Progress in Nonlinear Differential Equations
and Their Applications
87

Filippo Santambrogio

Optimal Transport for Applied Mathematicians

Calculus of Variations, PDEs, and Modeling

Birkhäuser

Progress in Nonlinear Differential Equations and Their Applications

Volume 87

More information about this series at http://www.springer.com/series/4889

Filippo Santambrogio

Optimal Transport for Applied Mathematicians

Calculus of Variations, PDEs, and Modeling

Birkhäuser

Filippo Santambrogio
Laboratoire de Mathématiques d'Orsay
Université Paris-Sud
Orsay, France

ISSN 1421-1750 ISSN 2374-0280 (electronic)
Progress in Nonlinear Differential Equations and Their Applications
ISBN 978-3-319-20827-5 ISBN 978-3-319-20828-2 (eBook)
DOI 10.1007/978-3-319-20828-2

Library of Congress Control Number: 2015946595

Mathematics Subject Classification (2010): 49J45, 90C05, 35J96, 35K10, 39B62, 90B06, 91A13, 90B20, 26B25, 49N15, 52A41, 49M29, 76M25

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

To my sweet wife Anna,
who is furious for my book was conceived
after hers but is ready long before…
but she seems to love me anyway

Preface

Why this book?

Why a new book on optimal transport? Were the two books by Fields Medalist Cédric Villani not enough? And what about the Bible of Gradient Flows, the book by Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, which also contains many advanced and general details about optimal transport?

The present text, very vaguely inspired by the classes that I gave in Orsay in 2011 and 2012 and by two short introductory courses that I gave earlier in a summer school in Grenoble [274, 275], would like to propose a different point of view and is partially addressed to a different audience. There is nowadays a large and expanding community working with optimal transport as a tool to do applied mathematics. We can think in particular of applications to image processing, economics, and evolution PDEs, in particular when modeling population dynamics in biology or social sciences, or fluid mechanics. More generally, in applied mathematics, optimal transport is both a technical tool to perform proofs, do estimates, and suggest numerical methods and a modeling tool to describe phenomena where distances, paths, and costs are involved.

For those who arrive at optimal transport from this side, some of the most important issues are how to handle numerically the computations of optimal maps and costs, which are the different formulations of the problem, so as to adapt them to new models and purposes; how to obtain in the most concrete ways the main results connecting transport and PDEs; and how the theory is used in the approach of existing models. We cannot say that the answers to these questions are not contained in the existing books, but it is true that probably none of them have been written with this purpose.

The first book by C. Villani [292] is the closest to our intentions. It is a wide introduction to the topic and its applications, suitable for every audience. Yet, some of the subjects that I decided to deal with here are unfortunately not described in [292] (e.g., the minimal flow problems discussed in Chapter 4 of this book). Also, since 2003, the theory has enormously advanced. Also the books by ST Rachev

and L Ruschednorf (two volumes, [257, 258]) should not be forgotten, as they cover many applications in probability and statistics. But their scopes diverge quite soon from ours, and we will not develop most of the applications nor the language developed in [257]. Indeed, we will mainly stick to a deterministic framework and to a variational taste.

If we look at what happened after [292], we should mention at least two beautiful books which appeared since then. The new book by C. Villani [293] has expanded the previous presentation into an almost-thousand-page volume, where most of the extra content actually deals with geometrical issues, in particular the notion of curvature. Optimal transport for the quadratic cost on a manifold becomes a central tool, and it is the starting point to study the case of metric measure spaces. The other reference book is the one by L. Ambrosio, N. Gigli, and G. Savaré [15], devoted to the study of gradient flow evolution in metric space and in particular in the Wasserstein space of probability measures. This topic has many interesting applications (e.g., the heat equation, Fokker-Planck equation, porous media, etc.), and the tools that are developed are very useful. Yet, the main point of view of the authors is the study of the hidden differential structures in these abstract spaces; modeling and applied mathematics were probably not their first concerns.

Some of the above references are very technical and develop powerful abstract machineries that can be applied in very general situations but could be difficult to use for many readers. As a consequence, some shorter surveys and more tractable lecture notes have appeared (e.g., [11] for a simplified version of some parts of [15] or [9] as a short introduction to the topic by L. Ambrosio and N. Gigli). Yet, a deeper analysis of the content of [9] shows that the first half deals with the general theory of optimal transport, with some variations, while the rest is devoted to gradient flows in metric spaces in their generality and to metric measure spaces with curvature bounds. We also mention two very recent survey, [63] and [228]: the former accounts for the main achievements on the theory on the occasion of the centennial of the birth of L. V. Kantorovich, and the second gives a general presentation of the topic with connections with geometry and economics.

In the meantime, many initiatives took place underlining the increasing interest in the applied side of optimal transport: publications,[1] schools, workshops, research projects, etc. The community is very lively since the 2010s in particular in France but also in Canada, Italy, Austria, the UK, etc. All these considerations suggested that a dedicated book could have been a good idea, and the one that you have in your hands now is the result of the work for the last three years.

[1] A special issue of ESAIM M2AN on "Optimal Transport in Applied Mathematics" is in preparation and some of the bibliographical references of the present book are taken from such an issue.

What about this book?

This book contains a rigorous description of the theory of optimal transport and of some neglected variants and explains the most important connections that it has with many topics in evolution PDEs, image processing, and economics.

I avoided as much as possible the most general frameworks and concentrated on the Euclidean case, except where statements in Polish spaces did not cost any more and happened to help in making the picture clearer. I skipped many difficulties by choosing to add compactness assumptions every time that this simplified the exposition without reducing too much the interest of the statement (to my personal taste). Also, in many cases, I first started from the easiest cases (e.g., with compactness or continuity) before generalizing.

When a choice was possible, I tried to prefer more "concrete" proofs, which I think are easier to figure out for the readers. As an example, the existence of velocity fields for Lipschitz curves in Wasserstein spaces has been proven by approximation and not via abstract functional analysis tools, as in [15], where the main point is a clever use of the Hahn-Banach theorem.

I did not search for an exhaustive survey of all possible topics, but I structured the book into eight chapters, more or less corresponding to one (long) lecture each. Obviously, I added a lot of material to what one could usually deal with during one single lecture, but the choice of the topics and their order really follows an eight-lecture course given in Orsay in 2011 (only exceptions: Chapter 1 and Chapter 5 took one lecture and a half, and the subsequent ones were shortened to half a lecture in 2011). The topics which are too far from those of the eight "lectures" have been omitted from the main body of the chapters. On the other hand, every chapter ends with a discussion section, where extensions, connections, side topics, and different points of view are presented. In some cases (congested and branched transport), these discussion sections correspond to a mini-survey on a related transport model. They are more informal; and sometimes statements are proven while at other times they are only sketched or evoked.

In order to enhance the readership and allow as many people as possible to access the content of this book, I decided to explain some notations in detail that I could have probably considered as known (but it is always better to recall them). Throughout the chapters, some notions are recalled via special boxes called *Memo* or *Important Notion*.[2] A third type of box, called *Good to Know!*, provides extra notions that are not usually part of the background of nonspecialized graduate students in mathematics. The density of the boxes, and of explanatory figures as well, decreases as the book goes on.

For the sake of simplicity, I also decided not to insist too much on at least one important issue: measurability. You can trust me that all the sets, functions, and maps

[2]The difference between the two is just that in the second case, *I* would rather consider it as a Memo, but *students* usually do not agree.

that are introduced throughout the text are indeed measurable as desired, but I did not underline this explicitly. Yet, actually, this is the only concession to sloppiness: proofs are rigorous (at least, I hope) throughout the book and could be used for sure by pure or applied mathematicians looking for a reference on the corresponding subjects. The only chapter where the presentation is a little bit informal is Chapter 6, on numerical methods, in the sense that we do not give proofs of convergence or precise implementation details.

Last but not least, there is a chapter on numerical methods! In particular those that are most linked to PDEs (continuous methods), while the most combinatorial and discrete ones are briefly described in the discussion section.

For whom is this book?

This book has been written with the point of view of an applied mathematician, and applied mathematicians are supposed to be the natural readership for it. Yet, the ambition is to speak to a much wider public. Pure mathematicians (whether this distinction between pure and applied makes sense is a matter of personal taste) are obviously welcome. They will find rigorous proofs, sometimes inspired by a different point of view. They could be interested in discovering where optimal transport can be used and how and to bring their own contributions.

More generally, the distinction can be moved to the level of people working *with* optimal transport rather than *on* optimal transport (instead of pure vs applied). The former are the natural readership, but the latter can find out they are interested in the content too. In the opposite direction, can we say that the text is also addressed to nonmathematicians (physicists, engineers, theoretical economists, etc.)? This raises the question of the mathematical background that the readers need in order to read it. Obviously, if they have enough mathematical background and if they work on fields close enough to the applications that are presented, it could be interesting for them to see what is behind those applications.

The question of how much mathematics is needed also concerns students. This is a graduate text in mathematics. Even if I tried to give tools to review the required background, it is true that some previous knowledge is required to fully profit from the content of the book. The main prerequisites are measure theory and functional analysis. I deliberately decided not to follow the advice of an "anonymous" referee, who suggested to include an appendix on measure theory. The idea is that mathematicians who want to approach optimal transport should already know something on these subjects (what is a measurable function, what is a measure, which are the conditions for the main convergence theorems, what about L^p and $W^{1,p}$ functions, what is weak convergence, etc.). The goal of the Memo Boxes is to help readers to not get lost. For nonmathematicians reading the book, I hope that the choice of a more concrete approach could help them in finding out what kind of properties is important and reasonable. On the other hand, these readers are

also expected to know some mathematical language, and for sure, they will need to put in extra effort to fully profit from it.

Concerning readership, the numerical part (Chapter 6) deserves being discussed a little bit more. Comparing in detail the different methods, their drawbacks, and their strengths, the smartest tricks for their implementation, and discussing the most recent algorithms are beyond the scopes of this book. Hence, this chapter is probably useless for people already working in this field. On the other hand, it can be of interest for people working with optimal transport without knowing numerical methods or for numericists who are not into optimal transport.

Also, I am certain that it will be possible to use some of the material that I present here for a graduate course on these topics because of the many boxes recalling the main background notions and the exercises at the end of the book.

What is in this book?

After this preface and a short introduction to optimal transport (where I mainly present the problem, its history, and its main connections with other part of mathematics), this book contains eight chapters. The two most important chapters, those which constitute the general theory of optimal transport, are Chapters 1 and 5. In the structure of the book, the first half of the text is devoted to the problem of the optimal way of transporting mass from a given measure to another (in the Monge-Kantorovich framework and then with a minimal flow approach), and Chapter 1 is the most important. Then, in the second half, I move to the case where measures vary, which is indeed the case in Chapter 5 and later in Chapters 7 and 8. Chapter 6 comes after Chapter 5 because of the connections of the Benamou-Brenier method with geodesics in the Wasserstein space.

Chapter 1 presents the relaxation that Kantorovich did of the original Monge problem and its duality issues (Kantorovich potentials, c-cyclical monotonicity, etc.). It uses these tools to provide the first theorem of existence of an optimal map (Brenier theorem). The discussion section as well mainly stems from the Kantorovich interpretation and duality.

Chapter 2 focuses on the unidimensional case, which is easier and already has many consequences. Then, the Knothe map is presented; it is a transport map built with 1D bricks, and its degenerate optimality is discussed. The main notion here is that of monotone transport. In the discussion section, 1D and monotone maps are used for applications in mathematics (isoperimetric inequalities) and outside mathematics (histogram equalization in image processing).

Chapter 3 deals with some limit cases, not covered in Chapter 1. Indeed, from the results of the first chapter, we know how to handle transport costs of the form $|x-y|^p$ for $p \in (1, +\infty)$, but not $p = 1$, which was the original question by Monge. This requires to use some extra ideas, in particular selecting a special minimizer via a secondary variational problem. Similar techniques are also needed for the other limit case, i.e., $p = \infty$, which is also detailed in the chapter. In the discussion

section we present the main challenges and methods to tackle the general problem of convex costs of the form $h(y-x)$ (without strict convexity and with possible infinite values), which has been a lively research subject in the last few years, and later we consider the case $0 < p < 1$, i.e., costs which are concave in the distance.

Chapter 4 presents alternative formulations, more Eulerian in spirit: how to describe a transportation phenomenon via a flow, i.e., a vector field $\mathbf{w}$ with prescribed divergence, and minimize the total cost via functionals involving $\mathbf{w}$. When we minimize the L^1 norm of $\mathbf{w}$, this turns out to be equivalent to the original problem by Monge. The main body of the chapter provides the dictionary to pass from Lagrangian to Eulerian frameworks and back and studies this minimization problem and its solutions. In the discussion section, two variants are proposed: traffic congestion (with strictly convex costs in $\mathbf{w}$) and branched transport (with concave costs in $\mathbf{w}$).

Chapter 5 introduces another essential tool in the theory of optimal transport: the distances (called Wasserstein distances) induced by transport costs on the space of measures. After studying their properties, we study the curves in these spaces, and in particular geodesics, and we underline the connection with the continuity equation. The discussion section makes a comparison between Wasserstein distances and other distances on probabilities and finally describes an application in terms of barycenters of measures.

Chapter 6 starts from the ideas presented in the previous chapter and uses them to propose numerical methods. Indeed, in the description of the Wasserstein space, one can see that finding the optimal transport is equivalent to minimizing a kinetic energy functional among solutions of the continuity equation. This provided the first numerical method for optimal transport called the *Benamou-Brenier* method. In the rest of the chapter, two other "continuous" numerical methods are described, and the discussion section deals with discrete and semidiscrete methods.

Chapter 7 contains a "bestiary" of useful functionals defined over measures and studies their properties, not necessarily in connection with optimal transport (convexity, semicontinuity, computing the first variation, etc.). The idea is that these functionals often appear in modeling issues accompanied by transport distances. Also, the notion of displacement convexity (i.e., convexity along geodesics of the Wasserstein space) is described in detail. The discussion section is quite heterogeneous, with applications to geometrical inequalities but also equilibria in urban regions.

Chapter 8 gives an original presentation of one of the most striking applications of optimal transport: gradient flows in Wasserstein spaces, which allow us to deal with many evolution equations, in particular of the parabolic type. The general framework is presented, and the Fokker-Planck equation is studied in detail. The discussion section presents other equations which have this gradient flow structure and also other evolution equations where optimal transport plays a role, without being gradient flows.

Before the detailed bibliography and the index which conclude the book, there is a list of 69 exercises from the various chapters and of different levels of difficulties. From students to senior researchers, the readers are invited to play with these exercises and enjoy the taste of optimal transport.

Orsay, France Filippo Santambrogio
May 2015

Introduction to optimal transport

The history of optimal transport began a long time ago in France, a few years before the revolution, when Gaspard Monge proposed the following problem in a report that he submitted to the *Académie des Sciences* [239].[3] Given two densities of mass $f,\ g \geq 0$ on $\mathbb{R}^d$, with $\int f(x)\,\mathrm{d}x = \int g(y)\,\mathrm{d}y = 1$, find a map $\mathrm{T} : \mathbb{R}^d \to \mathbb{R}^d$, pushing the first one onto the other, i.e. such that

$$\int_A g(y)\,\mathrm{d}y = \int_{\mathrm{T}^{-1}(A)} f(x)\,\mathrm{d}x \quad \text{for any Borel subset } A \subset \mathbb{R}^d, \tag{1}$$

and minimizing the quantity

$$M(\mathrm{T}) := \int_{\mathbb{R}^d} |\mathrm{T}(x) - x| f(x)\,\mathrm{d}x$$

among all the maps satisfying this condition. This means that we have a collection of particles, distributed according to the density f on $\mathbb{R}^d$, that have to be moved so that they form a new distribution whose density is prescribed and is g. The movement has to be chosen so as to minimize the average displacement. In the description of Monge, the starting density f represented a distribution of sand that had to be moved to a target configuration g. These two configurations correspond to what was called in French *déblais* and *remblais*. Obviously, the dimension of the space was only supposed to be $d = 2$ or 3. The map T describes the movement (that we must choose in an optimal way), and $\mathrm{T}(x)$ represents the destination of the particle originally located at x.

In the following, we will often use the image measure of a measure μ on X (measures will indeed replace the densities f and g in the most general formulation of the problem) through a measurable map $\mathrm{T} : X \to Y$: it is the measure on Y denoted by $\mathrm{T}_\#\mu$ and characterized, as in (1), by

$$(\mathrm{T}_\#\mu)(A) = \mu(\mathrm{T}^{-1}(A)) \quad \text{for every measurable set } A$$
$$\text{or } \int_Y \phi\,\mathrm{d}\,(\mathrm{T}_\#\mu) = \int_X (\phi \circ \mathrm{T})\,\mathrm{d}\mu \quad \text{for every measurable function } \phi.$$

More generally, we can consider the problem

$$\text{(MP)} \qquad \min\{M(\mathrm{T}) := \int c(x, \mathrm{T}(x))\,\mathrm{d}\mu(x) \ :\ \mathrm{T}_\#\mu = \nu\},$$

for a more general transport cost $c : X \times Y \to \mathbb{R}$.

[3] This happened in 1781, but we translate his problem into modern mathematical language.

When we stay in the Euclidean setting, with two measures μ, ν induced by densities f, g, it is easy – just by a change-of-variables formula – to transform the equality $\nu = \mathrm{T}_{\#}\mu$ into the PDE

$$g(\mathrm{T}(x)) \det(D\mathrm{T}(x)) = f(x), \tag{2}$$

if we suppose f, g and T to be regular enough and T to be injective.

Yet, this equation is highly nonlinear in T, and this is one of the difficulties preventing an easy analysis of the Monge problem. For instance: how do we prove the existence of a minimizer? Usually, what one does is the following: take a minimizing sequence T_n, find a bound on it giving compactness in some topology (here, if the support of ν is compact, the maps T_n take value in a common bounded set, $\mathrm{spt}(\nu)$, and so one can get compactness of T_n in the weak-* L^∞ convergence), take a limit $\mathrm{T}_n \rightharpoonup \mathrm{T}$, and prove that T is a minimizer. This requires semicontinuity of the functional M with respect to this convergence (which is true in many cases, for instance, if c is convex in its second variable): we need $\mathrm{T}_n \rightharpoonup \mathrm{T} \Rightarrow \liminf_n M(\mathrm{T}_n) \geq M(\mathrm{T})$), but we also need that the limit T still satisfies the constraint. Yet, the nonlinearity of the PDE prevents us from proving this stability when we only have weak convergence (the reader can find an example of a weakly converging sequence such that the corresponding image measures do not converge as an exercise; it is actually **Ex**(1) in the list of exercises).

In [239], Monge analyzed fine questions on the geometric properties of the solution to this problem, and he underlined several important ideas that we will see in Chapter 3: the fact that transport rays do not meet, that they are orthogonal to a particular family of surfaces, and that a natural choice along transport rays is to order the points in a monotone way. Yet, he did not really solve the problem. The question of the existence of a minimizer was not even addressed. In the next 150 years, the optimal transport problem mainly remained intimately French, and the *Académie des Sciences* offered a prize on this question. The first prize was won by P. Appell [21] with a long mémoire which improved some points but was far from being satisfactory (and did not address the existence issue[4]).

The problem of Monge has stayed with no solution (does a minimizer exist? how to characterize it?) until progress was made in the 1940s. Indeed, only with the work by Kantorovich (1942, see [200]), it was inserted into a suitable framework which gave the possibility to attack it and, later, to provide solutions and study them. The problem has then been widely generalized, with very general cost functions $c(x, y)$ instead of the Euclidean distance $|x - y|$ and more general measures and spaces. The main idea by Kantorovich is that of looking at Monge's problem as connected to linear programming. Kantorovich indeed decided to change the point of view, and to describe the movement of the particles via a measure γ on $X \times Y$, satisfying $(\pi_x)_{\#}\gamma = \mu$ and $(\pi_y)_{\#}\gamma = \nu$. These probability measures over $X \times Y$

[4]The reader can see [181] – in French, sorry – for more details on these historical questions about the work by Monge and the content of the papers presented for this prize.

are an alternative way to describing the displacement of the particles of μ: instead of giving, for each x, the destination $\mathrm{T}(x)$ of the particle originally located at x, we give for each pair (x, y) the number of particles going from x to y. It is clear that this description allows for more general movements, since from a single point x, particles can a priori move to different destinations y. If multiple destinations really occur, then this movement cannot be described through a map T. The cost to be minimized becomes simply $\int_{X\times Y} c\,\mathrm{d}\gamma$. We have now a linear problem, under linear constraints. It is possible to prove existence of a solution and to characterize it by using techniques from convex optimization, such as *duality*, in order to characterize the optimal solution (see Chapter 1).

In some cases, and in particular if $c(x, y) = |x - y|^2$ (another very natural cost, with many applications in physical modeling because of its connection with kinetic energy), it is even possible to prove that the optimal γ does not allow this splitting of masses. Particles at x are only sent to a unique destination $\mathrm{T}(x)$, thus providing a solution to the original problem by Monge. This is what is done by Brenier in [82], where he also proves a very special form for the optimal map: the optimal T is of the form $\mathrm{T}(x) = \nabla u(x)$, for a convex function u. This makes, by the way, a strong connection with the Monge-Ampère equation. Indeed, from (2), we get

$$\det(D^2 u(x)) = \frac{f(x)}{g(\nabla u(x))},$$

which is an (degenerate and nonlinear) elliptic equation exactly in the class of convex functions. Brenier also uses this result to provide an original *polar factorization* theorem for vector maps (see Section 1.7.2): vector fields can be written as the composition of the gradient of a convex function and of a measure-preserving map. This generalizes the fact that matrices can be written as the product of a symmetric positive-definite matrix and an orthogonal one.

The results by Brenier can be easily adapted to other costs, strictly convex functions of the difference $x-y$. They have also been adapted to the squared distance on Riemannian manifolds (see [231]). But the original cost proposed by Monge, the distance itself, was much more difficult.

After the French school, it was time for the Russian mathematicians. From the precise approach introduced by Kantorovich, Sudakov [290] proposed a solution for the original Monge problem (MP). The optimal transport plan γ in the Kantorovich problem with cost $|x - y|$ has the following property: the space $\mathbb{R}^d$ can be decomposed in an essentially disjoint union of segments that are preserved by γ (i.e., γ is concentrated on pairs (x, y) belonging to the same segment). These segments are built from a Lipschitz function, whose level sets are the surfaces "foreseen" by Monge. Then, it is enough to reduce the problem to a family of 1D problems. If $\mu \ll \mathscr{L}^d$, the measures that μ induces on each segment should also be absolutely continuous and have no atoms. And in dimension one, as soon as the source measure has no atoms, one can define a monotone increasing transport, which is optimal for any convex cost of the difference $x - y$.

The strategy is clear, but there is a small drawback: the absolute continuity of the disintegrations of μ along segments, which sounds like a Fubini-type theorem, fails for arbitrary families of segments. Some regularity on their directions is needed. This has been observed by Ambrosio and fixed in [8, 10]. In the meantime, other proofs were obtained by Evans-Gangbo [159] (via a method which is linked to what we will see in Chapter 4 and unfortunately under stronger assumptions on the densities) and by Caffarelli-Feldman-McCann [101] and, independently, Trudinger-Wang [291], via an approximation through strictly convex costs.

After much effort on the existence of an optimal map, its regularity properties have also been studied: the main reference in this framework is Caffarelli, who proved regularity in the quadratic case, thanks to a study of the Monge-Ampère equation above. Surprisingly, at the beginning of the present century, Ma-Trudinger-Wang [219] found out the key for the regularity under different costs. In particular, they found a condition on costs $c \in C^4$ on $\mathbb{R}^d$ (some inequalities on their fourth-order derivatives) which ensured regularity. It can be adapted to the case of squared distances on smooth manifolds, where the assumption becomes a condition on the curvature of the underlying manifolds. These conditions have later been proven to be sharp by Loeper in [216]. Regularity is a beautiful and delicate matter, which cannot have in this book all the attention that it would deserve (refer to Section 1.7.6 for more details and references).

But the theory of optimal transport cannot be reduced to the existence and the properties of optimal maps. The success of this theory can be associated to the many connections it has with many other branches of mathematics. Some of these connections pass through the use of the optimal map: think of some geometric and functional inequalities that can be proven (or reproven) in this way. In this book, we only present the isoperimetric inequality (Section 2.5.3) and the Brunn-Minkowski inequality (Section 7.4.2). We stress that one of the most refined advances in quantitative isoperimetric inequalities is a result by Figalli-Maggi-Pratelli, which strongly uses optimal transport tools in the proof [167].

On the other hand, many applications of optimal transport pass, instead, through the distances they defines (Wasserstein distances; see Chapter 5). Indeed, it is possible to associate to every pair of measure a quantity, denoted by $W_p(\mu, \nu)$, based on the minimal cost to transport μ onto ν for the cost $|x - y|^p$. It can be proven to be a distance and to metrize the weak convergence of probability measures (at least on compact spaces). This distance is very useful in the study of some PDEs. Some evolution equations, in particular of parabolic type, possibly with nonlocal terms, can be interpreted as gradient flows (curves of steepest descent; see Chapter 8) for this distance. This idea has first been underlined in [198, 246]. It can be used to provide existence results or approximation of the solutions. For other PDEs, the Wasserstein distance, differentiated along two solutions, may be a technical tool to give stability and uniqueness results or rate of convergence to a steady state (see Section 5.3.5). Finally, other evolution models are connected either to the minimization of some actions involving the kinetic energy, as it is standard

in physics (the speed of a curve of densities computed w.r.t. the W_2 distance is exactly a form of kinetic energy), or to the gradient of some special convex functions appearing in the PDE (see Section 8.4.4).

The structure of the space $\mathbb{W}_p$ of probability measures endowed with the distance W_p also received and still receives a lot of attention. In particular, the study of its geodesics and of the convexity properties of some functionals along these geodesics has been important, both because they play a crucial role in the metric theory of gradient flows developed in the reference book [15] and because of their geometrical consequences. The fact that the entropy $E(\varrho) := \int \varrho \log \varrho$ is or not convex along these geodesics turned out to be equivalent, on Riemannian manifolds, to lower bounds on the Ricci curvature of the manifold. This gave rise to a wide theory of analysis in metric measure spaces, where this convexity property was chosen as a definition for the curvature bounds (see [218, 288, 289]). This theory underwent much progress in the last few years, thanks to the many recent results by Ambrosio, Gigli, Savaré, Kuwada, Ohta, and their collaborators (see, as an example, [18, 183]).

From the point of view of modeling, optimal transport may appear in many fields more or less linked to spatial economics, traffic, networks, and collective motions, but the pertinence of the Monge-Kantorovich model can be questioned, at least for some specific applications. This leads to the study of an alternative model, either more "convex" (for traffic congestion) or more "concave" (for the organization of an efficient transport network). These models are typically built under the form of a flow minimization under divergence constraints and are somehow a variant of the original Monge cost. Indeed, the original Monge problem (optimal transport from μ to ν with cost $|x-y|$) is also equivalent to the problem of minimizing the L^1 norm of a vector field $\mathbf{w}$ under the condition $\nabla \cdot \mathbf{w} = \mu - \nu$. This very last problem is presented in Chapter 4, and the traffic congestion and branched transport models are presented as a variant in Section 4.4.

Finally, in particular due to its applications in image processing (see Sections 2.5.1 and 5.5.5), it has recently become crucial to have efficient ways of computing, or approximating, the optimal transport or the Wasserstein distances between two measures. This is a new and very lively field of investigation: the methods that are presented in Chapter 6 are only some classical ones. This book does not aim to be exhaustive on this point but sheds some light on these subjects.

It is not our intention to build a separating wall between two sides of optimal transportation, the "pure" one and the "applied one." Both have been progressing at an impressive rate in the last several years. This book is devoted to those topics in the theory that could be more interesting for the reader who looks at modeling issues in economics, image processing, social and biological evolutionary phenomena, and fluid mechanics; at the applications to PDEs; and at numerical issues. It would be impossible to summarize the new directions that these topics are exploring in this short introduction, and also the book cannot do it completely.

We will only try to give a taste of these topics as well as a rigorous analysis of the mathematics which are behind them.

Acknowledgments

Since the beginning of this project, in which my goal is to communicate some ideas on optimal transport, it was very clear in my mind whom I should thank first. Indeed, my first contact with this fascinating subject dates back to a course that Luigi Ambrosio gave at SNS in Pisa in 2002, when I was an undergraduate student. I really want to thank him very much for what I learned from him[5]: the contents of his course have for long been the true basis of what I know about optimal transport.

I also want to thank two persons who accompanied me throughout all of my career and gave me the taste for optimization, modeling, and applications: my former advisor Giuseppe Buttazzo and my close collaborator Guillaume Carlier. I thank them for their support, their advice, their ideas, and their friendship.

Another person that I bothered a lot, and who explained to me many ideas and applications, is Yann Brenier. He devoted a lot of time to me in the last several months and deserves a special thanks as well. I also acknowledge the long-standing and helpful support of Lorenzo Brasco (who has been there since the very beginning of this project and gave very useful suggestions through his "anonymous" referee reports, for a total of 28 pages) and Bertrand Maury (who observed, followed, commented, and appreciated this project from its very birth and motivated some of the applied research which is behind it).

Many other people have contributed with fruitful discussions and suggestions and by reading some parts of this book. I thank in particular my PhD students Nicolas Bonnotte, Jean Louet, Alpár Mészáros, Antonin Monteil, and Paul Pegon (Paul also wrote a preliminary version of Section 4.4.2), whom I involved a lot in the creation and the evolution of this book. Let me also thank (in alphabetical order) the other colleagues and friends who contributed by their comments on various versions of the manuscript: Jean-David Benamou, Aude Chupin, Maria Colombo, Guido De Philippis, Simone Di Marino, Augusto Gerolin, Chloé Jimenez, Quentin Mérigot, Mircea Petrache, and Davide Piazzoli. Thanks also to Bruno Lévy, Gabriel Peyré,

[5] And I will not complain for the grade that I got in this course, I promise, even if it was the worst in my studies in Pisa.

Julien Rabin, and Luca Nenna for their (hyper)active help with numerical figures and to José-Antonio Carrillo, Marco Cuturi, Nicola Gigli, Robert McCann, and Mike Cullen for fruitful discussions and interesting references.

I would like to acknowledge the warm hospitality of the Maths Department of the University of Athens (Greece, not Georgia) and of the local colleagues where I progressed on the manuscript while delivering a short course in May 2014, and of the Fields Institute at Toronto, where I drafted the final version. I also acknowledge the support of CNRS, which allowed me to devote more time to the manuscript via a *délégation* (a sort of sabbatical semester) in 2013/2014.

Contents

Notation

The following are standard symbols used throughout the book without always recalling their meaning.

- "domain": a nonempty connected set in $\mathbb{R}^d$, equal to the closure of its interior.
- $\mathbb{R}_+$: nonnegative real numbers, i.e., $[0, +\infty[$.
- log: the natural neperian logarithm of a positive number.
- $\lim_n$, $\liminf_n$, $\limsup_n$ (but n could be replaced by $k, h, j \dots$): limit, inferior limit (liminf), superior limit (limsup) as $n \to \infty$ (or $k, h, j \cdots \to \infty$).
- ∇ and $\nabla\cdot$ denote gradients and divergence, respectively.
- Δ denotes the Laplacian: $\Delta u := \nabla \cdot (\nabla u)$ (and not minus it).
- Δ_p denotes the p-Laplacian: $\Delta_p u := \nabla \cdot (|\nabla u|^{p-2}\nabla u)$.
- $D^2 u$: Hessian of the scalar function u.
- $\mathscr{P}(X), \mathscr{M}(X), \mathscr{M}_+(X), \mathscr{M}^d(X)$: the spaces of probabilities, finite measures, positive finite measures, and vector measures (valued in $\mathbb{R}^d$) on X.
- $\mathscr{M}^d_{\rm div}(\Omega)$: on $\Omega \subset \mathbb{R}^d$, the space of measures $\mathbf{w} \in \mathscr{M}^d(\Omega)$ with $\nabla\cdot\mathbf{w} \in \mathscr{M}(\Omega)$.
- $\mathbb{R}^d, \mathbb{T}^d, \mathbb{S}^d$: the d-dimensional Euclidean space, flat torus, and sphere.
- $C(X), C_b(X), C_c(X), C_0(X)$: continuous, bounded continuous, compactly supported continuous, and continuous vanishing at infinity functions on X.
- $L^\infty_c(\Omega)$: L^∞ functions with compact support in Ω.
- δ_a: the Dirac mass concentrated at point a.
- $\mathbb{1}_A$: the indicator function of a set A, equal to 1 on A and 0 on A^c.
- $\wedge, \vee$: the min and max operators, i.e., $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$.
- $|A|, \mathscr{L}^d(A)$: the Lebesgue measure of a set $A \subset \mathbb{R}^d$; integration w.r.t. this measure is denoted by $\mathrm{d}x$ or $\mathrm{d}\mathscr{L}^d$.
- ω_d: the measure of the unit ball in $\mathbb{R}^d$.
- $\mathscr{H}^k$: the k-dimensional Hausdorff measure.
- $\mathrm{Per}(A)$: the perimeter of a set A in $\mathbb{R}^d$ (equal to $\mathscr{H}^{d-1}(\partial A)$ for smooth sets A).
- $\mu \ll \nu$: the measure μ is absolutely continuous w.r.t. ν.
- $\mu_n \rightharpoonup \mu$: the sequence of probabilities μ_n converges to μ in duality with C_b.
- $f \cdot \mu$: the measure with density f w.r.t. μ.

- $\mu \llcorner A$: the measure μ restricted to a set A (i.e., $\mathbb{1}_A \cdot \mu$).
- $f_{|A}$: the restriction of a function f to a set A.
- $\mathrm{T}_\# \mu$: the image measure of μ through the map T.
- $\mathrm{M}^{k \times h}$: the set of real matrices with k lines and h columns.
- I: the identity matrix.
- Tr: the trace operator on matrices.
- M^t: the transpose of the matrix M.
- $\mathrm{cof}(M)$: the cofactor matrix of M, such that $\mathrm{cof}(M) \cdot M = \det(M)\mathrm{I}$.
- id: the identity map.
- If $\mathrm{T} : X \to Y$, the map $(\mathrm{id}, \mathrm{T})$ goes from X to $X \times Y$ and is given by $x \mapsto (x, \mathrm{T}(x))$.
- $\Pi(\mu, \nu)$: the set of transport plans from μ to ν.
- $\mu \otimes \nu$: the product measure of μ and ν (s.t. $\mu \otimes \nu(A \times B) = \mu(A)\nu(B)$).
- γ_T: the transport plan in $\Pi(\mu, \nu)$ associated to a map T with $\mathrm{T}_\# \mu = \nu$.
- $M(\mathrm{T})$, $K(\gamma)$: the Monge cost of a map T and the Kantorovich cost of a plan γ.
- π_x, π_y, π_i: the projection of a product space onto its components.
- $\mathrm{AC}(\Omega)$ (or $\mathscr{C}$, if there is no ambiguity): the space of absolutely continuous curves, parametrized on $[0, 1]$ and valued in Ω.
- $L(\omega), L_k(\omega)$: the length or weighted length (with coefficient k) of the curve ω.
- $e_t : \mathrm{AC}(\Omega) \to \Omega$: the evaluation map at time t, i.e., $e_t(\omega) := \omega(t)$.
- $a^i_j, a^{ij}_{kh} \ldots$: superscripts are components (of vectors, matrices, etc.) and subscripts denote derivatives. No distinction between vectors and covectors is performed.
- $x_1, \ldots, x_d$: coordinates of points in the Euclidean space are written as subscripts.
- $\rho_t, \mathbf{v}_t \ldots$: the subscript t denotes the value at time t, not a derivative in time, which is rather denoted by ∂_t or $\frac{d}{dt}$.
- $\mathbf{n}$: the outward normal vector to a given domain.
- $W_p, \mathbb{W}_p$: Wasserstein distance and Wasserstein space of order p, respectively.
- $\mathscr{T}_c(\mu, \nu)$: the minimal transport cost from μ to ν for the cost c.
- $\mathrm{Lip}_1(\Omega)$: the set of 1-Lipschitz functions.
- $\frac{\delta F}{\delta \rho}$: first variation of $F : \mathscr{P}(\Omega) \to \mathbb{R}$, defined via $\frac{d}{d\varepsilon} F(\rho + \varepsilon\chi)_{|\varepsilon=0} = \int \frac{\delta F}{\delta \rho} \mathrm{d}\chi$.

—

The following, instead, are standard choices of notations.

- The dimension of the ambient space is d; we use $\mathbb{R}^N$ when N stands for a number of particles, of points in a discretization...
- T is a transport map, while T is typically a final time.
- ω is usually a curve (but sometimes a modulus of continuity).
- Ω is usually a domain in $\mathbb{R}^d$, and general metric spaces are usually called X.
- $\mathscr{X}$ is typically an abstract Banach space.
- ξ is usually a vector function (often a test function).
- Q is typically a measure on $\mathscr{C}$.

- ϕ is typically a test function, while φ is a Kantorovich potential or similar.
- u is typically the Kantorovich potential for the cost $|x-y|$ or the convex potential $\mathrm{T} = \nabla u$ in the $|x-y|^2$ case.
- Velocity fields are typically denoted by $\mathbf{v}$, momentum fields by $\mathbf{w}$ (when they are not time dependent) or E (when they could be time dependent).

Chapter 1
Primal and dual problems

In this chapter, we will start with generalities on the transport problem from a measure μ on a space X to another measure ν on a space Y. In general, X and Y can be complete and separable metric spaces, but soon we will focus on the case where they are the same subset $\Omega \subset \mathbb{R}^d$ (often compact). The cost function $c : X \times Y \to [0, +\infty]$ will be assumed to be continuous or semi-continuous, and then we will analyze particular cases (such as $c(x, y) = h(x - y)$ for strictly convex h).

For the sake of the exposition, the structure of this chapter is somewhat involved and deserves an explanation. In Section 1.1, we present the problems by Monge and Kantorovich and prove existence for the Kantorovich problem (KP). In Section 1.2, we present the dual problem (DP), but we do not prove the duality result $\min\,(\mathrm{KP}) = \sup\,(\mathrm{DP})$. In Section 1.3, taking this duality result as proven, we discuss cases where the solution of (KP) turns out to be induced by a transport map, hence solving (MP). Sections 1.4 and 1.5 are devoted to counterexamples to the existence for (MP) and to the equality $\min\,(\mathrm{MP}) = \min\,(\mathrm{KP})$. In Section 1.6, we introduce the notion of cyclical monotonicity, which allows us to prove the duality $\min\,(\mathrm{KP}) = \sup\,(\mathrm{DP})$, as well as stability results and sufficient optimality conditions. We also give an independent proof of duality, based on a notion of convex analysis that we actually introduce in a Memo Box in the same section. The chapter is concluded by a long discussion in Section 1.7.

1.1 Kantorovich and Monge problems

The starting point of optimal transport is the classical problem by Monge [239] which reads in its most general version, and in modern language, as follows.

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2_1

Problem 1.1. Given two probability measures $\mu \in \mathscr{P}(X)$ and $\nu \in \mathscr{P}(Y)$ and a cost function $c : X \times Y \to [0, +\infty]$, solve

$$\text{(MP)} \qquad \inf \left\{ M(\mathrm{T}) := \int c(x, \mathrm{T}(x))\, \mathrm{d}\mu(x) \ : \ \mathrm{T}_{\#}\mu = \nu \right\}, \tag{1.1}$$

where we recall that the measure denoted by $\mathrm{T}_{\#}\mu$ is defined through $(\mathrm{T}_{\#}\mu)(A) := \mu(\mathrm{T}^{-1}(A))$ for every A and is called the *image measure* or *pushforward* of μ through T.

As we pointed out in the introduction, Problem (MP)[1] is difficult because of its constraint. In particular, this constraint on T is not closed under weak convergence: see **Ex**(1) or consider $\mathrm{T}_n(x) = \sin(nx)$ on $[0, 2\pi]$.

Because of this difficulty, we will forget (MP) for a while and pass to the generalization that appears as natural from the work of Kantorovich [200]:

Problem 1.2. Given $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$, and $c : X \times Y \to [0, +\infty]$, we consider the problem

$$\text{(KP)} \quad \inf \left\{ K(\gamma) := \int_{X \times Y} c\, \mathrm{d}\gamma \ : \ \gamma \in \Pi(\mu, \nu) \right\}, \tag{1.2}$$

where $\Pi(\mu, \nu)$ is the set of the so-called *transport plans*, i.e.,

$$\Pi(\mu, \nu) = \{\gamma \in \mathscr{P}(X \times Y) : (\pi_x)_{\#}\gamma = \mu,\ (\pi_y)_{\#}\gamma = \nu\},$$

where π_x and π_y are the two projections of $X \times Y$ onto X and Y, respectively. These probability measures over $X \times Y$ are an alternative way to describe the displacement of the particles of μ: instead of specifying, for each x, which is the destination $\mathrm{T}(x)$ of the particle originally located at x, we specify for each pair (x, y) how many particles go from x to y. More precisely, the value $\gamma(A \times B)$ denotes the amount of mass moving from A to B. It is clear that this description allows for more general movements, since from a single point x, particles can a priori move to different destinations y. If multiple destinations really occur, then this movement cannot be described through a map T. Note that the constraints on $(\pi_x)_{\#}\gamma$ and $(\pi_y)_{\#}\gamma$ exactly mean that we restrict our attention to the movements that really take particles distributed according to the distribution μ and move them onto the distribution ν.

The minimizers for this problem are called *optimal transport plans* between μ and ν. Should γ be of the form $(\mathrm{id}, \mathrm{T})_{\#}\mu$ for a measurable map $\mathrm{T} : X \to Y$ (i.e., when no splitting of the mass occurs), the map T would be called the *optimal transport map* from μ to ν.

[1] To clarify the notation, let us stress that we use (MP) to denote the name *of the minimization problem*, as well as (KP), (BP), and many others later on. For its minimal value, we write min(MP).

Remark 1.3. It can be easily checked that $(\mathrm{id}, \mathrm{T})_\# \mu$ (this transport plan will be denoted by γ_{T}) belongs to $\Pi(\mu, \nu)$ if and only if T pushes μ onto ν (i.e., $\nu(A) = \mu(\mathrm{T}^{-1}(A))$ for any Borel set A) and the functional takes the form $\int c(x, \mathrm{T}(x))\mathrm{d}\mu(x)$, thus generalizing the Monge problem.

This generalized problem by Kantorovich is much easier to handle than the original one proposed by Monge; for instance, in the Monge case, we would need existence of at least a map T satisfying the constraints. This is not verified when $\mu = \delta_0$, if ν is not a single Dirac mass (see Section 1.4). On the contrary, there always exist transport plans in $\Pi(\mu, \nu)$ (for instance, $\mu \otimes \nu \in \Pi(\mu, \nu)$). Moreover, one can state that (KP) is the relaxation of the original problem by Monge (something which will be made clear in Section 1.5, and it means, roughly speaking, that (KP) is somehow the minimal extension of (MP) which has some chances to admit a minimizer).

Anyway, it is important to note that an easy use of the direct method in calculus of variations proves that a minimum does exist. This means that we take a minimizing sequence, we say that it is compact in some topology (here it is the weak convergence of probability measures), and we find a limit and prove semi-continuity (or continuity) of the functional we minimize, so that the limit is a minimizer.

Box 1.1. ***Memo**: Weierstrass criterion for the existence of minimizers, semi-continuity*

The most common way to prove that a function admits a minimizer is called the "direct method in calculus of variations." It simply consists of the classic Weierstrass theorem, possibly replacing continuity with semi-continuity.

Definition. On a metric space X, a function $f : X \to \mathbb{R} \cup \{+\infty\}$ is said to be lower semi-continuous (l.s.c. in short) if for every sequence $x_n \to x$ we have $f(x) \leq \liminf_n f(x_n)$.

Definition. A metric space X is said to be compact if from any sequence x_n, we can extract a converging subsequence $x_{n_k} \to x \in X$.

Theorem (Weierstrass). If $f : X \to \mathbb{R} \cup \{+\infty\}$ is lower semi-continuous and X is compact, then there exists $\bar{x} \in X$ such that $f(\bar{x}) = \min\{f(x) : x \in X\}$.

Proof. Define $\ell := \inf\{f(x) : x \in X\} \in \mathbb{R} \cup \{-\infty\}$ ($\ell = +\infty$ only if f is identically $+\infty$, but in this case, any point in X minimizes f). By definition, there exists a minimizing sequence x_n, i.e., points in X such that $f(x_n) \to \ell$. By compactness, we can assume $x_n \to \bar{x}$. By lower semi-continuity, we have $f(\bar{x}) \leq \liminf_n f(x_n) = \ell$. On the other hand, we have $f(\bar{x}) \geq \ell$ since ℓ is the infimum. This proves $\ell = f(\bar{x}) \in \mathbb{R}$, and this value is the minimum of f, realized at $\bar{x}$.

Box 1.2. ***Memo**: Weak compactness in dual spaces*

Definition. A sequence x_n in a Banach space $\mathscr{X}$ is said to be weakly converging to x, and we write $x_n \rightharpoonup x$, if for every $\xi \in \mathscr{X}'$ (where $\mathscr{X}'$ is the topological dual of $\mathscr{X}$ and $\langle \cdot, \cdot \rangle$ stands for the duality product between these spaces) we have $\langle \xi, x_n \rangle \to \langle \xi, x \rangle$. A sequence $\xi_n \in \mathscr{X}'$ is said to be weakly-* converging to ξ, and we write $\xi_n \overset{*}{\rightharpoonup} \xi$, if for every $x \in \mathscr{X}$ we have $\langle \xi_n, x \rangle \to \langle \xi, x \rangle$.

(continued)

Box 1.2. (continued)

Theorem (Banach-Alaoglu). If $\mathscr{X}$ is separable and ξ_n is a bounded sequence in $\mathscr{X}'$, then there exists a subsequence ξ_{n_k} weakly converging to some $\xi \in \mathscr{X}'$.

We refer, for instance, to [90] for all the details on functional analysis.

