



**HAL**  
open science

# Quantum Optimal Transport: Quantum Couplings and Many-Body Problems

François Golse

► **To cite this version:**

François Golse. Quantum Optimal Transport: Quantum Couplings and Many-Body Problems. 2023. hal-04184518

**HAL Id: hal-04184518**

**<https://polytechnique.hal.science/hal-04184518v1>**

Preprint submitted on 22 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# QUANTUM OPTIMAL TRANSPORT: QUANTUM COUPLINGS AND MANY-BODY PROBLEMS

FRANÇOIS GOLSE

ABSTRACT. This text is a set of lecture notes for a 4.5-hour course given at the Erdős Center (Rényi Institute, Budapest) during the Summer School “Optimal Transport on Quantum Structures” (September 19th-23rd, 2023). Lecture I introduces the quantum analogue of the Wasserstein distance of exponent 2 defined in [F. Golse, C. Mouhot, T. Paul: *Comm. Math. Phys.* **343** (2016), 165–205], and in [F. Golse, T. Paul: *Arch. Ration. Mech. Anal.* **223** (2017) 57–94]. Lecture II discusses various applications of this quantum analogue of the Wasserstein distance of exponent 2, while Lecture III discusses several of its most important properties, such as the triangle inequality, and the Kantorovich duality in the quantum setting, together with some of their implications.

## INTRODUCTION

Optimal transport has become a thriving field of mathematical research in the last decade of the 20th century. Nowadays, it is used in a variety of subjects which the founder of this theory (G. Monge, 1781) could obviously not have foreseen: statistics, machine learning, probability theory, fluid mechanics, besides more classical applications such as the calculus of variations, partial differential equations, geometry. In his book [63], C. Villani speaks of a “revival” of optimal transport following Y. Brenier’s remarkable paper [15].

Among the many applications of optimal transport is a quite fascinating observation by Dobrushin [30], who realized that one special kind of approximation used in classical nonequilibrium statistical mechanics, namely the time-dependent mean-field limit, can be proved rigorously by using a notion of distance on the set of probability measures which is defined in terms of optimal transport.

The same mean-field limit is routinely used in quantum dynamics, but, until relatively recently, the methods of proof used in the rigorous justification of this limit were radically different in the quantum and in the classical settings. This state of affairs was slightly disturbing, since one would expect that it should be possible to take the classical limit of quantum mechanics and the mean-field limit of both classical and quantum dynamics in any order — in other words, one expects that both limits should be represented by a commutative diagram.

There have been some attempts in that direction at the beginning of the 2000s [43, 51], and, slightly later, this suggested the natural idea of “lifting” Dobrushin’s

---

*Date:* August 22, 2023.

*2020 Mathematics Subject Classification.* 49Q22, 49N15, 81S30, 35Q41, 81V70, 82C10, 35Q55, 81Q20, 35Q83.

*Key words and phrases.* Wasserstein distance, Kantorovich duality, Quantum dynamics, Mean-field limit, Classical limit of quantum mechanics, Observation inequality, Time-splitting methods for quantum dynamics.

approach [30] to the quantum setting. Since Dobrushin used an optimal transport metric to compare  $N$ -particle densities and their mean-field limits, it became a very natural motivation for defining an analogous optimal transport “metric” for the purpose of comparing quantum states. Ideally, this “metric” should converge to the metric used by Dobrushin in the classical limit.

I started working on this problem with C. Mouhot and T. Paul, after a visit to C. Mouhot in Cambridge in September 2014. At this point, we realized that mean-field quantum dynamics satisfies an analogous continuous dependence on the initial data defined in terms of a quantum optimal transport problem as in Dobrushin’s analysis. Our initial contribution [37] was followed by a series of works involving other collaborations [38, 19, 20, 36, 39, 40, 41], and exploring both our quantum analogue of optimal transport and its applications to various problems in quantum dynamics. It is precisely this approach to quantum optimal transport which is presented in these lectures.

But there are other, possibly (most likely?) unrelated approaches to quantum optimal transport.

One such approach, due to E. Carlen and J. Maas is based on the Benamou-Brenier formula (Theorem 8.1 in [63]): see [22]. Another approach, based on the notion of quantum channel and closer in spirit to the one presented in this course, yet different, is due to G. De Palma and D. Trevisan [26] — see also Dario Trevisan’s course [62] in this volume. Some of these different approaches to a theory of quantum optimal transport (see in particular Eric Carlen’s lectures [21]) will be presented in this school. Still another approach to the problem of quantum optimal transport is the very early reference [65] (see also the beautiful book [11] by I. Bengtsson and K. Życzkowski, and especially section 7.7 there).

Since this text is a set of lecture notes, several quizzes/exercises are proposed to the reader. Some of these exercises review classical material which the reader is expected to master before going further; some others discuss natural extensions of the material presented in this course. In any case, solving these exercises is strongly recommended in order to gain familiarity with the notions presented in these notes.

I am indebted to several colleagues for the mathematics discussed in these lectures, in the first place to C. Mouhot, T. Paul, E. Caglioti, and S. Jin, with whom I had the pleasure to work on various problems related to quantum optimal transport. The observation inequality for the Schrödinger equation discussed in Lecture II comes from questions posed by C. Bardos. I owe my first acquaintance with Dobrushin’s remarkable paper [30] to M. Pulvirenti, who gave a most lucid account of it in a lecture at École normale supérieure in 1997. I also benefited from numerous discussions on optimal transport with Y. Brenier. Most of these lectures are based on the numerous analogies between classical and quantum optimal transport. However, some very fundamental properties of classical optimal transport may fail to have quantum analogues. I am very grateful to D. Serre, who kindly showed me an example where the quantum and the classical theory significantly differ (see Quiz 31 in Lecture III).

Finally, I wish to express my gratitude to the organizers of this 2022 Summer School on Optimal Transport on Quantum Structures at the Erdős Center, J. Maas, S. Rademacher, T. Titkos and D. Virosztek for their kind invitation, and more generally to the Rényi Institute for its most enjoyable hospitality.

## TABLE OF CONTENTS

**Lecture I:** Extending the Wasserstein distance of exponent 2 to density operators.  
**Lecture II:** Applying the quantum Wasserstein pseudometric to particle dynamics.  
**Lecture III:** Triangle inequalities and optimal transport in the quantum setting.

1. LECTURE I: EXTENDING THE WASSERSTEIN DISTANCE OF EXPONENT 2  
TO DENSITY OPERATORS

Our general purpose is to extend optimal transport (Wasserstein) distances, defined on Borel probability measures on phase space, i.e.  $\mathbf{R}^d \times \mathbf{R}^d$ , to their quantum analogue, i.e. to density operators on the Hilbert space  $L^2(\mathbf{R}^d)$ .

In this first lecture, we

- recall some fundamental results on classical optimal transport (section 1.1),
- recall some material on trace-class and Hilbert-Schmidt operators (section 1.2),
- introduce one first noncommutative extension of optimal transport (section 1.3),
- present our quantum extension of the Wasserstein metric  $\mathcal{W}_2$  (section 1.4), and
- discuss some basic estimates and examples of computations (section 1.5).

**1.1. A Crash-Course on Classical Optimal Transport.** Before embarking on a description of a quantum analogue of the Wasserstein distance of exponent 2 for density operators, we need to recall some fundamental notions and results pertaining to the classical theory of optimal transport. There are many excellent reference textbooks on optimal transport, such as [63, 4, 64, 58, 33], where the interested reader will find the proofs of all the statements in this section — together with many fascinating applications of optimal transport in various areas of mathematics. Of course, Alessio Figalli’s course [32] in this school is strongly recommended as a general introduction to optimal transport.

1.1.1. *The Monge and the Kantorovich Problems.* Optimal transport grew from Monge’s celebrated “Mémoire<sup>1</sup> sur la théorie des déblais et des remblais”.

In modern mathematical terminology, Monge’s problem can be stated as follows.

**Monge’s problem.** For all  $\mu, \nu \in \mathcal{P}_1(\mathbf{R}^n)$ , find  $T : \mathbf{R}^n \rightarrow \mathbf{R}^n$  measurable such that  $T\#\mu = \nu$  and

$$\begin{aligned} & \int_{\mathbf{R}^n} |T(x) - x| \mu(dx) \\ & = \inf \left\{ \int_{\mathbf{R}^n} |F(x) - x| \mu(dx) \text{ with } F : \mathbf{R}^n \rightarrow \mathbf{R}^n \text{ measurable s.t. } F\#\mu = \nu \right\}, \end{aligned}$$

where  $\mathcal{P}(\mathbf{R}^n)$  designates the set of all Borel probability measures on  $\mathbf{R}^n$ , while

$$\mathcal{P}_k(\mathbf{R}^n) := \left\{ \mu \in \mathcal{P}(\mathbf{R}^n) \text{ s.t. } \int_{\mathbf{R}^n} |x|^k \mu(dx) < \infty \right\},$$

and  $T\#\mu$  is the push-forward of the measure  $\mu$  by the transformation  $T$ , defined by the formula

$$T\#\mu(B) := \mu(T^{-1}B), \quad \text{for all Borel } B \subset \mathbf{R}^n.$$

---

<sup>1</sup>“Déblai” and “remblai” are technical terms for earthwork in French. “Déblai” means excavation, whereas “remblai” is the French word for embankment or backfill. A first version of Monge’s memoir was read on February 7th 1776 at the Académie des sciences, followed by a second version, read on March 27th 1781, and finally published in 1784.

With such a general formulation, Monge's problem does not always have at least one solution. (For example, set  $n = 1$ , choose  $\mu := \delta_0$  and  $\nu := \frac{1}{2}(\delta_{+1} + \delta_{-1})$ . Then there does not exist any map  $T : \mathbf{R} \rightarrow \mathbf{R}$  such that  $T\#\mu = \nu$ .)

However, if  $\mu$  is absolutely continuous with respect to the Lebesgue measure of  $\mathbf{R}^n$ , Monge's problem always has at least one solution. This was proved more than 200 years after Monge's first version of his memoir, in 1979, by Sudakov [61] (see also section 6 of [3], where a gap in Sudakov's original argument is fixed).

Before Sudakov's proof, Kantorovich proposed in [47] a relaxed version of Monge's problem, for which the existence of a solution is elementary.

**The Kantorovich relaxation of Monge's problem.** For all  $\mu, \nu \in \mathcal{P}_1(\mathbf{R}^n)$ , find

$$\mathcal{W}_1(\mu, \nu) := \min_{\rho \in \mathcal{C}(\mu, \nu)} \iint_{\mathbf{R}^n \times \mathbf{R}^n} |x - y| \rho(dx dy),$$

where  $\mathcal{C}(\mu, \nu)$  is the set of "couplings", or "transport plans" between  $\mu$  and  $\nu$ , defined as follows:

$$\mathcal{C}(\mu, \nu) := \left\{ \rho \in \mathcal{P}(\mathbf{R}^{2n}) \left| \begin{array}{l} \rho(A \times \mathbf{R}^n) = \mu(A), \\ \rho(\mathbf{R}^n \times A) = \nu(A), \end{array} \right. \text{ for all Borel } A \subset \mathbf{R}^n \right\}.$$

Observe that  $\mu \otimes \nu \in \mathcal{C}(\mu, \nu)$ , so that  $\mathcal{C}(\mu, \nu) \neq \emptyset$ .

That Kantorovich's problem is indeed a relaxation of Monge's problem is explained by the following elementary observation.

**Remark.** If  $T : \mathbf{R}^n \rightarrow \mathbf{R}^n$  is a measurable map such that  $T\#\mu = \nu$ , the probability measure  $\rho(dx dy) = \mu(dx) \delta_{T(x)}(dy)$  belongs to  $\mathcal{C}(\mu, \nu)$  (this is trivial). Moreover, for all  $\mu, \nu \in \mathcal{P}_1(\mathbf{R}^n)$ , one has

$$\mathcal{W}_1(\mu, \nu) = \inf \left\{ \int_{\mathbf{R}^n} |F(x) - x| \mu(dx) \text{ with } F : \mathbf{R}^n \rightarrow \mathbf{R}^n \text{ measurable s.t. } F\#\mu = \nu \right\}$$

provided that  $\mu$  has no atom (this is not obvious: see Theorem 2.1 in [3]).

**Quiz 1.** Prove the existence of a solution to the Kantorovich problem. (Hint: if  $\mu, \nu \in \mathcal{P}_1(\mathbf{R}^n)$ , then  $\mathcal{C}(\mu, \nu) \subset \mathcal{P}_1(\mathbf{R}^n \times \mathbf{R}^n)$ , and  $\mathcal{C}(\mu, \nu)$  is weakly relatively compact in  $\mathcal{P}(\mathbf{R}^n \times \mathbf{R}^n)$  by Prokhorov's Theorem.)

The main result in [47] is the following equivalent variational formula for  $\mathcal{W}_1(\mu, \nu)$ , which can be deduced from the definition above by convex duality, applying the Fenchel-Moreau-Rockafellar duality theorem (see Theorem 1.12 in [17]).

**Kantorovich(-Rubinstein) duality.** For all  $\mu, \nu \in \mathcal{P}_1(\mathbf{R}^n)$ ,

$$\mathcal{W}_1(\mu, \nu) = \sup_{\substack{\chi \in \text{Lip}(\mathbf{R}^n, \mathbf{R}) \\ \text{Lip}(\chi) \leq 1}} \left| \int_{\mathbf{R}^n} \chi(z) \mu(dz) - \int_{\mathbf{R}^n} \chi(z) \nu(dz) \right|.$$

For a proof of Kantorovich duality, see chapter 1 of [63].

The Kantorovich duality formula for  $\mathcal{W}_1$  has several important applications to the topology of  $\mathcal{P}_1(\mathbf{R}^n)$ , listed below.

**Consequences of the Kantorovich(-Rubinstein) duality.**

- (1) The functional  $\mathcal{W}_1$  is a metric on  $\mathcal{P}_1(\mathbf{R}^n)$ .
- (2) Let  $\mu \in \mathcal{P}_1(\mathbf{R}^n)$  and  $\mu_j$  be a sequence of elements of  $\mathcal{P}_1(\mathbf{R}^n)$ . Then the three conditions below are equivalent
  - (a)  $\mathcal{W}_1(\mu_j, \mu) \rightarrow 0$  as  $j \rightarrow \infty$ ,

(b)  $\mu_j \rightarrow \mu$  weakly in the sense of probability measures as  $j \rightarrow \infty$  and

$$\lim_{R \rightarrow \infty} \sup_{j \geq 1} \int_{|x| > R} |x| \mu_j(dx) = 0,$$

(c)  $\mu_j \rightarrow \mu$  weakly in the sense of probability measures as  $j \rightarrow \infty$  and

$$\lim_{j \rightarrow \infty} \int_{\mathbf{R}^n} |x| \mu_j(dx) = \int_{\mathbf{R}^n} |x| \mu(dx).$$

Statement (2) is Theorem 7.12 in [63]. As for (1), the Kantorovich(-Rubinstein) duality formula obviously implies that  $\mathcal{W}_1$  is nonnegative, symmetric in both its arguments, and satisfies the triangle inequality. Finally, if  $\mathcal{W}_1(\mu, \nu) = 0$ , then

$$\int_{\mathbf{R}^n} \psi(z) \mu(dz) = \int_{\mathbf{R}^n} \psi(z) \nu(dz)$$

for each  $\psi \in C_c^\infty(\mathbf{R}^n)$ , since  $\phi := \psi / (1 + \|\text{grad } \psi\|_{L^\infty(\mathbf{R}^n)})$  satisfies  $\text{Lip}(\phi) \leq 1$ . Hence  $\mu = \nu$  (viewing  $\mu$  and  $\nu$  as distributions of order 0 on  $\mathbf{R}^n$ ).

The functional  $\mathcal{W}_1$  is usually referred to as the Wasserstein distance of exponent 1, and sometimes as the Monge-Kantorovich or the Kantorovich-Rubinstein distance.

1.1.2. *The Wasserstein Distance of Exponent 2.* In Monge's own words "Le prix du transport d'une molécule [est], toute choses égales d'ailleurs, proportionnel à son poids & à l'espace qu'on lui fait parcourir" (all else being equal, the cost of transport for one molecule is proportional to its weight and to the distance over which it is transported).

But one could assume instead that the cost of transport is proportional to some power of the distance. In this section, we consider the special case where the cost of transport is proportional to the *square* distance. In that case, the Monge and the Kantorovich problems are as follows.

**Monge's problem.** For all  $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^n)$ , find  $T : \mathbf{R}^n \rightarrow \mathbf{R}^n$  measurable such that  $T\#\mu = \nu$  and

$$\int_{\mathbf{R}^n} |T(x) - x|^2 \mu(dx)$$

$$= \inf \left\{ \int_{\mathbf{R}^n} |F(x) - x|^2 \mu(dx) \text{ with } F : \mathbf{R}^n \rightarrow \mathbf{R}^n \text{ measurable s.t. } F\#\mu = \nu \right\}.$$

**The Kantorovich problem.** For all  $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^n)$ , find

$$\mathcal{W}_2(\mu, \nu) := \left( \min_{\rho \in \mathcal{C}(\mu, \nu)} \iint_{\mathbf{R}^n \times \mathbf{R}^n} |x - y|^2 \rho(dxdy) \right)^{1/2}.$$

The existence of a solution to the Kantorovich problem is proved by exactly the same weak compactness argument as in Quiz 1 above.

**Remark.** For each  $p \in (1, +\infty)$ , there exists an analogue of the functional  $\mathcal{W}_2$ , denoted  $\mathcal{W}_p$ , in the case where the cost of transport is proportional to the  $p$ th power of the distance. We have chosen to restrict our attention to the cases  $p = 1$  and  $p = 2$ , for which "quantum" analogues have been defined.

The Kantorovich duality in that case is, at first sight, different from the case of exponent 1.

**Kantorovich duality for  $\mathcal{W}_2$ .** For all  $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^n)$ ,

$$\mathcal{W}_2(\mu, \nu)^2 = \sup_{\substack{a(x)+b(y) \leq |x-y|^2 \\ a, b \in \mathcal{C}_b(\mathbf{R}^n)}} \int_{\mathbf{R}^n} a(z)\mu(dz) + \int_{\mathbf{R}^n} b(z)\nu(dz).$$

Moreover, there exist two l.s.c. proper convex functions  $\alpha, \beta$  such that  $\alpha \in L^1(\mathbf{R}^n, \mu)$  and  $\beta \in L^1(\mathbf{R}^n, \nu)$ , satisfying<sup>2</sup>  $\alpha^* = \beta$  and  $\beta^* = \alpha$ , and such that

$$\mathcal{W}_2(\mu, \nu)^2 = \int_{\mathbf{R}^n} (|x|^2 - 2\alpha(x))\mu(dx) + \int_{\mathbf{R}^n} (|y|^2 - 2\beta(y))\nu(dy).$$

The first equality is Theorem 1.3 in [63], while the existence of an optimal pair  $\alpha, \beta$  is Theorem 2.9 in [63].

Perhaps the most important consequence of the Kantorovich duality for  $\mathcal{W}_2$  is the structure of optimal couplings for the Kantorovich problem.

**Optimal couplings.**

(a) (Knott-Smith Theorem.) Let  $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^n)$ . A probability measure  $\rho \in \mathcal{C}(\mu, \nu)$  is an optimal coupling for  $\mathcal{W}_2$  if and only if there exists  $\Phi : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{+\infty\}$ , l.s.c. proper and convex, such that

$$\text{supp}(\rho) \subset \text{graph}(\partial\Phi),$$

where  $\partial\Phi$  is the subdifferential<sup>3</sup> of  $\Phi$ .

(b) (Brenier's Theorem.) Let  $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^n)$ . If  $\mu(S) = 0$  for each Borel set  $S \subset \mathbf{R}^n$  of Hausdorff dimension  $\mathcal{H} - \dim(S) \leq n - 1$ , there exists a unique optimal coupling for  $\mathcal{W}_2$ , of the form

$$\rho(dxdy) = \mu(dx)\delta_{\text{grad}\Phi(x)}(dy)$$

with  $\Phi : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{+\infty\}$  convex and such that  $\Phi \in L^1(\mathbf{R}^n, \mu)$ .

See Theorem 2.12 in [63], and more generally chapter 2 of [63] for a proof of the Knott-Smith and the Brenier theorems.

Since  $\Phi \in L^1(\mathbf{R}^n, \mu)$ , one has  $\mu(\Phi^{-1}(\{+\infty\})) = 0$ , on the other hand  $\Phi$  is locally Lipschitz continuous on the interior of its domain  $\text{Dom}(\Phi) := \mathbf{R}^n \setminus \Phi^{-1}(\{+\infty\})$ . Then (1)  $\text{Dom}(\Phi)$  is a convex subset of  $\mathbf{R}^n$ , and hence  $\partial\text{Dom}(\Phi)$  has Hausdorff dimension  $\leq n - 1$ , and (2) the set of points  $x$  in the interior of  $\text{Dom}(\Phi)$  such that  $\Phi$  is not differentiable at  $x$  has Hausdorff dimension  $\leq n - 1$  (see [2]). Since

$$(\text{grad}\Phi)^{-1}(B) := \{x \in \mathbf{R}^n \text{ s.t. } \Phi \text{ is differentiable at } x \text{ and } \text{grad}\Phi(x) \in B\},$$

<sup>2</sup>A convex function  $\alpha : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{+\infty\}$  is said to be proper if there exists at least one point  $x \in \mathbf{R}^n$  such that  $\alpha(x) < +\infty$ . Its Legendre dual is

$$\alpha^*(p) := \sup_{x \in \mathbf{R}^n} (p \cdot x - \alpha(x)), \quad p \in \mathbf{R}^n.$$

If  $\alpha$  is a l.s.c. proper convex function on  $\mathbf{R}^n$ , then the function  $\alpha^*$  is also l.s.c. proper and convex on  $\mathbf{R}^n$ , and one has  $\alpha^{**} = \alpha$ . See section 1.4 and Theorem 1.11 in [17].

<sup>3</sup>If  $\Phi : \mathbf{R}^n \rightarrow \mathbf{R} \cup \{+\infty\}$  is a proper convex function, its subdifferential at  $x \in \mathbf{R}^n$  is the set

$$\partial\Phi(x) := \{\xi \in \mathbf{R}^n \text{ s.t. } \Phi(y) \geq \Phi(x) + \xi \cdot (y - x), \quad y \in \mathbf{R}^n\}.$$

See Example 2.1.4 in [16]. The subdifferential of a proper convex function is an example of monotone operator: if  $x_1, x_2 \in \mathbf{R}^n$  and if  $\xi_1 \in \partial\Phi(x_1)$  and  $\xi_2 \in \partial\Phi(x_2)$ , then

$$(\xi_2 - \xi_1) \cdot (x_2 - x_1) \geq 0.$$

One can check that (1)  $\Phi$  is differentiable at  $x \in \mathbf{R}^n$  if and only if  $\partial\Phi(x)$  contains a single element, which is  $\text{grad}\Phi(x)$ , and (2) if a l.s.c. proper convex function  $\Phi$  on  $\mathbf{R}^n$  is strictly convex in a neighborhood of  $x \in \mathbf{R}^n$ , its Legendre dual  $\Phi^*$  is differentiable on  $\partial\Phi(x)$ , and  $\text{grad}\Phi^*(\xi) = x$  for all  $\xi \in \partial\Phi(x)$ .

and since  $\text{grad } \Phi(x)$  exists for  $\mu$ -a.e.  $x \in \mathbf{R}^n$ , the push-forward measure  $(\text{grad } \Phi)_\# \mu$  is well-defined.

We shall conclude this section with a brief list of the most important topological properties of  $\mathcal{W}_2$ .

**Properties of  $\mathcal{W}_2$ .**

(1) The functional  $\mathcal{W}_2$  is a metric on  $\mathcal{P}_2(\mathbf{R}^n)$ .

In particular

$$\mathcal{W}_2(\mu, \nu) = 0 \iff \mu = \nu;$$

the optimal coupling in that case is  $\rho(dx dy) = \mu(dx) \delta_x(dy)$ , and the Brenier transport map is the identity, which is the gradient of the convex function  $x \mapsto \frac{1}{2}|x|^2$ .

(2) Let  $\mu \in \mathcal{P}_2(\mathbf{R}^n)$  and  $\mu_j$  be a sequence of elements of  $\mathcal{P}_2(\mathbf{R}^n)$ . Then the two conditions below are equivalent:

- (a)  $\mathcal{W}_2(\mu_j, \mu) \rightarrow 0$  as  $j \rightarrow \infty$ ,
- (b)  $\mu_j \rightarrow \mu$  weakly in the sense of probability measures, and

$$\lim_{R \rightarrow \infty} \sup_{j \geq 1} \int_{|x| > R} |x|^2 \mu_j(dx) = 0.$$

Statement (1) is Theorem 7.3 in [63], while statement (2) is Theorem 7.12 in [63]. The proof of the triangle inequality is definitely nontrivial, at variance with the case of  $\mathcal{W}_1$ , for which the triangle inequality follows from the expression of  $\mathcal{W}_1$  using the Kantorovich duality.

Here is a nontrivial example where  $\mathcal{W}_2(\mu, \nu)$  can be computed explicitly. See [34] for a proof of the formula below.

**Example.** Let  $G_1, G_2$  be Gaussian laws on  $\mathbf{R}^n$  with means  $m_1, m_2$  and covariance matrices  $A_1, A_2$ . Then

$$\mathcal{W}_2(G_1, G_2)^2 = |m_1 - m_2|^2 + \text{trace} \left( A_1 + A_2 - 2 \left( \sqrt{A_1} A_2 \sqrt{A_1} \right)^{\frac{1}{2}} \right)$$

**Quiz 2.** For all  $m_1, m_2 \in \mathbf{R}^n$ , compute  $\mathcal{W}_p(\delta_{m_1}, \delta_{m_2})$  for  $p = 1$  and  $p = 2$ . Does there exist optimal transport map(s) in both cases?

In most reference textbooks on optimal transport, the proof of the triangle inequality for  $\mathcal{W}_p$  with  $1 < p < \infty$  is based on a nontrivial construction referred to as “glueing” couplings having a common marginal (Lemma 7.6 in [63], or Lemma 5.3.2 in [4]). This procedure is itself based on the notion of “disintegration” of a probability measure on a Cartesian product with respect to one of its marginals (obviously related to the notion of conditional probability): see Theorem 5.3.1 in [4]. There is an alternative to the approach based on the disintegration theorem, which uses instead the Hahn-Banach theorem: see Exercise 7.9 in [63].

In the case where there exist optimal transport maps (e.g. when the sets of Hausdorff codimension  $\geq 1$  are negligible for the probability measures considered in the triangle inequality, according to the Brenier theorem), the proof of the triangle inequality is very simple: see Lemma 5.3 in [58]. In the setting considered here (where the underlying metric space is  $\mathbf{R}^n$  with its canonical Euclidean distance), one can always reduce the triangle inequality for general probability measures to this simple case by an approximation argument (Lemma 5.2 in [58]).

However, we shall see later that the very notion of a transport map in the quantum setting remains to be clarified. Likewise, the possibility of “glueing”



couplings with a common marginal in the quantum setting seems to be an open problem at the time of this writing.

For these reasons, a third approach to the triangle inequality for  $\mathcal{W}_2$ , entirely based on the Kantorovich duality formula, is proposed in the exercise below<sup>4</sup>. Perhaps this could be useful later, for instance in the quantum setting.

**Quiz 3.** Let  $\mu, \nu, \rho \in \mathcal{P}_2(\mathbf{R}^n)$ , and let  $A, \Gamma$  be optimal functions in the Kantorovich duality formula for  $\mathcal{W}_2(\mu, \rho)$  such that  $A \in L^1(\mathbf{R}^n, \mu)$  and  $\Gamma \in L^1(\mathbf{R}^n, \rho)$ , i.e.

$$\mathcal{W}_2(\mu, \rho)^2 = \int_{\mathbf{R}^n} A(x)\mu(dx) + \int_{\mathbf{R}^n} \Gamma(z)\rho(dz).$$

while  $\alpha$  and  $\gamma$  defined by  $\alpha(x) := \frac{1}{2}(|x|^2 - A(x))$  and  $\gamma(z) := \frac{1}{2}(|z|^2 - \Gamma(z))$  are l.s.c. convex functions such that  $\alpha^* = \gamma$  and  $\gamma^* = \alpha$ . For each  $\eta > 0$ , set

$$B_\eta(y) := \left(1 + \frac{1}{\eta}\right) \inf_{z \in \mathbf{R}^n} \left( |y - z|^2 - \frac{\Gamma(z)}{1 + \frac{1}{\eta}} \right), \quad y \in \mathbf{R}^n.$$

(1) Prove that the function

$$y \mapsto \frac{1}{2} \left( |y|^2 - \frac{B_\eta(y)}{1 + \frac{1}{\eta}} \right)$$

is the Legendre transform of a l.s.c. proper convex function to be computed in terms of  $\Gamma$ .

(2) Prove that  $A(x) - B_\eta(y) \leq (1 + \eta)|x - y|^2$  for all  $x, y \in \mathbf{R}^d$ .

(3) Prove that  $B_\eta \in L^1(\mathbf{R}^n, \nu)$  for each  $\eta > 0$ .

(4) Prove that, for each  $\eta > 0$

$$\int_{\mathbf{R}^n} A(x)\mu(dx) - \int_{\mathbf{R}^n} B_\eta(y)\nu(dy) \leq (1 + \eta)\mathcal{W}_2(\mu, \nu)^2.$$

(5) Prove that, for each  $\eta > 0$

$$\int_{\mathbf{R}^n} B_\eta(y)\nu(dy) + \int_{\mathbf{R}^n} \Gamma(z)\rho(dz) \leq \left(1 + \frac{1}{\eta}\right)\mathcal{W}_2(\nu, \rho)^2.$$

(6) Prove that, for each  $\eta > 0$

$$\mathcal{W}_2(\mu, \rho)^2 \leq (1 + \eta)\mathcal{W}_2(\mu, \nu)^2 + \left(1 + \frac{1}{\eta}\right)\mathcal{W}_2(\nu, \rho)^2.$$

(7) Prove that  $\mathcal{W}_2$  satisfies the triangle inequality.

**1.2. Density Operators in Quantum Mechanics.** In classical mechanics, the state of a point particle is completely defined by its position  $q \in \mathbf{R}^d$  and its momentum  $p \in \mathbf{R}^d$ . The point particle phase space is therefore  $\mathbf{R}^d \times \mathbf{R}^d$ , and it is natural to study the probability of finding a point particle in subsets of the phase space. This is precisely the statistical formalism used by Maxwell and Boltzmann in the kinetic theory of gases.

In quantum mechanics, the analogous formalism involves density operators, a special class of operators on the Hilbert space  $L^2(\mathbf{R}^d)$ . Before studying density operators, we need to recall some fundamental notions in the theory of operators on Hilbert spaces.

<sup>4</sup>I came up with this proof during the week of the summer school at the Rényi Institute. The same method applies to the Wasserstein distance of exponent  $p$  for all  $p \in (1, \infty)$ , but the computations (which I shall publish elsewhere) are more involved than in the case  $p = 2$ , which is the only one of interest here. I have not seen this proof in any of the reference textbooks on optimal transport that I have been using, and I do not know whether it is original.

1.2.1. *Trace of an Operator.* Let  $\mathfrak{H}$  be a complex, separable Hilbert space with inner product denoted by  $(\cdot|\cdot)_{\mathfrak{H}}$ , and let  $\mathcal{L}(\mathfrak{H})$  designate the algebra of bounded operators on the Hilbert space  $\mathfrak{H}$ .

**Positive operators.** An operator  $T \in \mathcal{L}(\mathfrak{H})$  is said to be positive if

$$T = T^*, \quad \text{and} \quad (x|Tx)_{\mathfrak{H}} \geq 0, \quad x \in \mathfrak{H}.$$

Equivalently, if  $T \in \mathcal{L}(\mathfrak{H})$ , then

$$T = T^* \geq 0 \iff \text{there exists } S \in \mathcal{L}(\mathfrak{H}) \text{ s.t. } T = S^*S.$$

**Trace of a positive operator.** For  $T \in \mathcal{L}(\mathfrak{H})$  such that  $T = T^* \geq 0$ , we define

$$\text{trace}_{\mathfrak{H}}(T) := \sum_{j \geq 1} (e_j|Te_j)_{\mathfrak{H}} \in [0, +\infty] \text{ for all Hilbert basis } (e_j)_{j \geq 0} \text{ of } \mathfrak{H}.$$

One easily checks that, if this sum is finite for *one* Hilbert basis  $(e_j)_{j \geq 0}$  of  $\mathfrak{H}$ , it is finite for *all* Hilbert basis of  $\mathfrak{H}$ , since the (infinite dimensional) “transition matrix” from any Hilbert basis of  $\mathfrak{H}$  to  $(e_j)_{j \geq 0}$  is a unitary operator on  $\mathfrak{H}$ .

The definition of the trace follows the definition of the Lebesgue integral on the real line: first we define trace of any self-adjoint positive operator, as we define the integral of any measurable positive function. Then we define the analogue of the Lebesgue space  $L^1$ , and extend the trace to this space by linearity, exactly in the same way as the integral is extended from the set of measurable positive functions to the Lebesgue space  $L^1$ .

**Trace-class operators.** The set of trace-class operators is

$$\mathcal{L}^1(\mathfrak{H}) := \{T \in \mathcal{L}(\mathfrak{H}) \text{ s.t. } \|T\|_1 := \text{trace}_{\mathfrak{H}}(|T|) < \infty\},$$

where  $|T| := \sqrt{T^*T}$  for each  $T \in \mathcal{L}(\mathfrak{H})$ .

The following properties of  $\mathcal{L}^1(\mathfrak{H})$  are well known.

**Properties of trace-class operators.**

(a) Trace-class operators are compact:

$$\mathcal{L}^1(\mathfrak{H}) \subset \mathcal{K}(\mathfrak{H}),$$

where  $\mathcal{K}(\mathfrak{H})$  is the set of compact operators on  $\mathfrak{H}$ , i.e. the operator-norm closure of the set of finite rank operators on  $\mathfrak{H}$ . (One easily checks that  $\mathcal{K}(\mathfrak{H})$  is a two-sided ideal of  $\mathcal{L}(\mathfrak{H})$ , stable by the involution  $T \mapsto T^*$ .)

(b) The set of trace-class operators  $\mathcal{L}^1(\mathfrak{H})$  is a two-sided ideal of  $\mathcal{L}(\mathfrak{H})$ , which is stable by the involution  $T \mapsto T^*$ :

$$\begin{aligned} A \in \mathcal{L}(\mathfrak{H}) \text{ and } T \in \mathcal{L}^1(\mathfrak{H}) &\implies AT \text{ and } TA \in \mathcal{L}^1(\mathfrak{H}), \\ T \in \mathcal{L}^1(\mathfrak{H}) &\implies T^* \in \mathcal{L}^1(\mathfrak{H}). \end{aligned}$$

(c) The trace, which is defined on the set of positive trace-class operators on  $\mathfrak{H}$ , extends as a linear functional on  $\mathcal{L}^1(\mathfrak{H})$  satisfying the properties

$$\text{trace}(T^*) = \overline{\text{trace}(T)}, \quad \text{trace}(AT) = \text{trace}(TA), \quad \text{and} \quad |\text{trace}(AT)| \leq \|A\| \|T\|_1,$$

for all  $T \in \mathcal{L}^1(\mathfrak{H})$  and all  $A \in \mathcal{L}(\mathfrak{H})$ .

(d) The  $\mathbf{C}$ -linear space  $\mathcal{L}^1(\mathfrak{H})$  is a Banach space for the trace norm  $T \mapsto \|T\|_1$ ; besides<sup>5</sup>

$$\mathcal{K}(\mathfrak{H})' = \mathcal{L}^1(\mathfrak{H}), \quad \text{and} \quad \mathcal{L}^1(\mathfrak{H})' = \mathcal{L}(\mathfrak{H}).$$

<sup>5</sup>The topological dual of a real or complex normed linear space  $E$ , i.e. the set of continuous linear functionals on  $E$  (with real or complex values), is denoted by  $E'$ .

In both equalities, the duality is defined by the trace, i.e.

$$\langle T, K \rangle_{\mathcal{L}^1(\mathfrak{H}), \mathcal{K}(\mathfrak{H})} := \text{trace}_{\mathfrak{H}}(TK), \quad \langle A, T \rangle_{\mathcal{L}(\mathfrak{H}), \mathcal{L}^1(\mathfrak{H})} := \text{trace}_{\mathfrak{H}}(AT).$$

**Partial Trace.** The construction of the trace explained before is strikingly similar to the construction of the Lebesgue integral. In the present section, we are going to study an analogue of the Fubini theorem.

For each  $T \in \mathcal{L}^1(\mathfrak{H}_1 \otimes \mathfrak{H}_2)$ , one defines  $T_1 = \text{trace}_2(T) \in \mathcal{L}^1(\mathfrak{H}_1)$  by the formula

$$\text{trace}_{\mathfrak{H}_1}(T_1 A) = \text{trace}_{\mathfrak{H}_1 \otimes \mathfrak{H}_2}(T(A \otimes I_{\mathfrak{H}_2})), \quad \text{for all } A \in \mathcal{L}(\mathfrak{H}).$$

There is a similar definition of  $\text{trace}_1(T) \in \mathcal{L}^1(\mathfrak{H}_2)$ .

Observe indeed that the map

$$\mathcal{K}(\mathfrak{H}_1) \ni A \mapsto \text{trace}_{\mathfrak{H}_1 \otimes \mathfrak{H}_2}(T(A \otimes I_{\mathfrak{H}_2})) \in \mathbf{C}$$

is a norm-continuous linear functional on  $\mathcal{K}(\mathfrak{H}_1)$ , and is therefore represented by a unique trace-class operator  $T_1$  on  $\mathfrak{H}_1$ . That the defining identity for  $\text{trace}_2(T)$  holds for all  $A \in \mathcal{L}(\mathfrak{H}_1)$  follows from the density of finite rank operators in  $\mathcal{L}^1(\mathfrak{H})$  for the trace-norm.

**Remark.** The analogue of this construction in the context of integration is the following form of the Fubini theorem: if  $f \in L^1(\mathbf{R}^m \times \mathbf{R}^n)$ , the function  $y \mapsto f(x, y)$  belongs to  $L^1(\mathbf{R}^n)$  for a.e.  $x \in \mathbf{R}^m$ , the function

$$F : x \mapsto \int_{\mathbf{R}^n} f(x, y) dy$$

belongs to  $L^1(\mathbf{R}^m)$ , and

$$\int_{\mathbf{R}^m \times \mathbf{R}^n} f(x, y) dx dy = \int_{\mathbf{R}^m} F(x) dx.$$

The function  $F$  in this statement is easily seen to be the analogue of  $\text{trace}_2(T)$  for  $T \in \mathcal{L}^1(\mathfrak{H}_1 \otimes \mathfrak{H}_2)$ .

1.2.2. *Hilbert-Schmidt Operators.* An operator  $T \in \mathcal{L}(\mathfrak{H})$  is said to be a Hilbert-Schmidt operator if

$$\text{trace}_{\mathfrak{H}}(T^*T) < \infty.$$

The Hilbert-Schmidt class is the set of Hilbert-Schmidt operators:

$$\mathcal{L}^2(\mathfrak{H}) := \{T \in \mathcal{L}(\mathfrak{H}) \text{ s.t. } \text{trace}_{\mathfrak{H}}(T^*T) < \infty\}.$$

The Hilbert-Schmidt class  $\mathcal{L}^2(\mathfrak{H})$  is a Hilbert space for the inner product

$$(T_1 | T_2)_2 := \text{trace}_{\mathfrak{H}}(T_1^* T_2)$$

defining the Hilbert-Schmidt norm

$$\|T\|_2 := \sqrt{\text{trace}_{\mathfrak{H}}(T^*T)}.$$

Moreover

$$\mathcal{L}^1(\mathfrak{H}) \subset \mathcal{L}^2(\mathfrak{H}) \subset \mathcal{K}(\mathfrak{H}) \subset \mathcal{L}(\mathfrak{H})$$

with continuous inclusions, and

$$\|T\| \leq \|T\|_2 \leq \|T\|_1, \quad T \in \mathcal{L}^1(\mathfrak{H}).$$

Besides  $\mathcal{L}^2(\mathfrak{H})$  is a two-sided ideal of  $\mathcal{L}(\mathfrak{H})$ .

If  $T = T^* \in \mathcal{L}^2(\mathfrak{H})$ , then  $T$  is a compact operator, so that there exists  $(e_j)_{j \geq 1}$ , a Hilbert basis of  $\mathfrak{H}$  and  $(\tau_j)_{j \geq 1} \in \ell^2(\mathbf{N}^*; \mathbf{R})$  such that

$$T = \sum_{j \geq 1} \tau_j P_j, \quad \|T\|_2^2 = \sum_{j \geq 1} \tau_j^2 \quad \text{with } P_j \phi := (e_j | \phi)_{\mathfrak{H}} e_j.$$

Besides, in the case where  $\mathfrak{H} = L^2(\mathbf{R}^d)$ , one has therefore

$$(1) \quad T\phi(x) = \int_{\mathbf{R}^d} t(x, y)\phi(y)dy, \quad \text{for all } \phi \in \mathfrak{H},$$

where

$$t(x, y) := \sum_{j \geq 1} \tau_j e_j(x) \overline{e_j(y)}.$$

In particular

$$(2) \quad \|T\|_2^2 = \sum_{j \geq 1} \tau_j^2 = \iint_{\mathbf{R}^d \times \mathbf{R}^d} |t(x, y)|^2 dx dy.$$

Conversely, an integral operator  $T$  of the form (1) belongs to  $\mathcal{L}^2(\mathfrak{H})$  if and only if  $t \in L^2(\mathbf{R}^d \times \mathbf{R}^d)$ , and the first left-hand side in (2) is equal to the last right-hand side of in (2).

