$ is monotone in $t$, we deduce in particular that if $F_T[Q(0)] \leq d$ and $Q(0) \leq d$ then (6.3.7) holds, while if $F_T[Q(0)] > d$ and $Q(0) \leq d$ then one can simply apply (6.3.8). Finally, if $Q(0) > d$ then we apply (6.3.6) to get

$$Q(t) \leq Q(0)\,e^{C_0\int_0^t A(s)\,ds}.$$

Combining these three estimates and recalling that $Q(0) = W_2(f_1(0), f_2(0))^2$ while $Q(t) \geq W_2(f_1(t), f_2(t))^2$, this concludes the proof. □

### 6.3.2 Control of the growth of the support in velocity in 2D

In this Section, $d = 2$. Our goal is to obtain estimates on the growth in time of the support in velocity. This allows us to get bounds for the $L^\infty$ norms of the local densities on some interval of time $[0, T]$. Recall that in the end, they will be used to apply the Wasserstein stability estimates proved in Section 6.3.1.

For $f_\varepsilon$ a solution of (6.1.1), define

$$V_\varepsilon(t) := \sup\left\{|v| \,:\, v \in \mathbb{R}^2,\, \exists x \in \mathbb{T}^2,\, f_\varepsilon(t,x,v) > 0\right\}.$$

The key point of this Section is the following Proposition.

**Proposition 6.3.3.** *Suppose that*

$$\|f_\varepsilon(0)\|_\infty \le C_0, \qquad \int \Big(|v|^2 + U_\varepsilon(0,x)\Big) f_\varepsilon(0,x,v)\, dv\, dx \le C_0.$$

*Assume that* $V_\varepsilon(0) \le C_0/\varepsilon^\gamma$*, for some* $\gamma > 0$*. Let* $T > 0$ *be fixed. For all* $\beta > 2$*, there is* $C_\beta > 0$*, such that we have for all* $\varepsilon \in (0,1)$ *and all* $t \in [0,T]$*,*

$$V_\varepsilon(t) \le C_\beta/\varepsilon^{\max\{\beta,\gamma\}}. \tag{6.3.9}$$

*Therefore, for all* $\beta > 2$*, there is* $C'_\beta > 0$*, such that, for all* $\varepsilon \in (0,1)$ *and all* $t \in [0,T]$*,*

$$\|\rho_\varepsilon\|_\infty \le C'_\beta/\varepsilon^{2\max\{\beta,\gamma\}}. \tag{6.3.10}$$

In order to prove this Proposition, we shall use for convenience the change of variables $(t,x,v) \mapsto (\frac{t}{\varepsilon}, \frac{x}{\varepsilon}, v)$. This leads us to consider, the following Vlasov-Poisson system, for $(x,v) \in \frac{1}{\varepsilon}\mathbb{T}^2 \times \mathbb{R}^2$:

$$\begin{cases} \partial_t g_\varepsilon + v \cdot \nabla_x g_\varepsilon + F_\varepsilon \cdot \nabla_v g_\varepsilon = 0, \\ F_\varepsilon = -\nabla_x \Phi_\varepsilon, \\ -\Delta_x \Phi_\varepsilon = \int_{\mathbb{R}^2} g_\varepsilon\, dv - \int_{\frac{1}{\varepsilon}\mathbb{T}^2 \times \mathbb{R}^2} g_\varepsilon\, dv\, dx, \\ g_\varepsilon|_{t=0} = g_{0,\varepsilon} \ge 0, \\ \int_{\frac{1}{\varepsilon}\mathbb{T}^2 \times \mathbb{R}^2} g_{0,\varepsilon}\, dx\, dv = \frac{1}{\varepsilon^2}. \end{cases} \tag{6.3.11}$$

We shall denote

$$\eta_\varepsilon := \int_{\mathbb{R}^2} g_\varepsilon\, dv$$

and define for all $t \ge 0$,

$$V(t) := \sup\left\{|v|,\, v \in \mathbb{R}^2,\, \exists x \in \frac{1}{\varepsilon}\mathbb{T}^2, g_\varepsilon(t,x,v) > 0\right\} = V_\varepsilon\left(\frac{t}{\varepsilon}\right). \tag{6.3.12}$$

In the following, for brevity, the notation $L^p$, for $p \in [1,+\infty]$, will stand for $L^p\left(\frac{1}{\varepsilon}\mathbb{T}^2 \times \mathbb{R}^2\right)$ or $L^p\left(\frac{1}{\varepsilon}\mathbb{T}^2\right)$, depending on the context.

The main goal is now to prove the following Proposition, from which it is straightforward to deduce Proposition 6.3.3 by applying the result for $t = \frac{T}{\varepsilon}$.

**Proposition 6.3.4.** *Let* $\beta > 0$*. Let* $C_0 > 0$ *such that for all* $\varepsilon \in (0,1)$*,*

$$\|g_\varepsilon(0)\|_{L^\infty} \le C_0, \quad \int_{\frac{1}{\varepsilon}\mathbb{T}^2 \times \mathbb{R}^2} \Big(|v|^2 + \Phi_\varepsilon(0,x)\Big) g_\varepsilon(0)\, dvdx \le \frac{C_0}{\varepsilon^2}. \tag{6.3.13}$$

*Then for all $\alpha \in (0,1)$, there exist $C_\alpha, C'_\alpha > 0$ such that for all $\varepsilon \in (0,1)$, and all $t \geq 0$,*

$$V(t) \leq \left( \frac{C_\alpha}{\varepsilon^{\alpha+1}} t + (1+V(0))^{1-\alpha} \right)^{1/(1-\alpha)} - 1. \tag{6.3.14}$$

We will use the following standard property of conservation of $L^p$ norms and energy for solutions to the Vlasov-Poisson system in the sense of Arsenev:

**Lemma 6.3.5.** *For all $t \geq 0$, we have*

$$\|g_\varepsilon(t)\|_{L^\infty} \leq C_0, \quad \int_{\frac{1}{\varepsilon}\mathbb{T}^2 \times \mathbb{R}^2} \Big( |v|^2 + \Phi_\varepsilon(t,x) \Big) g_\varepsilon(t)\, dvdx \leq \frac{C_0}{\varepsilon^2}. \tag{6.3.15}$$

We also rely on the following Lemma about the Green kernel of the Laplacian on $\frac{1}{\varepsilon}\mathbb{T}^2$ (obtained from standard results on the Green kernel of the Laplacian on $\mathbb{T}^2$, after rescaling). We refer for instance to Caglioti and Marchioro [28].

**Lemma 6.3.6.** *There exists $K_0 \in C^\infty(\mathbb{T}^2; \mathbb{R}^2)$ such that, denoting*

$$K_\varepsilon(x) = \frac{1}{2\pi} \frac{x}{|x|^2} + \varepsilon K_0(\varepsilon x),$$

*we have for all $x \in [-1/\varepsilon, 1/\varepsilon]^2$,*

$$F_\varepsilon(x) = \int_{[-1/\varepsilon, 1/\varepsilon]^2} K_\varepsilon(x-y) \big[\eta_\varepsilon(y) - 1\big]\, dy.$$

The key ingredient is the following Lemma, in which we obtain some appropriate $L^\infty$ bound for the electric field, which allows us to control the growth of the support in velocity.