Box 1.3. ***Memo**: Duality between C_0 and $\mathscr{M}$*

Definition. A finite signed measure λ on a metric space X is a map associating to every Borel subset $A \subset X$ a value $\lambda(A) \in \mathbb{R}$ (we will see in Chapter 4 the case of vector measures, where λ is valued in $\mathbb{R}^d$) such that, for every countable disjoint union $A = \bigcup_i A_i$ (with $A_i \cap A_j = \emptyset$ for $i \neq j$), we have

$$\sum_i |\lambda(A_i)| < +\infty \quad \text{and} \quad \lambda(A) = \sum_i \lambda(A_i).$$

We denote by $\mathscr{M}(X)$ the set of finite signed measures on X. To such measures, we can associate a positive scalar measure $|\lambda| \in \mathscr{M}_+(X)$ through

$$|\lambda|(A) := \sup \left\{ \sum_i |\lambda(A_i)| \ : \ A = \bigcup_i A_i \ \text{ with } A_i \cap A_j = \emptyset \text{ for } i \neq j \right\}.$$

Theorem (Riesz representation theorem). Suppose that X is a separable and locally compact metric space. Let $\mathscr{X} = C_0(X)$ be the space of continuous function on X vanishing at infinity, i.e., $f \in C_0(X) \iff f \in C(X)$, and for every $\varepsilon > 0$, there exists a compact subset $K \subset X$ such that $|f| < \varepsilon$ on $X \setminus K$. Let us endow this space with the sup norm since $C_0(X) \subset C_b(X)$ (this last space being the space of bounded continuous functions on X). Note that $C_0(X)$ is a Banach space and that it is a closed subset $C_b(X)$. Then every element of $\mathscr{X}'$ is represented in a unique way as an element of $\mathscr{M}(X)$: for all $\xi \in \mathscr{X}'$, there exists a unique $\lambda \in \mathscr{M}(X)$ such that $\langle \xi, \phi \rangle = \int \phi \, \mathrm{d}\lambda$ for every $\phi \in \mathscr{X}$; moreover, $\mathscr{X}'$ is isomorphic to $\mathscr{M}(X)$ endowed with the norm $||\lambda|| := |\lambda|(X)$.

For signed measures of $\mathscr{M}(X)$, we should call weak-* convergence the convergence in the duality with $C_0(X)$. Yet, another interesting notion of convergence is that in duality with $C_b(X)$. We will call it (by abuse of notation) weak convergence and denote it through the symbol $\rightharpoonup$: $\mu_n \rightharpoonup \mu$ if and only if for every $\phi \in C_b(X)$ we have $\int \phi \, \mathrm{d}\mu_n \to \int \phi \, \mathrm{d}\mu$ (note that, taking $\phi = 1$, we also have $\mu_n(X) \to \mu(X)$, which is not the case for the $\overset{*}{\rightharpoonup}$ convergence). Note that $C_0(X) = C_b(X) = C(X)$ if X is compact, and in this case, the two notions of convergence are the same. On the other hand, for non-compact X, the space $\mathscr{M}(X)$ is not the dual of $C_b(X)$: by Hahn-Banach's theorem, it is possible to produce (see, for instance, Section 1.3 in [292]) elements of $C_b(X)'$ that only look at the behavior of functions of $C_b(X)$ "out of compact subsets" (i.e., at infinity or on the boundary). The notion of weak convergence in duality with C_b is also sometimes called *narrow convergence*. For all details about measure theory, we refer, for instance, to [268].

Box 1.4. ***Memo**: Weak convergence of probability measures*

Probability measures are particular measures in $\mathscr{M}(X)$: $\mu \in \mathscr{P}(X) \iff \mu \in \mathscr{M}_+(X)$ and $\mu(X) = 1$ (note that for positive measures, μ and $|\mu|$ coincide).

Definition. A sequence μ_n of probability measures over X is said to be tight if for every $\varepsilon > 0$, there exists a compact subset $K \subset X$ such that $\mu_n(X \setminus K) < \varepsilon$ for every n.

Theorem (Prokhorov). Suppose that μ_n is a tight sequence of probability measures over a complete and separable metric space X (these spaces are also called *Polish spaces*). Then there exists $\mu \in \mathscr{P}(X)$ and a subsequence μ_{n_k} such that $\mu_{n_k} \rightharpoonup \mu$ (in duality with $C_b(X)$). Conversely, every sequence $\mu_n \rightharpoonup \mu$ is necessarily tight.

Sketch of proof (of the direct implication). For every compact $K \subset X$, the measures $(\mu_n) \llcorner K$ admit a converging subsequence (in duality with $C(K)$). From tightness, we have an increasing sequence of compact sets K_i such that $\mu_n(K_i^c) < \varepsilon_i = 1/i$ for every i and every n. By a diagonal argument, it is possible to extract a subsequence μ_{n_h} such that $(\mu_{n_h}) \llcorner K_i \rightharpoonup \nu_i$ (weak convergence as $n \to \infty$ in duality with $C(K_i)$). The measures ν_i are increasing in i, and define a measure $\mu = \sup_i \nu_i$ (i.e., $\mu(A) = \sup_i \nu_i(A \cap K_i)$). In order to prove $\mu_{n_h} \rightharpoonup \mu$, take $\phi \in C_b(X)$ and write $\int_X \phi \, \mathrm{d}(\mu_{n_h} - \mu) \leq 2\varepsilon_i + \int_{K_i} \phi \, \mathrm{d}(\mu_{n_h} - \nu_i)$. This allows us to prove the convergence. Proving $\mu \in \mathscr{P}(X)$, we only need to check $\mu(X) = 1$, by testing with $\phi = 1$.

We are now ready to state some existence results.

Theorem 1.4. *Let X and Y be compact metric spaces, $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$, and $c : X \times Y \to \mathbb{R}$ a continuous function. Then* (KP) *admits a solution.*

Proof. We just need to show that the set $\Pi(\mu, \nu)$ is compact and that $\gamma \mapsto K(\gamma) = \int c \, \mathrm{d}\gamma$ is continuous and apply Weierstrass's theorem. We have to choose a notion of convergence for that and we choose to use the weak convergence of probability measures (in duality with $C_b(X \times Y)$, which is the same here as $C(X \times Y)$ or $C_0(X \times Y)$). This gives continuity of K by definition, since $c \in C(X \times Y)$.

As for the compactness, take a sequence $\gamma_n \in \Pi(\mu, \nu)$. They are probability measures, so that their mass is 1, and hence they are bounded in the dual of $C(X \times Y)$. Hence, usual weak-* compactness in dual spaces guarantees the existence of a subsequence $\gamma_{n_k} \rightharpoonup \gamma$ converging to a probability γ. We just need to check $\gamma \in \Pi(\mu, \nu)$. This may be done by fixing $\phi \in C(X)$ and using $\int \phi(x) \, \mathrm{d}\gamma_{n_k} = \int \phi \, \mathrm{d}\mu$ and passing to the limit, which gives $\int \phi(x) \, \mathrm{d}\gamma = \int \phi \, \mathrm{d}\mu$. This shows $(\pi_x)_\# \gamma = \mu$. The same may be done for π_y. More generally, the image measure through continuous maps preserves weak convergence (and here we use the map $(x, y) \mapsto x$ or $(x, y) \mapsto y$). □

Theorem 1.5. *Let X and Y be compact metric spaces, $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$, and $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ be lower semi-continuous and bounded from below. Then* (KP) *admits a solution.*

Proof. Only difference: K is no more continuous; it is l.s.c. for the weak convergence of probabilities. This is a consequence of the following lemma, applied to $f = c$ on the space $X \times Y$. □

Lemma 1.6. *If $f : X \to \mathbb{R} \cup \{+\infty\}$ is a lower semi-continuous function, bounded from below, on a metric space X, then the functional $J : \mathscr{M}_+(X) \to \mathbb{R} \cup \{+\infty\}$ defined on positive measures on X through $J(\mu) := \int f \, \mathrm{d}\mu$ is lower semi-continuous for the weak convergence of measures.*

Proof. Consider a sequence f_k of continuous and bounded functions converging increasingly to f. Then write $J(\mu) = \sup_k J_k(\mu) := \int f_k \, \mathrm{d}\mu$ (actually $J_k \leq J$ and $J_k(\mu) \to J(\mu)$ for every μ by monotone convergence). Every J_k is continuous for the weak convergence, and hence, J is l.s.c. as a supremum of continuous functionals. □

Box 1.5. ***Memo**: l.s.c. functions as suprema of Lipschitz functions*

Theorem. If f_α is an arbitrary family of lower semi-continuous functions on X, then $f = \sup_\alpha f_\alpha$ (i.e., $f(x) := \sup_\alpha f_\alpha(x)$) is also l.s.c.

Proof. Take $x_n \to x$ and write $f_\alpha(x) \leq \liminf_n f_\alpha(x_n) \leq \liminf_n f(x_n)$. Then pass to the sup in α and get $f(x) \leq \liminf_n f(x_n)$. It is also possible to check the same fact using epigraphs: indeed, a function is l.s.c. if and only if its epigraph $\{(x,t) : t \geq f(x)\} \subset X \times \mathbb{R}$ is closed, and the epigraph of the sup is the intersection of the epigraphs.

Theorem. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a function bounded from below. Then f is l.s.c. if and only if there exists a sequence f_k of k-Lipschitz functions such that for every $x \in X$, $f_k(x)$ converges increasingly to $f(x)$.

Proof. One implication is easy, since the functions f_k are continuous, hence lower semi-continuous, and f is the sup of f_k. The other is more delicate. Given f lower semi-continuous and bounded from below, let us define

$$f_k(x) = \inf_y \ (f(y) + k d(x,y)) \, .$$

These functions are k-Lipschitz continuous since $x \mapsto f(y) + kd(x,y)$ is k-Lipschitz. For fixed x, the sequence $f_k(x)$ is increasing and we have $\inf f \leq f_k(x) \leq f(x)$. We just need to prove that $\ell := \lim_k f_k(x) = \sup_k f_k(x) = f(x)$. Suppose by contradiction $\ell < f(x)$, which implies in particular $\ell < +\infty$. For every k, let us choose a point y_k such that $f(y_k) + kd(y_k, x) < f_k(x) + 1/k$. We get $d(y_k, x) \leq \frac{\ell + 1/k - f(y_k)}{k} \leq \frac{C}{k}$, thanks to the lower bound on f and to $\ell < \infty$. Hence, we know $y_k \to x$. Yet, we have $f_k(x) + 1/k \geq f(y_k)$ and we get $\lim_k f_k(x) \geq \liminf_k f(y_k) \geq f(x)$. This proves $\ell \geq f(x)$. Finally, the functions f_k may be made bounded by taking $f_k \wedge k$.

Theorem 1.7. *Let X and Y be Polish spaces, i.e., complete and separable metric spaces, $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$, and $c : X \times Y \to [0, +\infty]$ lower semi-continuous. Then* (KP) *admits a solution.*

Proof. It is now the compactness which is no more evident. We need to use the Prokhorov theorem. This means showing that any sequence in $\Pi(\mu, \nu)$ is tight. To do that, fix $\varepsilon > 0$ and find two compact sets $K_X \subset X$ and $K_Y \subset Y$ such that $\mu(X \setminus K_X), \nu(Y \setminus K_Y) < \frac{1}{2}\varepsilon$ (this is possible thanks to the converse implication in the Prokhorov theorem, since a single measure is always tight). Then the set $K_X \times K_Y$ is compact in $X \times Y$ and, for any $\gamma_n \in \Pi(\mu, \nu)$, we have

$$\gamma_n((X \times Y) \setminus (K_X \times K_Y)) \le \gamma_n((X \setminus K_X) \times Y) + \gamma_n(X \times (Y \setminus K_Y)) \\ = \mu(X \setminus K_X) + \nu(Y \setminus K_Y) < \varepsilon.$$

This shows tightness (and hence compactness) of all sequences in $\Pi(\mu, \nu)$. □

We add to this section an improvement of the continuity and semi-continuity results above, which could be useful when the cost functions are not continuous.

Lemma 1.8. *Let $\gamma_n, \gamma \in \Pi(\mu, \nu)$ be probabilities on $X \times Y$ and $a : X \to \tilde{X}$ and $b : Y \to \tilde{Y}$ be measurable maps valued in two separable metric spaces $\tilde{X}$ and $\tilde{Y}$. Let $c : \tilde{X} \times \tilde{Y} \to \mathbb{R}_+$ be a continuous function with $c(a, b) \le f(a) + g(b)$ with f, g continuous and $\int (f \circ a)\, \mathrm{d}\mu$, $\int (g \circ b)\, \mathrm{d}\nu < +\infty$. Then*

$$\gamma_n \rightharpoonup \gamma \Rightarrow \int_{X \times Y} c(a(x), b(y))\, \mathrm{d}\gamma_n \to \int_{X \times Y} c(a(x), b(y))\, \mathrm{d}\gamma.$$

Proof. We start from the case where c is bounded, say $0 \le c \le M$. We can apply the weak version of Lusin's theorem (see the observations in the next Memo 1.6) to maps valued in $\tilde{X}$ and $\tilde{Y}$. Let us fix $\delta > 0$ and find two compact sets $K_X \subset X$, $K_Y \subset Y$, with $\mu(X \setminus K_X) < \delta$ and $\nu(Y \setminus K_Y) < \delta$, such that a and b are continuous when restricted to K_X and K_Y, respectively. Let us set $K := K_X \times K_Y \subset X \times Y$, which is a compact set in the product.

We can write

$$\int c(a, b)\, \mathrm{d}\gamma_n \le \int \mathbb{1}_K c(a, b)\, \mathrm{d}\gamma_n + 2M\delta,$$

and the function $\mathbb{1}_K c(a, b)$ is upper semi-continuous on $X \times Y$ (since it is continuous and positive on a closed set, and vanishes outside it). This implies

$$\limsup_n \int c(a, b)\, \mathrm{d}\gamma_n \le \int \mathbb{1}_K c(a, b)\, \mathrm{d}\gamma + 2M\delta \le \int c(a, b)\, \mathrm{d}\gamma + 2M\delta$$

and, since δ is arbitrary, $\limsup_n \int c(a, b)\, \mathrm{d}\gamma_n \le \int c(a, b)\, \mathrm{d}\gamma$. This proves upper semi-continuity of the integral functional when c is bounded by M. An analogous computation with $M - c$ instead of c proves lower semi-continuity.

If c is positive but unbounded, just approximate it from below with its truncations $c_M = c \wedge M$, and lower semi-continuity is proven for the integral functional, which will be a sup of lower semi-continuous functionals. By replacing c with the function $\tilde{X} \times \tilde{Y} \ni (\tilde{x}, \tilde{y}) \mapsto f(\tilde{x}) + g(\tilde{y}) - c(\tilde{x}, \tilde{y})$, upper semi-continuity is proven as well. □

Box 1.6. ***Memo**: Lusin's theorem*

A well-known theorem in measure theory states that every measurable function f on a reasonable measure space (X, μ) is actually continuous on a set K with $\mu(X \setminus K)$ small. This set K can be taken compact. Actually, there can be at least two statements: either we want f to be merely continuous on K or we want f to coincide on K with a continuous function defined on X. This theorem is usually stated for real-valued functions, but we happen to need it for functions valued in more general spaces. Let us be more precise: take a topological space X endowed with a finite regular measure μ (i.e., any Borel set $A \subset X$ satisfies $\mu(A) = \sup\{\mu(K) : K \subset A, K \text{ compact}\} = \inf\{\mu(B) : B \supset A, B \text{ open}\}$). The arrival space Y will be supposed to be second countable (i.e., it admits a countable family $(B_i)_i$ of open sets such that any other open set $B \subset Y$ may be expressed as a union of B_i; for instance, separable metric spaces are second countable).

Theorem (weak Lusin). Under the above assumptions on X, Y, μ, if $f : X \to Y$ is measurable, then for every $\varepsilon > 0$, there exists a compact set $K \subset X$ such that $\mu(X \setminus K) < \varepsilon$ and the restriction of f to K is continuous.

Proof. For every $i \in \mathbb{N}$, set $A_i^+ = f^{-1}(B_i)$ and $A_i^- = f^{-1}(B_i^c)$. Consider compact sets $K_i^\pm \subset A_i^\pm$ such that $\mu(A_i^\pm \setminus K_i^\pm) < \varepsilon 2^{-i}$. Set $K_i = K_i^+ \cup K_i^-$ and $K = \bigcap_i K_i$. For each i, we have $\mu(X \setminus K_i) < \varepsilon 2^{1-i}$. By construction, K is compact and $\mu(X \setminus K) < 4\varepsilon$. To prove that f is continuous on K, it is sufficient to check that $f^{-1}(B) \cap K$ is relatively open in K for each open set B, and it is enough to check this for $B = B_i$. Equivalently, it is enough to prove that $f^{-1}(B_i^c) \cap K$ is closed, and this is true since it coincides with $K_i^- \cap K$.

Theorem (strong Lusin). Under the same assumptions on X, if $f : X \to \mathbb{R}$ is measurable, then for every $\varepsilon > 0$, there exists a compact set $K \subset X$ and a continuous function $g : X \to \mathbb{R}$ such that $\mu(X \setminus K) < \varepsilon$ and $f = g$ on K.

Proof. First apply weak Lusin's theorem, since $\mathbb{R}$ is second countable. Then we just need to extend $f_{|K}$ to a continuous function g on the whole X. This is possible since $f_{|K}$ is uniformly continuous (as a continuous function on a compact set) and hence has a modulus of continuity ω: $|f(x) - f(x')| \le \omega(d(x, x'))$ (the function ω can be taken subadditive and continuous). Then define $g(x) = \inf\{f(x') + \omega(d(x, x')) : x' \in K\}$. It can be easily checked that g is continuous and coincides with f on K.

Note that this last proof strongly uses the fact that the arrival space is $\mathbb{R}$. It could be adapted to the case of $\mathbb{R}^d$ just by extending componentwise. On the other hand, it is clear that the strong version of Lusin's theorem cannot hold for any space Y: just take X connected and Y disconnected. A measurable function $f : X \to Y$ taking values in two different connected components on two sets of positive measure cannot be approximated by continuous functions in the sense of the strong Lusin's theorem.

The consequence of all these continuity, semi-continuity, and compactness results is the existence, under very mild assumptions on the cost and the space, of an optimal transport plan γ. Then, if one is interested in the problem of Monge, the question may become, "does this minimal γ come from a transport map T?". Should the answer to this question be yes, then (MP) would have a solution, which also solves a wider problem, that of minimizing among transport plans. This is the object of the next two sections, Sections 1.2 and 1.3. On the other hand, in some cases, proving that the optimal transport plan comes from a transport map (or proving that there exists at least one optimal plan coming from a map) is equivalent

to proving that (MP) has a solution, since very often the infimum among transport plans and among transport maps is the same. This depends on the presence of atoms (see Sections 1.4 and 1.5).

1.2 Duality

The problem (KP) is a linear optimization under convex constraints, given by linear equalities or inequalities. Hence, an important tool will be duality theory, which is typically used for convex problems. We will find a dual problem (DP) for (KP) and exploit the relations between dual and primal.

The first thing we will do is finding a formal dual problem, by means of an inf-sup exchange.

Let us express the constraint $\gamma \in \Pi(\mu,\nu)$ in the following way: note that, if $\gamma \in \mathscr{M}_+(X\times Y)$, then we have

$$\sup_{\varphi,\psi} \int_X \varphi\,\mathrm{d}\mu + \int_Y \psi\,\mathrm{d}\nu - \int_{X\times Y} (\varphi(x)+\psi(y))\ \mathrm{d}\gamma = \begin{cases} 0 & \text{if } \gamma\in\Pi(\mu,\nu),\\ +\infty & \text{otherwise,}\end{cases}$$

where the supremum is taken among bounded and continuous functions φ,ψ.

Hence, we can remove the constraints on γ if we add the previous sup, since if they are satisfied, nothing has been added, and if they are not, we get $+\infty$ (which will be avoided by the minimization). Hence, we may look at the problem, we get

$$\min_{\gamma} \int_{X\times Y} c\,\mathrm{d}\gamma + \sup_{\varphi,\psi} \int_X \varphi\,\mathrm{d}\mu + \int_Y \psi\,\mathrm{d}\nu - \int_{X\times Y} (\varphi(x)+\psi(y))\,\mathrm{d}\gamma \tag{1.3}$$

and consider interchanging sup and inf:

$$\sup_{\varphi,\psi} \int_X \varphi\,\mathrm{d}\mu + \int_Y \psi\,\mathrm{d}\nu + \inf_{\gamma} \int_{X\times Y} (c(x,y) - (\varphi(x)+\psi(y)))\ \mathrm{d}\gamma.$$

We would like the two above optimization problems ("inf sup" and "sup inf") to be equivalent and the value to be the same. This is not always possible, and the main tool to do it is a theorem by Rockafellar (see [260], Section 37) requiring concavity in one variable, convexity in the other one, and some compactness assumption. Yet, Rockafellar's statement concerns finite-dimensional spaces, which is not the case here. To handle infinite-dimensional situations, one needs to use a more general mini-max theorems[2].

For now, we prefer not to investigate anymore the question of obtaining the duality equality. The result is true (under suitable assumptions), and we will see later on how to prove it. For the moment, let us accept it as true.

[2]We will give a proof in this spirit in Section 1.6.3.

If we come back to the maximization over (φ, ψ), one can rewrite the inf in γ as a constraint on φ and ψ:

$$\inf_{\gamma \geq 0} \int_{X\times Y} (c - \varphi \oplus \psi))\, \mathrm{d}\gamma = \begin{cases} 0 & \text{if } \varphi \oplus \psi \leq c \text{ on } X \times Y \\ -\infty & \text{otherwise} \end{cases},$$

where $\varphi \oplus \psi$ denotes the function defined through $(\varphi \oplus \psi)(x, y) := \varphi(x) + \psi(y)$. The above equality is easy to see: if $\varphi \oplus \psi > c$ somewhere, then use measures γ concentrated on the set where this strict inequality holds, with mass tending to ∞ and the integral tends to $-\infty$. This leads to the following dual optimization problem.

Problem 1.9. Given $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$, and the cost function $c : X \times Y \to [0, +\infty[$ we consider the problem

$$\text{(DP)} \quad \max \left\{ \int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu \ : \ \varphi \in C_b(X), \psi \in C_b(Y) : \varphi \oplus \psi \leq c \right\}. \tag{1.4}$$

First of all, we notice that $\sup(\text{DP}) \leq \min(\text{KP})$: it is enough to integrate the condition $\varphi \oplus \psi \leq c$ according to γ, to get

$$\int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu = \int_{X\times Y} (\varphi \oplus \psi)\, \mathrm{d}\gamma \leq \int_{X\times Y} c \, \mathrm{d}\gamma.$$

This is valid for every admissible (φ, ψ) and every admissible γ and proves the desired inequality.

Yet, (DP) does not admit a straightforward existence result, since the class of admissible functions lacks compactness. Let us recall the main result concerning compactness in the space of continuous functions.

Box 1.7. ***Memo**: Compactness for the uniform convergence*

Theorem (Ascoli-Arzelà). If X is a compact metric space and $f_n : X \to \mathbb{R}$ are equicontinuous (i.e., for every $\varepsilon > 0$, there exists a common $\delta > 0$ such that $|f_n(x) - f_n(y)| < \varepsilon$ for all pairs x, y with $d(x, y) < \delta$ and for all n) and equibounded (i.e., there is a common constant C with $|f_n(x)| \leq C$ for all $x \in X$ and all n), then the sequence (f_n) admits a subsequence f_{n_k} uniformly converging to a continuous function $f : X \to \mathbb{R}$.

Conversely, a subset of $C(X)$ is relatively compact for the uniform convergence (if and) only if its elements are equicontinuous and equibounded.

The same result is true if the arrival space $\mathbb{R}$ and the equiboundedness assumption are replaced with an arrival space which is a compact metric space.

Definition 1.10. Given a function $\chi : X \to \overline{\mathbb{R}}$, we define its *c-transform* (also called c-conjugate function) $\chi^c : Y \to \overline{\mathbb{R}}$ by

$$\chi^c(y) = \inf_{x\in X} c(x, y) - \chi(x).$$

We also define the $\bar{c}$-transform of $\zeta : Y \to \overline{\mathbb{R}}$ by

$$\zeta^{\bar{c}}(x) = \inf_{y \in Y} c(x, y) - \zeta(y).$$

Moreover, we say that a function ψ defined on Y is $\bar{c}$-*concave* if there exists χ such that $\psi = \chi^c$ (and, analogously, a function φ on X is said to be c-concave if there is $\zeta : Y \to \overline{\mathbb{R}}$ such that $\varphi = \zeta^{\bar{c}}$), and we denote by $c{-}conc(X)$ and $\bar{c}{-}conc(Y)$ the sets of c- and $\bar{c}$-concave functions, respectively (when $X = Y$ and c is symmetric, this distinction between c and $\bar{c}$ will play no more any role and will be dropped as soon as possible).

It is important to note that the notion of c-concavity implies a bound on the modulus of continuity.

Box 1.8. ***Memo****: Continuity of functions defined as an* inf *or* sup

Proposition. Let $(f_\alpha)_\alpha$ be a family (finite, infinite, countable, uncountable, etc.) of functions, all satisfying the same condition

$$|f_\alpha(x) - f_\alpha(x')| \le \omega(d(x, x')).$$

Consider f defined through $f(x) := \inf_\alpha f_\alpha(x)$. Then f also satisfies the same estimate.

This can be easily seen from $f_\alpha(x) \le f_\alpha(x') + \omega(d(x, x'))$, which implies $f(x) \le f_\alpha(x') + \omega(d(x, x'))$ since $f \le f_\alpha$. Then, taking the infimum over α at the r.h.s., one gets $f(x) \le f(x') + \omega(d(x, x'))$. Interchanging x and x', one obtains

$$|f(x) - f(x')| \le \omega(d(x, x')).$$

In particular, if the function $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies $\lim_{t \to 0} \omega(t) = 0$ (which means that the family $(f_\alpha)_\alpha$ is equicontinuous), then f has the same modulus of continuity (i.e., the same function ω) as the functions f_α. The same idea obviously works for the supremum instead of the infimum.

In our case, if c is continuous and finite on a compact set, and hence uniformly continuous, this means that there exists an increasing continuous function $\omega : \mathbb{R}_+ \to \mathbb{R}_+$ with $\omega(0) = 0$ such that

$$|c(x, y) - c(x', y')| \le \omega(d(x, x') + d(y, y')).$$

Hence, when we take the definition of χ^c, we have $\chi^c(y) = \inf_x g_x(y)$ with $g_x(y) := c(x, y) - \chi(x)$, and the functions g_x satisfy $|g_x(y) - g_x(y')| \le \omega(d(y, y'))$. This proves that χ^c shares the same continuity modulus also.

It is quite easy to realize that, given a pair (φ, ψ) in the maximization problem (DP), one can always replace it with (φ, φ^c) and then with $(\varphi^{c\bar{c}}, \varphi^c)$, and the constraints are preserved and the integrals increased. Actually one could go on,

but we will see that $\varphi^{c\bar{c}c} = \varphi^c$ for any φ (see Section 1.6). The goal of these transformations is to "improve" the maximizing sequence so as to get a uniform bound on its continuity.

A consequence of these considerations is the following existence result.

Proposition 1.11. *Suppose that X and Y are compact and c is continuous. Then there exists a solution (φ, ψ) to problem (DP) and it has the form $\varphi \in c-conc(X)$, $\psi \in \bar{c}-conc(Y)$, and $\psi = \varphi^c$. In particular*

$$\max\,(\mathrm{DP}) = \max_{\varphi \in c-conc(X)} \int_X \varphi \,\mathrm{d}\mu + \int_Y \varphi^c \,\mathrm{d}\nu.$$

Proof. From the considerations above, we can take a maximizing sequence (φ_n, ψ_n) and improve it, by means of c- and $\bar{c}$-transforms, so that we can assume a uniform bound on the continuity of these functions (the same modulus of continuity as c). Instead of renaming the sequence, we will still call (φ_n, ψ_n) the new sequence obtained after these transforms. We only need to check equiboundedness so as to apply Ascoli-Arzelà's theorem. This may be done if we note that adding a constant to φ and subtracting it to ψ is always possible: the value of the functional does not change, nor the constraints are affected. Hence, since φ_n is continuous on a compact set and hence bounded, we can always subtract its minimum and suppose without loss of generality that $\min \varphi_n = 0$. We get $\max \varphi_n \leq \omega(\mathrm{diam}\, X)$ (since the oscillation of a function is always less than its modulus of continuity computed at the highest possible distance in the set). So, if we have chosen $\psi_n = \varphi_n^c$, we also have $\psi_n(y) = \inf_x c(x, y) - \varphi_n(x) \in [\min c - \omega(\mathrm{diam}\, X), \max c]$. This gives uniform bounds on φ_n and ψ_n and allows us to apply Ascoli-Arzelà's theorem.

Passing to a subsequence, we can assume $\varphi_n \to \varphi$ and $\psi_n \to \psi$, both convergences being uniform. It is easy to see that

$$\int_X \varphi_n \,\mathrm{d}\mu + \int_Y \psi_n \,\mathrm{d}\nu \to \int_X \varphi \,\mathrm{d}\mu + \int_Y \psi \,\mathrm{d}\nu,$$

as a consequence of uniform convergence. Moreover,

$$\varphi_n(x) + \psi_n(y) \leq c(x, y) \;\Rightarrow\; \varphi(x) + \psi(y) \leq c(x, y)$$

(here, pointwise convergence would have been enough). This shows that (φ, ψ) is an admissible pair for (DP) and that it is optimal. □

If we admit the duality result $\min\,(\mathrm{KP}) = \max\,(\mathrm{DP})$ (the proof is postponed to Section 1.6), then we also have

$$\min\,(\mathrm{KP}) = \max_{\varphi \in c-conc(X)} \int_X \varphi \,\mathrm{d}\mu + \int_Y \varphi^c \,\mathrm{d}\nu,$$

which also shows that the minimum value of (KP) is a convex function of (μ, ν), as it is a supremum of linear functionals.

Definition 1.12. The functions φ realizing the maximum in (3.1) are called *Kantorovich potentials* for the transport from μ to ν. This is in fact a small abuse, because traditionally, this term was used only in the case $c(x, y) = |x - y|$, but it is nowadays understood in the general case as well.

Remark 1.13. Note that the Kantorovich potential φ is not necessarily c-concave, but coincides with the c-concave function $\varphi^{c\bar{c}}$ μ-a.e. However, it is always possible to choose a c-concave Kantorovich potential, and we will do it often.

1.3 The case $c(x, y) = h(x - y)$ for h strictly convex and the existence of an optimal T

This section is mainly devoted to the results that we can obtain in the case where $X = Y = \Omega \subset \mathbb{R}^d$ and the cost c is of the form $c(x, y) = h(x - y)$, for a strictly convex function h. We will also assume Ω to be compact for simplicity. This case allows for very strong results, and in particular we will find existence, as well as a representation formula, for the optimal T. Anyway, the first few lines of this section will be concerned with a more general case: that of costs functions c satisfying a *twist condition*, the most remarkable case being exactly those of the form $h(x - y)$ with h strictly convex.[3]

The main tool is the duality result. More precisely, we use Theorem 1.39, which applies to the case of compact domains with continuous cost functions c. It guarantees max (DP) = min (KP).

From the equality between the minimum of (KP) and the maximum of (DP) and the fact that both extremal values are realized, one can consider an optimal transport plan γ and a Kantorovich potential φ and write

$$\varphi(x) + \varphi^c(y) \leq c(x, y) \text{ on } \Omega \times \Omega \text{ and } \varphi(x) + \varphi^c(y) = c(x, y) \text{ on } \operatorname{spt}(\gamma).$$

The equality on $\operatorname{spt}(\gamma)$ is a consequence of the inequality which is valid everywhere and of

$$\int_{\Omega\times\Omega} c \, \mathrm{d}\gamma = \int_\Omega \varphi \, \mathrm{d}\mu + \int_\Omega \varphi^c \, \mathrm{d}\nu = \int_{\Omega\times\Omega} (\varphi(x) + \varphi^c(y)) \, \mathrm{d}\gamma,$$

which implies equality γ-a.e. These functions being continuous, the equality is satisfied on a closed set, i.e., on the support of the measure γ (let us recall the definition of support of a measure, not to be confused with sets where it is concentrated).

[3] The general results of this section are essentially due to [176], as a generalization of [82, 84].

Definition 1.14. On a separable metric space X, the support of a measure γ is defined as the smallest closed set on which γ is concentrated, i.e.,

$$\mathrm{spt}(\gamma) := \bigcap \{A \,:\, A \text{ is closed and } \gamma(X \setminus A) = 0\}\,.$$

This is well defined since the intersection may be taken countable, due to the separability assumption. Moreover, there exists also this characterization:

$$\mathrm{spt}(\gamma) = \{x \in X \,:\, \gamma(B(x,r)) > 0 \text{ for all } r > 0\}\,.$$

Once we have that, let us fix a point $(x_0, y_0) \in \mathrm{spt}(\gamma)$. One may deduce from the previous computations that

$$x \mapsto \varphi(x) - c(x, y_0) \qquad \text{is minimal at } x = x_0$$

and, if φ and $c(\cdot, y_0)$ are differentiable at x_0 and $x_0 \notin \partial\Omega$, one gets $\nabla\varphi(x_0) = \nabla_x c(x_0, y_0)$. We resume this fact in a very short statement (where we do not put the sharpest assumptions on c) since we will use it much later on.

Proposition 1.15. *If c is C^1, φ is a Kantorovich potential for the cost c in the transport from μ to ν, and (x_0, y_0) belongs to the support of an optimal transport plan γ, then $\nabla\varphi(x_0) = \nabla_x c(x_0, y_0)$, provided φ is differentiable at x_0. In particular, the gradients of two different Kantorovich potentials coincide on every point $x_0 \in \mathrm{spt}(\mu)$ where both the potentials are differentiable.*

Proof. The proof is contained in the above considerations. □

The equality $\nabla\varphi = \nabla_x c$ is particularly useful when c satisfies the following definition.

Definition 1.16. For $\Omega \subset \mathbb{R}^d$, we say that $c : \Omega \times \Omega \to \mathbb{R}$ satisfies the twist condition whenever c is differentiable w.r.t. x at every point and the map $y \mapsto \nabla_x c(x_0, y)$ is injective for every x_0. This condition is also known in economics as Spence-Mirrlees condition (see, for instance, [261]). For "nice" domains and cost functions, it corresponds to $\det\left(\frac{\partial^2 c}{\partial y_i \partial x_j}\right) \neq 0$.

The goal of this condition is to deduce from $(x_0, y_0) \in \mathrm{spt}(\gamma)$ that y_0 is indeed uniquely defined from x_0. This shows that γ is concentrated on a graph, that of the map associating y_0 to each x_0, and this map will be the optimal transport. Since this map has been constructed using φ and c only, and not γ, it also provides uniqueness for the optimal γ.

We will see this strategy with more details in the particular case where $c(x, y) = h(x - y)$, with h strictly convex, but the reader can see how to translate it into the most general case.

For this choice of c, if φ and h are differentiable at x_0 and $x_0 - y_0$, respectively, and $x_0 \notin \partial\Omega$, one gets $\nabla\varphi(x_0) = \nabla h(x_0 - y_0)$. This works if the function h

is differentiable; if it is not, we shall write $\nabla\varphi(x_0) \in \partial h(x_0 - y_0)$ (using the subdifferential of h; see Box 1.12 in Section 1.6). For a strictly convex function h, one may inverse the relation passing to $(\nabla h)^{-1}$, thus getting

$$x_0 - y_0 = (\nabla h)^{-1}(\nabla\varphi(x_0)).$$

Notice that the expression $(\nabla h)^{-1}$ makes sense for strictly convex functions h, independently of $h \in C^1$, thanks to the considerations on the invertibility of ∂h in Box 1.12.

This formula gives the solution to the transport problem with this cost, provided φ is differentiable a.e. with respect to μ. This is usually guaranteed by requiring μ to be absolutely continuous with respect to the Lebesgue measure and using the fact that φ may be proven to be Lipschitz.

Box 1.9. ***Memo**: Differentiability of Lipschitz functions*

Theorem (Rademacher). Let $f : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz continuous function. Then the set of points where f is not differentiable is negligible for the Lebesgue measure.

We do not provide a proof of this fact, but one can look at Chapter 3 of [12].

Then, one may use the previous computation to deduce that, for every x_0, the point y_0 (whenever it exists) such that $(x_0, y_0) \in \mathrm{spt}(\gamma)$ is unique (i.e., γ is of the form $\gamma_{\mathrm{T}} := (\mathrm{id}, \mathrm{T})_{\#}\mu$ where $\mathrm{T}(x_0) = y_0$). Moreover, this also gives uniqueness of the optimal transport plan and of the gradient of the Kantorovich potential.

We may summarize everything in the following theorem:

Theorem 1.17. *Given μ and ν probability measures on a compact domain $\Omega \subset \mathbb{R}^d$, there exists an optimal transport plan γ for the cost $c(x, y) = h(x - y)$ with h strictly convex. It is unique and of the form $(\mathrm{id}, \mathrm{T})_{\#}\mu$, provided μ is absolutely continuous and $\partial\Omega$ is negligible. Moreover, there exists a Kantorovich potential φ, and T and the potentials φ are linked by*

$$\mathrm{T}(x) = x - (\nabla h)^{-1}(\nabla\varphi(x)).$$

Proof. The previous theorems give the existence of an optimal γ and an optimal φ. The previous considerations show that if we take a point $(x_0, y_0) \in \mathrm{spt}(\gamma)$ where $x_0 \notin \partial\Omega$ and $\nabla\varphi(x_0)$ exists, then necessarily we have $y_0 = x_0 - (\nabla h)^{-1}(\nabla\varphi(x_0))$. The points x_0 on the boundary are negligible by assumption. The points where the differentiability fails are Lebesgue-negligible by Rademacher's theorem. Indeed, φ shares the same modulus of continuity of c, which is a Lipschitz function on $\Omega \times \Omega$ since h is locally Lipschitz continuous and Ω is bounded. Hence, φ is also Lipschitz. From the absolute continuity assumption on μ, these two sets of "bad" points (the boundary and the non-differentiability points of φ) are μ-negligible as well. This shows at the same time that every optimal transport plan is induced by a transport

map and that this transport map is $x \mapsto x - (\nabla h)^{-1}(\nabla\varphi(x))$. Hence, it is uniquely determined (since the potential φ does not depend on γ). As a consequence, we also have uniqueness of the optimal γ. □

Remark 1.18. All the costs of the form $c(x, y) = |x - y|^p$ with $p > 1$ can be dealt with via Theorem 1.17.

Remark 1.19. In the previous theorem, we showed the uniqueness of the optimal plan by giving an explicit expression for the optimal map. Yet, it is possible to use a more general argument: every time that we know that any optimal γ must be induced by a map T, then we have uniqueness. Indeed, suppose that two different plans $\gamma_1 = \gamma_{T_1}$, $\gamma_2 = \gamma_{T_2}$ are optimal: consider $\frac{1}{2}\gamma_1 + \frac{1}{2}\gamma_2$, which is optimal as well by convexity. This last transport plan cannot be induced by a map unless $T_1 = T_2$ μ-a.e., which gives a contradiction.

Remark 1.20. Theorem 1.17 states that the optimal γ is of the form $(\mathrm{id}, T)_\#\mu$, under some assumptions on μ. If the same assumptions are also satisfied by ν, then we can also say that there is an optimal map the other way around, i.e., $\gamma = (S, \mathrm{id})_\#\nu$. In particular γ-a.e. we have $y = T(x)$ and $x = S(y)$, which means $S(T(x)) = x$ γ-a.e., i.e. for μ-a.e. point x. Hence, T is invertible on a set of full measure and its inverse is the optimal map from ν to μ.

1.3.1 The quadratic case in $\mathbb{R}^d$

The case where the cost is given by $c(x, y) = \frac{1}{2}|x - y|^2$ in $\mathbb{R}^d$ deserves a special attention. The main reason for this special attention is the connection between c-concavity and the usual notion of convexity.

Proposition 1.21. *Given a function $\chi : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$, let us define $u_\chi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ through $u_\chi(x) = \frac{1}{2}|x|^2 - \chi(x)$. Then we have $u_{\chi^c} = (u_\chi)^*$. In particular, a function ζ is c-concave if and only if $x \mapsto \frac{1}{2}|x|^2 - \zeta(x)$ is convex and l.s.c.*

Proof. Just compute

$$u_{\chi^c}(x) = \frac{1}{2}|x|^2 - \chi^c(x) = \sup_y \frac{1}{2}|x|^2 - \frac{1}{2}|x-y|^2 + \chi(y) = \sup_y x\cdot y - \left(\frac{1}{2}|y|^2 - \chi(y)\right).$$

This proves the first part of the statement. Moreover, since c-concave functions are characterized by the fact that they are c-transforms and convex l.s.c. functions by the fact that they are sup of affine functions (see Box 1.11 in Section 1.6), the second part of the statement follows. □

As a consequence of the above proposition, we can particularize Theorem 1.17 to the quadratic case $c(x, y) = \frac{1}{2}|x - y|^2$, thus getting the existence of an optimal transport map

$$\mathrm{T}(x) = x - \nabla\varphi(x) = \nabla\left(\frac{x^2}{2} - \varphi(x)\right) = \nabla u(x)$$

for a convex function u. Since we will also see the converse implication (sufficient optimality conditions), this will also prove the existence and the uniqueness of a gradient of a convex function transporting μ onto ν. This well-known fact has been investigated first by Brenier (see [82]) and is usually referred to as the Brenier theorem (however, see also [227]). Section 1.7.2 will present the original approach by Brenier, called "polar factorization."

Let us moreover note that a specific approach for the case $|x-y|^2$, based on the fact that we can withdraw the parts of the cost depending on x or y only and maximize $\int x\cdot y\,\mathrm{d}\gamma$, gives the same result in an easier way: we actually get $\varphi(x_0) + \varphi^*(y_0) = x_0 \cdot y_0$ for a convex function φ and its Legendre transform φ^*, and we deduce $y_0 \in \partial\varphi(x_0)$ (see Box 1.12 in Section 1.6).

We note now that the existence of an optimal transport map is true under weaker assumptions on μ: we can replace the condition of being absolutely continuous with the condition "$\mu(A) = 0$ for any $A \subset \mathbb{R}^d$ such that $\mathscr{H}^{d-1}(A) < +\infty$" or with any condition which ensures that the non-differentiability set of u is negligible.

In [4, 5], it is proven that the set where a convex function is not differentiable is $(d-1)$-rectifiable (i.e., it is contained, up to $\mathscr{H}^{d-1}$-negligible sets, in a countable union of Lipschitz images of $\mathbb{R}^{d-1}$); more generally, the set of points where the subdifferential has dimension at least k is $(d-k)$-rectifiable. Lipschitz $(d-k)$ surfaces can be replaced with C^1 surfaces, thanks to Lusin-type approximation results of Lipschitz functions via C^1 functions (see [160], Chapter 6).

Also, it is possible to give an even stronger result about the non-differentiability set: it is contained $\mathscr{H}^{d-1}$-a.e. in a countable union of $(d-1)$ surfaces of class C^2, as it is proven in [3]. Moreover, every set which is $(d-1)$-rectifiable in this C^2 sense is contained in the non-differentiability set of a convex function[4].

After these subtle details on the sharpest assumptions on μ which guarantee the existence of an optimal transport map T (essentially, the fact that convex functions must be differentiable μ-a.e.), we want now to adapt our analysis to the case of unbounded domains. This case is not covered by the general results of this section and has not been detailed so far. Yet, it can be handled replacing Theorem 1.39 with Theorem 1.40, which is valid for the cost $c(x,y) = \frac{1}{2}|x-y|^2$ without compactness assumptions on the domain (we just need $\int |x|^2\,\mathrm{d}x, \int |y|^2\,\mathrm{d}y < +\infty$). It provides the existence of solution (φ, ψ) to the following variant of (DP):

[4]These considerations allowed N. Gigli in [182] to prove the following characterization theorem, extending 1.17: given $\mu, \nu \in \mathscr{P}(\Omega)$, suppose $\mu(A) = 0$ for every C^2 surface A of co-dimension 1; then the optimal transport plan $\gamma \in \Pi(\mu,\nu)$ for the quadratic cost $|x-y|^2$ is unique and induced by a transport map, whatever ν is. Moreover, the condition "$\mu(A) = 0$ for every C^2 surface A of co-dimension 1" characterizes the class of measures μ such that the optimal transport plans (for the quadratic cost) from μ are unique and induced by maps for every target measure ν.

$$(\mathrm{DP}-\mathrm{var}) \quad \sup\left\{\int_{\mathbb{R}^d}\varphi\,\mathrm{d}\mu+\int_{\mathbb{R}^d}\psi\,\mathrm{d}\nu\ :\ \varphi\in L^1(\mu),\psi\in L^1(\nu),\,\varphi\oplus\psi\le c\right\},$$

together with $\max(\mathrm{DP}-\mathrm{var}) = \min(\mathrm{KP})$. The sharpest result in the unbounded case is detailed in the following theorem.

Theorem 1.22. *Let μ,ν be probabilities over $\mathbb{R}^d$ and $c(x,y)=\frac{1}{2}|x-y|^2$. Suppose $\int|x|^2\,\mathrm{d}x,\int|y|^2\,\mathrm{d}y<+\infty$, which implies $\min(\mathrm{KP})<+\infty$ and suppose that μ gives no mass to $(d-1)$ surfaces of class C^2. Then there exists, unique, an optimal transport map T from μ to ν, and it is of the form $\mathrm{T}=\nabla u$ for a convex function u.*

Proof. We follow again the same scheme as before: an optimal γ exists, and Theorem 1.40 also gives the existence of an optimal dual pair $(\varphi,\psi)\in L^1(\mu)\times L^1(\nu)$ and guarantees that we have no duality gap. Then, we can see that γ is concentrated on the graph of $x\mapsto x-\nabla\varphi(x):=\nabla u(x)$ provided u (or φ, which is equivalent) is differentiable μ-a.e. Since $\varphi\in L^1(\mu)$, we infer that u is finite μ-a.e. and, since u is convex, we get $\mathrm{spt}(\mu)\subset\overline{\{u<+\infty\}}$, which is a convex set. Note that $\partial(\{u<+\infty\})$ can be expressed locally as the graph of a concave function; hence, it is $(d-1)-C^2$ rectifiable set and is μ-negligible by assumption. In the interior of $\{u<+\infty\}$, u is differentiable μ-a.e. (again, because of the assumption on μ and of u being convex). □

Remark 1.23. As a consequence of all these considerations, the quadratic case gives a very interesting result in dimension one. Suppose that $\mu\in\mathscr{P}(\mathbb{R})$ is atomless. Then every convex function is differentiable μ-a.e., since we know that the set of non-differentiability points of a convex function is at most countable (this is a consequence of the fact that if ψ is convex, then the intervals $]\psi_l'(x),\psi_r'(x)[$, where ψ_l' and ψ_r' denote the left and right derivatives, are all nonempty and disjoint when x ranges among non-differentiability points). This implies the existence of an optimal transport map for the quadratic cost between μ and any measure $\nu\in\mathscr{P}(\mathbb{R})$. This transport map will be the derivative of a convex function, i.e., a nondecreasing map.

Remark 1.24. Some of the properties of the quadratic case stay true if we pass to a general cost $c(x,y)$ which is $C^{1,1}$ and satisfies the twist condition of Definition 1.16. Obviously, the fact that the transport is the gradient of a convex function is no more true; on the other hand, the differentiability properties of Kantorovich potentials are the same as for convex functions, since one can prove that φ is semi-concave (i.e., concave, up to subtracting a quadratic function), and this allows us to use the same assumptions on μ as in Theorem 1.22.