The interested reader will discover additional important properties of trace-class and Hilbert-Schmidt operators in the next two exercises.

**Quiz 4.** In this exercise,  $\mathfrak{H} := L^2(\mathbf{R}^d)$ .

- (1) Prove that any  $T \in L^1(\mathfrak{H})$  can be put in the form  $T = T_1 T_2$  with  $T_1, T_2 \in \mathcal{L}^2(\mathfrak{H})$ , and that  $\|T\|_1 \leq \|T_1\|_2 \|T_2\|_2$ .  
 (2) For all  $T \in L^1(\mathfrak{H})$ , can one find  $T_1, T_2 \in \mathcal{L}^2(\mathfrak{H})$  such that

$$T = T_1 T_2 \quad \text{and} \quad \|T\|_1 = \|T_1\|_2 \|T_2\|_2?$$

- (3) Prove that for each  $T \in L^1(\mathfrak{H})$ , there exists  $t \equiv t(x, y)$  such that  $z \mapsto t(x + z, x)$  belongs to  $C_b(\mathbf{R}_z^d; L^1(\mathbf{R}_x^d))$ , which is an integral kernel for  $T$ , in the sense that

$$T\phi(x) = \int_{\mathbf{R}^d} t(x, y)\phi(y)dy, \quad \text{for all } \phi \in \mathfrak{H},$$

and that

$$\text{trace}_{\mathfrak{H}}(T) = \int_{\mathbf{R}^d} t(x, x)dx.$$

Question (3) suggests that  $t(x, y)$  is the analogue for  $T$  of the entries of a matrix if the infinite dimensional Hilbert space  $\mathfrak{H}$  is replaced with  $\mathbf{C}^n$ . One might therefore believe that an integral operator is trace-class if the restriction of its integral kernel to the diagonal is summable. This is not the case, as shown by the next example.

**Quiz 5.** Consider the Volterra operator  $V$  defined on  $L^2([0, 1])$  by the formula

$$V\phi(x) := \int_0^x \phi(y)dy, \quad \phi \in L^2([0, 1]).$$

- (1) Prove that  $V$  is the operator defined by the integral kernel  $v(x, y) = \mathbf{1}_{0 \leq y \leq x}$ .  
 (2) Is  $V$  a Hilbert-Schmidt operator on  $L^2([0, 1])$ ?  
 (3) Does the function  $x \mapsto v(x, x)$  belong to  $L^1([0, 1])$ ?  
 (4) Is  $V$  a trace-class operator on  $L^2([0, 1])$ ?  
 (5) What are the eigenvalues of  $V$ ?  
 (6) What is the spectral radius of  $V$ ?

This last exercise shows the importance of the continuity condition in question (3) of Quiz 3. Any trace-class operator  $T$  on  $L^2(\mathbf{R}^d)$ , being a Hilbert-Schmidt

operator, is an integral operator defined by a unique integral kernel in  $L^2(\mathbf{R}^d \times \mathbf{R}^d)$ . Since the diagonal is a Lebesgue-negligible set in  $\mathbf{R}^d \times \mathbf{R}^d$ , the restriction to the diagonal of this integral kernel is a priori not a well defined function. Yet, any trace-class operator  $T$  on  $L^2(\mathbf{R}^d)$  has an integral kernel satisfying the continuity condition of (3) in Quiz 3, which is a representative of the unique  $L^2(\mathbf{R}^d \times \mathbf{R}^d)$  integral kernel of  $T$  viewed as a Hilbert-Schmidt operator on  $L^2(\mathbf{R}^d)$ , and the trace of  $T$  is indeed the integral of the restriction of this particular integral kernel to the diagonal, which is a well-defined element of  $L^1(\mathbf{R}^d)$ .

1.2.3. *Density Operators.* As mentioned above, density operators on  $L^2(\mathbf{R}^d)$  are the quantum analogue of Borel probability measures on the single-particle phase space  $\mathbf{R}^d \times \mathbf{R}^d$ . The positivity of a probability measure  $\mu$  on  $\mathbf{R}^d \times \mathbf{R}^d$  becomes the positivity of an operator  $R$  on  $L^2(\mathbf{R}^d)$ , while the analogue of the normalization condition  $\mu(\mathbf{R}^d \times \mathbf{R}^d) = 1$  in the quantum setting is the condition  $\text{trace}_{\mathfrak{H}}(R) = 1$ .

Throughout this section, we set  $\mathfrak{H} = L^2(\mathbf{R}^d)$ .

**(Quantum) Density operators.** A density operator on  $\mathfrak{H}$  is an element of

$$\mathcal{D}(\mathfrak{H}) := \{T \in \mathcal{L}(\mathfrak{H}) \text{ s.t. } T = T^* \geq 0 \text{ and } \text{trace}_{\mathfrak{H}}(T) = 1\} \subset \mathcal{L}^1(\mathfrak{H}).$$

As explained above,  $\mathcal{D}(\mathfrak{H})$  is the quantum analogue of  $\mathcal{P}(\mathbf{R}^d \times \mathbf{R}^d)$ .

When dealing with computations involving quantum states, it will be especially convenient to use the notation involving bras and kets, which is recalled below.

**Dirac bra-ket notation.** For  $\phi, \psi \in \mathfrak{H}$ , we denote by  $|\psi\rangle$  the vector  $\psi$ , while

$$\langle\phi| \text{ denotes the linear functional } \psi \mapsto \int_{\mathbf{R}^d} \overline{\phi(x)}\psi(x)dx = \langle\phi|\psi\rangle.$$

With this notation, one easily checks that

$$\psi \in \mathfrak{H} \text{ and } \|\psi\|_{\mathfrak{H}} = 1 \implies |\psi\rangle\langle\psi| = \text{orthogonal projection on } \mathbf{C}\psi.$$

**Example: Schrödinger's coherent state.** For  $q, p \in \mathbf{R}^d$ , set

$$|q, p\rangle(x) := (2\pi\hbar)^{-d/4} \exp\left(-\frac{1}{2\hbar}|x - q|^2\right) \exp\left(\frac{i}{\hbar}p \cdot \left(x - \frac{q}{2}\right)\right).$$

One easily checks that  $\| |q, p\rangle \|_{\mathfrak{H}} = 1$  so that  $|q, p\rangle\langle q, p| \in \mathcal{D}(\mathfrak{H})$ .

The density operator  $|q, p\rangle\langle q, p|$  defined above is a quantum analogue (and by no means the only one) of the probability measure  $\delta_{(q,p)} \in \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^d)$ .

Since density operators are the quantum analogue of phase space (Borel) probability measures, a natural problem is that of comparing two density operators by some procedure which corresponds to a comparison between phase space (Borel) probability measures in the classical limit of quantum mechanics. In the sequel, we shall consider an important example, involving Schrödinger's coherent states, for which explicit computations are very easy.

**Key example.** For  $(q_1, p_1) \neq (q_2, p_2) \in \mathbf{R}^d \times \mathbf{R}^d$ , set

$$R_1 := |q_1, p_1\rangle\langle q_1, p_1| \quad \text{and} \quad R_2 = |q_2, p_2\rangle\langle q_2, p_2|.$$

Then  $R_1 - R_2$  is a self-adjoint, rank-2 operator, such that  $\text{trace}(R_1 - R_2) = 0$ , and hence there exists  $\lambda > 0$  and an orthonormal basis  $(e, f)$  of  $\text{Ran}(R_1 - R_2)$  for which

$$R_1 - R_2 = \lambda|e\rangle\langle e| - \lambda|f\rangle\langle f|.$$

Therefore

$$\|R_1 - R_2\|_1 = 2\lambda \quad \text{and} \quad \|R_1 - R_2\|_2 = \sqrt{2}\lambda,$$

so that

$$\|R_1 - R_2\|_1 = \sqrt{2}\|R_1 - R_2\|_2.$$

On the other hand,  $R_1 - R_2$  is the Hilbert-Schmidt integral operator with integral kernel

$$\begin{aligned} r(x, y) := & (2\pi\hbar)^{-d/2} \exp\left(-\frac{1}{2\hbar}(|x - q_1|^2 + |y - q_1|^2)\right) \exp\left(\frac{i}{\hbar}p_1 \cdot (x - y)\right) \\ & - (2\pi\hbar)^{-d/2} \exp\left(-\frac{1}{2\hbar}(|x - q_2|^2 + |y - q_2|^2)\right) \exp\left(\frac{i}{\hbar}p_2 \cdot (x - y)\right), \end{aligned}$$

so that

$$\|R_1 - R_2\|_2^2 = \iint_{\mathbf{R}^d \times \mathbf{R}^d} |r(x, y)|^2 dx dy = 2(1 - e^{-(|q_1 - q_2|^2 + |p_1 - p_2|^2)/2\hbar}).$$

Thus

$$\|R_1 - R_2\|_1 = 2\sqrt{1 - e^{-(|q_1 - q_2|^2 + |p_1 - p_2|^2)/2\hbar}} \rightarrow 2 \quad \text{as } \hbar \rightarrow 0^+$$

since  $(q_1, p_1) \neq (q_2, p_2)$ . In other words, passing to the limit as  $\hbar \rightarrow 0$ ,

$$\| |q_1, p_1\rangle\langle q_1, p_1| - |q_2, p_2\rangle\langle q_2, p_2| \|_1 \rightarrow \|\delta_{(q_1, p_1)} - \delta_{(q_2, p_2)}\|_{TV} = \begin{cases} 2 & \text{if } (q_1, p_1) \neq (q_2, p_2), \\ 0 & \text{if } (q_1, p_1) = (q_2, p_2), \end{cases}$$

In particular,  $\|\cdot\|_1$  fails to discriminate between density operators concentrating on phase space points at a distance  $\gg O(\hbar^{1/2})$  of each other in the classical limit.

**1.3. The Connes Distance in Noncommutative Geometry.** It is well known that one of the main differences between quantum and classical mechanics is that the product of phase space coordinates of a point particle in classical mechanics is a commutative operation, whereas the quantum analogues of these quantities are operators on a Hilbert space, and their product is in general noncommutative. For example, think of a point particle in space dimension 1, and let  $q, p \in \mathbf{R}$  designate respectively its position and momentum in classical mechanics. Obviously

$$pq - qp = 0.$$

In quantum mechanics, the real-valued functions  $(q, p) \mapsto q$  and  $(q, p) \mapsto p$  are replaced with (unbounded) operators  $\hat{q}$  and  $\hat{p}$  on  $L^2(\mathbf{R})$  satisfying the canonical commutation relation (CCR)

$$[\hat{p}, \hat{q}] = \hat{p}\hat{q} - \hat{q}\hat{p} = -i\hbar \text{Id}.$$

To the best of our knowledge, the first attempt at extending the notions of optimal transport to the noncommutative setting is due to Connes [24]. We shall briefly describe his work on this topic in this section.

Let  $\mathcal{A}$  be a unital  $C^*$ -algebra. (We recall that a  $C^*$ -algebra is a complex Banach algebra endowed with a linear involution  $x \mapsto x^*$  such that  $(\alpha x)^* = \bar{\alpha}x^*$  for all  $\alpha \in \mathbf{C}$  and  $\|x^*x\| = \|x\|^2$  for all  $x$  in the algebra. The latter condition is known as “the  $C^*$  identity”. This is a very strong condition, which connects the norm with the algebraic structure. For instance, it implies that  $\|x\|^2$  is the spectral radius of  $x^*x$  for each  $x$  in the algebra.) Here are a few examples of  $C^*$ -algebras:

(a)  $\mathcal{A} = C(X, \mathbf{C})$  with  $X$  compact; in this example

$$f^*(x) := \overline{f(x)} \text{ for all } x \in X, \quad \|f\| := \sup_{x \in X} |f(x)|.$$

This is the prototype of a unital commutative  $C^*$ -algebra, the unit being the constant function  $x \mapsto 1$ .

(b)  $\mathcal{A} = \mathcal{L}(\mathfrak{H})$  where  $\mathfrak{H}$  is a separable complex Hilbert space; in this example, the involution is  $T \mapsto T^*$  where  $T^*$  is the adjoint of the operator  $T$  on  $\mathfrak{H}$ , while  $\|T\|$

is the operator-norm of  $T$ , i.e.  $\|T\| := \sup\{\|Tx\|_{\mathfrak{H}} : \|x\|_{\mathfrak{H}} = 1\}$ . This is a unital  $C^*$ -algebra, with unit  $\text{Id}_{\mathfrak{H}}$ , and it is not separable for the norm topology.

(c)  $\mathcal{A} = \mathcal{K}(\mathfrak{H})$ , the set of compact operators on an infinite dimensional, separable complex Hilbert space  $\mathfrak{H}$ ; this is a non unital  $C^*$ -subalgebra of  $\mathcal{L}(\mathfrak{H})$ ; in fact  $\mathcal{K}(\mathfrak{H})$  is the only norm-closed two-sided ideal of  $\mathcal{L}(\mathfrak{H})$ .

(d)  $\mathcal{A} = \mathcal{L}(\mathfrak{H})/\mathcal{K}(\mathfrak{H})$  where  $\mathfrak{H}$  is an infinite dimensional, separable complex Hilbert space; this a unital  $C^*$ -algebra, which is simple — meaning that its only closed two-sided ideals are  $\{0\}$  and  $\mathcal{A}$  — and known as the Calkin algebra; obviously  $\mathcal{A}$  is not of the form  $\mathcal{L}(\mathfrak{H}_1)$  for any separable Hilbert space.

A *state* on a  $C^*$ -algebra  $\mathcal{A}$  is a positive linear functional on  $\mathcal{A}$  — positive meaning that  $\omega(a^*a) \geq 0$  — of norm 1 — meaning that

$$\|\omega\| := \sup\{|\omega(x)| : x \in \mathcal{A} \text{ and } \|x\| \leq 1\} = 1.$$

It is a classical exercise<sup>6</sup> to check that, if  $\omega$  is continuous linear functional on a unital  $C^*$ -algebra  $\mathcal{A}$ , then

$$\omega \text{ is positive} \iff \|\omega\| = \omega(1).$$

### Examples of states

- $\mathcal{A} = C(X; \mathbf{C})$  with  $X$  compact, and  $\omega(f) = f(x_0) = \langle \delta_{x_0}, f \rangle$ , for some  $x_0 \in X$ ;
- $\mathcal{A} = \mathcal{L}(\mathfrak{H})$  and  $\omega(A) := \langle \psi | A | \psi \rangle$  for some  $\psi \in \mathfrak{H}$  with  $\|\psi\|_{\mathfrak{H}} = 1$ ;
- $\mathcal{A} = \mathcal{L}(\mathfrak{H})$  and  $\omega(A) := \text{trace}_{\mathfrak{H}}(RA)$  for some  $R \in \mathcal{D}(\mathfrak{H})$ .

Let  $(\mathfrak{H}, D)$  be a Fredholm module on  $\mathcal{A}$ , meaning that

- (a) there is a  $*$ -linear representation<sup>7</sup>  $\pi$  of  $\mathcal{A}$  in  $\mathfrak{H}$ ,
- (b)  $D = D^*$  is a self-adjoint unbounded operator on  $\mathfrak{H}$  such that

$$(I + D^2)^{-1} \in \mathcal{K}(\mathfrak{H}),$$

- (c)  $\{a \in \mathcal{A} \text{ s.t. } [D, \pi(a)] \in \mathcal{L}(\mathfrak{H})\}$  is norm-dense in  $\mathcal{A}$ .

**Theorem (Connes [24]).** Assume that

$$\{a \in \mathcal{A} \text{ s.t. } \|[D, \pi(a)]\|_{\mathfrak{H}} \leq 1\} / \mathbf{C}1 \text{ is bounded.}$$

Then, the following formula metrizes the set of states on  $\mathcal{A}$ :

$$\text{dist}_C(\omega_1, \omega_2) := \sup\{|\omega_1(a) - \omega_2(a)| \text{ s.t. } \|[D, \pi(a)]\|_{\mathfrak{H}} \leq 1\}.$$

There is an obvious similarity between Connes' definition and the Kantorovich(-Rubinstein) duality for  $\mathcal{W}_1$ : it clearly suggests to think of Connes' distance as the analogue of the metric  $\mathcal{W}_1$  in noncommutative geometry.

It is interesting to see how Connes' definition can be applied to a commutative setting, corresponding to example (a) above of a  $C^*$ -algebra.

**Example 1: the Dirac operator as a Fredholm module.** Set  $\mathcal{A} = C(M)$  where  $M$  is a compact spin Riemannian manifold (see for instance [13]), with Riemannian metric  $g$ . Let  $S$  be the spinor bundle on  $M$ , and set  $\mathfrak{H} := L^2(M; S)$ , the Hilbert space of  $L^2$  sections of  $S$ . Let  $\mathcal{A}$  act on  $\mathfrak{H}$  by scalar multiplication — to avoid

<sup>6</sup>Here is a hint for the interested reader to prove the direct implication. If  $\omega$  is a positive linear functional on  $\mathcal{A}$ , check that  $(x, y) \mapsto \omega(x^*y)$  is a positive sesquilinear form on  $\mathcal{A}$ , and use the Cauchy-Schwarz inequality to check that  $\omega(x^*) = \overline{\omega(x)}$  and that  $\omega$  is continuous on  $\mathcal{A}$  with norm  $\|\omega\| = \omega(1)$ .

<sup>7</sup>I.e.  $\pi$  is a morphism of algebras from  $\mathcal{A}$  to  $\mathcal{L}(\mathfrak{H})$  satisfying the condition  $\pi(a^*) = \pi(a)^*$

unnecessary complications in the notation, we write  $a\xi$  instead of  $\pi(a)\xi$  everywhere in this example. Finally, let  $D$  be the Dirac operator on  $M$ .

Then, the geodesic distance  $\text{dist}_g$  on  $M$  satisfies the following property:

$$\text{dist}_g(x, y) = \sup\{|a(x) - a(y)| : a \in C(M) \text{ s.t. } \|[D, a]\| \leq 1\} = \text{dist}_C(\delta_x, \delta_y).$$

*Proof.* Denote by  $\gamma(v)\zeta$  the Clifford multiplication of  $\zeta \in S_x$  by  $v \in T_x M$ . Since  $D$  is a differential operator of order 1, one easily checks that

$$([D, a]\xi)_x = \gamma((\text{grad } a)_x)\xi_x, \quad \xi \in \mathfrak{H},$$

for all  $a \in C^1(M)$ , so that

$$\|[D, a]\| = \|\text{grad } a\|_{L^\infty(M)}.$$

Hence

$$\text{dist}_C(\delta_x, \delta_y) = \sup_{\text{Lip}(a) \leq 1} |a(x) - a(y)|,$$

and

$$\text{dist}_g(x, y) \leq \text{dist}_C(\delta_x, \delta_y) \leq \text{dist}_g(x, y).$$

The upper bound is obvious by definition of the Lipschitz constant; as for the lower bound, it suffices to pick the function  $a(z) := \text{dist}_g(z, y)$ .  $\square$

**Example 2: the word length on a discrete group as a Fredholm module.**

Let  $\Gamma$  be a discrete group, with reduced  $C^*$ -algebra  $C_{red}^*(\Gamma)$  defined as the  $C^*$ -algebra generated by the left regular representation  $\lambda$  on  $\mathfrak{H} := \ell^2(\Gamma)$ . (We recall that  $(\lambda(g)\xi)_h := \xi_{g^{-1}h}$  for each  $\xi = (\xi_h)_{h \in \Gamma} \in \ell^2(\Gamma)$ .) Let  $L : \Gamma \rightarrow \mathbf{R}_+$  be a length function (for instance the word length with respect to a system of generators of  $\Gamma$ ). In other words, we assume that

$$L(1) = 0, \quad L(g^{-1})L(g), \quad L(gh) \leq L(g) + L(h), \quad g, h \in \Gamma.$$

Assume that  $L(g) \rightarrow +\infty$  as  $g \rightarrow \infty$ , and set

$$D\xi := (L(g)\xi_g)_{g \in \Gamma} \quad \text{for all } \xi = (\xi_g)_{g \in \Gamma} \in \ell^2(\Gamma).$$

Then  $(\mathfrak{H}, D)$  is an unbounded Fredholm module on  $C_{red}^*(\Gamma)$ . As in the preceding example, we seek to compute  $\|[D, a]\|$  for  $a \in C_{red}^*(\Gamma)$ . One finds that

$$\|[D, \lambda(g)]\| = L(g).$$

(Indeed, one has

$$(\lambda(g)D\lambda(g^{-1})\xi)_h = (D\lambda(g^{-1})\xi)_{g^{-1}h} = L(g^{-1}h)(\lambda(g^{-1})\xi)_{g^{-1}h} = L(g^{-1}h)\xi_h,$$

so that

$$((\lambda(g)D\lambda(g^{-1}) - D)\xi)_h = (L(g^{-1}h) - L(h))\xi_h,$$

and

$$\|[\lambda(g), D]\| = \|\lambda(g)D\lambda(g^{-1}) - D\| = \sup_{h \in \Gamma} |L(g^{-1}h) - L(h)| = L(g).$$

Connes' distance in noncommutative geometry is defined via an analogue of the Kantorovich(-Rubinstein) duality formula. Whether there exists a formula involving a notion of noncommutative coupling or transport plan seems to be an open question at the time of this writing — see however an interesting contribution to this problem by D'Andrea and Martinetti [25].

In spite of its great interest, the Connes distance is not exactly the quantum analogue of the Wasserstein distance which we are looking for. The first example presented above suggests that the Connes distance is a noncommutative analogue



of the classical distance on the space of positions, and not of a phase space distance. Moreover, the Connes distance is clearly an analogue of the Monge, or Wasserstein distance of exponent 1, instead of the Wasserstein distance of exponent 2, for which the optimal transport problem seems to have more structure (in particular by the Knott-Smith or the Brenier theorems).

**1.4. A Quantum Analogue of  $\mathcal{W}_2$ .** This section gathers together analogues of the Wasserstein distance  $\mathcal{W}_2$  in the quantum setting introduced in [37, 38]. The presentation given here is closer in spirit to [40].

Other approaches to the problem of generalizing  $\mathcal{W}_2$  to the quantum setting have been proposed by other authors: see for instance [26] and the earlier reference [65].

**1.4.1. Transport Cost.** The first step in extending the Wasserstein distance  $\mathcal{W}_2$  to the quantum setting is obviously to find some appropriate definition of the transport cost.

In classical mechanics, the phase space coordinates of a point particle are its position  $q \in \mathbf{R}^d$  and its momentum  $p \in \mathbf{R}^d$ .

In quantum mechanics, these coordinates, or functions thereof, must be replaced with appropriate operators. This procedure — associating operators on a Hilbert space to functions of the classical phase space coordinates — is called “quantization”.

The simplest quantization procedure

(i) associates to each function of the position variable only a multiplication operator

$$a(q) \rightarrow \text{multiplication by } a(y) \text{ in } L^2(\mathbf{R}_y^d);$$

(ii) associates to the classical momentum variable  $p$  the momentum operator

$$p \mapsto -i\hbar\nabla_y \text{ viewed as an unbounded self-adjoint operator on } L^2(\mathbf{R}_y^d).$$

The classical transport cost is a function on the Cartesian product of the phase space with itself which is the square Euclidean distance from the phase space point  $(x, \xi)$  to the phase space point  $(q, p)$  in  $\mathbf{R}^d \times \mathbf{R}^d$ , i.e.

$$|x - q|^2 + |\xi - p|^2.$$

The most natural thing to do is to quantize this expression in the variables  $(q, p)$  to measure the cost of transporting a particle from the classical phase space point  $(x, \xi)$  to a quantum state with position  $y$  and momentum  $-i\hbar\nabla_y$  (whatever it means). This leads to the

**Classical-to-quantum** transport cost, which is an operator on  $L^2(\mathbf{R}_y^d)$ :

$$c_\hbar(x, \xi) := |x - y|^2 + |\xi + i\hbar\nabla_y|^2.$$

This is “the” quantization<sup>8</sup> in  $(y, \eta)$  of  $(q, p) \mapsto |x - q|^2 + |\xi - p|^2$ .

One can also quantize this expression in both variables  $(x, \xi)$  and  $(q, p)$  to measure the cost of transporting a quantum particle from a quantum state with position  $x$  and momentum  $-i\hbar\nabla_x$  to a quantum state with position  $y$  and momentum  $-i\hbar\nabla_y$ . This leads to the

**Quantum-to-quantum** transport cost, which is an operator on  $L^2(\mathbf{R}_x^d \times \mathbf{R}_y^d)$ :

$$C_\hbar := |x - y|^2 - \hbar^2(\nabla_x - \nabla_y) \cdot (\nabla_x - \nabla_y).$$

---

<sup>8</sup>In truth, there is more than one quantization procedure; we shall see later in this lecture a notion of “Toeplitz quantization”, which in this case would give a slightly different result, with a difference of order  $O(\hbar)$ .

This is “the” quantization of the function  $(x, \xi, q, p) \mapsto |x - q|^2 + |\xi - p|^2$ .

Notice that  $c_{\hbar}(x, \xi)$  is the Hamiltonian of the harmonic oscillator, shifted in phase space by  $(x, \xi)$ , while  $C_{\hbar}$  is the Hamiltonian of a harmonic oscillator in the variable  $x - y$ .

Observe that

$$c_{\hbar}(x, \xi) \geq d\hbar I_{\mathfrak{H}}, \quad \text{for all } x, \xi \in \mathbf{R}^d,$$

while

$$C_{\hbar} \geq 2d\hbar I_{\mathfrak{H} \otimes \mathfrak{H}},$$

where  $\mathfrak{H} = L^2(\mathbf{R}^d)$ . These lower bounds are implied by Heisenberg’s uncertainty inequalities.

The exercise below gathers together several important facts related to the Hamiltonian of the harmonic oscillator  $y^2 - \hbar^2 \partial_y^2$  on the real line.

**Quiz 6.** In this exercise  $\mathfrak{H} = L^2(\mathbf{R})$ . Set

$$\omega(x) := \pi^{-1/4} e^{-x^2/2},$$

and

$$a := \frac{1}{\sqrt{2}}(x + \partial_x), \quad a^* := \frac{1}{\sqrt{2}}(x - \partial_x).$$

(1) Find  $\text{Ker}(a)$ , and prove that  $\text{Ker}(a) \subset \mathfrak{H}$ . Compute the commutator

$$[a, a^*] = aa^* - a^*a$$

together with the operators  $aa^*$  and  $a^*a$ .

Set  $\mathcal{V} := \{\psi \in H^1(\mathbf{R}) \text{ s.t. } y \mapsto y\psi(y) \in L^2(\mathbf{R})\}$ .

(2) Find

$$\inf_{\psi \in \mathcal{V}, \|\psi\|_{L^2} = 1} \int_{\mathbf{R}} \overline{\psi(y)} (y^2 - \hbar^2 \partial_y^2) \psi(y) dy.$$

(3) Find

$$\inf_{\psi \in \mathcal{V}, \|\psi\|_{L^2} = 1} \left( \int_{\mathbf{R}} |y|^2 |\psi(y)|^2 dy \right)^{1/2} \left( \int_{\mathbf{R}} |\hbar \partial_y \psi(y)|^2 dy \right)^{1/2}.$$

(Hint: change  $\hbar$  in  $\epsilon\hbar$  in question (2) where  $\epsilon > 0$  is arbitrary, and conclude by a minimization argument in  $\epsilon > 0$ .)

Question (3) leads to an inequality which is a mathematical formulation of Heisenberg’s uncertainty principle (for its physical interpretation, see for instance §16 in [48] or chapter I.C.3 in [23]).

(4) Find the spectrum of  $a^*a$ . (Hint: find  $\text{Ker}(a^*a)$  by using question (1). Then, argue as in the solution of the following classical exercise in algebra: if  $\mathcal{A}$  is a unital algebra with unit denoted by 1, and if  $a, b \in \mathcal{A}$ , then  $1 - ab$  is invertible in  $\mathcal{A}$  iff  $1 - ba$  is invertible in  $\mathcal{A}$ . To solve this exercise, the idea is to guess a formula relating  $(1 - ab)^{-1}$  and  $(1 - ba)^{-1}$ , which can be done easily by writing  $(1 - x)^{-1}$  as a formal series in powers of  $x$ .)

(5) Compute  $a^* a f_n$  for each integer  $n \geq 0$ , where  $f_n := (a^*)^n \omega$ .

(6) Set  $\phi_n = f_n / \sqrt{n!}$ . Prove that  $(\phi_n)_{n \geq 0}$  is an orthonormal system of  $\mathfrak{H}$ .

(7) Prove that  $\phi_n = (-1)^n \omega H_n / \sqrt{2^n n!}$  for each integer  $n \geq 0$ , where

$$H_n = (-1)^n \omega^{-2} \partial_x^n \omega^2$$

is the  $n$ -th Hermite polynomial in the so-called “physical form”. What is the leading coefficient in  $H_n$ ?

(8) Prove that the orthonormal system  $(\phi_n)_{n \geq 0}$  is complete in  $\mathfrak{H}$ .

(9) Consider the Fourier transform scaled as follows:

$$\mathbf{F}\psi(\xi) := \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} \psi(x) e^{-i\xi x} dx.$$

Compute  $\mathbf{F}(\phi_n)$  for each integer  $n \geq 0$ .

1.4.2. *Finite Energy Density Operators.* Before going further in the definition of a quantum analogue of the Wasserstein distance  $\mathcal{W}_2$ , we need to define the analogue of  $\mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$ .

Using the basic quantization procedure given in the previous section, shows that the phase space Euclidean norm is transformed into the Hamiltonian of a quantum harmonic oscillator:

$$\underbrace{|q|^2 + |p|^2}_{\text{phase space Euclidean norm}} \rightarrow \underbrace{|x|^2 - \hbar^2 \Delta_x}_{\text{harmonic oscillator}}.$$

This suggests the following definition of a quantum analogue to  $\mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$  (the set of Borel probability measures on phase space with finite second order moments):

$$\mathcal{D}_2(\mathfrak{H}) := \{R \in \mathcal{D}(\mathfrak{H}) \text{ s.t. } \text{trace}_{\mathfrak{H}}(R^{\frac{1}{2}}(|x|^2 - \hbar^2 \Delta_x)R^{\frac{1}{2}}) < \infty\}.$$

In other words,  $\mathcal{D}_2(\mathfrak{H})$  is the set of density operators with finite energy for the quantum harmonic oscillator.

If  $\{\psi_n \in L^2(\mathbf{R}^d, |x|^2 dx) \cap H^1(\mathbf{R}^d)\}$  is an orthonormal system in  $\mathfrak{H} = L^2(\mathbf{R}^d)$ , then

$$R = \sum_{n \geq 1} \rho_n |\psi_n\rangle\langle\psi_n| \in \mathcal{D}_2(\mathfrak{H}) \iff \begin{cases} \rho_n \geq 0 \text{ and } \sum_{n \geq 1} \rho_n = 1, \\ \sum_{n \geq 1} \rho_n (\|x\psi_n\|_{\mathfrak{H}}^2 + \hbar^2 \|\nabla\psi_n\|_{\mathfrak{H}}^2) < \infty. \end{cases}$$

The very simple verification of this statement is left to the reader.

1.4.3. *Couplings.* We already know the notion of coupling between two probability densities  $\mu$  and  $\nu$  belonging to  $\mathcal{P}(\mathbf{R}^d \times \mathbf{R}^d)$ : the set of such couplings is denoted by  $\mathcal{C}(\mu, \nu)$ .

Similarly, we define the set of **couplings of 2 quantum density operators**  $R, S \in \mathcal{D}(\mathfrak{H})$ :

$$\mathcal{C}(R, S) := \{T \in \mathcal{D}(\mathfrak{H} \otimes \mathfrak{H}) \text{ s.t. } \text{trace}_{\mathfrak{H}}(T(A \otimes I + I \otimes B)) = \text{trace}_{\mathfrak{H}}(RA + SB), \\ \text{for all } A, B \in \mathcal{L}(\mathfrak{H})\}.$$

The condition involving the test operators  $A, B \in \mathcal{L}(\mathfrak{H})$  can be equivalently replaced by conditions on partial traces:

$$\text{trace}_2(T) = R \quad \text{and} \quad \text{trace}_1(T) = S.$$

Finally, we define a notion of **coupling of a classical probability density  $f$  and a quantum density operator  $R$** . Let  $f(x, \xi)$  be a probability density on  $\mathbf{R}^{2d}$  and let  $R \in \mathcal{D}(\mathfrak{H})$ . A coupling of  $f$  and  $R$  is a measurable<sup>9</sup> operator-valued

<sup>9</sup>Since  $f \in L^1(\mathbf{R}^d \times \mathbf{R}^d)$  and  $\text{trace}_{\mathfrak{H}}(Q(x, \xi)) = \|Q(x, \xi)\|_1 = f(x, \xi) < \infty$  for a.e.  $(x, \xi) \in \mathbf{R}^d \times \mathbf{R}^d$ , one has  $Q(x, \xi) \ni \mathcal{L}^1(\mathfrak{H})$  for a.e.  $(x, \xi) \in \mathbf{R}^d \times \mathbf{R}^d$ . Since  $\mathcal{L}^1(\mathfrak{H})$  is separable, the map  $Q$  is weakly measurable if and only if it is strongly measurable by the Pettis Theorem (see Theorem 2 in chapter II of [28]). In other words, there are no measurability issues with  $Q$ .

map  $Q$  defined on  $\mathbf{R}^d \times \mathbf{R}^d$ , with values in  $\mathcal{L}(\mathfrak{H})$ , satisfying the following properties:

$$\begin{aligned} \mathbf{R}^{2d} \ni (x, \xi) &\mapsto Q(x, \xi) = Q(x, \xi)^* \in \mathcal{L}(\mathfrak{H}) \text{ s.t. } Q(x, \xi) \geq 0 \text{ a.e.} \\ \text{trace}_{\mathfrak{H}}(Q(x, \xi)) &= f(x, \xi) \text{ a.e.,} \quad \text{and} \quad \int_{\mathbf{R}^{2d}} Q(x, \xi) dx d\xi = R. \end{aligned}$$

In this case again, the set of couplings of  $f$  with  $R$  will be denoted by  $\mathcal{C}(f, R)$ .

Here are trivial examples of couplings (usually not the most clever couplings in optimal transport).

**Examples of couplings.**

- (a) For all  $R, S \in \mathcal{D}(\mathfrak{H})$ , the tensor product  $R \otimes S$  belongs to  $\mathcal{C}(R, S)$ .  
 (b) For each  $f$  probability density on  $\mathbf{R}^{2d}$ , and each density operator  $R$  on  $\mathfrak{H}$ , the operator-valued map

$$fR = f \otimes_{\mathbf{C}} R: (x, \xi) \mapsto f(x, \xi)R$$

belongs to  $\mathcal{C}(f, R)$ .

In particular

$$\mathcal{C}(R, S) \neq \emptyset \quad \text{and} \quad \mathcal{C}(f, R) \neq \emptyset.$$

(We have already observed that the set of couplings of two probability measures is never empty, since it always contains the tensor product of these two measures.)

1.4.4. *Extending the Wasserstein Distance to  $\mathfrak{D} := \mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d) \cup \mathcal{D}_2(\mathfrak{H})$ .* We are now ready to define the most important object in these lectures, namely the extension of the Wasserstein  $\mathcal{W}_2$  distance to the (disjoint) union of the sets of classical and quantum densities.

First we define the (disjoint) union of the set of (classical) Borel probability measures on phase space with finite second order moments, and of the set of finite energy (quantum) density operators:

$$\mathfrak{D} := \mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d) \cup \mathcal{D}_2(\mathfrak{H}).$$

**Definition of  $\mathfrak{d}$  on  $\mathfrak{D} \times \mathfrak{D}$ .**

- (1) For each  $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^{2d})$ , set

$$\mathfrak{d}(\mu, \nu) := \mathcal{W}_2(\mu, \nu).$$

- (2) For each phase space probability density  $f$  such that  $f(x, \xi) dx d\xi \in \mathcal{P}_2(\mathbf{R}^{2d})$  and each  $R \in \mathcal{D}_2(\mathfrak{H})$ , set

$$\mathfrak{d}(f, R) := \inf_{Q \in \mathcal{C}(f, R)} \left( \int_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q(x, \xi)^{\frac{1}{2}} c_{\hbar}(x, \xi) Q(x, \xi)^{\frac{1}{2}}) dx d\xi \right)^{\frac{1}{2}}.$$

- (3) For each  $R, S \in \mathcal{D}_2(\mathfrak{H})$ , set

$$\mathfrak{d}(R, S) := \inf_{T \in \mathcal{C}(R, S)} \left( \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(T^{\frac{1}{2}} C_{\hbar} T^{\frac{1}{2}}) \right)^{\frac{1}{2}}.$$

**Remark.** For all  $f(x, \xi) dx d\xi \in \mathcal{P}_2(\mathbf{R}^{2d})$  and all  $R, S \in \mathcal{D}_2(\mathfrak{H})$ , one has

$$\mathfrak{d}(f, R) \geq \sqrt{d\hbar} \quad \text{and} \quad \mathfrak{d}(R, S) \geq \sqrt{2d\hbar}.$$

In particular  $\mathfrak{d}(R, R) > 0$ , so that  $\mathfrak{d}$  is not a bona fide metric on  $\mathfrak{D}$ .

**Quiz 7.** Let  $\mathfrak{H}$  be a (complex) separable Hilbert space.

- (1) Let  $A, B \in \mathcal{L}(\mathfrak{H})$ . Prove that

$$A^*B + B^*A \leq |A|^2 + |B|^2.$$

(We recall that  $|A|^2 := A^*A$ .)

(2) Prove that, for all  $\epsilon > 0$  and all  $A, B \in \mathcal{L}(\mathfrak{H})$ , one has

$$A^*B + B^*A \leq \epsilon|A|^2 + \frac{1}{\epsilon}|B|^2.$$

(3) Prove that, for all  $\epsilon > 0$  and all  $A, B \in \mathcal{L}(\mathfrak{H})$ , one has

$$|A + B|^2 \leq (1 + \epsilon)|A|^2 + (1 + \frac{1}{\epsilon})|B|^2.$$

(4) Let  $A, B$  be Hilbert-Schmidt operators on  $\mathfrak{H}$ . Deduce from (2) that

$$|\text{trace}_{\mathfrak{H}}(B^*A)| \leq \sqrt{\text{trace}_{\mathfrak{H}}(|A|^2)}\sqrt{\text{trace}_{\mathfrak{H}}(|B|^2)}.$$

(5) Let  $R = R^* \geq 0$  be a trace-class operator on  $\mathfrak{H}$ . Prove that

$$|\text{trace}_{\mathfrak{H}}(B^*AR)| \leq \sqrt{\text{trace}_{\mathfrak{H}}(|A|^2R)}\sqrt{\text{trace}_{\mathfrak{H}}(|B|^2R)}.$$

(6) Let  $f$  be a convex function on  $\mathbf{R}$  (in particular  $f \in C(\mathbf{R})$ ), let  $R \in \mathcal{D}(\mathfrak{H})$ , and let  $A = A^* \in \mathcal{L}(\mathfrak{H})$ . Prove that

$$f(\text{trace}_{\mathfrak{H}}(AR)) \leq \text{trace}_{\mathfrak{H}}(f(A)R).$$

(Hint: prove that, for each  $z, m \in \mathbf{R}$  and for all  $\lambda \in [f'_g(m), f'_d(m)]$ , one has the inequality  $f(z) \geq f(m) + \lambda(z - m)$ . Using the spectral measure  $(\xi|E(dz)\xi)$  of  $A$ , where  $\xi \in \mathfrak{H}$  and  $E$  is the spectral decomposition<sup>10</sup> of  $A$ , prove that  $f(A) \geq f(m)I + \lambda(A - mI)$ . Conclude by choosing  $m$  appropriately.)

(7) How should one modify (1) if  $A$  and  $B$  are unbounded operators on  $\mathfrak{H}$ ?

(8) Prove that

$$\mathfrak{d}(f, R) + \mathfrak{d}(R, S) < \infty$$

for each probability density  $f$  with finite second order moments on  $\mathbf{R}^d \times \mathbf{R}^d$ , and all  $R, S \in \mathcal{D}_2(\mathfrak{H})$ .

In practice, the “pseudometric”  $\mathfrak{d}$  is not easy to compute, except in a few cases. But before discussing these cases, we need to return to the question of “quantization” — i.e. associating an operator on  $L^2(\mathbf{R}^d)$  to a function on phase space (i.e.  $\mathbf{R}^d \times \mathbf{R}^d$ ).

1.4.5. *Toeplitz Operators and Husimi Transform.* We begin with the definition of a quantum analogue of the Dirac mass at the phase space point  $(q, p) \in \mathbf{R}^d \times \mathbf{R}^d$ .

**Gaussian wave packet (Schrödinger coherent state).** For all  $q, p \in \mathbf{R}^d$ , set

$$|q, p\rangle(x) := (\pi\hbar)^{-d/4} \exp(-\frac{1}{2\hbar}|x - q|^2) \exp(\frac{i}{\hbar}p \cdot (x - \frac{q}{2})).$$

This is a plane wave oscillating at frequency  $|p|/\hbar$  in the direction  $p/|p|$ , modulated by a Gaussian envelope of width  $O(\sqrt{\hbar})$ , centered at the position  $q$ . Therefore, the oscillating profile of the wave function  $|q, p\rangle$  encodes the momentum vector  $p$ , while the envelope of the oscillations encodes the position  $q \in \mathbf{R}^d$ .

**Toeplitz map.** To  $m$ , a (complex) Radon measure on the phase space  $\mathbf{R}^d \times \mathbf{R}^d$ , one associates the operator

$$\mathcal{T}[m] := \int_{\mathbf{R}^d} |q, p\rangle\langle q, p| m(dqdp).$$

The form-domain of  $\mathcal{T}[m]$  is the set of  $\phi \in \mathfrak{H}$  such that the function  $(q, p) \mapsto \langle q, p|\phi\rangle$  belongs to  $L^2(\mathbf{R}^{2d}; m)$ .