**Lemma 6.3.7.** *We have the following bounds.*

1. *There is a constant $C_1 > 0$ such that for all $\varepsilon \in (0,1)$, and all $t \geq 0$,*

$$\|\eta_\varepsilon(t)\|_{L^2} \leq \frac{C_1}{\varepsilon}, \quad \|\eta_\varepsilon(t)\|_{L^\infty} \leq C_1 V(t)^2. \tag{6.3.16}$$

2. *There is a constant $C_2 > 0$ such that for all $\varepsilon \in (0,1)$, and all $t \geq 0$,*

$$\|F_\varepsilon(t,\cdot)\|_{L^\infty} \leq C_2 \left( 1 + \frac{1}{\varepsilon} \left[ \log \frac{1}{\varepsilon}(1+V(t)) \right]^{1/2} \right). \tag{6.3.17}$$

3. *For any $\alpha > 0$, there is a constant $C_\alpha$ such that for all $0 \le t' \le t$,*

$$V(t) \le V(t') + \frac{C_\alpha}{\varepsilon^{1+\alpha}} \int_{t'}^{t} (1 + V(s))^\alpha \, ds. \tag{6.3.18}$$

*Proof of Lemma 6.3.7.* In this proof, $C > 0$ will stand for an universal constant that may change from line to line.

1. By the following interpolation argument, we have for all $R > 0$

$$\eta_\varepsilon = \int_{\mathbb{R}^2} g_\varepsilon \, dv = \int_{|v| \le R} g_\varepsilon \, dv + \int_{|v|>R} g_\varepsilon \, dv \le \|g_\varepsilon\|_\infty R^2 + \frac{1}{R^2} \int_{\mathbb{R}^2} |v|^2 g_\varepsilon \, dv,$$

so by optimizing with respect to $R$ we deduce that there is a $C > 0$ such that for all $t \ge 0, x \in \frac{1}{\varepsilon}\mathbb{T}^2$,

$$|\eta_\varepsilon|(t,x) \le C \left( \int g_\varepsilon(t,x,v) |v|^2 \, dv \right)^{1/2}$$

By (6.3.15), we deduce the first estimate of (6.3.16).

For what concerns the $L^\infty$ estimate for $\eta_\varepsilon$, it is a plain consequence of the definition of $V(t)$, which controls the support in velocity.

2. By Lemma 6.3.6, there holds for all $t \ge 0, x \in [-1/\varepsilon, 1/\varepsilon]^2$,

$$|F_\varepsilon|(t,x) \le \frac{1}{2\pi} \int_{[-1/\varepsilon,1/\varepsilon]^2} \frac{1}{|x-x'|} |\eta_\varepsilon(t,x') - 1| \, dx' + \varepsilon \|K_0\|_\infty \|(\eta_\varepsilon - 1)\|_{L^1}.$$

We observe that, since $\|(\eta_\varepsilon - 1)\|_{L^1(\frac{1}{\varepsilon}\mathbb{T}^2)} \le \|\eta_\varepsilon\|_{L^1(\frac{1}{\varepsilon}\mathbb{T}^2)} + \|1\|_{L^1(\frac{1}{\varepsilon}\mathbb{T}^2)} = \frac{2}{\varepsilon^2}$,

$$\varepsilon \|K_0\|_\infty \|(\eta_\varepsilon - 1)\|_{L^1} \le \frac{C}{\varepsilon}.$$

Let $R > 0$ to be fixed later. We have, using the Cauchy-Schwarz inequality,

$$\begin{aligned}
&\int_{[-1/\varepsilon,1/\varepsilon]^2} \frac{1}{|x-x'|} |\eta_\varepsilon(t,x') - 1| \, dx' \\
&\qquad = \int_{|x-x'|<R} \frac{1}{|x-x'|} |\eta_\varepsilon(t,x') - 1| \, dx' + \int_{|x-x'| \ge R} \frac{1}{|x-x'|} |\eta_\varepsilon(t,x') - 1| \, dx' \\
&\qquad \le CR(\|\eta_\varepsilon\|_{L^\infty} + 1) + \left( \|\eta_\varepsilon\|_{L^2} + \frac{1}{\varepsilon} \right) \left( \int_{|x'| \ge R,\, x' \in [-1/\varepsilon,1/\varepsilon]^2} \frac{1}{|x'|^2} \, dx' \right)^{1/2} \\
&\qquad \le CR(\|\eta_\varepsilon\|_{L^\infty} + 1) + \left( \|\eta_\varepsilon\|_{L^2} + \frac{1}{\varepsilon} \right) \left( \log \frac{C}{\varepsilon R} \right)^{1/2}.
\end{aligned}$$

We choose $R = \frac{1}{\|\eta_\varepsilon\|_{L^\infty}+1}$, which yields,

$$\|F_\varepsilon\|_{L^\infty} \leq C\left(1 + \left(\|\eta_\varepsilon\|_{L^2} + \frac{1}{\varepsilon}\right)\left(\log\frac{1}{\varepsilon}(1+\|\eta\|_{L^\infty})\right)^{1/2}\right).$$

Using Point 1., we obtain the claimed estimate.

3. Let $\alpha > 0$. Let $0 \leq t' \leq t$ and let $(x,v)$ such that $g_\varepsilon(t',x,v) \neq 0$. Introduce the characteristics $(X(s,t',x,v), \xi(s,s,x,v))$ satisfying for $s \geq t'$ for the system of ODEs

$$\begin{cases} \dfrac{d}{ds}X(s,t',x,v) = \xi(s,t',x,v), \quad X(t',t',x,v) = x, \\ \dfrac{d}{ds}\xi(s,t',x,v) = F_\varepsilon(s, X(s,t',x,v)), \quad \xi(t',t',x,v) = v. \end{cases} \tag{6.3.19}$$

We have

$$\xi(t,t',x,v) = v + \int_{t'}^t F_\varepsilon(s, X(s,t',x,v))\,ds.$$

Therefore, we have, using Point 2.,

$$\begin{aligned} |\xi|(t,t',x,v) &\leq |v| + \int_{t'}^t |F_\varepsilon|(s, X(s,t',x,v))\,ds \\ &\leq |v| + \frac{C_2}{\varepsilon}\int_{t'}^t \left(1 + \left[\log\frac{1}{\varepsilon}(1+V(s))\right]^{1/2}\right) ds. \end{aligned}$$

Since $g_\varepsilon$ satisfies (6.3.11), it is thus constant along the characteristics (6.3.19), and we have

$$g_\varepsilon(t, X(t,t',x,v), \xi(t,t',x,v)) = g_\varepsilon(t',x,v).$$

By definition of $V(t)$ and $V(t')$, we obtain

$$V(t) \leq V(t') + \frac{C_2}{\varepsilon}\int_{t'}^t \left(1 + \left[\log\frac{1}{\varepsilon}(1+V(s))\right]^{1/2}\right) ds.$$

In order to get the polynomial bound on $V(t)$, let $r_\alpha > 0$ such that for all $x \geq 1$,

$$(\log x)^{1/2} \leq 1 + r_\alpha x^\alpha.$$

We thus get

$$V(t) \leq V(t') + \frac{C_2(2+r_\alpha)}{\varepsilon^{1+\alpha}}\int_{t'}^t (1+V(s))^\alpha\,ds,$$

which proves our claim, taking $C_\alpha := C_2(2+r_\alpha)$.

□

Equipped with this result, we can finally proceed with the proof of Proposition 6.3.4.