1.3.2 The quadratic case on the flat torus

It is useful and interesting to see what happens if we replace measures on $\mathbb{R}^d$ with measures on the flat torus $\mathbb{T}^d=\mathbb{R}^d/\mathbb{Z}^d$ and the cost $\frac{1}{2}|x-y|^2$ with the squared

distance on the torus, i.e., $c(x,y) = \frac{1}{2}|[x-y]|^2$, where

$$|[z]| := \min\{|z+k| \,:\, k \in \mathbb{Z}^d\}. \tag{1.5}$$

This is the simplest example of optimal transport on a manifold, for which a general existence theorem has been established by McCann in [231]. Yet, the case of the torus, first studied in [127], is much simpler, because of the structure that $\mathbb{T}^d$ inherits from $\mathbb{R}^d$. In many cases (we will see in Chapters 5 and 6, in particular), studying what happens on the torus allows us to understand the main qualitative features of the problem, getting rid of difficulties arising from lack of compactness and/or boundary issues. In the case of compactly supported measures, it is possible to view their supports as part of a large cube, to be identified with a torus, and this at no price. Indeed, if the measures are far from the boundary of the cube, the optimal transport maps on the torus and on the Euclidean space will coincide.

Note that an optimal $\bar{k}$ in (1.5) always exists. If we take $z \in Q := [-\frac{1}{2}, \frac{1}{2}]^d$, then $\bar{k} \in \{-1, 0, +1\}^d$, and it is unique (and equal to 0) unless $z \in \partial Q$. Moreover, the function $\mathbb{R}^d \ni z \mapsto \frac{1}{2}|[z]|^2$ is semi-concave, as it is the infimum of a family of functions with Hessian equal to the identity matrix I. It is differentiable at every point where the optimal $\bar{k}$ is unique. For a given point x, we call *cut locus* of x the set of points y such that the squared distance function, viewed as a function of the pair (x,y), is not differentiable at (x,y). This is a general notion in differential geometry, and we denote it by $cut(x)$ (see [294] for its use in optimal transport), but in the case of the flat torus, it simply coincides with the set of points y such that $y - x \in \partial Q + \mathbb{Z}^d$.

Theorem 1.25. *Take $\mu, \nu \in \mathscr{P}(\mathbb{T}^d)$, with $\mu \ll \mathscr{L}^d$ and $c(x,y) = \frac{1}{2}|[x-y]|^2$. Then there exists a unique optimal transport plan $\gamma \in \Pi(\mu,\nu)$. It has the form $\gamma = \gamma_{\mathrm{T}}$ and the optimal map T is given by $\mathrm{T}(x) = x - \nabla\varphi(x)$ for a.e. x, where the sum $x - \nabla\varphi(x)$ is to be intended modulo $\mathbb{Z}^d$. Here, the function φ is a Kantorovich potential, solution to the dual problem, in the transport from μ to ν for the cost c. Moreover, for a.e. $x \in \mathbb{T}^d$, the point $\mathrm{T}(x)$ does not belong to $cut(x)$.*

Proof. The strategy is the same as for the costs of the form $h(x-y)$. We take an optimal transport plan γ, together with a pair (φ, ψ), solution of the dual problem, and we fix $(x_0, y_0) \in \mathrm{spt}(\gamma)$. We have

$$\frac{1}{2}|[x-y_0]|^2 - \varphi(x) \geq \psi(y_0), \quad \frac{1}{2}|[x_0-y_0]|^2 - \varphi(x_0) = \psi(y_0).$$

Now, take $k_0 \in \mathbb{Z}^d$ such that $|[x_0 - y_0]| = |x_0 - y_0 - k_0|$ and observe that, from the inequality $|[x-y_0]| \leq |x - y_0 - k_0|$, we have

$$\frac{1}{2}|x-y_0-k_0|^2 - \varphi(x) \geq \psi(y_0), \quad \frac{1}{2}|x_0-y_0-k_0|^2 - \varphi(x_0) = \psi(y_0).$$

This means that the function $x \mapsto \frac{1}{2}|x - y_0 - k_0|^2 - \varphi(x)$ is minimal at $x = x_0$. Yet, $\varphi(x) = \inf_y \frac{1}{2}|[x - y]|^2 - \psi(y)$ is a semi-concave function, as it is defined as an infimum of uniformly semi-concave functions. Hence, it is differentiable a.e. and μ-a.e.

If we assume that x_0 is such that $\nabla\varphi(x_0)$ exists, we deduce from the minimality above that $x_0 - y_0 - k_0 = \nabla\varphi(x_0)$. This implies that k_0 is uniquely determined by x_0 and y_0, and hence $y_0 \notin cut(x_0)$. Moreover, y_0 is uniquely determined by x_0, and we get the usual result for the existence and uniqueness of the optimal map. □

Remark 1.26. In the proof of the previous theorem, we can also observe that we have $k_0 = 0$ if $x_0 - y_0 \in Q$; this means in particular that whenever $\mathrm{spt}(\mu), \mathrm{spt}(\nu) \subset \frac{1}{2}Q$, then the transport is completely identical to what we would have with the classical quadratic cost in $\mathbb{R}^d$ instead of using the periodic framework of $\mathbb{T}^d$.

1.4 Counterexamples to existence

We want to give at least two examples that are not included in the statement of Theorem 1.17 and where an optimal transport does not exist. We concentrate on examples where the cost $c(x, y)$ would usually allow for existence results (take, for instance, $c(x, y) = |x-y|^2$), but the reason for this lack of existence is the measure μ. Other examples where this is due to the cost and not to the measures are possible; see, for instance, **Ex**(58) and **Ex**(59).

No transport may exist

The first one is very easy. Consider the case $\mu = \delta_a$, $a \in X$, and suppose that ν is not a Dirac mass. In this case, there is no transport map at all. Indeed, it is easy to check that $\mathrm{T}_\#\delta_a = \delta_{T(a)}$, and hence no transport map $\mathrm{T} : X \to Y$ can exist if ν is not of the form δ_b for some $b \in Y$.

More generally, we can say that the image measure $\mathrm{T}_\#\mu$ always includes an atom of mass at least $\mu(\{a\})$ for every atom a of μ. This implies in particular that measures with atoms cannot be sent through a transport map onto measures without atoms. For these reasons, the absence of atoms is a typical assumption on the source measure μ when one wants to solve (MP).

On the contrary, the problem (KP) by Kantorovich still makes sense even for atomic measures. In particular, the situation is very easy when $\mu = \delta_a$. In this case, the set $\Pi(\mu, \nu)$ contains a unique element, which is $\gamma = \mu \otimes \nu = \delta_a \otimes \nu$.

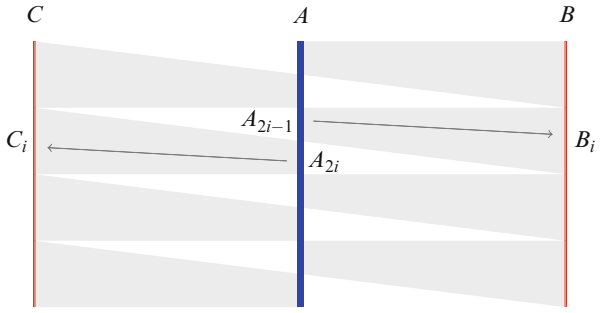
Transport can exist, but no optimal one

We consider the quadratic cost $c(x,y) = |x-y|^2$ (but any other cost strictly increasing in the distance would give the same result). Set

$$\mu = \mathscr{H}^1 \llcorner A \quad \text{and} \quad \nu = \frac{\mathscr{H}^1 \llcorner B + \mathscr{H}^1 \llcorner C}{2}$$

where A, B, and C are three vertical parallel segments in $\mathbb{R}^2$ whose vertices lie on the two lines $y=0$ and $y=1$ and the abscissas are 0, 1, and -1, respectively, and $\mathscr{H}^1$ is the 1-dimensional Hausdorff measure (this example is essentially due to [227]). It is clear that no transport plan may realize a cost better than 1 since, horizontally, every point needs to be displaced of a distance 1. Moreover, one can get a sequence of maps $\mathrm{T}_n : A \to B \cup C$ by dividing A into $2n$ equal segments $(A_i)_{i=1,\dots,2n}$ and B and C into n segments each, $(B_i)_{i=1,\dots,n}$ and $(C_i)_{i=1,\dots,n}$ (all ordered downward). Then define T_n as a piecewise affine map which sends A_{2i-1} onto B_i and A_{2i} onto C_i. In this way, the cost of the map T_n is less than $1+1/n$, which implies that the infimum of the Kantorovich problem is 1, as well as the infimum on transport maps only. Yet, no map T may obtain a cost 1, as this would imply that all points are sent horizontally, but this cannot respect the push-forward constraint. On the other hand, the transport plans associated to T_n weakly converge to the transport plan $\frac{1}{2}\gamma_{\mathrm{T}}^+ + \frac{1}{2}\gamma_{\mathrm{T}}^-$ where $\mathrm{T}^\pm(x) = x \pm e$ and $e = (1,0)$. This transport plan turns out to be the only optimal transport plan, and its cost is 1.

Note that in this last example, we also saw that the infimum among transport maps was equal to the infimum (i.e., the minimum) among transport plans. This is a general fact, which relies on the notion of relaxation, as we can see in the following section.



1.5 Kantorovich as a relaxation of Monge

Let us set again $K(\gamma) := \int_{\Omega\times\Omega} c \, \mathrm{d}\gamma$. Since we know that for any map T, we have $M(\mathrm{T}) = \int_\Omega c(x,\mathrm{T}(x))\,\mathrm{d}\mu = \int_{\Omega\times\Omega} c \, \mathrm{d}\gamma_{\mathrm{T}} = K(\gamma_{\mathrm{T}})$, the Monge problem may be

rewritten as

$$\min\{J(\gamma) \,:\, \gamma \in \Pi(\mu,\nu)\},$$

where

$$J(\gamma) = \begin{cases} K(\gamma) = M(\mathrm{T}) & \text{if } \gamma = \gamma_{\mathrm{T}}, \\ +\infty & \text{otherwise.} \end{cases}$$

This is simple to understand: the definition of J forces to restrict the minimization to those plan induced by a transport map. This fact is useful in order to consider (MP) and (KP) as two problems on the *same set of admissible objects*, where the only difference is the functional to be minimized, J or K.

The question is now: why did Kantorovich decide to replace J with K? Can we easily prove that $\inf K = \inf J$? This is obviously true when, by chance, the minimizer of K is of the form $\gamma = \gamma_{\mathrm{T}}$, since in this case, we would have equality of the two minima. But is it possible to justify the procedure in general? The main mathematical justification comes from the following notion of *relaxation*.

Box 1.10. ***Memo**: Relaxation*

Let $F : X \to \mathbb{R} \cup \{+\infty\}$ be a given functional on a metric space X, and suppose that it is bounded from below. We define the relaxation of F as the functional $\overline{F} : X \to \mathbb{R} \cup \{+\infty\}$ which is the maximal functional among those $G : X \to \mathbb{R} \cup \{+\infty\}$ which are lower semi-continuous and such that $G \le F$. This functional exists since the supremum of an arbitrary family of l.s.c. functions is also l.s.c. Moreover, we also have a representation formula, which is easy to prove:

$$\overline{F}(x) = \inf\{\liminf_n F(x_n) \,:\, x_n \to x\}.$$

A consequence of the definition is also that $\inf F = \inf \overline{F}$ (this latter infimum, that of $\overline{F}$, being often a minimum, when X is compact). This is easy to check: $F \ge \overline{F}$ implies $\inf F \ge \inf \overline{F}$; on the other hand, we have $F \ge \ell$ where $\ell := \inf F$ and a constant function is l.s.c. Hence $\overline{F} \ge \ell$ and $\inf \overline{F} \ge \inf F$.

Here we claim that, under some assumptions, K is actually the relaxation of J. It will happen, in this case, by chance, that this relaxation is also continuous, instead of only semi-continuous, and that it coincides with J on $\{J < +\infty\}$.

The assumptions are the following: we take $\Omega \subset \mathbb{R}^d$ to be compact, c continuous, and μ atomless (i.e., for every $x \in \Omega$, we have $\mu(\{x\}) = 0$). Compactness is supposed for simplicity, while the result could be true under more general assumptions. The references on this subject are [253] and [8].

We need some preliminary results. The first concerns the one-dimensional case, we will analyze in detail in Chapter 2, but can be stated right now.

Lemma 1.27. *If μ, ν are two probability measures on the real line $\mathbb{R}$ and μ is atomless, then there exists at least a transport map* T *such that* $\mathrm{T}_{\#}\mu = \nu$.

Proof. Just consider the monotone increasing map T provided by Remark 1.23. This map also optimizes the quadratic cost, but here we don't care about it. □

Lemma 1.28. *There exists a Borel map $\sigma_d : \mathbb{R}^d \to \mathbb{R}$ which is injective, its image is a Borel subset of $\mathbb{R}$, and its inverse map is Borel measurable as well.*

Proof. First, note that it is sufficient to prove this result for $d = 2$, since then one can proceed by induction: if a map σ_{d-1} is given on $\mathbb{R}^{d-1}$, then one can produce a map σ_d by considering $\sigma_d(x_1, x_2, \dots, x_d) = \sigma_2(x_1, \sigma_{d-1}(x_2, x_3, \dots, x_d))$.

Then, note also that it is enough to define such a map on $]0,1[^2$, since one can go from $\mathbb{R}^2$ to $]0,1[^2$ by considering $(x, y) \mapsto (\frac{1}{2} + \frac{1}{\pi}\arctan x, \frac{1}{2} + \frac{1}{\pi}\arctan y)$.

Then, consider the map which associates to the pair (x, y), where $x = 0, x_1x_2x_3\dots$ and $y = 0, y_1y_2y_3\dots$ in decimal (or binary) notation, the point $0, x_1y_1x_2y_2x_3y_3\dots$. In order to avoid ambiguities, we can decide that no decimal notation is allowed to end with a periodic 9 (i.e., $0.347299999\dots$ has to be written as 0.3473). This is why the image of this map will not be the whole interval, since the points like $0.39393939\dots$ are not obtained through this map. But this set of points is actually Borel measurable.

It is not difficult neither to check that the map is Borel measurable, as well as its inverse, since the pre-image of every interval defined by prescribing the first $2k$ digits of a number in $\mathbb{R}$ is just a rectangle in $\mathbb{R}^2$, the product of two intervals defined by prescribing the first k digits of every component. These particular intervals being a base for the Borel tribe proves the measurability we need. □

Corollary 1.29. *If μ, ν are two probability measures on $\mathbb{R}^d$ and μ is atomless, then there exists at least a transport map* T *such that* $\mathrm{T}_{\#}\mu = \nu$.

Proof. This is just obtained by considering a transport map T from $(\sigma_d)_{\#}\mu$ to $(\sigma_d)_{\#}\nu$ and then composing with σ_d and $(\sigma_d)^{-1}$. □

Remark 1.30. We provided here, through Lemma 1.28, an explicit way of constructing a transport map between an atomless measure μ and an arbitrary measure ν, when the ambient space is $\mathbb{R}^d$. Actually, this fact is much more general, since it is well known that any measure space endowed with an atomless measure is isomorphic to $[0, 1]$ with the Lebesgue measure[5]. Yet, we do not want to introduce this kind of arguments when an explicit construction is sufficient to deal with the Euclidean case, which already meets our scopes.

[5]Note that this was the argument used in [175] to obtain the existence of transport maps in general spaces, and [253] points out that, when composing with the isomorphisms sending arbitrary spaces to $[0, 1]$, the cost function is no more continuous. Yet, this could be handled using Lemma 1.8, which allows us to use costs of this form: the composition of a continuous function with measurable maps of x and y separately.

A last lemma:

Lemma 1.31. *Consider on a compact metric space X, endowed with a probability $\varrho \in \mathscr{P}(X)$, a sequence of partitions G_n, each G_n being a family of disjoint subsets $C_{i,n}$ such that $\bigcup_{i\in I_n} C_{i,n} = X$ for every n. Suppose that* $\operatorname{size}(G_n) := \max_i \operatorname{diam}(C_{i,n})$ *tends to 0 as $n \to \infty$ and consider a sequence of probability measures ϱ_n on X such that, for every n and $i \in I_n$, we have $\varrho_n(C_{i,n}) = \varrho(C_{i,n})$. Then $\varrho_n \rightharpoonup \varrho$.*

Proof. Set $m_{i,n} := \varrho_n(C_{i,n}) = \varrho(C_{i,n})$. It is sufficient to take a continuous function $\phi \in C(X)$ and note that

$$\left|\int_X \phi \,\mathrm{d}\varrho_n - \int_X \phi \,\mathrm{d}\varrho\right| \le \sum_{i\in I_n} \left|\int_{C_{i,n}} \phi \,\mathrm{d}\varrho_n - \int_{C_{i,n}} \phi \,\mathrm{d}\varrho\right|$$
$$\le \omega(\operatorname{diam}(C_{i,n})) \sum_{i\in I_n} m_{i,n} = \omega(\operatorname{diam}(C_{i,n})) \to 0,$$

where ω is the modulus of continuity of ϕ. This is justified by the fact that, whenever two measures have the same mass on a set $C \subset X$, since the oscillation of ϕ on the same set does not exceed $\omega(\operatorname{diam}(C))$, the difference of the two integrals is no more than this number times the common mass.

This proves $\int \phi \,\mathrm{d}\varrho_n \to \int \phi \,\mathrm{d}\varrho$ and hence $\varrho_n \to \varrho$. □

We can now prove the following.

Theorem 1.32. *On a compact subset $\Omega \subset \mathbb{R}^d$, the set of plans γ_{T} induced by a transport is dense in the set of plans $\Pi(\mu,\nu)$ whenever μ is atomless.*

Proof. Fix n and consider any partition of Ω into sets $K_{i,n}$ of diameter smaller than $1/(2n)$ (for instance, small cubes). The sets $C_{i,j,n} := K_{i,n} \times K_{j,n}$ make a partition of $\Omega \times \Omega$ with size smaller than $1/n$.

Let us now take any measure $\gamma \in \Pi(\mu,\nu)$. Thanks to Lemma 1.31, we only need to build a transport T sending μ to ν, such that γ_{T} gives the same mass as γ to each one of the sets $C_{i,j,n}$. To do this, define the columns $Col_{i,n} := K_{i,n} \times \Omega$ and denote by $\gamma_{i,n}$ the restriction of γ on $Col_{i,n}$. Its marginal will be denoted by $\mu_{i,n}$ and $\nu_{i,n}$. Consider now, for each i, a transport map $\mathrm{T}_{i,n}$ sending $\mu_{i,n}$ to $\nu_{i,n}$. It exists thanks to Corollary 1.29, since for each (i,n), we have $\mu_{i,n} \le \mu$, which makes these measures atomless. Since the $\mu_{i,n}$ are concentrated on disjoint sets, by "gluing" the transports $\mathrm{T}_{i,n}$, we get a transport T sending μ to ν (using $\sum_i \mu_{i,n} = \mu$ and $\sum_i \nu_{i,n} = \nu$).

It is enough to check that γ_{T} gives the same mass[6] as γ to every $C_{i,j,n}$, but it is easy to prove. Indeed, this mass equals that of $\gamma_{\mathrm{T}_{i,n}}$ and $\gamma_{\mathrm{T}_{i,n}}(C_{i,j,n}) = \mu_{i,n}(\{x : x \in K_{i,n},\ \mathrm{T}_{i,n}(x) \in K_{j,n}\}) = \mu_{i,n}(\{x : \mathrm{T}_{i,n}(x) \in K_{j,n}\}) = \nu_{i,n}(K_{j,n}) = \gamma(K_{i,n} \times K_{j,n})$. □

[6]Note that [8] and [253] proved the same fact in a more complicated way: indeed, the use of a unique transport map for every column is enough. The same idea has been used in [146] for a different density result, concerning Lipschitz transport maps.

The relaxation result is just a consequence.

Theorem 1.33. *Under the above-mentioned assumptions, K is the relaxation of J. In particular,* $\inf J = \min K$*, and hence Monge and Kantorovich problems have the same infimum.*

Proof. First note that, since K is continuous, then it is l.s.c., and since we have $K \leq J$, then K is necessarily smaller than the relaxation of J. We only need to prove that, for each γ, we can find a sequence of transports T_n such that $\gamma_{T_n} \rightharpoonup \gamma$ and $J(\gamma_{T_n}) \to K(\gamma)$, so that the infimum in the sequential characterization of the relaxed functional (see definition) will be smaller than K, thus proving the equality. Actually, since for $\gamma = \gamma_{T_n}$ the two functionals K and J coincide, and since K is continuous, we only need to produce a sequence T_n such that $\gamma_{T_n} \rightharpoonup \gamma$. This is possible thanks to Theorem 1.32. □

1.6 Convexity, c-concavity, cyclical monotonicity, duality, and optimality

1.6.1 Convex and c-concave functions

In this section, we analyze properties of c-concave functions in comparison with convex functions, which are better known. We start from recalling some notions from convex analysis.

Box 1.11. ***Memo**: Convex functions*

The definition of convex function is always the same: $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex if and only if for all $x, y \in \mathbb{R}^d$ and $t \in [0, 1]$, we have $f((1-t)x+ty) \leq (1-t)f(x)+tf(y)$. We do not care about convex functions defined on subsets $\Omega \subset \mathbb{R}^d$ (which should be convex, for the definition to make sense) since one can always extend f outside Ω by setting it to $+\infty$, preserving convexity.

An easy stability property is the following: if f_α is a family of convex functions, then f defined through $f(x) := \sup_\alpha f_\alpha(x)$ is also convex. This can be checked easily by writing $f_\alpha((1-t)x + ty) \leq (1-t)f_\alpha(x) + tf_\alpha(y) \leq (1-t)f(x) + tf(y)$ and passing to the sup in α. Otherwise, we can also use the epigraphs (and a function is convex if and only if its epigraph is convex) and note that the intersection of convex sets is still convex.

Real-valued convex functions defined on $\mathbb{R}^d$ are automatically continuous and locally Lipschitz (and in general, they are continuous and locally Lipschitz in the interior of their domain $\{f < +\infty\}$). However, on the boundary between $\{f < +\infty\}$ and $\{f = +\infty\}$, there could be discontinuities. This is why typically we require at least lower semi-continuity.

Theorem. A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex and l.s.c. if and only if there exists a family of affine functions f_α such that $f(x) := \sup_\alpha f_\alpha(x)$. This family can also be chosen to be the family of all affine functions smaller than f.

(continued)

Box 1.11. (continued)

One implication is easy (if f is a sup of affine functions, since affine implies convex, then it is convex). The geometrical intuition behind the converse one is the fact that the epigraph of a convex and l.s.c. function is a closed convex set in $\mathbb{R}^d \times \mathbb{R}$ and can be written as the intersection of the half-spaces which contain it. A precise proof requires the use of the Hahn-Banach theorem in its geometrical form.

Box 1.12. ***Memo**: Legendre transform, subdifferential*

Definition (Legendre-Flenchel transform). For any given function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, we can define its transform $f^*(y) := \sup_x x \cdot y - f(x)$.

Example. The Legendre transform of $x \mapsto \frac{1}{p}|x|^p$ is $x \mapsto \frac{1}{q}|x|^q$, where $\frac{1}{p} + \frac{1}{q} = 1$.

A function f is convex and l.s.c. if and only if there exists g such that $f = g^*$ (since the class of functions that are Legendre transform of something exactly agrees with that of functions which are expressed as suprema of affine functions). To be convinced that every supremum of affine functions can be expressed as g^*, just take $f(y) = \sup_\alpha x_\alpha \cdot y + b_\alpha$; then, for each vector x, set $g(x) = -\sup\{b_\alpha : x_\alpha = x\}$ (by setting $g(x) = +\infty$ if no α is such that $x_\alpha = x$) and check that we have $f = g^*$.

Proposition. A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex and l.s.c. if and only if $f^{**} = f$.

The proof of this fact can be seen as an application of the Hahn-Banach theorem. For a complete proof, even in infinite dimension, see, for instance, [23].

Definition (subdifferential). For every convex function f, we define its subdifferential at x as the set $\partial f(x) = \{p \in \mathbb{R}^d : f(y) \geq f(x) + p \cdot (y - x)\ \forall y \in \mathbb{R}^d\}$.

It is possible to prove that $\partial f(x)$ is never empty if x lies in the interior of the set $\{f < +\infty\}$. At every point where the function f is differentiable, then ∂f reduces to the singleton $\{\nabla f\}$.

For the subdifferential of the convex functions f and f^*, we have

$$p \in \partial f(x) \Leftrightarrow x \in \partial f^*(p) \Leftrightarrow f(x) + f^*(p) = x \cdot p.$$

This helps in proving the following equivalence: take two conjugate functions f and f^* (with $f = f^{**}$); then f is C^1 if and only if f^* is strictly convex. This comes from the fact that C^1 means that there is at most one p in every set $\partial f(x)$, while strictly convex means that the same vector cannot belong to the subdifferential of more than one point. In particular, if h is strictly convex, then ∂h, which is a multivalued map, can be inverted, and is univalued, thus getting a map $(\partial h)^{-1}$ that should be used in the statement of Theorem 1.17 instead of $(\nabla h)^{-1}$.

Finally, note that the subdifferential of a convex function satisfies this monotonicity property: if $p_1 \in \partial f(x_1)$ and $p_2 \in \partial f(x_2)$, then $(p_1 - p_2) \cdot (x_1 - x_2) \geq 0$.

Note the analogy with c-concave functions (or $\bar{c}$-concave, this distinction being meaningless in the symmetric case), which were defined as being the $\bar{c}$-transform (or c-transform) of something. Indeed, up to the change of sign, c-concavity has exactly been defined as a generalization of convexity. Note that we admit the value $+\infty$ for convex functions and $-\infty$ for concave and c-concave ones. Yet, we do not like the convex function which is identically $+\infty$ nor any c-concave function which is identically $-\infty$.

Instead of proving the result on $f^{**} = f$, we may prove this more general one concerning c-concave functions[7].

Proposition 1.34. *Suppose that c is real valued. For any $\varphi : X \to \mathbb{R} \cup \{-\infty\}$, we have $\varphi^{c\bar{c}} \geq \varphi$. We have the equality $\varphi^{c\bar{c}} = \varphi$ if and only if φ is c-concave; in general, $\varphi^{c\bar{c}}$ is the smallest c-concave function larger than φ.*

Proof. First we prove $\varphi^{c\bar{c}} \geq \varphi$. Write

$$\varphi^{c\bar{c}}(x) = \inf_y c(x,y) - \varphi^c(y) = \inf_y c(x,y) - \inf_{x'} c(x',y) - \varphi(x').$$

Consider that $\inf_{x'} c(x',y) - \varphi(x') \leq c(x,y) - \varphi(x)$, and hence

$$\varphi^{c\bar{c}}(x) \geq \inf_y c(x,y) - c(x,y) + \varphi(x) = \varphi(x).$$

Analogously, we have $\xi^{\bar{c}c} \geq \xi$ for any $\xi : Y \to \mathbb{R} \cup \{-\infty\}$.

Then, let us prove $\varphi^{c\bar{c}} = \varphi$ if φ is c-concave. In such a case, we may write

$$\varphi = \zeta^{\bar{c}} \Rightarrow \varphi^c = \zeta^{\bar{c}c} \geq \zeta \Rightarrow \varphi^{c\bar{c}} \leq \zeta^{\bar{c}} = \varphi,$$

where the last inequality is obtained by noting that c- and $\bar{c}$-transforms revert the inequalities between functions (due to the minus sign in the definition). This proves that in such a case, we have $\varphi^{c\bar{c}} \leq \varphi$, and hence $\varphi^{c\bar{c}} = \varphi$.

Finally, we can prove that for any φ, the function $\varphi^{c\bar{c}}$ is the smallest c-concave function larger than φ. To prove that, take $\tilde{\varphi} = \chi^{\bar{c}}$ any c-concave function and suppose $\tilde{\varphi} \geq \varphi$. Then consider

$$\chi^{\bar{c}} \geq \varphi \Rightarrow \chi^{\bar{c}c} \leq \varphi^c \Rightarrow \chi \leq \varphi^c \Rightarrow \tilde{\varphi} = \chi^{\bar{c}} \geq \varphi^{c\bar{c}},$$

which proves the desired inequality. □

We finish this section with a last concept about convex functions, which we will then translate into the framework of c-concave functions later on.

Let us define the graph of the subdifferential of a convex function as

$$\mathrm{Graph}(\partial f) := \{(x,p) \,:\, p \in \partial f(x)\} = \{(x,p) \,:\, f(x) + f^*(p) = x \cdot p\}.$$

We already know that this graph is monotone in the sense that $(x_i, p_i) \in \mathrm{Graph}(\partial f)$ for $i = 1, 2$ implies

$$(p_2 - p_1) \cdot (x_2 - x_1) \geq 0.$$

[7]Note that this result is easy to prove, since c-concave functions are exactly defined as c-transform of something, while usually convex functions are not defined via sup of affine functions, but via the convexity inequality. This is why the corresponding theorem for convex functions usually requires the use of the Hahn-Banach theorem.

Yet, not all monotone graphs are the graphs of the subdifferential of a convex function, nor they are contained in one of such graphs.[8]

Hence, monotonicity is not enough to characterize gradients and subdifferential of convex functions (to be more precise, gradient vector functions that are monotone are gradient of convex functions, but monotonicity alone does not imply *gradientness*).

A stronger notion is that of cyclical monotonicity.

Definition 1.35. A set $A \subset \mathbb{R}^d \times \mathbb{R}^d$ is said to be cyclically monotone if for every $k \in \mathbb{N}$, every permutation σ, and every finite family of points $(x_1, p_1), \dots, (x_k, p_k) \in A$ we have

$$\sum_{i=1}^{k} x_i \cdot p_i \geq \sum_{i=1}^{k} x_i \cdot p_{\sigma(i)}.$$

The word "cyclical" refers to the fact that, since every σ is the disjoint composition of cycles, it is enough to check this property for cyclical permutations, i.e., replacing $\sum_{i=1}^{k} x_i \cdot p_{\sigma(i)}$ with $\sum_{i=1}^{k} x_i \cdot p_{i+1}$ in the definition (with the obvious convention $p_{k+1} = p_1$).

Note that if we take $k = 2$ we get the usual definition of monotonicity, since

$$x_1 \cdot p_1 + x_2 \cdot p_2 \geq x_1 \cdot p_2 + x_2 \cdot p_1 \iff (p_1 - p_2) \cdot (x_1 - x_2) \geq 0.$$

A famous theorem by Rockafellar ([260], Theorem 24.8) states that every cyclically monotone set is contained in the graph of the subdifferential of a convex function. We will not prove this theorem here as will see it as a particular case of a theorem on c-concave functions.

1.6.2 c-Cyclical monotonicity and duality

We start from the translation to the c-concave case of the definition of cyclical monotonicity.

Definition 1.36. Once a function $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is given, we say that a set $\Gamma \subset X \times Y$ is c-cyclically monotone (briefly c-CM) if for every $k \in \mathbb{N}$, every permutation σ and every finite family of points $(x_1, y_1), \dots, (x_k, y_k) \in \Gamma$ we have

[8] Take, for instance, a 90° rotation R in $\mathbb{R}^2$ and consider the set $A = \{(x, Rx),\ x \in \mathbb{R}^d\}$. This set satisfies the monotonicity inequality since we have $(Rx_1 - Rx_2) \cdot (x_1 - x_2) = 0$ for any x_1 and x_2. Yet, the map $x \mapsto Rx$ is not the gradient of convex functions (it is not a gradient at all), nor can it be contained in the graph of the subdifferential of a convex function.

$$\sum_{i=1}^{k} c(x_i, y_i) \leq \sum_{i=1}^{k} c(x_i, y_{\sigma(i)}).$$

As for the convex case, the word "cyclical" refers to the fact that we can restrict our attention to cyclical permutations. The word "monotone" is a leftover from the case $c(x, y) = -x \cdot y$.

It is useful to recall the following theorem, which is a generalization of a theorem by Rockafellar in convex analysis. The classical reference for this proof is [269], but it is interesting to see that it can actually be traced back to much older papers, such as [261] (in the framework of rationalizability of actions in economics, Theorem 1 in the paper) or [91] (for applications to liquid crystals, as a part of a proof of a discrete Kantorovich duality for distance costs, see Lemma 2 in the paper).

Theorem 1.37. *If $\Gamma \neq \emptyset$ is a c-CM set in $X \times Y$ and $c : X \times Y \to \mathbb{R}$ (note that c is required not to take the value $+\infty$), then there exists a c-concave function $\varphi : X \to \mathbb{R} \cup \{-\infty\}$ (different from the constant $-\infty$ function) such that*

$$\Gamma \subset \{(x, y) \in X \times Y \,:\, \varphi(x) + \varphi^c(y) = c(x, y)\}.$$

Proof. We will give an explicit formula for the function φ and prove that it is well defined and that it satisfies the properties that we want to impose.

Let us fix a point $(x_0, y_0) \in \Gamma$: for $x \in X$, set

$$\varphi(x) = \inf \{c(x, y_n) - c(x_n, y_n) + c(x_n, y_{n-1}) - c(x_{n-1}, y_{n-1}) + \cdots + \\ + c(x_1, y_0) - c(x_0, y_0) \,:\, n \in \mathbb{N},\ (x_i, y_i) \in \Gamma \text{ for all } i = 1, \ldots, n\}.$$

Since c is real valued and Γ is nonempty, φ never takes the value $+\infty$.

If we set, for $y \in Y$,

$$-\psi(y) = \inf \{-c(x_n, y) + c(x_n, y_{n-1}) - c(x_{n-1}, y_{n-1}) + \cdots + c(x_1, y_0) + \\ - c(x_0, y_0) \,:\, n \in \mathbb{N},\ (x_i, y_i) \in \Gamma \text{ for all } i = 1, \ldots, n,\ y_n = y\}.$$

Note that from the definition, we have $\psi(y) > -\infty$ if and only if $y \in (\pi_y)(\Gamma)$. Moreover, by construction we have $\varphi = \psi^{\bar{c}}$. This proves that φ is c-concave[9]. The fact that φ is not constantly $-\infty$ can be seen from $\varphi(x_0) \geq 0$: indeed, if we take $x = x_0$, then for any chain of points $(x_i, y_i) \in \Gamma$, we have

[9]Note that we did not spend words here on the measurability of φ: from its c-concavity, this is straightforward, as it is also continuous, but this is based on the assumption that c is uniformly continuous. Measurability and integrability in the general case are trickier.

$$\sum_{i=0}^{n} c(x_{i+1}, y_i) \geq \sum_{i=0}^{n} c(x_i, y_i),$$

where we consider $x_{n+1} = x_0$. This shows that the infimum in the definition of $\varphi(x_0)$ is nonnegative.

To prove $\varphi(x) + \varphi^c(y) = c(x, y)$ on Γ, it is enough to prove the inequality $\varphi(x) + \varphi^c(y) \geq c(x, y)$ on the same set, since by definition of c-transform, the opposite inequality is always true. Moreover, since $\varphi^c = \psi^{\bar{c}c}$ and $\psi^{\bar{c}c} \geq \psi$, it is enough to check $\varphi(x) + \psi(y) \geq c(x, y)$.

Suppose $(x, y) \in \Gamma$ and fix $\varepsilon > 0$. From $\varphi = \psi^{\bar{c}}$, one can find a point $\bar{y} \in \pi_y(\Gamma)$ such that $c(x, \bar{y}) - \psi(\bar{y}) < \varphi(x) + \varepsilon$. In particular, $\psi(\bar{y}) \neq \pm\infty$. From the definition of ψ, one has the inequality $-\psi(y) \leq -c(x, y) + c(x, \bar{y}) - \psi(\bar{y})$ (since every chain starting from $\bar{y}$ may be completed adding the point $(x, y) \in \Gamma$).

Putting together this information, one gets

$$-\psi(y) \leq -c(x, y) + c(x, \bar{y}) - \psi(\bar{y}) < -c(x, y) + \varphi(x) + \varepsilon,$$

which implies the inequality $c(x, y) \leq \varphi(x) + \psi(y)$ since ε is arbitrary. □

We can now prove the following.

Theorem 1.38. *If γ is an optimal transport plan for the cost c and c is continuous, then* $\mathrm{spt}(\gamma)$ *is a c-CM set.*

Proof. Suppose by contradiction that there exist k, σ, and $(x_1, y_1), \ldots, (x_k, y_k) \in \mathrm{spt}(\gamma)$ such that

$$\sum_{i=1}^{k} c(x_i, y_i) > \sum_{i=1}^{k} c(x_i, y_{\sigma(i)}).$$

Take now $\varepsilon < \frac{1}{2k}\left(\sum_{i=1}^{k} c(x_i, y_i) - \sum_{i=1}^{k} c(x_i, y_{\sigma(i)})\right)$. By continuity of c, there exists r such that for all $i = 1, \ldots, k$ and for all $(x, y) \in B(x_i, r) \times B(y_i, r)$, we have $c(x, y) > c(x_i, y_i) - \varepsilon$ and for all $(x, y) \in B(x_i, r) \times B(y_{\sigma(i)}, r)$, we have $c(x, y) < c(x_i, y_{\sigma(i)}) + \varepsilon$.

Now consider $V_i := B(x_i, r) \times B(y_i, r)$ and note that $\gamma(V_i) > 0$ for every i, because $(x_i, y_i) \in \mathrm{spt}(\gamma)$. Define the measures $\gamma_i := \gamma \llcorner V_i / \gamma(V_i)$ and $\mu_i := (\pi_x)_{\#}\gamma_i$, $\nu_i := (\pi_y)_{\#}\gamma_i$. Take $\varepsilon_0 < \frac{1}{k} \min_i \gamma(V_i)$.

For every i, build a measure $\tilde{\gamma}_i \in \Pi(\mu_i, \nu_{\sigma(i)})$ at will (for instance, take $\tilde{\gamma}_i = \mu_i \otimes \nu_{\sigma(i)}$).

Now define

$$\tilde{\gamma} := \gamma - \varepsilon_0 \sum_{i=1}^{k} \gamma_i + \varepsilon_0 \sum_{i=1}^{k} \tilde{\gamma}_i.$$

We want to find a contradiction by proving that $\tilde{\gamma}$ is a better competitor than γ in the transport problem, i.e., $\tilde{\gamma} \in \Pi(\mu, \nu)$ and $\int c\, d\tilde{\gamma} < \int c\, \mathrm{d}\gamma$.

First we check that $\tilde{\gamma}$ is a positive measure. It is sufficient to check that $\gamma - \varepsilon_0 \sum_{i=1}^k \gamma_i$ is positive, and, for that, the condition $\varepsilon_0 \gamma_i < \frac{1}{k}\gamma$ will be enough. This condition is satisfied since $\varepsilon_0 \gamma_i = (\varepsilon_0/\gamma(V_i))\gamma \llcorner V_i$ and $\varepsilon_0/\gamma(V_i) \le \frac{1}{k}$.

Now, let us check the condition on the marginals of $\tilde{\gamma}$. We have

$$(\pi_x)_\# \tilde{\gamma} = \mu - \varepsilon_0 \sum_{i=1}^k (\pi_x)_\# \gamma_i + \varepsilon_0 \sum_{i=1}^k (\pi_x)_\# \tilde{\gamma}_i = \mu - \varepsilon_0 \sum_{i=1}^k \mu_i + \varepsilon_0 \sum_{i=1}^k \mu_i = \mu,$$

$$(\pi_y)_\# \tilde{\gamma} = \nu - \varepsilon_0 \sum_{i=1}^k (\pi_y)_\# \gamma_i + \varepsilon_0 \sum_{i=1}^k (\pi_y)_\# \tilde{\gamma}_i = \nu - \varepsilon_0 \sum_{i=1}^k \nu_i + \varepsilon_0 \sum_{i=1}^k \nu_{\sigma(i)} = \nu.$$

Finally, let us estimate $\int c\, \mathrm{d}\gamma - \int c\, d\tilde{\gamma}$ and prove that it is positive, thus concluding the proof. We have

$$\begin{aligned}
\int c\, \mathrm{d}\gamma - \int c\, d\tilde{\gamma} &= \varepsilon_0 \sum_{i=1}^k \int c\, \mathrm{d}\gamma_i - \varepsilon_0 \sum_{i=1}^k \int c\, d\tilde{\gamma}_i \\
&\ge \varepsilon_0 \sum_{i=1}^k (c(x_i, y_i) - \varepsilon) - \varepsilon_0 \sum_{i=1}^k (c(x_i, y_{\sigma(i)}) + \varepsilon) \\
&= \varepsilon_0 \left(\sum_{i=1}^k c(x_i, y_i) - \sum_{i=1}^k c(x_i, y_{\sigma(i)}) - 2k\varepsilon \right) > 0,
\end{aligned}$$

where we used the fact that γ_i is concentrated on $B(x_i, r) \times B(y_i, r)$, $\tilde{\gamma}_i$ on $B(x_i, r) \times B(y_{\sigma(i)}, r)$ and that they have unit mass (due to the rescaling by $\gamma(V_i)$). □

From the previous theorem, together with Theorem 1.37, we get the duality result that we were waiting for[10] from Section 1.2.

Theorem 1.39. *Suppose that X and Y are Polish spaces and that $c : X \times Y \to \mathbb{R}$ is uniformly continuous and bounded. Then the problem* (DP) *admits a solution* (φ, φ^c) *and we have* max (DP) = min (KP).

Proof. First consider the minimization problem (KP). Since c is continuous, it admits a solution γ. Moreover, the support Γ of γ is c-cyclically monotone, thanks to Theorem 1.38. Then, since c is real valued, we can apply Theorem 1.37 and obtain the existence of a c-concave function φ such that

$$\Gamma \subset \{(x, y) \in \Omega \times \Omega \,:\, \varphi(x) + \varphi^c(y) = c(x, y)\}.$$

[10] We insist that this is just a possible path: we prove duality via cyclical monotonicity; instead, one could first prove duality (see Section 1.6.3) and then prove cyclical monotonicity through arguments as in Theorem 1.42.

From their c- (and $\bar{c}$-)concavity, φ and φ^c are continuous. Moreover, from $\varphi^c(y) = \inf_x c(x,y) - \varphi(x)$, we obtain an upper bound on φ^c (since c is supposed to be bounded), which turns into a lower bound on φ. Symmetrically, we can also obtain upper bounds on φ and lower on φ^c, which prove that φ and φ^c are both continuous and bounded.

Hence, we can use (φ, φ^c) as an admissible pair in (DP). Consider now

$$\int_X \varphi \, \mathrm{d}\mu + \int_Y \varphi^c \, \mathrm{d}\nu = \int_{X\times Y} (\varphi(x) + \varphi^c(y)) \, \mathrm{d}\gamma = \int_{X\times Y} c(x,y) \, \mathrm{d}\gamma,$$

where the last equality is due to the fact that γ is concentrated on Γ and there we have $\varphi(x) + \varphi^c(y) = c(x,y)$. The equality just before comes from the fact that the marginals of γ are μ and ν, respectively. This finally shows, using the optimality of γ,

$$(\mathrm{DP}) \geq \int \varphi \, \mathrm{d}\mu + \int \varphi^c \, \mathrm{d}\nu = \int c(x,y) \, \mathrm{d}\gamma = (\mathrm{KP})$$

and implies (DP)=(KP), since we already know (DP)$\leq$ (KP). As a by-product of this proof, the pair (φ, φ^c) turns out to be optimal for (DP). □

We also want to deal with some cases which are not covered by the previous theorem. First we want to give a short statement about the quadratic case in $\mathbb{R}^d$, in order to comply with Section 1.3.1. The difference compared to Theorem 1.39 lies in the fact that the cost $c(x,y) = \frac{1}{2}|x-y|^2$ is neither bounded nor uniformly continuous.

Theorem 1.40. *Let μ, ν be probabilities over $\mathbb{R}^d$ and $c(x,y) = \frac{1}{2}|x-y|^2$. Suppose $\int |x|^2 \, \mathrm{d}x$, $\int |y|^2 \, \mathrm{d}y < +\infty$. Consider the following variant of* (DP)*:*

$$(\mathrm{DP}-\mathrm{var}) \quad \sup \left\{ \int_{\mathbb{R}^d} \varphi \, \mathrm{d}\mu + \int_{\mathbb{R}^d} \psi \, \mathrm{d}\nu \ : \ \varphi \in L^1(\mu), \psi \in L^1(\nu), \ \varphi \oplus \psi \leq c \right\}.$$

Then (DP-var) *admits a solution (φ, ψ), and the functions $x \mapsto \frac{1}{2}|x|^2 - \varphi(x)$ and $y \mapsto \frac{1}{2}|y|^2 - \psi(y)$ are convex and conjugate to each other for the Legendre transform. Moreover, we have* $\max(\mathrm{DP}-\mathrm{var}) = \min(\mathrm{KP})$.

Proof. Applying the same considerations as in Theorem 1.39, we get the existence of a pair (φ, ψ), with φ c-concave and $\psi = \varphi^c$ such that $\varphi(x) + \varphi^c(y) = c(x,y)$ for every (x,y) in the support of an optimal transport plan γ. From Proposition 1.21, we deduce that $x \mapsto \frac{1}{2}|x|^2 - \varphi(x)$ and $y \mapsto \frac{1}{2}|y|^2 - \psi(y)$ are convex and conjugate to each other. In particular, $\frac{1}{2}|x|^2 - \varphi(x)$ is bounded from below by a linear function, and hence φ is bounded from above by a second-order polynomial. This proves $\varphi_+ \in L^1(\mu)$, from the assumption on μ. The same can be said on ψ, i.e., $\psi_+ \in L^1(\nu)$. We can now integrate $\varphi \oplus \psi$ w.r.t. γ, thus obtaining

$$\int_{\mathbb{R}^d} \varphi \, \mathrm{d}\mu + \int_{\mathbb{R}^d} \psi \, \mathrm{d}\nu = \int_{\mathbb{R}^d \times \mathbb{R}^d} \varphi \oplus \psi \, \mathrm{d}\gamma = \int_{\mathbb{R}^d \times \mathbb{R}^d} c \, \mathrm{d}\gamma \geq 0.$$

This proves $\int_{\mathbb{R}^d} \varphi \, \mathrm{d}\mu, \int_{\mathbb{R}^d} \psi \, \mathrm{d}\nu > -\infty$, and hence $\varphi \in L^1(\mu)$ and $\psi \in L^1(\nu)$.

The conclusion follows as before. □

Then, we also give some details about the case where c is not continuous but only l.s.c. We recall that this assumption is sufficient for the existence of an optimal transport plan.

The first result concerns the validity of the duality formula, i.e.,

$$\min (\mathrm{KP}) = \sup \left\{ \int_X \varphi \, \mathrm{d}\mu + \int_Y \psi \, \mathrm{d}\nu \ : \ \varphi \in C(X), \psi \in C(Y), \ \varphi \oplus \psi \leq c \right\}. \tag{1.6}$$

By now, we have established this equality when c is uniformly continuous and bounded, also proving that the dual problem admits a maximizing pair. We also know that an inequality is always true: the minimum on the left is always larger than the maximum on the right. More precisely, we are able to deal with the uniformly continuous case (since we want to guarantee continuity of c-concave functions of the form $\varphi(x) = \inf_y c(x, y) - \psi(y)$).