Observe that  $|q, p\rangle\langle q, p|$  is a self-adjoint positive operator (specifically, a rank-1 projection in  $\mathfrak{H}$ ). Therefore, if the measure  $m$  is real-valued, the operator  $\mathcal{T}[m]$  (with some appropriate domain) is expected to be self-adjoint, and if  $m$  is a positive

<sup>10</sup>See for instance chapter 12 in [57].

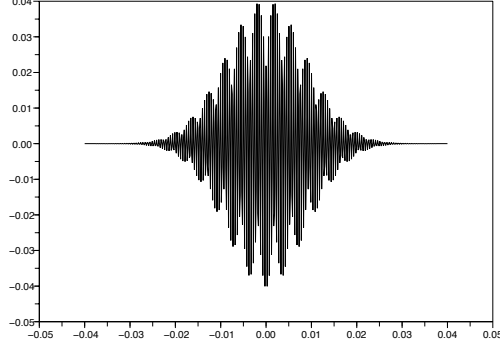
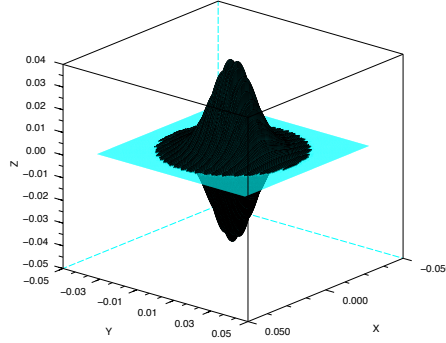


FIGURE 1. Oscillating structure of a Gaussian wave-packet

FIGURE 2. With  $\hbar = 8 \cdot 10^{-5}$ , plot of  $Z = \text{real part of the coherent state}$  centered at  $q = (0, 0)$  with momentum  $p = (1, 0)$  with space variable  $(X, Y) \in \mathbf{R}^2$ 

measure on phase space, the associated Toeplitz operator  $\mathcal{T}[m]$  is expected to be a positive operator on  $\mathfrak{H}$  (possibly unbounded, with some appropriate domain).

**Basic properties of the Toeplitz map.**

(1) The set of Schrödinger coherent states is a resolution of the identity:

$$\mathcal{T}[1] = \int_{\mathbf{R}^{2d}} |q, p\rangle \langle q, p| dq dp = (2\pi\hbar)^d I_{\mathfrak{H}} .$$

(2) The image by the Toeplitz map  $\mathcal{T}$  of the set of Borel phase space probability measure on  $\mathbf{R}^d \times \mathbf{R}^d$  is included in  $\mathcal{D}(L^2(\mathbf{R}^d))$ :

$$m \in \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^d) \implies \mathcal{T}[m] \in \mathcal{D}(\mathfrak{H}), \quad \text{with } \mathfrak{H} = L^2(\mathbf{R}^d) .$$

(3) One has

$$\mathcal{T}[q] = (2\pi\hbar)^d x \text{ (position operator),}$$

while

$$\mathcal{T}[p] = (2\pi\hbar)^d (-i\hbar\nabla_x) \text{ (momentum operator).}$$

(4) If  $f$  is a quadratic form on  $\mathbf{R}^d$ , then

$$\begin{cases} \mathcal{T}[f(q)] = (2\pi\hbar)^d \left( f(x) + \frac{1}{4}\hbar(\Delta f)I_{\mathfrak{H}} \right), \text{ and} \\ \mathcal{T}[f(p)] = (2\pi\hbar)^d \left( f(-i\hbar\nabla_x) + \frac{1}{4}\hbar(\Delta f)I_{\mathfrak{H}} \right). \end{cases}$$

**Quiz 8.** Prove the statements (1)-(4) above. (Hint: recall the formulas for moments of order  $\leq 2$  of Gaussian distributions, together with the oscillating integrals

$$\frac{1}{2\pi} \int_{\mathbf{R}^d} e^{ip(x-y)} dp = \delta_0(x-y), \quad \frac{1}{2\pi} \int_{\mathbf{R}^d} p_j e^{ip(x-y)} dp = -i\partial_{x_j} \delta_0(x-y),$$

and

$$\frac{1}{2\pi} \int_{\mathbf{R}^d} p_j p_k e^{ip(x-y)} dp = -\partial_{x_j} \partial_{x_k} \delta_0(x-y), \quad j, k = 1, \dots, d$$

— which are to be understood in the sense of tempered distributions. For instance, the first formula above is equivalent to the Fourier inversion formula on the set of tempered distributions.)

The Toeplitz map associates an operator on  $\mathfrak{H} = L^2(\mathbf{R}^d)$  to a function on the phase space  $\mathbf{R}^d \times \mathbf{R}^d$ . Conversely, given an operator on  $\mathfrak{H}$ , we seek to associate a function on phase space. There are various ways of doing this, one of which is the Husimi transform.

**Husimi Transform.** To  $T \in \mathcal{L}(\mathfrak{H})$ , one associates its Husimi transform

$$\mathcal{H}[T](q, p) := \frac{1}{(2\pi\hbar)^d} \langle q, p | T | q, p \rangle.$$

This definition can be extended to all unbounded operators on  $T$  such that the Gaussian wave packet  $|q, p\rangle$  belongs to the form domain of  $T$  for each  $q, p \in \mathbf{R}^d$ .

**Basic properties of the Husimi transform.**

(1) For each  $T \in \mathcal{L}(\mathfrak{H})$ ,

$$T = T^* \implies \mathcal{H}[T](q, p) \in \mathbf{R}, \text{ and } T \geq 0 \implies \mathcal{H}[T] \geq 0.$$

(2) The Husimi transform is an “almost inverse” of the Toeplitz map: for each  $m \in \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^d)$

$$\mathcal{H}[\mathcal{T}[m]] = e^{\frac{\hbar}{2}\Delta_{q,p}} m,$$

since

$$\langle q, p | q', p' \rangle = e^{-\frac{1}{4\hbar}(|q-q'|^2 + |p-p'|^2)} e^{-\frac{i}{\hbar}(p \cdot q' - q \cdot p')}.$$

(3) One has

$$\mathcal{H}[I] = (2\pi\hbar)^{-d},$$

while

$$\begin{cases} \mathcal{H}[f(x)](q, p) = (2\pi\hbar)^{-d} \left( I + \frac{1}{4}\hbar\Delta \right) f(q), \text{ and} \\ \mathcal{H}[f(-i\hbar\nabla_x)](q, p) = (2\pi\hbar)^{-d} \left( I + \frac{1}{4}\hbar\Delta \right) f(p). \end{cases}$$

(4) One has

$$\text{trace}_{\mathfrak{H}}(R^* \mathcal{T}[f]) = (2\pi\hbar)^d \iint_{\mathbf{R}^d \times \mathbf{R}^d} \overline{\mathcal{H}[R](q, p)} f(q, p) dq dp.$$

**Quiz 9.** Prove the statements (1)-(4) above.

An important property of the Husimi transform is that it is a one-to-one transformation. In other words, the Husimi transform of an operator specifies it completely.

**Quiz 10.** Set  $\mathfrak{H} := L^2(\mathbf{R}^d)$ .

(1) Let  $R \in \mathcal{D}_2(\mathfrak{H})$ . Prove that  $\mathcal{H}[R]$  is a probability density, and compute

$$\iint_{\mathbf{R}^d \times \mathbf{R}^d} (|q|^2 + |p|^2) \mathcal{H}[R](q, p) dq dp.$$

(2) Let  $R, S \in \mathcal{D}_2(\mathfrak{H})$ , and assume that  $\mathcal{H}[R] = \mathcal{H}[S]$ . Prove that  $R = S$ . (Hint: let  $r \equiv r(y, y')$  be an integral kernel of  $R$ . Set

$$J(x, \xi) = \iint_{\mathbf{R}^d \times \mathbf{R}^d} r(y, y') e^{-(|y|^2 + |y'|^2)/2\hbar} e^{x \cdot (y+y') - i\xi \cdot (y-y')/\hbar} dy dy'.$$

Prove that  $J$  extends as a holomorphic function on  $\mathbf{C}^d \times \mathbf{C}^d$ , and therefore is uniquely determined by its restriction to  $\mathbf{R}^d \times \mathbf{R}^d$ . Conclude by (a) computing the formula relating  $\mathcal{H}[R]$  to  $J$ , and (b) by computing the integral kernel  $r$  of  $R$  in terms of  $J$ .)

Here is another way of associating a function on the phase space  $\mathbf{R}^d \times \mathbf{R}^d$  to an operator on  $L^2(\mathbf{R}^d)$ .

**Quiz 11.** Set  $\mathfrak{H} := L^2(\mathbf{R}^d)$ . To each  $A \in \mathcal{L}^1(\mathfrak{H})$  with integral kernel  $a \equiv a(x, y)$  such that  $z \mapsto a(x+z, x)$  belongs to  $C_b(\mathbf{R}^d; L^1(\mathbf{R}^d))$  (see Quiz 4), we associate its Wigner transform

$$W[A](x, \xi) := \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} a(x + \frac{1}{2}\hbar y, x - \frac{1}{2}\hbar y) e^{-i\xi \cdot y} dy$$

(where the integral above is to be understood as the partial Fourier transform of the continuous bounded function  $y \mapsto a(x + \frac{1}{2}\hbar y, x - \frac{1}{2}\hbar y)$  with values in  $L^1(\mathbf{R}_x^d)$ , which is therefore a tempered distribution).

(1) Prove that

$$\overline{W[A]} = W[A^*],$$

and that, for each  $A, B \in \mathcal{L}^1(\mathfrak{H})$

$$\text{trace}_{\mathfrak{H}}(A^* B) = (2\pi\hbar)^d \iint_{\mathbf{R}^d \times \mathbf{R}^d} \overline{W[A](x, \xi)} W[B](x, \xi) dx d\xi,$$

and

$$\|A\|_2 = (2\pi\hbar)^{d/2} \|W[A]\|_{L^2(\mathbf{R}^d \times \mathbf{R}^d)}.$$

Prove that the Wigner transform has a unique extension to  $\mathcal{L}^2(\mathfrak{H})$ .

(2) Prove that, for each  $A \in \mathcal{L}^1(\mathfrak{H})$ ,

$$\text{trace}_{\mathfrak{H}}(A) = \iint_{\mathbf{R}^d \times \mathbf{R}^d} W[A](x, \xi) dx d\xi.$$

(3) Let  $t \mapsto R(t)$  be a time-dependent density operator, solution of the von Neumann equation

$$i\hbar \partial_t R(t) = [-\frac{1}{2}\hbar^2 \Delta + V, R(t)].$$

Prove that  $W[R(t)]$  is a solution of the Wigner equation

$$(\partial_t + \xi \cdot \nabla_x) W[R(t)](x, \xi) + \Theta[V] W[R(t)](x, \xi) = 0,$$

where  $\Theta[V]$  is the linear operator with distribution kernel

$$\frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \frac{1}{i\hbar} (V(x + \frac{1}{2}\hbar y) - V(x - \frac{1}{2}\hbar y)) e^{iy \cdot (\eta - \xi)} dy.$$

Prove that

$$\Theta[V] = -\nabla V(x) \cdot \nabla_\xi$$

in the case where  $V$  is a polynomial of degree 2.



In that case (for a quadratic potential  $V$ ), the Wigner equation coincides with the classical Liouville equation on the phase space  $\mathbf{R}^d \times \mathbf{R}^d$

$$\partial_t W[R(t)](x, \xi) + \left\{ \frac{1}{2} |\xi|^2 + V(x), W[R(t)](x, \xi) \right\} = 0.$$

(4) However, one cannot think of  $W[A]$  as a distribution function as in the kinetic theory of gases. Indeed

$$0 \leq A = A^* \in \mathcal{L}^1(\mathfrak{H}) \text{ does not imply } W[A] \geq 0.$$

To see this, compute  $W[|\psi\rangle\langle\psi|](0, 0)$  where  $\psi(x) = \sqrt{2}\pi^{-1/4}xe^{-x^2/2}$  for  $x \in \mathbf{R}$ .

(5) The relation between the Wigner and the Husimi transform is given by the following formula: for each  $A \in \mathcal{L}^1(\mathfrak{H})$ , one has

$$\mathcal{H}[A](x, \xi) = \exp\left(\frac{\hbar}{4}\Delta_{x, \xi}\right)W[A](x, \xi).$$

(6) Let  $\psi \in L^2(\mathbf{R})$  satisfy  $\|\psi\|_{L^2(\mathbf{R})} = 1$  and  $W[|\psi\rangle\langle\psi|] \geq 0$ . Prove that there exist  $q_0, p_0 \in \mathbf{R}^d$  and  $u \in \mathbf{C}$  with  $|u| = 1$  such that  $\psi = u|q_0, p_0\rangle$ . (Hint: prove that

$$F(z) := \int_{\mathbf{R}} \psi(x)e^{-\frac{1}{2}x^2 - zx} dx$$

defines an entire function on  $\mathbf{C}$ , that

$$0 < |F(z)|^2 \leq Ce^{\Re(z)^2}, \quad z \in \mathbf{C},$$

for some constant  $C > 0$ . Conclude by Hadamard's theorem<sup>11</sup> that  $F(z) = e^{g(z)}$  where  $g$  is a polynomial of degree 2.)

1.4.6. *Explicit Computations/Estimates.* We have gathered together in this section several useful explicit computations, or bounds, on the ‘‘pseudometric’’  $\mathfrak{d}$ .

**Theorem 1.**

(1) For all  $f, g$  probability densities on  $\mathbf{R}^{2d}$  with finite 2nd order moments,

$$\begin{aligned} \mathfrak{d}(\mathcal{T}[f], \mathcal{T}[g])^2 &\leq \mathcal{W}_2(f, g)^2 + 2d\hbar, & \mathfrak{d}(\mathcal{T}[f], \mathcal{T}[f]) &= \sqrt{2d\hbar}, \\ \mathfrak{d}(f, \mathcal{T}[g])^2 &\leq \mathcal{W}_2(f, g)^2 + d\hbar, & \mathfrak{d}(f, \mathcal{T}[f]) &= \sqrt{d\hbar}. \end{aligned}$$

(2) For all  $R, S \in \mathcal{D}_2(\mathfrak{H})$  and all probability density  $f$  on  $\mathbf{R}^{2d}$  with finite 2nd order moments,

$$\begin{aligned} \mathcal{W}_2(\mathcal{H}[R], \mathcal{H}[S])^2 &\leq \mathfrak{d}(R, S)^2 + 2d\hbar, \\ \mathcal{W}_2(f, \mathcal{H}[R])^2 &\leq \mathfrak{d}(f, R)^2 + d\hbar. \end{aligned}$$

(3) Moreover, if  $\text{rank}(R) = 1$ , then

$$\begin{aligned} \mathfrak{d}(R, S) &= \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}((R \otimes S)^{\frac{1}{2}} C_{\hbar}(R \otimes S)^{\frac{1}{2}})^{\frac{1}{2}}, \quad \text{and} \\ \mathfrak{d}(f, R) &= \left( \int_{\mathbf{R}^{2d}} f(x, \xi) \text{trace}_{\mathfrak{H}}(R^{\frac{1}{2}} c_{\hbar}(x, \xi) R^{\frac{1}{2}}) dx d\xi \right)^{\frac{1}{2}}. \end{aligned}$$

**Remark.** The second inequality in (1) can be recast as

$$\mathfrak{d}(f, \mathcal{T}[g])^2 \leq \mathfrak{d}(f, g)^2 + \mathfrak{d}(g, \mathcal{T}[g])^2$$

since

$$\mathfrak{d}(g, \mathcal{T}[g]) = \sqrt{d\hbar} = \min \mathfrak{d}.$$

This suggests that

(1) ‘‘the segment  $[g, \mathcal{T}[g]]$  is orthogonal to the set of classical densities’’, and

<sup>11</sup>See for instance chapter 5.3.2 in [1].

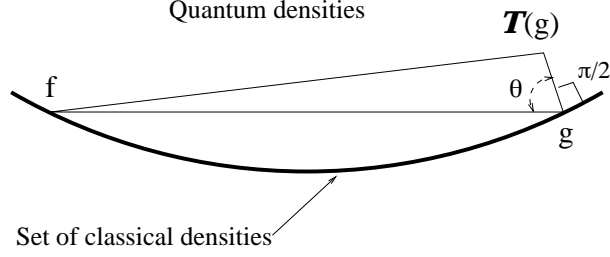


FIGURE 3. A geometric interpretation of Theorem 1 (1).

(2) the “angle”  $\theta$  between the “segment”  $[g, f]$  and the “segment”  $[g, \mathcal{T}(g)]$  in  $\mathfrak{D}_2$  is acute.

Therefore, one could think of  $\mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$  as a limit set (for the classical limit  $\hbar \rightarrow 0$ ) — or boundary — of  $\mathfrak{D}$ , and that the set  $\mathcal{D}(\mathfrak{H})$  of quantum densities lies on the “concave” side of the set of classical densities.

*Proof of Theorem 1.*

Proof of (1). Let  $\nabla\Phi$  (with  $\Phi$  convex) be the Brenier map pushing  $f$  to  $g$ . The optimal coupling of  $f$  and  $g$  for  $\mathcal{W}_2$  is

$$\Lambda := f(x, \xi) \delta_{\nabla\Phi(x, \xi)}(dyd\eta) dx d\xi.$$

Hence

$$\mathcal{T}[\Lambda] \in \mathcal{C}(\mathcal{T}[f], \mathcal{T}[g]) \quad \text{and} \quad (x, \xi) \mapsto f(x, \xi) \mathcal{T}[\delta_{\nabla\Phi(x, \xi)}] \in \mathcal{C}(f, \mathcal{T}[g]).$$

On the other hand (see Appendix B, and especially formulas (52)-(53), in [37])

$$\mathcal{H}[C_\hbar](q, p, q', p') = (2\pi\hbar)^{-2d} (|q - q'|^2 + |p - p'|^2 + 2d\hbar),$$

$$\mathcal{H}[c_\hbar(x, \xi)](q, p) = (2\pi\hbar)^{-d} (|x - q|^2 + |\xi - p|^2 + d\hbar).$$

Therefore

$$\begin{aligned} \mathfrak{d}(\mathcal{T}[f], \mathcal{T}[g])^2 &\leq \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(\mathcal{T}[\Lambda]^{\frac{1}{2}} C_\hbar \mathcal{T}[\Lambda]^{\frac{1}{2}}) \\ &= \int_{\mathbf{R}^{4d}} \underbrace{(|q - q'|^2 + |p - p'|^2 + 2d\hbar)}_{=(2\pi\hbar)^{2d} \overline{\mathcal{H}[C_\hbar](q, p)}} \Lambda(dq dp dq' dp') \\ &= \mathcal{W}_2(f, g)^2 + 2d\hbar, \end{aligned}$$

and

$$\begin{aligned} \mathfrak{d}(f, \mathcal{T}[g])^2 &\leq \int_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(\mathcal{T}[\delta_{\nabla\Phi(x, \xi)}]^{\frac{1}{2}} c_\hbar(x, \xi) \mathcal{T}[\delta_{\nabla\Phi(x, \xi)}]^{\frac{1}{2}}) f(x, \xi) dx d\xi \\ &= \int_{\mathbf{R}^{2d}} \underbrace{(|(x, \xi) - \nabla\Phi(x, \xi)|^2 + d\hbar)}_{=(2\pi\hbar)^d \overline{\mathcal{H}[c_\hbar(x, \xi)](\nabla\Phi(x, \xi))}} f(x, \xi) dx d\xi \\ &= \mathcal{W}_2(f, g)^2 + d\hbar. \end{aligned}$$

Proof of (2). Pick sequences  $a_n, b_n \in C_b(\mathbf{R}^{2d}; \mathbf{R})$  such that

$$\begin{aligned} a_n(q, p) + b_n(q', p') &\leq |q - q'|^2 + |p - p'|^2, \quad \text{and} \\ \underbrace{\int_{\mathbf{R}^{2d}} a_n(q, p) \mathcal{H}[R](q, p) dq dp}_{=(2\pi\hbar)^{-d} \text{trace}_{\mathfrak{H}}(\mathcal{T}[a_n]R)} &+ \underbrace{\int_{\mathbf{R}^{2d}} b_n(q', p') \mathcal{H}[S](q', p') dq' dp'}_{=(2\pi\hbar)^{-d} \text{trace}_{\mathfrak{H}}(\mathcal{T}[b_n]S)} \\ &\rightarrow \mathcal{W}_2(\mathcal{H}[R], \mathcal{H}[S])^2 \end{aligned}$$

as  $n \rightarrow \infty$ . That such sequences exist is a consequence of the Kantorovich duality formula for the Wasserstein distance  $\mathcal{W}_2$ .

On the other hand, for each  $T \in \mathcal{C}(R, S)$ , one has

$$\begin{aligned} &(2\pi\hbar)^{-d} \text{trace}_{\mathfrak{H}}(\mathcal{T}[a_n]R) + (2\pi\hbar)^{-d} \text{trace}_{\mathfrak{H}}(\mathcal{T}[b_n]S) \\ &= (2\pi\hbar)^{-d} \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(T^{\frac{1}{2}}(\mathcal{T}[a_n] \otimes I + I \otimes \mathcal{T}[b_n])T^{\frac{1}{2}}) \\ &= (2\pi\hbar)^{-2d} \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(T^{\frac{1}{2}} \mathcal{T}[a_n \otimes 1 + 1 \otimes b_n]T^{\frac{1}{2}}) \\ &\leq (2\pi\hbar)^{-2d} \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(T^{\frac{1}{2}} \mathcal{T}[|q - q'|^2 + |p - p'|^2]T^{\frac{1}{2}}). \end{aligned}$$

Now, one has (see the basic properties of the Toeplitz map and Quiz 8 above)

$$\mathcal{T}[|q - q'|^2 + |p - p'|^2] = (2\pi\hbar)^{2d} (C_\hbar + 2d\hbar I_{\mathfrak{H} \otimes \mathfrak{H}}).$$

Thus, for all  $T \in \mathcal{C}(R, S)$ , one has

$$\begin{aligned} \mathcal{W}_2(\mathcal{H}[R], \mathcal{H}[S])^2 &= \lim_{n \rightarrow \infty} (2\pi\hbar)^{-d} (\text{trace}_{\mathfrak{H}}(\mathcal{T}[a_n]R) + \text{trace}_{\mathfrak{H}}(\mathcal{T}[b_n]S)) \\ &\leq \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(T^{\frac{1}{2}}(C_\hbar + 2d\hbar I_{\mathfrak{H} \otimes \mathfrak{H}})T^{\frac{1}{2}}) \\ &= \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(T^{\frac{1}{2}} C_\hbar T^{\frac{1}{2}}) + 2d\hbar. \end{aligned}$$

Minimizing the r.h.s. in  $T \in \mathcal{C}(R, S)$  leads to

$$\mathcal{W}_2(\mathcal{H}[R], \mathcal{H}[S])^2 \leq \mathfrak{d}(R, S)^2 + 2d\hbar.$$

Proof of (3). We begin with a question of a rather fundamental nature in quantum mechanics.

**Question.** What is the structure of couplings for rank-1 density operators?

This question is answered by the following lemma.

**Lemma 2.** Let  $R \in \mathcal{D}(\mathfrak{H})$ . Then

$$\text{rank}(R) = 1 \implies \begin{cases} \mathcal{C}(f, R) = \{fR\}, & f \in \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^d), \\ \mathcal{C}(R, S) = \{R \otimes S\}, & S \in \mathcal{D}(\mathfrak{H}). \end{cases}$$

Obviously Lemma 2 implies statement (3) in Theorem 1. □

*Proof of Lemma 2.* Since  $\text{rank}(R) = 1$ , it is of the form  $R = |\phi\rangle\langle\phi|$ , with  $\|\phi\|_{\mathfrak{H}} = 1$ . We shall prove the second statement in the lemma. Let  $Q \in \mathcal{C}(R, S)$ . Then

$$\begin{aligned} \text{trace}(((I - R) \otimes I)Q((I - R) \otimes I)) &= \text{trace}(Q((I - R)^2 \otimes I)) \\ &= \text{trace}(Q((I - R) \otimes I)) = \text{trace}(R(I - R)) = 0, \end{aligned}$$

Since

$$((I - R) \otimes I)Q((I - R) \otimes I) \geq 0,$$

one has

$$((I - R) \otimes I)Q((I - R) \otimes I) = 0.$$

Next, we deduce from the Cauchy-Schwarz inequality that

$$\begin{aligned} & | \langle \psi_1 \otimes \psi_2 | (R \otimes I)Q((I - R) \otimes I) \psi'_1 \otimes \psi'_2 \rangle |^2 \\ & \leq \langle \psi'_1 \otimes \psi'_2 | ((I - R) \otimes I)Q((I - R) \otimes I) \psi'_1 \otimes \psi'_2 \rangle \\ & \quad \times \langle \psi_1 \otimes \psi_2 | (R \otimes I)Q(R \otimes I) \psi_1 \otimes \psi_2 \rangle. \end{aligned}$$

Hence

$$\begin{aligned} (R \otimes I)Q((I - R) \otimes I) = 0 &= ((R \otimes I)Q((I - R) \otimes I))^* \\ &= ((I - R) \otimes I)Q(R \otimes I), \end{aligned}$$

Since we already know that  $((I - R) \otimes I)Q((I - R) \otimes I) = 0$ , this implies that

$$Q = (R \otimes I)Q(R \otimes I).$$

Therefore  $Q = R \otimes T$ , where

$$\langle \psi | T | \psi' \rangle := \langle \phi \otimes \psi | Q | \phi \otimes \psi' \rangle.$$

Finally,  $T = S$ , since, for all  $A \in \mathcal{L}(\mathfrak{H})$ , one has

$$\text{trace}(SA) = \text{trace}(Q(I \otimes A)) = \text{trace}((R \otimes T)(I \otimes A)) = \text{trace}(TA).$$

This proves the second statement in the lemma.  $\square$

**Quiz 12.** Complete the proof of Lemma 2: prove that  $\mathcal{C}(f, R) = \{f \otimes_{\mathbf{C}} R\}$ , in the case where  $f$  is a probability density on  $\mathbf{R}^d \times \mathbf{R}^d$  and  $R$  is a rank-1 density operator on  $\mathfrak{H} = L^2(\mathbf{R}^d)$ .

**Remark.** It is well known that, for all  $\mu \in \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^d)$  and all  $(q, p) \in \mathbf{R}^d \times \mathbf{R}^d$ , the set of couplings of  $\mu$  with  $\delta_{(q,p)}$  contains only one element:

$$\mathcal{C}(\mu, \delta_{(q,p)}) = \{\mu \otimes \delta_{(q,p)}\}.$$

Lemma 2 suggests that **all** rank-1 density operators are quantum analogues of the Dirac mass in phase space. Thus the Schrödinger equation governing the evolution of the wave function  $\psi(t, x)$  — or the von Neumann equation specialized to  $|\psi(t, \cdot)\rangle\langle\psi(t, \cdot)|$  — is the quantum analogue of Newton's second law of motion in classical mechanics, which can be viewed as the equation governing  $\delta_{(q(t), p(t))}$ , where  $q(t)$  and  $p(t)$  are respectively the position and the momentum of a moving classical particle. This analogy also explains why the wave function is a purely quantum object, which has no classical analogue. Indeed, if a classical analogue of the wave function existed, it could be thought of as a “square root” of the phase space Dirac measure  $\delta_{(q(t), p(t))}$ .

**Quiz 13: another proof of Theorem 1 (2).**

(1) Prove that

$$\mathcal{T}[|q - q'|^2 + |p - p'|^2] = (2\pi\hbar)^{2d}(C_h + 2d\hbar I_{\mathfrak{H} \otimes \mathfrak{H}}).$$

(2) For each  $T \in \mathcal{C}(R, S)$ , prove that

$$\begin{aligned} \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}} \left( T^{\frac{1}{2}} C_h T^{\frac{1}{2}} \right) + 2d\hbar &\geq \frac{1}{(2\pi\hbar)^d} \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}} \left( T \mathcal{T} \left[ \frac{|q - q'|^2 + |p - p'|^2}{1 + \epsilon|q - q'|^2 + \epsilon|p - p'|^2} \right] \right) \\ &= \int_{\mathbf{R}^{4d}} \mathcal{H}[T](q, p, q', p') \frac{|q - q'|^2 + |p - p'|^2}{1 + \epsilon|q - q'|^2 + \epsilon|p - p'|^2} dq dp dq' dp'. \end{aligned}$$

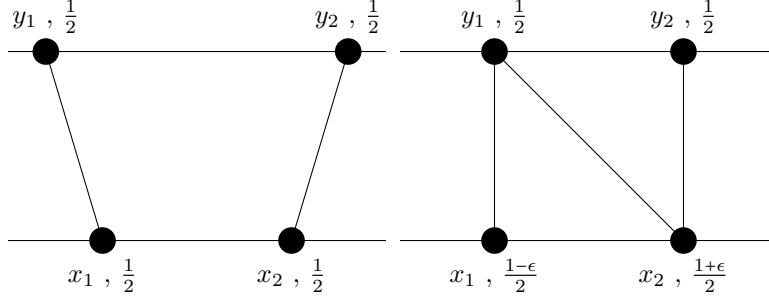


FIGURE 4. Left: equal masses; Right: unequal mass case

(3) Conclude by monotone convergence in the right-hand side of the inequality above as  $\epsilon \rightarrow 0^+$ , after observing that  $\mathcal{H}[T]$  is a coupling of  $\mathcal{H}[R]$  and  $\mathcal{H}[S]$ .

**1.5. Quantum Optimal Transport is Cheaper!** To conclude this first lecture, we shall study an example where  $\mathfrak{d}$  can be computed essentially explicitly, with interesting implications on the structure of optimal couplings. The material in this section — together with the somewhat provocative title — is taken from [19].

We begin with a simple lemma, which can be viewed as an amplification of statement (1) in Theorem 1.

**Lemma 3.** For  $\rho_1, \rho_2 \in \mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$  with optimal coupling  $\Pi$  for  $\mathcal{W}_2$ , one has

$$\begin{aligned} \mathfrak{d}(\mathcal{T}[\rho_1], \mathcal{T}[\rho_2])^2 &= \mathcal{W}_2(\rho_1, \rho_2)^2 + 2d\hbar \\ \iff \mathcal{T}[\Pi] \in \mathcal{C}(\mathcal{T}[\rho_1], \mathcal{T}[\rho_2]) &\text{ is a quantum optimal coupling for } \mathfrak{d}. \end{aligned}$$

The proof of this lemma is left to the reader as an easy exercise (see the proof of Theorem 1).

Here is an example of this kind of situation. With  $d = 1$  and  $0 < a < b$ , set

$$\mu := \frac{1}{2}(\delta_{(+a,0)} + \delta_{(-a,0)}) \quad \text{and} \quad \nu := \frac{1}{2}(\delta_{(+b,0)} + \delta_{(-b,0)}) \in \mathcal{P}_2(\mathbf{R} \times \mathbf{R}).$$

In other words, we have equal masses ( $1/2$ ), but different positions — since  $0 < a < b$ .

**Proposition 4.** One has

$$\mathfrak{d}(\mathcal{T}[\mu], \mathcal{T}[\nu])^2 = \mathcal{W}_2(\mu, \nu)^2 + 2\hbar.$$

The classical optimal transport in this case is obvious: send mass  $1/2$  from  $-a$  to  $-b$ , and mass  $1/2$  from  $+a$  to  $+b$ .

Next we consider the following example, with identical locations, but unequal masses: for  $0 < \epsilon < 1$ , set

$$\mu = \frac{1}{2}(\delta_{(+a,0)} + \delta_{(-a,0)}) \quad \text{and} \quad \rho_\epsilon = \frac{1+\epsilon}{2}\delta_{(+a,0)} + \frac{1-\epsilon}{2}\delta_{(-a,0)} \in \mathcal{P}_2(\mathbf{R} \times \mathbf{R}).$$

**Proposition 5.** For each  $\epsilon \in (0, 1)$ , one has

$$\mathfrak{d}(\mathcal{T}[\mu], \mathcal{T}[\rho_\epsilon])^2 < \mathcal{W}_2(\mu, \rho_\epsilon)^2 + 2\hbar.$$

We shall not prove Propositions 4 and 5, which rely on rather intricate computations, and refer instead the interested reader to the article [19]. However, we shall discuss both results.

Because of Lemma 3, an optimal coupling for  $\mathcal{T}[\mu]$  and  $\mathcal{T}[\nu]$  is  $\mathcal{T}[\lambda]$ , where  $\lambda$  is the optimal coupling of  $\mu$  and  $\nu$ , i.e.

$$\lambda := \frac{1}{2}(\delta_{(+a,0)} \otimes \delta_{(+b,0)} + \delta_{(-a,0)} \otimes \delta_{(-b,0)}),$$

so that

$$\lambda := \frac{1}{2}(|+a, 0, +b, 0\rangle\langle +a, 0, +b, 0| + |-a, 0, +b, 0\rangle\langle -a, 0, +b, 0|),$$

with the notation

$$|q, p, q', p'\rangle(x, y) = |q, p\rangle \otimes |q', p'\rangle(x, y) = |q, p\rangle(x) |q', p'\rangle(y).$$

In the unequal mass case, it is proved in [19] that there exists a quantum coupling of the form

$$\begin{aligned} T = & \sum_{k, l \in \{\pm\}} \tau_{klkl} |ka, 0, la, 0\rangle\langle ka, 0, la, 0| \\ & + \sum_{(k, l) \neq (m, n) \in \{\pm\}} \tau_{klmn} |ka, 0, la, 0\rangle\langle ma, 0, na, 0| \in \mathcal{C}(\mathcal{T}[\mu], \mathcal{T}[\rho_\epsilon]), \end{aligned}$$

with

$$\sum_{(k, l) \neq (m, n) \in \{\pm\}} |\tau_{klmn}|^2 > 0,$$

and

$$\text{trace}(T^{1/2} C_h T^{1/2}) < \mathcal{W}_2(\mu, \rho_\epsilon)^2 + 2\hbar.$$

Clearly, any coupling of  $\mu$  and  $\rho_\epsilon$  must be of the form

$$\sum_{k, l \in \{\pm\}} q_{kl} \delta_{(ka, 0)} \otimes \delta_{(la, 0)},$$

and therefore belongs to the 4-dimensional linear space

$$\text{span}\{\delta_{(\pm a, 0)} \otimes \delta_{(\pm a, 0)}\}.$$

On the contrary, couplings of two rank-2 operators with

$$R = \frac{1}{2}(|+a, 0\rangle\langle +a, 0| + |-a, 0\rangle\langle -a, 0|)$$

and

$$S_\epsilon = \frac{1+\epsilon}{2} |+a, 0\rangle\langle +a, 0| + \frac{1-\epsilon}{2} |-a, 0\rangle\langle -a, 0|$$

with  $0 < \epsilon < 1$  belong to the 16-dimensional linear space

$$\text{span}\{|ka, 0, la, 0\rangle\langle m, a, 0, na, 0| : k, l, m, n \in \{\pm\}\}.$$

Therefore, one could summarize the results in [19] as follows: since there are more degrees of freedom in the set of quantum couplings than in the set of classical couplings, it is natural to surmise that quantum optimal transport is cheaper than classical optimal transport, since more couplings are allowed in the quantum case than in the classical case. However, this simple argument is not sufficient to prove a *strict* inequality as in [19]. Putting together Proposition 5 and Lemma 3 shows that any optimal coupling for  $\mathfrak{d}(\mathcal{T}[\mu], \mathcal{T}[\rho_\epsilon])$  must be of the form

$$\begin{aligned} T = & \sum_{k, l \in \{\pm\}} \tau_{klkl} |ka, 0, la, 0\rangle\langle ka, 0, la, 0| \\ & + \sum_{(k, l) \neq (m, n) \in \{\pm\}} \tau_{klmn} |ka, 0, la, 0\rangle\langle ma, 0, na, 0| \in \mathcal{C}(\mathcal{T}[\mu], \mathcal{T}[\rho_\epsilon]), \end{aligned}$$

with

$$\sum_{(k, l) \neq (m, n) \in \{\pm\}} |\tau_{klmn}|^2 > 0.$$

While the term

$$\sum_{k,l \in \{\pm\}} \tau_{klkl} |ka, 0, la, 0\rangle \langle ka, 0, la, 0|$$

has a classical interpretation, since it is the image of a bounded, positive Radon measure on phase space by the Toeplitz map, the term

$$\sum_{(k,l) \neq (m,n) \in \{\pm\}} \tau_{klmn} |ka, 0, la, 0\rangle \langle ma, 0, na, 0| \neq 0$$

does not have any classical interpretation.

At the time of this writing, the structure of optimal couplings for  $\mathfrak{d}(R_1, R_2)$  with  $R_1, R_2 \in \mathcal{D}_2(\mathfrak{H})$  is not very well understood in general, at variance with the classical case (see however the discussion of this point in [20]), which is completely clarified by the Knott-Smith and the Brenier theorems. The very simple examples discussed in [19] show that optimal couplings between quantum densities are much more involved than in the classical case.

## 2. LECTURE II: APPLYING THE QUANTUM WASSERSTEIN PSEUDOMETRIC TO PARTICLE DYNAMICS

In this lecture, we shall discuss several applications of the quantum Wasserstein pseudometric  $\mathfrak{d}$  introduced in Lecture I. These applications include

- various limits of many-body problems in quantum mechanics (section 2.3),
- proofs of the uniform in  $\hbar$  convergence of some numerical schemes for quantum dynamics (section 2.4), and
- observation inequalities for the Schrödinger and for the von Neumann equations (section 2.4).

**2.1. Basics of Quantum Dynamics.** As already explained, one does not need to be an expert in quantum mechanics to read these notes.

However, some familiarity with the most elementary notions of quantum mechanics is required in order to understand the simple ideas behind the computations and the mathematical apparatus studied in this course. An excellent reference to learn quantum mechanics is [10]; see also [23] for a more detailed, yet equally lucid presentation. The reference [44] is interesting for mathematicians, but contains few physical explanations, at variance with [10, 23].

**2.1.1. Classical Mechanics.** We recall the fundamentals of classical dynamics for a point particle of mass  $m$ , in Hamiltonian form.

The position of this point particle at time  $t$  is denoted by  $q(t) \in \mathbf{R}^d$ , while its momentum at time  $t$  is denoted by  $p(t) \in \mathbf{R}^d$ .

The **Hamiltonian** for the point particle is the total energy of that particle, expressed in terms of its position  $q$  and momentum  $p$  — a word of caveat: it is essential at this point to use these variables, and not other variables, say for instance the velocity instead of the momentum. For a point particle with mass  $m$

$$H(q, p) = \underbrace{|p|^2/2m}_{\text{kinetic}} + \underbrace{V(q)}_{\text{potential}} = \text{total energy.}$$

With this, we can write **Newton's 2nd law of motion in Hamiltonian form:**

$$\dot{q}(t) = \partial H / \partial p = p(t) / m, \quad \dot{p}(t) = -\partial H / \partial q = -\nabla V(q(t)).$$

(In fact, the first equation is of a kinematic nature, since it can be viewed as a definition of the momentum  $p(t)$  in terms of the particle mass  $m$  and of the velocity  $\dot{q}(t)$ . Only the second equation corresponds to Newton's second law.)

Newton's second law of motion can be viewed as the governing equation for the special phase space probability measure  $\delta_{(q(t), p(t))}$ . But of course, it is equally interesting to consider the dynamics of phase space Borel probability measures more general than  $\delta_{(q(t), p(t))}$ .

This is precisely the purpose of the **Liouville equation**: if  $f \equiv f(t, x, \xi)$  is the probability density of finding the point particle at the position  $x \in \mathbf{R}^d$  with momentum  $\xi \in \mathbf{R}^d$  at time  $t$ , it satisfies the equation

$$\partial_t f(t, x, \xi) + \frac{1}{m} \xi \cdot \nabla_x f(t, x, \xi) - \nabla V(x) \cdot \nabla_\xi f(t, x, \xi) = 0.$$

This equation can be recast in terms of the **Poisson bracket**:

$$\{H(x, \xi), f(x, \xi)\} := \nabla_\xi H(x, \xi) \cdot \nabla_x f(x, \xi) - \nabla_x H(x, \xi) \cdot \nabla_\xi f(x, \xi),$$

as follows:

$$\partial_t f(t, x, \xi) + \{H(x, \xi), f(t, x, \xi)\} = 0.$$

Newton's second law of motion in Hamiltonian form is a system of ODEs to which one can apply the Cauchy-Lipschitz theorem. Specifically, the local existence of a unique solution of the Cauchy problem for the Hamiltonian formulation of Newton's second law of motion is implied by the assumption

$$V \in C^{1,1}(\mathbf{R}^d)$$

— meaning that  $V \in C^1(\mathbf{R}^d)$  and  $\nabla V$  is Lipschitz continuous on  $\mathbf{R}^d$ . Since

$$H(q(t), p(t)) = H(q(0), p(0))$$

(the verification of this is left to the reader as an exercise), which corresponds to the conservation of total energy by the dynamics deduced from Newton's second law of motion, it is easily seen that all solutions of the Cauchy problem for Newton's second law of motion are defined for all  $t \in \mathbf{R}$  under the condition

$$\lim_{|x| \rightarrow +\infty} V(x) = +\infty.$$

(Indeed, in that case, the map  $(q, p) \mapsto H(q, p)$  is proper on  $\mathbf{R}^d \times \mathbf{R}^d$ , i.e. the inverse image of any compact subset of  $\mathbf{R}$  is compact in  $\mathbf{R}^d \times \mathbf{R}^d$ ; this implies that the Hamiltonian flow generated by Newton's second law of motion is global by the most elementary continuation argument for ODEs; see for instance (10.5.5) in section 5 of chapter X in [29].)