*Proof of Proposition 6.3.4.* We begin by observing that dividing by $1/(t-t')$ in both sides of (6.3.18) and letting $t' \to t$ we deduce that

$$\frac{d}{dt}V(t) \leq \frac{C_\alpha}{\varepsilon^{1+\alpha}}(1+V(t))^\alpha. \tag{6.3.20}$$

We will obtain the claimed bound by a comparison principle. To this end, introduce a small parameter $\mu > 0$ and define

$$W_\mu(t) := \left(\frac{2[C_\alpha + \mu](1-\alpha)}{\varepsilon^{\alpha+1}}t + (1+V(0)+\mu)^{1-\alpha}\right)^{\frac{1}{1-\alpha}}.$$

By construction, it satisfies for $t \geq 0$

$$\frac{d}{dt}W_\mu(t) = \frac{C_\alpha + \mu}{\varepsilon^{1+\alpha}}W_\mu(t)^\alpha. \tag{6.3.21}$$

and

$$W_\mu(0) = 1 + V(0) + \mu > 1 + V(0).$$

We claim that $W_\mu(t) \geq 1 + V(t)$ for all $t$. Indeed, let

$$t_0 := \inf\{t > 0 \,:\, W_\mu(t) < 1 + V(t)\},$$

and assume by contradiction that $t_0 < +\infty$. Notice that because $\mu > 0$ we have $t_0 > 0$. Then by continuity at the time $t_0$ we get

$$W_\mu(t_0) = 1 + V(t_0).$$

By (6.3.20) and (6.3.21), we have

$$\begin{aligned}\frac{d}{dt}(V(t) - W_\mu(t) - 1)_{|t=t_0} &\leq \frac{C_\alpha}{\varepsilon^{1+\alpha}}(1+V(t_0))^\alpha - \frac{C_\alpha+\mu}{\varepsilon^{1+\alpha}}W_\mu(t_0)^\alpha \\ &\leq -\frac{\mu}{\varepsilon^{1+\alpha}}W_\mu(t_0)^\alpha < 0.\end{aligned}$$

This is a contradiction with the definition of $t_0$.

Hence we obtained that for all $\mu > 0$ and all $t \geq 0$,

$$1 + V(t) \leq W_\mu(t).$$

By taking the limit $\mu \to 0$, one finally gets for all all $t \geq 0$

$$1 + V(t) \leq \left( \frac{2C_\alpha(1-\alpha)}{\varepsilon^{\alpha+1}} t + (1+V(0))^{1-\alpha} \right)^{\frac{1}{1-\alpha}},$$

which proves the Proposition. □

### 6.3.3 Control of the growth of the support in velocity in 3D using Batt and Rein's estimates

In this Section, we deal with the case $d = 3$. We consider as before, for $f_\varepsilon$ a solution of (6.1.1),

$$V_\varepsilon(t) := \sup \left\{ |v|,\, v \in \mathbb{R}^3,\, \exists x \in \mathbb{T}^3,\, f_\varepsilon(t,x,v) > 0 \right\}.$$

We have in 3D the analogue of the key Proposition 6.3.3 in 2D.

**Proposition 6.3.8.** *Suppose that*

$$\|f_\varepsilon(0)\|_\infty \leq C_0, \qquad \int \Big( |v|^2 + U_\varepsilon(0,x) \Big) f_\varepsilon(0,x,v)\, dv\, dx \leq C_0.$$

*Assume that $V_\varepsilon(0) \leq C_0/\varepsilon^\gamma$, for some $\gamma > 0$. Let $T > 0$ be fixed. There is $C' > 0$, such that we have for all $\varepsilon \in (0,1)$ and all $t \in [0,T]$,*

$$V_\varepsilon(t) \leq \frac{C'}{\varepsilon^{\max\{38/3,\gamma\}}}. \tag{6.3.22}$$

*Then*

$$\|\rho_\varepsilon\|_\infty \leq \frac{C'}{\varepsilon^{\max\{38,3\gamma\}}}. \tag{6.3.23}$$

Note that this result involves exponents which are "more degenerate" than in the 2-D case. We shall prove this result as an application of the estimates obtained by Batt and Rein in [8]. The result of [8] is an adaptation to the case of the torus $\mathbb{T}^3$ of the fundamental contribution of Pfaffelmoser [86] (see also [89, 62]), which allowed to build global classical solutions of the Vlasov-Poisson system in $\mathbb{R}^3 \times \mathbb{R}^3$. In $\mathbb{R}^3 \times \mathbb{R}^3$, it may be possible to get better estimates than (6.3.22) (i.e. with smaller exponents) by using dispersive effects, see [86, 89, 62] and more recently [84].

In order to prove this Proposition, we shall use the change of variables $(t,x,v) \mapsto (\frac{t}{\varepsilon}, \frac{x}{\varepsilon}, v)$. This leads us to consider, the following Vlasov-Poisson system, for $(x,v) \in \frac{1}{\varepsilon}\mathbb{T}^3 \times \mathbb{R}^3$:

$$\begin{cases} \partial_t g_\varepsilon + v \cdot \nabla_x g_\varepsilon + F_\varepsilon \cdot \nabla_v g_\varepsilon = 0, \\ F_\varepsilon = -\nabla_x \Phi_\varepsilon, \\ -\Delta_x \Phi_\varepsilon = \int_{\mathbb{R}^3} g_\varepsilon \, dv - \int_{\frac{1}{\varepsilon}\mathbb{T}^3 \times \mathbb{R}^3} g_\varepsilon \, dv \, dx, \\ g_\varepsilon|_{t=0} = g_{0,\varepsilon} \geq 0, \quad \int_{\frac{1}{\varepsilon}\mathbb{T}^3 \times \mathbb{R}^3} g_{0,\varepsilon} \, dx \, dv = \frac{1}{\varepsilon^3}. \end{cases} \tag{6.3.24}$$

We shall denote as in the 2D case

$$\eta_\varepsilon := \int_{\mathbb{R}^3} g_\varepsilon \, dv,$$

and define for all $t \geq 0$,

$$V(t) := \sup \left\{ |v|, \, v \in \mathbb{R}^3, \, \exists x \in \frac{1}{\varepsilon}\mathbb{T}^3, g_\varepsilon(t,x,v) > 0 \right\}. \tag{6.3.25}$$

As before, in what follows we use the notation $L^p$, for $p \in [1, +\infty]$, will stand for $L^p\left(\frac{1}{\varepsilon}\mathbb{T}^3 \times \mathbb{R}^3\right)$ or $L^p\left(\frac{1}{\varepsilon}\mathbb{T}^3\right)$, depending on the context.

The main goal is now to prove the following Proposition, from which we deduce Proposition 6.3.8 by choosing $t = \frac{T}{\varepsilon}$.

**Proposition 6.3.9.** *Suppose that*

$$\|f_\varepsilon(0)\|_{L^\infty} \leq C_0, \qquad \int_{\frac{1}{\varepsilon}\mathbb{T}^3 \times \mathbb{R}^3} \Big(|v|^2 + \Phi_\varepsilon(0,x)\Big) f_\varepsilon(0,x,v) \, dv \, dx \leq C_0.$$

*Let $\gamma > 0$. Let $C_0 > 0$ such that for all $\varepsilon \in (0,1)$,*

$$V(0) \leq \frac{C_0}{\varepsilon^\gamma}. \tag{6.3.26}$$

*Then there exists $C_1 > 0$ that for all $\varepsilon \in (0,1)$, and all $t \in [0,T]$,*

$$V(t) \leq \max \left\{ \frac{C_0}{\varepsilon^\gamma} + \left[ -\frac{C_1}{\varepsilon^{32/3}} + \sqrt{\frac{C_1^2}{\varepsilon^{64/3}} T^4 + 4 \frac{C_1}{\varepsilon^{32/3+\gamma}}} \right], \frac{C_0}{\varepsilon^\gamma} + T^{-7/2} \right\}. \tag{6.3.27}$$

*Proof of Proposition 6.3.9.* Consider the usual notations for characteristics of (6.3.19) and introduce, as in [8],

$$h_1(t) := \sup\{\|\eta_\varepsilon(s)\|_{L^\infty}, \, 0 \leq s \leq t\} + 1,$$
$$h_2(t) := \sup\{|\xi(s,\tau,x,v) - v|, \, 0 \leq s, \tau \leq t, \, (x,v) \in \frac{1}{\varepsilon}\mathbb{T}^3 \times \mathbb{R}^3\}.$$

We look for a bound on $h_2$, which will imply the control on $V(t)$. To this end, we crucially rely on the key bootstrap result in the paper of Batt and Rein [8], which we recall in the form of a lemma for the reader's convenience.