To deal with a l.s.c. cost c bounded from below, we will use the fact that there exists a sequence c_k of continuous functions (each one being k-Lipschitz) increasingly converging to c. We need the following lemma.

Lemma 1.41. *Suppose that c_k and c are l.s.c. and bounded from below and that c_k converges increasingly to c. Then*

$$\lim_{k \to \infty} \min \left\{ \int c_k \, \mathrm{d}\gamma \ : \ \gamma \in \Pi(\mu, \nu) \right\} = \min \left\{ \int c \, \mathrm{d}\gamma \ : \ \gamma \in \Pi(\mu, \nu) \right\}.$$

Proof. Due to the increasing limit condition, we have $c_k \leq c$, and hence the limit on the left (which exists by monotonicity) is obviously smaller than the quantity on the right. Now consider a sequence $\gamma_k \in \Pi(\mu, \nu)$, built by picking an optimizer for each cost c_k. Up to subsequences, due to the tightness of $\Pi(\mu, \nu)$, we can suppose $\gamma_k \rightharpoonup \overline{\gamma}$. Fix now an index j. Since for $k \geq j$ we have $c_k \geq c_j$, we have

$$\lim_k \min \left\{ \int c_k \, \mathrm{d}\gamma \ : \ \gamma \in \Pi(\mu, \nu) \right\} = \lim_k \int c_k \, \mathrm{d}\gamma_k \geq \liminf_k \int c_j \, \mathrm{d}\gamma_k.$$

By semi-continuity of the integral cost c_j, we have

$$\liminf_k \int c_j \, \mathrm{d}\gamma_k \geq \int c_j \, \mathrm{d}\overline{\gamma}.$$

Hence, we have obtained

$$\lim_k \min\left\{\int c_k \,\mathrm{d}\gamma,\ \gamma \in \Pi(\mu,\nu)\right\} \geq \int c_j \,\mathrm{d}\overline{\gamma}.$$

Since j is arbitrary and $\lim_j \int c_j \,\mathrm{d}\overline{\gamma} = \int c \,\mathrm{d}\overline{\gamma}$ by monotone convergence, we also have

$$\lim_k \min\left\{\int c_k \,\mathrm{d}\gamma,\ \gamma \in \Pi(\mu,\nu)\right\} \geq \int c \,\mathrm{d}\overline{\gamma} \geq \min\left\{\int c \,\mathrm{d}\gamma,\ \gamma \in \Pi(\mu,\nu)\right\}.$$

This concludes the proof. Note that it also gives, as a by-product, the optimality of $\overline{\gamma}$ for the limit cost c. □

We can now establish the validity of the duality formula for semi-continuous costs.

Theorem 1.42. *If X, Y are Polish spaces and $c : X \times Y \to \mathbb{R} \cup \{+\infty\}$ is l.s.c. and bounded from below, then the duality formula* $\min\,(\mathrm{KP}) = \sup\,(\mathrm{DP})$ *holds.*

Proof. Consider a sequence c_k of k-Lipschitz functions approaching c increasingly. Then the same duality formula holds for c_k, and hence, we have

$$\begin{aligned}\min\left\{\int c_k \,\mathrm{d}\gamma : \gamma \in \Pi(\mu,\nu)\right\} &= \max\left\{\int \varphi \,\mathrm{d}\mu + \int \psi \,\mathrm{d}\nu \ :\ \varphi \oplus \psi \leq c_k\right\}\\ &\leq \sup\left\{\int \varphi \,\mathrm{d}\mu + \int \psi \,\mathrm{d}\nu \ :\ \varphi \oplus \psi \leq c\right\},\end{aligned}$$

where the max and the sup are computed among C_b functions φ, ψ. The inequality is justified by the fact that $c_k \leq c$, and hence, every pair (φ, ψ) satisfying $\varphi(x) + \psi(y) \leq c_k(x,y)$ also satisfies $\varphi(x) + \psi(y) \leq c(x,y)$. The conclusion follows by letting $k \to +\infty$, using Lemma 1.41. Note that for the cost c, we cannot guarantee the existence of a maximizing pair (φ, ψ). □

The duality formula also allows us to prove the following c-cyclical monotonicity theorem.

Theorem 1.43. *If c is l.s.c. and γ is an optimal transport plan, then γ is concentrated on a c-CM set Γ (which will not be closed in general).*

Proof. Thanks to the previous theorem, the duality formula holds. This means that, if we take a sequence of maximizing pairs (φ_h, ψ_h) in the dual problem, we have

$$\int (\varphi_h(x) + \psi_h(y))\,\mathrm{d}\gamma = \int \varphi_h \,\mathrm{d}\mu + \int \psi_h \,\mathrm{d}\nu \to \int c \,\mathrm{d}\gamma,$$

since the value of $\int c \,\mathrm{d}\gamma$ is the minimum of the primal problem, which is also the maximum of the dual. Moreover, we have $c(x,y) - \varphi_h(x) + \psi_h(y) \geq 0$, which implies

that the functions $f_h := c(x,y) - \varphi_h(x) - \psi_h(y)$, defined on $X \times Y$, converge to 0 in $L^1(X \times Y, \gamma)$ (since they are positive and their integral tends to 0). As a consequence, up to a subsequence (not relabeled), they also converge pointwisely γ-a.e. to 0. Let $\Gamma \subset X \times Y$ be a set with $\gamma(\Gamma) = 1$ where the convergence holds.

Take any k, σ, and $(x_1, y_1), \dots, (x_k, y_k) \in \Gamma$: we have

$$\sum_{i=1}^{k} c(x_i, y_i) = \lim_h \sum_{i=1}^{k} \varphi_h(x_i) + \psi_h(y_i) = \lim_h \sum_{i=1}^{k} \varphi_h(x_i) + \psi_h(y_{\sigma(i)})$$
$$\leq \sum_{i=1}^{k} c(x_i, y_{\sigma(i)}),$$

which proves that this set is c-CM. □

Remark 1.44. The duality formula that we proved for l.s.c. cost c differs from that for continuous costs in that there is no existence for the dual problem (DP). Actually, if one restates the dual problem as in (DP-var), then one can produce an optimizer by first applying Theorem 1.43 to say that the optimal γ is concentrated on a c-CM set Γ and then build a potential φ through Theorem 1.37. This works under the assumption that c is real valued and does not depend on its continuity[11].

1.6.3 A direct proof of duality

In Section 1.6.2, we paid off the debt that we contracted with duality in Section 1.2: we proved, via c-cyclical monotonicity arguments, that min (KP) = sup (DP), with no duality gap, under some very general assumptions.

Here we want to give an alternative proof, independent of c-cyclical monotonicity and based on a simple convex analysis trick[12]. We will restrict to the compact case and only use the following fact: if H is convex and l.s.c., then $H^{**} = H$ (see Box 1.12 in Section 1.6.1).

Suppose that X and Y are compact metric spaces and $c : X \times Y \to \mathbb{R}$ is continuous. For every $p \in C(X \times Y)$, define

[11] Yet, note that the construction requires to use an optimal γ, which is guaranteed in the case where c is l.s.c., but could be the case in some special situations as well, such as when we can apply Lemma 1.8. Also, for nonuniformly continuous costs c, there are measurability issues on the function φ of Theorem 1.37.

[12] Even if not strictly necessary, we prefer to give it for completeness, and because it is different from what is presented in other texts, such as [292], where the main tool is a theorem from [90], Chapter 1, applied to well-chosen functions and space. The present proof, suggested by C. Jimenez, is essentially adapted from Section 4 in [70].

$$H(p) := -\max\left\{ \int_X \varphi \,\mathrm{d}\mu + \int_Y \psi \,\mathrm{d}\nu \;:\; \varphi(x)+\psi(y) \le c(x,y)-p(x,y) \right\},$$

which coincides with the opposite of the value of (DP) with cost $c-p$. The fact that this value is attained is proven in Proposition 1.11, which also provides bounds on the modulus of continuity of the optimal φ and ψ (the same as $c-p$) and on their oscillation (if we assume $\varphi(x_0)=0$, then both φ and ψ are bounded by a constant only depending on $\sup|c-p|$).

Lemma 1.45. *The function $H : C(X\times Y) \to \mathbb{R}$ is convex and l.s.c. for the uniform convergence on the compact space $X\times Y$.*

Proof. For convexity, take p_0 and p_1 with their optimal potentials (φ_0,ψ_0) and (φ_1,ψ_1). For $t\in[0,1]$, define $p_t = (1-t)p_0 + tp_1$, $\varphi_t = (1-t)\varphi_0 + t\varphi_1$, and $\psi_t = (1-t)\psi_0 + t\psi_1$. The pair (φ_t,ψ_t) is admissible in the max defining $-H(p_t)$. We have

$$H(p_t) \le -\left(\int_X \varphi_t \,\mathrm{d}\mu + \int_Y \psi_t \,\mathrm{d}\nu\right) = (1-t)H(p_0) + tH(p_1),$$

which shows convexity.

For semi-continuity, take $p_n \to p$. Extract a subsequence p_{n_k}, realizing the liminf of $H(p_n)$. From uniform convergence, the sequence p_{n_k} is equicontinuous and bounded (use the converse of the Ascoli-Arzelà theorem). Hence, the corresponding optimal potential $(\varphi_{n_k},\psi_{n_k})$ are also equicontinuous and bounded and we can assume $\varphi_{n_k}\to\varphi$, $\psi_{n_k}\to\psi$ uniformly, up to extracting again. Obviously $\varphi_{n_k}(x)+\psi_{n_k}(y) \le c(x,y)-p_{n_k}(x,y)$ implies $\varphi(x)+\psi(y)\le c(x,y)-p(x,y)$. Hence,

$$H(p) \le -\left(\int_X \varphi \,\mathrm{d}\mu + \int_Y \psi \,\mathrm{d}\nu\right) = \lim_k H(p_{n_k}) = \liminf_n H(p_n),$$

which proves lower semi-continuity. □

Let us now compute $H^* : \mathscr{M}(X\times Y) \to \mathbb{R}\cup\{+\infty\}$. For $\gamma\in\mathscr{M}(X\times Y)$, we have

$$H^*(\gamma) := \sup_p \int_{X\times Y} p\,\mathrm{d}\gamma + \sup\left\{ \int_X \varphi\,\mathrm{d}\mu + \int_Y \psi\,\mathrm{d}\nu : \varphi(x)+\psi(y)\le c(x,y)-p(x,y)\right\},$$

which can be written as a unique sup over p,φ,ψ. Note that, if $\gamma\notin\mathscr{M}_+(X\times Y)$, then there is $p\le 0$ such that $\int p_0 \mathrm{d}\gamma > 0$, and one can take $\varphi=0$, $\psi=0$, $p=c+np_0$, and, for $n\to\infty$, we get $H^*(\gamma)=+\infty$. On the contrary, if $\gamma\in\mathscr{M}_+(X\times Y)$, we should choose the largest possible p, i.e., $p(x,y)=c(x,y)-\varphi(x)-\psi(y)$. This gives

$$H^*(\gamma) = \sup_{\varphi,\psi} \int_{X\times Y} c(x,y)\,\mathrm{d}\gamma + \int_X \varphi\,\mathrm{d}\mu - \int_{X\times Y}\varphi(x)\,\mathrm{d}\gamma + \int_Y \psi\,\mathrm{d}\nu - \int_{X\times Y}\psi(x)\,\mathrm{d}\gamma.$$

But this is exactly the expression that we introduced in Section 1.2 (see Equation (1.3)) to rewrite the constraints in (KP). Indeed we have

$$H^*(\gamma) = \begin{cases} K(\gamma) & \text{if } \gamma \in \Pi(\mu,\nu), \\ +\infty & \text{if not.} \end{cases}$$

The duality theorem follows.

Theorem 1.46. *If X, Y are compact spaces and c continuous, then* $\min$ (KP) $=$ $\sup$ (DP) (proof by convex analysis).

Proof. We have $\max$ (DP) $= -H(0) = -H^{**}(0)$ (since H is convex and l.s.c., see Box 1.12). Moreover, $H^{**}(0) = \sup_\gamma \langle 0, \gamma\rangle - H^*(\gamma) = -\inf H^* = -\min$ (KP). □

1.6.4 *Sufficient conditions for optimality and stability*

The considerations of Sections 1.6.1 and 1.6.2 about c-transforms allow us to prove sufficient conditions for optimality in optimal transportation, at least when the cost satisfies the twist condition of Definition 1.16. As we did in Section 1.3, we will first give a general statement under compactness assumptions, and we will then give a different proof in the quadratic case, with no need of compactness.

Theorem 1.47. *Let $\Omega \subset \mathbb{R}^d$ be compact and c be a C^1 cost function satisfying the twist condition on $\Omega \times \Omega$. Suppose that $\mu \in \mathscr{P}(\Omega)$ and $\varphi \in c-conc(\Omega)$ are given, that φ is differentiable μ-a.e., and that $\mu(\partial\Omega) = 0$. Suppose that the map* T *satisfies $\nabla_x c(x, \mathrm{T}(x)) = \nabla\varphi(x)$. Then* T *is optimal for the transport cost c between the measures μ and $\nu := \mathrm{T}_\#\mu$.*

Proof. Consider the function φ, which may be written as $\varphi(x) = \inf_y c(x,y) - \psi(y)$ for a certain function $\psi : \Omega \to \mathbb{R}$. The function ψ may be supposed to be $\bar{c}$-concave and hence continuous (actually, we can take $\psi = \varphi^c$, from the considerations of the previous section). Fix now $x_0 \in \Omega$ such that $\nabla\varphi(x_0)$ exists and $x_0 \notin \partial\Omega$. By compactness and continuity, one can say that $\inf_y c(x,y) - \psi(y)$ is realized by a certain point y_0. This gives

$$\varphi(x) \le h(x - y_0) - \psi(y_0) \quad \text{for every } x \qquad \varphi(x_0) = h(x_0 - y_0) - \psi(y_0),$$

and hence $x \mapsto h(x - y_0) - \varphi(x)$ is minimal at $x = x_0$ (note that we defined y_0 by the optimality in y, but now we use the optimality in x). As a consequence, we get

$$\nabla\varphi(x_0) = \nabla_x c(x_0, y_0).$$

By assumption, $y \mapsto \nabla_x c(x_0, y)$ is injective: this implies $y_0 = \mathrm{T}(x_0)$. This proves $\varphi(x_0) + \psi(\mathrm{T}(x_0)) = c(x_0, \mathrm{T}(x_0))$, and this same equality is true for μ-a.e. x_0. If we

integrate with respect to μ, we get

$$\int \varphi \, \mathrm{d}\mu + \int \psi \, \mathrm{d}\nu = \int \varphi \, \mathrm{d}\mu + \int \psi \circ \mathrm{T} \, \mathrm{d}\mu = \int c(x, \mathrm{T}(x)) \mathrm{d}\mu(x),$$

which proves the optimality of T since the last integral equals the cost of T in the problem of Monge, and we have $\int \varphi \, \mathrm{d}\mu + \int \psi \, \mathrm{d}\nu \leq (\mathrm{DP}) \leq (\mathrm{KP})$, since (φ, ψ) is admissible in (DP). □

As we promised, here is a different proof for the case $c(x, y) = \frac{1}{2}|x - y|^2$.

Theorem 1.48. *Suppose that $\mu \in \mathscr{P}(\mathbb{R}^d)$ is such that $\int |x|^2 \, \mathrm{d}\mu(x) < \infty$ and that $u : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is convex and differentiable μ-a.e. Set $\mathrm{T} = \nabla u$ and suppose $\int |\mathrm{T}(x)|^2 \mathrm{d}\mu(x) < +\infty$. Then* T *is optimal for the transport cost $c(x, y) := \frac{1}{2}|x - y|^2$ between the measures μ and $\nu := \mathrm{T}_{\#}\mu$.*

Proof. First note that, for a convex function u, we have the following properties:

$$u(x) + u^*(y) \geq x \cdot y \ \text{ for all } x, y \in \mathbb{R}^d, \quad u(x) + u^*(y) = x \cdot y \ \text{ if } y = \nabla u(x).$$

Now, consider any transport plan $\gamma \in \Pi(\mu, \nu)$ and write

$$\begin{aligned} \int_{\mathbb{R}^d \times \mathbb{R}^d} (x \cdot y) \, \mathrm{d}\gamma(x, y) &\leq \int_{\mathbb{R}^d \times \mathbb{R}^d} \big(u(x) + u^*(y)\big) \, \mathrm{d}\gamma(x, y) \\ &= \int_{\mathbb{R}^d} u(x) \, \mathrm{d}\mu(x) + \int_{\mathbb{R}^d} u^*(\mathrm{T}(x)) \, \mathrm{d}\mu(x) = \int_{\mathbb{R}^d} (x \cdot \mathrm{T}(x)) \, \mathrm{d}\mu(x). \end{aligned}$$

This proves $\int (x \cdot y) \, \mathrm{d}\gamma \leq \int (x \cdot y) \, \mathrm{d}\gamma_{\mathrm{T}}$. Subtracting these quantities from the integral, which only depends on μ and ν, $\int \frac{1}{2}(|x|^2 + |y|^2) \, \mathrm{d}\gamma = \int \frac{1}{2}(|x|^2 + |y|^2) \, \mathrm{d}\gamma_{\mathrm{T}}$, we get

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{1}{2}|x - y|^2 \, \mathrm{d}\gamma \geq \int_{\mathbb{R}^d \times \mathbb{R}^d} \frac{1}{2}|x - y|^2 \mathrm{d}\gamma_{\mathrm{T}},$$

which proves the claim. □

Let us spend some words on the two different proofs that we just gave: we needed to prove some optimality properties of the point $\mathrm{T}(x)$ among possible points y. In the first case, we proved that an optimal point y_0 existed, and we deduced from the first-order optimality conditions $y_0 = \mathrm{T}(x_0)$. In the second case, we used the characterization of minimizers for convex functions: whenever a gradient vanishes, we are at a minimizer. This is the idea behind the property $u(x) + u^*(y) = x \cdot y$ if $y = \nabla u(x)$.

We now pass to a more general criterion that can be used for almost arbitrary costs. The main idea is that we proved that optimal plans are necessarily concentrated on c-CM sets, but the converse is also true: a plan which is concentrated on a c-CM set is optimal, at least in the most reasonable cases. We will prove this fact in the easiest case, i.e., when c is uniformly continuous and bounded. Actually, we

will see in the next proof that the main ingredient is Theorem (1.37), which only requires finiteness of the cost, but continuity and compactness are needed to avoid integrability issues (in this case, all the functions are measurable and bounded). Moreover, we will apply this sufficient criterion to another interesting problem, i.e., stability of the optimal plans, which will require these assumptions for other reasons. The fact that this converse implication stays true when these assumptions are withdrawn (and in particular for infinite costs) is a delicate matter, and we refer to [10] for a counterexample when the cost takes the value $+\infty$ and to [30, 60, 254] for interesting positive proofs.

Theorem 1.49. *Suppose that $\gamma \in \mathscr{P}(X \times Y)$ is given, that X and Y are Polish spaces, that $c : X \times Y \to \mathbb{R}$ is uniformly continuous and bounded, and that* $\mathrm{spt}(\gamma)$ *is c-CM. Then γ is an optimal transport plan between its marginals $\mu = (\pi_x)_\# \gamma$ and $\nu = (\pi_y)_\# \gamma$ for the cost c.*

Proof. Theorem (1.37) gives the existence of a c-concave function φ such that $\mathrm{spt}(\gamma)$ is contained in the set $\{(x, y) : \varphi(x) + \varphi^c(y) = c(x, y)\}$. Both φ and φ^c are continuous thanks to the continuity of c, and hence bounded on X and Y, respectively.

Thanks to what we know about duality, we have

$$(\mathrm{KP}) \leq \int c(x, y)\, \mathrm{d}\gamma = \int (\varphi(x) + \varphi^c(y))\, \mathrm{d}\gamma = \int \varphi\, \mathrm{d}\mu + \int \varphi^c\, \mathrm{d}\nu \leq (\mathrm{DP}) = (\mathrm{KP})$$

which shows that γ is optimal (and that (φ, φ^c) solves the dual problem). □

We are now able to prove the following stability result that we only state in the compact case.

Theorem 1.50. *Suppose that X and Y are compact metric spaces and that $c : X \times Y \to \mathbb{R}$ is continuous. Suppose that $\gamma_n \in \mathscr{P}(X \times Y)$ is a sequence of transport plan which are optimal for the cost c between their own marginals $\mu_n := (\pi_x)_\# \gamma_n$ and $\nu_n := (\pi_y)_\# \gamma_n$, and suppose $\gamma_n \rightharpoonup \gamma$; then $\mu_n \rightharpoonup \mu := (\pi_x)_\# \gamma$, $\nu_n \rightharpoonup \nu := (\pi_y)_\# \gamma$ and γ is optimal in the transport between μ and ν.*

Proof. Set $\Gamma_n := \mathrm{spt}(\gamma_n)$. Up to subsequences, we can assume $\Gamma_n \to \Gamma$ in the Hausdorff topology (see below). Each support Γ_n is a c-CM set (Theorem (1.38)) and the Hausdorff limit of c-CM sets is also c-CM. Indeed, if one fixes $(x_1, y_1), \dots, (x_k, y_k) \in \Gamma$, there are points $(x_1^n, y_1^n), \dots, (x_k^n, y_k^n) \in \Gamma_n$ such that, for each $i = 1, \dots, k$, we have $(x_i^n, y_i^n) \to (x_i, y_i)$. The cyclical monotonicity of Γ_n gives $\sum_{i=1}^k c(x_i^n, y_i^n) \leq \sum_{i=1}^k c(x_i^n, y_{\sigma(i)}^n)$, which implies, taking the limit $n \to \infty$

$$\sum_{i=1}^k c(x_i, y_i) \leq \sum_{i=1}^k c(x_i, y_{\sigma(i)}).$$

This proves that Γ is c-CM. Moreover, we know from the convergence $\gamma_n \rightharpoonup \gamma$ together with $\Gamma_n \to \Gamma$ that $\mathrm{spt}(\gamma) \subset \Gamma$. This shows that $\mathrm{spt}(\gamma)$ is c-CM and implies the optimality of γ. □

We finish with an easy but useful consequence of Theorem 1.50. To fix the notations, for a given cost $c : X \times Y \to \mathbb{R}$ and $\mu \in \mathscr{P}(X)$, $\nu \in \mathscr{P}(Y)$, let us define

$$\mathscr{T}_c(\mu,\nu) := \min\left\{ \int_{X\times Y} c \,\mathrm{d}\gamma \,:\, \gamma \in \Pi(\mu,\nu)\right\}.$$

Theorem 1.51. *Suppose that X and Y are compact metric spaces and that $c : X \times Y \to \mathbb{R}$ is continuous. Suppose that $\mu_n \in \mathscr{P}(X)$ and $\nu_n \in \mathscr{P}(Y)$ are two sequences of probability measures, with $\mu_n \rightharpoonup \mu$ and $\nu_n \rightharpoonup \nu$. Then we have $\mathscr{T}_c(\mu_n,\nu_n) \to \mathscr{T}_c(\mu,\nu)$.*

Proof. Let γ_n be an optimal transport plan from μ_n to ν_n for the cost c. Up to subsequences, we can assume $\gamma_n \rightharpoonup \gamma$. Theorem 1.50 provides the optimality of γ. This means that we have (along this subsequence, but since it is arbitrary)

$$\mathscr{T}_c(\mu_n,\nu_n) = \int_{X\times Y} c \,\mathrm{d}\gamma_n \to \int_{X\times Y} c \,\mathrm{d}\gamma = \mathscr{T}_c(\mu,\nu)$$

which proves the claim. □

Box 1.13. ***Memo**: Hausdorff convergence*

Definition. In a compact metric space X, we define the Hausdorff distance on pair of compact subsets of X by setting

$$d_H(A,B) := \max\left\{\max\{d(x,A) \,:\, x\in B\}, \max\{d(x,B) \,:\, x\in A\}\right\}.$$

Properties. We have the following equivalent definition:

1) $d_H(A,B) = \max\{|d(x,A) - d(x,B)| \,:\, x \in X\}$
2) $d_H(A,B) = \inf\{\varepsilon > 0 \,:\, A \subset B_\varepsilon,\, B \subset A_\varepsilon\}$, where A_ε and B_ε stand for the ε-neighborhood of A and B, respectively.

Theorem (Blaschke). d_H is a distance: it is positive and symmetric, it only vanishes if the two sets coincide, and it satisfies triangle inequality. With this distance, the set of compact subsets of X becomes a compact metric space itself.

We refer to [12] for a detailed discussion (with proofs) of this topic. Here we only prove a simple fact that we need.

Proposition. If $d_H(A_n,A) \to 0$ and μ_n is a sequence of positive measures such that $\mathrm{spt}(\mu_n) \subset A_n$ with $\mu_n \rightharpoonup \mu$, then $\mathrm{spt}(\mu) \subset A$.

Sketch of proof. For each n, we have $\int d(x,A_n)\,\mathrm{d}\mu_n = 0$, since μ_n is supported on A_n. Since $d(x,A_n) \to d(x,A)$ uniformly and $\mu_n \rightharpoonup \mu$, thanks to the duality between uniform convergence and weak converge of measures, we get $\int d(x,A_n)\,\mathrm{d}\mu_n \to \int d(x,A)\,\mathrm{d}\mu$. This implies $\int d(x,A)\,\mathrm{d}\mu = 0$, and hence μ is concentrated on A.

Let us also conclude with a result on the stability of the Kantorovich potentials.

Theorem 1.52. *Suppose that X and Y are compact metric spaces and that $c : X \times Y \to \mathbb{R}$ is continuous. Suppose that $\mu_n \in \mathscr{P}(X)$ and $\nu_n \in \mathscr{P}(Y)$ are two sequences of probability measures, with $\mu_n \rightharpoonup \mu$ and $\nu_n \rightharpoonup \nu$. Let (φ_n, ψ_n) be, for each n, a pair of c-concave Kantorovich potentials for the cost c in the transport from μ_n to ν_n. Then, up to subsequences, we have $\varphi_n \to \varphi$, $\psi_n \to \psi$, where the convergence is uniform and (φ, ψ) is a pair of Kantorovich potentials for μ and ν.*

Proof. We know that c-concave functions have the same modulus of continuity as c and hence, up to translating by a constant, we can apply the Ascoli-Arzelà theorem. We have $\varphi_n \to \tilde{\varphi}$ and $\psi_n \to \tilde{\psi}$, the convergences being uniform. From $\varphi_n(x) + \psi_n(y) \leq c(x, y)$, we deduce, as $n \to \infty$, the inequality $\tilde{\varphi}(x) + \tilde{\psi}(y) \leq c(x, y)$. Moreover, we have

$$\mathscr{T}_c(\mu_n, \nu_n) = \int \varphi_n \, \mathrm{d}\mu_n + \int \psi_n \, \mathrm{d}\nu_n \to \int \tilde{\varphi} \, \mathrm{d}\mu + \int \tilde{\psi} \, \mathrm{d}\nu.$$

Yet, we also have (Theorem 1.50) $\mathscr{T}_c(\mu_n, \nu_n) \to \mathscr{T}_c(\mu, \nu)$. We infer that the pair $(\tilde{\varphi}, \tilde{\psi})$ is admissible in the dual problem and realizes the maximal value $\int \tilde{\varphi} \, \mathrm{d}\mu + \int \tilde{\psi} \, \mathrm{d}\nu = \mathscr{T}_c(\mu, \nu)$. Hence, they are Kantorovich potentials.

The convergence is obviously true on the full sequence (without passing to a subsequence) in case of uniqueness of the Kantorovich potentials at the limit. □

1.7 Discussion

This section presents alternative point of views (probabilistic language, polar factorization), connections with economics and finance, generalizations and a brief presentation of the regularity issues on the optimal transport map, and the optimal potentials.

1.7.1 Probabilistic interpretation

We completely skipped up to now any interpretations involving probabilities, which are actually natural in this setting. By changing for a while our language, let us describe the situation in the following way.

Two probability measures μ and ν on some spaces (often, $\mathbb{R}^d$) are given and can be interpreted as the laws of two random variables. Yet, we do not prescribe the joint law (which corresponds to γ) of these two random variables and we consider the optimization problem

$$\min\{\mathbb{E}[c(X, Y)] \ : \ X \sim \mu, \ Y \sim \nu\}$$

where $\mathbb{E}$ denotes the expected value (according to a probability $\mathbb{P}$ on a probability space Ω, which is not relevant for the minimization, but could be considered either "large enough" or to be part of the optimization). This expected value obviously depends on the joint law of (X, Y), which is the main unknown.

The particular case of real (or vector) valued r.v. with $c(X, Y) = |X - Y|^p$ reads

$$\min\{||X - Y||_{L^p} \ : \ X \sim \mu, \ Y \sim \nu\}.$$

More interesting is the case $p = 2$, where the problem can be expressed in terms of covariance. Indeed, let us set $x_0 = \mathbb{E}[X] = \int x \, \mathrm{d}\mu$ and $y_0 = \mathbb{E}[Y] = \int y \, \mathrm{d}\nu$, these two values being the mean values of X and Y. We have

$$\begin{aligned}\mathbb{E}[|X - Y|^2] &= \mathbb{E}[|(X - x_0) - (Y - y_0) + (x_0 - y_0)|^2] \\ &= \mathbb{E}[|X - x_0|^2] + \mathbb{E}[|Y - y_0|^2] + |x_0 - y_0|^2 \\ &\quad +2\mathbb{E}[X - x_0] \cdot (x_0 - y_0) - 2\mathbb{E}[Y - y_0] \cdot (x_0 - y_0) \\ &\quad -2\mathbb{E}[(X - x_0) \cdot (Y - y_0)].\end{aligned}$$

In this expression, the three first terms only depend on the laws of X and Y separately (the first being the variance of X, the second the variance of Y, and the third the squared distance between the two mean values), and the next two terms vanish (since the mean value of $X - x_0$ is 0 and so for the mean value of $Y - y_0$). The problem is hence reduced to the maximization of $\mathbb{E}[(X - x_0) \cdot (Y - y_0)]$. This means that we need to find the joint law which guarantees maximal covariance (i.e., somehow maximal dependence) of two r.v. with given laws. In the case of real-valued r.v. (see next chapter), the answer will be that the optimal coupling is obtained in the case where X and Y are completely dependent and one is an increasing function of the other. The multidimensional case obviously replaces this increasing behavior with other monotone behaviors.

1.7.2 Polar factorization

A classical result in linear algebra states that every matrix $A \in \mathrm{M}^{N \times N}$ can be decomposed as a product $A = S \cdot U$, where S is symmetric and positive semidefinite, and U is a unitary matrix, i.e., $U \cdot U^t = I$. The decomposition is unique if A is nonsingular (otherwise, U is not uniquely defined), and in such a case, S is positive definite. Also, one can see that the matrix U of this decomposition is also a solution (the unique one if A is nonsingular) of

$$\max\{A : R \ : \ R \cdot R^t = I\},$$

where $A : R$ stands for the scalar product between matrices, defined as $A : R := \mathrm{Tr}(A \cdot R^t)$.

Indeed, one can write

$$A : R = (S \cdot U) : R = \mathrm{Tr}(S \cdot U \cdot R^t) = \sum_{i,j} S^{ii} U^{ij} R^{ij},$$

where the coordinates are chosen so that S is diagonal (and hence $S^{ii} \geq 0$). Let us observe that $U \cdot U^t = I$ imposes that the vectors $u(i)$ with components $u(i)^j := U^{ij}$ are unit vectors (let us also define the unit vectors $r(i)$ in the same way from the matrix R). Thus

$$\begin{aligned} A : R = \sum_{i,j} S^{ii} U^{ij} R^{ij} &= \sum_i S^{ii} u(i) \cdot r(i) \\ &\leq \sum_i S^{ii} = \mathrm{Tr}(S) = \mathrm{Tr}(S \cdot U \cdot U^t) = \mathrm{Tr}(A \cdot U^t) = A : U. \end{aligned}$$

Also note that maximizing $A : R$ is the same as minimizing $||A - R||^2$ (where the norm of a matrix is defined as the square root of its scalar product with itself). Indeed, we have $||A - R||^2 = ||A||^2 - 2A : R + ||R||^2$, and both $||A||^2$ and $||R||^2 = \mathrm{Tr}(R \cdot R^t) = \mathrm{Tr}(\mathrm{I}) = N$ are constants.

Analogously, in his first works about the quadratic optimal transport, Y. Brenier noted that Monge-Kantorovich's theory allowed us to provide a similar decomposition for vector fields instead of linear maps (see [82, 84] and also [174]).

The statement is the following:

Theorem 1.53. *Given a vector map $\xi : \Omega \to \mathbb{R}^d$ with $\Omega \subset \mathbb{R}^d$, consider the rescaled Lebesgue measure $\mathscr{L}_\Omega$ on Ω and suppose that $\xi_\# \mathscr{L}_\Omega$ is absolutely continuous; then, one can find a convex function $u : \Omega \to \mathbb{R}$ and a measure-preserving map $s : \Omega \to \Omega$ (i.e., such that $s_\# \mathscr{L}_\Omega = \mathscr{L}_\Omega$) such that $\xi = (\nabla u) \circ s$. Moreover, both s and ∇u are uniquely defined a.e., and s solves*

$$\max \left\{ \int \xi(x) \cdot r(x) \, \mathrm{d}x \ : \ r_\# \mathscr{L}_\Omega = \mathscr{L}_\Omega \right\}.$$

Note that the statement concerning nonsingular matrices exactly corresponds to this one when one takes $\Omega = B(0, 1)$, since the assumption on the image measure corresponds to the non-degeneracy of the linear map $x \mapsto Ax$, the matrix S is the gradient of the convex function $x \mapsto \frac{1}{2}(Sx) \cdot x$, and the unitary matrix U is measure preserving.

We give a proof of this statement; see also [84].

Proof. Consider $\mu = \xi_\# \mathscr{L}_\Omega$ and take the optimal transport for the quadratic cost between μ and $\mathscr{L}_\Omega$. This transport is the gradient of a convex function that we will call u^*. Set $s := (\nabla u^*) \circ \xi$, and note that $s_\# \mathscr{L}_\Omega = (\nabla u^*)_\# (\xi_\# \mathscr{L}_\Omega) = (\nabla u^*)_\# \mu = \mathscr{L}_\Omega$; hence s is measure preserving. Then, take the Legendre transform u of u^* and note $(\nabla u) \circ s = (\nabla u) \circ (\nabla u^*) \circ \xi = \xi$ (indeed, the gradients of both u and u^*

are defined a.e. and we always have $(\nabla u) \circ (\nabla u^*) = \mathrm{id}$). In this way, we have the desired decomposition.

The uniqueness of ∇u comes from the fact that $\xi = (\nabla u) \circ s$ with s measure preserving implies $(\nabla u)_{\#}\mathscr{L}_\Omega = \mu$, and there is only one gradient of a convex function transporting a given measure (here $\mathscr{L}_\Omega$) to another given measure (here, μ), and it is the optimal transport between them.

Once ∇u is unique, the uniqueness of s is obtained by composing with ∇u^*.

Concerning the optimality of s, use the optimality of ∇u. The maximization of the scalar product instead of the minimization of the quadratic cost, we get

$$\int (\nabla u(x) \cdot x)\, \mathrm{d}x \geq \int_{\Omega \times \mathbb{R}^d} (y \cdot x)\, \mathrm{d}\gamma(x, y)$$

for every $\gamma \in \Pi(\mathscr{L}_\Omega, \mu)$. If we consider a measure-preserving map $r : \Omega \to \Omega$ and build $\gamma = (r, \xi)_{\#}\mathscr{L}_\Omega$, we get

$$\begin{aligned} \int (\xi(x) \cdot r(x))\, \mathrm{d}x &= \int (y \cdot x)\, \mathrm{d}\gamma(x, y) \\ &\leq \int (\nabla u(x) \cdot x)\, \mathrm{d}x = \int (\nabla u(s(x)) \cdot s(x))\, \mathrm{d}x = \int (\xi(x) \cdot s(x))\, \mathrm{d}x, \end{aligned}$$

which provides the optimality of s. □

1.7.3 Matching problems and economic interpretations

Optimal transport problems have several "abstract" economic interpretations, where the role of transport plan γ is that of matchings between different actors of an economy, and the function $c(x, y)$ does not represent anymore a cost for moving from x to y but rather a compatibility of the two objects x and y, or the opposite of the "utility" for x and y to be coupled.[13] Consider that, in economics, it is typical to maximize utility, instead of minimizing costs.

A first easy and well-known example is that of the maximization of the productivity. A company has a certain number of employees of different types (let us use the variable x for the types and the measure μ for the distribution of these types, i.e., the quantity of employees for each type) and some tasks to attribute (we use y for the different kinds of tasks and ν for the distribution of different tasks), and if the productivity $p(x, y)$ of the employees of type x when they work out the task y is known, then the goal of the company is to solve

[13] For a comprehensive introduction to some applications of otpimal transport to economics, see also [228].

$$\max\left\{\int p(x,y)\,\mathrm{d}\gamma \;:\; \gamma\in\Pi(\mu,\nu)\right\}.$$

This problem is easy to understand, but we want to analyze some other ones where also the Kantorovich potentials also play a role.

The first example is that of stable marriages. Given two populations W of women and M of men, represented by a measure μ on W and a measure ν on M, we denote by $u_w(x,y)$ the interest of Ms x for Mr y and $u_m(x,y)$ that of Mr y for Ms x. The problem is that of finding a stable set of marriages, i.e., a measure γ on the set $W\times M$ giving how many women of type x get married with how many men of type y. This γ must obviously belong to $\Pi(\mu,\nu)$, and we want to select such a γ so that it is stable, i.e., no new couple (x,y) will decide to divorce (each one from his/her current partner) to go together. The case that we are interested in is that of *transferable utility*: it means that we suppose that, once x and y get married, they decide how to split their total utility $u_w(x,y)+u_m(x,y)$. They divide into a quantity $\varphi(x)$, which is the utility surplus for Ms x (now Mrs y, but let us call her with her maiden name) and $\psi(y)$ for Mr y. Note that in this case, only the sum $U(x,y):=u_w(x,y)+u_m(x,y)$ really plays a role. A stable marriage is a triple (γ,φ,ψ) such that $U(x,y)=\varphi(x)+\psi(y)$ γ-a.e. (this represents the fact that φ and ψ are a splitting of the total utility) and $U(x,y)\le\varphi(x)+\psi(y)$ for all (x,y) (this represents the fact that no pair (x,y) will be interested in quitting the status quo). It is easy to see that this corresponds to an optimal γ with its potentials. We mention [125] for an interesting discussion of this problem. On the other hand, the case of nontransferable utility is much more involved and does not correspond to a transport problem.

If, in the previous example, the potentials represented split utilities, we can think of other cases where they are really prices. Indeed, consider that the variable x represents the goods that are available on the market and that μ is the distribution (how many of each type) of these goods, which we consider as fixed. The variable y plays the role of the type of consumers and ν is their distribution. Let $u(x,y)$ be the utility of the consumer y when he/she buys the good x. The goal is to determine the prices of the goods and who buys what. Suppose for a while that the price $\varphi(x)$ of each good is known; then, each consumer will choose what to buy by solving $\max_x u(x,y)-\varphi(x)$. Let us denote (by abuse of notation, since usually we used minimization instead of maximization) this quantity as $\varphi^u(y)$. We describe the choices of the consumers through a measure $\gamma\in\Pi(\mu,\nu)$ where $\gamma(A\times B)$ stands for the number of consumers of type $y\in B$ buying a good $x\in A$. The constraint $\gamma\in\Pi(\mu,\nu)$ stands for the constraints given by the supply and the demand on the market (we say that the market is cleared). Another natural condition to impose is the fact that each consumer only buys goods which are optimal for him, i.e., that γ is concentrated over the set of pairs (x,y) with $\varphi^u(y)=u(x,y)-\varphi(x)$, i.e., such that x is an optimal choice, given φ, for y.

This means that we are led to the following problem:

$$\text{find } (\gamma,\varphi) \text{ such that } \gamma\in\Pi(\mu,\nu) \text{ and } \varphi(x)+\varphi^u(y)=u(x,y)\ \gamma-a.e.,$$

and we can see again an optimal transport problem for the maximization of the cost u. Indeed, the pairs (γ, φ) can be characterized as the solutions of the Kantorovich problem

$$\max\left\{\int u\ \mathrm{d}\gamma\ :\ \gamma\in\Pi(\mu,\nu)\right\}$$

and of the dual problem

$$\min\left\{\int \varphi\ \mathrm{d}\mu+\int \varphi^u\ \mathrm{d}\nu\right\}.$$

By the way, this optimal transport interpretation shows that a simple equilibrium condition (clearing the market and only using rational choices of the consumers) implies good news, i.e., the fact that the general satisfaction $\int u\ \mathrm{d}\gamma$ is maximized.

One could remark that the values of the functions φ and φ^u are only defined up to additive constants, which means that this procedure only selects the relative differences of prices between goods and not the complete price chart. Yet, natural assumptions guarantee uniqueness up to additive constants of the solution of the dual problem (it is the case, for instance, when x and y belong to Euclidean spaces, $u(x, y)$ is differentiable and one of the two measures has strictly positive density a.e.). In this case, as soon as the price of one special good $\bar{x}$ is known, then everything is uniquely determined. A typical example can be obtained if the "empty" good is included in the market. Let us denote by x_0 a special good which corresponds to "not buying anything at all." We can normalize utilities by setting $u(y, x_0) = 0$ for every y, but this is not really important. However, we can assume that no seller will charge a price different from 0 when selling this empty good. This fixes $\varphi(x_0) = 0$ and allows for computing the other prices. If the measure μ is such that $\mu(\{x_0\}) > 0$, then this means that there are not enough goods on the market so as to satisfy all the demands and that some of the consumers will stay "out of the market," i.e., they will buy the empty good.

Another interesting problem is obtained when the measure μ on the sets of goods is not fixed, and we consider a class of goods which are sold by a unique company, acting as a monopolist on this market. We suppose that the set X of feasible goods (those that the company can technically produce) is known and that it includes the empty good x_0. The goal of the company is, hence, to select at the same time a measure μ (how much production for each type of goods) and a price list $\varphi : X \to \mathbb{R}$, satisfying $\varphi(x_0) = 0$, so as to maximize its profit, assuming that each consumer will buy according to the maximization of $x \mapsto u(x - y) - \varphi(x)$.

This problem may be expressed in two ways: indeed, the equilibrium condition of the previous problem (when μ was fixed) induces a relationship between μ and φ, and this allows us to consider either one or the other as the variable of the problem.

Probably the easiest way to describe the problem is to think that the company chooses the price list φ, that every consumer y selects its optimal good $X(y) \in \operatorname{argmin}_x u(x, y) - \varphi(x)$, and that the total income (supposing zero production costs)

of the company is $\int \varphi(X(y))\,\mathrm{d}\nu(y) = \int \varphi\,\mathrm{d}\mu$ for $\mu = X_\#\nu$. The measure μ is also the measure of the real production of goods that the company will implement. In other words, it will adapt its production to the choice of the consumers. Anyway φ has then to be chosen so as to maximize $\int \varphi(X(y))\,\mathrm{d}\nu(y)$, taking into account that the map X depends on φ. The problem can also be adapted so as to take into account production costs, in the form of a function $c : X \to \mathbb{R}$, and the maximization becomes that of $\int(\varphi - c)(X(y))\,\mathrm{d}\nu(y) = \int(\varphi - c)\,\mathrm{d}\mu$. This formulation should be revisited in case the optimal point $X(y)$ is not uniquely defined (since different optimizers for the consumer could lead to very different incomes for the company). We refer to [178], where different relaxed formulations (defined as "optimistic" and "pessimistic") are proposed. Anyway, we will see in a while that in a very reasonable case, there is no ambiguity.

The other approach is complementary to this one: we can give the company the right to select the production measure μ; then the market is cleared thanks to the previous considerations, thus determining a measure $\gamma \in \Pi(\mu, \nu)$ and a potential φ. The goal of the company is to maximize $\int(\varphi - c)\,\mathrm{d}\mu$, where μ is unknown and φ depends on μ. Note that this is an optimization problem in the space of measures which involves the Kantorovich potential φ in the transport from a given measure ν to μ.

The case which is studied the most is the case where both x and y belong to convex sets $X, Y \subset \mathbb{R}^d$ and the function $u(x, y) = x \cdot y$ is a scalar product. This is natural if one thinks that goods can be described through a set of parameters (think of cars: we can use number of seats, size of the luggage van, fuel consumption, maximal speed, etc.) and that consumers are described through the importance that they give to each one of these parameters. Customers of type $y = (y_1, \dots, y_n)$ are ready to spend y_i extra unit of money for every unit increase in the ith parameter of the good. This means that the values of the coefficients y_i give the relative value of each feature of the good compared to money (in the case of cars, we can expect that they will be higher for richer people, who care less about money, or for people who like very much cars, or need them absolutely). The empty good can be chosen as the point $x_0 = 0 \in \mathbb{R}^d$.

In this case, it is easy to check that φ^u is simply the Legendre transform φ^* of convex analysis. Standard considerations in convex analysis imply that if φ^* is differentiable at y, then the optimal choice x in $\max_x x \cdot y - \varphi(x)$ is exactly $x = \nabla\varphi^*(y)$ and gives uniqueness a.e. in case ν is absolutely continuous (due to differentiability properties of convex functions). Then, the map $\nabla\varphi^*(y)$ sends ν onto μ and, conversely, $\nabla\varphi$ sends μ onto ν. Hence, the optimization problem reads

$$\max\left\{\int \varphi_\mu\,\mathrm{d}\mu \;:\; \mu \in \mathscr{P}(X)\right\},$$

where, for every μ, φ_μ is the unique convex function with $(\nabla\varphi)_\#\mu = \nu$ and $\varphi(x_0) = 0$. It can also be rephrased as an optimization problem in the class of convex functions, if one takes $\psi = \varphi^*$ as unknown, since in this case, one should maximize $\int \psi^*\, d(\nabla\psi)_\#\nu = \int(\psi^*(\nabla\psi(y)) - c(\nabla\psi(y)))\,\mathrm{d}\nu$, but we know

$\psi^*(\nabla\psi(y)) + \psi(y) = y \cdot \nabla\psi(y)$, which turns the problem into

$$\max\left\{ \int (y \cdot \nabla\psi(y) - \psi(y) - c(\nabla\psi(y)))\, \mathrm{d}\nu(y) \ : \ \min\psi = 0, \psi \ \text{convex} \right\},$$

the constraint $\min\psi = 0$ coming from $0 = \psi^*(0) = \sup_y 0 \cdot y - \psi(y)$. This is a standard calculus of variations problem but considered in the restricted class of convex functions. It has been deeply studied in [262], and it is usually known as *principal-agent problem*.

For more general couplings $u(x, y)$, [168] studied fine features on the problem and in particular which assumptions on the utility u provide a convex optimization problem. Interestingly, the condition to guarantee this convexity is connected to the MTW assumption that we will mention in Section 1.7.6 for regularity.

1.7.4 Multi-marginal transport problems

This recent part of the theory of optimal transport is becoming more and more popular and finds applications in many different fields. We will only give a brief sketch of what is done or can be done on this topic.