**2.1.2. Quantum Mechanics.** The state of a quantum particle at time  $t$  is given by its **wave function**:

$$\psi \equiv \psi(t, x) \in L^2(\mathbf{R}^d; \mathbf{C}) =: \mathfrak{H}$$

such that

$$\|\psi(t, \cdot)\|_{\mathfrak{H}} = 1.$$

In classical mechanics, the Hamiltonian is a function on the phase space  $\mathbf{R}^d \times \mathbf{R}^d$ ; in quantum mechanics, the **quantum Hamiltonian** is an (unbounded) self-adjoint operator on  $\mathfrak{H}$ :

$$\mathbf{H} = -\frac{\hbar^2}{2m} \Delta_x + V(x) = \mathbf{H}^*.$$

(Here, the real-valued potential operator  $V$  is to be understood as a multiplication operator, i.e.  $\psi(x) \mapsto V(x)\psi(x)$ .)



The reason for considering this specific operator by analogy with the case of a point particle in classical mechanics is

**The correspondence principle.**

$$V(q) \rightarrow \text{multiplication by } V(x) \text{ and } p_j \rightarrow -i\hbar\partial_{q_j} = \hbar D_{q_j}.$$

With these mathematical objects, the quantum analogue of Newton's second law of motion is

**The Schrödinger equation.**

$$i\hbar\partial_t\psi(t, x) = \mathbf{H}\psi(t, x).$$

Assuming that  $\mathbf{H}$  is an unbounded self-adjoint operator on  $\mathfrak{H}$ , it generates a unitary group on  $\mathfrak{H}$  by Stone's theorem, denoted by  $e^{-it\mathbf{H}/\hbar}$ , so that

$$\psi(t, \cdot) = e^{-it\mathbf{H}/\hbar}\psi(0, \cdot), \quad t \in \mathbf{R}.$$

As explained in Lecture I, the quantum analogue of Borel probability measures on  $\mathbf{R}^d \times \mathbf{R}^d$  are density operators on  $\mathfrak{H} = L^2(\mathbf{R}^d)$ . The quantum analogue of the Liouville equation, defining the dynamics of phase space distribution functions is

**The von Neumann equation** for  $R(t) \in \mathcal{D}(\mathfrak{H})$ ,

$$i\hbar\partial_t R(t) = \underbrace{\mathbf{H}R(t) - R(t)\mathbf{H}}_{=:[\mathbf{H}, R(t)]}.$$

One easily checks that, under the assumption that  $\mathbf{H}$  is an unbounded self-adjoint operator on  $\mathfrak{H}$ ,

$$R(t) = e^{-it\mathbf{H}/\hbar}R(0)e^{it\mathbf{H}/\hbar}.$$

For instance, one easily checks that, if  $\psi(t, \cdot)$  is a solution of the Schrödinger equation, then the rank-1 density operator  $R(t) := |\psi(t, \cdot)\rangle\langle\psi(t, \cdot)|$  is a solution of the von Neumann equation.

This brings forward a further analogy in the correspondence principle, between the Poisson bracket  $\{f, g\}$  of two  $C^1$  functions defined on  $\mathbf{R}^d \times \mathbf{R}^d$ , and the commutator  $[A, B]$  of two operators on  $L^2(\mathbf{R}^d)$ :

$$\{\cdot, \cdot\} \rightarrow \frac{i}{\hbar}[\cdot, \cdot].$$

We conclude this section with a quick discussion of sufficient conditions on the potential  $V$  under which the quantum Hamiltonian  $\mathbf{H} = -\frac{\hbar^2}{2m}\Delta + V$  is self-adjoint on  $\mathfrak{H} = L^2(\mathbf{R}^d)$ . As explained above, if  $\mathbf{H}$  is self-adjoint, by Stone's theorem, it generates a quantum dynamics, namely the unitary group  $e^{-it\mathbf{H}/\hbar}$  on  $\mathfrak{H}$ .

**Self-adjointness of  $\mathbf{H}$ .** A first procedure for generating self-adjoint Hamiltonians of the form  $-\Delta + V$  is based on the associated (sesquilinear) quadratic form  $b$  defined as follows:

$$(\phi, \psi) \mapsto b(\phi, \psi) = \int_{\mathbf{R}^d} \nabla\overline{\phi(x)} \cdot \nabla\psi(x)dx + \int_{\mathbf{R}^d} V(x)\overline{\phi(x)}\psi(x)dx$$

for all  $\phi, \psi \in \mathcal{Q}(b)$ , where

$$\mathcal{Q}(b) := \{\phi \in H^1(\mathbf{R}^d) \text{ s.t. } |V|^{1/2}\phi \in L^2(\mathbf{R}^d)\}.$$

Assume that  $V$  is real-valued, and satisfies the following condition:

$$V \in L_{loc}^\infty(\mathbf{R}^d), \quad \text{and there exists } M \geq 0 \text{ s.t. } V(x) \geq -M \text{ for a.e. } x \in \mathbf{R}^d.$$

Then  $C_c^\infty(\mathbf{R}^d) \subset \mathcal{Q}(b)$  which is therefore dense in  $\mathfrak{H} = L^2(\mathbf{R}^d)$ . Then, the quadratic form  $b$  is semi-bounded, since

$$\phi \in \mathcal{Q}(b) \implies b(\phi, \phi) \geq -M\|\phi\|_{\mathfrak{H}}^2.$$

Besides, the quadratic form is **closed**, meaning that, for each sequence  $\phi_n \in \mathcal{Q}(b)$  such that

$$\phi_n \rightarrow \phi \text{ in } \mathfrak{H} \text{ as } n \rightarrow \infty, \quad \text{and } b(\phi_n - \phi_m, \phi_n - \phi_m) \rightarrow 0 \text{ as } m, n \rightarrow \infty,$$

one has

$$\phi \in \mathcal{Q}(b), \quad \text{and } b(\phi_n - \phi, \phi_n - \phi) \rightarrow 0 \text{ as } n \rightarrow \infty.$$

(Indeed, one easily checks that  $\nabla\phi_n$  is a Cauchy sequence in  $L^2(\mathbf{R}^d; \mathbf{R}^d)$ , and therefore converges towards an  $L^2$  vector field  $\xi$  on  $\mathbf{R}^d$ ; since  $\nabla\phi_n \rightarrow \nabla\phi$  in the sense of distributions on  $\mathbf{R}^d$ , one has  $\phi \in H^1(\mathbf{R}^d)$  and  $\xi = \nabla\phi$ . Similarly,  $(1+M+V)^{1/2}\phi_n$  is a Cauchy sequence in  $\mathfrak{H}$ , and therefore converges to some limit  $\ell \in \mathfrak{H}$ ; then, one easily checks that  $\ell/\sqrt{1+M+V} \in \mathfrak{H}$  and that  $\phi_n \rightarrow \ell/\sqrt{1+M+V}$  as  $n \rightarrow \infty$ , so that  $\phi = \ell/\sqrt{1+M+V}$  by uniqueness of the limit in  $\mathfrak{H}$ . Hence  $\sqrt{1+M+V}\phi \in \mathfrak{H}$ , so that  $\phi \in \mathcal{Q}(b)$ . The remaining part of the proof is routine, and left to the reader.) By Theorem VIII.15 of [53], there exists an unbounded self-adjoint operator  $A$  with domain  $D(A) \subset \mathcal{Q}(b)$  such that

$$\text{for all } \phi, \psi \in D(A), \quad b(\phi, \psi) = \langle \psi | A\phi \rangle.$$

Obviously  $C_c^\infty(\mathbf{R}^d) \subset D(A)$  and  $A$  coincides with  $-\Delta + V$  on  $C_c^\infty(\mathbf{R}^d)$ .

A second procedure is the **Kato-Rellich Theorem** (see Theorem X.12 of [54]). Assume that  $A$  is a self-adjoint operator on  $\mathfrak{H}$  with domain  $D(A)$ , and  $B$  is a symmetric<sup>12</sup> operator on  $\mathfrak{H}$  with domain  $D(B)$ . Assume that  $D(A) \subset D(B)$ , and that there exists  $a \in [0, 1)$  and  $b \geq 0$  such that

$$\|B\phi\| \leq a\|A\phi\| + b\|\phi\|, \quad \phi \in D(A).$$

Then  $A + B$  is self-adjoint on  $D(A)$ .

With this result, one can prove that  $-\Delta + V$  is self-adjoint on  $D(-\Delta) = H^2(\mathbf{R}^3)$  provided that

$$V \in L^2(\mathbf{R}^3) + L^\infty(\mathbf{R}^3).$$

This result is particularly important in atomic physics, since the Coulomb potential

$$V(x) = \pm \frac{1}{|x - x_0|} = \pm \left( \frac{\mathbf{1}_{|x-x_0| \leq 1}}{|x - x_0|} + \frac{\mathbf{1}_{|x-x_0| > 1}}{|x - x_0|} \right) \in L^2(\mathbf{R}^3) + L^\infty(\mathbf{R}^3).$$

Here,  $V$  is the potential energy of an electron at the position  $x$  interacting with a nucleus located at the position  $x_0$ .

One can also combine these two result to treat the case of  $-\Delta + V + V_\infty$  where  $V \in L^2(\mathbf{R}^3) + L^\infty(\mathbf{R}^3)$ , while  $V_0 \geq 0$  belongs to  $L_{loc}^\infty(\mathbf{R}^3)$  and  $V_0(x) \rightarrow +\infty$  as  $|x| \rightarrow +\infty$  (in other words,  $V_0$  is a confining potential).

In all these cases, one can see that much less regularity is required on the potential in order to define the quantum dynamics, than in the case of the classical dynamics. Perhaps the reason for this difference is that classical mechanics deals with the dynamics of much more singular objects (i.e.  $\delta_{(q(t), p(t))}$ ) than the wave function  $\psi(t, \cdot) \in L^2(\mathbf{R}^d)$  in quantum mechanics.

<sup>12</sup>I.e.  $\langle \phi | B\psi \rangle = \langle B\phi | \psi \rangle$  for all  $\phi, \psi \in D(B)$ . In other words,  $D(B) \subset D(B^*)$  and  $B^*|_{D(B)} = B$ .

**Quiz 14.** Explain how one can use the Kato-Rellich theorem to prove that

$$V \in L^2(\mathbf{R}^3) + L^\infty(\mathbf{R}^3) \implies -\Delta + V \text{ is self-adjoint with domain } H^2(\mathbf{R}^3).$$

(Hint: use the Sobolev embedding.)

In all the situations described above, one starts from the operator  $-\frac{\hbar^2}{2m}\Delta + V$ , which is well defined on  $C_c^2(\mathbf{R}^d) \subset L^2(\mathbf{R}^d)$ . Both methods described above (either the method involving a quadratic form, or the Kato-Rellich Theorem) produce an unbounded self-adjoint operator on  $L^2(\mathbf{R}^d)$ , which is therefore densely defined in  $L^2(\mathbf{R}^d)$ . The domain of this self-adjoint operator contains  $C_c^2(\mathbf{R}^d)$ , and this self-adjoint operator coincides with the differential operator  $-\frac{\hbar^2}{2m}\Delta + V$  on  $C_c^2(\mathbf{R}^d)$ . This self-adjoint operator is therefore an extension of the differential operator  $-\frac{\hbar^2}{2m}\Delta + V$ , defined on  $C_c^2(\mathbf{R}^d) \subset L^2(\mathbf{R}^d)$ , and we shall keep the notation  $-\frac{\hbar^2}{2m}\Delta + V$  to designate this extension.

**2.1.3. The Classical Limit of Quantum Mechanics.** Consider a point particle of mass  $m$ , moving at a speed  $v$ ; its de Broglie wavelength is the ratio  $2\pi\hbar/mv$ . For instance a dust particle of diameter  $1\mu$  with mass  $m = 10^{-6}\mu\text{g}$  moving at speed  $1\text{mm/s}$ , has a de Broglie wavelength  $6.6 \cdot 10^{-6}\text{\AA} \ll 1\mu$ , the size of the dust particle. (This example is taken from chapter I, complement A in [23].) If the de Broglie wavelength of a particle is negligible when compared to its size, or to the typical length scale of the experiment, one expects that the laws of classical mechanics should be sufficient to describe its behavior.

There are various ways of describing the classical limit of quantum mechanics.

As explained in §6 of [48], in the quasi-classical regime, the phase of the wave function of a particle is proportional to the mechanical action of that particle, and the constant of proportionality is Planck's constant  $\hbar$ . One could therefore study solutions of the Schrödinger equation

$$i\hbar\partial_t\psi = -\left(\frac{\hbar^2}{2m}\Delta + V\right)\psi$$

in the form of a **WKB ansatz** i.e. a formal series

$$\psi(t, x) = \sum_{n \geq 0} \hbar^n a_n(t, x) e^{iS(t, x)/\hbar}, \quad S(t, x) \text{ and } a_n(t, x) \in \mathbf{R},$$

where the phase  $S$  and the amplitude coefficients  $a_0, a_1, \dots$  are smooth.

Assuming that  $a_0(t, x) \neq 0$  for all  $(t, x)$ , one finds that

(a) the phase  $S$  is a solution of the *eikonal equation*, that is a Hamilton-Jacobi equation:

$$\partial_t S + H(\nabla_x S, x) = 0;$$

(b) at leading order, the amplitude  $a_0$  is a solution of the *transport equation*:

$$\partial_t a_0^2 + \text{div}_x(a_0^2 \nabla_x S(t, x)) = 0.$$

However, this description usually fails after some finite time, for the following reason: the graph of the map  $x \mapsto \nabla_x S(t, x)$  is the image of the graph of the map  $x \mapsto \nabla_x S(0, \cdot)$  by the flow of the classical Hamiltonian  $H(x, \xi)$ . In general, the graph of  $\nabla_x S(0, \cdot)$  becomes folded after some finite time in such a way that it is no longer the graph of map from  $\mathbf{R}^d$  to  $\mathbf{R}^d$ . The image of these folds by the projection  $\mathbf{R}^d \times \mathbf{R}^d \ni (x, \xi) \mapsto x \in \mathbf{R}^d$  is referred to as the “caustic”, by analogy with geometric optics. The appearance of caustics is the reason why the WKB ansatz is in general only local in time.

**Quiz 15.** Write the WKB ansatz for the free Schrödinger equation

$$i\hbar\partial_t\psi(t, x) = -\frac{\hbar^2}{2m}\partial_x^2\psi(t, x), \quad \psi(0, x) = a^{in}(x)e^{iS^{in}(x)/\hbar}, \quad x \in \mathbf{R},$$

where

$$S^{in}(x) = \begin{cases} -x^{1/3} & \text{for } x \geq 0, \\ +|x|^{1/3} & \text{for } x < 0. \end{cases}$$

Study the dynamics for all  $t \in \mathbf{R}$  (in the past  $t < 0$  as well as in the future  $t > 0$ ), and describe the caustic in this case.

Another approach to the classical limit of quantum mechanics involves the **Wigner Transform** already introduced in Quiz 11. We briefly recall the essentials: start from some integral operator  $R \in \mathcal{L}(\mathfrak{H})$  of the form

$$R\phi(x) = \int_{\mathbf{R}^d} r(x, y)\phi(y)dy.$$

For  $R =$  projection on  $\mathbf{C}\psi$  with  $\|\psi\|_{\mathfrak{H}} = 1$ , written  $R = |\psi\rangle\langle\psi|$  in Dirac's notation:

$$R\phi(x) = \underbrace{\left( \int_{\mathbf{R}^d} \overline{\psi(y)}\phi(y)dy \right)}_{=:\langle\psi|\phi\rangle} \psi(x) \implies r(x, y) = \psi(x)\overline{\psi(y)}.$$

We consider its Wigner transform at scale  $\hbar$ , given by the formula

$$W_{\hbar}[R](q, p) := \frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} r\left(q + \frac{1}{2}\hbar y, q - \frac{1}{2}\hbar y\right) e^{ip \cdot y} dy.$$

Some assumptions are needed on the integral kernel  $r$  for the Wigner transform to make sense. For instance, if  $R \in \mathcal{L}^2(\mathfrak{H})$  is a Hilbert-Schmidt operator, its integral kernel  $r$  belongs to  $L^2(\mathbf{R}^d \times \mathbf{R}^d)$ , so that the map

$$(q, y) \mapsto r\left(q + \frac{1}{2}\hbar y, q - \frac{1}{2}\hbar y\right)$$

belongs to  $L^2(\mathbf{R}_q^d \times \mathbf{R}_y^d)$ , since the Jacobian of the transformation

$$(q, y) \mapsto \left(q + \frac{1}{2}\hbar y, q - \frac{1}{2}\hbar y\right)$$

is  $(-\hbar)^d$ , which is in particular independent of  $(q, y)$ . In that case, the Wigner transform  $W_{\hbar}[R]$  belongs to  $L^2(\mathbf{R}_q^d \times \mathbf{R}_p^d)$  by the Plancherel theorem, as the partial Fourier transform of a square integrable measurable function.

Observe that

$$R = R^* \implies W_{\hbar}[R](q, p) \in \mathbf{R},$$

but

$$R \geq 0 \text{ does not imply that } W_{\hbar}[R] \geq 0.$$

(For instance, if  $\psi$  is an odd wave function, then  $W_{\hbar}[|\psi\rangle\langle\psi|](0, 0) < 0$ .)

We have seen in Quiz 11 that, if a continuous, time-dependent density operator  $R(t)$  is a weak solution of the von Neumann equation

$$i\hbar\partial_t R(t) = \left[-\frac{\hbar^2}{2m}\Delta + V, R(t)\right],$$

its Wigner transform  $W_{\hbar}[R(t)]$  is a weak solution of the Wigner equation

$$(\partial_t + q \cdot \nabla_q)W_{\hbar}[R(t)](q, p) + \Theta[V]W_{\hbar}[R(t)](q, p) = 0,$$

where  $\Theta(V)$  is the nonlocal (linear) operator with distribution kernel

$$\frac{1}{(2\pi)^d} \int_{\mathbf{R}^d} \frac{1}{i\hbar} \left( (V\left(q + \frac{1}{2}\hbar y\right) - V\left(q - \frac{1}{2}\hbar y\right)) e^{iy \cdot (p' - p)} \right) dy.$$

If  $V$  is a polynomial of degree  $\leq 2$ , one easily checks that

$$\Theta[V] = -\nabla V(q) \cdot \nabla_p.$$

The classical limit of quantum mechanics can be formulated in terms of the Wigner function, in the following manner.

**Theorem (Lions-Paul).** Assume that  $V$  is a real-valued function satisfying the assumptions

$$\inf_{q \in \mathbf{R}^d} V(q) > -\infty, \quad V \in C^{1,1}(\mathbf{R}^d) \quad \text{and} \quad V(q) = O(|q|^n) \quad \text{for some } n \geq 0 \quad \text{as } |q| \rightarrow \infty,$$

and let

$$\mathbf{H} = -\frac{\hbar^2}{2m} \Delta + V$$

be the quantum Hamiltonian, with domain  $H^2(\mathbf{R}^d) \cap L^2(\mathbf{R}^d; V(x)_+^2 dx)$ . Let  $R_\hbar^{in}$  be a family of  $\mathcal{D}(\mathfrak{H})$  such that

$$W_\hbar[R_\hbar^{in}] \rightarrow f^{in} \quad \text{in } \mathcal{S}'(\mathbf{R}^d \times \mathbf{R}^d) \quad \text{as } \hbar \rightarrow 0.$$

Then

$$f^{in} \text{ is a probability density on } \mathbf{R}^d \times \mathbf{R}^d,$$

and

$$W_\hbar[e^{-it\mathbf{H}/\hbar} R_\hbar^{in} e^{it\mathbf{H}/\hbar}] \rightarrow f(t, \cdot, \cdot) \quad \text{in } \mathcal{S}'(\mathbf{R}^d \times \mathbf{R}^d) \quad \text{as } \hbar \rightarrow 0$$

uniformly in  $t \in [0, T]$  for each  $T > 0$ , where  $f$  is the probability density solution to the Liouville equation

$$\partial_t f(t, q, p) + \left\{ \frac{1}{2m} |p|^2 + V(q), f(t, q, p) \right\} = 0, \quad f|_{t=0} = f^{in}.$$

This is Theorem IV.1 in [50], and the interested reader is referred to this article for its proof, together with several other interesting examples involving the Wigner transform.

The connection between these two approaches to the classical limit of quantum mechanics is made clear by the following example, which is left to the reader as an exercise.

**Quiz 16.** Consider a **WKB wave function**

$$\psi_\hbar^{in}(x) = a^{in}(x) e^{iS^{in}(x)/\hbar},$$

with  $\|a^{in}\|_{L^2} = 1$  and  $S \in \text{Lip}(\mathbf{R}^d; \mathbf{R})$ . Prove that

$$W_\hbar[|\psi_\hbar^{in}\rangle\langle\psi_\hbar^{in}|](q, p) \rightarrow |a^{in}(q)|^2 \delta(p - \nabla S^{in}(q))$$

in the sense of tempered distributions as  $\hbar \rightarrow 0$ . Explain how the eikonal and the transport equation predicted in statements (a)-(b) above in this section emerge from the dynamics of  $f^{in}$  predicted by the Lions-Paul Theorem.

**2.2. Amplification of the Pseudometric  $\mathfrak{d}$  by Hamiltonian Dynamics.** This is the core of the present lecture. Our goal is to control amplifications of the pseudometric  $\mathfrak{d}$  constructed in Lecture I by Hamiltonian dynamics (classical or quantum).

2.2.1. *Pair Dispersion in Classical Mechanics.* In this brief section, we seek to compare the evolution of two different initial data, with two different potentials, in classical and in quantum mechanics. (The term “pair dispersion” is used in Lagrangian fluid mechanics, and is perhaps not so common in the present setting.) The results obtained here should be thought of as a warm-up — and a motivation — for the study of the dynamical amplification of  $\mathfrak{d}$ .

**Classical dynamics.** We seek to compare two solutions of Newton’s equations,  $t \mapsto (X, \Xi)(t)$  and  $t \mapsto (Y, H)(t)$  with two different potentials  $V$  and  $W$  belonging to  $C^{1,1}(\mathbf{R}^d)$ :

$$\begin{cases} \dot{X} = \frac{1}{m}\Xi, \\ \dot{\Xi} = -\nabla V(X), \end{cases} \quad \text{and} \quad \begin{cases} \dot{Y} = \frac{1}{m}H, \\ \dot{H} = -\nabla W(Y), \end{cases}$$

with initial data

$$(X, \Xi)(0) = (X^{in}, \Xi^{in}), \quad (Y, H)(0) = (Y^{in}, H^{in}).$$

Setting  $L := \text{Lip}(\nabla V)$ , we compute

$$\begin{aligned} & \frac{d}{dt}(|X - Y|^2 + |\Xi - H|^2) \\ &= \frac{2}{m}(\Xi - H) \cdot (X - Y) - 2(\nabla V(X) - \nabla W(Y)) \cdot (\Xi - H) \\ &= \frac{2}{m}(\Xi - H) \cdot (X - Y) - 2(\nabla V(X) - \nabla V(Y)) \cdot (\Xi - H) \\ & \quad + 2(\nabla W(Y) - \nabla V(Y)) \cdot (\Xi - H) \\ & \leq \left(\frac{1}{m} + L\right)(|X - Y|^2 + |\Xi - H|^2) \\ & \quad + 2\|\nabla(V - W)\|_{L^\infty(\mathbf{R}^d)}(|\Xi| + |H|) \\ & \leq \left(\frac{1}{m} + L\right)(|X - Y|^2 + |\Xi - H|^2) \\ & \quad + 2\|\nabla(V - W)\|_{L^\infty(\mathbf{R}^d)}\sqrt{|\Xi|^2 + 2V(X) + 2\|V\|_{L^\infty(\mathbf{R}^d)}} \\ & \quad + 2\|\nabla(V - W)\|_{L^\infty(\mathbf{R}^d)}\sqrt{|H|^2 + 2W(Y) + 2\|W\|_{L^\infty(\mathbf{R}^d)}} \\ & = \left(\frac{1}{m} + L\right)(|X - Y|^2 + |\Xi - H|^2) \\ & \quad + 2\|\nabla(V - W)\|_{L^\infty(\mathbf{R}^d)}\sqrt{|\Xi^{in}|^2 + 2V(X^{in}) + 2\|V\|_{L^\infty(\mathbf{R}^d)}} \\ & \quad + 2\|\nabla(V - W)\|_{L^\infty(\mathbf{R}^d)}\sqrt{|H^{in}|^2 + 2W(Y^{in}) + 2\|W\|_{L^\infty(\mathbf{R}^d)}}, \end{aligned}$$

where the last equality follows from the energy conservation. By Gronwall’s inequality

$$\begin{aligned} (|X(t) - Y(t)|^2 + |\Xi(t) - H(t)|^2) &\leq (|X^{in} - Y^{in}|^2 + |\Xi^{in} - H^{in}|^2)e^{\left(\frac{1}{m} + L\right)t} \\ & \quad + M \frac{e^{\left(\frac{1}{m} + L\right)t} - 1}{\left(\frac{1}{m} + L\right)} \|\nabla(V - W)\|_{L^\infty(\mathbf{R}^d)}, \end{aligned}$$

where

$$M := 2 \left( \sqrt{|\Xi^{in}|^2 + 2V(X^{in}) + 2\|V\|_{L^\infty(\mathbf{R}^d)}} + \sqrt{|H^{in}|^2 + 2W(Y^{in}) + 2\|W\|_{L^\infty(\mathbf{R}^d)}} \right).$$

**Quantum dynamics.** Consider the quantum dynamics of two different wave functions, driven by two potentials  $V$  and  $W$  such that both quantum Hamiltonians

$-\frac{\hbar^2}{2m}\Delta + V$  and  $-\frac{\hbar^2}{2m}\Delta + W$  have self-adjoint extensions to  $\mathfrak{H} = L^2(\mathbf{R}^d)$ :

$$\begin{aligned} i\hbar\partial_t\phi &= \left(-\frac{\hbar^2}{2m}\Delta + V\right)\phi, & \phi|_{t=0} &= \phi^{in}, \\ i\hbar\partial_t\psi &= \left(-\frac{\hbar^2}{2m}\Delta + W\right)\psi, & \psi|_{t=0} &= \psi^{in}. \end{aligned}$$

Thus

$$\partial_t(\phi - \psi) = \frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\Delta + V\right)(\phi - \psi) + \frac{1}{i\hbar}(V - W)\psi,$$

so that

$$\begin{aligned} \|(\phi - \psi)(t, \cdot)\|_{\mathfrak{H}} &\leq \|\phi^{in} - \psi^{in}\|_{\mathfrak{H}} + \frac{1}{\hbar} \int_0^t \|V - W\|_{L^\infty(\mathbf{R}^d)} \|\psi(s, \cdot)\|_{L^2(\mathbf{R}^d)} ds \\ &= \|\phi^{in} - \psi^{in}\|_{\mathfrak{H}} + \frac{t}{\hbar} \|V - W\|_{L^\infty(\mathbf{R}^d)}. \end{aligned}$$

If we compare these two estimates, we immediately see the following differences:

- (a) the  $L^2$  bound on the difference of wave functions of the quantum particle involves the difference of potentials in sup norm, whereas the difference in the position and momenta of the classical particle involves the difference of the force fields, i.e. of the gradients of the potentials, in sup norm;
- (b) there is no exponential amplification of the  $L^2$  norm of the difference of wave functions in the quantum case, whereas the bound on the difference in positions and momenta of the classical particle involves an amplification factor  $e^{Lt}$ , where  $L$  is the Lipschitz constant of one of the force fields; yet
- (c) the  $L^2$  estimate on the difference of wave functions is not uniform in  $\hbar$  as  $\hbar \rightarrow 0^+$ .

Notice however that the  $L^2$  bound on this difference is uniform in  $\hbar$  in the very special case where both potentials are equal. This is perhaps of limited interest in the context of numerical analysis, since numerical schemes typically replace the true potential  $V$  by an approximation thereof.

Also, the  $L^2$  bound on the difference of wave functions is an estimate of the same kind as a bound in trace, or Hilbert-Schmidt norm for the difference of the density operators, since

$$\| |\phi\rangle\langle\phi| - |\psi\rangle\langle\psi| \|_2^2 \leq 2\|\phi - \psi\|_{\mathfrak{H}}^2 \|\phi\|_{\mathfrak{H}}^2 + 2\|\psi\|_{\mathfrak{H}}^2 \|\phi - \psi\|_{\mathfrak{H}}^2 = 4\|\phi - \psi\|_{\mathfrak{H}}^2.$$

We have seen in Lecture I, in the case where  $\phi = |q_1, p_1\rangle$  and  $\psi = |q_2, p_2\rangle$ , that the Hilbert-Schmidt norm for the difference between such density operators converges to  $\|\delta_{(q_1, p_1)} - \delta_{(q_2, p_2)}\|_{TV} = 2$  unless  $(q_1, p_1) = (q_2, p_2)$  in the limit as  $\hbar \rightarrow 0^+$ . Therefore, in the case where  $V = W$ , the fact that the quantum dynamics is unitary would typically result in the perfectly true, but uninteresting inequality  $2 \leq 4$  in the small  $\hbar$  limit — assuming that  $\phi(t, \cdot) = |q_1, p_1\rangle$  and  $\psi(t, \cdot) = |q_2, p_2\rangle$ , corresponding to  $\phi(0, \cdot) = U(t)^*|q_1, p_1\rangle$  and  $\psi(0, \cdot) = U(t)^*|q_2, p_2\rangle$ , with  $U(t) := \exp(-it(-\frac{\hbar^2}{2m}\Delta + V)\hbar)$ .

On the contrary, the pair dispersion estimate in the classical setting can be understood as

$$\begin{aligned} \mathcal{W}_2(\delta_{(X(t), \Xi(t))}, \delta_{(Y(t), H(t))})^2 &\leq \mathcal{W}_2(\delta_{(X^{in}, \Xi^{in})}, \delta_{(Y^{in}, H^{in})})^2 e^{(\frac{1}{m} + L)t} \\ &\quad + M \frac{e^{(\frac{1}{m} + L)t} - 1}{(\frac{1}{m} + L)} \|\nabla(V - W)\|_{L^\infty(\mathbf{R}^d)}. \end{aligned}$$

It is an easy exercise (left to the reader) to extend this estimate to arbitrary phase space probability measures that are weak solutions of the Liouville equation — not necessary of the form  $\delta_{(z(t), \zeta(t))}$ , where  $t \mapsto (z(t), \zeta(t))$  is a solution of Newton's motion equations.

This suggests the following question: is there an estimate analogous to the classical pair dispersion estimate in the quantum setting for  $\mathfrak{d}(R_1(t), R_2(t))$ , where  $R_1$  and  $R_2$  are time-dependent density operators on  $\mathfrak{H} = L^2(\mathbf{R}^d)$  whose dynamics is governed by the von Neumann equation with two different potentials? Should such an estimate exist, it would be

- uniform in  $\hbar$  as  $\hbar \rightarrow 0^+$ , and
- should involve the same amplification factor  $e^{(\frac{1}{m}+L)t}$  as in the classical case, and a similar error term of order

$$e^{(\frac{1}{m}+L)t} \|\nabla(V-W)\|_{L^\infty(\mathbf{R}^d)} \quad \text{instead of} \quad \frac{1}{\hbar} \|V-W\|_{L^\infty(\mathbf{R}^d)}.$$

Of course, such an estimate, should it exist, would require more regularity on the potentials (typically the same kind of regularity as in the classical setting) than what is needed for the quantum Hamiltonians

$$-\frac{\hbar^2}{2m}\Delta + V \quad \text{and} \quad -\frac{\hbar^2}{2m}\Delta + W$$

to have (unbounded) self-adjoint extensions to  $\mathfrak{H} = L^2(\mathbf{R}^d)$ .

**2.2.2. Propagation Estimate for  $\mathfrak{d}$ .** It will be convenient to consider, instead of the pseudometric  $\mathfrak{d}$  introduced in the previous lecture, a deformation thereof, henceforth designated by  $\mathfrak{d}_\lambda$ , where  $\lambda > 0$  is the deformation parameter.

**Pseudometric  $\mathfrak{d}_\lambda$ .** For all  $R, S \in \mathcal{D}_2(\mathfrak{H})$ , and all probability density  $f$  on the phase space  $\mathbf{R}^d \times \mathbf{R}^d$  with finite 2nd order moment

$$\begin{aligned} \mathfrak{d}_\lambda(f, R)^2 &:= \inf_{Q \in \mathcal{C}(f, R)} \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q(x, \xi)^{\frac{1}{2}} c_{\lambda, \hbar}(x, \xi) Q(x, \xi)^{\frac{1}{2}}) dx d\xi, \\ \mathfrak{d}_\lambda(R, S)^2 &:= \inf_{Q \in \mathcal{C}(R, S)} \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(Q^{1/2} C_{\lambda, \hbar} Q^{1/2}), \end{aligned}$$

where  $c_{\lambda, \hbar}$  and  $C_{\lambda, \hbar}$  are the differential operators defined on  $\mathbf{R}^d$  and  $\mathbf{R}^d \times \mathbf{R}^d$  respectively by the formulas

$$\begin{aligned} c_{\lambda, \hbar}(x, \xi) \phi(y) &:= (\lambda^2 |x-y|^2 + |\xi + i\hbar \nabla_y|^2) \phi(y), \\ C_{\lambda, \hbar} \Phi(x, y) &:= (\lambda^2 |x-y|^2 - \hbar^2 (\nabla_x - \nabla_y) \cdot (\nabla_x - \nabla_y)) \Phi(x, y). \end{aligned}$$

We leave it to the reader as an easy exercise to check that

$$c_\hbar(x, \xi) \geq \lambda d\hbar I_{\mathfrak{H}}, \quad C_\hbar \geq 2\lambda d\hbar I_{\mathfrak{H} \otimes \mathfrak{H}}.$$

**Theorem 6.** Assume that  $V \in C^{1,1}(\mathbf{R}^d)$  satisfies

$$V(y) \rightarrow +\infty \text{ as } |y| \rightarrow \infty, \quad \text{and } \text{Lip}(\nabla V) < \infty.$$

Hence  $\mathbf{H} := -\frac{\hbar^2}{2m}\Delta + V$  has a self-adjoint extension to  $\mathfrak{H}$ , and defines a quantum dynamics via the unitary group  $U(t) := e^{-it\mathbf{H}/\hbar}$ . On the other hand, let  $\Phi(t; \cdot, \cdot)$  be the flow of the classical Hamiltonian  $H(x, \xi) := \frac{1}{2m}|\xi|^2 + V(x)$ , which is defined for all  $t \in \mathbf{R}$ , since  $V$  is confining (tends to  $+\infty$  at infinity). Then, for each  $R_1^{in}, R_2^{in} \in \mathcal{D}_2(\mathfrak{H})$  and each probability density  $f$  on the phase space  $\mathbf{R}^d \times \mathbf{R}^d$  with finite 2nd order moment, one has

$$\begin{aligned} \mathfrak{d}_\lambda(f^{in} \circ \Phi(-t, \cdot, \cdot), U(t)R_1^{in}U(t)^*) &\leq \mathfrak{d}_\lambda(f^{in}, R_1^{in})e^{L|t|}, \\ \mathfrak{d}_\lambda(U(t)R_1^{in}U(t)^*, U(t)R_2^{in}U(t)^*) &\leq \mathfrak{d}_\lambda(R_1^{in}, R_2^{in})e^{L|t|}, \end{aligned}$$



for all  $t \in \mathbf{R}$ , with

$$L := \frac{1}{2} \left( \frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda} \right).$$

*Proof.* Let  $Q^{in} \in \mathcal{C}(f^{in}, R_1^{in})$ ; set

$$Q(t, x, \xi) := U(t)Q^{in}(\Phi(t, x, \xi))U(t)^*.$$

One easily checks that

$$\begin{aligned} \text{trace}_{\mathfrak{H}}(Q(t, x, \xi)) &= \text{trace}_{\mathfrak{H}}(U(t)Q^{in}(\Phi(-t, x, \xi))U(t)^*) \\ &= \text{trace}_{\mathfrak{H}}(Q^{in}(\Phi(-t, x, \xi))) = f^{in}(\Phi(-t, x, \xi)), \end{aligned}$$

while

$$\begin{aligned} \iint_{\mathbf{R}^{2d}} U(t)Q^{in}(\Phi(-t, x, \xi))U(t)^* dx d\xi &= U(t) \left( \iint_{\mathbf{R}^{2d}} Q^{in}(\Phi(-t, x, \xi)) dx d\xi \right) U(t)^* \\ &= U(t) \left( \iint_{\mathbf{R}^{2d}} Q^{in}(X, \Xi) dX d\Xi \right) U(t)^* = U(t)R^{in}U(t)^*, \end{aligned}$$

so that

$$Q(t, \cdot, \cdot) \in \mathcal{C}(f^{in} \circ \Phi(-t, \cdot, \cdot), U(t)R_1^{in}U(t)^*).$$

Thus

$$\begin{aligned} &\mathfrak{d}_\lambda(f^{in} \circ \Phi(-t, \cdot, \cdot), U(t)R_1^{in}U(t)^*)^2 \\ &\leq \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q(t, x, \xi))^{1/2} c_{\lambda, \hbar}(x, \xi) Q(t, x, \xi)^{1/2} dx d\xi \\ &= \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q^{in}(X, \Xi))^{1/2} U(t)^* c_{\lambda, \hbar}(\Phi(t, X, \Xi)) U(t) Q^{in}(X, \Xi)^{1/2} dX d\xi, \end{aligned}$$

and

$$\begin{aligned} &\frac{d}{dt} \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q(t, x, \xi))^{1/2} c_{\lambda, \hbar}(x, \xi) Q(t, x, \xi)^{1/2} dx d\xi \\ &= \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q(t, x, \xi))^{1/2} \left\{ \frac{1}{2m} |\xi|^2 + V(x), c_{\lambda, \hbar}(x, \xi) \right\} Q(t, x, \xi)^{1/2} dx d\xi \\ &\quad + \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q(t, x, \xi))^{1/2} \left[ -\frac{\hbar^2}{2m} \Delta + V, c_{\lambda, \hbar}(x, \xi) \right] Q(t, x, \xi)^{1/2} dx d\xi. \end{aligned}$$

Now, we compute

$$\begin{aligned} &\left\{ \frac{1}{2m} |\xi|^2 + V(x), c_{\lambda, \hbar}(x, \xi) \right\} + \frac{i}{\hbar} \left[ -\frac{\hbar^2}{2m} \Delta + V, c_{\lambda, \hbar}(x, \xi) \right] \\ &= \frac{\lambda^2}{m} ((\xi + i\hbar \nabla_y) \cdot (x - y) + (x - y) \cdot (\xi + i\hbar \nabla_y)) \\ &\quad - (\xi + i\hbar \nabla_y) \cdot (\nabla V(x) - \nabla V(y)) - (\nabla V(x) - \nabla V(y)) \cdot (\xi + i\hbar \nabla_y) \\ &= \frac{\lambda}{m} ((\xi + i\hbar \nabla_y) \cdot \lambda(x - y) + \lambda(x - y) \cdot (\xi + i\hbar \nabla_y)) \\ &\quad - \frac{\text{Lip}(\nabla V)}{\lambda} ((\xi + i\hbar \nabla_y) \cdot \lambda \frac{\nabla V(x) - \nabla V(y)}{\text{Lip}(\nabla V)} + \lambda \frac{\nabla V(x) - \nabla V(y)}{\text{Lip}(\nabla V)} \cdot (\xi + i\hbar \nabla_y)) \\ &\leq \left( \frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda} \right) (\lambda^2 |x - y| + |\xi + i\hbar \nabla_y|^2). \end{aligned}$$

Therefore

$$\begin{aligned} &\frac{d}{dt} \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q(t, x, \xi))^{1/2} c_{\lambda, \hbar}(x, \xi) Q(t, x, \xi)^{1/2} dx d\xi \\ &\leq \left( \frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda} \right) \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q(t, x, \xi))^{1/2} c_{\lambda, \hbar}(x, \xi) Q(t, x, \xi)^{1/2} dx d\xi, \end{aligned}$$

and hence

$$\begin{aligned} & \mathfrak{d}_\lambda(f^{in} \circ \Phi(-t, \cdot, \cdot), U(t)R_1^{in}U(t)^*)^2 \\ & \leq \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q(t, x, \xi)^{1/2} c_{\lambda, \hbar}(x, \xi) Q(t, x, \xi)^{1/2}) dx d\xi \\ & \leq \exp(|t|(\frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda})) \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q^{in}(x, \xi)^{1/2} c_{\lambda, \hbar}(x, \xi) Q^{in}(x, \xi)^{1/2}) dx d\xi. \end{aligned}$$

Choosing a minimizing sequence  $Q_n^{in}$  of elements of  $\mathcal{C}(f^{in}, R_1^{in})$  so that

$$\iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q_n^{in}(x, \xi)^{1/2} c_{\lambda, \hbar}(x, \xi) Q_n^{in}(x, \xi)^{1/2}) dx d\xi \rightarrow \mathfrak{d}_\lambda(f^{in}, R_1^{in})^2$$

as  $n \rightarrow \infty$  leads to the first inequality.  $\square$

**Quiz 17.** Prove the second inequality in Theorem 6 (the argument follows the one presented above for the first inequality).

Theorem 6 will often be used together with Theorem 1 from Lecture I, in the following manner. For instance, one could start from Toeplitz density operators, for which the pseudometric  $\mathfrak{d}_\lambda$  is very well-known, by using Theorem 1 (1). Usually, the quantum dynamics fails to preserve the Toeplitz structure of the density operator, but at time  $t$ , one can use the lower bound for  $\mathfrak{d}_\lambda$  deduced from Theorem 1 (2) to compare the Husimi transforms of the (quantum) density operators by means of the classical Wasserstein  $\mathcal{W}_2$  metric. The resulting statement is as follows.