**Lemma 6.3.10** (Batt, Rein)**.** *Assume that there is $C^* > 0$ and $\beta > 0$ such that*

$$h_2(t) \leq C^* t h_1(t)^\beta$$

*then for some universal constant $C > 0$ (that hereafter may change from line to line),*

$$h_2(t) \leq Ct\Big(C^{*\,4/3} h_1^{2\beta/3}(t) + \frac{1}{\varepsilon^3}\left(h_1^{1/6}(t) + \frac{1}{C^*}\right)\Big), \tag{6.3.28}$$

*if $h_1(t)^{-\beta/2} \leq t$.*

By using [8, Eq. (5), Section 4, p.414], there is $C > 0$ independent of $\varepsilon$ such that for all $\varepsilon > 0$ and $t \geq 0$,

$$h_2(t) \leq C t h_1(t)^{4/9}. \tag{6.3.29}$$

We deduce from Lemma 6.3.10 and (6.3.28) that

$$h_2(t) \leq \frac{C}{\varepsilon^3} t h_1(t)^{8/27}, \quad \text{if } h_1(t)^{-2/9} \leq t.$$

Using (6.3.28) twice, we finally obtain

$$h_2(t) \leq \frac{C}{\varepsilon^4} t h_1(t)^{16/81}, \quad \text{if } h_1(t)^{-4/27} \leq t$$

and since $\frac{32}{243} < \frac{1}{6}$, we get

$$h_2(t) \leq \frac{C}{\varepsilon^{16/3}} t h_1(t)^{1/6}, \quad \text{if } h_1(t)^{-8/81} \leq t,$$

Using the straightforward bound

$$h_1(t) \leq C\left(V(0) + h_2(t)\right)^3,$$

we deduce that for all $\varepsilon \in (0,1)$ and $t > 0$,

$$h_2(t) \leq \frac{C_1}{\varepsilon^{16/3}} t \left(\frac{1}{\varepsilon^\gamma} + h_2(t)\right)^{1/2},$$

if $h_1(t)^{-8/81} \le t$. On the other hand, if $h_1(t)^{-8/81} > t$ then,

$$h_1(t) < \frac{1}{t^{8/81}},$$

and thus, by (6.3.29), we get

$$h_2(t) \le Ct^{-7/2}.$$

We conclude that for all $\varepsilon \in (0,1)$ and $t > 0$,

$$h_2(t) \le C \max\left\{\frac{1}{\varepsilon^{16/3}} t \left(\frac{1}{\varepsilon^\gamma} + h_2(t)\right)^{1/2},\ t^{-7/2}\right\},$$

which yields, taking $t = T$,

$$h_2(T) \le \max\left\{\frac{1}{2}\left[-\frac{C_1}{\varepsilon^{32/3}} + \sqrt{\frac{C_1^2}{\varepsilon^{64/3}}T^4 + 4\frac{C_1}{\varepsilon^{32/3+\gamma}}}\right],\ T^{-7/2}\right\}.$$

Therefore, for $t \in [0,T]$ we get

$$V(t) \le \max\left\{\frac{C_0}{\varepsilon^\gamma} + \left[-\frac{C_1}{\varepsilon^{32/3}} + \sqrt{\frac{C_1^2}{\varepsilon^{64/3}}T^4 + 4\frac{C_1}{\varepsilon^{32/3+\gamma}}}\right],\ \frac{C_0}{\varepsilon^\gamma} + T^{-7/2}\right\},$$

which proves Proposition 6.3.9. □

## 6.4 Proof of Theorem 6.1.3

We prove the main Theorem by a perturbation argument, relying on the Wasserstein stability estimates of Theorem 6.3.1.

Before that, we first state a Lemma about the effect of $x$-dependent translations in the velocity variable on the $W_1$ distance.

**Lemma 6.4.1.** *Let $\mu, \nu$ be probability measures on $\mathcal{P}_1(\mathbb{T}^d \times \mathbb{R}^d)$ and let $\tilde\mu, \tilde\nu$ be probability measures on $\mathcal{P}_1(\mathbb{T}^d \times \mathbb{R}^d)$ defined as follows:*

$$< \tilde\mu, \phi(x,v) > := < \mu, \phi(x, v - C(x)) > \qquad \textit{for all } \phi \in Lip\,(\mathbb{T}^d \times \mathbb{R}^d);$$

$$< \tilde\nu, \phi(x,v) > := < \nu, \phi(x, v - C(x)) > \qquad \textit{for all } \phi \in Lip\,(\mathbb{T}^d \times \mathbb{R}^d).$$

*Then*

$$W_1(\tilde\mu, \tilde\nu) \le (1 + \|D_x C\|_{L^\infty})\, W_1(\mu,\nu) \qquad \textit{where } D_x C := \big(\partial_{x_i} C_j(x)\big)_{0 \le i,j \le d}. \tag{6.4.1}$$

*Proof.* By the Kantorovich duality we have the following expression:

$$W_1(\tilde{\mu},\tilde{\nu}) = \sup_{\|\varphi\|_{Lip}\le 1}\left[\int \varphi(x,v+C(x))d\,\mu(x) - \varphi(x,v+C(x))d\,\nu(x)\right].$$

Let us denote $\psi(x,v)=\varphi(x,v+C(x))$ and computing the gradient of $\psi$ we deduce that

$$\|\psi\|_{Lip} \le \big(1+\|D_xC\|_{L^\infty}\big)\|\varphi\|_{Lip} \le \big(1+\|D_xC\|_{L^\infty}\big)$$

from which we obtain (6.4.1). □

We can now proceed with the proof of Theorem 6.1.3. Let $f_{0,\varepsilon}, g_{0,\varepsilon}, h_{0,\varepsilon}$ satisfy the hypotheses of Theorem 6.1.3. Using the same notations of the statement, we want to show that for some $T>0$,

$$\lim_{\varepsilon\to 0}\sup_{t\in[0,T]} W_1(\widetilde{f}_\varepsilon(t), g(t)) = 0.$$

In analogy with the definition of $\widetilde{f}_\varepsilon$ we define

$$g_\varepsilon(t,x,v) = \int_{\mathcal{M}} \rho^\theta_\varepsilon(t,x)\delta_{v=v^\theta_\varepsilon(t,x)}\,d\mu(\theta)$$

and

$$\widetilde{g}_\varepsilon(t,x,v) = \int_{\mathcal{M}} \rho^\theta_\varepsilon(t,x)\delta_{v=v^\theta_\varepsilon(t,x)+C_\varepsilon(t,x)}\,d\mu(\theta)$$

where $C_\varepsilon(t,x) := -\frac{1}{i}(d_+(t,x)e^{\frac{it}{\sqrt{\varepsilon}}} - d_-(t,x)e^{-\frac{it}{\sqrt{\varepsilon}}})$.