It is not difficult to imagine that, instead of using measures $\gamma \in \mathscr{P}(X \times Y)$ with prescribed marginals on the two factors X and Y, one could consider more than two marginals. Take some spaces $X_1, X_2, \dots, X_N$; define $\mathscr{X} := X_1 \times X_2 \times \cdots \times X_N$; take a cost function $c : \mathscr{X} \to \mathbb{R} \cup \{+\infty\}$, some measures $\mu_i \in \mathscr{P}(X_i)$, and solve

$$\min\left\{ \int_{\mathscr{X}} c(x_1, \dots, x_N)\, \mathrm{d}\gamma \ : \ \gamma \in \mathscr{P}(\mathscr{X}),\ (\pi_i)_\#\gamma = \mu_i \right\},$$

where $\pi_i : \mathscr{X} \to X_i$ is the canonical projection map. Duality could be performed in the same way as for the two-marginal case, getting to problems such as

$$\max\left\{ \sum_i \int_{X_i} \varphi_i \, \mathrm{d}\mu_i \ : \ \varphi_i : X_i \to \mathbb{R},\ \sum_i \varphi_i(x_i) \le c(x_1, \dots, x_N) \right\}.$$

The existence of a transport map (in the sense that there exist maps $\mathrm{T}_i : X_1 \to X_i$, i$\geq 2$, such that the optimal γ is of the form $\gamma = (\mathrm{id}, \mathrm{T}_2, \dots, \mathrm{T}_N)_\#\mu_1$) is a much more delicate question than the usual case. In some sense, it corresponds to considering the above transport problem as a two-marginal Monge problem on the spaces X_1 and $X_2 \times \cdots \times X_N$. Yet, the difference in "dimensions" between the source and target spaces makes the existence of an optimal map trickier. On the other hand, the image measure on $X_2 \times \cdots \times X_N$ is not really prescribed, but only its marginals on $X_2, \dots, X_N$ are fixed. This means that extra optimality properties should be taken into account.

This kind of problems may have several different interpretations. For a comprehensive survey about multi-marginal transport problems and their applications in physics and economics, see [249].

Among the first examples that have been treated in this setting, we cite the cost

$$c(x_1,\dots,x_N) := \sum_{i,j} |x_i - x_j|^2,$$

which has many similar properties to the 2-marginal case (see [177]), and the determinant cost, where the goal is to maximize

$$c(x_1,\dots,x_d) := \det(x_1,\dots,x_d) \quad \text{or} \quad c(x_1,\dots,x_d) := |\det(x_1,\dots,x_d)|,$$

where the determinant of a family of d vectors in $\mathbb{R}^d$ is intended as the determinant of the matrix where these vectors are the columns (see [109]). The problem is interesting for $d > 2$ since in 2D, from $\det(x, y) = x \cdot Ry$ (where R is the rotation of 90°), one comes back to Brenier problem via a change of variable (also compare to **Ex**(2)). The case of [177] will also appear in Section 5.5.5 in connection with barycenters.

Yet, we prefer to present some few cases where important applications naturally lead to a multi-marginal transport problem. For instance, in economics (see [108]), one can think at contracts with more than two agents (when a seller given a good to a buyer in exchange of money, this is a contract between two agents; when somebody wants to build a house, he/she buys some land and hires a carpenter and an architect, there are at least four agents). A point in the space $\mathscr{X}$ stands in this case for a possible contract between the agents $x_1, x_2, \dots, x_N$ and $c(x_1,\dots,x_N)$ is the global utility of this contract.

A completely different setting is found in recent works from physics, in the framework of Electronic Density Functional Theory (see [97] and [128]). To describe these issues in a very simplified way, consider a family of N electrons moving around a nucleus. In quantum physics, the configuration of the electrons is described through a wave function $\psi : \Omega^N \to \mathbb{C}$, where Ω is the physical region where the electrons live (typically $\Omega = \mathbb{R}^3$). For simplicity, we will ignore spin issues and the fact that the symmetry properties of ψ depend on the distinction between fermionic and bosonic particles. We will only consider that the quantity $|\psi|^2$ is invariant under permutations of the variables. This same scalar function denotes the joint law of the position of the electrons. Hence, the probability measure $\varrho \in \mathscr{P}(\mathbb{R}^3)$ given by $(\pi_1)_\#(|\psi|^2)$ represents the law of the position of a single electron. The energy of the state is represented by the following sum:

$$\hbar \int_{\Omega^N} |\nabla\psi|^2 \,\mathrm{d}x + \int_{\Omega^N} (V(x_1) + \cdots + V(x_N))|\psi|^2 \,\mathrm{d}x + \int_{\Omega^N} \left(\sum_{i<j} \frac{1}{|x_i - x_j|} \right) |\psi|^2 \,\mathrm{d}x,$$

where the first term represents the kinetic energy, the second the potential energy (created, for instance, by the nucleus), and the last the electric interaction energy between the electrons. As far as the parameter N is concerned, it is meant to be the atomic number of an atom, which means that it ranges from two to one or two hundreds. The minimization of the above energy becomes a multi-marginal transport problem as soon as one takes the semiclassical limit $\hbar \to 0$ (see [128, 129]; the rigorous limit for $N > 2$ is an ongoing work). This allows us to get rid of the first term in the energy, and the second is easily seen to depend on ϱ only. Hence, the only unknown is the joint law of the positions of the N electrons, i.e., the probability $\gamma \in \mathscr{P}((\Omega)^N)$ given by $\gamma = |\psi|^2 \mathrm{d}x_1 \dots \mathrm{d}x_N$, with marginals all equal to ϱ. The cost c is given by

$$c(x_1, \dots, x_N) := \sum_{i<j} \frac{1}{|x_i - x_j|},$$

so that γ should minimize $\int c \, \mathrm{d}\gamma$ under marginal constraints. The fact that c is neither bounded nor finite (it goes to $+\infty$ as soon as two electrons approach each other) is a strong difficulty of this problem even in the "easy" case $N = 2$. It is by the way an interesting transport problem where configurations with equalities $x_i = x_j$ cost the most, so that the problem is meaningful and intriguing even and in particular when the marginals are all equal (which would not be the case for costs increasingly depending on the distances).

We observe that we already faced several examples of "symmetric" cost functions, meaningful when the marginals are identical (which is not the case for the costs in [177] and [109]). This raises several questions about symmetric minimizers. Indeed, that all these problems can be restricted to the set of plans γ which are invariant under permutations of the components. The theory of symmetric transport problems, in connection with a variant of polar factorization theory and with self-dual convex problems, has been investigated a lot by N. Ghoussoub and collaborators, and we refer to [179, 180] without entering into extra details.

We finish this section by underlining another point: in evolution problems, and particularly in fluid mechanics, a natural Lagrangian point of view is the description of the global movement by using a measure on the set of possible trajectories. This typically requires the use of probabilities on the set of continuous or Lipschitz paths on $[0, 1]$, but any kind of time discretization of such a setting can be interpreted in terms of measures on Ω^N, where Ω is the state space and N represents the number of time steps, each one of length $\tau = 1/N$. In this case, the marginals are somehow "ordered" and the cost takes into account this feature: for instance, natural costs could be

$$c(x_1, \dots, x_N) = \sum_{i=1}^{N-1} |x_i - x_{i+1}| \quad \text{or} \quad c(x_1, \dots, x_N) = \sum_{i=1}^{N-1} \frac{|x_i - x_{i+1}|^2}{\tau},$$

which are, respectively, the total length of a discrete path or the total kinetic energy (the term $|x-y|^2/\tau$ standing for the integration over a time interval of length τ of a speed $|x-y|/\tau$). Let us mention that the continuous counterpart of this last cost, i.e., $c(\omega) := \int_0^1 |\omega'(t)|^2 \mathrm{d}t$ for $\omega \in \mathscr{C} := H^1([0,1];\Omega)$, appears in Brenier's variational formulation of the incompressible Euler equation (see [83]), which gives rise to a minimization problem such as

$$\min \left\{ \int_H c(\omega)\,\mathrm{d}Q(\omega) \;:\; Q \in \mathscr{P}(\mathscr{C}),\ (e_t)_\# Q = \mathscr{L}_\Omega,\ (e_0,e_1)_\# Q = \gamma_{0,1} \right\},$$

where $e_t : \mathscr{C} \to \Omega$ is the evaluation map $e_t(\omega) := \omega(t)$, $\mathscr{L}_\Omega$ is the rescaled Lebesgue measure over Ω; and $\gamma_{0,1}$ is a fixed measure in $\mathscr{P}(\Omega \times \Omega)$ with both marginals equal to $\mathscr{L}_\Omega$.

1.7.5 Martingale optimal transport and financial applications

We present in this section two problems from mathematical finance involving multi-marginal optimal transport. To analyze them, one should know what is an option. An option is a contract giving the owner the right to buy or sell some financial assets at some given conditions, which could bring a monetary advantage. The most typical case is the European Call option, saying "at time T you will have the right to buy this asset at price K": should the price be S_T, higher than K, then you gain $S_T - K$; should it be lower, then simply do not use the option, so that finally the gain is $(S_T - K)_+$. More generally, we call option any financial contract giving the owner a gain which is a function of the value S_T at time T of a given asset (a share, for instance). This value is a random variable (based on a probability $\mathbb{P}$) evolving in time. More exotic options such that their payoff depends on the whole history $(S_t)_t$ of the asset value also exist. One of the main issues of financial mathematics is to give formulas to compute the correct price for these contracts. This is based on the no-arbitrage assumption: the only reasonable price for a contract is the one which avoids the existence of arbitrage[14] opportunities on the market, i.e., the possibility of buying and selling this contract together with the related underlying asset and to produce a positive amount of money out of nothing (more precisely, since all the values are random variables: to produce with probability 1 a nonnegative amount of money, which is strictly positive with strictly positive probability). A general theorem in financial mathematics (see, for instance, [139] for references) states that all these no-arbitrage prices are actually the expected value of the contract according to a probability $\mathbb{Q}$ which is in general not $\mathbb{P}$ (but $\mathbb{P} \ll \mathbb{Q}$ and $\mathbb{Q} \ll \mathbb{P}$, so that they have the same negligible sets) and which has the property that all the asset values on the market are martingales under $\mathbb{Q}$. This means, for instance, that the price of

[14] The absence of arbitrage is usually referred to as "no free lunch."

a European Call option should be $\mathbb{E}^{\mathbb{Q}}[(S_T - K)_+]$. Note that the knowledge of this price for every K implies the knowledge of the law of S_T under $\mathbb{Q}$.

A first problem that one could find in finance is the following: consider a complicated option with payoff depending on the values at time T of several assets S^i_T at the same time T. One should compute $\mathbb{E}^{\mathbb{Q}}[f(S^1_T, S^2_T, \dots, S^N_T)]$. Yet, this depends on the joint law of the different assets under the unknown probability $\mathbb{Q}$. If we suppose that this option is new on the market, but that all the European Calls on each of the assets S^i are regularly sold on the market for each strike price K, then this means that we know the marginal laws μ^i under $\mathbb{Q}$ of each asset, and we only lack the joint law. We can have an estimate of the correct price of the option by computing

$$\min / \max \left\{ \int f(x_1, x_2, \dots, x_N) \, \mathrm{d}\gamma \ : \ (\pi_i)_\# \gamma = \mu_i \right\}$$

which is nothing but a transport problem. We ignored here the constraints on the fact that $\mathbb{Q}$ should make the asset value process a martingale since they are already included in the law of each asset and do not involve the joint law. Indeed, in this case, the different marginals represent different financial assets, and not different instants of time, and the martingale condition does not play any role[15].

Another problem, much more intriguing, can be obtained when one considers a single asset but uses several marginals for several time steps. This allows us to consider exotic options depending on the whole history of the asset value. Consider a process $(S_t)_t$ defined for various instants of time (we will consider discrete time models for the sake of simplicity: $t = 0, 1, \dots, N$). Suppose that an option pays $f(S_0, S_1, \dots, S_N)$. Suppose that European Calls are traded in the market for each maturity time t and each strike price K, which provides the law of S_t under the unknown martingale measure $\mathbb{Q}$. Then, the price of the option may be estimated by solving

$$\min / \max \left\{ \mathbb{E}^{\mathbb{Q}}[f(S_0, S_1, \dots, S_N)] \ : \ (S_t)_\# \mathbb{Q} = \mu^t, \ \ (S_t)_t \text{ is a martingale under } \mathbb{Q} \right\}.$$

This can also be expressed in terms of the measure $\gamma = (S_0, S_1, \dots, S_N)_\# \mathbb{Q}$ as

$$\min / \max \left\{ \int f(x_1, x_2, \dots, x_N) \, \mathrm{d}\gamma \ : \ (\pi_i)_\# \gamma = \mu_i, \ \gamma \in \mathrm{Mart}_N \right\},$$

where the set Mart_T is the set of discrete time martingale measures γ, satisfying for each $t \leq N - 1$

[15] We could have presented this problem in Section 1.7.4, but we preferred to do it here in order to introduce the martingale setting.

$$\int x_{t+1}\phi(x_0, x_1, \dots, x_t)\,\mathrm{d}\gamma = \int x_t\phi(x_0, x_1, \dots, x_t))\,\mathrm{d}\gamma. \tag{1.7}$$

These problems are now called martingale optimal transport problems and are out of the framework of usual transport problems that we presented before. By the way, in general, there is no reason to hope for the existence of an optimal transport map. Indeed, even in the easiest case, i.e., two marginals, this will not be the case. The reason is the fact that the only martingale measure which is of the form $\gamma_{\mathrm{T}} = (\mathrm{id}, \mathrm{T})_{\#}\mu$ is that with $\mathrm{T} = \mathrm{id}$. This can be seen from Equation (1.7), since it gives

$$\int \mathrm{T}(x)\phi(x)\,\mathrm{d}\mu = \int y\phi(x)\,\mathrm{d}\gamma_{\mathrm{T}} = \int x\phi(x)\,\mathrm{d}\gamma_{\mathrm{T}} = \int x\phi(x)\;\mathrm{d}\mu,$$

which implies, ϕ being an arbitrary test function, $\mathrm{T}(x) = x$ μ-a.e. This means that, unless $\nu = \mu$, the martingale transport problem with two marginals admits no admissible transport plan issued by a transport map. On the contrary, it is possible to realize γ as a combination of two transport maps, and optimality results of this kind are proven in [31, 193]. For a general presentation of the martingale optimal transport problem, see [32] and [173].

We finish this section presenting an alternative approach, leading to the same problems, and based on the idea of optimal hedging. Think that $f(S_0, \dots, S_N)$ is the loss that a company would face, depending on the prices of the asset S (i.e., it is the amount of money that the company needs to pay, because of a debt or other contracts). The company wants to "cover" (or "hedge") this amount, by using the possibilities in the market. In order to hedge this amount, the company can

- buy usual options, based on the value of the asset at time t, i.e., buy a contract whose value will be $\phi(S_t)$ and whose price is known to be $\int \phi\,\mathrm{d}\mu_t$;
- buy some amounts of the asset itself: at time $t \leq N-1$, it is possible to but for a price S_t and to resell at price S_{t+1}; the number of assets to buy can only be chosen according to the information that are available at time t, which means that one can buy (and then resell) an amount of the form $\psi(S_0, S_1, \dots, S_t)$; this results in covering a quantity $\psi(S_0, S_1, \dots, S_t)(S_{t+1} - S_t)$, at cost 0.

The optimal hedging problem becomes

$$\min\left\{\sum_{t=0}^{N}\int \varphi_t\,\mathrm{d}\mu_t \;:\; \sum_{t=0}^{N}\varphi_t(x_t) + \sum_{t=0}^{N-1}\psi(x_0, x_1, \dots, x_t)(x_{t+1} - x_t) \geq f(x_0, x_1, \dots, x_N)\right\}.$$

It is not difficult to check, thanks to the marginal constraints and to the martingale condition (1.7), that this problem is exactly the dual of the above martingale transport problem.

1.7.6 Monge-Ampère equations and regularity

We saw in Section 1.3.1 that the optimal transport for the quadratic case is of the form $\mathrm{T}(x) = \nabla u(x)$, for a convex function u. If this function is smooth and strictly convex, and the two measures μ and ν are absolutely continuous (say, with densities f and g), then we can write the condition $\mathrm{T}_\#\mu = \nu$ in a PDE form. Indeed, this condition is equivalent, if $\mu = f(x)\,\mathrm{d}x$ and $\nu = g(y)\,\mathrm{d}y$ and if T is C^1 and injective, to the Jacobian condition

$$\det(D\mathrm{T}(x)) = \frac{f(x)}{g(\mathrm{T}(x))}$$

which can be obtained as a consequence of a simple change-of-variable computation, as explained here below.

Box 1.14. ***Memo**: Change-of-variable and image measures*

Proposition. Suppose that $\varrho \in L^1(\Omega)$ is a positive density on $\Omega \subset \mathbb{R}^d$ and $\mathrm{T} : \Omega \to \mathbb{R}^d$ is a Lipschitz injective map, which is thus differentiable a.e. We suppose that $\det(D\mathrm{T}) \neq 0$ a.e. on $\{\varrho > 0\}$. Then the image measure $\mathrm{T}_\#\varrho$ is absolutely continuous and its density ϱ_T is given by

$$\varrho_\mathrm{T}(y) = \frac{\varrho(\mathrm{T}^{-1}(y))}{\det(D\mathrm{T}(\mathrm{T}^{-1}(y)))}.$$

If T is non-injective, the formula becomes $\mathrm{T}_\#\varrho = \varrho_\mathrm{T} \cdot \mathscr{L}^d$ with ϱ_T given by

$$\varrho_\mathrm{T}(y) = \sum_{x:\mathrm{T}(x)=y} \frac{\varrho(x)}{\det(D\mathrm{T}(x))}.$$

The same formulas stay true if T is not Lipschitz, but if there are measurable sets E_i such that $|\Omega \setminus \bigcup_i E_i| = 0$ and T is Lipschitz continuous on each set E_i (see also Definition 3.13 for countably Lipschitz functions), with the differential $D\mathrm{T}$ which is actually the differential of the restriction of T to each set where it is Lipschitz continuous (and coincides thus with the approximate differential of T, see Section 3.3.2).

In the particular case where $\mathrm{T} = \nabla u$ (where we suppose u to be strictly convex just to guarantee injectivity of T), this becomes the so-called Monge-Ampère equation

$$\det(D^2u(x)) = \frac{f(x)}{g(\nabla u(x))}, \tag{1.8}$$

which is a nonlinear PDE of elliptic type. Nonlinearity comes both from the right-hand side with the term $g(\nabla u)$ but also from the higher-order term $\det(D^2u)$. This Jacobian term is nonlinear but has some monotonicity property when we assume the

matrix D^2u to be positive (which is the case for convex functions), as we have $0 \le A \le B \Rightarrow \det(A) \le \det(B)$. This monotonicity is anyway very much degenerate, but enough to classify this nonlinear PDE as a fully nonlinear (degenerate) elliptic equation. We refer, for instance, to [190] for the general theory about this equation.

In our case, it is important to understand the meaning and the boundary conditions to be associated to this equation.

First, let us consider the boundary conditions. In view of the transportation meaning of this PDE, if we consider $\mu \in \mathscr{P}(\Omega)$ and $\nu \in \mathscr{P}(\Omega')$, there is no point in imposing the behavior of u or ∇u on $\partial\Omega$. On the contrary, the natural condition is just $\nabla u(\Omega) = \Omega'$, which is a global condition on u. This corresponds, roughly speaking and only when u is a homeomorphism, to $u(\partial\Omega) = \partial\Omega'$. This particular sort of "boundary" condition is usually called *second boundary value problem* for the Monge-Ampère equation.

We define various notions of solutions for (1.8):

- we say that u satisfies (1.8) in the Brenier sense if $(\nabla u)_\#(f \cdot \mathscr{L}^d) = g \cdot \mathscr{L}^d$ (and this is actually the sense to be given to this equation in optimal transport);
- we say that u satisfies (1.8) in the Alexandroff sense if the measure ϱ defined via $\varrho(A) := |\bigcup_{x\in A} \partial u(x)|$ is absolutely continuous and its density equals the right-hand side of (1.8) a.e. (warning: the fact that ϱ is a measure is not straightforward, one has to use properties of convex functions, as in Lemma 1.1.12 and Theorem 1.1.13 in [190]);
- we say that u satisfies (1.8) in the viscosity sense if it satisfies the usual comparison properties required by viscosity theory but restricting the comparisons to regular convex test functions (this is possible because of the monotonicity properties of the determinant);
- we say that u satisfies (1.8) in the classical sense if it is of class C^2 and the equation holds pointwise.

Note that all these notions except the first may also be applied to the more general equation $\det D^2u = f$ (for fixed f, without the structure $f/g \circ \mathrm{T}$), while the first one just applies to this specific transportation case.

The regularity of the solutions of (1.8), which implies regularity results for the optimal transport map, has been studied in the 1990s by Caffaelli, with very strong results. The results we want to use are well summarized in Theorem 50 of [292]:

Theorem 1.54. *If f and g are $C^{0,\alpha}(\Omega)$ and are both bounded from above and from below on the whole Ω by positive constants and Ω is a convex open set, then the unique Brenier solution u of (1.8) belongs to $C^{2,\alpha}(\Omega) \cap C^{1,\alpha}(\overline{\Omega})$, and u satisfies the equation in the classical sense (hence also in the Alexandroff and viscosity senses).*

We just detail a possible bibliographical path to arrive at this result. It is not easy to deal with Brenier solutions, so the idea is to consider viscosity solutions, for which it is in general easy to prove existence by Perron's method. Then prove some regularity result on viscosity solutions, up to getting a classical solution. After that, once we have a classical convex solution to Monge-Ampère equation, this will be a Brenier solution too. Since this is unique (up to additive constants), we have got a regularity

statement for Brenier solutions. We can find results on viscosity solutions in [98, 100] and [99]. In [98], some conditions are given so as to ensure strict convexity of the solution of $\det D^2 u = f$ when f is bounded from above and below. In [100], for the same equation it is proved $C^{1,\alpha}$ regularity provided we have strict convexity. In this way, the term $f/g(\nabla u)$ becomes a $C^{0,\alpha}$ function, and in [99], it is proved $C^{2,\alpha}$ regularity for solutions of $\det D^2 u = f$ with $f \in C^{0,\alpha}$.

Some recent improvements on this regularity theory are due to De Philippis and Figalli, who worked on this natural question: if f and g are bounded from above and below, then the transport map is a BV map (since it is the gradient of a convex function), which happens to be continuous (since $u \in C^{1,\alpha}$); its derivative has no jump part (because of continuity), and the question arises whether $\mathrm{T} \in W^{1,1}$, i.e., $u \in W^{2,1}$. The answer is yes, as it is shown in [147, 148, 150], and one can even prove sharper estimates ($D^2 u$ is not only L^1 but slightly better, including some $L^{1+\varepsilon}$ results, also proven by Schmidt in [281]). For a comprehensive survey of the relations of the Monge-Ampére equation with optimal transport and its regularity properties, see [149].

The situation becomes much trickier when dealing with different costs than the quadratic one. If the cost c satisfies the twist condition, it is still possible to express the optimal map T in terms of a gradient, $\mathrm{T}(x) = (\nabla_x c(x,\cdot))^{-1}(\nabla\varphi(x))$. If we write $\nabla_x c(x, \mathrm{T}(x)) = \nabla\varphi(x)$, differentiating this equality we get

$$D^2_{xx} c(x, \mathrm{T}(x)) + D^2_{xy} c(x, \mathrm{T}(x)) D\mathrm{T}(x) = D^2\varphi(x),$$

which means that the equation on $\det(D\mathrm{T})$ reads

$$\det(D^2_{xy} c(x, \mathrm{T}(x))) \det(D\mathrm{T}(x)) = \det(D^2\varphi(x) - D^2_{xx} c(x, \mathrm{T}(x))),$$

i.e.,

$$\det(D^2\varphi(x) - D^2_{xx} c(x, \mathrm{T}(x))) = \det(D^2_{xy} c(x, \mathrm{T}(x))) \frac{f(x)}{g(\mathrm{T}(x))},$$

which is a Monge-Ampère-type equation of the form

$$\det(D^2\varphi(x) - A(x, \nabla\varphi(x))) = F(x, \nabla\varphi(x)).$$

The regularity problem of the Kantorovich potentials for general costs has been a long-standing problem, and it was listed as open in [292]. It is now well clarified after a remarkable paper by Ma, Trudinger, and Wang, who found out in [219] a key assumption on the cost c so as to be able to develop regularity estimates, based on some Pogorelov-type methods. The assumption on c is very technical and requires some inequalities on its fourth-order derivatives. We do not want to write them here, but it is worth noting that these assumptions have later been proven to be sharp by Loeper in [216]. In particular, the case $c(x, y) = |x - y|^2$ is covered by this assumption but is a limit case, and no other power cost $|x - y|^p$ for $p \in]1, +\infty[$

satisfies it. The cost $\sqrt{\varepsilon^2 + |x-y|^2}$ also satisfies the MTW assumption for every $\varepsilon > 0$, and this allowed some approximation result for the linear cost $|x-y|$ (see [210]). The convexity of the domain Ω is also replaced by a suitable notion of c-convexity that we do not detail here. The connection with convexity properties related to c-concave functions is also underlined in [168], as we saw in Section 1.7.3. The theory has been deeply studied after this discovery, and many costs have been considered: the case of the squared distance on a manifold can also be investigated, and the condition translates into a curvature condition on the manifold. We refer to [164] for an interesting survey of the whole theory.

Chapter 2
One-dimensional issues

This chapter is devoted to the special features of the 1D case and to a less known transport map in the higher-dimensional case, called the *Knothe map*, built from 1D constructions and related to some degenerate transport cost. The discussion section deals with applications of 1D transport maps to image processing and geometric inequalities.

2.1 Monotone transport maps and plans in 1D

We already discussed, in Section 1.3, the quadratic cost in 1D: as soon as the source measure μ has no atoms, there exists an optimal map, which is monotone nondecreasing.

In this section we want to discuss monotone (in the following "monotone" means "monotone nondecreasing") transport maps between two given measures in terms of their cumulative distribution functions and generalize to the case where μ may have atoms. This discussion is independent of optimal transport considerations.

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2_2

Box 2.1. ***Memo**: Cumulative distribution function*

Definition. Given a probability measure $\mu \in \mathscr{P}(\mathbb{R})$, we define its cumulative distribution function F_μ through

$$F_\mu(x) = \mu((-\infty, x]).$$

The cumulative distribution function F_μ is easily seen to be nondecreasing and right continuous since if $x_n \to x$ with $x_n > x$, then $(-\infty, x] = \cap_n (-\infty, x_n]$ and hence $\mu((-\infty, x]) = \lim_n \mu((-\infty, x_n]) = \inf_n \mu((-\infty, x_n])$. It is continuous at any point where μ has no atom since if $x_n \to x$ with $x_n < x$, then $\lim_n \mu((-\infty, x_n]) = \mu((-\infty, x[)$.

The CDF F_μ is important because it characterizes the measure μ. Indeed, the sets of the form $(-\infty, x]$ are enough to generate all open sets: we have $\mu(]a, b]) = F_\mu(b) - F_\mu(a)$, $\mu(]a, b[) = \sup\{F_\mu(t) - F_\mu(a) \,:\, t < b\}$ and the measure of any open set is obtained via countable disjoint unions of open intervals.

In $\mathbb{R}^d$ the situation is trickier, but the knowledge of $\mu(]-\infty, x_1] \times \cdots \times]-\infty, x_d])$ for every $x = (x_1, \dots, x_d)$ is enough, as in 1D, to characterize μ (by finite differences one could get all the semi-open rectangles $]x_1^-, x_1^+] \times \cdots \times]x_d^-, x_d^+]$ and by countable disjoint unions all the open sets).

Unfortunately, the CDF above cannot always be inverted, as it is not always strictly increasing, but we can define a pseudo-inverse.

Definition 2.1. Given a nondecreasing and right-continuous function $F : \mathbb{R} \to [0, 1]$, its pseudo-inverse is the function $F^{[-1]} : [0, 1] \to \overline{\mathbb{R}}$ given by

$$F^{[-1]}(x) := \inf\{t \in \mathbb{R} \,:\, F(t) \geq x\},$$

where the infimum is a minimum as soon as the set is nonempty (otherwise it is $+\infty$) and bounded from below (otherwise it is $-\infty$), thanks to right continuity of F.

Note, as a simple consequence of the definition of pseudo-inverse, that we have

$$F^{[-1]}(x) \leq a \Leftrightarrow F(a) \geq x\,; \quad F^{[-1]}(x) > a \Leftrightarrow F(a) < x. \tag{2.1}$$

We look now at some properties of the pseudo-inverse of cumulative distribution functions.

Proposition 2.2. *If $\mu \in \mathscr{P}(\mathbb{R})$ and $F_\mu^{[-1]}$ is the pseudo-inverse of its cumulative distribution function F_μ, then $(F_\mu^{[-1]})_\#(\mathscr{L}^1 \llcorner [0, 1]) = \mu$.*

Moreover, given $\mu, \nu \in \mathscr{P}(\mathbb{R})$, if we set $\eta := (F_\mu^{[-1]}, F_\nu^{[-1]})_\#(\mathscr{L}^1 \llcorner [0, 1])$, then $\eta \in \Pi(\mu, \nu)$ and $\eta((-\infty, a] \times (-\infty, b]) = F_\mu(a) \wedge F_\nu(b)$.

Proof. For the first part of the statement, using (2.1), we write

$$|\{x \in [0,1] \,:\, F_\mu^{[-1]}(x) \le a\}| = |\{x \in [0,1] \,:\, F_\mu(a) \ge x\}| = F_\mu(a),$$

and this proves that the image measure is μ, using the characterization of a measure through its CDF (see the Memo Box 2.1 above).

For the second part of the statement, $\eta \in \Pi(\mu,\nu)$ is just a consequence of the first. Then, let us compute

$$\begin{aligned}\eta((-\infty,a]\times(-\infty,b]) &= |\{x \in [0,1] \,:\, F_\mu^{[-1]}(x) \le a,\, F_\nu^{[-1]}(x) \le b\}| \\ &= |\{x \in [0,1] \,:\, F_\mu(a) \ge x,\, F_\nu(b) \ge x\}| = F_\mu(a) \wedge F_\nu(b),\end{aligned}$$

which is the desired equality. □

Definition 2.3. We will call the transport plan $\eta := (F_\mu^{[-1]}, F_\nu^{[-1]})_\#(\mathscr{L}^1 \llcorner [0,1])$ the co-monotone transport plan between μ and ν and denote it by γ_{mon}.

Now, consider two measures $\mu, \nu \in \mathscr{P}(\mathbb{R})$: we want to build a monotone transport *map* (and not only a plan) sending μ onto ν, provided μ is atomless.

We first need a simple lemma on the CDF of atomless measures.

Lemma 2.4. *If $\mu \in \mathscr{P}(\mathbb{R})$ is atomless, then $(F_\mu)_\#\mu = \mathscr{L}^1 \llcorner [0,1]$. As a consequence, for every $\ell \in [0,1]$, the set $\{x \,:\, F_\mu(x) = \ell\}$ is μ-negligible.*

Proof. First note that F_μ is continuous because μ is atomless. Hence, for $a \in]0,1[$, the set $\{x \,:\, F_\mu(x) \le a\}$ is a closed interval of the form $]-\infty, x_a]$, with $F_\mu(x_a) = a$. Hence, $\mu(\{x \,:\, F_\mu(x) \le a\}) = F_\mu(x_a) = a$, which proves $(F_\mu)_\#\mu = \mathscr{L}^1 \llcorner [0,1]$.

As a consequence, for $\ell \in [0,1]$, the sets $\{x \,:\, F_\mu(x) = \ell\}$ are μ-negligible, since otherwise the image measure $(F_\mu)_\#\mu$ would have an atom at ℓ, which contradicts the first part of the statement. □

Theorem 2.5. *Given $\mu, \nu \in \mathscr{P}(\mathbb{R})$, suppose that μ is atomless. Then, there exists a unique nondecreasing map $\mathrm{T}_{\text{mon}} : \mathbb{R} \to \mathbb{R}$ such that $(\mathrm{T}_{\text{mon}})_\#\mu = \nu$.*

Proof. First, let us build one such a map. Let us consider the cumulative distribution functions F_μ and F_ν and define the map T_{mon} through

$$\mathrm{T}_{\text{mon}}(x) := F_\nu^{[-1]}(F_\mu(x)).$$

This quantity is well defined and belongs to $\mathbb{R}$ provided $F_\mu(x) \in]0,1[$ (so that the set on which we take the infimum in Definition 2.1 is neither empty nor unbounded from below). The sets $\{x \,:\, F_\mu(x) = 0\}$ and $\{x \,:\, F_\mu(x) = 1\}$ are μ-negligible thanks to Lemma 2.4. Hence, T_{mon} is well-defined μ-a.e.

The fact that T_{mon} is monotone nondecreasing is obvious; we just have to prove $(\mathrm{T}_{\text{mon}})_\#\mu = \nu$. Since we already know $(F_\nu^{[-1]})_\#(\mathscr{L}^1 \llcorner [0,1]) = \nu$, by composition we just need to use Lemma 2.4, which proves $(F_\mu)_\#\mu = \mathscr{L}^1 \llcorner [0,1]$.

We now pass to the proof of uniqueness. Consider any nondecreasing function T such that $T_{\#}\mu = \nu$.

From monotonicity we have $T^{-1}(]-\infty, T(x)]) \supset]-\infty, x]$. We deduce

$$F_\mu(x) = \mu(]-\infty, x]) \le \mu(T^{-1}(]-\infty, T(x)])) = \nu(]-\infty, T(x)]) = F_\nu(T(x)),$$

which means $T(x) \ge F_\nu^{[-1]}(F_\mu(x))$. Suppose that the inequality is strict. This means that there exists $\varepsilon_0 > 0$ such that $F_\nu(T(x) - \varepsilon) \ge F_\mu(x)$ for every $\varepsilon \in]0, \varepsilon_0[$. On the other hand, from $T^{-1}(]-\infty, T(x) - \varepsilon[) \subset]-\infty, x[$, we get $F_\nu(T(x) - \varepsilon) \le F_\mu(x)$.

Hence, we obtain $F_\nu(T(x) - \varepsilon) = F_\mu(x)$ for every $\varepsilon \in]0, \varepsilon_0[$. Thus, $F_\nu(T(x) - \varepsilon)$ is the value that F_ν takes on an interval where it is constant. These intervals are a countable quantity; hence the values of F_ν on those intervals are also countable. We call ℓ_i these values. As a consequence, the points x where $T(x) > F_\nu^{[-1]}(F_\mu(x))$ are contained in $\bigcup_i \{x : F_\mu(x) = \ell_i\}$. Lemma 2.4 proves that this set is negligible. This allows us to conclude $T(x) = F_\nu^{[-1]}(F_\mu(x))$ μ-a.e. □

Remark 2.6. Note that the previous proof was complicated by the possibility that the cumulative distribution functions could be either discontinuous or not strictly increasing. Should F_ν be continuous and strictly monotone (which means that ν is atomless and supported on the whole $\mathbb{R}$), then one would simply have

$$T_{\text{mon}} = (F_\nu)^{-1} \circ F_\mu.$$

Remark 2.7. As a consequence of the explicit formula of the previous remark, one can also study the regularity of this map T_{mon} depending on the regularity of μ and ν. Indeed, as soon as these measures are fully supported and have no atoms, the two functions F_μ and F_ν are homeomorphisms, and so is T_{mon}. Moreover, if μ and ν are absolutely continuous, with continuous densities which do not vanish, then they are also diffeomorphisms, and so is T_{mon}. In general, the regularity of T_{mon} is one degree higher than that of the two measures, as it is the case for F_μ, F_ν, F_μ^{-1}, and F_ν^{-1}.

In the next section we will see that the map T_{mon} that we have just built optimizes a whole class of transport costs. To prove it, we will need the following characterization of γ_{mon} and T_{mon}.

Lemma 2.8. *Let $\gamma \in \Pi(\mu, \nu)$ be a transport plan between two measures μ, $\nu \in \mathscr{P}(\mathbb{R})$. Suppose that it satisfies the property*

$$(x, y), (x', y') \in \text{spt}(\gamma),\ x < x' \ \Rightarrow\ y \le y'. \tag{2.2}$$

Then, we have $\gamma = \gamma_{\text{mon}}$. In particular there is a unique γ satisfying (2.2).

Moreover, if μ is atomless, then $\gamma = \gamma_{T_{\text{mon}}}$.

Proof. For the first part of the statement, we just need to prove

$$\gamma((-\infty, a] \times (-\infty, b]) = F_\mu(a) \wedge F_\nu(b).$$

Indeed, from Proposition 2.2, this condition is satisfied by γ_{mon}, and we saw that this is enough to characterize a measure on $\mathbb{R}^2$.

Consider the two sets $A =]-\infty, a] \times]b, +\infty[$ and $B =]a, +\infty[\times]-\infty, b]$. From the assumption on γ, it is not possible to have both $\gamma(A) > 0$ and $\gamma(B) > 0$ (otherwise we would have points in $\text{spt}(\gamma)$ violating condition (2.2)). Hence, we can write

$$\gamma((-\infty, a] \times (-\infty, b]) = \gamma(((-\infty, a] \times (-\infty, b]) \cup A) \wedge \gamma(((-\infty, a] \times (-\infty, b]) \cup B).$$

Yet, we have

$$\gamma(((-\infty, a] \times (-\infty, b]) \cup A) = \gamma((-\infty, a] \times \mathbb{R}) = F_\mu(a)$$

and

$$\gamma(((-\infty, a] \times (-\infty, b]) \cup B) = \gamma(\mathbb{R} \times (-\infty, b]) = F_\nu(b).$$

This proves the first part of the claim. For the second part, we suppose μ to be atomless. For any point $x \in \mathbb{R}$, one can define the interval I_x as the minimal interval I such that $\text{spt}(\gamma) \cap (\{x\} \times \mathbb{R}) \subset \{x\} \times I$. This interval can obviously be reduced to a singleton. The assumption on γ implies that the interior of all these intervals is disjoint (and ordered). In particular there can be at most a countable quantity of points such that I_x is not a singleton. Since these points are μ-negligible (as a consequence of μ being atomless), we can define μ-a.e. a map T such that γ is concentrated on the graph of T. This map will be monotone nondecreasing because of (2.2), and this gives $\text{T} = \text{T}_{\text{mon}}$ since we already know the uniqueness of a nondecreasing map with fixed marginals (Theorem 2.5). □

2.2 The optimality of the monotone map

Now that we know quite well the properties, definitions, and characterizations of the map T_{mon} and of the plan γ_{mon}, we can see that they are, in the 1D case, optimal for several different costs and not only for the quadratic one. This is really specific to the 1D case; it will not be true in higher dimension. The costs that we will consider will be convex functions of the difference $x-y$ and, to stress the possible asymmetric behavior of these costs, we prefer to write $h(y-x)$ instead of $h(x-y)$.

Theorem 2.9. *Let $h : \mathbb{R} \to \mathbb{R}_+$ be a strictly convex function and $\mu, \nu \in \mathscr{P}(\mathbb{R})$ be probability measures. Consider the cost $c(x, y) = h(y-x)$ and suppose that* (KP) *has a finite value. Then,* (KP) *has a unique solution, which is given by γ_{mon}. In the case where μ is atomless, this optimal plan is induced by the map* T_{mon}.

Moreover, if the strict convexity assumption is withdrawn and h is only convex, then the same $\gamma_{\rm mon}$ is actually an optimal transport plan, but no uniqueness is guaranteed anymore.

Proof. We will use the fact that the support of any optimal γ is a c-CM set Γ. This means in particular that for $(x,y),(x',y')\in\Gamma$, we have

$$h(y-x)+h(y'-x')\le h(y'-x)+h(y-x'). \tag{2.3}$$

We only need to show that this implies (in the strictly convex case) a monotone behavior: we will actually deduce from (2.3) that $x<x'$ implies $y\le y'$, and this will allow us to conclude from Lemma 2.8.

To prove $y\le y'$, suppose by contradiction $y>y'$ and denote $a=y-x$, $b=y'-x'$ and $\delta=x'-x>0$. Condition (2.3) reads $h(a)+h(b)\le h(b+\delta)+h(a-\delta)$. Moreover, the assumption $y'<y$ implies $b+\delta<a$. Hence, $b+\delta$ and $a-\delta$ are located in the segment between b and a (and $b<a$). More precisely, we have

$$b+\delta=(1-t)b+ta,\qquad a-\delta=tb+(1-t)a,\qquad \text{for } t=\frac{\delta}{a-b}\in]0,1[.$$

Thus, strict convexity yields

$$\begin{aligned}h(a)+h(b)&\le h(b+\delta)+h(a-\delta)\\&<(1-t)h(b)+th(a)+th(b)+(1-t)h(a)=h(a)+h(b).\end{aligned}$$

This gives a contradiction and proves the statement in the strictly convex case.

The statement when h is only convex is trivial if h is constant (since every γ is optimal), and, if not, it is obtained by approximation. Lemma 2.10 below proves that there exists a sequence of strictly convex functions h_ε such that $h\le h_\varepsilon\le(1+\varepsilon)h+\varepsilon$. Let us take the transport cost $c_\varepsilon(x,y):=h_\varepsilon(y-x)$. In this case we know that $\gamma_{\rm mon}$ optimizes the cost $\int c_\varepsilon\,{\rm d}\gamma$ and hence

$$\int h(y-x)\,{\rm d}\gamma_{\rm mon}\le\int h_\varepsilon(y-x)\,{\rm d}\gamma_{\rm mon}\le\int h_\varepsilon(y-x)\,{\rm d}\gamma\le\varepsilon+(1+\varepsilon)\int h_\varepsilon(y-x)\,{\rm d}\gamma$$

for all $\gamma\in\Pi(\mu,\nu)$. Passing to the limit as $\varepsilon\to0$, we get that $\gamma_{\rm mon}$ also optimizes the cost c. □

Lemma 2.10. *For every nonconstant convex and positive function $h:\mathbb{R}\to\mathbb{R}_+$ and every $\varepsilon>0$, there exists $h_\varepsilon:\mathbb{R}\to\mathbb{R}_+$ strictly convex, such that $h\le h_\varepsilon\le(1+\varepsilon)h+\varepsilon$.*

Proof. We just need to prove that there is a strictly convex function $f:\mathbb{R}\to\mathbb{R}_+$ such that $f\le h+1$. Then, we will take $h_\varepsilon:=h+\varepsilon f$. From the fact that h is convex, it is bounded from below by an affine function, and we have $h(t)\ge(at+b)_+$ (we take the positive part since we also know $h\ge0$). It can be checked that

$$f(t) := \frac{1}{2}\sqrt{4 + (at+b)^2} + \frac{1}{2}(at+b)$$

is strictly convex and satisfies $f(t) \le \frac{1}{2}(2 + |at+b| + (at+b)) = 1 + (at+b)_+$. □

Remark 2.11. Positivity of the function h is not really necessary, as soon as μ and ν satisfy some integrability conditions so that $\inf(\mathrm{KP}) > -\infty$. For instance, if we have $\int |x| \mathrm{d}\mu(x), \int |y| \mathrm{d}\nu(y) < +\infty$, we can add an affine function to h and make it convex, and the cost of this affine function does not depend on γ and is finite (see Example 2.14 below).

Remark 2.12. We stress that a completely analogous proof could be used to prove that optimal maps for strictly concave costs (of the form $h(y-x)$ for h strictly concave) are monotone decreasing instead of monotone increasing. Compare with what is described in Section 3.3.2.

Remark 2.13. Also, the optimality of γ_{mon} is true under more general assumptions, i.e., for some costs which are not of the form $h(x-y)$ but satisfy the *twist condition*. See **Ex**(10).

Easy examples where c is convex but not strictly convex and $\mathrm{T}_{\mathrm{mon}}$ is not the unique optimal transport map may be built as follows.

Example 2.14 (Linear costs). Suppose that $c(x,y) = L(x-y)$, the map $a : \mathbb{R}^d \to \mathbb{R}$ being linear. In this case any transport plan γ is optimal and any transport map as well. This can be easily seen if one writes

$$\int L(x-y)\ \mathrm{d}\gamma = \int L(x)\ \mathrm{d}\gamma - \int L(y)\ \mathrm{d}\gamma = \int L(x)\ \mathrm{d}\mu - \int L(y)\ \mathrm{d}\nu,$$

which shows that the result does not depend on γ but only on its marginals. This general example works for μ, ν compactly supported (so that we do not have any problem of integrability of $L(x)$ and $L(y)$) and in any dimension. Hence, also in 1D.

Example 2.15 (Distance costs on the line). Suppose that $c(x,y) = |x-y|$ and that $\mu, \nu \in \mathscr{P}(\mathbb{R})$ are such that $\sup \mathrm{spt}(\mu) < \inf \mathrm{spt}(\nu)$. In this case any transport plan γ is optimal and any transport map is optimal as well. This can be seen by noting that for every $(x,y) \in \mathrm{spt}(\mu) \times \mathrm{spt}(\nu)$, we have $c(x,y) = y - x$, which is again a linear cost.

Example 2.16 (Book shifting). Consider $c(x,y) = |x-y|$, $\mu = \frac{1}{2}\mathscr{L}^1_{[0,2]}$, and $\nu = \frac{1}{2}\mathscr{L}^1_{[1,3]}$. Then $\mathrm{T}_{\mathrm{mon}}(x) = x+1$ is the monotone transport plan transporting μ onto ν. Its cost is $M(\mathrm{T}) = \int |\mathrm{T}_{\mathrm{mon}}(x) - x|\ \mathrm{d}\mu = 1$. Yet, the transport map T given by

$$\mathrm{T}(x) = \begin{cases} x+2 & \text{if } x \le 1 \\ x & \text{if } x > 1 \end{cases}$$

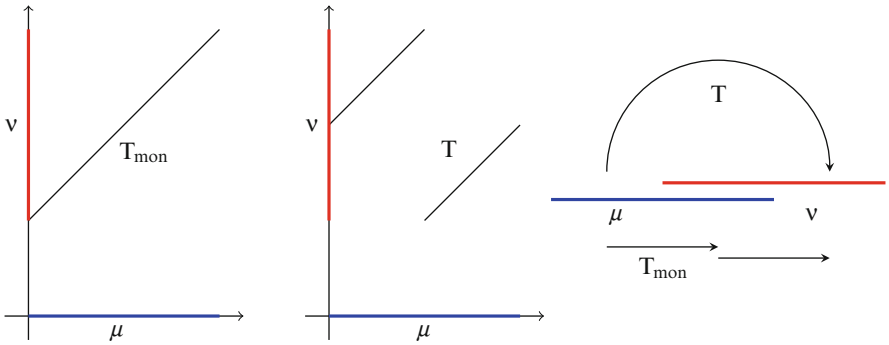


Fig. 2.1 The transport maps in the book-shifting example

also satisfies $\mathrm{T}_{\#}\mu = \nu$ and $\int |\mathrm{T}(x) - x| \, \mathrm{d}\mu = \frac{1}{2}\int_0^1 2 \, \mathrm{d}x = 1$ and is optimal as well (Figure 2.1).