**Corollary 7.** Under the same assumptions as in Theorem 6, let  $f^{in}$  be a probability density with finite second order moments on  $\mathbf{R}^d \times \mathbf{R}^d$ , and let  $g_1^{in}, g_2^{in} \in \mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$ . Set  $R_1^{in} := \mathcal{T}[g_1^{in}]$  and  $R_2^{in} := \mathcal{T}[g_2^{in}]$ . Then

$$\begin{aligned} \mathcal{W}_2(f^{in} \circ \Phi(-t, \cdot, \cdot), \mathcal{H}[U(t)R_1^{in}U(t)^*])^2 & \leq e^{L|t|} \frac{\max(1, \lambda^2)}{\min(1, \lambda^2)} \mathcal{W}_2(f^{in}, g_1^{in})^2 \\ & \quad + \frac{(1 + \lambda^2)d\hbar}{2 \min(1, \lambda^2)} (e^{L|t|} + 1), \\ \mathcal{W}_2(\mathcal{H}[U(t)R_1^{in}U(t)^*], \mathcal{H}[U(t)R_2^{in}U(t)^*])^2 & \leq e^{L|t|} \frac{\max(1, \lambda^2)}{\min(1, \lambda^2)} \mathcal{W}_2(g_1^{in}, g_2^{in})^2 \\ & \quad + \frac{(1 + \lambda^2)d\hbar}{\min(1, \lambda^2)} (e^{L|t|} + 1), \end{aligned}$$

where  $L := \frac{1}{2}(\frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda})$ .

**Quiz 18.** Prove Corollary 7. (Hint: compute the operators

$$\iint_{\mathbf{R}^d \times \mathbf{R}^d} F(p)|q, p\rangle\langle q, p|dqdp \quad \text{and} \quad \iint_{\mathbf{R}^d \times \mathbf{R}^d} F(q)|q, p\rangle\langle q, p|dqdp,$$

where  $F$  is a polynomial of degree 2 on  $\mathbf{R}^d$ . See Appendix B of [37], or the basic properties of the Toeplitz map in Lecture I, and Quiz 8.)

In the remaining part of this lecture, we shall study three (more or less direct) applications of the propagation bound in Theorem 6:

Application 1: mean-field and classical limits of quantum mechanics (section 2.3);

Application 2: time-splitting schemes for quantum mechanics (section 2.4);

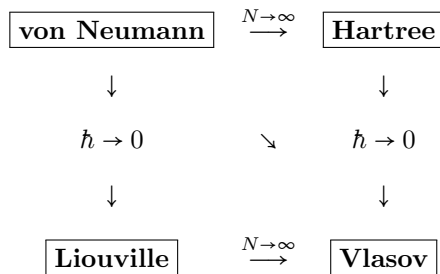
Application 3: observation inequalities for quantum dynamics (section 2.5).

**2.3. Mean-Field and Classical Limits of Quantum Mechanics.** So far, we have considered the quantum dynamics of a single particle. In this section, we consider the quantum dynamics of  $N$  identical particles. We are concerned with two different limiting regimes:

- the large  $N$  limit, and
- the semiclassical regime.

The large  $N$  limit is of special interest: in that case, the problem is set on a space of large dimension (typically  $3N$ , the number of position variables for  $N$  points in space dimension 3). Numerical simulations for problems of this type are very often untractable.

For this reason, one seeks to replace the original equation governing the dynamics of  $N$ -particle systems with “reduced models”, where the space dimension does not increase with the number of particles considered. The situation studied in this section is summarized by the following diagram.



The upper horizontal arrow (quantum mean-field limit) has been proved by Spohn [60] — see also [8] — in the case of bounded potentials; the case of a Coulomb potential was treated subsequently by [31] — see also [7], by the method of the “BBGKY hierarchy” (see for instance [35] for an elementary introduction to the BBGKY hierarchy). Other approaches involve the method of second quantization [56] — see also an original method due to Pickl [52], which makes use of some notions originating from second quantization without the full machinery of Fock spaces. Both [56] and [52] include a treatment of potentials with a singularity at the origin of the same type as for the Coulomb potential.

The lower horizontal arrow is the mean-field limit in classical mechanics; it has been proved by Braun-Hepp [14] by using the notion of “Klimontovich” solutions — i.e. phase-space empirical measures — of the Vlasov equation, with a convergence rate obtained by Dobrushin [30], who used optimal transport distances for the first time on this kind of problem. Dobrushin’s analysis is our first motivation for defining a quantum analogue of the Wasserstein distance, and for the analysis in the present section.

The limit corresponding to the left vertical arrow follows from the Lions-Paul Theorem mentioned above (Theorem IV.1 in [50]). The limit corresponding to the right vertical arrow follows from Theorem IV.2 in [50]. It should be mentioned that the case of the classical limit of the Hartree equation for singular potentials including the Coulomb case is treated in Theorem IV.4 of [50] — one should however keep in mind that this proof assumes that the Wigner function of the density operator solution of the Hartree equation is assumed to be bounded in  $L^2$ , which

excludes the case of pure states, such as solutions of the Schrödinger equation with WKB initial data, for instance.

Since we are interested in a situation involving two small parameters (specifically  $\frac{1}{N}$  and  $\hbar$ ), it is natural to investigate the uniformity of the mean-field (large  $N$ ) limit in the semiclassical (small  $\hbar$ ) regime. It seems that the first result in that direction is [43] — see also [51].

To conclude this introduction, one should also mention that the mean-field limit in classical mechanics (the lower horizontal arrow) remains an open problem in the case of the Coulomb potential — see however [45, 46] in the case of interactions with a singularity at the origin weaker than that of the Coulomb potential, and [59] in the case of the Coulomb potential itself, but for a restricted class of initial data (specifically for monokinetic data). The quantum analogue of [59], i.e. the joint mean-field and classical limit of the  $N$ -particle quantum dynamics in the case of monokinetic data, which is based on Serfaty's remarkable inequality on the Coulomb potential (Proposition 2.3 in [59]).

**2.3.1. Quantum  $N$ -Particle Dynamics.** The state of a quantum  $N$ -particle system at time  $t$  is represented by a density operator  $R(t) \in \mathcal{D}(\mathfrak{H}_N)$ , where  $\mathfrak{H}_N = \mathfrak{H}^{\otimes N}$  is the  $N$ -particle Hilbert space. If  $\mathfrak{H} = L^2(\mathbf{R}^d)$ , then it is easily seen that  $\mathfrak{H}_N = L^2(\mathbf{R}^{dN})$ .

Since the  $N$  particles are indistinguishable, their density operator should commute with permutations of the particle labels. More precisely, for each  $\sigma \in \mathfrak{S}_N$ , we define the map

$$U_\sigma : \mathfrak{H}_N \ni \Psi \equiv \Psi(x_1, \dots, x_N) \mapsto (U_\sigma \Psi) \equiv \Psi(x_{\sigma^{-1}(1)}, \dots, x_{\sigma^{-1}(N)}) \in \mathfrak{H}_N.$$

One easily checks that the map  $\sigma \mapsto U_\sigma$  is a unitary representation of the symmetric group  $\mathfrak{S}_N$  on the set of  $N$  elements in  $\mathfrak{H}_N$ . A density operator for a system of  $N$  identical particles should satisfy the relation

$$U_\sigma R(t) U_\sigma^* = R(t), \quad \sigma \in \mathfrak{S}_N, t \in \mathbf{R}.$$

Henceforth, we denote by  $\mathcal{D}^s(\mathfrak{H}_N)$  the set of density operators on  $\mathfrak{H}_N$  satisfying this symmetry property:

$$\mathcal{D}^s(\mathfrak{H}_N) := \{R \in \mathcal{D}(\mathfrak{H}_N) \text{ s.t. } U_\sigma R U_\sigma^* = R \text{ for all } \sigma \in \mathfrak{S}_N\}.$$

**Remark.** One should avoid confusing this symmetry, corresponding to indistinguishable particles, with the symmetries corresponding to the Bose-Einstein, or the Fermi-Dirac statistics. The Bose-Einstein statistics applies to particles with integer spin, referred to as bosons, such as photons,  ${}^4\text{He}$  nuclei ( $\alpha$  particles). At low temperature, large numbers of bosons can condense in a single energy state, thereby forming a Bose-Einstein condensate. The Fermi-Dirac statistics applies to particles with half-integer spin, referred to as fermions, such as electrons, protons, neutrons,  ${}^3\text{He}$  atoms. Fermions satisfy the Pauli exclusion principle: two (or more than two) fermions in a given quantum system cannot occupy simultaneously the same quantum state. The difference between bosons and fermions can be read on their  $N$ -particle wave functions: for all  $\sigma \in \mathfrak{S}_N$ , one has

$$\begin{aligned} U_\sigma \Psi_N &= \Psi_N, & \text{if the } N \text{ particles are bosons,} \\ U_\sigma \Psi_N &= (-1)^{\text{sign}(\sigma)} \Psi_N, & \text{if the } N \text{ particles are fermions.} \end{aligned}$$

Notice that, in both cases, one has

$$U_\sigma |\Psi_N\rangle \langle \Psi_N| U_\sigma^* = |U_\sigma \Psi_N\rangle \langle U_\sigma \Psi_N| = |\Psi_N\rangle \langle \Psi_N|, \quad \sigma \in \mathfrak{S}_N,$$

so that the density operator  $|\Psi_N\rangle\langle\Psi_N| \in \mathcal{D}^s(\mathfrak{H}_N)$ .

**The quantum  $N$ -particle Hamiltonian** is the unbounded operator on the  $N$ -particle Hilbert space  $\mathfrak{H}_N = L^2(\mathbf{R}^{dN})$

$$\sum_{k=1}^N -\frac{\hbar^2}{2m} \Delta_{x_k} + \sum_{1 \leq k < l \leq N} V(x_k - x_l).$$

Of course, it is assumed that  $N \geq 2$ .

Since the  $N$  particles are identical, the total mass of the system is  $M = Nm$ , so that the “energy per particle” is

$$\frac{1}{N} \left( \sum_{k=1}^N -\frac{\hbar^2}{2m} \Delta_{x_k} + \sum_{1 \leq k < l \leq N} V(x_k - x_l) \right) = \sum_{k=1}^N -\frac{\hbar^2}{2M} \Delta_{x_k} + \frac{1}{N} \sum_{1 \leq k < l \leq N} V(x_k - x_l).$$

Henceforth we set  $M = 1$  without loss of generality, and consider as the quantum Hamiltonian the energy per particle, i.e.

$$\mathbf{H}_N := \sum_{k=1}^N -\frac{\hbar^2}{2} \Delta_{x_k} + \frac{1}{N} \sum_{1 \leq k < l \leq N} V(x_k - x_l).$$

We shall adopt the following **assumptions on the interaction potential**:

$$V \in C^{1,1}(\mathbf{R}^d) \quad \text{with } V^- \in L^\infty(\mathbf{R}^d) \quad \text{and} \quad V(z) = V(-z) \in \mathbf{R}.$$

Hence

$$\mathbf{H}_N \text{ is self-adjoint on } \mathfrak{H}_N \text{ with domain } \text{Dom}(\mathbf{H}_N) \supset H^2(\mathbf{R}^{dN}).$$

By Stone’s theorem, the quantum Hamiltonian  $\mathbf{H}_N$  generates a unitary group  $\mathcal{U}_N(t) := e^{-it\mathbf{H}_N/\hbar}$  on  $\mathfrak{H}_N$ .

The operator  $R_{\hbar,N}(t) = \mathcal{U}_N(t)R_N^{in}\mathcal{U}_N(t)^*$  solves **the  $N$ -particle von Neumann equation**

$$i\hbar\partial_t R_{\hbar,N}(t) = [\mathbf{H}_N, R_{\hbar,N}(t)], \quad R_{\hbar,N}(0) = R_N^{in} \in \mathcal{D}(\mathfrak{H}_N).$$

If  $R_N^{in}$  is a pure state, meaning that  $R_N^{in} = |\Psi_N^{in}\rangle\langle\Psi_N^{in}|$ , then

$$R_{\hbar,N}(t) = |\Psi_{\hbar,N}(t)\rangle\langle\Psi_{\hbar,N}(t)|$$

where  $\Psi_{\hbar,N}(t, \cdot) := \mathcal{U}_N(t)\Psi_N^{in}$  is the solution of the  $N$ -particle Schrödinger equation

$$i\hbar\Psi_{\hbar,N}(t, \cdot) = \mathbf{H}_N\Psi_{\hbar,N}(t, \cdot), \quad \Psi_{\hbar,N}(0, \cdot) = \Psi_N^{in}.$$

We leave it to the reader, as an easy exercise, to check that

$$U_\sigma\mathcal{U}_N(t) = \mathcal{U}_N(t)U_\sigma \quad \text{for all } \sigma \in \mathfrak{S}_N \text{ and all } t \in \mathbf{R}.$$

Therefore

$$R_N^{in} \in \mathcal{D}^s(\mathfrak{H}_N) \implies R_{\hbar,N}(t) = \mathcal{U}_N(t)R_N^{in}\mathcal{U}_N(t)^* \in \mathcal{D}^s(\mathfrak{H}_N) \quad \text{for all } t \in \mathbf{R}.$$

Henceforth, we are interested in situations where  $N \gg 1$  (mean-field regime) and  $\hbar \ll 1$  (semiclassical regime).

2.3.2. *Mean-Field Equations.* The purpose of the mean-field limit is to replace the description of a  $N$ -particle system by the quantum Hamiltonian  $\mathbf{H}_N$  — involving functions of  $N$  variable in  $L^2(\mathbf{R}^{dN})$ , therefore functions of  $dN \gg 1$  variables, by an equation, or a system of equations posed on the single-particle phase space  $\mathfrak{H}$  instead of  $\mathfrak{H}_N$ . The physical idea leading to these mean-field equations can be expressed as follows: one seeks to write an equation governing the evolution of a single, “typical” particle. This typical particle is driven by the interaction with the  $N - 1$  other particles, approximated as follows: call  $\rho(t, x)$  the single-particle **density function**, i.e.

$$\rho(t, x) = r(t, x, x)$$

where  $r(t, x, y)$  is an integral kernel of  $R(t) \in \mathcal{D}(\mathfrak{H})$ , the density operator of the typical, single particle, such that  $z \mapsto r(t, x + z, x)$  belongs to  $C_b(\mathbf{R}^d; L^1(\mathbf{R}^d))$  for each  $t \in \mathbf{R}$  (see Lecture I, Quiz 4). Then, one expects that

$$\begin{aligned} \frac{1}{N} \sum_{1 \leq j \leq N} V(x_j - y) &\simeq \int_{\mathbf{R}^d} V(x - y) \rho(t, x) dx = (V \star \rho(t, \cdot))(y) \\ &= \text{trace}(V(\cdot - y)R(t)) =: V_{R(t)}(y) \end{aligned}$$

as  $N \rightarrow \infty$ . This time-dependent potential  $V_{R(t)}$  is usually referred to as the mean-field, self-consistent potential. Then one can write the equation governing the (quantum) evolution of  $R(t)$ : this is a von Neumann equation where the potential is the mean-field potential, i.e.

$$i\hbar \partial_t R(t) = [-\frac{1}{2}\hbar^2 \Delta + V_{R(t)}, R(t)].$$

This is a quantum dynamical equation, referred to as the (time-dependent) **Hartree equation**, set on the single-particle Hilbert space  $\mathfrak{H} = L^2(\mathbf{R}^d)$ , instead of the  $N$ -particle Hilbert space  $\mathfrak{H}_N = L^2(\mathbf{R}^{dN})$ . The advantage of this description is obvious: one has to manipulate wave functions depending on  $d$  space variables ( $d \leq 3$  in practice) instead of  $Nd$  space variables, which is untractable. The drawback is that the mean-field equation is nonlinear (but the nonlinearity is relatively mild since it involves a convolution), and that it is only an approximation of the true,  $N$ -particle dynamics.

In the case of pure states, the (time-dependent) Hartree equation takes the form

$$i\hbar \partial_t \psi(t, x) = -\frac{1}{2}\hbar^2 \Delta_x \psi(t, x) + \psi(t, x) V \star_x |\psi|^2(t, x), \quad x \in \mathbf{R}^d,$$

and  $R(t) = |\psi(t, \cdot)\rangle\langle\psi(t, \cdot)|$ .

In the classical setting, the idea is the same: the classical mean-field potential is defined in terms of the single-particle distribution function  $f \equiv f(t, x, \xi)$  by the formula

$$V_f(t, x) := \iint_{\mathbf{R}^d \times \mathbf{R}^d} V(x - y) f(t, y, \eta) dy d\eta = (V \star f(t, \cdot, \cdot))(x).$$

Then, the mean-field equation, known as the **Vlasov equation** is written in terms of the mean-field Hamiltonian with the usual single particle Poisson bracket:

$$\begin{aligned} \partial_t f(t, x, \xi) + \underbrace{\xi \cdot \nabla_x f(t, x, \xi) - \nabla_x V_f(t, x) \cdot \nabla_\xi f(t, x, \xi)}_{= \{\frac{1}{2}|\xi|^2 + V_f(t, x), f(t, x, \xi)\}} &= 0. \end{aligned}$$

2.3.3. *Uniform in  $\hbar$  Error Bounds for the Quantum Mean-Field Limit.* In this section, we present error bounds comparing the solution of the  $N$ -particle von Neumann equation and the solution of the quantum (Hartree), or the classical (Vlasov) mean-field equations.

We begin with the error bound for the quantum mean-field limit (from the  $N$ -particle von Neumann equation to the Hartree equation). This is the upper horizontal arrow in the diagram above.

**Theorem 8.** Assume that  $V \in C^{1,1}(\mathbf{R}^d)$  is a real-valued function such that

$$V(z) = V(-z) \geq -M \quad \text{for some } M > 0,$$

and set

$$L := 2(1 + 4\text{Lip}(\nabla V)^2).$$

Choose an initial (1-particle) distribution function  $f^{in}$  such that  $f^{in} dx d\xi \in \mathcal{P}_2(\mathbf{R}^{2d})$ . Set the Hartree initial data to be  $R^{in} = \mathcal{T}[f^{in}]$ , and the  $N$ -particle initial data to be  $R_N^{in} = (R^{in})^{\otimes N} = \mathcal{T}[(f^{in})^{\otimes N}]$ . Let  $t \mapsto R(t)$  be the solution of the time-dependent Hartree equation with initial data  $R^{in}$ , while  $t \mapsto \mathcal{U}_N(t)\mathcal{T}[(f^{in})^{\otimes N}]\mathcal{U}_N(t)^*$  is the solution of the  $N$ -particle von Neumann equation with initial data  $R_N^{in}$ .

Then, for each  $t > 0$ ,

$$\frac{\mathfrak{d}(R(t)^{\otimes N}, \mathcal{U}_N(t)\mathcal{T}[(f^{in})^{\otimes N}]\mathcal{U}_N(t)^*)^2}{N} \leq 2d\hbar e^{Lt} + \frac{8\|\nabla V\|_{L^\infty}}{N-1} \frac{e^{Lt}-1}{L}.$$

Next we study the error bound for the joint mean-field and classical limits (from the  $N$ -particle von Neumann equation to the Vlasov equation). This is the diagonal arrow in the diagram above.

**Theorem 9.** Under the same assumptions as in Theorem 8, choose an initial (1-particle) distribution function  $f^{in}$  such that  $f^{in} dx d\xi \in \mathcal{P}_2(\mathbf{R}^{2d})$ . Set the  $N$ -particle initial data to be  $R_N^{in} = \mathcal{T}[(f^{in})^{\otimes N}]$ . Let  $t \mapsto f(t, \cdot, \cdot)$  be the solution of the Vlasov equation with initial data  $f^{in}$ , while the solution of the  $N$ -particle von Neumann equation with initial data  $R_N^{in}$  is  $t \mapsto \mathcal{U}_N(t)\mathcal{T}[(f^{in})^{\otimes N}]\mathcal{U}_N(t)^*$ .

Then, for each  $t > 0$ ,

$$\frac{\mathfrak{d}(f(t, \cdot, \cdot)^{\otimes N}, \mathcal{U}_N(t)\mathcal{T}[(f^{in})^{\otimes N}]\mathcal{U}_N(t)^*)^2}{N} \leq d\hbar e^{Lt} + \frac{8\|\nabla V\|_{L^\infty}}{N-1} \frac{e^{Lt}-1}{L}.$$

Several comments are in order before going further.

**Remarks.**

(1) In both Theorems 8 and 9, we have restricted our attention to Toeplitz initial density operators. The reason for using such well-prepared initial data comes from statement (1) in Theorem 1: if  $T$  is a Toeplitz operator in  $\mathcal{D}_2(\mathfrak{H})$ , then

$$\mathfrak{d}(T, T) = \sqrt{2d\hbar} = \min_{R \in \mathcal{D}_2(\mathfrak{H})} \mathfrak{d}(R, R).$$

Similarly, if  $f$  is the symbol of  $T$  — i.e. if  $T = \mathcal{T}[f]$  — then

$$\mathfrak{d}(f, T) = \sqrt{d\hbar} = \min_{\substack{g \in \mathcal{P}_2(\mathbf{R}^{2d}) \\ R \in \mathcal{D}_2(\mathfrak{H})}} \mathfrak{d}(g, R).$$

Therefore, for this choice of initial data, the first terms in the right-hand side of the upper bounds in both theorems are as small as possible.

(2) The reason for considering the expressions

$$\frac{\mathfrak{d}(R(t)^{\otimes N}, \mathcal{U}_N(t) \mathcal{T}[(f^{in})^{\otimes N}] \mathcal{U}_N(t)^*)^2}{N}$$

and

$$\frac{\mathfrak{d}(f(t, \cdot, \cdot))^{\otimes N}, \mathcal{U}_N(t) \mathcal{T}[(f^{in})^{\otimes N}] \mathcal{U}_N(t)^*)^2}{N},$$

instead of

$$\mathfrak{d}(R(t)^{\otimes N}, \mathcal{U}_N(t) \mathcal{T}[(f^{in})^{\otimes N}] \mathcal{U}_N(t)^*)^2$$

and

$$\mathfrak{d}(f(t, \cdot, \cdot))^{\otimes N}, \mathcal{U}_N(t) \mathcal{T}[(f^{in})^{\otimes N}] \mathcal{U}_N(t)^*)^2$$

comes from the fact that the quantum-to-quantum transport cost and the classical-to-quantum transport cost in  $\mathfrak{H}_N$  are sums of  $N$  terms, respectively

$$\sum_{k=1}^N I_{\mathfrak{H} \otimes \mathfrak{H}}^{\otimes(k-1)} \otimes C_h \otimes I_{\mathfrak{H} \otimes \mathfrak{H}}^{\otimes(N-k)},$$

and

$$\sum_{k=1}^N I_{\mathfrak{H}}^{\otimes(k-1)} \otimes c_h(x_k, \xi_k) \otimes I_{\mathfrak{H}}^{\otimes(N-k)}.$$

(3) For all  $R_N \in \mathcal{D}^s(\mathfrak{H}_N)$ , we consider its  $k$ -th marginal  $R_{N:k} \in \mathcal{D}^s(\mathfrak{H}_k)$  defined by the following prescription

$$\text{trace}_{\mathfrak{H}_k}(R_{N:k} A) = \text{trace}_{\mathfrak{H}_N}(R_N(A \otimes I_{\mathfrak{H}_{N-k}}^{N-k})), \quad \text{for all } A \in \mathcal{L}(\mathfrak{H}_k).$$

**Quiz 19.** Let  $R_N, S_N \in \mathcal{D}_2^s(\mathfrak{H}_N)$  and let  $f_N$  be a symmetric probability density on  $(\mathbf{R}^d \times \mathbf{R}^d)^N$  with finite second order moments. Prove that

$$\mathfrak{d}(R_{N:1}, S_{N:1})^2 \leq \dots \leq \frac{\mathfrak{d}(R_{N:k}, S_{N:k})^2}{k} \leq \dots \leq \frac{\mathfrak{d}(R_N, S_N)^2}{N}, \quad 1 \leq k \leq N,$$

and that

$$\mathfrak{d}(f_{N:1}, S_{N:1})^2 \leq \dots \leq \frac{\mathfrak{d}(f_{N:k}, S_{N:k})^2}{k} \leq \dots \leq \frac{\mathfrak{d}(f_N, S_N)^2}{N}, \quad 1 \leq k \leq N,$$

where

$$f_{N:k}(x_1, \xi_1, \dots, x_k, \xi_k) := \int_{(\mathbf{R}^d \times \mathbf{R}^d)^{N-k}} f_N(x_1, \xi_1, \dots, x_N, \xi_N) dx_{k+1} d\xi_{k+1} \dots dx_N d\xi_N.$$

(Hint: use the formulas for the quantum-to-quantum and the classical-to-quantum costs in (2)).

In particular, Theorems 8 and 9 imply the following result on the first marginal of the  $N$ -particle density operator.

**Corollary 10.** Under the same assumptions as in Theorems 8 and 9, one has

$$\begin{aligned} \mathfrak{d}(R(t), (\mathcal{U}_N(t) \mathcal{T}[(f^{in})^{\otimes N}] \mathcal{U}_N(t)^*)_{:1})^2 &\leq 2d\hbar e^{Lt} + \frac{8\|\nabla V\|_{L^\infty}}{N-1} \frac{e^{Lt}-1}{L}, \\ \mathfrak{d}(f(t, \cdot, \cdot), (\mathcal{U}_N(t) \mathcal{T}[(f^{in})^{\otimes N}] \mathcal{U}_N(t)^*)_{:1})^2 &\leq d\hbar e^{Lt} + \frac{8\|\nabla V\|_{L^\infty}}{N-1} \frac{e^{Lt}-1}{L}. \end{aligned}$$

(4) One can obviously use Theorems 8 and 9, together with Corollary 10 and Theorem 1 (2) in order to compare the Vlasov solution and the Husimi transform of the Hartree solution to the Husimi transform of the first marginal of the  $N$ -particle density operator.



**Corollary 11.** Under the same assumptions as in Theorems 8 and 9, one has

$$\begin{aligned} \mathcal{W}_2(\mathcal{H}[R(t)], \mathcal{H}[(\mathcal{U}_N(t)\mathcal{T}[(f^{in})^{\otimes N}]\mathcal{U}_N(t)^*)_{:1}])^2 &\leq 2d\hbar(e^{Lt} + 1) + \frac{8\|\nabla V\|_{L^\infty}}{N-1} \frac{e^{Lt}-1}{L}, \\ \mathcal{W}_2(f(t, \cdot, \cdot), \mathcal{H}[(\mathcal{U}_N(t)\mathcal{T}[(f^{in})^{\otimes N}]\mathcal{U}_N(t)^*)_{:1}])^2 &\leq d\hbar(e^{Lt} + 1) + \frac{8\|\nabla V\|_{L^\infty}}{N-1} \frac{e^{Lt}-1}{L}. \end{aligned}$$

(5) Since density operators are characterized by their Husimi transforms (see Quiz 20 below), and since the Husimi transform of a density operator is a probability density, a natural idea to define optimal transport distances between density operators is to use the formula

$$d_{ZS}(\rho_1, \rho_2) = \mathcal{W}_2(\mathcal{H}[\rho_1], \mathcal{H}[\rho_2]), \quad \rho_1, \rho_2 \in \mathcal{D}_2(\mathfrak{H}).$$

This approach has been proposed by K. Życzkowski and W. Słomczyński [65].

This definition has some advantages over the pseudometric  $\mathfrak{d}$  discussed in these lectures

- (a) it is a bona fide metric (see quiz below), and
- (b) one is always dealing with probability densities, i.e. functions on phase space, which are easier to manipulate than operators.

Similarly, one might prefer considering

$$\mathcal{W}_2(f, \mathcal{H}[R]) \quad \text{instead of} \quad \mathfrak{d}(f, R)$$

for  $R \in \mathcal{D}_2(\mathfrak{H})$  and  $f$  a probability density on  $\mathcal{P}(\mathbf{R}^d \times \mathbf{R}^d)$  with finite second order moments. Indeed, with the first quantity, one transforms the density operator  $R$  into a probability density by means of the Husimi transform, and then compares this probability density with  $f$  by means of the Wasserstein distance. On the contrary, the second quantity compares two very different objects (a probability density and a density operator), which is like comparing apples with pears.

**Quiz 20: the Husimi transform is one-to-one.** Set  $\mathfrak{H} = L^2(\mathbf{R}^d)$ .

- (1) Let  $R \in \mathcal{D}_2(\mathfrak{H})$ . Prove that  $\mathcal{H}[R]$  is a probability density, and compute

$$\iint_{\mathbf{R}^d \times \mathbf{R}^d} (|q|^2 + |p|^2) \mathcal{H}[R](q, p) dq dp$$

in terms of  $R$ .

- (2) Let  $R, S \in \mathcal{D}_2(\mathfrak{H})$ , and assume that  $\mathcal{H}[R] = \mathcal{H}[S]$ . Prove that  $R = S$ . (Hint: let  $r \equiv r(y, y')$  be an integral kernel of  $R$ . Set

$$J(x, \xi) := \iint_{\mathbf{R}^d \times \mathbf{R}^d} r(y, y') e^{-(|y|^2 + |y'|^2)/2\hbar} e^{x \cdot (y+y') - i\xi \cdot (y-y')/\hbar} dy dy'.$$

Prove that  $J$  extends as a holomorphic function on  $\mathbf{C}^d \times \mathbf{C}^d$ , and therefore is uniquely determined by its restriction to  $\mathbf{R}^d \times \mathbf{R}^d$ . Conclude (a) by finding the formula relating  $\mathcal{H}[R]$  to  $J$ , and (b) by computing the integral kernel  $r$  of  $R$  in terms of  $J$ .)

However, there is a rather heavy price to pay with this approach, which is that the Husimi transform, and therefore  $d_{ZS}$  is not easy to propagate by usual quantum dynamics. This is easily explained if one returns to Quiz 11 (3) and (5): if  $t \mapsto R(t)$  is a time-dependent density operator, solution of the von Neumann equation

$$i\hbar \partial_t R(t) = [-\frac{1}{2}\hbar^2 \Delta + V, R(t)], \quad R(0) = R^{in},$$

we have seen that its Wigner transform  $W_\hbar[R(t)]$  satisfies the Wigner equation

$$(\partial_t + \xi \cdot \nabla_x) W_\hbar[R(t)](x, \xi) + \Theta[V] W_\hbar[R(t)](x, \xi) = 0,$$

and that

$$\mathcal{H}[R(t)] = \exp(\frac{\hbar}{4}\Delta_{x,\xi})W_{\hbar}[R(t)].$$

Therefore, the evolution of  $\mathcal{H}[R(t)]$  can be described as follows:

$$\mathcal{H}[R^{in}] \mapsto W[R^{in}] \mapsto W_{\hbar}[R(t)] \mapsto \mathcal{H}[R(t)].$$

The second arrow is the value at time  $t$  of the group generated by the Wigner equation, and the third arrow is  $\exp(\frac{\hbar}{4}\Delta_{x,\xi})$ . But the first arrow corresponds to expressing  $W_{\hbar}[R^{in}]$  in terms of  $\mathcal{H}[R^{in}]$  by solving

$$\exp(\frac{\hbar}{4}\Delta_{x,\xi})W_{\hbar}[R^{in}] = \mathcal{H}[R^{in}].$$

This first step corresponds to inverting the heat flow at time  $\hbar/4$ , hence the difficulty of this problem.

A more thorough discussion of the evolution of the Husimi transform of a solution of the von Neumann equation can be found in [5].

This discussion explains one of the advantages of the pseudometric  $\mathfrak{d}$  over  $d_{KS}$ , that is the remarkable simplicity of the propagation estimate stated in Theorem 6, and the fact that the bounds reported in that theorem are exactly the same as the pair dispersion estimate in classical mechanics (with the same potential in both dynamics). In other words, when compared by means of the pseudometric  $\mathfrak{d}$ , quantum particles behave as if they were points moving on classical phase-space trajectories. This picture of quantum dynamics is known to be wrong in general — see for instance Young's double-slit experiment and its discussion in chapter I, section 2 of [23] — but turns out to be true as far as the pseudodistance  $\mathfrak{d}$ , or its variants  $\mathfrak{d}_{\lambda}$  discussed in Theorem 6, are concerned. Accordingly, the assumptions on the potential used in Theorem 6 are those used in the definition of a classical dynamics by means of the Cauchy-Lipschitz theorem, which require in particular that  $V \in C^{1,1}(\mathbf{R}^d)$ , which are much more restrictive than the conditions on  $V$  which imply that  $-\frac{\hbar^2}{2m}\Delta + V$  has a self-adjoint extension as an unbounded operator on  $L^2(\mathbf{R}^d)$ .

After this long list of remarks, here is a quick sketch of the proof of Theorems 8. Theorem 9 is proved in essentially the same way, and the necessary modifications are left to the reader as a (relatively) easy exercise.

*Sketch of the proof of Theorem 8.* This proof is split in several steps.

Step 1. Pick  $\mathcal{Q}_N^{in} \in \mathcal{C}(R_N^{in}, (R^{in})^{\otimes N})$  and solve for  $\mathcal{Q}_N(t) \in \mathcal{D}(\mathfrak{H}_N \otimes \mathfrak{H}_N)$  the linear von Neumann equation with time-dependent potential

$$\begin{cases} i\hbar\partial_t \mathcal{Q}_N(t) = \left[ \mathbf{H}_N \otimes I_{\mathfrak{H}_N} + I_{\mathfrak{H}_N} \otimes \sum_{k=1}^N J_{k,N}(-\frac{\hbar^2}{2}\Delta + V_{R(t)}), \mathcal{Q}_N(t) \right], \\ \mathcal{Q}_N(0) = \mathcal{Q}_N^{in}, \end{cases}$$

with the notation

$$J_{k,N}A := I_{\mathfrak{H}}^{\otimes(k-1)} \otimes A \otimes I_{\mathfrak{H}}^{\otimes(N-k)}.$$

One checks, by taking partial traces and using uniqueness for the solution of the von Neumann equation with time-dependent potential, that

$$\mathcal{Q}_N(t) \in \mathcal{C}(R_N(t), R(t)^{\otimes N}), \quad \text{for all } t \geq 0,$$

where  $R(t)$  is the Hartree solution, while

$$R_N(t) := \mathcal{U}_N(t)R_N^{in}\mathcal{U}_N(t)^*.$$

Step 2. Let

$$D_N(t) := \frac{1}{N} \text{trace}_{\mathfrak{H}_N \otimes \mathfrak{H}_N} (\mathcal{Q}_N(t)^{\frac{1}{2}} C_h \mathcal{Q}_N(t)^{\frac{1}{2}}),$$

with  $C_h$  defined for all  $\Phi \equiv \Phi(X_N, Y_N) \in \mathcal{S}(\mathbf{R}^{2dN})$  by

$$C_h \Phi := \sum_{j=1}^N (|x_j - y_j|^2 \Phi - \hbar^2 (\text{div}_{x_j} - \text{div}_{y_j}) ((\nabla_{x_j} - \nabla_{y_j}) \Phi))(X_N, Y_N).$$

Step 3. By definition of  $D_N(t)$ , Step 1 shows that

$$D_N(t) \geq \frac{1}{N} \mathfrak{d}(R_N(t), R(t)^{\otimes N})^2, \quad \text{for all } t \geq 0.$$

Step 4. On the other hand

$$\begin{aligned} i\hbar \frac{dD_N}{dt} &= \frac{1}{N} \text{trace}_{\mathfrak{H}_N \otimes \mathfrak{H}_N} (\mathcal{Q}_N(t)^{\frac{1}{2}} [\mathbf{H}_N \otimes I_{\mathfrak{H}_N}, C_h] \mathcal{Q}_N(t)^{\frac{1}{2}}) \\ &+ \sum_{k=1}^N \frac{1}{N} \text{trace}_{\mathfrak{H}_N \otimes \mathfrak{H}_N} (\mathcal{Q}_N(t)^{\frac{1}{2}} [I_{\mathfrak{H}_N} \otimes J_{k,N} (-\frac{1}{2}\hbar^2 \Delta + V_{R(t)}), C_h] \mathcal{Q}_N(t)^{\frac{1}{2}}), \end{aligned}$$

and it remains to compute

$$Z_N = -\frac{i}{\hbar} \left[ \mathbf{H}_N \otimes I_{\mathfrak{H}_N} + I_{\mathfrak{H}_N} \otimes \sum_{k=1}^N J_{k,N} (-\frac{1}{2}\hbar^2 \Delta + V_{R(t)}), C_h \right].$$

Step 5. Using the notation  $A \vee B := AB + BA$ , one finds that

$$\begin{aligned} Z_N &= \sum_{j=1}^N (x_j - y_j) \vee (-i\hbar \nabla_{x_j} + i\hbar \nabla_{y_j}) \\ &+ \sum_{j=1}^N \frac{1}{N} \sum_{k=1}^N (-\nabla V(x_j - x_k) + \nabla V_{R(t)}(y_j)) \vee (-i\hbar \nabla_{x_j} + i\hbar \nabla_{y_j}), \end{aligned}$$

and one uses the elementary operator inequality (see Quiz 7 (2))

$$AB^* + BA^* \leq AA^* + BB^*$$

to prove that

$$Z_N \leq 2C_h + \sum_{j=1}^N \left| \nabla V_{R(t)}(y_j) - \frac{1}{N} \sum_{k=1}^N \nabla V(x_j - x_k) \right|^2.$$

Split the summand so as to involve the difference between the  $N$ -body and the mean-field potentials on the  $y_j$  variables only

$$\begin{aligned} Z_N &\leq 2C_h + 2 \sum_{j=1}^N \left| \frac{1}{N} \sum_{k=1}^N (\nabla V(y_j - y_k) - \nabla V(x_j - x_k)) \right|^2 \\ &+ 2 \sum_{j=1}^N \left| \nabla V_{R(t)}(y_j) - \frac{1}{N} \sum_{k=1}^N \nabla V(y_j - y_k) \right|^2 \\ &\leq 2C_h + \frac{2}{N} \text{Lip}(\nabla V)^2 \underbrace{\sum_{j,k=1}^N |(y_j - y_k) - (x_j - x_k)|^2}_{\leq 4NC_h} \\ &+ 2 \sum_{j=1}^N \left| \nabla V_{R(t)}(y_j) - \frac{1}{N} \sum_{k=1}^N \nabla V(y_j - y_k) \right|^2. \end{aligned}$$

Step 6. It remains to bound

$$\begin{aligned} & \sum_{j=1}^N \text{trace}_{\mathfrak{H}_N \otimes \mathfrak{H}_N} \left( \mathcal{Q}_N(t)^{\frac{1}{2}} \left| \nabla V_{R(t)}(y_j) - \frac{1}{N} \sum_{k=1}^N \nabla V(y_j - y_k) \right|^2 \mathcal{Q}_N(t)^{\frac{1}{2}} \right) \\ &= \sum_{j=1}^N \text{trace}_{\mathfrak{H}_N} \left( \left| \nabla V_{R(t)}(y_j) - \frac{1}{N} \sum_{k=1}^N \nabla V(y_j - y_k) \right|^2 R(t)^{\otimes N} \right). \end{aligned}$$

This is easy since all the potentials considered involve only the variables  $y_j$  for  $j = 1, \dots, N$ , so that  $\mathcal{Q}_N(t)$  can be replaced with the **factorized** density  $R(t)^{\otimes N}$ , where  $R(t)$  is the Hartree solution. Here is the argument.

$$\begin{aligned} & \text{trace}_{\mathfrak{H}_N} \left( \left| \nabla V_{R(t)}(y_1) - \frac{1}{N} \sum_{k=1}^N \nabla V(y_1 - y_k) \right|^2 R(t)^{\otimes N} \right) \\ &= \text{trace}_{\mathfrak{H}_N} \left( \left| \frac{1}{N} \sum_{k=1}^N (\nabla V_{R(t)}(y_1) - \nabla V(y_1 - y_k)) \right|^2 R(t)^{\otimes N} \right) \\ &= \frac{1}{N^2} \text{trace}_{\mathfrak{H}_N} \left( \sum_{k=1}^N |\nabla V_{R(t)}(y_1) - \nabla V(y_1 - y_k)|^2 R(t)^{\otimes N} \right) \\ &+ 2 \sum_{1 \leq k < m \leq N} \text{trace}_{\mathfrak{H}_N} \left( \frac{\nabla V_{R(t)}(y_1) - \nabla V(y_1 - y_k)}{N} \cdot \frac{\nabla V_{R(t)}(y_1) - \nabla V(y_1 - y_m)}{N} R(t)^{\otimes N} \right) \\ &= \frac{1}{N^2} \text{trace}_{\mathfrak{H}_N} \left( \sum_{k=1}^N |\nabla V_{R(t)}(y_1) - \nabla V(y_1 - y_k)|^2 R(t)^{\otimes N} \right) \leq \frac{4}{N} \|\nabla V\|_{L^\infty(\mathbf{R}^d)}^2, \end{aligned}$$

since

$$\text{trace}_{\mathfrak{H}_N} \left( (\nabla V_{R(t)}(y_1) - \nabla V(y_1 - y_k)) \cdot (\nabla V_{R(t)}(y_1) - \nabla V(y_1 - y_m)) R(t)^{\otimes N} \right) = 0$$

for all  $m > k \geq 1$ .