We now prove the following estimate:

$$W_1(\widetilde{f}_\varepsilon, g) \le W_1(\widetilde{f}_\varepsilon, \widetilde{g}_\varepsilon) + W_1(\widetilde{g}_\varepsilon, g). \tag{6.4.2}$$

We first consider the second term in the right hand side. The uniform convergence to 0 follows from Proposition 6.1.1, and the Sobolev embedding theorem. We have indeed for some $T>0$, for all $t\in[0,T]$:

$$\begin{aligned}
W_1(\widetilde{g}_\varepsilon(t), g(t)) &= \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1} \langle \widetilde{g}_\varepsilon - g,\, \varphi\rangle \\
&= \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1}\left\{\int_{\mathbb{T}^d}\int_{\mathcal{M}} (\rho^\theta_\varepsilon(t,x)\varphi(x,v^\theta_\varepsilon(t,x)+C_\varepsilon(t,x)) - \rho^\theta(t,x)\varphi(x,v^\theta(t,x)))\,d\mu(\theta)\,dx\right\} \\
&= \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1}\left\{\int_{\mathbb{T}^d}\int_{\mathcal{M}} \rho^\theta_\varepsilon(t,x)(\varphi(x,v^\theta_\varepsilon(t,x)+C_\varepsilon(t,x)) - \varphi(x,v^\theta(t,x))\,d\mu(\theta)\,dx\right\} \\
&+ \sup_{\|\varphi\|_{\mathrm{Lip}}\le 1}\left\{\int_{\mathbb{T}^d}\int_{\mathcal{M}} (\rho^\theta_\varepsilon(t,x) - \rho^\theta(t,x))\varphi(x,v^\theta(t,x))\,d\mu(\theta)\,dx\right\}.
\end{aligned}$$

Thus, we deduce the estimate

$$W_1(\widetilde{g}_\varepsilon(t), g(t)) \leq \sup_{\|\varphi\|_{\mathrm{Lip}}\leq 1} \sup_{\varepsilon\in(0,1),\,\theta\in\mathcal{M}} \|\rho_\varepsilon^\theta\|_\infty \|\varphi\|_{\mathrm{Lip}} \int_{\mathcal{M}} \|v_\varepsilon^\theta(t,x) + C_\varepsilon(t,x) - v^\theta(t,x)\|_\infty \, d\mu(\theta)$$
$$+ \sup_{\|\varphi\|_{\mathrm{Lip}}\leq 1} \int_{\mathcal{M}} \|\rho_\varepsilon^\theta - \rho_\theta\|_\infty \, d\mu(\theta) \|\varphi\|_{\mathrm{Lip}} \left(1/2 + \sup_{\theta\in\mathcal{M}} \|v^\theta(t,x)\|_\infty\right)$$
$$+ \sup_{\|\varphi\|_{\mathrm{Lip}}\leq 1} \left\{ \int_{\mathbb{T}^d\times\mathbb{R}} \int_{\mathcal{M}} (\rho_\varepsilon^\theta(t,x) - \rho^\theta(t,x))\varphi(0,0) \, d\mu(\theta)\, dx \right\}.$$

We notice that the last term is equal to 0 since for all $t \geq 0$,

$$\int_{\mathbb{T}^d} \int_{\mathcal{M}} \rho_\varepsilon^\theta(t,x)\, d\mu(\theta)\, dx = \int_{\mathbb{T}^d} \int_{\mathcal{M}} \rho_\varepsilon^\theta(t,x)\, d\mu(\theta)\, dx = 1,$$

by conservation of the total mass. Considering the supremum in time, we see that the other two terms converge to 0, using (6.1.6), so that we get

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} W_1(\widetilde{g}_\varepsilon, g) = 0.$$

We thus focus on the first term of the right hand side of (6.4.2). First we use Lemma 6.4.1 to see that

$$W_1(\widetilde{f}_\varepsilon, \widetilde{g}_\varepsilon) \leq (1 + \|D_x C_\varepsilon(t,x)\|_{L^\infty})\, W_1(f_\varepsilon, g_\varepsilon).$$

Observe from the definition of the corrector $C_\varepsilon$, there is $C_T > 0$ independent from $\varepsilon$ such that for all $t \in [0,T]$,

$$\|D_x C_\varepsilon(t,\cdot)\|_{L^\infty} \leq C_T.$$

We therefore have to study $W_1(f_\varepsilon, g_\varepsilon)$. We first use the rough bound

$$W_1(f_\varepsilon, g_\varepsilon) \leq W_2(f_\varepsilon, g_\varepsilon),$$

then use Theorem 6.3.1 to get the estimate

$$\sup_{t\in[0,T]} W_2(f_\varepsilon, g_\varepsilon) \leq 16\, d\, \exp\left\{ \log\left(\frac{\varphi(\varepsilon)}{16\, d}\right) \exp\left[ C_0 T \frac{1}{\varepsilon^2}\left(1 + \|\rho_{f_\varepsilon}\|_{L^\infty([0,T];L^\infty_x)} + \|\rho_{g_\varepsilon}\|_{L^\infty([0,T];L^\infty_x)}\right)\right]\right\},$$

where

$$\rho_{f_\varepsilon} := \int_{\mathbb{R}^d} f_\varepsilon \, dv, \quad \rho_{g_\varepsilon} = \int_{\mathcal{M}} \rho_\varepsilon^\theta \, d\mu(\theta).$$

Recalling (6.1.6), we have for some $C > 0$ independent of $\varepsilon$ that

$$\|\rho_{g_\varepsilon}\|_{L^\infty([0,T];L^\infty_x)} \leq C.$$

For what concerns $\rho_{f_\varepsilon}$ we apply

- in two dimensions, (6.3.10) in Proposition 6.3.3 to infer that for all $\beta > 2$, there is some $C_\beta > 0$ independent of $\varepsilon$ such that

$$\|\rho_\varepsilon\|_\infty \leq \frac{C_\beta}{\varepsilon^{2\max\{\beta,\gamma\}}};$$

- in three dimensions, (6.3.23) in Proposition 6.3.8 to infer that for some $C > 0$ independent of $\varepsilon$ such that

$$\|\rho_\varepsilon\|_\infty \leq \frac{C}{\varepsilon^{\max\{38,3\gamma\}}}.$$

We deduce that choosing

- in two dimensions, $\varphi(\varepsilon) = \exp\left[\exp\left(-\frac{K}{\varepsilon^{2(1+\max(\beta,\gamma))}}\right)\right]$, for some constant $K > 0$;
- in three dimensions, $\varphi(\varepsilon) = \exp\left[\exp\left(-\frac{K}{\varepsilon^{2+\max(38,3\gamma))}}\right)\right]$, for some constant $K > 0$,

up to take a smaller time interval of convergence $[0, T]$,

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} W_2(f_\varepsilon, g_\varepsilon) = 0.$$

We conclude that

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} W_1(\tilde{f}_\varepsilon, g) = 0$$

and the proof of Theorem 6.1.3 is complete.

# Appendix A

# Jacobi fields

Given an $n$-dimensional $C^\infty$ differentiable manifolds $\mathcal{M}$, for each $x \in \mathcal{M}$ we denote by $T_x\mathcal{M}$ the tangent space to $\mathcal{M}$ at $x$, and by $T\mathcal{M} := \cup_{x\in\mathcal{M}}(\{x\} \times T_x\mathcal{M})$ the whole tangent bundle of $\mathcal{M}$. On each tangent space $T_x\mathcal{M}$, we assume that is given a symmetric positive definite quadratic form $g_x : T_x\mathcal{M} \times T_x\mathcal{M} \to \mathbb{R}$ which depends smoothly on $x$; $g = (g_x)_{x\in\mathcal{M}}$ is called a *Riemannian metric*, and $(\mathcal{M}, g)$ is a *Riemannian manifold*.