Starting from the fact that the optimal transport for all these costs is the monotone one, we can express the cost for sending a given measure μ onto another measure ν in terms of their cumulative distribution functions F_μ and F_ν.

Proposition 2.17. *Given $\mu, \nu \in \mathscr{P}(\mathbb{R})$, consider the cost $c(x, y) = h(x-y)$, where h is a convex function. Then, we have*

$$\mathscr{T}_c(\mu, \nu) = \int_0^1 h(F_\nu^{[-1]} - F_\mu^{[-1]}) d\mathscr{L}^1.$$

If $h(z) = |z|$, then this also coincides with $\int_{\mathbb{R}} |F_\mu(t) - F_\nu(t)| \, \mathrm{d}t$.

Proof. The first part of the statement is just a consequence of the optimality of γ_{mon} (Theorem 2.9).

The particular case of $h(z) = |z|$ may be treated by geometric consideration: indeed, the integral $\int_0^1 \left|(F_\nu)^{-1} - (F_\mu)^{-1}\right|$ equals the area of the part of the strip $[0, 1] \times \mathbb{R}$ bounded by the graphs of $(F_\mu)^{-1}$ and $(F_\nu)^{-1}$. In order to pass from the inverse functions to the direct ones, it is enough to turn the head and look at the same strip from the variable t instead of x. But, if we want to prove it through computations, we have

$$\begin{aligned}
&\int |F_\nu^{[-1]}(x) - F_\mu^{[-1]}(x)| \, \mathrm{d}x \\
&= \mathscr{L}^2\big(\{(x,t) \in [0,1] \times \mathbb{R} : F_\mu^{[-1]}(x) \le t < F_\nu^{[-1]}(x) \text{ or } F_\nu^{[-1]}(x) \le t < F_\mu^{[-1]}(x)\}\big) \\
&\qquad = \mathscr{L}^2\big(\{(x,t) \in [0,1] \times \mathbb{R} : F_\mu^{[-1]}(x) \le t < F_\nu^{[-1]}(x)\}\big) \\
&\qquad\qquad + \mathscr{L}^2\big(\{(x,t) \in [0,1] \times \mathbb{R} : F_\nu^{[-1]}(x) \le t < F_\mu^{[-1]}(x)\}\big).
\end{aligned}$$

Then, by Fubini's theorem, we have

$$\begin{aligned}\mathscr{L}^2\big(\{(x,t)\in[0,1]\times\mathbb{R}: F_\mu^{[-1]}(x)\le t<F_\nu^{[-1]}(x)\}\big)&\\ =\int_{\mathbb{R}}\mathscr{L}^1\big(\{x\in[0,1]: x\le F_\mu(t)\text{ and }F_\nu(t)<x\}\big)\mathrm{d}t&\\ =\int_{\mathbb{R}}\mathscr{L}^1\big(\{x\in[0,1]: F_\nu(t)<x\le F_\mu(t)\}\big)\mathrm{d}t.&\end{aligned}$$

Analogously,

$$\begin{aligned}\mathscr{L}^2\big(\{(x,t)\in[0,1]\times\mathbb{R}: F_\nu^{[-1]}(x)\le t<F_\mu^{[-1]}(x)\}\big)&\\ =\int_{\mathbb{R}}\mathscr{L}^1\big(\{x\in[0,1]: F_\mu(t)<x\le F_\nu(t)\}\big)\mathrm{d}t,&\end{aligned}$$

and, summing up,

$$\begin{aligned}\mathscr{L}^2\big(\{(x,t)\in[0,1]\times\mathbb{R}: F_\mu^{[-1]}(x)\le t<F_\nu^{[-1]}(x)\text{ or }F_\nu^{[-1]}(x)\le t<F_\mu^{[-1]}(x)\}\big)&\\ =\int_{\mathbb{R}}\mathscr{L}^1\big(\{x\in[0,1]: F_\nu(t)<x\le F_\mu(t)\text{ or }F_\mu(t)<x\le F_\nu(t)\}\big)\mathrm{d}t&\\ =\int_{\mathbb{R}}|F_\mu(t)-F_\nu(t)|\mathrm{d}t.&\end{aligned}$$

This concludes the proof (Figure 2.2). □

2.3 The Knothe transport

The Knothe transport, also known as Knothe-Rosenblatt rearrangement, is a special transport map, which has a priori nothing to do with optimal transport, that may be associated to two given measures $\mu,\nu\in\mathscr{P}(\mathbb{R}^d)$. It was independently proposed

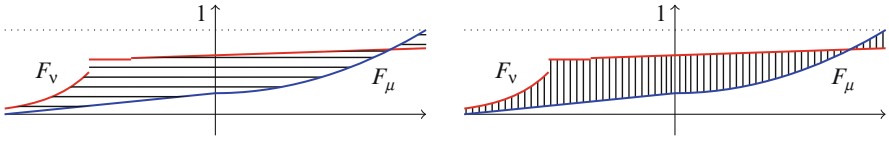


Fig. 2.2 The computation of the areas in the last part of Proposition 2.17: the integral of $|F_\nu^{[-1]}-F_\mu^{[-1]}|$ corresponds to seeing "horizontally" the graph on the left, while if seen "vertically," we get the integral of $|F_\mu-F_\nu|$

by Rosenblatt [264] for statistical purposes and by Knothe [203] for applications to geometric inequalities (see also Section 2.5.3). The main interesting point of such a map is its computational simplicity. We will explain the principle of this transport map in the simpler case where the two measures are absolutely continuous: $\mu = f(x)\,\mathrm{d}x$, $\nu = g(y)\,\mathrm{d}y$.

Let us first define the densities $\hat{f}^d$ and $\hat{g}^d$ via

$$\hat{f}^d(x_d) = \int f(t_1, t_2, \dots, t_{d-1}, x_d)\,\mathrm{d}t_1\mathrm{d}t_2 \dots \mathrm{d}t_{d-1},$$

$$\hat{g}^d(y_d) = \int g(s_1, s_2, \dots, s_{d-1}, y_d)\,\mathrm{d}s_1\mathrm{d}s_2 \dots \mathrm{d}s_{d-1}$$

as well as, for $k < d$,

$$\hat{f}^k(x_k, x_{k+1}, \dots, x_d) = \int f(t_1, \dots, t_{k-1}, x_k, \dots, x_d)\,\mathrm{d}t_1 \dots \mathrm{d}t_{k-1},$$

$$\hat{g}^k(y_k, y_{k+1}, \dots, y_d) = \int f(s_1, \dots, s_{k-1}, y_k, \dots, y_d)\,\mathrm{d}s_1 \dots \mathrm{d}s_{k-1}.$$

It is easy to check that $\hat{f}^d$ and $\hat{g}^d$ are the densities of $(\pi_d)_\#\mu$ and $(\pi_d)_\#\nu$, where $\pi_d : \mathbb{R}^d \to \mathbb{R}$ is the projection on the last variable. More generally, $\hat{f}^k$ and $\hat{g}^k$ are the densities of $\mu^k = (\pi_{k,d})_\#\mu$ and $\nu^k = (\pi_{k,d})_\#\nu$, where $\pi_{k,d} : \mathbb{R}^d \to \mathbb{R}^{d-k+1}$ is the map given by the projection onto the variables from k to d:

$$\pi_{k,d}(x_1, \dots, x_{k-1}, x_k, x_{k+1}, \dots, x_d) = (x_k, x_{k+1}, \dots, x_d).$$

Then, we define, for $k = 1, \dots, d-1$,

$$f^k(x_k, x_{k+1}, \dots, x_d) = \frac{\hat{f}^k(x_k, x_{k+1}, \dots, x_d)}{\hat{f}^{k+1}(x_{k+1}, \dots, x_d)},$$

$$g^k(y_k, y_{k+1}, \dots, y_d) = \frac{\hat{g}^k(y_k, y_{k+1}, \dots, y_d)}{\hat{g}^{k+1}(y_{k+1}, \dots, y_d)}.$$

(Note that all these functions will only be used on the set of points where the denominator does not vanish.)

The function f^k, considered as a function of x_k with parameters $(x_{k+1}, \dots, x_d)$, can be actually seen as the density of the disintegration of μ^k according to the variables $(x_{k+1}, \dots, x_d)$, and g^k is, in the same terms, the density of the disintegration of ν^k. We briefly sketch below the main notions about the disintegration of measures. With this language, the Knothe transport that we are going to define could be defined even in the case of non-absolutely continuous measures, under some assumptions on the absence of atoms. Anyway, for the sake of simplicity, we will give most of the results in the case of absolutely continuous measures.

Box 2.2. ***Important notion**: Disintegrations of measures*

Definition. Consider a measure space X endowed with a Borel measure μ and a map $f : X \to Y$ valued in a topological space Y. We say that a family $(\mu_y)_{y\in Y}$ is a disintegration of μ according to f if every μ_y is a probability measure concentrated on $f^{-1}(\{y\})$, and for every test function $\phi \in C(X)$, the map $y \mapsto \int_X \phi \, \mathrm{d}\mu_y$ is Borel measurable and

$$\int_X \phi \, \mathrm{d}\mu = \int_Y \mathrm{d}\nu(y) \int_X \phi \, \mathrm{d}\mu_y,$$

where $\nu = f_\# \mu$.

The disintegration of μ is also often written (by abuse of notation) as $\mu = \mu_y \otimes \nu$.

In the particular case where $X = Y \times Z$ and f is the projection on the Y factor, we usually identify the measures μ_y, which are "officially" defined as measures on $Y \times Z$ concentrated on $\{y\} \times Z$, with measures on Z, so that we get

$$\int_{Y\times Z} \phi(y,z) \, \mathrm{d}\mu(y,z) = \int_Y \mathrm{d}\nu(y) \int_Z \phi(y,z) \, \mathrm{d}\mu_y(z).$$

The disintegration of a measure μ exactly corresponds to the conditional law in probability. The reader is invited to consult [152] to find proofs about conditional probabilities and then to translate them into the disintegration language. Indeed, in probability we usually speak of the conditional law of a random variable X knowing $Y = y$. This means that the probability $\mathbb{P}$ on the probability space Ω is disintegrated according to the map $Y : \Omega \to E$ (where E is the image space of Y) into probabilities $\mathbb{P}_y$ and that we take the law of X under $\mathbb{P}_y$. The existence and the uniqueness of the disintegration depend on some assumptions on the spaces, but are true if $X = \mathbb{R}^d$.

In order to define the Knothe rearrangement, let us start from $k = d$ and define the transport map $\mathrm{T}^d : \mathbb{R} \to \mathbb{R}$ as the monotone nondecreasing transport map sending f^d onto g^d (these functions being considered as probability densities). We know that this map is well defined and we know how to compute it in terms of cumulative distribution functions. We will now define a family of maps $\mathrm{T}^k : \mathbb{R}^{d-k+1} \to \mathbb{R}$, where the variables of T^k will be $(x_k, x_{k+1}, \dots, x_d)$. The first one is the map T^d that we just defined. For the sake of notations, we also define some maps $\hat{\mathrm{T}}^k : \mathbb{R}^{d-k+1} \to \mathbb{R}^{d-k+1}$ given by

$$\hat{\mathrm{T}}^k(x_k, x_{k+1}, \dots, x_d) = (\mathrm{T}^k(x_k, x_{k+1}, \dots, x_d), \mathrm{T}^{k+1}(x_{k+1}, \dots, x_d), \dots, \mathrm{T}^d(x_d)).$$

Obviously T^d and $\hat{\mathrm{T}}^d$ coincide. Now, if we write $f^k_{(x_{k+1},\dots,x_d)}$ and $g^k_{(y_{k+1},\dots,y_d)}$ for the functions $x_k \mapsto f^k(x_k, x_{k+1}, \dots, x_d)$ and $y_k \mapsto g^k(y_k, y_{k+1}, \dots, y_d)$ and we interpret them as densities of probability measures, we can define the map T^k, if we suppose that we have already defined T^j for $j > k$, in the following way: take $\mathrm{T}^k_{(x_{k+1},\dots,x_d)}$ to be the monotone nondecreasing transport map sending $f^k_{(x_{k+1},\dots,x_d)}$ onto $g^k_{\hat{\mathrm{T}}^{k+1}(x_{k+1},\dots,x_d)}$.

Finally, the Knothe-Rosenblatt rearrangement T is defined by $\mathrm{T} = \hat{\mathrm{T}}^1$.

We want to check that this map T is a transport map from μ to ν.

Proposition 2.18. *The Knothe-Rosenblatt rearrangement map* T *defined above satisfies* $\mathrm{T}_\#\mu = \nu$.

Proof. We will prove by induction that $\hat{\mathrm{T}}^k$ satisfies

$$\hat{\mathrm{T}}^k_\#\mu^k = \nu^k \tag{2.4}$$

(let us recall that $\mu^k = (\pi_{k,d})_\#\mu$ and $\nu^k = (\pi_{k,d})_\#\nu$ are the marginals of μ and ν onto the last $d-k+1$ variables). Equality (2.4) fact is evident by construction for $k = d$, and if we get it for $k = 1$, we have proven $\mathrm{T}_\#\mu = \nu$.

We only need to prove that if the claim is true for a given $k+1$, then it will be true for k. To check equality (2.4), we just need to use test functions $\chi(y_k, \dots, y_d)$ and check

$$\int \chi(\mathrm{T}^k(x_k, x_{k+1}, \dots, x_d), \mathrm{T}^{k+1}(x_{k+1}, \dots, x_d), \dots, \mathrm{T}^d(x_d))\, \mathrm{d}\mu = \int \chi(y_k, \dots, y_d)\ \mathrm{d}\nu.$$

Moreover, by density (see below for Stone-Weierstrass's theorem), it is enough to check the equality above on functions χ which are separable, i.e., of the form $\chi(y_k, \dots, y_d) = \phi(y_k)\psi(y_{k+1}, \dots, y_d)$. To ease the notation, for fixed k, we will denote by $\bar{x}$ the vector $\pi_{k+1,d}(x)$ and $\bar{y} = \pi_{k+1,d}(y)$. In this case we should check

$$\int \psi \circ \hat{\mathrm{T}}^{k+1} \left(\int \phi \circ \mathrm{T}^k f^k \, \mathrm{d}x_k \right) \hat{f}^{k+1} \, \mathrm{d}\bar{x} = \int \psi(\bar{y}) \left(\int \phi(y_k)\, g^k \, \mathrm{d}y_k \right) \hat{g}^{k+1} \, \mathrm{d}\bar{y}.$$

In order to get this equality, we just need to use the definition of T^k, so that we get, for every x,

$$\int \phi \circ \mathrm{T}^k f^k \, \mathrm{d}x_k = \int \phi(y_k)\, g^k_{\hat{\mathrm{T}}^{k+1}(\bar{x})}(y_k) \, \mathrm{d}y_k.$$

If we define $G : \mathbb{R}^{d-k} \to \mathbb{R}$ the function given by $G(\bar{y}) = \int \phi(y_k)\, g^k_{\bar{y}}(y_k) \, \mathrm{d}y_k$, we have now

$$\int \psi \circ \hat{\mathrm{T}}^{k+1} \left(\int \phi \circ \mathrm{T}^k f^k \, \mathrm{d}x_k \right) \hat{f}^{k+1} \, \mathrm{d}\bar{x} = \int \psi \circ \hat{\mathrm{T}}^{k+1}\, G \circ \hat{\mathrm{T}}^{k+1} \hat{f}^{k+1} \, \mathrm{d}\bar{x}.$$

By taking the image measure of $\hat{f}^{k+1}$ under $\hat{\mathrm{T}}^{k+1}$, the last expression is equal to $\int \psi(\bar{y}) G(\bar{y}) \hat{g}^{k+1}(\bar{y}) \, \mathrm{d}\bar{y}$, which is in turn equal to

$$\int \psi(\bar{y}) \left(\int \phi(y_k)\, g^k_{\bar{y}}(y_k) \, \mathrm{d}y_k \right) \hat{g}^{k+1}(\bar{y}) \, \mathrm{d}\bar{y}$$

and proves the claim. □

Box 2.3. ***Memo**: Stone-Weierstrass's Theorem*

Theorem. Suppose that X is a compact space and that $E \subset C(X)$ is a subset of the space of continuous functions on X satisfying

i) constant functions belong to E
ii) E is an algebra, i.e., it is stable by sum and product: $f, g \in E \Rightarrow f + g, fg \in E$
iii) E separates the points of X, i.e., for all $x \neq y \in X$, there is $f \in E$ such that $f(x) \neq f(y)$.

Then E is dense in $C(X)$ for the uniform convergence (see, for instance, [268]).

This theorem is usually evoked to justify the density of polynomials when $X \subset \mathbb{R}^d$. Trigonometric polynomials are also dense for the same reason (and, by the way, this may be used to prove the L^2 convergence of the Fourier series). In product spaces $X = Y \times Z$, linear combinations of separable functions of the form $\sum_i \varphi_i(y)\psi_i(z)$ are dense because of this very same theorem.

We just proved that the map T is a transport map between μ and ν. Moreover, by construction, should it be a regular map, the Knothe transport T has a triangular Jacobian matrix with nonnegative entries on its diagonal. Compare this with the Brenier optimal map for the quadratic cost: its Jacobian matrix is the Hessian matrix of a convex function, which is symmetric (instead of triangular) and has also positive eigenvalues. Moreover, we underline that assuming absolute continuity as we did, all the monotone optimal transports that we use in defining T are invertible, since not only the source measures are atomless but also the arrival ones are atomless (they are all absolutely continuous). As a consequence, each map $\hat{\mathrm{T}}^k$ is also invertible.

Remark 2.19. It is interesting to note that, due to its explicit definition, the Knothe map automatically inherits some regularity properties from the measures μ and ν. In general, T has the same regularity as μ and ν and not better. Indeed, each T^k is built as a monotone transport, which gives one extra derivative than the regularity of μ and ν (see Remark 2.7), but only in the direction of the variable x_k. With respect to the other variables, the explicit formula of monotone transports with cumulative distribution functions allows us to prove that the dependence on $(x_{k+1}, \dots, x_d)$ is as regular as the densities are.

Remark 2.20. It is also interesting to note that the Knothe map defined above is the (unique) transport map from μ to ν which is monotone for the lexicographic order, i.e., the order where $x < y$ is defined in the following way: "there exists $k \in \{1, 2, \dots, d\}$ such that $x_j = y_j$ for $j > k$ and $x_k < y_k$."

Remark 2.21. How to define the Knothe transport for μ and ν which are not absolutely continuous? The construction uses the notion of disintegration of measures (Box 2.2).

First we define the transport map $\mathrm{T}^d : \mathbb{R} \to \mathbb{R}$ as the monotone nondecreasing transport map sending $(\pi_d)_\# \mu$ onto $(\pi_d)_\# \nu$. We need to suppose that $(\pi_d)_\# \mu$ has no atoms. Then, for each k we consider $(\pi_{k,d})_\# \mu$ and we disintegrate it with respect

to $(\pi_{k+1,d})_\#\mu$, and we do the same for $(\pi_{k,d})_\#\nu$ with $(\pi_{k+1,d})_\#\nu$. We obtain some families of probabilities on the real line that we can call $\mu^k_{x_{k+1},\dots,x_n}$ and $\nu^k_{y_{k+1},\dots,y_n}$, respectively. They are measures on the variables x_k and y_k and they replace the densities $f^k(\cdot, x_{k+1},\dots,x_d)$ and $g^k(\cdot, y_{k+1},\dots,y_d)$. Again, we need to suppose that each measure $\mu^k_{x_{k+1},\dots,x_n}$ is atomless. The maps $\hat{\mathrm{T}}^k(x_k, x_{k+1},\dots,x_d)$ are defined as before, using 1D monotone increasing maps.

2.4 Knothe transport as a limit of quadratic optimal transports

Let us slightly modify the quadratic cost that we discussed in Section 1.3.1 and replace it with the weighted quadratic cost

$$c_\varepsilon(x,y) := \sum_{i=1}^d \lambda_i(\varepsilon)(x_i - y_i)^2$$

where the $\lambda_i(\varepsilon)$s are positive scalars depending on a parameter $\varepsilon > 0$. If μ is absolutely continuous with respect to the Lebesgue measure, the corresponding optimal transportation problem admits a unique solution T_ε. For simplicity, in this section we will only consider the case where the supports of μ and ν are compact (but everything could be adapted to the case $\int |x|^2 \mathrm{d}\mu, \int |y|^2 \mathrm{d}\nu < +\infty$).

When in addition, for all $k \in \{1,\dots,d-1\}$, $\lambda_k(\varepsilon)/\lambda_{k+1}(\varepsilon) \to 0$ as $\varepsilon \to 0$, it is natural to expect the convergence of T_ε to the Knothe transport T. This convergence result was conjectured by Y. Brenier as a very natural one.

We will show it under the absolute continuity assumption of the previous section. On the other hand, [114] proves the same result under a slightly weaker assumption, namely, the absence of atoms in all the disintegrated measures. The curios point is that absence of atoms is also needed on the target measure ν and that a counterexample to the convergence result exists if it is not satisfied.

Example 2.22 (Explicit computations for Gaussians). To illustrate the problem in a particular case where explicit solutions are available, take $d = 2$, and μ and ν two Gaussian measures where $\mu = N(0, I)$ and $\nu = N\left(0, \begin{pmatrix} a & b \\ b & c \end{pmatrix}\right)$ (with $ac > b^2$, $a > 0$; the notation $N(a, A)$ stands for the Gaussian density with covariance matrix A, centered at a). Take $\lambda_1(\varepsilon) = \varepsilon$ and $\lambda_2(\varepsilon) = 1$. Then it can be verified (see also **Ex**(11)) that T_ε is linear and that its matrix in the canonical basis of $\mathbb{R}^2$ is

$$\mathrm{T}_\varepsilon = \frac{1}{\sqrt{a\varepsilon^2 + c + 2\varepsilon\sqrt{ac - b^2}}} \begin{pmatrix} a\varepsilon + \sqrt{ac - b^2} & b \\ b\varepsilon & c + \varepsilon\sqrt{ac - b^2} \end{pmatrix}$$

which converges as $\varepsilon \to 0$ to $\mathrm{T} = \begin{pmatrix} \sqrt{a - b^2/c} & b/\sqrt{c} \\ 0 & \sqrt{c} \end{pmatrix}$, i.e., the matrix of the Knothe transport from μ to ν.

We directly state our first result, whose proof, in the spirit of Γ-convergence developments (see [77] and Box 4.6), will require several steps.

Theorem 2.23. *Let μ and ν be two absolutely continuous probability measures on $\mathbb{R}^d$ with compact supports and $\gamma_\varepsilon \in \Pi(\mu, \nu)$ be an optimal transport plan for the costs $c_\varepsilon(x, y) = \sum_{i=1}^d \lambda_i(\varepsilon)(x_i - y_i)^2$, for some weights $\lambda_k(\varepsilon) > 0$. Suppose that for all $k \in \{1, \dots, d-1\}$, $\lambda_k(\varepsilon)/\lambda_{k+1}(\varepsilon) \to 0$ as $\varepsilon \to 0$. Let T be the Knothe-Rosenblatt map between μ and ν and $\gamma_K \in \mathscr{P}(\mathbb{R}^d \times \mathbb{R}^d)$ the associated transport plan (i.e., $\gamma_K := (\mathrm{id}, \mathrm{T})\#\mu$). Then $\gamma_\varepsilon \rightharpoonup \gamma_K$ as $\varepsilon \to 0$.*

Moreover, the plans γ_ε are induced by transport maps T_ε, and these maps converge to T in $L^2(\mu)$ as $\varepsilon \to 0$.

The proof will roughly stick to the following strategy. We will take a limit point of γ_ε and show that it shares the same (x_d, y_d) marginal as γ_K. Then, we will disintegrate with respect to (x_d, y_d) and prove that the conditional (x_{d-1}, y_{d-1})-marginals coincide. We will end the proof by iterating this argument.

Proof. Without loss of generality, we assume $\lambda_d(\varepsilon) = 1$ and $\lambda_i(\varepsilon)/\lambda_{i+1}(\varepsilon) \to 0$.

Take the optimal plans γ_ε and suppose (which is possible, up to subsequences) $\gamma_\varepsilon \rightharpoonup \gamma$. We want to prove $\gamma = \gamma_K$. From the uniqueness of the Knothe transport, the convergence will hold on the full sequence and not only on subsequences.

By comparing γ_ε to γ_K and using optimality, we first get

$$\int c_\varepsilon \, \mathrm{d}\gamma_\varepsilon \leq \int c_\varepsilon \, \mathrm{d}\gamma_K \tag{2.5}$$

and, passing to the limit as $\varepsilon \to 0$, since c_ε converges uniformly to $c^{(d)}(x, y) = (x_d - y_d)^2$ (we will use the general notation $c^{(k)}(x, y) = |x_k - y_k|^2$), we get

$$\int c^{(d)} \, \mathrm{d}\gamma \leq \int c^{(d)} \, \mathrm{d}\gamma_K.$$

Yet, the function $c^{(d)}$ only depends on the variables x_d and y_d and this shows that the measure $(\pi_d^{x,y})_\# \gamma$ gets a result at least as good as $(\pi_d^{x,y})_\# \gamma_K$ with respect to the quadratic cost ($\pi_d^{x,y}$ being the projection onto the last coordinates of x and y, i.e., $\pi_d^{x,y}(x, y) = (x_d, y_d)$). Yet, the measure γ_K has been chosen on purpose to be optimal from μ^d to ν^d for this cost, and this optimal transport plan is unique. Then $(\pi_d^{x,y})_\# \gamma = (\pi_d^{x,y})_\# \gamma_K$. Let us call γ^d this common measure.

We go back to (2.5) and go on by noting that all the measures γ_ε have the same marginals as γ_K and hence their (separate) projections onto x_d and y_d are μ^d and ν^d, respectively. This implies that $(\pi_d^{x,y})_\# \gamma_\varepsilon$ must realize a result which is not better than $(\pi_d^{x,y})_\# \gamma_K$ as far as the 1D quadratic cost is concerned. Consequently, we have

$$\int |x_d - y_d|^2 \mathrm{d}(\pi_d^{x,y})_\# \gamma_K(x_d, y_d) + \sum_{i=1}^{d-1} \lambda_i(\varepsilon) \int (x_i - y_i)^2 \,\mathrm{d}\gamma_\varepsilon$$
$$\leq \int c_\varepsilon \,\mathrm{d}\gamma_\varepsilon \leq \int c_\varepsilon \,\mathrm{d}\gamma_K$$
$$= \int |x_d - y_d|^2 \mathrm{d}(\pi_d^{x,y})_\# \gamma_K(x_d, y_d) + \sum_{i=1}^{d-1} \lambda_i(\varepsilon) \int (x_i - y_i)^2 \,\mathrm{d}\gamma_K,$$

which implies, by simplifying the common term, dividing by $\lambda_{d-1}(\varepsilon)$, and passing to the limit,

$$\int c^{(d-1)} \,\mathrm{d}\gamma \leq \int c^{(d-1)} \,\mathrm{d}\gamma_K. \tag{2.6}$$

We can note that both integrals depend on the variables x_{d-1} and y_{d-1} only. Anyway, we can project onto the variables (x_{d-1}, x_d) and (y_{d-1}, y_d) (obtaining measures $(\pi_{d-1,d}^{x,y})_\#\gamma$ and $(\pi_{d-1,d}^{x,y})_\#\gamma_K$) so that we disintegrate with respect to the measure γ^d. We have

$$\int \mathrm{d}\gamma^d(x_d, y_d) \int |x_{d-1} - y_{d-1}|^2 \,\mathrm{d}\gamma_{(x_d,y_d)}^{d-1}(x_{d-1}, y_{d-1})$$
$$\leq \int \mathrm{d}\gamma^d(x_d, y_d) \int |x_{d-1} - y_{d-1}|^2 \,\mathrm{d}\gamma_{(x_d,y_d),K}^{d-1}(x_{d-1}, y_{d-1}). \tag{2.7}$$

We want to prove that the measures $\gamma_{(x_d,y_d)}^{d-1}$ share the same marginals on x_{d-1} and y_{d-1} of the corresponding $\gamma_{(x_d,y_d),K}^{d-1}$. Should this be the case, their quadratic cost would be not better than the corresponding cost of $\gamma_{(x_d,y_d),K}^{d-1}$ (because the Knothe measure has been chosen exactly with the intention of being quadratically optimal on (x_{d-1}, y_{d-1}) once x_d and y_d are fixed). Yet, (2.7) shows that, on average, the result given by those measures is not worse than the results of the optimal ones. Thus, the two results coincide for almost any pair (x_d, y_d), and by uniqueness of the optimal transports in the 1D case, we get $\gamma_{(x_d,y_d)}^{d-1} = \gamma_{(x_d,y_d),K}^{d-1}$. To let this proof work, it is sufficient to prove that the marginals of the two measures coincide for γ^d-a.e. pair (x_d, y_d). For fixed (x_d, y_d), we would like to prove, for any ϕ,

$$\int \phi(x_{d-1}) \,\mathrm{d}\gamma_{(x_d,y_d)}^{d-1} = \int \phi(x_{d-1}) \,\mathrm{d}\gamma_{(x_d,y_d),K}^{d-1}$$

(and to prove an analogous equality for functions of y_{d-1}). Since we want to prove it for a.e. pair (x_d, y_d), it is sufficient to prove this equality:

$$\int \mathrm{d}\gamma^d(x_d,y_d)\psi(x_d,y_d)\int \phi(x_{d-1})\,\mathrm{d}\gamma^{d-1}_{(x_d,y_d)}$$
$$=\int \mathrm{d}\gamma^d(x_d,y_d)\psi(x_d,y_d)\int \phi(x_{d-1})\,\mathrm{d}\gamma^{d-1}_{(x_d,y_d),K}$$

for any ϕ and any ψ. This means proving

$$\int \psi(x_d,y_d)\phi(x_{d-1})\,\mathrm{d}\gamma^{d-1}=\int \psi(x_d,y_d)\phi(x_{d-1})\,\mathrm{d}\gamma^{d-1}_K,$$

which is not trivial. As far as now, we only know that the two measures γ^{d-1} and γ^{d-1}_K have the same marginals with respect to the pairs (x_{d-1},x_d), (y_{d-1},y_d) (since they have the same projections onto x and onto y) and (x_d,y_d) (since we just proved it). But here there is a function of the three variables (x_{d-1},x_d,y_d). Yet, we know that the measure γ^d is concentrated on the set $y_d=\mathrm{T}_d(x_d)$ for a certain map T_d (here we use that the optimal plan is indeed a map), and this allows us to replace the expression of y_d, thus getting rid of one variable. This proves that the function $\psi(x_d,y_d)\phi(x_{d-1})$ is actually a function of (x_{d-1},x_d) only and that equality holds when passing from γ to γ_K. The same can be performed on functions $\psi(x_d,y_d)\phi(y_{d-1})$, but we have in this case to ensure that we can replace x_d with a function of y_d, i.e., that we can invert T_d. This is possible thanks to the absolute continuity of ν^d, since T_d is the optimal transport from μ^d to ν^d, but an optimal transport exists in the other direction as well and it gives the same optimal plan (thanks to uniqueness). These facts prove that the measures $\gamma^{d-1}_{(x_d,y_d)}$ and $\gamma^{d-1}_{(x_d,y_d),K}$ have the same marginals and hence, since they are both optimal, they coincide for a.e. pair (x_d,y_d). This implies $(\pi^{x,y}_{d-1,d})_\#\gamma=(\pi^{x,y}_{d-1,d})_\#\gamma_K$, and we will call this common measure γ^{d-1}.

In the case $d=2$, the proof would be finished. Now, it is possible to go on by induction in the following way: we prove, by induction on N, that the statement of the theorem is true if $d\le N$ and that, whatever the dimension is, we have $(\pi^{x,y}_{k,d})_\#\gamma=(\pi^{x,y}_{k,d})_\#\gamma_K$ (i.e., that we have convergence in what concerns the components from the kth to the dth) for every $k\ge d-(N-1)$. Both facts are proven for $N=2$.

Suppose the result proven for a certain value of N. Take $d>N$ and $k=d-(N-1)$. We want to prove $(\pi^{x,y}_{k-1,d})_\#\gamma=(\pi^{x,y}_{k-1,d})_\#\gamma_K$. If we can do it, we are done, since this also proves that the result is true in dimension $d=N+1$ (since in this case we have $k=2$ and with our extra step we conclude the proof).

Write $x=(x^-,x^+)$ and $y=(y^-,y^+)$ with $x^-,y^-\in\mathbb{R}^{k-1}$ and $x^+,y^+\in\mathbb{R}^{d-k+1}$ (decomposing into the components up to $k-1$ and after k). Then, define a competitor γ^ε_K in the following way: γ^ε_K is a measure on $\mathbb{R}^d\times\mathbb{R}^d$ with the following marginals:

$$(\pi_{x^-,x^+})_\#\gamma^\varepsilon_K=\mu,\ (\pi_{y^-,y^+})_\#\gamma^\varepsilon_K=\nu,\ (\pi_{x^+,y^+})_\#\gamma^\varepsilon_K=\eta_{\varepsilon,k},$$

where $\eta_{\varepsilon,k}$ is the measure which optimizes the transport cost $\sum_{i=k}^d\lambda_i(\varepsilon)c^{(i)}$ between the marginals $\mu^k=(\pi_{x^+})_\#\mu$ and $\nu^k=(\pi_{y^+})_\#\nu$. Thanks to the recursive structure of the proof, we know that $\eta_{\varepsilon,k}$ converges to $(\pi_{x^+,y^+})_\#\gamma_K$ (since we suppose the result to be true in dimension N).

We need to specify how to build the measure γ_K^ε (which is likely not to be unique). One easy way to construct such a measure is the following: disintegrate μ and ν according to π_{x^+} and π_{y^+}, respectively, thus getting two family of measures μ^{x^+} and ν^{y^+}. Then, for every pair (x^+, y^+) pick a transport plan γ^{x^+,y^+} between them. The measure $\gamma^{x^+,y^+} \otimes \eta_{\varepsilon,k}$ satisfies the required properties. Moreover, we will choose for γ^{x^+,y^+} the Knothe transport between those two measures. Indeed, for the sake of this step of the proof, the only important point is that, in what concerns the $(k-1)$th component, we choose for any (x^+, y^+) the monotone map from $\mu_{k-1}^{x^+} := (\pi_{x_{k-1}})_\# \mu^{x^+}$ to $\nu_{k-1}^{y^+} := (\pi_{y_{k-1}})_\# \nu^{y^+}$, so that $\int c^{(k-1)} \, \mathrm{d}\gamma^{x^+,y^+} = \mathscr{T}_c(\mu_{k-1}^{x^+}; \nu_{k-1}^{y^+})$ (in this last expression $\mathscr{T}_c$ denotes the quadratic transport cost on the real line).

With this choice in mind, we write

$$\begin{aligned}
&\sum_{i\geq h} \lambda_i(\varepsilon) \int |x_i - y_i|^2 \, \mathrm{d}\gamma_K^\varepsilon + \sum_{i=1}^{k-1} \lambda_i(\varepsilon) \int (x_i - y_i)^2 \, \mathrm{d}\gamma_\varepsilon \\
&\leq \int c_\varepsilon \, \mathrm{d}\gamma_\varepsilon \leq \int c_\varepsilon \, \mathrm{d}\gamma_K^\varepsilon \\
&= \sum_{i\geq h} \lambda_i(\varepsilon) \int |x_i - y_i|^2 \, \mathrm{d}\gamma_K^\varepsilon + \sum_{i=1}^{k-1} \lambda_i(\varepsilon) \int (x_i - y_i)^2 \, \mathrm{d}\gamma_K^\varepsilon,
\end{aligned}$$

and consequently by erasing the common terms and dividing by $\lambda_{k-1}(\varepsilon)$, we get

$$\int c^{(k-1)} \, \mathrm{d}\gamma_\varepsilon \leq \int c^{(k-1)} \, \mathrm{d}\gamma_K^\varepsilon = \int \mathscr{T}_c(\mu_{k-1}^{x^+}; \nu_{k-1}^{y^+}) d\eta_{\varepsilon,h}.$$

Then we pass to the limit as $\varepsilon \to 0$. The left-hand side of the last inequality tends to $\int c^{(k-1)} \, \mathrm{d}\gamma$, while the right-hand side converges to

$$\int \mathscr{T}_c(\mu_{k-1}^{x^+}; \nu_{k-1}^{y^+}) \mathrm{d}(\pi_{x^+,y^+})_\# \gamma_K = \int c^{(k-1)} \, \mathrm{d}\gamma_K.$$

This convergence result is justified by the following Corollary 2.24 of Lemma 1.8. In the end we get

$$\int c^{(k-1)} \, \mathrm{d}\gamma \leq \int c^{(k-1)} \, \mathrm{d}\gamma_K$$

and we can go on. This last inequality was obtained thanks to a procedure which was similar to what we did for getting (2.6), but it required some more work since γ_K is not exactly optimal for all the components $i \geq k$ as it was the case for $k = d$.

We disintegrate with respect to $\pi_{k,d}^{x,y}$ and we act exactly as before: proving that the marginals of the disintegrations coincide are sufficient to prove equality of the measures. Here we will use test functions of the form

$$\psi(x_h, x_{k+1}, \dots, x_d, y_h, y_{k+1}, \dots, y_d)\varphi(x_{k-1})$$

and

$$\psi(x_h, x_{k+1}, \dots, x_d, y_h, y_{k+1}, \dots, y_d)\varphi(y_{k-1}).$$

The same trick as before, i.e., replacing the variables y with functions of the variables x, is again possible. To invert the trick and replace x with y, one needs to invert part of Knothe's transport. This is possible since, as we noted, our assumptions implied that all the monotone transports we get are invertible. In the end we get, as before, $\gamma^{k-1} = \gamma_K^{k-1}$. This proves that we can move from step N to step $(N+1)$ in the induction and allows us to conclude.

We have now proven $\gamma_\varepsilon \rightharpoonup \gamma_K$. Yet, if all these transport plans come from transport maps, then $(\mathrm{id}, \mathrm{T}_\varepsilon)_\# \mu \rightharpoonup (\mathrm{id}, \mathrm{T})_\# \mu$ implies $\mathrm{T}_\varepsilon \to \mathrm{T}$ in $L^2(\mu)$ (see Lemma 2.25 below). □

In the proof we used the following result, which is a corollary of Lemma 1.8.

Corollary 2.24. *If μ and ν are two probabilities over $\mathbb{R}^d = \mathbb{R}^k \times \mathbb{R}^{d-k}$ with compact support which disintegrate as $\mu = \mu^+ \otimes \mu^{x^+}$ and $\nu = \nu^+ \otimes \nu^{y^+}$ and c is a continuous cost, then the functional*

$$\eta \mapsto \int \mathscr{T}_c(\mu^{x^+}, \nu^{y^+})\, \mathrm{d}\eta(x^+, y^+)$$

is continuous over the transport plans between μ^+ and ν^+.

Proof. It is sufficient to apply Lemma 1.8 to the following setting: $X = Y = \mathbb{R}^k$, $\tilde{X} = \tilde{Y} = \mathscr{P}(\Omega)$, where $\Omega \subset \mathbb{R}^{d-k}$ is such that $\mathrm{spt}(\mu), \mathrm{spt}(\nu) \subset \mathbb{R}^k \times \Omega$. The space $\mathscr{P}(\Omega)$ is endowed with any distance metrizing weak convergence and making this space a separable metric space (see, for instance, Chapter 5 for a choice of distances with this property, or build it with standard methods for the weak convergence in the unit ball of a Banach space). For the cost function which appears in the statement of Lemma 1.8, we take $(a, b) \mapsto \mathscr{T}_c(a, b)$ and $a(x^+) := \mu^{x^+}$, $b(y^+) := \nu^{y^+}$. Since this quantity is bounded by $\max |c|$ (as we are on compact supports), there is no difficulty in choosing two bounded functions f, g so that $\mathscr{T}_c(a, b) \le f(a) + g(b)$. □

Another easy lemma that we needed is the following.

Lemma 2.25. *Given $\mu \in \mathscr{P}(X)$ and $\nu_n \in \mathscr{P}(\Omega)$, where $\Omega \subset \mathbb{R}^d$ is a compact set, suppose that $\gamma_n := \gamma_{\mathrm{T}_n} \in \Pi(\mu, \nu_n)$ is such that $\gamma_n \to \gamma = \gamma_{\mathrm{T}}$. Then $\mathrm{T}_n \to \mathrm{T}$ in $L^2(X, \mu)$.*

Proof. Since the functions T_n are valued in Ω, they are equibounded and they admit a weak limit in $L^2(X, \mu)$ up to subsequences. Let us call S this limit. To prove $S = \mathrm{T}$, take a continuous function $\xi \in C(X; \mathbb{R}^d)$ and consider the limit of

$$\int (\xi(x)\cdot S(x))\mathrm{d}\mu(x) = \lim_n \int (\xi(x)\cdot \mathrm{T}_n(x))\mathrm{d}\mu(x) = \lim_n \int (\xi(x)\cdot y)\ \mathrm{d}\gamma_n(x,y)$$
$$= \int (\xi(x)\cdot y)\ \mathrm{d}\gamma_{\mathrm{T}}(x,y) = \int (\xi(x)\cdot \mathrm{T}(x))\mathrm{d}\mu(x).$$

This proves $S = \mathrm{T}$.

In order to prove strong convergence, just consider

$$\int |\mathrm{T}_n|^2\,\mathrm{d}\mu = \int |y|^2\,\mathrm{d}\gamma_n \to \int |y|^2\,\mathrm{d}\gamma = \int |\mathrm{T}|^2\,\mathrm{d}\mu.$$

This proves that we do not only have weak convergence in L^2 but also convergence of the L^2 norm. This implies strong convergence, as usual in Hilbert spaces. □

Note that the above proof also works in other spaces L^p and that the boundedness assumption (i.e., compact supports) could be replaced by suitable integrability bounds.

Let us remark here that if instead of considering the quadratic cost c_ε, one considers the more general separable cost

$$c_\varepsilon(x,y) := \sum_{i=1}^d \lambda_i(\varepsilon)c_i(x_i - y_i)$$

where each c_i is a smooth strictly convex function (with suitable growth), then the previous convergence proof carries over.

Example 2.26 (A counterexample when the measures have atoms). We now show that interestingly, the absolute continuity hypothesis in Theorem 2.23 is necessary not only for μ but also for ν (see Remark 2.21 to check the definition of the Knothe map in such a case). We propose a very simple example in $\mathbb{R}^2$ where μ is absolutely continuous with respect to the Lebesgue measure but ν is not, and we show that the conclusion of the theorem fails. On the square $\Omega := [-1,1]\times[-1,1]$, define $\mu = \frac{1}{2}\mathbb{1}_E \cdot \mathscr{L}^2$, where $E = \{(x_1,x_2) \in [-1,1]\times[-1,1] \ :\ x_1x_2 < 0\}$, so that the measure μ is uniformly spread on the upper left and the lower right quadrants, and $\nu = \frac{1}{2}\mathscr{H}^1_{|S}$, S being the segment $[-1,1]\times\{0\}$.

The Knothe-Rosenblatt map is easily computed as

$$\mathrm{T}(x) := (2x_1 + \mathrm{sign}(x_2)), 0)$$

(indeed, the measure ν^2 is a Dirac mass at 0; hence $\mathrm{T}^2(x) = 0$). The solution of any transportation problem with cost $|x_1 - y_1|^2 + \varepsilon|x_2 - y_2|^2$ is $\mathrm{T}^0(x) := (x_1, 0)$ (no transport may do better than this one, which projects on the support of ν). Therefore, in this example the optimal transportation maps fail to tend to the Knothe-Rosenblatt map. The reason is the atom in the measure $\nu^2 = \delta_0$ (Figure 2.3).

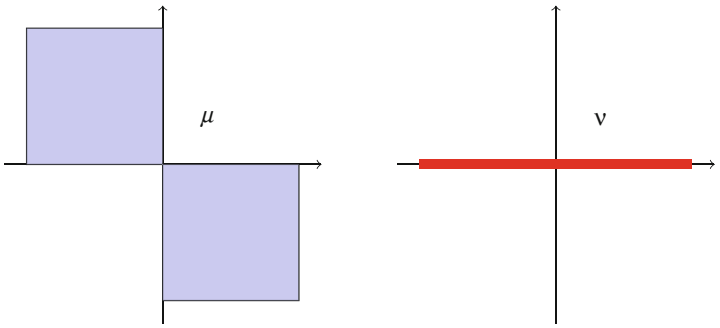


Fig. 2.3 The measures μ and ν in the counterexample

2.5 Discussion

2.5.1 Histogram equalization

A very typical problem where we encounter monotone transport maps in image processing is histogram equalization. The setting is easy to explain: take a black-and-white digital picture, and consider all possible grey levels; these levels are numbered from 0 to 255; for each $i = 0, \ldots, 255$, count how many pixels have exactly such a grey level and make a histogram of that. In some situations, it could be interesting to transform the given picture into a different one, which should represent again the same subject, but have a different histogram. Why? A typical example is the following: you have two pictures A and B which are supposed to represent the same object, with minor modifications (something has moved), at different instants of time; you want to compare them, but B happens to be much darker than A (this means that the histogram of B is much more concentrated on the right-hand part of the graph than that of A). Then, before comparing them, one prefers to transform B into a new picture B', where the darkness level of B' is comparable to that of A, i.e., the histogram of B' is equal to that of A. Similarly, one can also consider the case where A and B represent two parts of a same object, and the two pictures must be merged, but it could happen that the two partial pictures have been taken from different angles, and the light is different.

In some other cases, we have only one picture, but its histogram is too much concentrated around one peak. This means that pixels have different grey levels, but not so different. In order to see better what is inside the picture, one would like to enhance the contrast, which corresponds to using a more widespread histogram. In such a case one fixes a target histogram, for instance, the uniform one, where every level is used by the same number of pixels. This is the best possible histograms in the sense that we do not waste information by concentrating too much around a same grey level.

In all these cases, we want to perform a transformation which does not act pixel by pixel, but grey level by grey level: we look for a map T from the set X of grey levels to itself (all pixels with grey level x will be turned into the grey level $\mathrm{T}(x)$), and we want to prescribe the histogram of the image after the transformation. This is the same as saying that we have a measure μ on X and we want to impose a condition of the form $\mathrm{T}_{\#}\mu = \nu$. Here the important fact is that the space X is a subset of $\mathbb{R}$, i.e., it is 1D and ordered. The natural condition on T is that it should be monotone increasing. This condition translates the fact that we do not want to invert darker and brighter pixels.

Up to the fact that, in this discrete setting (the grey level only takes value in a finite set of integers), the two measures μ and ν are atomic (and hence a transport map could not exist), the situation is exactly the one described in this chapter. The problem of nonexistence of transport maps could be solved in several ways in practice: one could introduce some rounding rules or use a transport plan which corresponds to use different colors on pixels which had the same colors on the original picture, i.e., splitting the mass.