Step 7. Hence

$$Z_N \leq 2(1 + 4 \text{Lip}(\nabla V)^2) C_h + 2 \sum_{j=1}^N \left| \nabla V_{R(t)}(y_j) - \frac{1}{N} \sum_{k=1}^N \nabla V(x_j - x_k) \right|^2,$$

and therefore

$$\begin{aligned} \frac{d}{dt} D_N(t) &= \text{trace}_{\mathfrak{H}_N \otimes \mathfrak{H}_N} (\mathcal{Q}(t)^{\frac{1}{2}} Z_N \mathcal{Q}(t)^{\frac{1}{2}}) \\ &\leq 2(1 + 4 \text{Lip}(\nabla V)^2) D_N(t) + \frac{8}{N} \|\nabla V\|_{L^\infty(\mathbf{R}^d)}^2. \end{aligned}$$

By Step 3 and the Gronwall inequality, one has

$$\frac{1}{N} \mathfrak{d}(R_N(t), R(t)^{\otimes N})^2 \leq D_N(t) \leq D_N(0) e^{Lt} + \frac{8}{N} \|\nabla V\|_{L^\infty(\mathbf{R}^d)}^2 \frac{e^{Lt} - 1}{L}.$$

This inequality holds for each initial coupling  $\mathcal{Q}_N^{in}$  of  $(R^{in})^{\otimes N}$  with  $R_N^{in}$ , so that

$$\frac{1}{N} \mathfrak{d}(R_N(t), R(t)^{\otimes N})^2 \leq \frac{1}{N} \mathfrak{d}(R_N^{in}, (R^{in})^{\otimes N})^2 e^{Lt} + \frac{8}{N} \|\nabla V\|_{L^\infty(\mathbf{R}^d)}^2 \frac{e^{Lt} - 1}{L}.$$

Specializing this inequality to

$$R^{in} = \mathcal{T}[f^{in}] \quad \text{and} \quad R_N^{in} = (R^{in})^{\otimes N}$$

and using Theorem 1 (1) leads to the announced result.  $\square$

**2.4. Time-Splitting Numerical Schemes for Quantum Mechanics.** In this section, we study the numerical error relative to some time-discretizations of quantum dynamical equations.

Consider the **von Neumann equation** with unknown  $R(t) \in \mathcal{D}(\mathfrak{H})$

$$i\hbar\partial_t R = [-\frac{1}{2}\hbar^2\Delta + V, R], \quad R|_{t=0} = R^{in},$$

where the potential  $V$  is chosen so that the operator

$$\mathbf{H}_\hbar := -\frac{1}{2}\hbar^2\Delta + V$$

has a self-adjoint (unbounded) extension to  $\mathfrak{H} = L^2(\mathbf{R}^d)$ .

Our purpose is to obtain uniform in  $\hbar$  error bounds for various time-splitting schemes for the von Neumann equation. The simplest example of time-splitting scheme for the von Neumann equation is as follows:

**Time-split Heisenberg equation.** Starting from  $R^0 = R^{in} \in \mathcal{D}(\mathfrak{H})$ , we define a sequence  $R^m(t) \in \mathcal{D}(\mathfrak{H})$  for  $m \in \frac{1}{2}\mathbf{N}$  by the following formulas:

$$\begin{aligned} R^{n+\frac{1}{2}} &= \exp(\frac{i\hbar\Delta t}{2}\Delta)R^n \exp(-\frac{i\hbar\Delta t}{2}\Delta), \\ R^{n+1} &= \exp(\frac{\Delta t}{i\hbar}V)R^{n+\frac{1}{2}} \exp(-\frac{\Delta t}{i\hbar}V). \end{aligned}$$

The analysis of time-splitting schemes like this one has been studied in detail, for instance by [27]. This particular time-splitting scheme is known as the **Lie-Trotter splitting method**.

**Error bound.** [27] Pick  $R^{in} \in \mathcal{D}(\mathfrak{H})$  such that

$$\|\langle \hbar D_x \rangle R^{in} \langle \hbar D_x \rangle\|_1 =: M < \infty \quad \text{with } \langle \hbar D_x \rangle := (1 - \hbar^2 \Delta_x)^{\frac{1}{2}}.$$

Then, for each integer  $n \geq 0$ , one has

$$\|R(n\Delta t) - R^n\|_1 \leq C(M, \|V\|_{W^{2,\infty}}) \frac{\Delta t}{\hbar}.$$

This error bound is obviously **not uniform** as  $\hbar \rightarrow 0$ . The convergence of this scheme requires that  $\Delta t \ll \hbar$ , which is obviously very costly in the semiclassical regime, where  $\hbar \ll 1$ .

In the sequel, we seek to obtain an error bound for the Lie-Trotter time-splitting scheme that is uniform in  $\hbar$  by the methodology of **asymptotic preserving (AP)** schemes. The idea is to use the Descombes-Thalhammer error bound for  $\hbar \geq 1$ , and, in the case where  $\hbar \ll 1$ , to use the same time-splitting scheme for the classical Liouville equation, which is the classical dynamical equation obtained by passing to the  $\hbar \rightarrow 0$  limit in the von Neumann equation, as explained for instance by the Lions-Paul theorem recalled above. This last error estimate involves comparing the numerical quantum and the classical solutions, and the exact classical and numerical solutions on the other hand. This leads to two different error bounds depending on  $\hbar$ , and one hopes to get a uniform in  $\hbar$  error estimate by optimizing in  $\hbar > 0$ . This is summarized by the diagram in Figure 5.

The first result on the uniform in  $\hbar$  convergence of the Lie-Trotter method for the von Neumann equation was obtained in [6], without any error bound. The goal of the present section is to provide such uniform in  $\hbar$  error estimates by means of the  $\mathfrak{d}$  pseudometric.

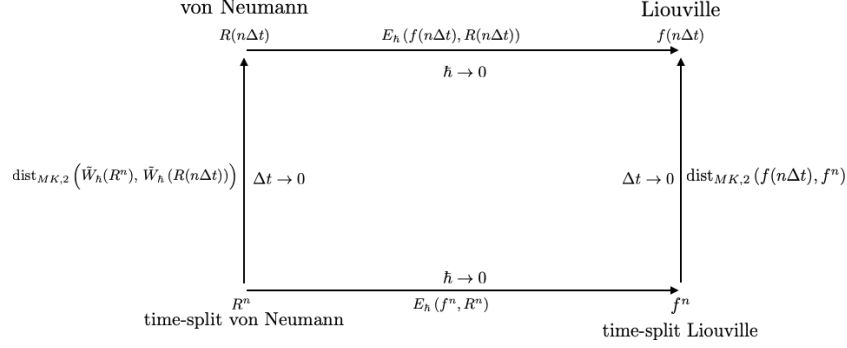


FIGURE 5. The horizontal arrows represent the semiclassical limit  $\hbar \ll 1$  and the vertical arrows the convergence of the numerical scheme  $\Delta t \ll 1$ .

2.4.1. *Lie-Trotter Time Splitting for the Liouville Equation.* As recalled above, the classical dynamical equation corresponding to the von Neumann equation is the **Liouville equation** with unknown  $f \equiv f(t, x, \xi) \geq 0$ , written as

$$\partial_t f(t, x, \xi) + \{H(x, \xi), f(t, x, \xi)\} = 0, \quad f|_{t=0} = f^{in},$$

where the classical Hamiltonian is

$$H(x, \xi) := \frac{1}{2}|\xi|^2 + V(x).$$

We shall assume here that  $V \in W^{2,\infty}(\mathbf{R}^d)$ , and that  $V \geq -M$  on  $\mathbf{R}^d$  for some constant  $M > 0$ .

This is a first order PDE that is solved by the **method of characteristics**. If one denotes by  $\Phi_t$  the Hamiltonian flow generated by  $H$ , one finds that

$$f(t, x, \xi) = f^{in}(\Phi_{-t}(x, \xi)), \quad x, \xi \in \mathbf{R}^d, \quad t \in \mathbf{R}.$$

The **Lie-Trotter time-splitting** method for the Liouville equation is

$$\begin{aligned} f^{n+\frac{1}{2}}(y, \eta) &= f^n \circ K_{-\Delta t}(y, \eta), \\ f^{n+1}(y, \eta) &= f^{n+\frac{1}{2}} \circ P_{-\Delta t}(y, \eta), \end{aligned}$$

starting from  $f^0 := f^{in}$ , where the maps  $K_{-\Delta t}$  and  $P_{-\Delta t}$  are defined by the following formulas:

$$K_t(y, \eta) := (y + t\eta, \eta), \quad P_t(y, \eta) := (y, \eta - t\nabla V(y)).$$

The **error bound for the Lie-Trotter splitting scheme for the Liouville equation** is formulated below in terms of the Wasserstein distance of exponent 2.

Setting

$$(X_t, \Xi_t) := \Phi_t(x, \xi) \quad \text{and} \quad (Y_t, H_t) := P_t \circ K_t(y, \eta),$$

one finds that

$$\begin{aligned} |X_t - Y_t|^2 + |\Xi_t - H_t|^2 &\leq (|x - y|^2 + |\xi - \eta|^2)e^{(2+\Lambda)|t|} \\ &\quad + \frac{e^{(2+\Lambda)|t|} - 1}{2 + \Lambda} \frac{9}{4} \Lambda^2 \left(\frac{1}{2} + \Lambda\right)^2 t^2 (1 + |y|^2 + |\eta|^2), \end{aligned}$$

with

$$\Lambda := \max(1, E, \|\nabla^2 V\|_{L^\infty}), \quad E := |\nabla V(0)|.$$

**Quiz 21.** Prove the inequality above. (Hint: write a differential system satisfied by  $t \mapsto (Y_t, H_t)$ , and compare the trajectories of that system with those of the Hamiltonian system corresponding to  $H(x, \xi)$  by means of the Gronwall inequality.)

The inequality above implies the following error estimate for the Lie-Trotter scheme in the case of the Liouville equation.

**Lemma 12.** Let  $V \in W^{2,\infty}(\mathbf{R}^d)$  satisfy  $V \geq -M$  for some  $M > 0$ . Assume that  $f^{in}$  is a probability density on  $\mathbf{R}^{2d}$  such that

$$\iint_{\mathbf{R}^d \times \mathbf{R}^d} (|x|^2 + |\xi|^2) f^{in}(x, \xi) dx d\xi < \infty.$$

Then, for each  $\Delta t \in (0, 1)$  and each  $n = 0, \dots, [T/\Delta t]$ , one has

$$\mathcal{W}_2(f^n, f(n\Delta t)) \leq C_T[\Lambda, E, f^{in}] \Delta t,$$

where  $C_T[\Lambda, E, f^{in}]$  is a positive constant depending only on the computing time  $T$ , on the constants  $\Lambda$  and  $E$  defined above, and on the initial data  $f^{in}$ .

**2.4.2. Lie-Trotter Time Splitting for the Liouville Equation.** At this point, we use Theorem 6 first with  $\lambda = m = 1$  and  $V = 0$ , so that

$$\mathfrak{d}(f^n \circ K_{-\Delta t}, R^{n+\frac{1}{2}}) \leq \mathfrak{d}(f^n, R^n) e^{\frac{\Delta t}{2}}.$$

Next we use Theorem 6, this time with  $\lambda = 1$  and  $m \rightarrow +\infty$ , so that

$$\mathfrak{d}(f^{n+\frac{1}{2}} \circ P_{-\Delta t}, R^{n+1}) \leq \mathfrak{d}(f^{n+\frac{1}{2}}, R^{n+\frac{1}{2}}) e^{\frac{\Delta t}{2} \text{Lip}(\nabla V)}.$$

Hence, for all integer  $n \geq 0$ , one has

$$\mathfrak{d}(f^n, R^n) \leq \mathfrak{d}(f^{in}, R^{in}) \exp\left(\frac{1}{2} n \Delta t (1 + \text{Lip}(\nabla V))\right).$$

On the other hand, applying Theorem 6 with  $m = 1$  shows that, for all  $t \geq 0$ ,

$$\begin{aligned} \mathfrak{d}(f(t, \cdot, \cdot), R(t)) &= \mathfrak{d}(f^{in} \circ \Phi_{-t}, e^{-it\mathbf{H}_h/h} R^{in} e^{it\mathbf{H}_h/h}) \\ &\leq \mathfrak{d}(f^{in}, R^{in}) \exp\left(\frac{1}{2} t (1 + \text{Lip}(\nabla V))\right). \end{aligned}$$

Putting together these two bounds, and the error estimate in the previous section, we arrive at the following result.

**Theorem 13.** Let  $V \in W^{2,\infty}(\mathbf{R}^d)$  satisfy  $V \geq -M$  for some  $M > 0$ . Let  $f^{in}$  be a probability density on  $\mathbf{R}^d \times \mathbf{R}^d$  such that

$$\iint_{\mathbf{R}^d \times \mathbf{R}^d} (|x|^2 + |\xi|^2) f^{in}(x, \xi) dx d\xi < \infty.$$

Set  $R^{in} = \mathcal{T}[f^{in}]$ . Then, the Lie-Trotter splitting scheme for the Heisenberg equation satisfies the uniform as  $\hbar \rightarrow 0$  error bound

$$\mathfrak{d}(R^n, R(n\Delta t)) \leq C_T[\Lambda, \|V\|_{W^{2,\infty}}, f^{in}] \Delta t + 2\sqrt{d\hbar} e^{\frac{T}{2}(1+\text{Lip}(\nabla V))}.$$

This implies the uniform in  $\hbar$  convergence rate

$$\sup_{\max(2\|\phi\|_{L^\infty}, \text{Lip}(\phi)) \leq 1} \left| \int_{\mathbf{R}^{2d}} \phi(x, \xi) (\mathcal{H}[R^n] - \mathcal{H}[R(n\Delta t)]) dx d\xi \right| \leq C'_T[\Lambda, \|V\|_{W^{2,\infty}}, f^{in}] \Delta t^{1/3}.$$

The uniform as  $\hbar \rightarrow 0$  error estimate follows from the two bounds obtained in this section before Theorem 13, and the error bound for the Lie-Trotter method for the Liouville equation in the previous section (Lemma 12). One uses also the following triangle inequality, which will be discussed in Lecture III:

$$\mathfrak{d}(R, S) \leq \mathfrak{d}(R, f) + \mathcal{W}_2(f, g) + \mathfrak{d}(g, S).$$

The uniform in  $\hbar$  convergence rate follows from optimizing between the uniform as  $\hbar \rightarrow 0$  bound and the Descombes-Thalhammer bound.

There is a similar result with higher order splitting formulas, such as Strang's splitting scheme, leading to a uniform  $O(\Delta t^{2/3})$  estimate. The interested reader is referred to [36] for these higher order estimates, and for the missing details on the case of the Lie-Trotter splitting method.

**2.5. Observation Inequalities for Quantum Dynamics.** The observation problem for a general PDE can be formulated as follows: let  $P(x, \partial_x)$  be a (linear) partial differential operator on some open set  $\Omega \subset \mathbf{R}^d$ , and let  $\omega$  be an open subset of  $\Omega$ . Is a solution  $u$  of the PDE  $P(x, \partial_x)u(x) = 0$  for all  $x \in \Omega$  completely determined by its restriction to  $\omega$ ?

**Example.** Here is an elementary example. Set  $d = 2$ , and identify  $\mathbf{R}^2$  to  $\mathbf{C}$  (by sending  $(x, y) \in \mathbf{R}^2$  to  $z = x + iy \in \mathbf{C}$ ). Set

$$P(x, y, D_x, D_y) = \bar{\partial} = \frac{1}{2}(\partial_x + i\partial_y),$$

(the Cauchy-Riemann operator) and set  $\omega \subset \Omega \subset \mathbf{C}$  be (nonempty) open sets in  $\mathbf{C}$ . Any distribution  $u$  on  $\Omega$  such that  $\bar{\partial}u = 0$  on  $\Omega$  is a holomorphic function on  $\Omega$ . The restriction of  $u$  to the connected component of  $\omega$  in  $\Omega$  is uniquely determined by the restriction  $u|_\omega$ . In particular, if  $u|_\omega = 0$ , then  $u = 0$  on  $\Omega$ .

In the context of quantum dynamics, the observation problem can be formulated as follows. Consider the **von Neumann equation** with unknown  $t \mapsto R(t) \in \mathcal{D}(\mathfrak{H})$

$$i\hbar \partial_t R = \left[-\frac{\hbar^2}{2m} \Delta + V, R\right], \quad R|_{t=0} = R^{in}.$$

The real-valued potential  $V$  is chosen so that the quantum Hamiltonian

$$\mathbf{H} := -\frac{\hbar^2}{2m} \Delta + V$$

has a self-adjoint (unbounded) extension to  $L^2(\mathbf{R}^d)$ .

Specialists of control usually consider  $R(t) = |\psi(t, \cdot)\rangle\langle\psi(t, \cdot)|$  with

$$i\hbar \partial_t \psi(t, x) = -\frac{1}{2} \hbar^2 \Delta_x \psi(t, x) + V(x) \psi(t, x).$$

A class  $\mathcal{K}$  of solutions  $t \mapsto R(t) \in \mathcal{D}(\mathfrak{H})$  of the von Neumann equation can be **observed** on a domain  $\Omega \subset \mathbf{R}^d$  during time  $T$  if there exists a constant  $C_{OBS} > 0$ , which may depend on  $\mathcal{K}$ , on  $T$  and on  $\Omega$ , but not on the specific solution  $R \in \mathcal{K}$ , such that

$$1 (= \|R(t)\|_1) \leq C_{OBS} \int_0^T \text{trace}_{\mathfrak{H}}(\mathbf{1}_\Omega R(t)) dt \quad \text{for all } R \in \mathcal{K}.$$



The reason for the interest in this notion is explained by the HUM method (Hilbert Uniqueness Method) introduced by J.-L. Lions: if  $\Omega$  and  $T$  are such that all solutions of the Schrödinger equations — meaning that  $\mathcal{K} = C_b(\mathbf{R}; L^2(\mathbf{R}^d))$  — can be observed on  $(0, T) \times \Omega$ , one can control the Schrödinger equation by acting only on the domain  $\Omega$  and on the time interval  $(0, T)$  via a source term in the Schrödinger equation. (As a matter of fact, Lions' initial interest was with the wave equation, or with the equations of linear elasticity [49], and the problem was to eliminate vibrations on large solid structures by acting with motors on small regions of the structure; the question was obviously to find the optimal way of placing the stabilizing motors.)

**Quiz 22: The Lions HUM method.** Let  $H$  be an unbounded self-adjoint operator on  $L^2(\mathbf{R}^d)$  — for instance the Hamiltonian  $H = -\frac{\hbar^2}{2m}\Delta + V$  under some appropriate conditions on the real-valued potential  $V$ . Consider the control problem

$$\begin{cases} i\partial_t \phi = H\phi + \mathbf{1}_{(0,T) \times \Omega} f, \\ \phi|_{t=T} = 0. \end{cases}$$

Can one drive any solution  $\phi \in C_b(\mathbf{R}; L^2(\mathbf{R}^d))$  to 0 at time  $T$  by acting on the domain  $\Omega$  only via the control  $f$ ? In other words, is the control operator

$$\mathcal{C} : L^2((0, T) \times \Omega) \ni f \mapsto -i\phi|_{t=0} \in L^2(\mathbf{R}^d)$$

onto?

To answer this question, one considers instead the observation problem

$$\begin{cases} i\partial_t \psi = H\psi, \\ \psi|_{t=0} = \psi^{in}, \end{cases}$$

and the observation operator

$$\mathcal{O} : L^2(\mathbf{R}^d) \ni \psi^{in} \mapsto \psi|_{(0,T) \times \Omega} \in L^2((0, T) \times \Omega).$$

If the full class  $C_b(\mathbf{R}; L^2(\mathbf{R}^d))$  of solutions of the Schrödinger equation is observable on  $(0, T) \times \Omega$  in the sense of the definition above, the operator  $\mathcal{O}$  is one-to-one.

- (1) Prove that  $\mathcal{C}$  and  $\mathcal{O}$  are bounded operators.
- (2) Prove that  $\mathcal{O}$  is the adjoint of  $\mathcal{C}$ .
- (3) Which statement about the control operator can be deduced from the fact that the class  $C_b(\mathbf{R}; L^2(\mathbf{R}^d))$  of solutions of the Schrödinger equation is observable on  $(0, T) \times \Omega$ ?

The controllability problem for the wave (or elasticity) equation has been studied by several authors under various specific conditions, until Bardos, Lebeau and Rauch [9] came up with a very satisfying answer based on the propagation of high frequency waves. As is well known, high frequency waves propagate according to the laws of geometric optics, and one approach to relating geometric optics to the wave equation is through the consideration of the wave front set in microlocal analysis. (There was a serious additional difficulty in the Bardos-Lebeau-Rauch paper, namely the fact that they sought to control the solution only at the boundary of the domain, which involved using the Melrose-Sjöstrand theory of propagation of the wave front set along “broken bicharacteristics”, in other words, optical rays from the theory of geometric optics interacting with the boundary.) However, the key idea in the Bardos-Lebeau-Rauch approach was a “geometric condition” saying that all the rays of geometric optics should hit the region of control at the boundary

at least once in the time interval  $(0, T)$ . Under this condition, all solutions of the wave equation can be driven to 0 after time  $T$  by some appropriate action on the region of control at the boundary of the domain.

In view of the analogy between geometric optics and the classical limit of quantum mechanics (see for instance §§6 and 46 in [48]), we shall consider the following geometric condition, which can be formulated in terms of the trajectories of the classical mechanics for the same potential as in the quantum problem for which the observation problem is posed.

Thus, to the quantum Hamiltonian  $\mathbf{H}$  above, we associate the classical Hamiltonian

$$H(x, \xi) := \frac{1}{2m}|\xi|^2 + V(x)$$

generating a flow  $\Phi(t; \cdot, \cdot) = (X(t; \cdot, \cdot), \Xi(t; \cdot, \cdot))$  on  $\mathbf{R}^d \times \mathbf{R}^d$  defined by the following prescription:

$$\begin{cases} \dot{X} = \frac{1}{m}\Xi, & X(0; x, \xi) = x, \\ \dot{\Xi} = -\nabla V(X), & \Xi(0; x, \xi) = \xi. \end{cases}$$

**Bardos-Lebeau-Rauch Geometric Condition.** Let  $K \subset \mathbf{R}^{2d}$  be compact, consider a domain  $\Omega \subset \mathbf{R}^d$  and let  $T > 0$ ; the triple  $(K, \Omega, T)$  is said to satisfy the Bardos-Lebeau-Rauch (BLR) geometric condition if

$$(GC) \quad \begin{cases} \text{for each } (x, \xi) \in K, \text{ there exists} \\ t \in (0, T) \text{ such that } X(t; x, \xi) \in \Omega. \end{cases}$$

**Lemma 14.** Assume that  $V \in C^{1,1}(\mathbf{R}^d)$  is real-valued and satisfies  $V \geq -M$  on  $\mathbf{R}^d$  for some  $M > 0$ . Let  $K$  be a compact subset of  $\mathbf{R}^d \times \mathbf{R}^d$ , let  $\Omega$  be an open set in  $\mathbf{R}^d$  and let  $T > 0$ . Assume that the triple  $(K, \Omega, T)$  satisfy (GC). Then

$$C[K, \Omega, T] := \inf_{(x, \xi) \in K} \int_0^T \mathbf{1}_\Omega(X(t; x, \xi)) dt > 0.$$

*Proof.* Since  $\Omega$  is open in  $\mathbf{R}^d$ , its indicator function  $\mathbf{1}_\Omega$  is lower semicontinuous (l.s.c.) on  $\mathbf{R}^d$ . By Fatou's lemma, the function

$$(x, \xi) \mapsto \int_0^T \mathbf{1}_\Omega(X(t; x, \xi)) dt \in (0, +\infty) \text{ is l.s.c. on } \mathbf{R}^d \times \mathbf{R}^d.$$

Since  $K$  is compact in  $\mathbf{R}^d \times \mathbf{R}^d$ , there exists  $(x^*, \xi^*) \in K$  such that

$$C[K, \Omega, T] = \int_0^T \mathbf{1}_\Omega(X(t; x^*, \xi^*)) dt.$$

By the geometric condition, there exists  $t^* \in (0, T)$  such that

$$\mathbf{1}_\Omega(X(t^*; x^*, \xi^*)) = 1.$$

Since  $t \mapsto \mathbf{1}_\Omega(X(t; x^*, \xi^*))$  is l.s.c. on  $(0, T)$ , there exists  $\eta \in (0, T)$  such that

$$|t - t^*| < \eta \implies \mathbf{1}_\Omega(X(t; x^*, \xi^*)) = 1.$$

Hence

$$C[K, \Omega, T] = \int_0^T \mathbf{1}_\Omega(X(t; x^*, \xi^*)) dt \geq \eta > 0.$$

□

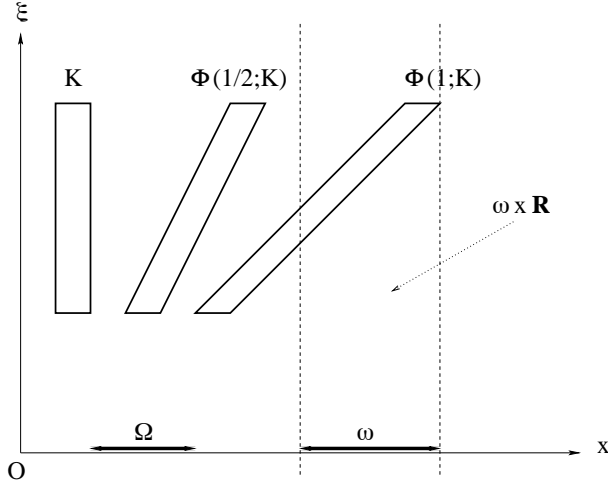


FIGURE 6. The geometric condition in space dimension  $d = 1$ , with  $V \equiv 0$ . The classical free flow is  $\Phi(t; x, \xi) := (X(t; x, \xi), \Xi(t; x, \xi)) = (x + t\xi, \xi)$ . The picture represents the image of the closed phase-space rectangle  $K$  by the map  $(x, \xi) \mapsto \Phi(t; x, \xi)$  at time  $t = \frac{1}{2}$  and  $t = 1$ . The interval  $\Omega$  satisfies the geometric condition with  $\bar{T} = 1$ , at variance with  $\omega$ . Indeed, phase-space points on the bottom side of  $K$  stay out of the strip  $\omega \times \mathbf{R}$  for all  $t \in [0, 1]$ .

With the propagation estimate for  $\mathfrak{d}$  presented in Theorem 6 above, one can formulate a quantitative observation inequality for the von Neumann equation, with rather explicit observation constants defined in terms of the classical dynamics. The following result has been obtained in [41].

**Theorem 15.** Let  $V \in C^{1,1}(\mathbf{R}^d)$  be a real-valued potential satisfying  $V \geq -M$  for some  $M > 0$ . Let  $K$  be a compact subset of  $\mathbf{R}^d \times \mathbf{R}^d$ , let  $\Omega$  be an open set in  $\mathbf{R}^d$  and let  $T > 0$ . Assume that the triple  $(K, \Omega, T)$  satisfy (GC). Then, for all initial density operator  $R^{in} \in \mathcal{D}_2(\mathfrak{H})$  and all  $\delta > 0$ , one has the “observation” inequality on the open set  $\Omega_\delta = \Omega + B(0, \delta)$

$$\int_0^T \text{trace}_{\mathfrak{H}}(\mathbf{1}_{\Omega_\delta} U(t) R^{in} U(t)^*) dt \geq C[K, \Omega, T] \frac{1}{\delta} \inf_{\lambda > 0} \frac{1}{l} \frac{\exp\left(\frac{1}{2} T \left(\frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda}\right)\right) - 1}{\frac{1}{2} \left(\frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda}\right)} \inf_{\text{supp}(f^{in}) \subset K} \mathfrak{d}_\lambda(f^{in}, R^{in}).$$

We have put the word observation between quotes in Theorem 15, since the inequality in that theorem is a bona fide observation inequality only if there exists  $\lambda > 0$  and a probability density  $f^{in}$  with support in  $K$  such that

$$\mathfrak{d}_\lambda(f^{in}, R^{in}) < C[K, \Omega, T] \frac{\frac{\delta \lambda}{2} \left(\frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda}\right)}{\exp\left(\frac{T}{2} \left(\frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda}\right)\right) - 1}.$$

Here are two examples where this inequality is known to be true.

**Example 1: Toeplitz initial data.** Assume that the initial density operator  $R^{in}$  is of the form

$$R^{in} := \mathcal{T}[\mu^{in}], \quad \text{where } \mu^{in} \in \mathcal{P}(\mathbf{R}^{2d}) \text{ satisfies } \text{supp}(\mu^{in}) \subset K.$$

By Theorem 1 (1) (see Lecture I)

$$\lambda d\hbar \leq \mathfrak{d}_\hbar(f^{in}, R^{in})^2 \leq \max(1, \lambda^2) W_2(f^{in}, \mu^{in})^2 + d\hbar,$$

so that

$$\text{supp}(\mu^{in}) \subset K \implies \inf_{\text{supp}(f^{in}) \subset K} \mathfrak{d}_\lambda(f^{in}, R^{in}) = \sqrt{\frac{1}{2}(\lambda^2 + 1)d\hbar}.$$

Indeed,

$$\text{supp}(\mu^{in}) \subset K \implies \inf_{\text{supp}(f^{in}) \subset K} \mathcal{W}_2(f^{in}, \mu^{in}) = 0,$$

since any probability measure  $\mu^{in}$  supported in  $K$  is the weak limit of a sequence of absolutely continuous (with respect to the Lebesgue measure of  $\mathbf{R}^{2d}$ ) probability measures supported in  $K$ . (To see this, let  $\zeta_\epsilon$  be a regularizing sequence on  $\mathbf{R}^{2d}$ , and set

$$f_\epsilon(x, \xi) := Z_\epsilon^{-1}(\mu^{in} \star \zeta_\epsilon)(x, \xi) \mathbf{1}_K(x, \xi), \quad \text{with } Z_\epsilon := \iint_K \mu^{in} \star \zeta_\epsilon(x, \xi) dx d\xi.$$

Then  $f_\epsilon$  is a probability density on  $\mathbf{R}^{2d}$  with support in  $K$ , and one easily check that

$$f_\epsilon \rightarrow \mu^{in} \text{ weakly in the sense of probability measures as } \epsilon \rightarrow 0,$$

and since  $\text{supp}(f_\epsilon) \subset K$  for each  $\epsilon > 0$ , this implies that  $\mathcal{W}_2(f_\epsilon, \mu^{in}) \rightarrow 0$  as  $\epsilon \rightarrow 0^+$ .

**Example 2: Pure state.** Assume now that  $R(t) = |U(t)\psi^{in}\rangle\langle U(t)\psi^{in}|$ , where  $U(t) = e^{-it\mathbf{H}/\hbar}$  is the Schrödinger group. (We recall that

$$\mathbf{H} := -\frac{\hbar^2}{2m}\Delta + V,$$

and the assumptions on  $V$  imply that  $\mathbf{H}$  has a self-adjoint extension to  $L^2(\mathbf{R}^d)$ .)

Choosing

$$f^{in}(q, p) := \frac{|\langle q, p | \psi^{in} \rangle|^2}{(2\pi\hbar)^d} = \mathcal{H}[|\psi^{in}\rangle\langle\psi^{in}|]$$

and setting  $\lambda = 1$  leads to

$$\frac{1}{C_{OBS}} = C[K, \Omega, T] \iint_K |\langle q, p | \psi^{in} \rangle|^2 \frac{dq dp}{(2\pi\hbar)^d} - D[T, \text{Lip}(\nabla V)] \frac{\Sigma[\psi^{in}]}{\delta},$$

where

$$D[T, L] := 4 \frac{e^{(\frac{1}{m}+L)T/2} - 1}{\frac{1}{m} + L},$$

$$\Sigma[\psi^{in}]^2 := \langle \psi^{in} | |x|^2 | \psi^{in} \rangle - |\langle \psi^{in} | x | \psi^{in} \rangle|^2$$

$$+ \langle \psi^{in} | -\hbar^2 \Delta_x | \psi^{in} \rangle - |\langle \psi^{in} | -i\hbar \nabla_x | \psi^{in} \rangle|^2.$$

**Quiz 23.** The missing argument to arrive at the formula for  $C_{OBS}$  in the pure state case is a good opportunity to revise some elementary facts about  $\mathfrak{d}$ .

(1) Find  $\mathcal{C}(\mathcal{H}[|\psi^{in}\rangle\langle\psi^{in}|], |\psi^{in}\rangle\langle\psi^{in}|)$ .

(2) Compute  $\mathfrak{d}(\mathcal{H}[|\psi^{in}\rangle\langle\psi^{in}|], |\psi^{in}\rangle\langle\psi^{in}|)$ . (Hint: observe that  $x_j$ ,  $-i\hbar\partial_{x_j}$ ,  $x_j^2$  and  $-\hbar^2\partial_{x_j}^2$  are Toeplitz operators, and compute their symbols.)

Here is a sketch of the proof of Theorem 15.

*Proof.* Call  $f(t, \cdot, \cdot) := f^{in} \circ \Phi(-t; \cdot, \cdot)$  and  $R(t) := U(t)R^{in}U(t)^*$ . For each coupling  $Q(t) \in \mathcal{C}(f(t, \cdot, \cdot), R(t))$ , one has

$$\begin{aligned} & \left| \text{trace}_{\mathfrak{H}}(\chi R(t)) - \iint_{\mathbf{R}^{2d}} \chi(x) f(t, x, \xi) dx d\xi \right| \\ &= \left| \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}((\chi(x) - \chi(y))Q(t, x, \xi) dx d\xi \right| \\ &\leq \frac{\text{Lip}(\chi)}{\lambda} \left( \iint_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q_t^{\frac{1}{2}}(\lambda^2|x-y|^2 + |\xi + i\hbar\nabla_y|^2)Q_t^{\frac{1}{2}}) dx d\xi \right)^{\frac{1}{2}}, \end{aligned}$$

so that, by the propagation estimate in Theorem 6,

$$\begin{aligned} \left| \text{trace}_{\mathfrak{H}}(\chi R(t)) - \iint_{\mathbf{R}^{2d}} \chi(x) f(t, x, \xi) dx d\xi \right| &\leq \frac{\text{Lip}(\chi)}{\lambda} \mathfrak{d}_{\lambda}(f(t, \cdot, \cdot), R(t)) \\ &\leq \frac{\text{Lip}(\chi)}{\lambda} \mathfrak{d}_{\lambda}(f^{in}, R^{in}) \exp\left(\frac{1}{2}t\left(\frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda}\right)\right). \end{aligned}$$

Since

$$\iint_{\mathbf{R}^{2d}} \chi(x) f(t, x, \xi) dx d\xi = \iint_{\mathbf{R}^{2d}} \chi(X(t; x, \xi)) f^{in}(x, \xi) dx d\xi,$$

one has

$$\begin{aligned} \int_0^T \text{trace}_{\mathfrak{H}}(\chi R(t)) dt &\geq \inf_{(x, \xi) \in K} \int_0^T \chi(X(t; x, \xi)) dt \iint_K f^{in}(x, \xi) dx d\xi \\ &\quad - \frac{\text{Lip}(\chi)}{\lambda} \frac{\exp\left(\frac{1}{2}T\left(\frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda}\right)\right) - 1}{\frac{1}{2}\left(\frac{\lambda}{m} + \frac{\text{Lip}(\nabla V)}{\lambda}\right)} \mathfrak{d}_{\lambda}(f^{in}, R^{in}). \end{aligned}$$

Conclude by choosing  $\chi(x) := \left(1 - \frac{\text{dist}(x, \Omega)}{\delta}\right)_+$ , so that  $\text{Lip}(\chi) = \frac{1}{\delta}$ .  $\square$

The main result in this lecture is obviously Theorem 6, which explains how the pseudometric  $\mathfrak{d}$ , or its variant  $\mathfrak{d}_{\lambda}$  are propagated by self-adjoint quantum Hamiltonians of the form

$$-\frac{\hbar^2}{2m}\Delta + V.$$

However, there are other quantum dynamics for which it may be useful to know how  $\mathfrak{d}$  is propagated. The example in the following exercise was communicated to us by E. Carlen.

**Quiz 24: Quantum Heat Equation.**

(1) Let  $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^n)$ . Prove that

$$\mathcal{W}_2(e^{t\Delta}\mu, e^{t\Delta}\nu) \leq \mathcal{W}_2(\mu, \nu).$$

(Hint: represent  $e^{t\Delta}\mu(x)$  by means of the Brownian motion, and consider the process  $(x + B_t, y + B_t) \in \mathbf{R}^n \times \mathbf{R}^n$ , with the **same** Brownian motion  $B_t$ .)

(2) Find another proof of (1) without appealing to the representation of the solution by means of the Brownian motion. (Hint: pick  $\rho_{in} \in \mathcal{C}(\mu, \nu)$  and propagate  $\rho_{in}$  by a degenerate diffusion operator  $A^*A$ , i.e. set

$$\partial_t \rho_t + A^* A \rho_t = 0, \quad \rho_0 = \rho_{in},$$

where  $A$  is a 1st order differential operator such that  $A|x-y|^2 = 0$ .)

(3) Set  $\mathfrak{H} := L^2(\mathbf{R})$  and  $q\psi(y) := y\psi(y)$  while  $p\psi(y) := -i\hbar \frac{d\psi}{dy}(y)$ . Consider the Quantum Heat Equation

$$\partial_t R = -\frac{1}{\hbar^2}[p, [p, R]] - \frac{1}{\hbar^2}[q, [q, R]], \quad R(0) = R^{in} \in \mathcal{D}_2(\mathfrak{H}).$$

Prove that the Cauchy problem above is solved by a contraction semigroup on  $\mathcal{L}^2(\mathfrak{H})$ , and that  $R(t) \in \mathcal{D}_2(\mathfrak{H})$  for all  $t \geq 0$ .

(4) Let  $R_1, R_2$  be the solutions of

$$\begin{aligned} \partial_t R_1 &= -\frac{1}{\hbar^2}[p, [p, R_1]] - \frac{1}{\hbar^2}[q, [q, R_1]], & R_1(0) &= R_1^{in} \in \mathcal{D}_2(\mathfrak{H}), \\ \partial_t R_2 &= -\frac{1}{\hbar^2}[p, [p, R_2]] - \frac{1}{\hbar^2}[q, [q, R_2]], & R_2(0) &= R_2^{in} \in \mathcal{D}_2(\mathfrak{H}). \end{aligned}$$

Prove that

$$\mathfrak{d}(R_1(t), R_2(t)) \leq \mathfrak{d}(R_1^{in}, R_2^{in}), \quad t \geq 0.$$

(Hint: consider the operators  $[p \otimes I + I \otimes p, q \otimes I - I \otimes q]$  and  $[q \otimes I + I \otimes q, p \otimes I - I \otimes p]$ .)

### 3. LECTURE III: TRIANGLE INEQUALITIES AND OPTIMAL TRANSPORT IN THE QUANTUM SETTING

We have seen in Lecture II that the pseudometric  $\mathfrak{d}$  defined in Lecture I behaves very satisfyingly under propagation by the most fundamental quantum dynamics (i.e. the unitary group generated by self-adjoint Hamiltonians of the form  $-\frac{\hbar^2}{2m}\Delta + V$ , and also by the quantum heat equation). We have not treated the interesting case of the Schrödinger equation with a magnetic field, leading to Hamiltonians of the form

$$\frac{1}{m} | -i\hbar\nabla + A|^2 + V$$

where  $A$  is the vector potential (so that the magnetic field is  $B = \text{curl}A$ ): see [12] for a comprehensive treatment of this case.

Although Lecture II gives (hopefully) convincing arguments, mostly based on applications, in favor of the pseudometric  $\mathfrak{d}$ , we have already seen in Lecture I, that  $\mathfrak{d}$  is not a bona fide metric, in particular because<sup>13</sup>  $\mathfrak{d}(T, T) > 0$  for all  $T \in \mathcal{D}_2(\mathfrak{H})$ .

In Lecture I, we have postponed the necessary task of exploring the properties of the pseudometric  $\mathfrak{d}$ , except those reported in Theorem 1, which have proved already very useful for applications of the pseudometric  $\mathfrak{d}$  to quantum dynamical problems. For instance, we have not discussed the triangle inequality for  $\mathfrak{d}$ , in spite of the fact that we used it in Lecture II in deriving a uniform in  $\hbar$  error estimate for the Lie-Trotter splitting scheme for the von Neumann equation.

In this last lecture we shall return in particular to this question. We shall also discuss further properties of the pseudometric  $\mathfrak{d}$  by analogy with the Wasserstein distance  $\mathcal{W}_2$ . Specifically, we shall study the following topics:

- a Kantorovich-type duality for quantum optimal transport (section 3.3)
- two kinds of triangle inequalities for the quantum pseudometric  $\mathfrak{d}$  on  $\mathfrak{D}$  (presented in section 3.1), and
- the structure of optimal couplings for the pseudometric  $\mathfrak{d}$  (section 3.5).

<sup>13</sup>One might think of changing the definition of  $\mathfrak{d}$ , for instance by replacing the quantum-to-quantum cost operator  $C_\hbar$  in Lecture I by  $C_\hbar - 2d\hbar I_{\mathfrak{H} \otimes \mathfrak{H}}$ . However, if you have understood the essence of the proof of Theorem 6, this is not going to help. Indeed, this proof is based on the operator inequality

$$[H \otimes I_{\mathfrak{H}} + I_{\mathfrak{H}} \otimes H, C_\hbar] \leq LC_\hbar$$

for some positive constant  $L$ , where  $H$  is a Hamiltonian of the form  $-\frac{\hbar^2}{2m}\Delta + V$ . Changing  $C_\hbar$  into  $C_\hbar - \alpha I_{\mathfrak{H} \otimes \mathfrak{H}}$  will obviously not change the commutator on the left-hand side, but changes the right-hand side.

**3.1. Restricted Triangle Inequalities.** In this section, we shall prove the triangle inequality for  $\mathfrak{d}$  in some special cases — hence the terminology of “restricted” triangle inequality used here. Our goal is the following statement.

**Theorem 16.** For all  $\rho_1, \rho_2, \rho_3 \in \mathfrak{D} = \mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d) \cup \mathcal{D}_2(\mathfrak{H})$ , one has

$$\mathfrak{d}(\rho_1, \rho_3) \leq \mathfrak{d}(\rho_1, \rho_2) + \mathfrak{d}(\rho_2, \rho_3),$$

provided that  $\rho_2$  is a probability density in  $\mathbf{R}^d \times \mathbf{R}^d$  or one of the  $\rho_j$ s is a rank-1 density operator on  $\mathfrak{H}$ .