A Riemannian metric defines a scalar product and a norm on each tangent space: for each $v, w \in T_x\mathcal{M}$

$$\langle v, w\rangle_x := g_x(v, w), \qquad |v|_x := \sqrt{g_x(v, v)}.$$

Let $U$ be an open subset of $\mathbb{R}^n$ and $\Phi : U \to \Phi(U) = V \subset \mathcal{M}$ a chart. Given $x = \Phi(x^1, \dots, x^n) \in V$, the vectors $\frac{\partial}{\partial x^i} := \frac{\partial \Phi}{\partial x_i}(x^1, \dots, x^n)$, $i = 1, \dots, n$, constitute a basis of $T_x\mathcal{M}$: any $v \in T_x\mathcal{M}$ can be written as $v = \sum_{i=1}^n v^i \frac{\partial}{\partial x^i}$. We can use this chart to write our metric $g$ in coordinates inside $V$:

$$g_x(v, v) = \sum_{i,j=1}^n g_x\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right) v^i v^j = \sum_{i,j=1}^n g_{ij}(x) v^i v^j,$$

where by definition $g_{ij}(x) := g_x\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right)$. We also denote by $g^{ij}$ the coordinates of the inverse of $g$: $g^{ij} = (g_{ij})^{-1}$; more precisely, $\sum_j g^{ij} g_{jk} = \delta^i_k$, where $\delta^i_k$ denotes Kronecker's delta:

$$\delta^i_k = \begin{cases} 1 & \text{if } i = k, \\ 0 & \text{if } i \neq k. \end{cases}$$

In the sequel we will use these coordinates to perform many computations. Einstein's convention of summation over repeated indices will be used systematically: $a_k b^k = \sum_k a_k b^k$, $g_{ij} v^i v^j = \sum_{i,j} g_{ij} v^i v^j$, etc.

On a Riemannian manifold, constant-speed minimizing geodesics satisfy a second order differential equation:

$$\ddot{\gamma}^k + \Gamma_{ij}^k \dot{\gamma}^i \dot{\gamma}^j = 0, \tag{A.0.1}$$

where $\Gamma_{ij}^k$ are the Christoffel symbols defined by

$$\Gamma_{ij}^k = \frac{1}{2} g^{k\ell} \left( \frac{\partial g_{j\ell}}{\partial x^i} + \frac{\partial g_{i\ell}}{\partial x^j} - \frac{\partial g_{ij}}{\partial x^\ell} \right).$$

Let us consider a family $(\gamma_\theta)_{-\varepsilon \le \theta \le \varepsilon}$ of constant-speed geodesics $\gamma_\theta : [0,1] \to \mathcal{M}$. Then, for each $t \in [0,1]$, we can consider the vector field

$$J(t) := \left. \frac{\partial}{\partial \theta} \right|_{\theta=0} \gamma_\theta(t) \in T_{\gamma(t)}\mathcal{M}.$$

The vector field $J$ is called a *Jacobi field* along $\gamma = \gamma_0$. By differentiating the geodesic equations with respect to $\theta$, we get a second order differential equation for $J$:

$$\frac{\partial}{\partial \theta} \left( \ddot{\gamma}_\theta^k + \Gamma_{ij}^k(\gamma_\theta) \dot{\gamma}_\theta^i \dot{\gamma}_\theta^j \right) = 0$$

gives

$$\ddot{J}^k + \frac{\partial \Gamma_{ij}^k}{\partial x^\ell} J^\ell \dot{\gamma}^i \dot{\gamma}^j + 2\Gamma_{ij}^k \dot{J}^i \dot{\gamma}^j = 0.$$

This complicated equation takes a nicer form if we choose time-dependent coordinates determined by a moving orthonormal basis $\{e_1(t), \dots, e_n(t)\}$ of $T_{\gamma(t)}\mathcal{M}$, such that

$$\dot{e}_\ell^k(t) + \Gamma_{ij}^k(\gamma(t)) e_\ell^i(t) \dot{\gamma}^j(t) = 0$$

(in this case, we say that the basis is *parallel transported* along $\gamma$). With this choice of the basis, defining $J^i(t) := \langle J(t), e_i(t) \rangle_{\gamma(t)}$ we get

$$\ddot{J}_i(t) + R_i^j(t) J_j(t) = 0.$$

For our purposes it suffices to know that $R_i^j$ is a symmetric matrix; in fact one can show that $R_i^j(t) = \langle \mathrm{Riem}(\dot{\gamma}, e_i) \cdot \dot{\gamma}, e_j \rangle$, where Riem denotes the *Riemann tensor* of $(\mathcal{M}, g)$.

We now write the *Jacobi equation* in matrix form: let $\boldsymbol{J}(t) = (J_1(t), \dots, J_n(t))$ be a matrix of Jacobi fields, and define $J_{ij}(t) := \langle J_i(t), e_j(t) \rangle_{\gamma(t)}$, with $\{e_1(t), \dots, e_n(t)\}$ parallel transported as before. Then

$$\ddot{\boldsymbol{J}}(t) + R(t)\boldsymbol{J}(t) = 0,$$

where $R(t)$ is a symmetrix matrix involving derivatives of the metric $g_{ij}(\gamma(t))$ up to the second order, and such that (up to identification) $R(t)\dot{\gamma}(t) = 0$.

Now, fix a point $x_0 \in \mathcal{M}$, consider $v, w \in \mathbb{S}^{d-1} \subset T_{x_0}\mathcal{M}$ with $v \perp w$, and consider the family of geodesics

$$\gamma_\theta(t) := \exp_{x_0}\big(t(v + \theta w)\big).$$

Then

$$J(t) := \frac{\partial}{\partial\theta}\bigg|_{\theta=0} \gamma_\theta(t) = d_{tv} \exp_{x_0}[tw],$$

and $J(t)$ solves (in a suitable system of coordinates)

$$\begin{cases} \ddot{J}_i(t) + R_i^j(t) J_j(t) = 0, \\ J(0) = 0, \\ \dot{J}(0) = w. \end{cases}$$

This fact shows the relation between the differential of the exponential map and the Jacobi fields, and it is at the basis of the proof of Rauch Comparison Theorem [73, Theorem 11.9] which was used in Section 3.1 to prove Corollary 3.1.6. Indeed, the rough idea is that if the sectional curvature of $\mathcal{M}$ is bounded by $-K$ ($K \geq 0$) then (since $w$ is orthogonal to the velocity $v$ of the geodesic $\gamma_0(t) = \exp_{x_0}(tv)$) the solutions of the above system is controlled by the solutions of

$$\begin{cases} \ddot{J}_i^K(t) + K J_i^K(t) = 0, \\ J^K(0) = 0, \\ \dot{J}^K(0) = w, \end{cases}$$

which is simply $w \frac{\sinh(Kt)}{K}$. More precisely, since $|w| = 1$, one gets (see for instance [73, Theorem 11.9])

$$|J(t)| \leq |J^K(t)| = \frac{\sinh(Kt)}{K},$$

which implies the control $A_{x_0}(r) \leq \frac{\sinh(Kr)}{K}$, where $A_{x_0}$ is the quantity defined (3.1.5).

# Bibliography

[1] L. Ambrosio. Transport equation and Cauchy problem for $BV$ vector fields. *Invent. Math.*, 158(2):227–260, 2004.

[2] L. Ambrosio, M. Colombo, and A. Figalli. On the Lagrangian structure of transport equations: the Vlasov-Poisson system. Preprint, 2014.

[3] L. Ambrosio and N. Gigli. *A user's guide to optimal transport.* Modelling and optimisation of flows on networks, 1-155, Lecture Notes in Math., 2062, Springer, Heidelberg, 2013.

[4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures. Second edition.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008.

[5] V. I. Arnold. On conditions for non-linear stability of plane stationary curvilinear flows of an ideal fluid. *Dokl. Akad. Nauk. SSSR* 162:975–978, 1965.

[6] V. I. Arnold. Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier (Grenoble)*, 16:319–361, 1966.

[7] A. A. Arsenev. Existence in the large of a weak solution of Vlasov's system of equations. *Ž. Vyčisl. Mat. i Mat. Fiz.*, 15:136–147, 276, 1975.

[8] J. Batt and G. Rein. Global classical solutions of the periodic Vlasov-Poisson system in three dimensions. *C. R. Acad. Sci. Paris Sér. I Math.*, 313(6):411–416, 1991.

[9] V. Barbu. *Nonlinear differential equations of monotone type in Banach spaces.* Springer Monographs in Mathematics, Springer, New York, 2010.