The situation becomes trickier when we take color images. In this case the space X of colors becomes 3D and the notion of monotone transport does not make sense any more. Obviously the quadratic optimal transport (Brenier's map) could play the role of a monotone map, since it is monotone in the vector sense $(\mathrm{T}(x) - T(y)) \cdot (x - y) \geq 0$. On the other hand, it is expensive to compute (we will see in Chapter 6 some numerical methods to do it and the associated difficulties) and some of its features are probably useless for this kind of application in image processing (the fact that it is a gradient, for instance).

Several solutions are now a matter of current research: looking for transport maps which are more regular but not so far from the optimal one [163], or producing maps defined from the composition/superposition of several 1D bricks, so that they keep some monotonicity properties and are easier to compute (see [252] and Section 2.5.2).

Also, in some cases all the components of the color space are not equally important, and one could decide to look only at one or two components or to treat them separately. In this case, colors are often described by different sets of parameters.

See, for instance, [140], where some applications of a 1D analysis on the circle are proposed for *hue* component of the color space (with the HSL decomposition into hue, saturation, and lightness).

The same space of colors can also be written as the combination of a 1D information (the *luma* signal, corresponding to a weighted average of the three components R, G, and B and standing for the brightness of the image) and of a 2D information (called *chrominance* or *chroma*, standing for the color information). In Figure 2.4 we see an application of optimal transport in 1D and 2D to the color transfer between two streets with very different colors.

For a wider discussion about 1D histogram equalization and 3D color transfer in connection with optimal transport, we refer to [163] and [255] and the references therein (in particular some older papers such as [242, 259]).

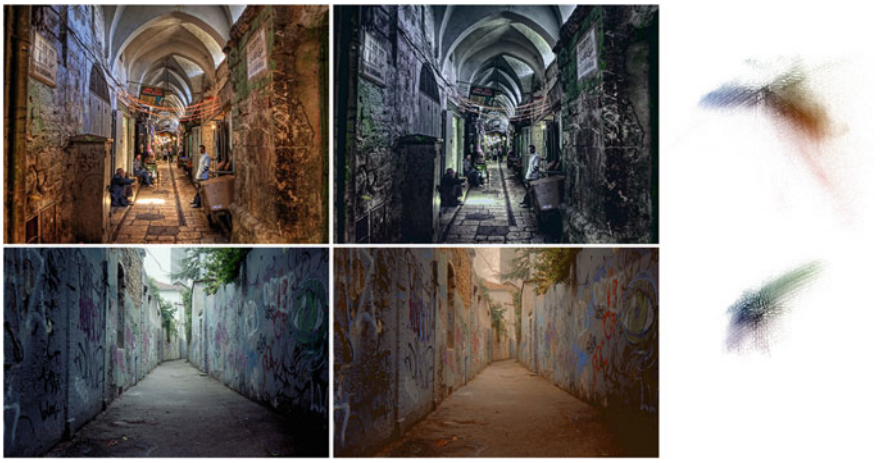


Fig. 2.4 First column: the images of a souq arcade and of a narrow Western street with tags on the walls, in different lights. Second column: the two images, each with the colors of the other. Third column: their chrominance histograms (in the CIE-lab representation). The transfer has been done independently on the 1D feature of the luma on the 2D feature of the chroma. Only the chroma histograms are represented. Pictures kindly provided by G. Peyré

2.5.2 *Efficient monotone maps from 1D constructions*

For many applications, in particular in image processing (as in the previous Section 2.5.1), people tried to build "good" transport maps, which are not necessarily optimal, but satisfy at least two criteria: they must be computationally easy to find, and they must have some "monotonicity" properties. A good idea to cope with these two requirements is to build these maps by using 1D bricks.

A very first idea that one could have is to use the Knothe map. Yet, the main drawback of such a choice is that it is highly anisotropic. The choice of the orthonormal basis is crucial, and the map gives very different roles to the different axes.

A different idea was contained in [252], where it is proposed the so-called *Iterative Distribution Transfer algorithm*: a source measure μ is given, together with a target measure ν; starting from $\mu_0 = \mu$, the authors build recursively a sequence of maps T_n and set $\mu_{n+1} = (\mathrm{T}_n)_{\#}\mu_n$. The sequence is built so that μ_n should converge to ν, which means that, for large n, the map $\mathrm{T}_n \circ \mathrm{T}_{n-1} \circ \ldots \mathrm{T}_1 \circ \mathrm{T}_0$ is a transport map from μ to a measure μ_n very close to ν.

In the construction of [252], the maps T_n are obtained in the following way: at every step n an orthonormal basis B_n is randomly selected; then, we define T_n^j as the monotone transport map from $(\pi_j)_{\#}\mu_n$ to $(\pi_j)_{\#}\nu$, where π_j is the projection on the jth coordinate of the selected basis; the map T_n is defined, in the corresponding basis, as $\mathrm{T}_n(x) = (\mathrm{T}_n^1(x_1), \ldots, \mathrm{T}_n^d(x_d))$.

There are two ideas behind this choice: first, all the maps T_n are easy to compute, composed of finitely many 1D monotone maps, and have monotonicity properties; second, if the bases B_n are chosen in a suitable diversified way, then the only possibility for the algorithm to stop on a constant measure μ_n is that μ_n and ν have the same projections on all directions. This implies $\mu_n = \nu$, thanks to the considerations on Radon and X-ray transforms of Box 2.4, and obviously suggests that the limit of μ_n should be ν.

Yet, this convergence is observed empirically but proven in [252] only if ν is Gaussian and the bases B_n are i.i.d. random variable, uniform on the set of orthonormal bases. See also Chapter 5 in [65] for a revised and more general proof.

Box 2.4. ***Good to know!**: The Radon or X-ray transform*

Definition. Given a measure μ on $\mathbb{R}^d$, we define the Radon transform $R\mu$ of μ as the family of measures (μ_e), parametrized by $e \in \mathbb{S}^{d-1}$ and defined as $(\pi_e)_\# \mu$, where $\pi_e(x) := x \cdot e$. When μ is absolutely continuous, these measures are also absolutely continuous, and we can define $R\mu$ as a function over $\mathbb{S}^{d-1} \times \mathbb{R}$, through $R\mu(e,t) = f_e(t)$ and $\mu_e = f_e(t)\mathrm{d}t$. In dimension $d > 2$, one could consider projections onto hyperplanes instead of lines (note that projecting onto all $(d-1)$-dimensional hyperplanes means considering the integral of the measure (or of its density) on arbitrary 1D lines). This transformation is called X-ray transform, but Radon and X-ray transforms coincide for $d = 2$. See, for instance, [192] for more details and applications.

Proposition. Given two measures μ, ν on $\mathbb{R}^d$, if their Radon transforms coincide, then $\mu = \nu$.

Proof. Note that $\int_{\mathbb{R}^d} e^{i\xi\cdot x}\,\mathrm{d}\mu = \int_{\mathbb{R}} e^{it|\xi|}\,\mathrm{d}\mu_{\xi/|\xi|}(t)\mathrm{d}t$. This means that the Radon transform uniquely determines the Fourier transform of a measure, so that $R\mu = R\nu \Rightarrow \hat{\mu} = \hat{\nu}$. The equality $\mu = \nu$ follows from standard results about the Fourier transform.

The Radon and the X-ray transforms are very well-known objects in applications such as tomography. Computer tomography devices in medical images are exactly based on the inversion of the X-ray transform.

Another interesting idea on how to build similar transport maps is due to M. Bernot. Many features are in common with the approach of [252].

Consider two measures $\mu, \nu \in \mathscr{P}(\mathbb{R}^d)$ and project them onto any 1D direction. For every $e \in \mathbb{S}^{d-1}$ (the unit sphere of $\mathbb{R}^d$), we take the map $\pi_e : \mathbb{R}^d \to \mathbb{R}$ given by $\pi_e(x) = x \cdot e$ and look at the image measures $(\pi_e)_\#\mu$ and $(\pi_e)_\#\nu$. Again, they are measures on the real line, and we call $T_e : \mathbb{R} \to \mathbb{R}$ the monotone optimal transport between them. The idea is that, as far as the direction e is concerned, every point x of $\mathbb{R}^d$ should be displaced of a vector $S_e(x) := (T_e(\pi_e(x)) - \pi_e(x))e$. To do a global displacement, consider $S(x) = \fint_{\mathbb{S}^{d-1}} S_e(x)\, d\mathscr{H}^{d-1}(e)$, where $\mathscr{H}^{d-1}$ is the uniform measure on the sphere.

Again, there is no reason to guarantee that $\mathrm{id} + S$ is a transport map from μ to ν. But if one fixes a small time step $\tau > 0$ and uses a displacement τS getting a measure $\mu_1 = (\mathrm{id} + \tau S)_\#\mu$, then it is possible to iterate the construction, as it is

Fig. 2.5 Optimal transport between a finite number of points on an annulus and on a duck. Colors represent the assignment of points in the annulus to points on the duck. Middle, 1D-sliced assignment; right, optimal map for the quadratic cost. The two results are very similar (but one can find few points with different colors, which means a slightly different assignment). Picture taken from [256] with permission

done in [252]. Only the definition of the transport map T_n is different. Again, one expects the sequence of measures μ_n that are built in this way to converge to ν, but the problem is not really solved. From the empirical point of view, the transport maps that are obtained in this way are quite satisfactory and have been tested in particular in the discrete case (a finite number of Dirac masses with equal mass, i.e., the so-called *assignment problem*). An example is given in Figure 2.5, where there is a comparison between the optimal transport obtained via linear programming methods (see Section 6.4.1) and the assignment obtained via these 1D-sliced ideas by M. Bernot.

Moreover, we will see in Section 5.5 that one can also build an interesting distance on the set of probabilities with this kind of idea (taking the average of a distance defined direction by direction: it will be much easier to compute than the corresponding multidimensional distance). In Section 8.4 we will also show that the iterative algorithm above corresponds indeed to a gradient flow evolution.

2.5.3 Isoperimetric inequality via Knothe or Brenier maps

For the last discussion section of this chapter, we review one of the most spectacular applications of optimal transport, i.e., its role in geometric inequalities. We will see here how the isoperimetric inequality can be proven either by using the optimal (Brenier) map or even the Knothe map.

It is interesting to learn that it was indeed one of the first motivations for the use of Knothe transport to deal with geometric inequalities (see [203]). However, the short proof of the isoperimetric inequality that we see here was first found by Gromov (see the appendix in [236]), while the original goal of Knothe in [203] was to establish other inequalities concerning convex sets.

The isoperimetric inequality states that every set E has a larger perimeter than the ball B with the same volume $|B| = |E|$. Thanks to the scaling properties of the volume and the perimeter, it is possible to write this fact in the following way:

$$\mathrm{Per}(E) \geq d\omega_d^{1/d}|E|^{1-1/d},$$

where ω_d is the volume of the unit ball in $\mathbb{R}^d$.

As usual in these discussion sections, we will not give full details on the ideas that we mention, and we will be quite sloppy concerning regularity. Anyway, think that E is a smooth domain in $\mathbb{R}^d$. By scaling, we can also suppose that E has the same volume of the unit ball $B = B(0, 1)$.

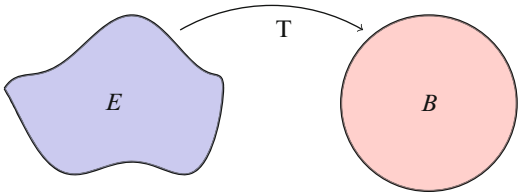
The idea to prove this result is the following: consider the densities $\mu = \mathscr{L}_E$ and $\nu = \mathscr{L}_B$, and use the Knothe transport between these densities, denoted by T. Due to the considerations on the Jacobian matrix of T, it has only positive eigenvalues and it satisfies the condition

$$\det(D\mathrm{T}) = \frac{|B|}{|E|} = 1$$

(with no absolute value at the determinant). Hence, we can write

$$\omega_d = |B| = |E| = \int_E (\det(D\mathrm{T}))^{\frac{1}{d}} \leq \frac{1}{d}\int_E \nabla\cdot\mathrm{T} = \frac{1}{d}\int_{\partial E} \mathrm{T}\cdot\mathbf{n} \leq \frac{1}{d}\mathrm{Per}(E), \qquad (2.8)$$

where the inequalities are obtained thanks to the arithmetic-geometric inequality $\frac{1}{d}\sum_i \lambda_i \geq (\lambda_1 \dots \lambda_d)^{\frac{1}{d}}$ applied to the eigenvalues of $D\mathrm{T}$ (which gives an inequality between the determinant and the trace, i.e., the divergence of T) and to the fact that T is valued in B, whence $\mathrm{T}\cdot\mathbf{n} \leq 1$. This shows $\mathrm{Per}(E) \geq d\omega_d = \mathrm{Per}(B)$. Note



that we only used the fact that the Jacobian matrix of T has positive eigenvalues and that it transports one measure onto the other. These properties are also satisfied by the Brenier map; simply this map was not known at the time this proof was first performed.

The same proof can also be used to prove the uniqueness of the optimal sets (i.e., if a set E has the same volume of the unit ball, and is such that $\mathrm{Per}(E) = d\omega_d$, then it is a unit ball, up to negligible sets).

The use of the Brenier map strongly simplifies the study of the equality cases. Suppose indeed to have equality in the above inequalities: this implies the equality $(\det(D\mathrm{T}))^{\frac{1}{d}} = (\nabla \cdot \mathrm{T})/d$ a.e., and from the cases of equality in the arithmetic-geometric inequality, we get that the eigenvalues of $D\mathrm{T}$ should all be equal. Thanks to the condition on the determinant, they are equal to 1. If the matrix is known to be symmetric (which is the case for Brenier's map, but not for Knothe's one), then we get $D\mathrm{T} = \mathrm{I}$ and $\mathrm{T}(x) = x + x_0$. Since T would be a translation, then we conclude $E = B$, up to translations.

It is slightly harder for the Knothe transport, since $D\mathrm{T}$ is not symmetric. Yet, if all the eigenvalues are 1, one deduces that T^d (the last component) is a translation. If we suppose from the beginning that (up to a fixed translation) E and B share the same barycenter, then this translation is the identity. This means $(\pi_d)_\# \mu = (\pi_d)_\# \nu$. This is not yet enough to get $\mu = \nu$, and hence $E = B$, unless one uses the fact that the Knothe transport is not isotropic (i.e., its definition depends on the choice of the axes) and that the directions of the axes are arbitrary. Hence, if E is a set giving equality in the isoperimetric inequality, then its projections onto any direction coincide with those of the ball. This implies $E \subset B$ and, by equality of the volumes, $E = B$ (the argument may also be done through the measures and the notion of Radon transform, since $(\pi_d)_\# \mu = (\pi_d)_\# \nu$ for arbitrary direction of the dth axis implies $\mu = \nu$).

We finish the section by stressing that if the isoperimetric inequality can obviously be proven in many other ways, one recent advanced refinement of it (i.e., the quantitative version of the anisotropic perimetric inequality, where the definition of perimeter is modified, and the optimal set is no more the ball but another convex set, unit ball of another norm) was only proven with the optimal exponent via optimal transport techniques (see [167] and [276] for a short review).

Chapter 3
L^1 and L^∞ theory

Chapter 1 gave, in the framework of the general theory of optimal transportation based on duality methods, an existence result for the optimal transport map when the cost is of the form $c(x, y) = h(x - y)$ and h is strictly convex. In particular using $c(x, y) = |x - y|^p$, this applies to the minimization problems

$$\min \left\{ ||\mathrm{T} - \mathrm{id}||_{L^p(\mu)} \ : \ \mathrm{T}_\# \mu = \nu \right\},$$

(expressed in Kantorovich terms as $\min \left\{ ||y - x||_{L^p(\gamma)} \ : \ \gamma \in \Pi(\mu, \nu) \right\}$), for $p \in]1, +\infty[$. We look in this chapter at the two limit cases $p = 1$ and $p = \infty$, which require additional techniques.

Then, the discussion section will go into two different directions: on the one hand, the L^1 and L^∞ cases introduced and motivated the study of convex costs which could be non-strictly convex or infinite valued somewhere, and the last developments on this topic will be debated in Section 3.3.1; on the other hand, one could wonder what is the situation for $0 < p < 1$ and more generally for costs which are concave increasing functions of the distance, which will be the subject of Section 3.3.2.

3.1 The Monge case, with cost $|x - y|$

This section will prove that, if $\mu, \nu \in \mathscr{P}(\Omega)$, where $\Omega \subset \mathbb{R}^d$ is a compact domain, and $\mu \ll \mathscr{L}^d$, then there exists an optimal transport map T for the cost $|x - y|$ (the Euclidean distance in $\mathbb{R}^d$). The proof that we present is essentially due to Ambrosio [8], as a modification from that of Sudakov, [290]. In the introduction, we gave references for the other alternative proofs of the existence in this.

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2_3

3.1.1 Duality for distance costs

Let us spend some words on the case where the cost $c(x,y)$ is actually a distance (thus satisfying the triangle inequality, vanishing if $x = y\dots$). Since the cost is symmetric, we will avoid the distinction between the $\bar{c}$- and the c-transforms.

Proposition 3.1. *If $c : X \times X \to \mathbb{R}$ is a distance, then a function $u : X \to \mathbb{R}$ is c-concave if and only if it is Lipschitz continuous with Lipschitz constant less than 1 w.r.t. the distance c. We will denote by Lip_1 the set of these functions. Moreover, for every $u \in \mathrm{Lip}_1$, we have $u^c = -u$.*

Proof. First take a c-concave function u. It can be written as

$$u(x) = \chi^c(x) = \inf_y \quad c(x,y) - \chi(y)$$

for some function $\chi : X \to \mathbb{R} \cup \{-\infty\}$. One can ignore the points y such that $\chi(y) = -\infty$ and, for the others, note that $x \mapsto c(x,y) - \chi(y)$ is a Lip_1 function, since $x \mapsto c(x,y)$ is Lip_1 as a consequence of the triangle inequality. Hence, $u \in \mathrm{Lip}_1$ since the infimum of a family of Lipschitz continuous functions with the same constant shares the same modulus of continuity.

Take now a function $u \in \mathrm{Lip}_1$. We claim that one can write

$$u(x) = \inf_y \quad c(x,y) + u(y).$$

Actually, it is clear that the infimum at the r.h.s. is not larger than $u(x)$, since one can always use $y = x$. On the other hand, $u \in \mathrm{Lip}_1$ implies $u(x) - u(y) \le c(x,y)$, i.e., $u(x) \le u(y) + c(x,y)$ for every y, and hence the infimum is not smaller than $u(x)$. This expression shows that $u = (-u)^c$ and that u is c-concave.

Applying this last formula to $-u$, which is also Lip_1, we get $u^c = -u$ and the last part of the claim follows. □

As a consequence, the duality formula, in the case of a distance cost function, gives

$$\min\left\{\int_{X\times X} c(x,y)\ \mathrm{d}\gamma\ :\ \gamma \in \Pi(\mu,\nu)\right\} = \max\left\{\int_X u\,\mathrm{d}(\mu-\nu)\ :\ u \in \mathrm{Lip}_1\right\}. \tag{3.1}$$

This also implies a useful property concerning the transport cost $\mathscr{T}_c$ when c is a distance, namely, the fact that it satisfies the triangle inequality. We will see it again in Chapter 5, where transport costs $\mathscr{T}_c$ for $c(x,y) = |x-y|^p$ are used to define distances over $\mathscr{P}(\Omega)$. We want to stress it here since we will need it later on in this chapter.

Corollary 3.2. *If $c : X \times X \to \mathbb{R}$ is a distance, given $\mu, \nu, \varrho \in \mathscr{P}(X)$, we have*

$$\mathscr{T}_c(\mu,\nu) \le \mathscr{T}_c(\mu,\varrho) + \mathscr{T}_c(\varrho,\nu).$$

Proof. Take an arbitrary function $u \in \mathrm{Lip}_1$ and write

$$\int u\,\mathrm{d}(\mu-\nu) = \int u\,\mathrm{d}(\mu-\varrho) + \int u\,\mathrm{d}(\varrho-\nu) \le \mathscr{T}_c(\mu,\varrho) + \mathscr{T}_c(\varrho,\nu).$$

The result follows by taking the supremum in u. □

We also note another useful property, typical of distances.

Corollary 3.3. *If $c : X\times X \to \mathbb{R}$ is a distance and μ,ν are such that $\mathscr{T}_c(\mu,\nu)=0$, then $\mu=\nu$.*

Proof. Many proofs can be given of this simple fact. For instance, the duality formula gives $\int u\,\mathrm{d}(\mu-\nu)=0$ for every $u\in\mathrm{Lip}_1$, which is enough to guarantee $\mu=\nu$. Otherwise, one can simply note that $\mathscr{T}_c(\mu,\nu)=0$ implies the existence of $\gamma\in\Pi(\mu,\nu)$ such that $\int c\,\mathrm{d}\gamma=0$, i.e., $c(x,y)=0$ γ-a.e. Since c is a distance, this means $x=y$ γ-a.e., which implies $\int\phi\,\mathrm{d}\mu=\int\phi(x)\,\mathrm{d}\gamma=\int\phi(y)\,\mathrm{d}\gamma=\int\phi\,\mathrm{d}\nu$ for every test function ϕ, and hence, $\mu=\nu$. □

3.1.2 Secondary variational problem

As we already saw in the book-shifting example 2.15, the optimal γ for a distance cost like $|x-y|$ in $\mathbb{R}$ is not unique in general. We want now to select a special optimizer γ, so as to be able to prove that it is induced by a transport map. For simplicity, as we will use very soon some Euclidean features of the problem, in connection with the quadratic cost $|x-y|^2$, we switch back to a concrete presentation in a domain Ω of the Euclidian space $\mathbb{R}^d$. In particular, we will not use general distances, and the norm will only be the Euclidean norm.

Let us call $O(\mu,\nu)$ the set of optimal transport plans for the cost $|x-y|$. To fix notation, let us denote by K_p the functional associating to $\gamma\in\mathscr{P}(\Omega\times\Omega)$, the quantity $\int|x-y|^p\,\mathrm{d}\gamma$, and m_p its minimal value on $\Pi(\mu,\nu)$. In this language

$$O(\mu,\nu) = \mathrm{argmin}_{\gamma\in\Pi(\mu,\nu)} K_1(\gamma) = \{\gamma\in\Pi(\mu,\nu)\,:\,K_1(\gamma)\le m_1\}.$$

Note that $O(\mu,\nu)$ is a closed subset (w.r.t. the weak convergence of measures) of $\Pi(\mu,\nu)$, which is compact. This is a general fact whenever we minimize a lower semi-continuous functional of γ.

From now on, we fix a Kantorovich potential u for the transport between μ and ν with cost $c(x,y)=|x-y|$, and we will use it as a tool to check optimality. Indeed, we have

$$\gamma\in O(\mu,\nu) \Leftrightarrow \mathrm{spt}(\gamma)\subset\{(x,y)\,:\,u(x)-u(y)=|x-y|\}. \tag{3.2}$$

This is true because optimality implies $\int (u(x) - u(y))\,\mathrm{d}\gamma = \int |x-y|\,\mathrm{d}\gamma$ and the global inequality $u(x)-u(y) \le |x-y|$ gives equality γ-a.e. All these functions being continuous, the equality finally holds on the whole support. Vice versa, equality on the support allows us to integrate it and prove that $K_1(\gamma)$ equals the value of the dual problem, which is the same of the primal; hence, one gets optimality.

As we said, we want to select a special minimizer, somehow better than the others, and prove that it is actually induced by a map.

Let us consider the problem

$$\min\{K_2(\gamma) \,:\, \gamma \in O(\mu,\nu)\}.$$

This problem has a solution $\overline{\gamma}$ since K_2 is continuous for the weak convergence and $O(\mu,\nu)$ is compact for the same convergence. We do not know a priori about the uniqueness of such a minimizer. It is interesting to note that the solution of this problem may also be obtained as the limits of solutions γ_ε of the transport problem

$$\min\{K_1(\gamma) + \varepsilon K_2(\gamma) : \gamma \in \Pi(\mu,\nu)\},$$

but we will not exploit this fact here (its proof can be done in the spirit of Γ-convergence developments, as in Section 2.4).

The goal is now to characterize this plan $\overline{\gamma}$ and prove that it is induced by a transport map.

The fact that the condition $\gamma \in O(\mu,\nu)$ may be rewritten as a condition on the support of γ is really useful since it allows us to state that $\overline{\gamma}$ also solves

$$\min\left\{\int c\,\mathrm{d}\gamma \,:\, \gamma \in \Pi(\mu,\nu)\right\}, \quad \text{where } c(x,y) = \begin{cases} |x-y|^2 & \text{if } u(x)-u(y) = |x-y|, \\ +\infty & \text{otherwise.} \end{cases}$$

Actually, minimizing this new cost implies being concentrated on the set where $u(x)-u(y) = |x-y|$ (i.e., belonging to $O(\mu,\nu)$) and minimizing the quadratic cost among those plans γ concentrated on the same set (i.e., among those $\gamma \in O(\mu,\nu)$).

Let us spend some words more in general on costs of the form

$$c(x,y) = \begin{cases} |x-y|^2 & \text{if } (x,y) \in A \\ +\infty & \text{otherwise,} \end{cases} \tag{3.3}$$

where A is a given closed subset of $\Omega \times \Omega$. First of all we note that such a cost is l.s.c. on $\Omega \times \Omega$.

Semi-continuity of the cost implies that optimal plans are concentrated on a set which is *c-CM* (Theorem 1.43). What does it mean in such a case? c-cyclical monotonicity is a condition which imposes an inequality for every k, every σ, and every family $(x_1,y_1),\dots,(x_k,y_k)$; here we will only need to use the condition for $k=2$, which gives the following result.

Lemma 3.4. *Suppose that $\Gamma \subset \Omega \times \Omega$ is c-CM for the cost c defined in (3.3). Then*

$$(x_1, y_1), (x_2, y_2) \in \Gamma, (x_1, y_2), (x_2, y_1) \in A \Rightarrow (x_1 - x_2) \cdot (y_1 - y_2) \geq 0.$$

Proof. From the definition of c-cyclical monotonicity, we have

$$(x_1, y_1), (x_2, y_2) \in \Gamma \Rightarrow c(x_1, y_1) + c(x_2, y_2) \leq c(x_1, y_2) + c(x_2, y_1).$$

For costs c of this form, this is only useful when both (x_1, y_2) and (x_2, y_1) belong to A (otherwise, we have $+\infty$ on the right-hand side of the inequality). If we also use the equivalence

$$|x_1 - y_1|^2 + |x_2 - y_2|^2 \leq |x_1 - y_2|^2 + |x_2 - y_1|^2 \Leftrightarrow (x_1 - x_2) \cdot (y_1 - y_2) \geq 0$$

(which can be obtained just by expanding the squares), the claim is proven. □

3.1.3 Geometric properties of transport rays

Let us consider for a while the role played by the function u. We collect some properties.

Lemma 3.5. *If $x, y \in \Omega$ are such that $u(x) - u(y) = |x - y|$, then u is affine on the whole segment $[x, y] := \{z = (1-t)x + ty,\ t \in [0, 1]\}$.*

Proof. Take $z = (1-t)x + ty$. Just consider that the Lip_1 condition implies

$$u(x) - u(z) \leq |x - z| = t|x - y|, \quad u(z) - u(y) \leq |z - y| = (1-t)|x - y|.$$

Summing up the two inequalities, we get $u(x) - u(y) \leq |x - y|$, but the assumption is that this should be an equality. Hence, we can infer that both inequalities are equalities, and in particular $u(z) = u(x) - t|x-y|$, i.e., $u((1-t)x+ty) = u(x) - t|x-y|$ is an affine function of t. □

Lemma 3.6. *If $z \in]x, y[:= \{z = (1-t)x + ty,\ t \in]0, 1[\}$, for a pair of points $x \neq y$ such that $u(x) - u(y) = |x - y|$, then u is differentiable at z and $\nabla u(z) = e := \frac{x-y}{|x-y|}$.*

Proof. First of all, let us give an estimate on the increment of u in directions orthogonal to e. Consider a unit vector h with $h \cdot e = 0$ and a point $z' \in]x, y[$, and let t_0 be such that $z' \pm t_0 e \in [x, y]$. Set $\delta := u(z' + th) - u(z')$. By using $u \in \mathrm{Lip}_1$, we can say that

$$u(z' + th) - u(z' - t_0 e) = u(z' + th) - u(z') + u(z') - u(z' - t_0 e) = \delta + t_0 \leq \sqrt{t^2 + t_0^2}$$

as well as

$$u(z'+th)-u(z'+t_0e) = u(z'+th)-u(z')+u(z')-u(z'+t_0e) = \delta - t_0 \le \sqrt{t^2+t_0^2}.$$

By raising these inequalities to power 2, we get

$$\delta^2 + t_0^2 \pm 2t_0\delta \le t^2 + t_0^2.$$

These two inequalities give $\pm 2t_0\delta \le \delta^2 \pm 2t_0\delta \le t^2$, and hence $2t_0|\delta| \le t^2$, i.e., $|\delta| \le t^2/2t_0$.

Consider now a point $z \in]x,y[$ and a number $t_0 < \min\{|z-x|, |z-y|\}$. Any point z'' sufficiently close to z may be written as $z'' = z' + th$ with h a unit vector orthogonal to e, $t \ll 1$, $z' \in]x,y[$ such that $z' \pm t_0 e \in [x,y]$. This allows us to write

$$u(z'')-u(z) = u(z'+th)-u(z')+u(z')-u(z) = (z'-z)\cdot e + O(t^2) = (z''-z)\cdot e + O(t^2).$$

Using $O(t^2) = o(t) = o(|z''-z|)$, we get $\nabla u(z) = e$ (Figure 3.1). □

Definition 3.7. We call transport ray any nontrivial (i.e., different from a singleton) segment $[x,y]$ such that $u(x)-u(y) = |x-y|$ that is maximal for the inclusion among segments of this form. The corresponding open segment $]x,y[$ is called the interior of the transport ray and x and y its boundary points. We call direction of a transport ray the unit vector $\frac{x-y}{|x-y|}$. We call $\mathrm{Trans}(u)$ the union of all nondegenerate transport rays, $\mathrm{Trans}^{(b)}(u)$ the union of their boundary points, and $\mathrm{Trans}(u)^{(i)}$ the union of their interiors. Moreover, let $\mathrm{Trans}^{(b+)}(u)$ be the set of upper boundary points of nondegenerate transport rays (i.e., those where u is minimal on the transport ray, say the points y in the definition $u(x) - u(y) = |x-y|$) and $\mathrm{Trans}^{(b-)}(u)$ the set of lower boundary points of nondegenerate transport rays (where u is maximal, i.e., the points x).

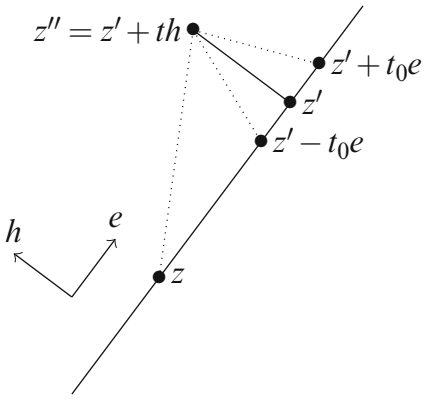


Fig. 3.1 The points used in the proof of Lemma 3.6

Corollary 3.8. *Two different transport rays can only meet at a point z which is a boundary point for both of them, and in such a case, u is not differentiable at z. In particular, if one removes the negligible set* $\mathrm{Sing}(u)$ *of non-differentiability points of u, the transport rays are disjoint.*

Proof. First of all, note that if a point z belongs to a transport ray with direction e, and u is differentiable at z, then necessarily we have $e = \nabla u(z)$. Indeed, close to z, there are points of the form $z + te$ which belong to the same transport ray, for arbitrarily small values of t (we do not care about the sign of t), and on these points, we can write

$$t = u(z + te) - u(z) = \nabla u(z) \cdot te + o(t).$$

This shows $\nabla u(z) \cdot e = 1$. Together with $|\nabla u(z)| \leq 1$, the only possibility is $\nabla u(z) = e$.

Now, suppose that two transport rays meet at a point z which is internal for both rays. In such a case, u must have two different gradients at z, following both the directions of the rays, which is impossible (recall that the different rays meeting at one point must have different directions, since otherwise they are the same transport ray, by maximality).

Suppose that they meet at z, which is in the interior of a transport ray with direction e but is a boundary point for another ray, with direction $e' \neq e$. In such a case, u should be differentiable at z, and again, $\nabla u(z)$ should take two distinct values, which is impossible.

Finally, suppose that the intersection point z is a boundary point for both segments, one with direction e and the other one with direction $e' \neq e$. In this case, there is no contradiction if u is not differentiable at z. On the other hand, if one supposes that u is differentiable at z, then we get $\nabla u(z) = e = e'$, which is, again, a contradiction. □

We will see later on that we need to say something more on the direction of the transport rays.

From this section on, we fix a transport plan $\overline{\gamma}$, optimal for the secondary variational problem, and we will try to prove that it is actually induced by a transport map. We use here that $\overline{\gamma}$ is actually concentrated on a set Γ which is c-CM (we recall that we are using the cost c defined as in (3.3)) and see how this interacts with transport rays. We can suppose $\Gamma \subset A = \{(x, y) : u(x) - u(y) = |x - y|\}$, since anyway $\overline{\gamma}$ must be concentrated on such a set, so as to have a finite value for $\int c \, \mathrm{d}\gamma$. We want to say that $\overline{\gamma}$ behaves, on each transport ray, as the monotone increasing transport. More precisely, the following is true.

Lemma 3.9. *Suppose that x_1, x_2, y_1, and y_2 are all points of a transport ray $[x, y]$ and that $(x_1, y_1), (x_2, y_2) \in \Gamma$. Define an order relation on such a transport ray through $x \leq x' \Leftrightarrow u(x) \geq u(x')$. Then if $x_1 < x_2$, we also have $y_1 \leq y_2$.*

Proof. We already know that $x_1 \le y_1$ and $x_2 \le y_2$ (thanks to $\Gamma \subset A$). Hence, the only case to be considered is the case where we have $x_1 < x_2 \le y_2 < y_1$. If we prove that this is not possible, then we have proven the claim. And this is not possible due to the fact that Γ is c-CM, since on this transport ray, due to the order relationship we have supposed and to the behavior of u on such a segment, the condition $(x_1, y_2), (x_2, y_1) \in A$ is guaranteed. This implies $(x_1 - x_2) \cdot (y_1 - y_2) \ge 0$. But this is the scalar product of two vectors parallel to e, and on the segment, this simply means that y_1 and y_2 must be ordered exactly as x_1 and x_2 are. □

From what we have seen in Chapter 2, we know that when s is a segment and $\Gamma' \subset s \times s$ is such that

$$(x_1, y_1), (x_2, y_2) \in \Gamma',\ x_1 < x_2 \Rightarrow y_1 \le y_2$$

(for the order relation on s), then Γ' is contained in the graph of a monotone increasing multivalued function, which associates to every point either a point or a segment. Yet, the interiors of these segments being disjoint, there is at most a countable number of points where the image is not a singleton (see the proof of Lemma 2.8). This means that Γ' is contained in a graph over s, up to a countable number of points of s. We will apply this idea to $\Gamma' = \Gamma \cap (s \times s)$, where s is a transport ray.

If we combine what we obtained on every transport ray, we obtain:

Proposition 3.10. *The optimal transport plan $\overline{\gamma}$ is concentrated on a set Γ with the following properties:*

- *if $(x, y) \in \Gamma$, then*
  - *either $x \in \mathrm{Sing}(u)$ (but this set of points is Lebesgue-negligible),*
  - *or $x \notin \mathrm{Trans}(u)$, i.e., it does not belong to a nondegenerate transport ray (and in this case, we necessarily have $y = x$, since otherwise, $[x, y]$ would be contained in a nondegenerate transport ray),*
  - *or $x \in \mathrm{Trans}^{(b+)}(u) \setminus \mathrm{Sing}(u)$ (and in this case, we necessarily have $y = x$, since x is contained in a unique transport ray s and it cannot have other images $y \in s$, due to the order relation $x \le y$),*
  - *or $x \in \mathrm{Trans}(u) \setminus (\mathrm{Trans}^{(b+)}(u) \cup \mathrm{Sing}(u))$ and y belongs to the same transport ray s of x (which is unique);*
- *on each transport ray s, $\Gamma \cap (s \times s)$ is contained in the graph of a monotone increasing multivalued function;*
- *on each transport ray s, the set*

$$N_s = \{x \in s \setminus \mathrm{Trans}^{(b+)}(u) : \#(\{y : (x, y) \in \Gamma\}) > 1\}$$

is countable.

It is clear that $\overline{\gamma}$ is induced by a transport map if $\mu(\bigcup_s N_s) = 0$, i.e., if we can get rid of a countable number of points on every transport ray.

This could also be expressed in terms of disintegration of measures (if μ is absolutely continuous, then all the measures μ_s given by the disintegrations of μ along the rays s are atomless; see Box 2.2 in Section 2.3 for the notion of disintegration), but we will try to avoid such an argument for the sake of simplicity. The only point that we need is the following (*Property N*, for negligibility).

Definition 3.11. We say that Property N holds for a given Kantorovich potential u if $B \subset \Omega$ is such that

- $B \subset \mathrm{Trans}(u) \setminus \mathrm{Trans}^{(b+)}(u)$,
- $B \cap s$ is at most countable for every transport ray s,

then $|B| = 0$.

This property is not always satisfied by any disjoint family of segments in $\mathbb{R}^d$, and there is an example (by Alberti, Kirchheim, and Preiss, later improved by Ambrosio, Kirchheim and Pratelli; see [14]) where a disjoint family of segments contained in a cube is such that the collection of their middle points has positive measure. We will prove that the direction of the transport rays satisfies additional properties, which guarantee *property N*.

Please be aware that, as usual, we are ignoring here measurability issues of the transport map T that we are constructing: this map is obtained by gluing the monotone maps on every segment, but this should be done in a measurable way[1].

It appears that the main tool to prove Property N is the Lipschitz regularity of the directions of the transport rays (which is the same as the direction of ∇u).

Theorem 3.12. *Property N holds if ∇u is Lipschitz continuous or if there exists a countable family of sets E_h such that ∇u is Lipschitz continuous when restricted to each E_h and $|\mathrm{Trans}(u) - \setminus \bigcup_h E_h| = 0$.*

Proof. First, suppose that ∇u is Lipschitz. Consider all the hyperplanes parallel to $d-1$ coordinate axes and with rational coordinates on the last coordinate. Consider a set B in the definition of Property N. By definition, the points of B belong to nondegenerate transport rays, i.e., to segments with positive length. Hence, every point of B belongs to a transport ray that meets at least one of the above hyperplanes at exactly one point of its interior. Since these hyperplanes are a countable quantity, up to countable unions, we can suppose that B is included in a collection S_Y of transport rays all meeting the same hyperplane Y. Moreover, we can also suppose that B does not contain any point which is a boundary point of two different transport rays, since we already know that those points are negligible. Now, let us fix such a hyperplane Y and let us consider a map $f : Y \times \mathbb{R} \to \mathbb{R}^d$ of the following form: for $y \in Y$ and $t \in \mathbb{R}$ the point $f(y,t)$ is defined as $y + t\nabla u(y)$. This map is well defined and injective on a set $A \subset Y \times \mathbb{R}$ which is the one we are interested in. This set A is

[1]Measurability could be proven, either by restricting to a σ-compact set Γ or by considering the disintegrations μ_s and ν_s and using the fact that, on each s, T is the monotone map sending μ_s onto ν_s (and hence it inherits some measurability properties of the dependence of μ_s and ν_s w.r.t. s, which are guaranteed by abstract disintegration theorems).

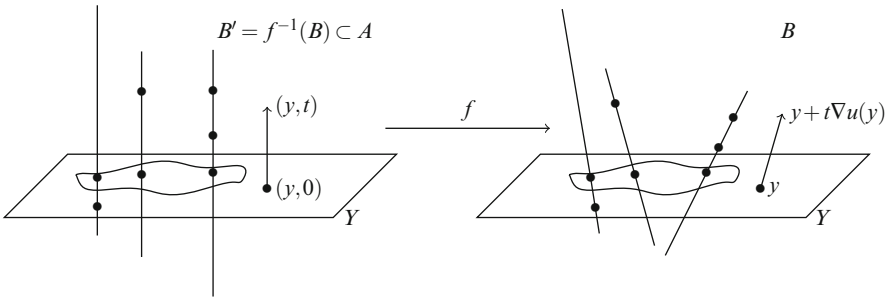


Fig. 3.2 The map f in Theorem 3.12. The points of B and B' are only a countable number on each line and represented by bullets

defined as those pairs (y,t) where y is in the interior of a transport ray of S_Y; hence, u is differentiable at such a point, and $y + t\nabla u(y)$ belongs to the interior of the same transport ray. The map f is injective, since getting the same point as the image of (y,t) and of (y',t') would mean that two different transport rays cross at such a point. B is contained in the image of f by construction, so that f is a bijection between B and $B' := f^{-1}(B)$. The map f is also Lipschitz continuous, as a consequence of the Lipschitz behavior of ∇u. Note that B' is a subset of $Y \times \mathbb{R}$ containing at most countably many points on every line $\{y\} \times \mathbb{R}$. By Fubini's theorem, this implies $|B'| = 0$. Then we have also $|B| = |f(B')| \le \mathrm{Lip}(f)^d |B'|$, which implies $|B| = 0$.

It is clear that the property is also true when ∇u is not Lipschitz but is Lipschitz continuous on each set E_h of a partition covering almost all the points of $\mathrm{Trans}(u)$, since one can apply the same kind of arguments to all the sets $B \cap E_h$ and then use countable unions (Figure 3.2). □

We now need to prove that ∇u is countably Lipschitz.

Definition 3.13. A function $f : \Omega \to \mathbb{R}^d$ is said to be countably Lipschitz if there exists a countable family of sets E_h such that f is Lipschitz continuous on each E_h and $|\Omega \setminus \bigcup_h E_h| = 0$.

Note that "being Lipschitz continuous on a set E" or "being the restriction to E of a Lipschitz continuous function defined on the whole $\mathbb{R}^d$" is actually the same property, due to Lipschitz extension theorems. Indeed, whenever $E \subset X$ are metric spaces, every L-Lipschitz function $f : E \to \mathbb{R}$ can be extended to an L-Lipschitz function over X. This is known as Kirszbraun theorem and can be easily obtained by using the functions f_k of the Memo Box 1.5, for $k \ge L$.

We want to prove that ∇u is countably Lipschitz. We will first prove that u coincides with some λ-convex or λ-concave functions on a sequence of sets covering almost everything. This requires a definition.

Definition 3.14. A function $f : \Omega \to \mathbb{R}$ is said to be λ-convex if $x \mapsto f(x) - \frac{\lambda}{2}|x|^2$ is convex and λ-concave if $x \mapsto f(x) + \frac{\lambda}{2}|x|^2$ is concave. Note that the number λ is not required to be positive, so that λ-convex functions for $\lambda > 0$ are strictly convex, and,

for $\lambda < 0$, they just have second derivatives (where they exist) which are bounded from below. Functions which are λ-convex (or λ-concave) for some values of λ are called *semi-convex* (or *semi-concave*).

Proposition 3.15. *There exist some sets E_h such that*

- $\bigcup_h E_h = \mathrm{Trans}(u) \setminus \mathrm{Trans}(u)^{(b+)}$,
- *on each set E_h the function u is the restriction of a λ-concave function, for a value λ depending on h.*

Proof. Let us define

$$E_h = \left\{ x \in \mathrm{Trans}(u) \,:\, \exists z \in \mathrm{Trans}(u) \text{ with } u(x) - u(z) = |x-z| > \frac{1}{h} \right\},$$

which is roughly speaking made of those points in the transport rays that are at least at a distance $\frac{1}{h}$ apart from the upper boundary point of the ray. It is clear that $E_h \subset \mathrm{Trans}(u) \setminus \mathrm{Trans}(u)^{(b+)}$. Moreover, we easily have $\bigcup_h E_h = \mathrm{Trans}(u) \setminus \mathrm{Trans}^{(b+)}(u)$.

Let us fix a function $c_h : \mathbb{R}^d \to \mathbb{R}$ with the following properties: $c_h \in C^2(\mathbb{R}^d)$, $|\nabla c_h| \le 1$, $c_h(z) \ge |z|$ for all $z \in \mathbb{R}^d$, $c_h(z) = |z|$ for all $z \notin B(0, 1/h)$. Set $\lambda_h := -||D^2 c_h||_{L^\infty}$ (note that we can take this value to be of the order of $-\frac{1}{h}$). It is easy to check that, if $x \in E_h$, one has

$$u(x) = \inf_{y \in \mathbb{R}^d} |x-y| + u(y) \le \inf_{y \in \mathbb{R}^d} c_h(x-y) + u(y) \le \inf_{y \notin B(x,1/h)} |x-y| + u(y) = u(x),$$

where the first inequality is a consequence of $|z| \le c_h(z)$ and the second is due to the restriction to $y \notin B(x, 1/h)$. The last equality is justified by the definition of E_h. This implies that all the inequalities are actually equalities and that $u(x) = u_h(x)$ for all $x \in E_h$, where

$$u_h(x) := \inf_{y \in \mathbb{R}^d} c_h(x-y) + u(y).$$

It is important to note that u_h is a λ_h-concave function.

Let us justify that u_h is λ_h-concave. With this choice of λ_h, c_h is obviously λ_h-concave. Consider

$$\begin{aligned} u_h(x) + \frac{\lambda_h}{2}|x|^2 &= \inf_{y \in \mathbb{R}^d} c_h(x-y) + \frac{\lambda_h}{2} 2|x|^2 + u(y) \\ &= \inf_{y \in \mathbb{R}^d} c_h(x-y) + \frac{\lambda_h}{2}|x-y|^2 + \lambda_h x \cdot y - \frac{\lambda_h}{2}|y|^2 + u(y). \end{aligned}$$

This last expression shows that $u_h(x) + \frac{\lambda_h}{2}|x|^2$ is concave in x, since it is expressed as an infimum of concave functions ($c_h(x-y) + \frac{\lambda_h}{2}|x-y|^2$ is concave, $\lambda_h x \cdot y$ is linear, and the other terms are constant in x). Hence, $u_h(x)$ is λ_h-concave. □

The previous theorem allows us to replace the function u with the functions u_h, which are more regular (since they are λ-concave, they share the same regularity of

convex functions). Unfortunately, this is not enough yet, since convex functions in general are not even differentiable. A new countable decomposition is needed and can be obtained from the following theorem that we do not prove.

Theorem 3.16. *If f is a convex function, then ∇f is countably Lipschitz.*

The proof of this theorem may be found in [13], in Theorem 5.34. It is also true when we replace ∇f with an arbitrary BV function, and this is the framework that one finds in [13].