See Theorems A and 3.1 in [40]; see also Theorem 3.5 in [39].

**3.1.1. Operator Inequalities.** As a preparation to the proof of Theorem 16, we prove several inequalities involving the quantum-to-quantum and the classical-to-quantum transport cost operators. These inequalities are of the same type as triangle inequalities, with some arbitrary parameter  $\alpha > 0$ .

First, we discuss inequalities where the intermediate “point” is an operator.

**Lemma 17.** For all  $\alpha > 0$ , one has

$$\begin{aligned} |x - z|^2 + |\xi - \zeta|^2 &\leq (1 + \alpha)c_{\hbar}(x, \xi; y, \hbar\nabla_y) \\ &\quad + (1 + \frac{1}{\alpha})c_{\hbar}(z, \zeta; y, \hbar\nabla_y), \\ c_{\hbar}(x, \xi; z, \hbar\nabla_z) &\leq (1 + \alpha)c_{\hbar}(x, \xi; y, \hbar\nabla_y) \\ &\quad + (1 + \frac{1}{\alpha})C_{\hbar}(y, \hbar\nabla_y, z, \hbar\nabla_z), \\ C_{\hbar}(x, \hbar\nabla_x, z, \hbar\nabla_z) &\leq (1 + \alpha)C_{\hbar}(x, \hbar\nabla_x, y, \hbar\nabla_y) \\ &\quad + (1 + \frac{1}{\alpha})C_{\hbar}(y, \hbar\nabla_y, z, \hbar\nabla_z). \end{aligned}$$

These operator inequalities mean that, for all  $\phi \in \mathcal{S}(\mathbf{R}_{x,\xi}^{2d} \times \mathbf{R}_y^d \times \mathbf{R}_{z,\zeta}^{2d})$ , or all  $\phi \in \mathcal{S}(\mathbf{R}_{x,\xi}^{2d} \times \mathbf{R}_y^d \times \mathbf{R}_z^d)$ , or all  $\phi \in \mathcal{S}(\mathbf{R}_x^d \times \mathbf{R}_y^d \times \mathbf{R}_z^d)$ ,

$$\langle \phi | r.h.s. - l.h.s. | \phi \rangle \geq 0.$$

*Proof.* Write

$$\begin{aligned} C_{\hbar}(x, \hbar\nabla_x, z, \hbar\nabla_z) &= |x - y + y - z|^2 - \hbar^2 |\nabla_x - \nabla_y + \nabla_y - \nabla_z|^2 \\ &= C_{\hbar}(x, \hbar\nabla_x, y, \hbar\nabla_y) + C_{\hbar}(y, \hbar\nabla_y, z, \hbar\nabla_z) \\ &\quad + 2(x - y) \cdot (y - z) - 2\hbar^2 (\nabla_x - \nabla_y) \cdot (\nabla_y - \nabla_z). \end{aligned}$$

Use the Peter-Paul elementary inequality

$$2(x - y) \cdot (y - z) \leq \alpha |x - y|^2 + \frac{1}{\alpha} |y - z|^2,$$

and, for operators  $A, B$ , the analogous inequality

$$A^* B + B^* A \leq \alpha |A|^2 + \frac{1}{\alpha} |B|^2$$

with  $A = A^* = -i\hbar(\partial_{x_j} - \partial_{y_j})$  and  $B = B^* = -i\hbar(\partial_{y_j} - \partial_{z_j})$  for all indices  $j = 1, \dots, d$ . (Observe that these operators commute, which is inessential here). The operator inequality comes from expanding

$$0 \leq \left| \alpha^{\frac{1}{2}} A - \alpha^{-\frac{1}{2}} B \right|^2 = \alpha |A|^2 + \frac{1}{\alpha} |B|^2 - A^* B - B^* A.$$

See also Quiz 7 in Lecture I.

Hence

$$\begin{aligned} &2(x - y) \cdot (y - z) - 2\hbar^2 (\nabla_x - \nabla_y) \cdot (\nabla_y - \nabla_z) \\ &\leq \alpha C_{\hbar}(x, \hbar\nabla_x, y, \hbar\nabla_y) + \frac{1}{\alpha} C_{\hbar}(y, \hbar\nabla_y, z, \hbar\nabla_z) \end{aligned}$$

With the previous equality involving  $C_{\hbar}(x, \hbar\nabla_x, z, \hbar\nabla_z)$ , we arrive at the 3rd inequality of Lemma 17.  $\square$

**Quiz 25.** Prove the first two inequalities in Lemma 17.

Next, we discuss inequalities similar to those of Lemma 17, but with classical phase-space point as intermediate term.

**Lemma 18.** For all  $\alpha > 0$ , one has

$$\begin{aligned} c_{\hbar}(x, \xi; z, \hbar\nabla_z) &\leq (1 + \alpha)(|x - y|^2 + |\xi - \eta|^2) \\ &\quad + (1 + \frac{1}{\alpha})c_{\hbar}(y, \eta; z, \hbar\nabla_z) \\ C_{\hbar}(x, \hbar\nabla_x, z, \hbar\nabla_z) &\leq (1 + \alpha)c_{\hbar}(x, \hbar\nabla_x, y, \eta) \\ &\quad + (1 + \frac{1}{\alpha})c_{\hbar}(y, \eta, z, \hbar\nabla_z). \end{aligned}$$

These operator inequalities mean that

$$\langle \phi | r.h.s. - l.h.s. | \phi \rangle \geq 0$$

for all  $\phi \in \mathcal{S}(\mathbf{R}_{x,\xi}^{2d} \times \mathbf{R}_{y,\eta}^d \times \mathbf{R}_{z,\zeta}^{2d})$ , or all  $\phi \in \mathcal{S}(\mathbf{R}_x^d \times \mathbf{R}_{y,\eta}^d \times \mathbf{R}_z^d)$ .

**Quiz 26.** Prove Lemma 18 (by the same method as in the proof of Lemma 17).

3.1.2. *The Rank-1 Case.* At this point, we can prove the part of Theorem 16 involving a rank-1 density operator.

*Proof of Theorem 16: sketch for the rank-1 case.* Assume for example that  $\rho_1$  and  $\rho_2 \in \mathcal{D}_2(\mathfrak{H})$  while  $\rho_3$  is a rank-1 density operator, and let  $Q \in \mathcal{C}(\rho_1, \rho_2)$ . Set

$$T := Q \otimes \rho_3, \quad T_{13} = \text{trace}_2 T \in \mathcal{C}(\rho_1, \rho_3).$$

Hence, by the 3rd inequality in Lemma 17,

$$\begin{aligned} \mathfrak{d}(\rho_1, \rho_3)^2 &\leq \text{trace}_{\mathfrak{H} \otimes 2}(T_{13}^{\frac{1}{2}} C_{\hbar}(x, \hbar\nabla_x, z, \hbar\nabla_z) T_{13}^{\frac{1}{2}}) \\ &= \text{trace}_{\mathfrak{H} \otimes 3}(T^{\frac{1}{2}} C_{\hbar}(x, \hbar\nabla_x, z, \hbar\nabla_z) T^{\frac{1}{2}}) \\ &\leq (1 + \alpha) \text{trace}_{\mathfrak{H} \otimes 3}(T^{\frac{1}{2}} C_{\hbar}(x, \hbar\nabla_x, y, \hbar\nabla_y) T^{\frac{1}{2}}) \\ &\quad + (1 + \frac{1}{\alpha}) \text{trace}_{\mathfrak{H} \otimes 3}(T^{\frac{1}{2}} C_{\hbar}(y, \hbar\nabla_y, z, \hbar\nabla_z) T^{\frac{1}{2}}) \\ &= (1 + \alpha) \text{trace}_{\mathfrak{H} \otimes 2}(Q^{\frac{1}{2}} C_{\hbar}(x, \hbar\nabla_x, y, \hbar\nabla_y) Q^{\frac{1}{2}}) \\ &\quad + (1 + \frac{1}{\alpha}) \underbrace{\text{trace}_{\mathfrak{H} \otimes 2}((\rho_2 \otimes \rho_3)^{\frac{1}{2}} C_{\hbar}(y, \hbar\nabla_y, z, \hbar\nabla_z) (\rho_2 \otimes \rho_3)^{\frac{1}{2}})}_{=\mathfrak{d}(\rho_2, \rho_3)^2}. \end{aligned}$$

That the last term on the r.h.s. is equal to  $\mathfrak{d}(\rho_2, \rho_3)^2$  follows from the fact that  $\mathcal{C}(\rho_2, \rho_3) = \{\rho_2 \otimes \rho_3\}$  since  $\rho_3$  has rank 1 (see Lemma 2 in Lecture I). Minimizing the last r.h.s. in  $Q \in \mathcal{C}(\rho_1, \rho_2)$  shows that

$$\mathfrak{d}(\rho_1, \rho_3)^2 \leq (1 + \alpha)\mathfrak{d}(\rho_1, \rho_2)^2 + (1 + \frac{1}{\alpha})\mathfrak{d}(\rho_2, \rho_3)^2.$$

Minimizing the r.h.s. in  $\alpha > 0$ , i.e. setting

$$\alpha := \frac{\mathfrak{d}(\rho_2, \rho_3)}{\mathfrak{d}(\rho_1, \rho_2)} \quad \text{assuming } \mathfrak{d}(\rho_1, \rho_2) > 0,$$

leads to

$$\mathfrak{d}(\rho_1, \rho_3)^2 \leq \mathfrak{d}(\rho_1, \rho_2)^2 + \mathfrak{d}(\rho_2, \rho_3)^2 + 2\mathfrak{d}(\rho_1, \rho_2)\mathfrak{d}(\rho_2, \rho_3).$$

Conclude by taking the square root of both sides of this inequality.  $\square$



**Quiz 27.** Complete the missing argument in the proof by justifying the equality

$$\begin{aligned} & \text{trace}_{\mathfrak{H}^{\otimes 2}}(T_{13}^{\frac{1}{2}}C_h(x, \hbar\nabla_x, z, \hbar\nabla_z)T_{13}^{\frac{1}{2}}) \\ &= \text{trace}_{\mathfrak{H}^{\otimes 3}}(T^{\frac{1}{2}}C_h(x, \hbar\nabla_x, z, \hbar\nabla_z)T^{\frac{1}{2}}). \end{aligned}$$

- (1) Prove this identity when  $C_h$  is replaced with  $(I_{\mathfrak{H}^{\otimes 2}} + \frac{1}{n}C_h)^{-1}C_h$ .  
(2) Using the Fatou lemma for trace-class operators, prove that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \text{trace}_{\mathfrak{H}^{\otimes 2}} \left( T_{13}^{\frac{1}{2}} \frac{C_h(x, \hbar\nabla_x, z, \hbar\nabla_z)}{I_{\mathfrak{H}^{\otimes 2}} + \frac{1}{n}C_h(x, \hbar\nabla_x, z, \hbar\nabla_z)} T_{13}^{\frac{1}{2}} \right) \\ &= \text{trace}_{\mathfrak{H}^{\otimes 2}} \left( T_{13}^{\frac{1}{2}} C_h(x, \hbar\nabla_x, z, \hbar\nabla_z) T_{13}^{\frac{1}{2}} \right), \\ & \lim_{n \rightarrow \infty} \text{trace}_{\mathfrak{H}^{\otimes 3}} \left( T^{\frac{1}{2}} \frac{C_h(x, \hbar\nabla_x, z, \hbar\nabla_z)}{I_{\mathfrak{H}^{\otimes 3}} + \frac{1}{n}C_h(x, \hbar\nabla_x, z, \hbar\nabla_z)} T^{\frac{1}{2}} \right) \\ &= \text{trace}_{\mathfrak{H}^{\otimes 3}} \left( T^{\frac{1}{2}} C_h(x, \hbar\nabla_x, z, \hbar\nabla_z) T^{\frac{1}{2}} \right). \end{aligned}$$

**Quiz 28.** Complete the proof of Theorem 16 by treating the missing cases where one of the  $\rho_j$ 's is a rank-1 density operator.

3.1.3. *The Case of a Classical Intermediate Density.* All the cases of the triangle inequality involving a rank-1 operator are easy, because of Lemma 2. Indeed, the set of couplings of any density (quantum or classical) with a rank-1 density operator is a singleton, and therefore the pseudometric  $\mathfrak{d}$  is easily computed explicitly in such a case.

In the present section, we shall discuss all the cases of the triangle inequality where the intermediate point is a classical density on phase-space. The triangle inequality in such cases is much more involved, but fortunately, the proof can be modelled on one of the proofs of the triangle inequality for  $\mathcal{W}_2$  (See the proof of Theorem 7.3 in [63].)

The key step is the following lemma, which explains how to *disintegrate* a coupling between a classical probability density on phase-space and a density operator. **Lemma 19.** Let  $f$  be a probability density on  $\mathbf{R}^d \times \mathbf{R}^d$ , let  $R \in \mathcal{D}(\mathfrak{H})$  and let  $Q \in \mathcal{C}(f, R)$ . There exists a weakly measurable map

$$\mathbf{R}^d \times \mathbf{R}^d \ni (x, \xi) \mapsto Q_f(x, \xi) \in \mathcal{L}^1(\mathfrak{H})$$

defined a.e. on  $\mathbf{R}^d \times \mathbf{R}^d$ , which satisfies

$$Q_f(x, \xi) = Q_f(x, \xi)^* \geq 0, \quad \text{trace}(Q(x, \xi)) = 1,$$

and

$$Q(x, \xi) = f(x, \xi)Q_f(x, \xi) \quad \text{a.e. in } (x, \xi) \in \mathbf{R}^d \times \mathbf{R}^d.$$

*Proof.* First replace  $f$  with a Borel representative, and consider the set

$$\mathcal{N} := f^{-1}(\{0\}),$$

which is Borel measurable. Pick  $u \in \mathfrak{H}$  such that  $\|u\|_{\mathfrak{H}} = 1$ , and set

$$Q_f(x, \xi) := \frac{Q(x, \xi) + \mathbf{1}_{\mathcal{N}}(x, \xi)|u\rangle\langle u|}{f(x, \xi) + \mathbf{1}_{\mathcal{N}}(x, \xi)} \in \mathcal{L}(\mathfrak{H}).$$

Obviously

$$Q(x, \xi) = Q(x, \xi)^* \geq 0 \quad \text{and} \quad f(x, \xi) \geq 0$$

and hence

$$Q_f(x, \xi) = Q_f(x, \xi)^* \geq 0, \quad \text{for a.e. } (x, \xi) \in \mathbf{R}^d \times \mathbf{R}^d.$$

Moreover

$$\text{trace}_{\mathfrak{H}}(Q(x, \xi) + \mathbf{1}_{\mathcal{N}}(x, \xi)|u\rangle\langle u|) = f(x, \xi) + \mathbf{1}_{\mathcal{N}}(x, \xi),$$

so that

$$\text{trace}_{\mathfrak{H}}(Q_f(x, \xi)) = 1, \quad \text{for a.e. } (x, \xi) \in \mathbf{R}^d \times \mathbf{R}^d.$$

Finally

$$f(x, \xi)Q_f(x, \xi) = \frac{f(x, \xi)Q(x, \xi)}{f(x, \xi) + \mathbf{1}_{\mathcal{N}}(x, \xi)} = Q(x, \xi)$$

for a.e.  $(x, \xi) \in \mathbf{R}^d \times \mathbf{R}^d$ . Indeed, whenever  $f(x, \xi) > 0$ , one has  $\mathbf{1}_{\mathcal{N}}(x, \xi) = 0$  and the claimed equality is obvious. On the other hand, since

$$Q(x, \xi) = Q(x, \xi)^* \geq 0 \quad \text{and} \quad \text{trace}_{\mathfrak{H}}(Q(x, \xi)) = f(x, \xi)$$

for a.e.  $(x, \xi) \in \mathbf{R}^d \times \mathbf{R}^d$ , it follows that

$$f(x, \xi) = 0 \implies Q(x, \xi) = 0 = f(x, \xi)Q_f(x, \xi).$$

□

With Lemma 19 at our disposal, the proof of Theorem 16 in the case where the intermediate point is a classical density follows the proof of Theorem 7.3 in [63].

*Proof of Theorem 16: the case of a classical intermediate density.* Consider for example the case where both  $\rho_1$  and  $\rho_3 \in \mathcal{D}_2(\mathfrak{H})$ , and assume that  $\rho_2 = f(y, \eta)dyd\eta \in \mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$ . Choose couplings  $Q^1 \in \mathcal{C}(\rho_1, f)$  while  $Q^3 \in \mathcal{C}(f, \rho_3)$ . Call  $Q_f^3$  the disintegration of  $Q^3$  with respect to  $f$  as in Lemma 19. Set

$$T(y, \eta) := Q^1(y, \eta) \otimes Q_f^3(y, \eta).$$

By construction

$$T(y, \eta) = T(y, \eta)^* \geq 0,$$

and

$$\begin{aligned} \text{trace}_1(T(y, \eta)) &= f(y, \eta)Q_f^3(y, \eta) = Q^3(y, \eta), \\ \text{trace}_3(T(y, \eta)) &= Q^1(y, \eta) \text{trace}_{\mathfrak{H}}(Q_f^3(y, \eta)) = Q^1(y, \eta). \end{aligned}$$

In particular

$$\int_{\mathbf{R}^{2d}} T(y, \eta)dyd\eta =: \mathcal{Q} \in \mathcal{C}(\rho_1, \rho_3).$$

By the second inequality in Lemma 17

$$\begin{aligned} \mathfrak{d}(\rho_1, \rho_3)^2 &\leq \text{trace}_{\mathfrak{H}^{\otimes 2}}(\mathcal{Q}^{\frac{1}{2}}C_{\hbar}(x, \hbar\nabla_x, z, \hbar\nabla_z)\mathcal{Q}^{\frac{1}{2}}) \\ &= \int_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}^{\otimes 2}}(T(y, \eta)^{\frac{1}{2}}C_{\hbar}(x, \hbar\nabla_x, z, \hbar\nabla_z)T(y, \eta)^{\frac{1}{2}})dyd\eta \\ &\leq (1 + \alpha) \int_{\mathbf{R}^{2d}} \text{trace}_1\left(\text{trace}_3(T(y, \eta)^{\frac{1}{2}}c_{\hbar}(x, \hbar\nabla_x, y, \eta)T(y, \eta)^{\frac{1}{2}})\right)dyd\eta \\ &\quad + (1 + \frac{1}{\alpha}) \int_{\mathbf{R}^{2d}} \text{trace}_3\left(\text{trace}_1(T(y, \eta)^{\frac{1}{2}}c_{\hbar}(x, \hbar\nabla_x, y, \eta)T(y, \eta)^{\frac{1}{2}})\right)dyd\eta \\ &\leq (1 + \alpha) \int_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}(Q^1(y, \eta)^{\frac{1}{2}}c_{\hbar}(x, \hbar\nabla_x, y, \eta)Q^1(y, \eta)^{\frac{1}{2}})dyd\eta \\ &\quad + (1 + \frac{1}{\alpha}) \int_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}}\left(Q_f^3(y, \eta)^{\frac{1}{2}}c_{\hbar}(x, \hbar\nabla_x, y, \eta)Q_f^3(y, \eta)^{\frac{1}{2}}\right)f(y, \eta)dyd\eta. \end{aligned}$$

Minimizing the last right-hand side in  $Q^1 \in \mathcal{C}(\rho_1, \rho_2)$  and in  $Q^3 \in \mathcal{C}(\rho_2, \rho_3)$  leads to the inequality

$$\mathfrak{d}(\rho_1, \rho_3)^2 \leq (1 + \alpha)\mathfrak{d}(\rho_1, \rho_2)^2 + (1 + \frac{1}{\alpha})\mathfrak{d}(\rho_2, \rho_3)^2,$$

and we conclude as in the rank-1 case.  $\square$

**Quiz 28.**

(1) Complete the missing details in the proof of Theorem 16 in the case where  $\rho_1, \rho_3 \in \mathcal{D}_2(\mathfrak{H})$  and  $\rho_2 = f(y, \eta)dyd\eta$ . In particular, prove the identity

$$\begin{aligned} & \text{trace}_{\mathfrak{H}^{\otimes 2}}(\mathcal{Q}^{\frac{1}{2}}C_{\hbar}(x, \hbar\nabla_x, z, \hbar\nabla_z)\mathcal{Q}^{\frac{1}{2}}) \\ &= \int_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}^{\otimes 2}}(T(y, \eta)^{\frac{1}{2}}C_{\hbar}(x, \hbar\nabla_x, z, \hbar\nabla_z)T(y, \eta)^{\frac{1}{2}})dyd\eta. \end{aligned}$$

(2) Write the proof of Theorem 16 in the missing cases.

**3.2. Applications of the Restricted Triangle Inequalities.** Before going further in our discussion of the triangle inequality for the pseudometric  $\mathfrak{d}$ , we shall present two easy applications of the (restricted) triangle inequalities already established in the previous section.

3.2.1. *Definition of  $\mathfrak{d}$  on  $\mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d) \times \mathcal{D}_2(\mathfrak{H})$ .* Our first application is of a quite fundamental nature, since it completes our definition of  $\mathfrak{d}$  in Lecture I. So far we have defined  $\mathfrak{d}(\mu, R) = \mathfrak{d}(R, \mu)$  for  $\mu \in \mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$  and  $R \in \mathcal{D}_2(\mathfrak{H})$  only when  $\mu = f(x, \xi)dxd\xi$ , with  $f$  a probability density on  $\mathbf{R}^d \times \mathbf{R}^d$  — i.e. only when  $\mu$  is absolutely continuous with respect to the phase-space Lebesgue measure, by the Radon-Nikodym theorem.

**Theorem 20.** For each  $R \in \mathcal{D}_2(\mathfrak{H})$ , the map  $f \mapsto \mathfrak{d}(f, R)$ , defined for all probability density  $f$  with finite second order moments on  $\mathbf{R}^d \times \mathbf{R}^d$  has a unique extension to  $\mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$  satisfying

$$|\mathfrak{d}(\mu, R) - \mathfrak{d}(\nu, R)| \leq \mathcal{W}_2(f, g), \quad \mu, \nu \in \mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d).$$

See section 3 in [40].

**Remark.** This extension of  $\mathfrak{d}$  obviously satisfies the restricted triangle inequality

$$\mathfrak{d}(\rho_1, \rho_3) \leq \mathfrak{d}(\rho_1, \rho_2) + \mathfrak{d}(\rho_2, \rho_3)$$

for all  $\rho_1, \rho_2, \rho_3 \in \mathfrak{D}$  provided that  $\rho_2 \in \mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$  or if one of the  $\rho_j$ s is a rank-1 density operator on  $\mathfrak{H} = L^2(\mathbf{R}^d)$ .

*Proof.* For all  $f, g$  probability densities with finite 2nd order moments on  $\mathbf{R}^d \times \mathbf{R}^d$ , one has the triangle inequality

$$\mathfrak{d}(f, R) \leq \mathfrak{d}(f, g) + \mathfrak{d}(g, R)$$

by Theorem 16, so that

$$\mathfrak{d}(f, R) - \mathfrak{d}(g, R) \leq \mathfrak{d}(f, g) = \mathcal{W}_2(f, g).$$

Exchanging  $f$  and  $g$  in the inequality above implies that

$$|\mathfrak{d}(f, R) - \mathfrak{d}(g, R)| \leq \mathcal{W}_2(f, g).$$

Thus the function  $f \mapsto \mathfrak{d}(f, R)$  is Lipschitz-continuous for the metric  $\mathcal{W}_2$ . It has therefore a unique Lipschitz-continuous extension to  $\mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$  by the density argument recalled in Lemma 21 below.  $\square$

**Lemma 21.** Let  $\mu \in \mathcal{P}_2(\mathbf{R}^n)$  and let  $\chi_\epsilon(x) = \chi(x/\epsilon)/\epsilon^n$  be an even  $C^\infty$  mollifier with support in  $B_\epsilon(0)$ . Then  $f_\epsilon := \chi_\epsilon \star \mu$  is a  $C^\infty$  probability density on  $\mathbf{R}^n$  and

$$\mathcal{W}_2(f_\epsilon, \mu) \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0.$$

This is Lemma 3.2 in [40].

*Proof.* For all  $\phi \in C_0(\mathbf{R}^n)$ , one has

$$\int_{\mathbf{R}^n} \phi(x) \mu(dx) - \int_{\mathbf{R}^n} f_\epsilon(x) \phi(x) dx = \int_{\mathbf{R}^n} (\phi(x) - \chi_\epsilon \star \phi(x)) \mu(dx)$$

since  $\chi_\epsilon$  is even, so that

$$\left| \int_{\mathbf{R}^n} \phi(x) \mu(dx) - \int_{\mathbf{R}^n} f_\epsilon(x) \phi(x) dx \right| \leq \|\phi - \phi \star \chi_\epsilon\|_{L^\infty(\mathbf{R}^n)} \rightarrow 0.$$

Hence  $f_\epsilon \rightarrow \mu$  weakly in  $\mathcal{P}(\mathbf{R}^n)$  as  $\epsilon \rightarrow 0$ .

It remains to establish the tightness property (see the properties of  $\mathcal{W}_2$  recalled in Lecture I). Since  $\chi_\epsilon$  is even,

$$\int_{\mathbf{R}^n} \mathbf{1}_{|x|>R} |x|^2 \chi_\epsilon \star \mu(x) dx = \int_{\mathbf{R}^n} \chi_\epsilon \star (\mathbf{1}_{|x|>R} |x|^2) \mu(dx).$$

On the other hand, for all  $\epsilon \in (0, 1)$

$$\begin{aligned} \chi_\epsilon \star (\mathbf{1}_{|x|>R} |x|^2) &\leq \mathbf{1}_{|x|+1 \geq R} \int_{\mathbf{R}^n} |x - \epsilon y|^2 \chi(y) dy \\ &\leq 2 \mathbf{1}_{|x|+1 \geq R} \left( |x|^2 + \underbrace{\epsilon^2 \int_{\mathbf{R}^n} |y|^2 \chi(y) dy}_{\leq 1} \right) \end{aligned}$$

Hence

$$\sup_{0 < \epsilon < 1} \int_{\mathbf{R}^n} \mathbf{1}_{|x|>R} |x|^2 \chi_\epsilon \star \mu(x) dx \leq 2 \int_{\mathbf{R}^n} \mathbf{1}_{|x|+1 > R} (|x|^2 + 1) \mu(dx) \rightarrow 0$$

as  $R \rightarrow \infty$ , by dominated convergence.

Therefore, by Theorem 7.12 of [63]

$$\mathcal{W}_2(\chi_\epsilon \star \mu, \mu) \rightarrow 0 \quad \text{as } \epsilon \rightarrow 0.$$

□

**3.2.2.  $\mathcal{W}_2$  is the Classical Limit of  $\mathfrak{d}$ .** Our next application of the restricted triangle inequalities presented in Theorem 16 can be thought of as a confirmation of the geometric picture proposed in Lecture I.

We recall the idea of considering  $\mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$  as a limit set, or boundary set, of  $\mathcal{D}_2(\mathfrak{H})$  in  $\mathfrak{D}$ . The next result completes this picture by showing that  $\mathcal{W}_2$  is the limiting metric deduced from the pseudometric  $\mathfrak{d}$  on  $\mathcal{D}(\mathfrak{H})$ .

**Theorem 22.** Let  $R_\hbar, S_\hbar \in \mathcal{D}_2(\mathfrak{H})$  and  $\mu, \nu \in \mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$ . Assume that  $\mu, \nu$  are the classical limits of  $R_\hbar, S_\hbar$  respectively, i.e.

$$\mathfrak{d}(\mu, R_\hbar) + \mathfrak{d}(\nu, S_\hbar) \rightarrow 0 \quad \text{as } \hbar \rightarrow 0.$$

Then

$$\lim_{\hbar \rightarrow 0} \mathfrak{d}(R_\hbar, S_\hbar) = \mathfrak{d}(\mu, \nu).$$

This is Theorem C (see also Theorem 5.5) in [40].

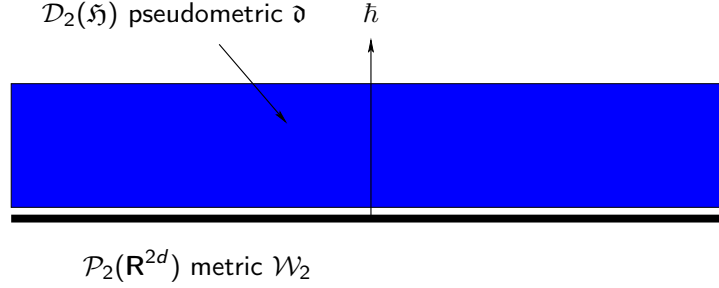


FIGURE 7. The Wasserstein distance  $\mathcal{W}_2$  is the classical limit of the pseudometric  $\mathfrak{d}$  as  $\hbar \rightarrow 0$ .

*Proof.* By the restricted triangle inequality

$$\mathfrak{d}(R_\hbar, S_\hbar) \leq \mathfrak{d}(R_\hbar, \mu) + \mathfrak{d}(\mu, \nu) + \mathfrak{d}(\nu, S_\hbar),$$

so that

$$\overline{\lim}_{\hbar \rightarrow 0} \mathfrak{d}(R_\hbar, S_\hbar) \leq \mathfrak{d}(\mu, \nu) = \mathcal{W}_2(\mu, \nu).$$

On the other hand, by Theorem 1 (2) of Lecture I,

$$\mathfrak{d}(R_\hbar, \mu)^2 \geq \mathcal{W}_2(\mathcal{H}[R_\hbar], \mu)^2 - d\hbar \implies \lim_{\hbar \rightarrow 0} \mathcal{W}_2(\mathcal{H}[R_\hbar], \mu) = 0,$$

$$\mathfrak{d}(S_\hbar, \nu)^2 \geq \mathcal{W}_2(\mathcal{H}[S_\hbar], \nu)^2 - d\hbar \implies \lim_{\hbar \rightarrow 0} \mathcal{W}_2(\mathcal{H}[S_\hbar], \nu) = 0.$$

Hence

$$\mathfrak{d}(R_\hbar, S_\hbar)^2 \geq \mathcal{W}_2(\mathcal{H}[R_\hbar], \mathcal{H}[S_\hbar])^2 - 2d\hbar,$$

implies that

$$\underline{\lim}_{\hbar \rightarrow 0} \mathfrak{d}(R_\hbar, S_\hbar) \geq \lim_{\hbar \rightarrow 0} \mathcal{W}_2(\mathcal{H}[R_\hbar], \mathcal{H}[S_\hbar]) = \mathcal{W}_2(\mu, \nu).$$

(This part of the argument does not appeal to the restricted triangle inequality, and had been already established in Theorem 2.3 (2) of [37], albeit in a slightly different form). Summarizing

$$\mathcal{W}_2(\mu, \nu) \leq \underline{\lim}_{\hbar \rightarrow 0} \mathfrak{d}(R_\hbar, S_\hbar) \leq \overline{\lim}_{\hbar \rightarrow 0} \mathfrak{d}(R_\hbar, S_\hbar) \leq \mathcal{W}_2(\mu, \nu).$$

□

**3.3. Quantum Kantorovich Duality.** In the classical setting, Kantorovich duality has several important consequences. The Knott-Smith and Brenier theorems, giving invaluable information on the structure of optimal couplings between two (Borel) probability measures on  $\mathbf{R}^n$  are among the most important applications of duality — at least in the case of the Wasserstein distance  $\mathcal{W}_2$ .

It is therefore natural to seek extensions of Kantorovich duality to the quantum setting. However, the examples studied at the end of Lecture I (in the section showing that “quantum optimal transport is cheaper”) suggest that the structure of quantum optimal couplings might differ significantly from that of classical optimal couplings.

First we consider the case of  $\mathfrak{d}(f, R)$ , where  $f$  is a probability density on  $\mathbf{R}^d \times \mathbf{R}^d$  with finite 2nd order moments and  $R \in \mathcal{D}_2(\mathfrak{H})$ .

Define the set  $\mathfrak{k}$  of test Kantorovich potentials, as follows:

$$\mathfrak{k} := \{(a, B) : a \in C_b(\mathbf{R}^d \times \mathbf{R}^d) \text{ and } B = B^* \in \mathcal{L}(\mathfrak{H}) \text{ s.t. } a(x, \xi)I_{\mathfrak{H}} + B \leq c_h(x, \xi)\}.$$

The operator inequality in the definition of  $\mathfrak{k}$  means that

$$a(x, \xi)\|\phi\|_{\mathfrak{H}}^2 + \langle \phi | B | \phi \rangle \leq \langle \phi | c_h(x, \xi) | \phi \rangle, \quad x, \xi \in \mathbf{R}^d,$$

for all  $\phi \in H^1(\mathbf{R}^d) \cap L^2(\mathbf{R}^d; |y|^2 dy)$ .

**Theorem 23.** Let  $f$  be a probability density on  $\mathbf{R}^d \times \mathbf{R}^d$  with finite second order moments and let  $R \in \mathcal{D}_2(\mathfrak{H})$ . Then

$$\mathfrak{d}(f, R)^2 = \sup_{(a, B) \in \mathfrak{k}} \left( \int_{\mathbf{R}^{2d}} a(x, \xi) f(x, \xi) dx d\xi + \text{trace}_{\mathfrak{H}}(BR) \right).$$

The proof of Theorem 23 is somewhat involved technically, and we shall not repeat it in these notes. We refer instead the interested reader to Theorem 4.1 and section 4 of [40].

However, it is a good idea to keep in mind the core argument in that proof, which is based on convex duality, exactly as in the classical setting.

Set  $\mathcal{E} := C_b(\mathbf{R}^{2d}; \mathcal{L}(\mathfrak{H}))$  with

$$\|T\|_{\mathcal{E}} := \sup_{x, \xi \in \mathbf{R}^d} \|T(x, \xi)\|_{\mathcal{L}(\mathfrak{H})}.$$

Define

$$G(T) := \begin{cases} 0 & \text{if } T(x, \xi) = T(x, \xi)^* \geq -c_h(x, \xi), \\ +\infty & \text{otherwise,} \end{cases}$$

and

$$H(T) := \begin{cases} \int_{\mathbf{R}^{2d}} \tilde{a} f(x, \xi) dx d\xi + \text{trace}_{\mathfrak{H}}(\tilde{B}R) & \text{if } \begin{cases} T(x, \xi) = T(x, \xi)^* \\ = \tilde{a}(x, \xi)I_{\mathfrak{H}} + \tilde{B}, \end{cases} \\ +\infty & \text{otherwise.} \end{cases}$$

Theorem 23 follows from the Fenchel-Rockafellar duality formula

$$\inf_{T \in \mathcal{E}} (G(T) + H(T)) = \max_{\Lambda \in \mathcal{E}'} (-G^*(-\Lambda) - H^*(\Lambda)).$$

The role of the functional  $G$  in this formulation is obviously to penalize the inequality constraint in the definition of  $\mathfrak{k}$ . (Sometimes, in convex analysis, the functional equal to 0 if a constraint is satisfied, and to  $+\infty$  otherwise is referred to as the “indicator function” of that constraint. This is not to be confused with the classical notion of indicator function of a set, equal to one on the set, and to zero outside.)

Next we consider the case of  $\mathfrak{d}(R, S)$  where  $R, S \in \mathcal{D}_2(\mathfrak{H})$ . Here again, there is an analogue of the classical Kantorovich duality theorem.

Define the set  $\mathfrak{K}$  of test Kantorovich potentials as follows

$$\mathfrak{K} := \{(A, B) : A = A^* \text{ and } B = B^* \in \mathcal{L}(\mathfrak{H}) \text{ s.t. } A \otimes I_{\mathfrak{H}} + I_{\mathfrak{H}} \otimes B \leq C_h\}.$$

The operator inequality means that for all  $\Phi \equiv \Phi(x, y) \in \mathfrak{H} \otimes \mathfrak{H}$  such that

$$(\nabla_x - \nabla_y)\Phi \in L^2(\mathbf{R}^d \times \mathbf{R}^d) \quad \text{and} \quad \Phi \in L^2(\mathbf{R}^d \times \mathbf{R}^d; |x - y|^2 dx dy),$$

it holds

$$\langle \Phi | A \otimes I_{\mathfrak{H}} + I_{\mathfrak{H}} \otimes B | \Phi \rangle \leq \langle \Phi | C_h | \Phi \rangle.$$

**Theorem 24.** For all  $R, S \in \mathcal{D}_2(\mathfrak{H})$ , one has

$$\mathfrak{d}(R, S)^2 = \sup_{(A, B) \in \mathfrak{K}} \text{trace}_{\mathfrak{H}}(AR + BS).$$

Here again, the proof of Theorem 24 is too technically involved to be of immediate interest for these lecture notes. We refer the interested reader to [20] for a detailed exposition of duality in the quantum-to-quantum setting, together with various applications thereof.

Although we have chosen to avoid reproducing the proofs of Theorems 23 and 24, we believe that both results are of key importance in the study of the pseudometric  $\mathfrak{d}$ , and in our approach of quantum optimal transport. As a matter of fact, the end of Lecture III will be focussed on two applications of Theorem 23, which we believe are of some importance.

**3.4. Generalized Triangle Inequalities.** In this section, we return to the question of the triangle inequality for  $\mathfrak{d}$  for  $\rho_1, \rho_2, \rho_3 \in \mathfrak{D}$ , and consider the case where the intermediate point  $\rho_2$  is a density operator and none of the points  $\rho_1, \rho_2, \rho_3$  are rank-1 density operators.

In that case, we do not know of any analogue of Lemma 19, and at the time of this writing we do not know how to “glue”  $Q_{12} \in \mathcal{C}(\rho_1, \rho_2)$  and  $Q_{23} \in \mathcal{C}(\rho_2, \rho_3)$  along  $\rho_2$  as in Lemma 7.6 in [63], to mimic the proof of Theorem 7.3 in [63] in the classical setting.

At the beginning of this lecture, we have proved “restricted” triangle inequalities for  $\mathfrak{d}$ , i.e. the triangle inequality under additional assumptions on  $\rho_1, \rho_2, \rho_3$  — specifically if one of the points  $\rho_1, \rho_2, \rho_3$  is a rank-1 density operator, or if the intermediate point  $\rho_2$  is an element of  $\mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$ .

In the sequel, we shall prove a “generalized triangle inequality”, which holds for all  $\rho_1, \rho_2, \rho_3 \in \mathfrak{D}$ , but includes a correction term of order  $\sqrt{\hbar}$  in the right-hand side.

**Theorem 25.** For all  $\rho_1, \rho_2, \rho_3 \in \mathfrak{D}$ , one has

$$\mathfrak{d}(\rho_1, \rho_3) < \mathfrak{d}(\rho_1, \rho_2) + \mathfrak{d}(\rho_2, \rho_3) + \sqrt{d\hbar}.$$

In particular

$$\mathfrak{d}(\rho_1, \rho_3) < \mathfrak{d}(\rho_1, \rho_2) + \mathfrak{d}(\rho_2, \rho_3) + \frac{1}{\sqrt{2}}\mathfrak{d}(\rho_2, \rho_2).$$

**Remark.** One should compare this result with the De Palma-Trevisan generalized triangle inequality for their distance — which is reminiscent of our generalization of the Wasserstein distance  $\mathcal{W}_2$  to  $\mathcal{D}_2(\mathfrak{H})$  in [37] (the restriction of  $\mathfrak{d}$  to  $\mathcal{D}_2(\mathfrak{H}) \times \mathcal{D}_2(\mathfrak{H})$ ). They arrive at the inequality

$$d_{dPT}(R, T) \leq d_{dPT}(R, S) + d_{dPT}(R, T) + d_{dPT}(S, S)$$

for all  $R, S, T \in \mathcal{D}_2(\mathfrak{H})$  for  $d_{dPT}$  defined in formula (38) of Definition 8, on p. 3208 in [26]: see formula (51) in Theorem 2 of [26] on p. 3210. Notice however that  $d_{dPT}$  differs from the restriction of  $\mathfrak{d}$  to  $\mathcal{D}_2(\mathfrak{H}) \times \mathcal{D}_2(\mathfrak{H})$  — in particular the definition of couplings in [26] uses the notion of quantum channel and is very different from the definition used in these lectures, which comes from [37].

**Quiz 29.** Can one slightly modify the definition of  $\mathfrak{d}$  so as to obtain a functional  $\tilde{\mathfrak{d}}$  defined on  $\mathfrak{D} \times \mathfrak{D}$  that is symmetric and satisfies the (genuine) triangle inequality? Same question for the DePalma-Trevisan functional  $d_{dPT}$ . (Of course, the functional  $\tilde{\mathfrak{d}}$  so obtained satisfies  $\tilde{\mathfrak{d}} > \mathfrak{d} > 0$  and therefore is not a bona fide metric

either, since  $\tilde{\mathfrak{d}}(\rho, \rho) > 0$  for each  $\rho \in \mathfrak{D}$ . The functional  $\tilde{\mathfrak{d}}$  obtained in this exercise is therefore of limited interest.)

The proof of Theorem 25 is very different from the proof of the triangle inequality for the original Wasserstein metric  $\mathcal{W}_2$  — although it uses at some point the restricted triangle inequality in Theorem 16, whose proof is modelled on the proof of Theorem 7.3 in [63].

A key step in the proof of Theorem 25 is the following lemma, which can be seen as a consequence of duality for the classical-to-quantum pseudometric presented in Theorem 23.

**Lemma 26.** For each  $R, S \in \mathcal{D}_2(\mathfrak{H})$ , one has

$$\mathfrak{d}(R, S)^2 \geq \mathfrak{d}(R, \mathcal{H}[S])^2 - d\hbar.$$

Taking this lemma for granted, we give a quick proof of Theorem 25.