[10] J. Binney and S. Tremaine. *Galactic Dynamics.* Princeton: Princeton University Press, 1987.

[11] A. Bohun, F. Bouchut, and G. Crippa. Lagrangian solutions to the Vlasov-Poisson equation with $L^1$ density. Preprint, 2014.

[12] C. Boldrighini. *Macroscopic limits of microscopic systems. Rend. Mat. Appl. (7)*, 16(1):1–107, 1996.

[13] M. Bonforte, J.L. Vazquez, *Positivity, local smoothing, and Harnack inequalities for very fast diffusion equations*, Advances in Math. 223 (2010), 529578.

[14] G. Bouchitté, C. Jimenez, and M. Rajesh. Asymptotique d'un problème de positionnement optimal. *C. R. Math. Acad. Sci. Paris*, 335(10):853–858, 2002.

[15] G. Bouchitté, C. Jimenez, and R. Mahadevan. Asymptotic analysis of a class of optimal location problems. *J. Math. Pures Appl. (9)*, 95(4):382–419, 2011.

[16] F. Bouchut, F. Golse, and M. Pulvirenti. *Kinetic Equations and Asymptotic Theory.* L. Desvillettes et B. Perthame eds. Series in Applied Mathematics, 4. Gauthier-Villars, Editions Scientifiques et Mèdicales Elsevier, Paris; North-Holland, Amsterdam, 2000

[17] A. Brancolini, G. Buttazzo, F. Santambrogio, and E. Stepanov. Long-term planning versus short-term planning in the asymptotical location problem. *ESAIM Control Optim. Calc. Var.*, 15(3):509–524, 2009.

[18] W. Braun and K. Hepp. The Vlasov dynamics and its fluctuations in the $1/N$ limit of interacting classical particles. *Commun. Math. Phys.* 56:125–146, 1977.

[19] J. Bedrossian and N. Masmoudi. Inviscid damping and the asymptotic stability of planar shear flows in the 2D Euler equations. Preprint, 2014.

[20] J. Bedrossian, N. Masmoudi, and C. Mouhot. Landau damping: paraproducts and Gevrey regularity. Preprint, 2015.

[21] F. Bouchut. Global weak solution of the Vlasov-Poisson system for small electrons mass. *Comm. Partial Differential Equations*, 16:1337–1365, 1991.

[22] Y. Brenier. A Vlasov-Poisson type formulation of the Euler equations for perfect incompressible fluids. *Rapport de recherche INRIA*, 1989.

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

[24] Y. Brenier. Convergence of the Vlasov-Poisson system to the incompressible Euler equations. *Comm. Partial Differential Equations*, 25(3-4):737–754, 2000.

[25] J. Bucklew and G. Wise. Multidimensional Asymptotic Quantization Theory with $r$-th Power Distortion Measures. *IEEE Inform. Theory* 28 (2):239–247, 1982.

[26] E. Caglioti, F. Golse, and M. Iacobelli. A gradient flow approach to quantization of measures. *Math. Models Methods Appl. Sci.*, to appear.

[27] E. Caglioti and C. Maffei. Time asymptotics for solutions of Vlasov-Poisson equation in a circle. *J. Stat. Phys.*, 92:301–323,1998.

[28] E. Caglioti and C. Marchioro. Bounds on the growth of the velocity support for the solutions of the Vlasov-Poisson equation in a torus. *J. Statist. Phys.*, 100(3-4):659–677, 2000.

[29] C. Cercignani, R. Illner, and M. Pulvirenti. *The mathematical theory of dilute gases.* Applied Mathematical Sciences, 106. Springer-Verlag, New York, 1994.

[30] F. Chen. *Introduction to plasma physics and controlled fusion.* Plenum Press, New York, 1984.

[31] D. L. Cohn. *Measure theory.* Birkhäuser, Boston, Mass., 1980.

[32] E. DiBenedetto. *Degenerate parabolic equations.* Universitext. Springer-Verlag, New York, 1993.

[33] R. J. DiPerna and P.-L. Lions. Solutions globales d'équations du type Vlasov-Poisson. (French) [Global solutions of Vlasov-Poisson type equations] *C. R. Acad. Sci. Paris Sér. I Math.*, 307:655-658, 1988.

[34] R. J. DiPerna and P.-L. Lions. Global weak solutions of kinetic equations. *Rend. Sem. Mat. Univ. Politec. Torino*, 46:259–288, 1988.

[35] R. J. DiPerna & P.-L. Lions. Global weak solutions of Vlasov-Maxwell systems. *Comm. Pure Appl. Math.*, 42:729–757, 1989.

[36] R. L. Dobrushin. Vlasov Equations Funktsional. *Anal.i Prilozhen.* 13: 48–58, 1979. ?

[37] L. C. Evans. *Partial differential equations.* Graduate studies in mathematics, 1998.

[38] L. C. Evans and R. Gariepy. *Measure Theory and Fine Properties of Functions.* Studies in Advanced Mathematics, 1992.

[39] A. Figalli. Existence and uniqueness of martingale solutions for SDEs with rough or degenerate coefficients. *J. Funct. Anal.*, 254(1):109–153, 2008.

[40] A. F. Filippov. *Differential equations with discontinuous righthand sides*, volume 18 of *Mathematics and its Applications (Soviet Series).* Kluwer Academic Publishers Group, Dordrecht, 1988. Translated from the Russian.

[41] A. Friedman. *Partial differential equations of parabolic type.* Prentice-Hall, Inc., Englewood Cliffs, N.J. 1964

[42] A. M. Fridman, V. L. Polyachenko. *Physics of Gravitating Systems I.* New York: Springer-Verlag 1984

[43] G. Gallavotti. *Statistical Mechanics: a Short Treatise.* Springer, Berlin-Heidelberg, 1999.

[44] R.T. Glassey. *The Cauchy problem in kinetic theory.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996

[45] F. Golse. On the Dynamics of Large Particle Systems in the Mean Field Limit. Preprint, available at `http://arxiv.org/pdf/1301.5494.pdf`

[46] F. Golse and L. Saint-Raymond. The Vlasov-Poisson system with strong magnetic field in quasineutral regime. *Math. Models Methods Appl. Sci.*, 13(5):661–714, 2003.

[47] S. Graf, H. Luschgy. *Foundations of Quantization for Probability Distributions.* Lecture Notes in Math. 1730, Springer-Verlag, Berlin Heidelberg, 2000.

[48] E. Grenier. Defect measures of the Vlasov-Poisson system in the quasineutral regime. *Comm. Partial Differential Equations*, 20(7-8):1189–1215, 1995.

[49] E. Grenier. Oscillations in quasineutral plasmas. *Comm. Partial Differential Equations*, 21(3-4):363–394, 1996.

[50] E. Grenier. Limite quasineutre en dimension 1. In *Journées "Équations aux Dérivées Partielles" (Saint-Jean-de-Monts, 1999)*, pages Exp. No. II, 8. Univ. Nantes, Nantes, 1999.

[51] M. K.-H. Kiessling. The "Jeans swindle": a true story—mathematically speaking. *Adv. in Appl. Math.* 31(1):132–149, 2003.

[52] B. Kloeckner. Approximation by finitely supported measures. *ESAIM Control Optim. Calc. Var.*, 18(2):343–359, 2012.

[53] N. V. Krylov. *Lectures on elliptic and parabolic equations in Hölder spaces*. Graduate Studies in Mathematics, 12. American Mathematical Society, Providence, RI, 1996.

[54] D. Han-Kwan. The three-dimensional finite Larmor radius approximation. *Asymptot. Anal.*, 66(1):9–33, 2010.