**Proof of Theorem 25.** Using  $\mathcal{H}[\rho_2]$  as intermediate point, the restricted triangle inequality implies that

$$\mathfrak{d}(\rho_1, \rho_3) \leq \mathfrak{d}(\rho_1, \mathcal{H}[\rho_2]) + \mathfrak{d}(\mathcal{H}[\rho_2], \rho_3).$$

Then, Lemma 26 implies that

$$\begin{aligned} \mathfrak{d}(\rho_1, \mathcal{H}[\rho_2]) &\leq \sqrt{\mathfrak{d}(\rho_1, \rho_2)^2 + d\hbar} < \mathfrak{d}(\rho_1, \rho_2) + \frac{1}{2}\sqrt{d\hbar}, \\ \mathfrak{d}(\mathcal{H}[\rho_2], \rho_3) &\leq \sqrt{\mathfrak{d}(\rho_2, \rho_3)^2 + d\hbar} < \mathfrak{d}(\rho_2, \rho_3) + \frac{1}{2}\sqrt{d\hbar}. \end{aligned}$$

The second inequalities above result from the following elementary observation

$$X > Y > 0 \implies \sqrt{X^2 + Y^2} \leq X + \frac{1}{2}Y,$$

whose proof is left to the reader as an (easy) exercise.

With the restricted triangle inequality above (Theorem 16), this implies the first generalized triangle inequality in Theorem 25.

To get the second inequality, observe that

$$\rho_2 \in \mathcal{D}_2(\mathfrak{H}) \implies \mathfrak{d}(\rho_2, \rho_2) \geq \sqrt{2d\hbar}.$$

□

**Remark.** In fact, we have proved the slightly more precise inequality

$$\mathfrak{d}(\rho_1, \rho_3) \leq \sqrt{\mathfrak{d}(\rho_1, \rho_2)^2 + d\hbar} + \sqrt{\mathfrak{d}(\rho_2, \rho_3)^2 + d\hbar}.$$

*Proof of Lemma 26.* For all  $a \in C_b(\mathbf{R}^d \times \mathbf{R}^d)$  and all  $B = B^* \in \mathcal{L}(\mathfrak{H})$  satisfying

$$a(x, \xi)I_{\mathfrak{H}} + B \leq c_{\hbar}(x, \xi),$$

one applies the Toeplitz map to both sides of the inequality above in the variables  $x, \xi$ , to find

$$\begin{aligned} \mathcal{T}[a] \otimes I_{\mathfrak{H}} + (2\pi\hbar)^d I_{\mathfrak{H}} \otimes B &\leq (2\pi\hbar)^d \int |q, p\rangle \langle q, p| c_{\hbar}(q, p) dq dp \\ &\leq (2\pi\hbar)^d (C_{\hbar} + d\hbar I_{\mathfrak{H} \otimes \mathfrak{H}}) \end{aligned}$$



(see the basic properties of the Toeplitz map and Quiz 8 in Lecture I). Thus, for all  $T \in \mathcal{C}(R, S)$ , one has

$$\begin{aligned} & (2\pi\hbar)^d \left( \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(T^{\frac{1}{2}} C_{\hbar} T^{\frac{1}{2}}) + d\hbar \right) \\ & \geq \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}} \left( T^{\frac{1}{2}} (\mathcal{T}[a] \otimes I_{\mathfrak{H}} + (2\pi\hbar)^d I_{\mathfrak{H}} \otimes B) T^{\frac{1}{2}} \right) \\ & = \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}} \left( T(\mathcal{T}[a] \otimes I_{\mathfrak{H}} + (2\pi\hbar)^d I_{\mathfrak{H}} \otimes B) \right) \\ & = \text{trace}_{\mathfrak{H}}(RT[a]) + (2\pi\hbar)^d \text{trace}_{\mathfrak{H}}(SB). \end{aligned}$$

Transforming  $\text{trace}_{\mathfrak{H}}(RT[a])$  into an integral involving the functions  $a$  and  $\mathcal{H}[R]$ , i.e. (see Lecture I, formula (4) on Husimi transforms)

$$\text{trace}_{\mathfrak{H}}(RT[a]) = (2\pi\hbar)^d \int_{\mathbf{R}^{2d}} \mathcal{H}[R](q, p) a(q, p) dq dp,$$

we arrive at the formula

$$\begin{aligned} & (2\pi\hbar)^d \left( \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(T^{\frac{1}{2}} C_{\hbar} T^{\frac{1}{2}}) + d\hbar \right) \\ & \geq (2\pi\hbar)^d \left( \int_{\mathbf{R}^{2d}} \mathcal{H}[R](q, p) a(q, p) dq dp + \text{trace}_{\mathfrak{H}}(SB) \right). \end{aligned}$$

Maximizing the right-hand side above in  $a \in C_b(\mathbf{R}^d \times \mathbf{R}^d)$  and  $B = B^* \in \mathcal{L}(\mathfrak{H})$  s.t.

$$a(x, \xi) I_{\mathfrak{H}} + B \leq c_{\hbar}(x, \xi),$$

and applying the duality formula in Theorem 16 shows that

$$(2\pi\hbar)^d \left( \text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(T^{\frac{1}{2}} C_{\hbar} T^{\frac{1}{2}}) + d\hbar \right) \geq (2\pi\hbar)^d \mathfrak{d}(\mathcal{H}[R], S)^2$$

i.e.

$$\text{trace}_{\mathfrak{H} \otimes \mathfrak{H}}(T^{\frac{1}{2}} C_{\hbar} T^{\frac{1}{2}}) \geq \mathfrak{d}(\mathcal{H}[R], S)^2 - d.$$

Minimizing the left-hand side of the inequality above in  $T \in \mathcal{C}(R, S)$  leads to the desired inequality.  $\square$

**Quiz 30.** Use Lemma 26 to recover the following result (already known as statement (2) in Theorem 1 of Lecture I)

$$\mathfrak{d}(R, S)^2 \geq \mathfrak{d}(\mathcal{H}[R], \mathcal{H}[S])^2 - 2d\hbar.$$

**Remarks.**

(1) If you include the proof of the duality formula in Theorem 16, this is the longest and most difficult proof of the inequality above... On the other hand, Lemma 26 is a (much) stronger statement — it is the key to the generalized triangle inequality. That its proof is more involved than the proof of the inequality (2) in Theorem 1 is only natural.

(2) Summarizing, in order to prove the triangle inequality for  $\mathfrak{d}$  when the intermediate point is not a classical density and none of the density operators involved are rank-1 projections, you

(i) first use the exact (restricted) triangle inequality from Theorem 16

$$\mathfrak{d}(\rho_1, \rho_3) \leq \mathfrak{d}(\rho_1, \mathcal{H}[\rho_2]) + \mathfrak{d}(\mathcal{H}[\rho_2], \rho_3),$$

(ii) and then pay the price for replacing  $\rho_2$  with its Husimi function

$$\mathfrak{d}(\rho_1, \rho_3) \leq \sqrt{\mathfrak{d}(\rho_1, \rho_2)^2 + d\hbar} + \sqrt{\mathfrak{d}(\rho_2, \rho_3)^2 + d\hbar}$$

which is the result in Lemma 26, based on the Kantorovich duality for the the classical-to-quantum distance (Theorem 23). The end of the proof is Kindergarten analysis.

The reason for the detour through  $\mathcal{H}[\rho_2]$  instead of  $\rho_2$  is due to the fact that question (5) in the following quiz is answered in the negative<sup>14</sup>.

Before working on this exercise, it is a good idea to review the proofs of Theorem 7.3 (the triangle inequality for  $\mathcal{W}_p$ ) and of Lemma 7.6 (i.e. the disintegration and the glueing of couplings) in [63].

**Quiz 31.** Pick  $\rho_1, \rho_2, \rho_3 \in \mathcal{D}_2(\mathfrak{H})$ , all of them or rank  $\geq 2$  — otherwise, there is nothing to prove. Pick  $R_{12}$  and  $R_{23}$  to be optimal couplings of  $\rho_1, \rho_2$  and  $\rho_2, \rho_3$  (recall briefly why such couplings exist).

(1) Assume there exists  $T \in \mathcal{D}(\mathfrak{H} \otimes \mathfrak{H} \otimes \mathfrak{H})$  such that

$$\text{trace}_1(T) = R_{23} \quad \text{and} \quad \text{trace}_3(T) = R_{12}.$$

Prove that

$$\mathfrak{d}(\rho_1, \rho_3) \leq \mathfrak{d}(\rho_1, \rho_2) + \mathfrak{d}(\rho_2, \rho_3).$$

(Hint: observe that  $\text{trace}_2(T) \in \mathcal{C}(\rho_1, \rho_3)$ .)

Therefore, proving the triangle inequality boils down to proving the existence of such a  $T$ . The classical analogue of this is precisely the content of Lemma 7.6 in Villani's book [63].

Let us consider this problem in finite dimension:  $\mathfrak{H} = \mathbf{C}^2$  — notice that 2 is the first interesting dimension, because if one of the densities  $\rho_j$  for  $j = 1, 2, 3$  has rank 1, the triangle inequality is already known (see Theorem 16 on the restricted triangle inequality).

(2) Let  $R, R' \in M_2(\mathbf{C})$ . Find a necessary and sufficient condition on  $R, R'$  such that there exists  $A, B, C \in M_2(\mathbf{C})$  for which the block-wise matrix

$$T := \begin{pmatrix} A & B \\ B^* & C \end{pmatrix}, \quad A = A^*, \quad C = C^*,$$

satisfies

$$\tau'(T) := A + C = R \quad \text{and} \quad \tau(T) := \begin{pmatrix} \text{trace}(A) & \text{trace}(B) \\ \text{trace}(B^*) & \text{trace}(C) \end{pmatrix} = R'.$$

(3) Assume now that  $R, R' \in M_2(M_2(\mathbf{C}))$ . Find a necessary and sufficient condition on  $R, R'$  such that there exists  $A, B, C \in M_2(M_2(\mathbf{C}))$  for which the block-wise matrix

$$T := \begin{pmatrix} A & B \\ B^* & C \end{pmatrix}, \quad A = A^*, \quad C = C^*$$

satisfies

$$\tau'(T) = A + C = R \quad \text{and} \quad \begin{pmatrix} \text{trace}_{M_2(\mathbf{C})}(A) & \text{trace}_{M_2(\mathbf{C})}(B) \\ \text{trace}_{M_2(\mathbf{C})}(B^*) & \text{trace}_{M_2(\mathbf{C})}(C) \end{pmatrix} = R'.$$

---

<sup>14</sup>While preparing the final version of these lecture notes, I showed the problem to Prof. Denis Serre, who found a counterexample. At first sight, Serre's counterexample does not seem to suggest that the triangle inequality itself (without the extra  $\sqrt{d\hbar}$  term on the right-hand side) should be wrong. However, it shows that the procedure of "glueing" classical couplings described in Lemma 7.6 of [63], which is key to proving the triangle inequality for  $\mathcal{W}_p$  with  $1 < p < \infty$ , does not have an analogue for general quantum couplings. This is a rather fundamental difference between classical and quantum optimal transport.

The notation needs being explained. An element of  $B \in M_2(M_2(\mathbf{C}))$  is of the form

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \quad \text{with } B_{kl} \in M_2(\mathbf{C}).$$

Then

$$B^* := \begin{pmatrix} \overline{B_{11}^T} & \overline{B_{21}^T} \\ \overline{B_{12}^T} & \overline{B_{22}^T} \end{pmatrix}$$

while

$$\text{trace}_{M_2(\mathbf{C})}(B) := B_{11} + B_{22}.$$

(4) Explain how (3) is related to the problem of finding  $T$  as in (1), in the case where  $\rho_1, \rho_2, \rho_3 \in \mathcal{D}(\mathbf{C}^2)$ .

Now, there's the rub...

(5) Assuming that  $R, R' \in \mathcal{D}(\mathbf{C}^2)$ , does (the) block-wise matrix (matrices)  $T$  obtained in (3) satisfy  $T = T^* \geq 0$ ?

**Remark.** A final observation on the generalized triangle inequality is in order. The presence of the additional term  $\sqrt{d\hbar}$  in the right-hand side of the generalized triangle inequality may be related to the fact that some points in  $\mathfrak{D}$  have positive thickness, in the sense that  $\mathfrak{d}(R, R) \geq \sqrt{2d\hbar}$  for all  $R \in \mathcal{D}_2(\mathfrak{H})$ . However, the idea of relating this additional term to the “thickness” of the intermediate point  $\rho_2$  in the triangle inequality could be misleading. In the first place, the genuine triangle inequality holds if  $\rho_2$  is a rank-1 density operator, although  $\mathfrak{d}(\rho_2, \rho_2) \geq \sqrt{2d\hbar}$  in that case.

Besides, the extra term  $\sqrt{d\hbar}$  on the right-hand side of the generalized triangle inequality seems more related to the method of proof of Theorem 25 (viz. the idea of replacing the intermediate point  $\rho_2$  with its Husimi transform  $\mathcal{H}[\rho_2]$ ) than to some intrinsic feature of  $\mathfrak{d}$ .

Ultimately, it could be that question (5) in Quiz 31 is answered in the negative, and yet the genuine triangle inequality holds on  $\mathfrak{D}$  because there always exists an optimal coupling with a special structure for each pair of finite energy density operators on  $\mathfrak{H}$ , and this special structure acts in favor of the genuine triangle inequality.

This suggests investigating the structure of optimal couplings for the “pseudo-metric”  $\mathfrak{d}$  — however, the example discussed in Proposition 5 suggests that this is not an easy task.

**3.5. Towards Quantum Optimal Transport.** Until now, we have not said much about quantum optimal transport *per se*, although it is the topic of this school.

We shall conclude these lectures with some (partial) remarks in that direction. The material in this section mostly comes from [40].

We recall that the proofs of the Brenier, or of the Knott-Smith theorems use some form of Kantorovich duality to obtain some additional information on the structure of optimal couplings. That an optimal coupling for  $\mathcal{W}_2$  is supported in the graph of the subdifferential of some appropriate function comes from the specific structure of the inequality constraint in the Kantorovich dual formulation of the  $\mathcal{W}_2$  metric, and from the specific structure of the quadratic transport cost.

It is therefore natural to study the quantum analogue of Kantorovich duality in order to obtain some information on the structure of optimal couplings for  $\mathfrak{d}$ . In the

sequel, we shall follow this approach with the classical-to-quantum duality formula in Theorem 23.

Our first task is to give a systematic procedure for constructing elements of the class  $\mathfrak{k}$  of test Kantorovich potentials defined before the statement of Theorem 23.

Set  $z := (x, \xi) \in \mathbf{R}^d \times \mathbf{R}^d$  and  $Z := (y, -i\hbar\nabla_y)$ , with the notation

$$z \cdot Z := x \cdot y - i\hbar\xi \cdot \nabla_y.$$

Thus  $c_\hbar(x, \xi) = |Z|^2 + |z|^2 I_{\mathfrak{H}} - 2z \cdot Z \geq d\hbar I_{\mathfrak{H}}$  and by Weyl's theorem (see Corollary 2 of Theorem XIII.14 in [55])

$$\tilde{B} \in \mathcal{L}(\mathfrak{H}) \implies c_\hbar(z)^{-1} \tilde{B} \in \mathcal{K}(\mathfrak{H}) \implies \text{ess-spec}(c_\hbar(z) - \tilde{B}) = \emptyset.$$

Assume that  $\tilde{B} = \tilde{B}^*$  is such that  $c_\hbar(z) - \tilde{B}$  has nondegenerate ground state (i.e. with geometric multiplicity 1) for each  $z \in \mathbf{R}^{2d}$ . (For instance, choose for  $\tilde{B}$  to be a bounded multiplication operator, and apply Theorem XIII.47 in [55]).

Define next

$$\begin{aligned} \tilde{a}(z) &:= \min \text{spec}(c_\hbar(z) - \tilde{B}) = \inf_{\|\phi\|_{\mathfrak{H}}=1} \langle \phi | c_\hbar(z) - \tilde{B} | \phi \rangle \\ &\implies c_\hbar(z) - \tilde{B} \geq \tilde{a}(z) I_{\mathfrak{H}}. \end{aligned}$$

Besides,  $z \mapsto \tilde{a}(z)$  is continuous (even real-analytic) by the Kato-Rellich theorem (Theorem XII.8 of [55]), and

$$\begin{aligned} \tilde{a}(z) &\leq \langle z | c_\hbar(z) - \tilde{B} | z \rangle = d\hbar - \langle z | \tilde{B} | z \rangle \leq d\hbar + \|\tilde{B}\|, \\ \tilde{a}(z) &\geq d\hbar + \inf_{\|\phi\|_{\mathfrak{H}}=1} \langle \phi | -\tilde{B} | \phi \rangle \geq d\hbar - \|\tilde{B}\|. \end{aligned}$$

Hence  $\tilde{a} \in C_b(\mathbf{R}^d \times \mathbf{R}^d)$ , and we have obtained in this way

$$(\tilde{a}, \tilde{B}) \in \mathfrak{k}.$$

The Kato-Rellich theorem recalled above also implies the existence of a continuous (even real-analytic) map

$$\mathbf{R}^d \times \mathbf{R}^d \ni z \mapsto \psi_z \in \mathfrak{H} \quad \text{s.t.} \quad \begin{cases} (c_\hbar(z) - \tilde{B})\psi_z = \tilde{a}(z)\psi_z, \\ \text{and } \|\psi_z\|_{\mathfrak{H}} = 1, \quad z \in \mathbf{R}^{2d}. \end{cases}$$

With this, we can define a notion of quantum optimal transport from  $\mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$  to  $\mathcal{D}_2(\mathfrak{H})$ .

**Theorem 27.** Let  $\tilde{B} = \tilde{B}^*$  be such that  $c_\hbar(z) - \tilde{B}$  has nondegenerate ground state, set  $\tilde{a}(z) := \min \text{spec}(c_\hbar(z) - \tilde{B})$ , and let  $z \mapsto \psi_z$  be a continuous map from  $\mathbf{R}^d \times \mathbf{R}^d$  to  $\mathfrak{H}$  such that  $\|\psi_z\|_{\mathfrak{H}} = 1$  and  $\psi_z \in \text{Ker}(c_\hbar(z) - \tilde{B} - \tilde{a}(z)I_{\mathfrak{H}})$ .

Then, for each probability density  $f$  with finite 2nd order moments, the map  $z \mapsto f(z)|\psi_z\rangle\langle\psi_z|$  is an optimal coupling for the pseudometric  $\mathfrak{d}$  between  $f$  and the operator

$$\mathcal{T}^{\tilde{B}}[f] := \int_{\mathbf{R}^{2d}} f(z) |\psi_z\rangle\langle\psi_z| dz \in \mathcal{D}_2(\mathfrak{H}).$$

**Example.** Take for example  $\tilde{B} = 0$ ; then, one easily checks that

$$\tilde{a}(z) = d\hbar, \quad \text{Ker}(c_\hbar(z) - d\hbar I_{\mathfrak{H}}) = \mathbf{C}|z\rangle,$$

where  $|z\rangle$  is the Schrödinger coherent state centered at  $z := q + ip$ , so that

$$\mathcal{T}^0[f] = \mathcal{T}[f]$$

is the Toeplitz operator of symbol  $f$ .

We already knew from Theorem 1 (1) in Lecture I that

$$\mathfrak{d}(f, \mathcal{T}[f]) = \sqrt{d\hbar} = \inf_{\mathcal{P}_2(\mathbf{R}^{2d}) \times \mathcal{D}_2(\mathfrak{H})} \mathfrak{d}.$$

*Proof of Theorem 27.* Set  $Q(z) := f(z)|\psi_z\rangle\langle\psi_z|$ , so that  $Q(z)^{\frac{1}{2}} = \sqrt{f(z)}|\psi_z\rangle\langle\psi_z|$ , and

$$\begin{aligned} \int_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}} \left( Q(z)^{\frac{1}{2}} c_{\hbar}(z) Q(z)^{\frac{1}{2}} \right) dz &= \int_{\mathbf{R}^{2d}} \langle \psi_z | c_{\hbar}(z) | \psi_z \rangle f(z) dz \\ &= \int_{\mathbf{R}^{2d}} (\tilde{a}(z) \langle \psi_z | \psi_z \rangle + \langle \psi_z | B | \psi_z \rangle) f(z) dz \\ &= \int_{\mathbf{R}^{2d}} \tilde{a}(z) f(z) dz + \text{trace}_{\mathfrak{H}} \left( \tilde{B} \mathcal{T}^{\tilde{B}}[f] \right). \end{aligned}$$

Since  $(\tilde{a}, \tilde{B}) \in \mathfrak{k}$  and  $Q \in \mathcal{C}(f, \mathcal{T}^{\tilde{B}}[f])$ , this implies that

$$\begin{aligned} &\int_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}} \left( Q(z)^{\frac{1}{2}} c_{\hbar}(z) Q(z)^{\frac{1}{2}} \right) dz \\ &= \min_{T \in \mathcal{C}(f, \mathcal{T}^{\tilde{B}}[f])} \int_{\mathbf{R}^{2d}} \text{trace}_{\mathfrak{H}} \left( T(z)^{\frac{1}{2}} c_{\hbar}(z) T(z)^{\frac{1}{2}} \right) dz = \mathfrak{d}(f, \mathcal{T}^{\tilde{B}}[f])^2. \end{aligned}$$

This also implies that

$$\begin{aligned} &\int_{\mathbf{R}^{2d}} \tilde{a}(z) f(z) dz + \text{trace}_{\mathfrak{H}} \left( \tilde{B} \mathcal{T}^{\tilde{B}}[f] \right) \\ &= \sup_{(a, B) \in \mathfrak{k}} \int_{\mathbf{R}^{2d}} a(z) f(z) dz + \text{trace}_{\mathfrak{H}} \left( B \mathcal{T}^{\tilde{B}}[f] \right). \end{aligned}$$

Therefore, in this case, the sup is attained in  $\mathfrak{k}$  (this is not true in general).  $\square$

### Remarks.

(1) Thus the optimal transport map for  $\mathfrak{d}$  between  $\mathcal{P}_2(\mathbf{R}^d \times \mathbf{R}^d)$  and  $\mathcal{D}_2(\mathfrak{H})$  can be thought of as a *deformation* of the Toeplitz quantization, at least when  $\tilde{B}$  is such that  $c_{\hbar}(z) - \tilde{B}$  has a ground state of geometric multiplicity 1.

(2) Notice that the starting point in Theorem 27 is the pair  $(\tilde{a}, \tilde{B}) \in \mathfrak{k}$ , and not the pair consisting of  $f$  (the probability density) and the density operator which are the arguments of  $\mathfrak{d}$ . This approach is vaguely reminiscent of the notion of geodesic in Riemannian geometry: the original definition of a geodesic curve on a Riemannian manifold is that of the shortest path between two points on the manifold. The calculus of variations shows that geodesics define local solutions of an ODE system set on the tangent, or cotangent bundle of the manifold. Conversely, solutions of this ODE system define *local* geodesic curves, which may differ from prescribing arbitrary end points and finding a shortest path between these points. There is a very loose analogy between the local theory of geodesics through the ODE system, and the definition of optimal couplings starting from elements of  $\mathfrak{k}$  as explained in Theorem 27. In other words, the transport map  $\mathcal{T}^{\tilde{B}}$  is independent of the choice of the endpoints — or more precisely the endpoints  $f$  and  $\mathcal{T}^{\tilde{B}}[f]$  follow from the optimal transport map, instead of the other way around.

This raises the following question: in Brenier's theorem, the (classical) optimal transport map is the gradient of a convex function. Is there some analogous property in the quantum setting?

*Operator Legendre Duality and Quantum Optimal Transport.* If  $(\tilde{a}, \tilde{B}) \in \mathfrak{k}$ , one has

$$\underbrace{|Z|^2 + |z|^2 I_{\mathfrak{H}} - 2z \cdot Z}_{=c_h(z)} \geq \tilde{a}(z) I_{\mathfrak{H}} + \tilde{B} \iff a(z) + B \geq z \cdot Z,$$

with

$$a(z) := \frac{1}{2}(|z|^2 - \tilde{a}(z)), \quad B = \frac{1}{2}(|Z|^2 - \tilde{B}).$$

Besides, one has

$$\text{Dom}(c_h(z)) = \text{Dom}(|Z|^2) = H^2(\mathbf{R}^d) \cap L^2(\mathbf{R}^d, |y|^4 dy) =: D.$$

After these preliminaries, we define a notion of Legendre transform of an (unbounded) operator on  $\mathfrak{H}$ .

**Definition.** Let  $B$  satisfy  $|Z|^2 - 2B \in \mathcal{L}(\mathfrak{H})$ . The Legendre dual of  $B$  is the convex function (upper envelope of affine functions)

$$B^L(z) := \sup_{\phi \in D, \|\phi\|_{\mathfrak{H}}=1} (z \cdot \langle \phi | Z | \phi \rangle - \langle \phi | B | \phi \rangle).$$

Indeed, we recall that, if  $T$  is an operator on  $\mathfrak{H}$ , one can think of  $\psi \mapsto \langle \psi | T | \psi \rangle$  as the noncommutative analogue of the evaluation at  $x \in \mathbf{R}^n$  of a real-valued function  $f$  defined on  $\mathbf{R}^n$ . Therefore, the definition above is analogous to the usual definition of the Legendre(-Fenchel) transform

$$\phi^*(\xi) = \sup_{x \in E} (\langle \xi, x \rangle_{E', E} - \phi(x)), \quad \xi \in E',$$

for all  $\phi : E \rightarrow (-\infty, +\infty]$  where  $E$  is a normed linear space on  $\mathbf{R}$ , and  $\phi$  is not identically equal to  $+\infty$ , and  $E'$  is the topological dual of  $E$ , i.e. the space of linear functionals on  $E$  that are continuous for the norm topology.

**Theorem 28.** Let  $\tilde{B} = \tilde{B}^*$  be such that  $c_h(z) - \tilde{B}$  has nondegenerate ground state, set  $\tilde{a}(z) := \min \text{spec}(c_h(z) - \tilde{B})$ , and let  $z \mapsto \psi_z$  be a continuous map from  $\mathbf{R}^d \times \mathbf{R}^d$  to  $\mathfrak{H}$  such that  $\|\psi_z\|_{\mathfrak{H}} = 1$  and  $\psi_z \in \text{Ker}(c_h(z) - \tilde{B} - \tilde{a}(z) I_{\mathfrak{H}})$ .

(1) Setting

$$a(z) := \frac{1}{2}(|z|^2 - \tilde{a}(z)), \quad B := \frac{1}{2}(|Z|^2 - \tilde{B}),$$

one has

$$a = B^L.$$

(2) Besides

$$\nabla a(z) = z - \nabla \tilde{a}(z) = \langle \psi_z | Z | \psi_z \rangle.$$

*Proof.* Statement (1) follows from the definition and the variational formula for the ground state. As for (2), differentiate in  $z$  the identity

$$B\psi_z - z \cdot Z\psi_z + a(z)\psi_z = 0,$$

and take the inner product with  $\psi_z$  to get

$$\underbrace{\langle \psi_z | B - z \cdot Z + a(z) | \psi_z \rangle}_{=0 \text{ since } B=B^*, Z=Z^*, a(z) \in \mathbf{R}} + \langle \psi_z | -Z + \nabla a(z) | \psi_z \rangle = 0.$$

□

**Remarks.**

(1) In the Knott-Smith theorem recalled in Lecture I, optimal couplings for  $\mathcal{W}_2$  are supported in the graph of the subdifferential of a l.s.c. convex function, while, in the Brenier theorem, the optimal transport map is the gradient of a convex function. In both results, the function is obtained from an optimal Kantorovich potential by

the same transformation as  $\tilde{a} \mapsto a$ . Theorem 28 (2) is a partial analogue of this crucial piece of information, except that, in the quantum setting, density operators are not “functions of  $Z$ ”.

(2) In classical optimal transport, there exist an optimal pair  $(a, b)$  of Kantorovich potentials; they are l.s.c. proper convex functions and are Legendre duals of each other, so that  $\nabla a \circ \nabla b = \text{Id}$ ; besides  $a \in L^1_\mu$  and  $b \in L^1_\nu$ . If one tries to proceed by analogy, in the present case, one should define some notion of “quantum gradient” of the operator  $B$ .

One idea to do so is to use the phase space symplectic structure. For a smooth function  $\alpha \equiv \alpha(x, \xi)$  on  $\mathbf{R}^d \times \mathbf{R}^d$ , one has

$$\partial_{x_j} \alpha = \{\xi_j, \alpha\}, \quad \partial_{\xi_j} \alpha = -\{x_j, \alpha\}, \quad j = 1, \dots, d.$$

This suggests to define “quantum derivatives” as follows,

$$\partial_{y_j}^Q B := \frac{i}{\hbar} [-i\hbar \partial_{y_j}, B], \quad \partial_{\eta_j}^Q B := -\frac{i}{\hbar} [y_j, B], \quad j = 1, \dots, d,$$

by using the correspondence principle and the analogy between commutator and Poisson bracket

$$\frac{i}{\hbar} [\cdot, \cdot] \rightarrow \{\cdot, \cdot\}$$

recalled in Lecture II. Since

$$B\psi_z = z \cdot Z\psi_z + a(z)\psi_z, \quad B = B^*, \quad Z = Z^* \text{ and } a(z) \in \mathbf{R},$$

one easily checks that

$$\begin{cases} x_j = \langle \psi_z | \partial_{y_j}^Q B | \psi_z \rangle \\ \xi_j = \langle \psi_z | \partial_{\eta_j}^Q B | \psi_z \rangle \end{cases}$$

This formula can be viewed as the inverse transform of Theorem 28 (2).

(3) Analogous ideas on a definition of an optimal transport “map” between elements of  $\mathcal{D}_2(\mathfrak{H})$  can be found in [20]. Partial results analogous to Theorem 28 have been obtained there, but much remains to be done.

Following the proof of Theorems 27–28 suggests viewing the operator

$$-\frac{1}{2}(|x|^2 - \hbar^2 \Delta_x - A)$$

as the “smallest eigenvalue” of the operator

$$\frac{1}{2}(|y|^2 - \hbar^2 \Delta_y - B) - x \cdot y + \hbar^2 \nabla_x \cdot \nabla_y,$$

viewed as a “matrix” whose entries are operators in the  $x$ -variables. However, inequalities between operators do not define a total order relation, so that even the notion of ground state in this setting does not seem to make much sense.

New ideas on this problem are obviously needed.

## REFERENCES

- [1] L.V. Ahlfors: “Complex Analysis. An Introduction to the Theory of Analytic Functions of One Complex Variable”. 3rd edition. McGraw Hill, Inc., 1979.
- [2] G. Alberti, L. Ambrosio, P. Cannarsa: On the Singularities of Convex Functions. *Manuscripta Math.* **76** (1992), 421–435.
- [3] L. Ambrosio: Lecture Notes on Optimal Transport Problems. In “Mathematical aspects of evolving interfaces”, P. Colli et J.F. Rodrigues eds. *Lect. Notes Math.* 1812, 1–52. Springer, Berlin, 2003.
- [4] L. Ambrosio, N. Gigli, G. Savaré: “Gradient Flows in Metric Spaces and in the Space of Probability Measures”, 2nd ed. Birkhäuser Verlag AG, Basel, Boston, Berlin, 2008.

- [5] A.G. Athanassoulis: Exact equations for smoothed Wigner transforms and homogenization of wave propagation. *Appl. Comput. Harmon. Anal.* **24** (2008), 378–392.
- [6] W. Bao, S. Jin, P.A. Markowich: On time-splitting spectral approximations for the Schrödinger equation in the semiclassical regime. *J. Comp. Phys.* **175** (2002), 487–524.
- [7] C. Bardos, L. Erdős, F. Golse, N. Mauser, H.-T. Yau: Derivation of the Schrödinger–Poisson equation from the quantum  $N$ -body problem. *C. R. Math. Acad. Sci. Paris* **334** (2002), 515–520.
- [8] C. Bardos, F. Golse, N. Mauser: Weak coupling limit of the  $N$ -particle Schrödinger equation. *Method. Appl. Anal.* **7** (2000), 275–293.
- [9] C. Bardos, G. Lebeau, J. Rauch: Sharp sufficient conditions for the observation, control and stabilization of waves from the boundary. *SIAM J. Control Optim.* **30** (1992), 1024–1065.
- [10] J.-L. Basdevant, J. Dalibard: “Quantum Mechanics”. Springer-Verlag, Berlin, Heidelberg, 2002.
- [11] I. Bengtsson, K. Życzkowski: “The Geometry of Quantum States. An Introduction to Quantum Entanglement”. 2nd edition. Cambridge University Press, 2017.
- [12] I. Ben Porath: PhD Thesis, Ecole polytechnique, 2022.
- [13] N. Berline, E. Getzler, M. Vergne: “Heat Kernels and Dirac Operators”. Springer-Verlag, Berlin, Heidelberg, 1992.
- [14] W. Braun, K. Hepp: The Vlasov dynamics and its fluctuations in the  $1/N$  limit of interacting classical particles. *Commun. Math. Phys.* **56** (1977), 101–113.
- [15] Y. Brenier: Polar Factorization and Monotone Rearrangement of Vector-Valued Functions. *Comm. on Pure and Applied Math.* **44** (1991), 375–417.
- [16] H. Brezis: “Opérateurs maximaux monotones et semi-groupes de contraction dans les espaces de Hilbert”. North Holland, Amsterdam, 1973.
- [17] H. Brezis: “Functional Analysis, Sobolev Spaces and Partial Differential Equations”. Springer Science+Business Media, LLC 2011.
- [18] H. Brezis: Remarks on the Monge-Kantorovich problem in the discrete setting. *C. R. Acad. Sci. Paris, Sér. I* **356** (2018), 207–213.
- [19] E. Caglioti, F. Golse, T. Paul: Quantum optimal transport is cheaper. *J. Statistical Phys.* **181** (2020), 149–162
- [20] E. Caglioti, F. Golse, T. Paul: Towards optimal transport for quantum densities. Preprint arXiv:2101.03256 [math-ph]. To appear in *Ann. Sc. Norm. Super. Pisa, Cl. Sci.* DOI: 10.2422/2036-2145.202106\_011.
- [21] E.A. Carlen: Quantum Optimal Transport: Dynamics. Course at the School on Optimal Transport on Quantum Structures. Budapest, September 19th-23rd, 2022.
- [22] E.A. Carlen, J. Maas: An Analog of the 2-Wasserstein Metric in Non-Commutative Probability Under Which the Fermionic Fokker-Planck Equation is Gradient Flow for the Entropy. *Commun. Math. Phys.* **331** (2014), 887–926.
- [23] C. Cohen Tannoudji, B. Diu, F. Laloë: “Quantum Mechanics. Volume I”. 2nd edition. Hermann, Paris, and J. Wiley & Sons, Inc., 1977.
- [24] A. Connes: Compact metric spaces, Fredholm modules, and hyperfiniteness. *Ergodic Theory Dyn. Syst.* **9** (1989), 207–220.
- [25] F. D’Andrea, P. Martinetti: A dual formula for the spectral distance in noncommutative geometry. *J. Geom. Phys.* **159** (2021), Art. 103920, 8 p.
- [26] G. De Palma, D. Trevisan: Quantum optimal transport with quantum channels. *Ann. H. Poincaré* **22** (2021), 3199–3234.
- [27] S. Descombes, M. Thalhammer: An exact local error representation of exponential operator splitting methods for evolutionary problems and applications to linear Schrödinger equations in the semiclassical regime. *Numer. Math.* **50** (2010), 729–749.
- [28] J. Diestel, J.J. Uhl Jr: “Vector Measures”. American Math. Soc., Providence, 1977.
- [29] J. Dieudonné: “Foundations of Modern Analysis. Volume 12.” Academic Press, New York, London, 1969.
- [30] R.L. Dobrushin: Vlasov equations. *Funct. Anal. Appl.* **13** (1979), 115–123.
- [31] L. Erdős, H.-T. Yau: Derivation of the nonlinear Schrödinger equation from a many body Coulomb system. *Adv. Theor. Math. Phys.* **5** (2001), 1169–1205.
- [32] A. Figalli: An Introduction to Classical Optimal Transport. Course at the School on Optimal Transport on Quantum Structures. Budapest, September 19th-23rd, 2022.



- [33] A. Figalli, F. Glaudo: “An Invitation to Optimal Transport, Wasserstein Distances, and Gradient Flows”. European Mathematical Society (EMS), Zürich, 2021.
- [34] C.R. Givens, R.M. Shortt: A Class of Wasserstein Metrics for Probability Distributions. *Michigan Math. J.* **31** (1984), 231–240.
- [35] F. Golse: On the Dynamics of Large Particle Systems in the Mean-Field Limit. In “Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity”, A. Muntean, J.D.M. Rademacher and A. Zagaris eds., 1–144, Springer International Publishing Switzerland 2016.
- [36] F. Golse, S. Jin, T. Paul: On the Convergence of Time Splitting Methods for Quantum Dynamics in the Semiclassical Regime. *Foundations of Computational Math.* **21** (2021), 613–647.
- [37] F. Golse, C. Mouhot, T. Paul: On the Mean Field and Classical Limits of Quantum Mechanics. *Commun. Math. Phys.* **343** (2016), 165–205.
- [38] F. Golse T. Paul: The Schrödinger Equation in the Mean-Field and Semiclassical Regime. *Arch. Ration. Mech. Anal.* **223** (2017) 57–94.
- [39] F. Golse, T. Paul: Semiclassical evolution with low regularity. *J. Math. Pures Appl.* **151** (2021), 257–311.
- [40] F. Golse T. Paul: Optimal transport pseudometrics for quantum and classical densities. *Journal of Functional Analysis* **282** (2022), Art. 109417, 53 p.
- [41] F. Golse, T. Paul: Quantitative observability for the Schrödinger and Heisenberg equations: An optimal transport approach. *Math. Models and Methods in Applied Sciences* **32** (2022), 941–963.
- [42] F. Golse, T. Paul: Mean-Field and Classical Limit for the  $N$ -Body Quantum Dynamics with Coulomb Interaction. *Comm. on Pure and Appl. Math.* **75** (2022), 1332–1376.
- [43] S. Graffi, A. Martinez, M. Pulvirenti: Mean-field approximation of quantum systems and classical limit. *Math. Models Methods Appl. Sci.* **13** (2003), 59–73.
- [44] B.C. Hall: “Quantum Theory for Mathematicians”. Springer Science+Business Media, New York, 2013.
- [45] M. Hauray, P.-E. Jabin:  $N$ -particle approximation of the Vlasov equations with singular potential. *Arch. Ration. Mech. Anal.* **183** (2007), 489–524.
- [46] M. Hauray, P.-E. Jabin: Particle approximation of Vlasov equations with singular forces: propagation of chaos. *Ann. Sci. Éc. Norm. Supér.* **48** (2015), 891–940.
- [47] L.V. Kantorovich: On the translocation of masses. *Dokl. Akad. Nauk SSSR* **37** (1942), 227–229.
- [48] L.D. Landau, E.M. Lifshitz: “Course of Theoretical Physics. Vol. 3. Quantum Mechanics”. 3rd edition. Pergamon Press Ltd, 1977.
- [49] J.-L. Lions: “Contrôlabilité Exacte, Perturbation et Stabilisation de Systèmes Distribués. Tome 1: Contrôlabilité Exacte. Masson, Paris, 1988.
- [50] P.-L. Lions, T. Paul: Sur les mesures de Wigner. *Rev. Mat. Iberoam.* **9** (1993), 553–618.
- [51] F. Pezzotti, M; Pulvirenti: Mean-field limit and semiclassical expansion of a quantum particle system. *Ann. Henri Poincaré* **10** (2009), 145–187.
- [52] P. Pickl: A simple derivation of mean field limits for quantum systems. *Lett. Math. Phys.* **97** (2011), 151–164.
- [53] M. Reed, B. Simon: “Methods of Modern Mathematical Physics. I: Functional Analysis”. Revised edition. Academic Press, Inc., 1980.
- [54] M. Reed, B. Simon: “Methods of Modern Mathematical Physics. II: Fourier Analysis, Self-Adjointness”. Academic Press, Inc., 1975.
- [55] M. Reed, B. Simon: “Methods of Modern Mathematical Physics. IV: Analysis of Operators”. Academic Press, Inc., 1978.
- [56] I. Rodnianski, B. Schlein: Quantum fluctuations and rate of convergence towards mean field dynamics. *Commun. Math. Phys.* **291** (2009), 31–61.
- [57] W. Rudin: “Functional Analysis”. 2nd edition. Mc Graw Hill, Inc., 1991.
- [58] F. Santambrogio: “Optimal Transport for Applied Mathematicians. Calculus of Variations, PDEs and Modeling”. Birkhäuser, Springer, Cham, Heidelberg, New York, Dordrecht, London, 2015.
- [59] S. Serfaty: Mean field limit for Coulomb-type flows. *Duke Math. J.* **169** (2020), 2887–2935.
- [60] H. Spohn: Kinetic equations from Hamiltonian dynamics. *Rev. Mod. Phys.* **52** (1980), 600–640.

- [61] V.N. Sudakov: Geometric problems in the theory of infinite dimensional distributions. Proc. Steklov Inst. Math. **141** (1979), 1–178.
- [62] D. Trevisan: Quantum Optimal Transport: Quantum Channels and qubits Course at the School on Optimal Transport on Quantum Structures. Budapest, September 19th-23rd, 2022.
- [63] C. Villani: “Topics in Optimal Transportation”. Amer. Math. Soc., Providence RI, 2003.
- [64] C. Villani: “Optimal Transport. Old and New”. Springer-Verlag, Berlin, Heidelberg, 2009.
- [65] K. Życzkowski, W. Słomiński: Monge distance between quantum states. J. Phys. A **31** (1998), 9095–9104.

(F.G.) CMLS, ÉCOLE POLYTECHNIQUE, IP PARIS, 91128 PALAISEAU CEDEX, FRANCE  
*Email address:* `francois.golse@polytechnique.edu`