[55] D. Han-Kwan. Quasineutral limit of the Vlasov-Poisson system with massless electrons. *Comm. Partial Differential Equations*, 36(8):1385–1425, 2011.

[56] D. Han-Kwan. Contribution à l'étude mathématique des plasmas fortement magnétisés. Phd thesis, 2011.

[57] D. Han-Kwan. From Vlasov-Poisson to Korteweg–de Vries and Zakharov-Kuznetsov. *Comm. Math. Phys.*, 324(3):961–993, 2013.

[58] D. Han-Kwan and M. Hauray. Stability issues in the quasineutral limit of the one-dimensional Vlasov-Poisson equation.

*Comm. Math. Phys. (online first), 110.1007/s00220-014-2217-4*, 2014.

[59] D. Han-Kwan and M. Iacobelli. The quasineutral limit of the Vlasov-Poisson equation in Wasserstein metric. Preprint, 2014.

[60] D. Han-Kwan and M. Iacobelli. Quasineutral limit for Vlasov-Poisson via Wasserstein stability estimates in higher dimension. Preprint, 2015.

[61] M. Hauray. Mean field limit for the one dimensional Vlasov-Poisson equation. *Séminaire Laurent Schwartz - EDP et applications*, 21:16 p., 2012-2013.

[62] E. Horst. On the asymptotic growth of the solutions of the Vlasov-Poisson system. *Math. Methods Appl. Sci.*, 16(2):75–86, 1993.

[63] E. Horst and R. Hunze. Weak solutions of the initial value problem for the unmodified nonlinear Vlasov equation. *Math. Methods Appl. Sci.* 6:262–279, 1984.

[64] H. J. Hwang and J. J. L. Velázquez. On the existence of exponentially decreasing solutions of the nonlinear Landau damping problem. *Indiana Univ. Math. J.*, 58(6):2623–2660, 2009.

[65] M. Iacobelli. Asymptotic quantization for probability measures on Riemannian manifolds. *ESAIM Control Optim. Calc. Var.*, to appear.

[66] S. V. Iordanskii. The Cauchy problem for the kinetic equation of plasma. *Trudy Mat. Inst. Steklov.*, 60:181–194, 1961.

[67] P.-E. Jabin. A review of the mean field limits for Vlasov equations. *Kinet. Relat. Models* 7(4): 661–711, 2014.

[68] J. Jeans. On the theory of star-streaming and the structure of the universe. *Mon. Not. R. Astron. Soc.* 76:70–84, 1915, and 76:552–567, 1916.

[69] A. N. Kaufman and R. L. Dewar. Canonical derivation of the Vlasov-Coulomb noncanonical Poisson structure. Fluids and plasmas: geometry and dynamics (Boulder, Colo., 1983), 51–54, *Contemp. Math.*, 28, Amer. Math. Soc., Providence, RI, 1984.

[70] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and quasilinear equations of parabolic type.* (Russian) Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23 American Mathematical Society, Providence, R.I. 1968

[71] L. D. Landau. On the vibrations of the electronic plasma. (Russian) *Akad. Nauk SSSR. Zhurnal Eksper. Teoret. Fiz.*, 16:574–586, 1946.

[72] L. D. Landau. On the vibrations of the electronic plasma. *Acad. Sci. USSR. J. Phys.*, 10:25–34, 1946.

[73] J. M. Lee. *Riemannian manifolds. An introduction to curvature.* Graduate Texts in Mathematics, 176. Springer-Verlag, New York, 1997.

[74] M. Lieberman and A. Lichtenberg. *Principles of plasma discharge and materials processing.* New-York, Wiley, 1994.

[75] P.-L. Lions and B. Perthame. Propagation of moments and regularity for the 3-dimensional Vlasov-Poisson system. *Invent. Math.*, 105: 415–430, 1991.

[76] G. Loeper. Uniqueness of the solution to the Vlasov-Poisson system with bounded density. *J. Math. Pures Appl. (9)*, 86(1):68–79, 2006.

[77] C. Marchioro and M. Pulvirenti. *Mathematical theory of incompressible fluids.* Springer-Verlag, New York, 1994.

[78] N. Masmoudi. From Vlasov-Poisson system to the incompressible Euler system. *Comm. Partial Differential Equations*, 26(9-10):1913–1928, 2001.

[79] P. J. Morrison. *Poisson Brackets for Fluids and Plasmas*, in Mathematical Methods in Hydrodynamics and Integrability in Dynamical Systems, eds. M. Tabor and Y. Treve, American Institute of Physics Conference Proceedings No. 88 (American Institute of Physics, New York, 1982) pp. 13–46.

[80] S. Mosconi and P. Tilli. Γ-Convergence for the Irrigation Problem. *J. of Conv. Anal.*, 12(1):145–158, 2005.

[81] C. Mouhot and C. Villan. On Landau damping. *Acta Math.* 207(1):29–201, 2011.

[82] H. Neunzert. *An introduction to the nonlinear Boltzmann-Vlasov equation.* In Kinetic theories and the Boltzmann equation, C. Cercignani, Ed., vol. 1048 of Lecture Notes in Math., Springer, Berlin, Heidelberg, 1984, pp. 60–110.

[83] P. J. Olver. *Applications of Lie groups to differential equations.* Second edition. Graduate Texts in Mathematics, 107. Springer-Verlag, New York, 1993.

[84] C. Pallard. Large velocities in a collisionless plasma. *J. Differential Equations*, 252(3):2864–2876, 2012.

[85] O. Penrose. Electrostatic instability of a uniform non-Maxwellian plasma. *Phys. Fluids*, 3:258–265, 1960.

[86] K. Pfaffelmoser. Global classical solutions of the Vlasov-Poisson system in three dimensions for general initial data. *J. Differential Equations*, 95(2):281–303, 1992.

[87] J. G. Ratcliffe. *Foundations of hyperbolic manifolds.* Second edition. Graduate Texts in Mathematics, 149. Springer, New York, 2006.

[88] G. Rein. *Collisionless kinetic equations from astrophysics: the Vlasov- Poisson system.* Handbook of differential equations: evolutionary equations. Vol. III, 383–476, Handb. Differ. Equ., Elsevier/North-Holland, Amsterdam, 2007.

[89] J. Schaeffer. Global existence of smooth solutions to the Vlasov-Poisson system in three dimensions. *Comm. Partial Differential Equations*, 16(8-9):1313–1335, 1991.

[90] S. Ukai and T. Okabe. On classical solutions in the large in time of two-dimensional Vlasov's equation. *Osaka J. Math.*, 15:245–261, 1978.

[91] J. L. Vazquez, Failure of the strong maximum principle in nonlinear diffusion. Existence of needles. *Comm. Partial Differential Equations*, 30(7-9):1263–1303, 2005.

[92] J. L. Vazquez. *Smoothing and Decay Estimates for Nonlinear Diffusion Equations.* Oxford Lect; Ser. in Math. and its Appl. 33, Oxford Univ. Press, New York, 2006.

[93] J. L. Vazquez. *The porous medium equation. Mathematical theory.* Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, Oxford, 2007.

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

[95] C. Villani. *Optimal Transport. Old and new.* Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 338. Springer-Verlag, Berlin, 2009.

[96] A. A. Vlasov. Zh. Eksper. *Teor. Fiz.* 8:291, 1938.

[97] A. A. Vlasov. Vlasov equation and plasma dispersion relation. *J. Phys. (U.S.S.R.)* 9:25, 1945.

[98] H. Ye and P. J. Morrison. Action principles for the Vlasov equation. *Phys. Fluids B*, 4(4):771–777, 1992.

[99] Y. Zheng and A. Majda. Existence of global weak solutions to one-component Vlasov-Poisson and Fokker-Planck-Poisson systems in one space dimension with measures as initial data. *Comm. Pure Appl. Math.*, 47(10):1365–1401, 1994.