Box 3.1. ***Memo**: BV functions in $\mathbb{R}^d$*

On a domain $\Omega \subset \mathbb{R}^d$, an L^1 function $f : \Omega \to \mathbb{R}$ is said to be a BV function if its distributional derivatives are finite measures. This means that we require

$$\int_\Omega f\,(\nabla \cdot \xi)\,\mathrm{d}x \le C||\xi||_{L^\infty},$$

for every C^1_c vector field $\xi : \Omega \to \mathbb{R}^d$. The space BV is hence a wider class than the Sobolev spaces $W^{1,p}$, where the distributional derivatives are supposed to belong to L^p (i.e., they are absolutely continuous measures with integrability properties).

It happens that the distributional derivative ∇f cannot be any measure, but must be of the form $\nabla f = \nabla^a f(x)\,\mathrm{d}x + D^s f + D^j f$, where $\nabla^a f(x)\,\mathrm{d}x$ is the absolutely continuous part of ∇f (with respect to the Lebesgue measure), $D^j f$ is the so-called jump part and has a density with respect to $\mathscr{H}^{d-1}$ (it is concentrated on the set J_f of those points where there is an hyperplane such that the function f has two different limits on the two sides of the hyperplane, in a suitable measure-theoretical sense, and the density of $D^j u$ with respect to $\mathscr{H}^{d-1}_{|J_f}$ is exactly the difference of these two limit values times the direction of the normal vector to the hyperplane: $D^j f = (f^+ - f^-)\mathbf{n}_{J_f} \cdot \mathscr{H}^{d-1}_{|J_f}$), and D^c is the so-called Cantor part, which is singular to the Lebesgue measure but vanishes on any $(d-1)$-dimensional set.

We denote by $BV_{loc}(\mathbb{R}^d)$ the space of functions which are locally BV in the sense that their derivatives are Radon measures, i.e., measures which are locally finite (in the definition with test functions ξ, we use $\xi \in C^1_c(\mathbb{R}^d)$ and the constant C may depend on the support of ξ). Vector-valued BV functions are just defined as functions $f = (f_1, \dots, f_k) : \mathbb{R}^d \to \mathbb{R}^k$ which are componentwise BV, i.e., $f_i \in BV$ for each $i = 1, \dots, k$. It is interesting that gradients $g = \nabla f$ of convex functions are always locally BV. This depends on the fact that the Jacobian matrix of the gradient of a convex function is indeed the Hessian of such a function and is positive definite. This means that the matrix-valued distribution $\nabla(\nabla f) = D^2 f$ is a positive distribution, and we know that positive distributions are necessarily positive measures (warning: this requires precise definitions when we work with matrix- and vector-valued functions).

BV functions satisfy several fine regularity properties almost everywhere for the Lebesgue or the $\mathscr{H}^{d-1}$-measure, and the reader may refer to [13] or [160]. In particular, we cite the following (not at all easy) result, which implies Theorem 3.16.

Theorem. If $f \in BV_{loc}(\mathbb{R}^d)$, then f is countably Lipschitz.

As a consequence of Proposition 3.15 and Theorem 3.16 one has:

Proposition 3.17. *If u is a Kantorovich potential, then $\nabla u : \mathrm{Trans}(u) \to \mathbb{R}^d$ is countably Lipschitz.*

Proof. It is clear that the countably Lipschitz regularity of Theorem 3.16 is also true for the gradients of λ-convex and λ-concave functions. This means that this is true for ∇u_h and, by countable union, for ∇u. □

3.1.4 Existence and nonexistence of an optimal transport map

From the analysis performed in the previous sections, we obtain the following result, essentially taken from [10, 14, 44].

Theorem 3.18. *Under the usual assumption $\mu \ll \mathscr{L}^d$, the secondary variational problem admits a unique solution $\overline{\gamma}$, which is induced by a transport map* T*, monotone nondecreasing on every transport ray.*

Proof. Proposition 3.17, together with Theorem 3.12, guarantees that Property N holds. Hence, Proposition 3.10 may be applied to get $\overline{\gamma} = \gamma_{\mathrm{T}}$. The uniqueness follows in the usual way: if two plans $\overline{\gamma}' = \gamma_{\mathrm{T}'}$ and $\overline{\gamma}'' = \gamma_{\mathrm{T}''}$ optimize the secondary variational problem, the same should be true for $\frac{1}{2}\gamma_{\mathrm{T}'} + \frac{1}{2}\gamma_{\mathrm{T}''}$. Yet, for this measure to be induced by a map, it is necessary to have $\mathrm{T}' = \mathrm{T}''$ a.e. (see Remark 1.19). □

Definition 3.19. The optimal transport plan $\overline{\gamma}$ will be called *ray-monotone* transport plan and the transport map which corresponds to it *ray-monotone* transport map.

Note that as a by-product of this analysis, we also obtain $|\mathrm{Trans}(u)^{(b)}| = 0$, since $\mathrm{Trans}(u)^{(b)}$ is a set meeting every transport ray in two points, and it is hence negligible by Property N. Yet we could not have proven it before, so unfortunately every strategy based on a decomposition of $\mathrm{Trans}(u)^{(i)}$ is not complete.

To complete this section, one could wonder whether the assumption $\mu \ll \mathscr{L}^d$ is really necessary for this existence result: would the condition "$\mu(A) = 0$ for every $(d-1)$-rectifiable set A" (or other conditions as in Section 1.3.1) be enough for the existence of an optimal map, as it happened for the quadratic case?

The answer is not, as it is shown in the following example.

Transport from a graph to a square

Consider a continuous function $B : [0,1] \to \mathbb{R}$ with the following property: for every Lipschitz function $f : \mathbb{R} \to \mathbb{R}$, the sets $\{t \in [0,1] : B(t) = f(t)\}$ and $\{t \in [0,1] : f(B(t)) = t\}$ are Lebesgue negligible. Then take $\mu = (\mathrm{id}, B)_{\#}(\mathscr{L}^1 \llcorner [0,1])$ and $\nu = \mathscr{L}^2 \llcorner Q$, where Q is the square $[0,1] \times [L, L+1]$ with $L = ||B||_{L^\infty}$.

In practice, μ is concentrated on the graph of B and ν on a square above the graph. With this choice, the measure μ gives zero mass to any rectifiable curve of the form $\{(t, f(t))\}$ or $\{(f(s), s)\}$, for f Lipschitz. On the other hand, it is not absolutely continuous, as it gives positive mass to a graph, which is $\mathscr{L}^2$-negligible. The goal is now to prove that no optimal transport map exists from μ to ν (Figure 3.3).

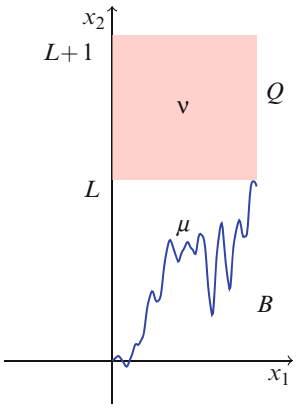


Fig. 3.3 The measure μ concentrated on a graph, with ν on a square

First consider that there is an obvious transport map S sending ν into μ. Just take $S(x_1, x_2) = (x_1, B(x_1))$ (i.e., each point moves vertically to the corresponding point on the graph of B). Obviously, S is not injective. The Monge cost of S is given by $\int_Q |x_2 - B(x_1)|\, \mathrm{d}x_1\, \mathrm{d}x_2 = \int_Q (x_2 - B(x_1))\, \mathrm{d}x_1\, \mathrm{d}x_2$. Moreover, the function $u(x_1, x_2) = -x_2$ is in Lip_1 and we have

$$\int_{\mathbb{R}^2} u\, \mathrm{d}(\mu - \nu) = -\int_0^1 B(x_1)\, \mathrm{d}x_1 + \int_Q x_2\, \mathrm{d}x_1\, \mathrm{d}x_2 = \int_Q |x_2 - B(x_1)|\, \mathrm{d}x_1\, \mathrm{d}x_2.$$

This proves that S is an optimal map from ν to μ and u is a Kantorovich potential (more precisely: u is a Kantorovich potential from μ to ν and $-u$ from ν to μ). But this also proves that any optimal transport plan between μ and ν must be concentrated on the set $\{(x, y) : y_2 - x_2 = |x - y|\}$, i.e., on $\{(x, y) : y_1 = x_1, \text{ and } x_2 \le y_2\}$. In particular the transport can only occur vertically, exactly as it happens for S. Yet, there is no transport map from μ to ν with this property: indeed, the image of $\mathrm{spt}(\mu)$ through such a map could only contain one point in each vertical segment of $\mathrm{spt}(\nu)$ and could not cover almost all Q.

To make the example complete[2], one only needs to produce such a function B. As the reader could imagine, the choice of the notation has been done on purpose: the typical trajectory of the Brownian motion satisfies this assumption. Indeed, it is known that the trajectories of a Brownian motion are almost surely never differentiable and even never approximately differentiable (see Box 3.3 in Section 3.3.2 for this notion, and see [222] for the behavior of the Brownian motion with this respect). By the way, it is also proven in [25] that, almost surely, there is no set A with dimension bigger than $1/2$ on which B could be Hölder continuous with exponent larger than $1/2$. This prevents the possibility that B coincides with a Lipschitz

[2] Note that the construction is essentially the same as in the example provided in [199], for a different goal. The regularity degree is slightly different, and we decided to handle by hands "vertical" Lipschitz curves in order to make a self-contained presentation.

function on a 1-dimensional set. We can even say that $\{t \in [0,1] \,:\, B(t) = f(t)\}$ has dimension less or equal than $1/2$.

For the other set (i.e., the points where $t = f(B(t))$), it is trickier. We can assume, by contradiction, that there exists a compact set K with $|K| > 0$ and $f \circ B = \mathrm{id}$ on K. Hence, the set $H = B(K)$ is also compact and must have positive measure. Otherwise, if $|H| = 0$, $K \subset f(H)$ would be negligible too, because f is Lipschitz. We can assume that f is C^1, if we restrict H. Indeed, for every Lipschitz function f and every n, there is a C^1 function $\tilde{f}_n$ with $|\{f \neq \tilde{f}_n\}| < \frac{1}{n}$ (see [160] for this fact). If we set $H_n = \{f = \tilde{f}_n\}$, we have $|H \setminus \bigcup_n H_n| = 0$ and $|f(H \setminus \bigcup_n H_n)| = 0$; this proves $|f(\bigcup_n H_n)| = |\bigcup_n f(H_n)| > 0$, and at least for some n, we must have $|H_n|, |f(H_n)| > 0$. Hence, we can suppose $f \in C^1$. Now, there are two possibilities: either $f' = 0$ on H or there is a small interval where f is a diffeomorphism. In this second case, we find $B = f^{-1}$ on a set of positive measure, which is a contradiction with the well-known properties of Brownian motions (since f^{-1} is Lipschitz, and even C^1). On the contrary, if $f' = 0$ on H, then the area formula provides $|K| = 0$, which is also a contradiction.

3.1.5 *Approximation issues for the ray-monotone optimal transport*

The present section deals with some natural questions concerning the ray-monotone optimal transport plan $\overline{\gamma}$, and in particular approximation questions. Indeed, we have seen (Theorem 1.50) that for any weakly converging sequence $\gamma_n \rightharpoonup \gamma$, if γ_n is optimal for a continuous cost between its marginal, then so is γ. Yet, we know that for $c(x,y) = |x-y|$, the optimal transport plan is not in general unique, and for many applications, one prefers to stick to the ray-monotone one. Hence, the question is how to select this special transport plan by approximation, when the measures and/or the cost varies.

For many applications (see in particular Chapter 4), it is interesting to approximate transport plans through transport maps sending a given measure to an atomic one. This is because these kinds of transport maps are actually composed of different homotheties defined on a partition of the domain.

Here is a useful approximation lemma in the spirit of Γ-convergence developments (see [77], Section 2.4, and Box 4.6 in Section 4.4.2).

For fixed measures $\mu, \nu \in \mathscr{P}(\Omega)$, with $\mu \ll \mathscr{L}^d$, consider the following family of minimization problems (P_ε):

$$(P_\varepsilon) = \min \Big\{ \mathscr{T}_c((\pi_y)_\# \gamma, \nu) + \varepsilon K_1(\gamma) + \varepsilon^2 K_2(\gamma) + \varepsilon^{3d+3} \#((\pi_y)_\# \gamma) \;:\; \gamma \in \mathscr{P}(\Omega \times \Omega),\ (\pi_x)_\# \gamma = \mu \Big\},$$

where $\mathscr{T}_c$ is the minimum value of the transport problem for the cost $c(x,y) = |x-y|$, $K_p(\gamma) = \int |x-y|^p \gamma(\mathrm{d}x, dy)$ for $p = 1, 2$, and the symbol # denotes the cardinality of the support of a measure. Note that $K_1(\gamma) \neq \mathscr{T}_c((\pi_x)_\#\gamma, (\pi_y)_\#\gamma)$ in general.

Actually, the minimization of (P_ε) consists in looking for a transport plan with first marginal equal to μ satisfying the following criteria with decreasing degree of importance: the second marginal must be close to ν; the K_1 cost of transportation should be small; the K_2 as well; and, finally, the second marginal must be atomic with not too many atoms.

This minimization problem has obviously at least a solution (by the direct method, since Ω is compact). We denote by γ_ε such a solution and $\nu_\varepsilon := (\pi_y)_\#\gamma_\varepsilon$ its second marginal. It is straightforward that ν_ε is an atomic measure and that γ_ε is the (unique, since $\mu \ll \mathscr{L}^d$) solution of $\min\{K_1(\gamma) + \varepsilon K_2(\gamma) \,:\, \gamma \in \Pi(\mu, \nu_\varepsilon)\}$, which is an optimal transport problem from μ to ν_ε for the cost $|x-y| + \varepsilon|x-y|^2$. Set

$$\overline{\gamma} = \operatorname{argmin}\{K_2(\gamma) \,:\, \gamma \text{ is a } K_1\text{-optimal transport plan from } \mu \text{ to } \nu\}. \tag{3.4}$$

This transport plan $\overline{\gamma}$ is unique, and it is the unique ray-monotone optimal transport plan from μ to ν (this is a consequence of Section 3.1.4); note that the functional K_2 could have been replaced by any functional $\gamma \mapsto \int h(x-y)\,\mathrm{d}\gamma$ for a strictly convex function h.

Lemma 3.20. *As $\varepsilon \to 0$, we have $\nu_\varepsilon \rightharpoonup \nu$ and $\gamma_\varepsilon \rightharpoonup \overline{\gamma}$.*

Proof. It is sufficient to prove that any possible limit of subsequences coincides with ν or $\overline{\gamma}$, respectively. Let γ_0 be one such a limit and $\nu_0 = (\pi_y)_\#\gamma_0$ the limit of the corresponding subsequence of ν_ε. Consider a regular grid $G^n \subset \Omega$ composed of approximately Cn^d points (take $G^n = \frac{1}{n}\mathbb{Z}^d \cap \Omega$), and let p_n be any measurable map from Ω to G^n, with the property $|p_n(x) - x| \leq 1/n$ (for instance, p_n can be the projection map, where it is well defined). Set $\nu^n := (p_n)_\#\nu$ and note $\#\nu^n \leq Cn^d$, as well as $\nu^n \rightharpoonup \nu$ and $\mathscr{T}_c(\nu^n, \nu) \leq \frac{1}{n}$.

First step: $\nu_0 = \nu$. Take γ^n any transport plan from μ to ν^n. By optimality of γ_ε we have

$$\mathscr{T}_c(\nu_\varepsilon, \nu) \leq \mathscr{T}_c(\nu^n, \nu) + \varepsilon K_1(\gamma^n) + \varepsilon^2 K_2(\gamma^n) + C\varepsilon^{3d+3} n^d.$$

Fix n, let ε go to 0, and get

$$\mathscr{T}_c(\nu_0, \nu) \leq \mathscr{T}_c(\nu^n, \nu) \leq \frac{1}{n}.$$

Then let $n \to \infty$ and get $\mathscr{T}_c(\nu_0, \nu) = 0$, which implies $\nu_0 = \nu$ (thanks to Corollary 3.3).

Second step: γ_0 is optimal for K_1 from μ to ν. Take any optimal transport plan γ^n (for the K_1 cost) from μ to ν^n (i.e., $K_1(\gamma^n) = \mathscr{T}_c(\mu, \nu^n)$). These plans converge to a certain optimal plan $\tilde{\gamma}$ from μ to ν (i.e., $K_1(\tilde{\gamma}) = \mathscr{T}_c(\mu, \nu)$). Then, by optimality of γ_ε, we have

$$\varepsilon K_1(\gamma_\varepsilon) \le \mathscr{T}_c(\nu^n,\nu) + \varepsilon K_1(\gamma^n) + \varepsilon^2 K_2(\gamma^n) + C\varepsilon^{3d+3}n^d$$
$$\le \frac{1}{n} + \varepsilon K_1(\gamma^n) + C\varepsilon^2 + C\varepsilon^{3d+3}n^d.$$

Then take $n \approx \varepsilon^{-2}$ and divide by ε:

$$K_1(\gamma_\varepsilon) \le \varepsilon + K_1(\gamma^n) + C\varepsilon + C\varepsilon^{d+2}.$$

Passing to the limit, we get

$$K_1(\gamma_0) \le K_1(\tilde{\gamma}) = \mathscr{T}_c(\mu,\nu),$$

which implies that γ_0 is also optimal.

Third step: $\gamma_0 = \overline{\gamma}$. Take any optimal transport plan γ (for the cost K_1) from μ to ν. Set $\gamma^n = (\mathrm{id}\times p_n)_\#\gamma$, where $(\mathrm{id}\times p_n)(x,y) := (x, p_n(y))$. We have $(\pi_y)_\#\gamma^n = \nu^n$. Then we have

$$\mathscr{T}_c(\nu_\varepsilon,\nu) + \varepsilon K_1(\gamma_\varepsilon) + \varepsilon^2 K_2(\gamma_\varepsilon) \le \mathscr{T}_c(\nu^n,\nu) + \varepsilon K_1(\gamma^n) + \varepsilon^2 K_2(\gamma^n) + C\varepsilon^{3d+3}n^d.$$

Moreover, using the triangle inequality of Corollary 3.2, we infer

$$K_1(\gamma_\varepsilon) \ge \mathscr{T}_c(\mu,\nu_\varepsilon) \ge \mathscr{T}_c(\mu,\nu) - \mathscr{T}_c(\nu_\varepsilon,\nu).$$

We also have

$$K_1(\gamma^n) \le K_1(\gamma) + \int |p_n(y)-y|\,\mathrm{d}\gamma(x,y) \le K_1(\gamma) + \frac{1}{n} = \mathscr{T}_c(\mu,\nu) + \frac{1}{n}.$$

Hence, we have

$$(1-\varepsilon)\mathscr{T}_c(\nu_\varepsilon,\nu) + \varepsilon\mathscr{T}_c(\mu,\nu) + \varepsilon^2 K_2(\gamma_\varepsilon) \le \frac{1}{n} + \varepsilon\mathscr{T}_c(\mu,\nu) + \frac{\varepsilon}{n} + \varepsilon^2 K_2(\gamma^n) + C\varepsilon^{3d+3}n^d.$$

Getting rid of the first term (which is positive) on the left-hand side, simplifying $\varepsilon\mathscr{T}_c(\mu,\nu)$ and dividing by ε^2, we get

$$K_2(\gamma_\varepsilon) \le \frac{1+\varepsilon}{n\varepsilon^2} + K_2(\gamma^n) + C\varepsilon^{3d+1}n^d.$$

Here it is sufficient to take $n \approx \varepsilon^{-3}$ and pass to the limit to get

$$K_2(\gamma_0) \le K_2(\gamma),$$

which is the condition characterizing $\overline{\gamma}$ (K_2-optimality among plans which are K_1-minimizers). □

This approximation result will be useful in Chapter 4. It is mainly based on the fact (already mentioned in Section 3.1.2) that, when we minimize the transport cost $|x-y|+\varepsilon|x-y|^2$, we converge to the solution of the secondary variational problem, i.e., to the ray-monotone transport map. As we said, any kind of strictly convex perturbation should do the same job as the quadratic one.

Yet, there are other approximations that are as natural as this one but are still open questions. We list two of them.

Open Problem (stability of the monotone transport): take γ_n to be the ray-monotone optimal transport plan for the cost $|x-y|$ between μ_n and ν_n. Suppose $\mu_n \rightharpoonup \mu, \nu_n \rightharpoonup \nu$, and $\gamma_n \rightharpoonup \gamma$. Is γ the ray-monotone optimal transport plan between μ and ν?

Open Problem (limit of $\sqrt{\varepsilon^2+|x-y|^2}$): take γ_ε to be the optimal transport plan for the cost $\sqrt{\varepsilon^2+|x-y|^2}$ between two given measures μ and ν. If $\mathrm{spt}(\mu)\cap \mathrm{spt}(\nu)=\emptyset$, then γ_ε can be easily proven to converge to the ray-monotone optimal transport plan (because, at the first nonvanishing order in ε, the perturbation of the cost is of the form $|x-y|+\varepsilon^2/|x-y|+o(\varepsilon^2)$, and the function $h(t)=1/t$ is strictly convex on $\{t>0\}$). Is the same true if the measures have intersecting (or identical) supports as well?

This last approximation is very useful when proving (or trying to prove) regularity results (see [210]).

3.2 The supremal case, L^∞

We consider now a different problem: instead of minimizing

$$K_p(\gamma)=\int |x-y|^p \,\mathrm{d}\gamma,$$

we want to minimize the maximal displacement, i.e., its L^∞ norm. This problem has been first addressed in [122], but the proof that we present here is essentially taken from [197].

Let us define

$$\begin{aligned} K_\infty(\gamma) := ||x-y||_{L^\infty(\gamma)} &= \inf\{m\in\mathbb{R} \,:\, |x-y|\le m \text{ for } \gamma-a.e.(x,y)\} \\ &= \max\{|x-y| \,:\, (x,y)\in \mathrm{spt}(\gamma)\} \end{aligned}$$

(where the last equality, between an L^∞ norm and a maximum on the support, is justified by the continuity of the function $|x-y|$).

Lemma 3.21. *For every $\gamma\in\mathscr{P}(\Omega\times\Omega)$, the quantities $K_p(\gamma)^{\frac{1}{p}}$ increasingly converge to $K_\infty(\gamma)$ as $p\to+\infty$. In particular, $K_\infty(\gamma)=\sup_{p\ge1}K_p(\gamma)^{\frac{1}{p}}$ and K_∞*

is l.s.c. for the weak convergence in $\mathscr{P}(\Omega \times \Omega)$. Thus, it admits a minimizer over $\Pi(\mu, \nu)$, which is compact.

Proof. It is well known that, on any finite measure space, the L^p norms converge to the L^∞ norm as $p \to \infty$, and we will not reprove it here. This may be applied to the function $(x, y) \mapsto |x - y|$ on $\Omega \times \Omega$, endowed with the measure γ, thus getting $K_p(\gamma)^{\frac{1}{p}} \to K_\infty(\gamma)$. It is important that this convergence is monotone here, and this is true when the measure has unit mass. In such a case, we have for $p < q$, using Hölder (or Jensen) inequality

$$\int |f|^p \, \mathrm{d}\gamma \leq \left(\int |f|^q \, \mathrm{d}\gamma \right)^{p/q} \left(\int 1 \, \mathrm{d}\gamma \right)^{1-p/q} = \left(\int |f|^q \, \mathrm{d}\gamma \right)^{p/q},$$

for every $f \in L^q(\gamma)$. This implies, by taking the pth root, $||f||_{L^p} \leq ||f||_{L^q}$. Applied to $f(x, y) = |x - y|$, this gives the desired monotonicity.

From this fact, we infer that K_∞ is the supremum of a family of functionals which are continuous for the weak convergence (since K_p is the integral of a bounded continuous function, Ω being compact, and taking the pth root does not break continuity). As a supremum of continuous functionals, it is l.s.c. and the conclusion follows. □

The goal now is to analyze the solution of

$$\min\{K_\infty(\gamma) \ : \ \gamma \in \Pi(\mu, \nu)\}$$

and to prove that there is at least a minimizer γ induced by a transport map. This map would solve

$$\min\{||\mathrm{T} - \mathrm{id}||_{L^\infty(\mu)} \ : \ \mathrm{T}_\# \mu = \nu\}.$$

Here as well, there will be no uniqueness (it is almost always the case when we minimize an L^∞ criterion); hence we define $O_\infty(\mu, \nu) = \mathrm{argmin}_{\gamma \in \Pi(\mu,\nu)} K_\infty(\gamma)$, the set of optimal transport plans for this L^∞ cost. Note that $O_\infty(\mu, \nu)$ is compact, since K_∞ is l.s.c. (as for $O(\mu, \nu)$). Set now $L := \min\{K_\infty(\gamma) \ : \ \gamma \in \Pi(\mu, \nu)\}$: we can write

$$\gamma \in O_\infty(\mu, \nu) \Leftrightarrow \mathrm{spt}(\gamma) \subset \{(x, y) : \ |x - y| \leq L\}$$

(since any transport plan γ concentrated on the pairs where $|x - y| \leq L$ satisfies $K_\infty(\gamma) \leq L$ and is hence optimal). We will suppose $L > 0$; otherwise, this means that it is possible to obtain ν from μ with no displacement, i.e., $\mu = \nu$ and the optimal displacement is the identity.

Consequently, exactly as for the L^1 case, we can define a secondary variational problem:

$$\min\{K_2(\gamma) \ : \ \gamma \in O_\infty(\mu, \nu)\}.$$

This problem has a solution $\overline{\gamma}$ since K_2 is continuous for the weak convergence and $O_\infty(\mu,\nu)$ is compact. We do not know for the moment whether we have uniqueness for this minimizer. Again, it is possible to say that $\overline{\gamma}$ also solves

$$\min\left\{\int_{\Omega\times\Omega} c\,\mathrm{d}\gamma \,:\, \gamma\in\Pi(\mu,\nu)\right\}, \quad \text{where } c(x,y)=\begin{cases}|x-y|^2 & \text{if } |x-y|\le L\\ +\infty & \text{otherwise.}\end{cases}$$

The arguments are the same as in the L^1 case. Moreover, also the form of the cost c is similar, and this cost is l.s.c. as well. Hence, $\overline{\gamma}$ is concentrated on a set $\Gamma\subset\Omega\times\Omega$ which is c-CM. This means

$$(x_1,y_1),(x_2,y_2)\in\Gamma, |x_1-y_2|,|x_2-y_1|\le L \Rightarrow (x_1-x_2)\cdot(y_1-y_2)\ge 0. \tag{3.5}$$

We can also suppose $\Gamma\subset\{(x,y)\,:\,|x-y|\le L\}$. We will try to "improve" the set Γ, by removing negligible sets and getting better properties. Then, we will show that the remaining set $\tilde{\Gamma}$ is contained in the graph of a map T, thus obtaining the result.

As usual, we will assume $\mu\ll\mathscr{L}^d$. Then, we need to recall the notion of Lebesgue points, which is specific to the Lebesgue measure.

Box 3.2. **– *Memo*** – *Density points*

Definition - For a measurable set $E\subset\mathbb{R}^d$ we call Lebesgue point of E a point $x\in\mathbb{R}^d$ such that

$$\lim_{r\to 0}\frac{|E\cap B(x,r)|}{|B(x,r)|}=1.$$

The set of Lebesgue points of E is denoted by $\mathrm{Leb}(E)$, and it is well known that $|E\setminus\mathrm{Leb}(E)|+|\mathrm{Leb}(E)\setminus E|=0$. This is actually a consequence of a more general fact: given a function $f\in L^1_{loc}(\mathbb{R}^d)$, a.e. point x is a Lebesgue point for f, in the sense that $\lim_{r\to 0}\fint_{B(x,r)}|f(y)-f(x)|dy=0$, which also implies $f(x)=\lim_{r\to 0}\fint_{B(x,r)}f(y)\,\mathrm{d}y$. If this is applied to $f=\mathbb{1}_E$, then one recovers the notion of Lebesgue points of a set (also called density points).

Lemma 3.22. *The plan $\overline{\gamma}$ is concentrated on a c-CM set $\tilde{\Gamma}$ such that for every $(x_0,y_0)\in\tilde{\Gamma}$ and for every $\varepsilon,\delta>0$, every unit vector ξ and every sufficiently small $r>0$, there are a point $x\in\left(B(x_0,r)\setminus B(x_0,\frac{r}{2})\right)\cap C(x_0,\xi,\delta)$ and a point $y\in B(y_0,\varepsilon)$ such that $(x,y)\in\tilde{\Gamma}$, where $C(x_0,\xi,\delta)$ is the following convex cone:*

$$C(x_0,\xi,\delta):=\{x:(x-x_0)\cdot\xi>(1-\delta)|x-x_0|\}.$$

Proof. First, take a c-CM set Γ such that $\overline{\gamma}(\Gamma)=1$. Then, let B_i be a countable family of balls in $\mathbb{R}^d$, generating the topology of $\mathbb{R}^d$ (for instance, all the balls such that the coordinates of the center and the radius are rational numbers). Consider now

$$A_i:=(\pi_x)(\Gamma\cap(\Omega\times B_i)),$$

i.e., the set of all points x such that at least a point y with $(x, y) \in \Gamma$ belongs to B_i (the points that have at least one "image" in B_i, if we imagine $\overline{\gamma}$ as a multivalued map). Then set $N_i := A_i \setminus \mathrm{Leb}(A_i)$. This set has zero Lebesgue measure, and hence it is μ-negligible. Also $\mu(\bigcup_i N_i) = 0$. As a consequence, one can define $\tilde{\Gamma} := \Gamma \setminus \big((\bigcup_i N_i) \times \Omega\big)$. The plan $\overline{\gamma}$ is still concentrated on $\tilde{\Gamma}$, since we only removed μ-negligible points. Obviously, $\tilde{\Gamma}$ stays c-CM and enjoys property (3.5), since it is a subset of Γ.

Moreover, $\tilde{\Gamma}$ has the property which is claimed in the statement. This is true since there is at least one of the balls B_i containing y_0 and contained in $B(y_0, \varepsilon)$. Since $x_0 \in A_i$ and we have removed N_i, this means that x_0 is a Lebesgue point for A_i. Since the region $\big(B(x_0, r) \setminus B(x_0, \frac{r}{2})\big) \cap C(x_0, \xi, \delta)$ is a portion of the ball $B(x_0, r)$ which takes a fixed proportion (depending on δ) of the volume of the whole ball, for $r \to 0$ it is clear that A_i (and also $\mathrm{Leb}(A_i)$) must meet it (otherwise x_0 would not be a Lebesgue point). It is then sufficient to pick a point in $\mathrm{Leb}(A_i) \cap \big(B(x_0, r) \setminus B(x_0, \frac{r}{2})\big) \cap C(x_0, \xi, \delta)$ and we are done. □

Lemma 3.23. *If (x_0, y_0) and (x_0, z_0) belong to $\tilde{\Gamma}$, then $y_0 = z_0$ (Figure 3.4).*

Proof. Suppose by contradiction $y_0 \neq z_0$. In order to fix the ideas, let us suppose $|x_0 - z_0| \leq |x_0 - y_0|$ (and in particular $y_0 \neq x_0$, since otherwise $|x_0 - z_0| = |x_0 - y_0| = 0$ and $z_0 = y_0$).

Now, use the property of $\tilde{\Gamma}$ and find $(x, y) \in \tilde{\Gamma}$ with $y \in B(y_0, \varepsilon)$ and $x \in \big(B(x_0, r) \setminus B(x_0, \frac{r}{2})\big) \cap C(x_0, \xi, \delta)$, for a vector ξ to be determined later. Use now the fact that $\tilde{\Gamma}$ is c-CM applied to (x_0, z_0) and (x, y).

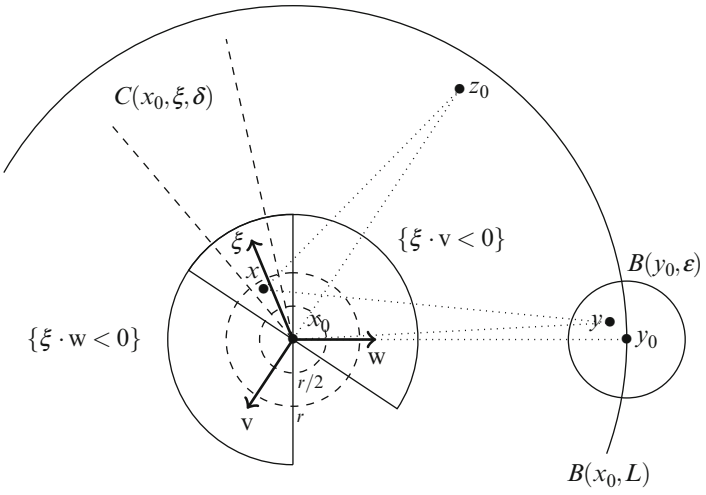


Fig. 3.4 The points in the proof of Lemma 3.23

If we can prove that $|x - z_0|, |x_0 - y| \le L$, then we should have

$$(x - x_0) \cdot (y - z_0) \ge 0$$

and we would use this to get a contradiction, by means of a suitable choice for ξ. Indeed, the direction of $x - x_0$ is almost that of ξ (up to an error of δ) and that of $y - z_0$ is almost that of $y_0 - z_0$ (up to an error of the order of $\varepsilon/|z_0 - y_0|$). If we choose ξ such that $\xi \cdot (y_0 - z_0) < 0$, this means that for small δ and ε, we would get a contradiction.

We need to guarantee $|x - z_0|, |x_0 - y| \le L$, in order to prove the claim. Let us compute $|x - z_0|^2$: we have

$$|x - z_0|^2 = |x_0 - z_0|^2 + |x - x_0|^2 + 2(x - x_0) \cdot (x_0 - z_0).$$

In this sum, we have

$$|x_0 - z_0|^2 \le L^2; \quad |x - x_0|^2 \le r^2; \quad 2(x - x_0) \cdot (x_0 - z_0) = |x - x_0|(\xi \cdot (x_0 - z_0) + O(\delta)).$$

Suppose now that we are in one of the following situations: either choose ξ such that $\xi \cdot (x_0 - z_0) < 0$ or $|x_0 - z_0|^2 < L^2$. In both cases, we get $|x - z_0| \le L$ for r and δ small enough. In the first case, since $|x - x_0| \ge \frac{r}{2}$, we have a negative term of the order of r and a positive one of the order of r^2; in the second, we add to $|x_0 - z_0|^2 < L^2$ some terms of the order of r or r^2.

Analogously, for what concerns $|x_0 - y|$, we have

$$|x_0 - y|^2 = |x_0 - x|^2 + |x - y|^2 + 2(x_0 - x) \cdot (x - y).$$

The three terms satisfy

$$|x_0 - x|^2 \le r^2; \quad |x - y|^2 \le L^2;$$
$$2(x_0 - x) \cdot (x - y) = |x - x_0|(-\xi \cdot (x_0 - y_0) + O(\delta + \varepsilon + r)).$$

In this case, either we have $|x_0 - y_0| < L$ (which would guarantee $|x_0 - y| < L$ for ε, δ small enough) or we need to impose $\xi \cdot (y_0 - x_0) < 0$.

Note that imposing $\xi \cdot (y_0 - x_0) < 0$ and $\xi \cdot (x_0 - z_0) < 0$ automatically gives $\xi \cdot (y_0 - z_0) < 0$, which is the desired condition so as to have a contradiction. Set $\mathbf{v} = y_0 - x_0$ and $\mathbf{w} = x_0 - z_0$. If it is possible to find ξ with $\xi \cdot \mathbf{v} < 0$ and $\xi \cdot \mathbf{w} < 0$, we are done. When is it the case that two vectors $\mathbf{v}$ and $\mathbf{w}$ do not admit the existence of a vector ξ with both scalar products that are negative? The only case is when they go in opposite directions. But this would mean that x_0, y_0, and z_0 are colinear, with z_0 between x_0 and y_0 (since we supposed $|x_0 - z_0| \le |x_0 - y_0|$). If we want z_0 and y_0 to be distinct points, we should have $|x_0 - z_0| < L$. Hence, in this case, we do not need to check $\xi \cdot (x_0 - z_0) < 0$. We only need a vector ξ satisfying $\xi \cdot (y_0 - x_0) < 0$ and $\xi \cdot (y_0 - z_0) < 0$, but the directions of $y_0 - x_0$ and $y_0 - z_0$ are the same, so that this

can be guaranteed by many choices of ξ, since we only need the scalar product with just one direction to be negative. Take, for instance, $\xi = -(y_0 - x_0)/|y_0 - x_0|$. □

Theorem 3.24. *The secondary variational problem*

$$\min\{K_2(\gamma) \,:\, \gamma \in O_\infty(\mu, \nu)\}$$

admits a unique solution $\overline{\gamma}$*, it is induced by a transport map* T*, and such a map is an optimal transport for the problem*

$$\min\{||\mathrm{T} - \mathrm{id}||_{L^\infty(\mu)} \,,\, \mathrm{T}_\#\mu = \nu\}.$$

Proof. We have already seen that $\overline{\gamma}$ is concentrated on a set $\tilde{\Gamma}$ satisfying some useful properties. Lemma 3.23 shows that $\tilde{\Gamma}$ is contained in a graph, since for any x_0, there is no more than one possible point y_0 such that $(x_0, y_0) \in \tilde{\Gamma}$. Let us consider such a point y_0 as the image of x_0 and call it $\mathrm{T}(x_0)$. Then $\overline{\gamma} = \gamma_{\mathrm{T}}$. The optimality of T in the "Monge" version of this L^∞ problem comes from the usual comparison with the Kantorovich version on plans γ. The uniqueness comes from standard arguments (since it is true as well that convex combinations of minimizers should be minimizers and this allows us to perform the proof in Remark 1.19). □

3.3 Discussion

3.3.1 Different norms and more general convex costs

The attentive reader will have observed that the proof in Section 3.1 about the case of the distance cost function was specific to the case of the distance induced by the Euclidean norm. Starting from the original problem by Monge and relying on the proof strategy by Sudakov [290] (which had a gap later solved by Ambrosio in [8, 10], but was meant to treat the case of an arbitrary norm), the case of different norms has been extensively studied in the last years. Note that the case of uniformly convex norms (i.e., those such that the Hessian of the square of the norm is bounded from below by a matrix which is positively definite) is the one which is most similar to the Euclidean case and can be treated in a similar way.

One of the first extensions was the case studied by Ambrosio, Kirchheim, and Pratelli about crystalline norms [14], i.e., norms such that their unit balls are convex polyhedra. One can guess that the fact that we have a finite number of faces makes the task easier, but it already required a huge approximation work. Indeed, if one uses duality, the gradient $\nabla u(x)$ of a Kantorovich potential is not enough to determine the direction of the displacement $y - x$ for a pair $(x, y) \in \mathrm{spt}(\gamma)$, because a crystalline norm has the same gradient on all the points of a same face (and not only on points on the same ray, as is the case for the Euclidean norm, or for other strictly convex norms). To handle this problem, [14] develops a strategy with a

double approximation, where $||z||$ is replaced by something like $||z|| + \varepsilon|z| + \varepsilon^2|z|^2$, $||z||$ is the norm we consider, $|z|$ is the Euclidean norm which is added so as to select a direction, and its square is added so as to get strict convexity and monotonicity on each ray (exactly as we saw in Section 3.1). In this way, the authors prove the existence of an optimal map for crystalline norms in any dimensions and for arbitrary norms in $\mathbb{R}^2$.

Later, the way was open to the generalization to other norms. The main tool is always a secondary variational problem, since for norm costs it is in general false that optimal plans all come from a map, and one needs to select a special one. The difficulty with secondary variational problems is that they correspond to a new cost c which is not finitely valued. This prevents from using duality on the secondary problem. Indeed, Proposition 1.42 proves that we have equality between the primal and the dual problem for all costs which are l.s.c., but does not guarantee the existence of an optimizer in the dual; existence is usually guaranteed either via Ascoli-Arzelà arguments, or through Theorem 1.37, which requires finiteness of the cost. This is why some methods avoiding duality and concentrating on c-cyclical monotonicity have been developed. The first time they appeared is in [122], for a different problem (an L^∞ case): it corresponds to the idea that we presented in Section 3.2 (looking for density points of this concentration set). Later, Champion and De Pascale managed to use the same tools to prove the existence of an optimal map first for arbitrary strictly convex norms (in [119], the same result being obtained differently at almost the same time by Caravenna in [103]) and then for general norms in [120]. This last result was more difficult to obtain and required some extra approximation tools inspired from [273] and in particular the selection of a special optimal transport plan via approximation through atomic marginals, as we did in Section 3.1.5.

But the history of optimal transport did not look only at norms. Many studies have been done for different cost functions (distances or squared distances on manifolds, geodesic distances on $\mathbb{R}^d$ if obstacle are present, etc.). In this book, we prefer to stick to the Euclidean case, and in this section, we only consider costs of the form $c(x, y) = h(x - y)$ for h convex. Even this case is far from being completely understood. In a paper with Carlier and De Pascale (see [113]), a general (straightforward) strategy of decomposition according to the "faces" of the cost function h is presented.

The decomposition is based on the following steps:

- Consider an optimal plan γ and look at the optimality conditions as in Section 1.3. For all $(x_0, y_0) \in \mathrm{spt}(\gamma)$, if x_0 is a differentiability point for the potential φ (we write $x_0 \in \mathrm{Diff}(\varphi)$), one gets $\nabla\varphi(x_0) \in \partial h(x_0 - y_0)$, which is equivalent to

$$x_0 - y_0 \in \partial h^*(\nabla\varphi(x_0)). \tag{3.6}$$

Let us define

$$\mathscr{F}_h := \{\partial h^*(p) \ : \ p \in \mathbb{R}^d\},$$

which is the set of all possible values of the subdifferential multi-map of h^*. These values are those convex sets where the function h is affine, and they will be called *faces* of h. The fact that φ is differentiable μ-a.e. is enforced by supposing $\mu \ll \mathscr{L}^d$ and h to be Lipschitz, so that φ is also Lipschitz.
Thanks to (3.6), for every fixed x, all the points y such that (x, y) belongs to the support of an optimal transport plan are such that the difference $x - y$ belongs to a same face of h. Classically, when these faces are singletons (i.e., when h^* is differentiable, which is the same as h being strictly convex), this is the way to obtain a transport map, since only one y is admitted for every x.
Equation (3.6) also enables one to classify the points x as follows. For every $K \in \mathscr{F}_h$, we define the set

$$X_K := \{x \in \mathrm{Diff}(\varphi) \ : \ \partial h^*(\nabla\varphi(x)) = K\}.$$

Hence, γ may be decomposed into several subplans γ_K according to the criterion $x \in X_K$. If K varies among all possible faces, this decomposition covers γ-almost all pairs (x, y). Moreover, if (x, y) belongs to $\mathrm{spt}(\gamma)$ and x to $\mathrm{Diff}(\varphi)$, then $x \in X_K$ implies $x - y \in K$.
If the set $\mathscr{F}_h$ is finite or countable, we define

$$\gamma_K := \gamma_{|X_K \times \mathbb{R}^d}.$$

In this case, the marginal measures μ_K and ν_K of γ_K (i.e., its images under the maps π_x and π_y, respectively) are sub-measures of μ and ν, respectively. In particular μ_K inherits the absolute continuity from μ. This is often useful for proving existence of transport maps.
If $\mathscr{F}_h$ is uncountable, in some cases, one can still rely on a countable decomposition by considering the set $\mathscr{F}_h^{multi} := \{K \in \mathscr{F}_h \ : \ K \text{ is not a singleton }\}$. If $\mathscr{F}_h^{multi}$ is countable, then one can separate those x such that $\partial h^*(\nabla(\varphi(x))$ is a singleton (where a transport already exists) and look at a decomposition for $K \in \mathscr{F}_h^{multi}$ only.

- This decomposition reduces the transport problem to a superposition of transport problems of the type

$$\min\left\{\int h(x-y)\,\mathrm{d}\gamma(x,y) \ : \ \gamma \in \Pi(\mu_K, \nu_K),\ \mathrm{spt}(\gamma) \subset \{x - y \in K\}\right\}.$$

The advantage is that the cost c restricted to K is easier to study. If K is a face of h, then h is affine on K and in this case, the transport cost does not depend any more on the transport plan.

- The problem is reduced to find a transport map from μ_K to ν_K satisfying the constraint $x - \mathrm{T}(x) \in K$, knowing a priori that a transport plan satisfying the same constraint exists.
  In some cases (e.g., if K is a convex compact set containing 0 in its interior), this problem may be reduced to an L^∞ transport problem. In fact if one denotes by $||\cdot||_K$ the (gauge-like) "norm" such that $K = \{x : ||x||_K \leq 1\}$, one has

$$\min\left\{\max\{||x-y||_K : (x,y) \in \mathrm{spt}(\gamma)\} : \gamma \in \Pi(\mu,\nu)\right\} \leq 1 \tag{3.7}$$

  and the question is whether the same inequality would be true if one restricted the admissible set to transport maps only (passing from Kantorovich to Monge, say). The answer would be yes if a solution of (3.7) were induced by a transport map T. This is what we presented in Section 3.2 in the case $K = \overline{B(0,1)}$ and which was first proven in [122] via a different strategy (instead of selecting a minimizer via a secondary variational problem, selecting the limit of the minimizers for the L^p norms or $p \to \infty$). Note also that this issue is easy in 1D, where the monotone transport solves all the L^p problems and hence the L^∞ as well (and this does not need the measure to be absolutely continuous, but just atomless).

Let us note that the assumption that the number of faces is countable is quite restrictive and is essentially used to guarantee the absolute continuity of μ_K, with no need of a disintegration argument (which could lead to difficulties at least as hard as those faced by Sudakov those faced by Sudakov, see also [104]). However, an interesting example that could be approached by finite decomposition is that of crystalline norms. In this case, the faces of the cost h are polyhedral cones but, if the support of the two measures are bounded, we can suppose that they are compact convex polyhedra. This means, thanks to the considerations above, that it is possible to perform a finite decomposition and to reduce the problem to some L^∞ minimizations for norms whose unit balls are polyhedra (the faces of the cone). In particular, the L^1 problem for crystalline norms is solved if we can solve L^∞ optimal transport problem for other crystalline norms.

It becomes then interesting to solve the L^∞ problem as in Section 3.2, replacing the unit ball constraint $|x-y| \leq 1$ with a more general constraint $x - y \in C$, the set C being a generic convex set, for instance, a polyhedron. This is studied in [197] by minimizing the quadratic cost

$$c(x,y) = \begin{cases} |x-y|^2 & \text{if } x-y \in C, \\ +\infty & \text{otherwise,} \end{cases}$$

which is also of the form $c(x,y) = h(x-y)$ for h strictly convex (but not real valued). The existence of an optimal map (better: the fact that any optimal γ for this problem is induced by a map) is proven when μ is absolutely continuous and C is either strictly convex or has a countable number of faces. This can be achieved by adapting the arguments of Section 3.2 and proving that, if (x_0, y_0) and (x_0, y_1) belong

to spt(γ) (and hence $y_0 - x_0, y_1 - x_0 \in C$), then the middle point $\frac{1}{2}(y_0 + y_1) - x_0$ cannot lie in the interior of C. If C is strictly convex, this proves $y_0 = y_1$. If not, this proves that the points $y_0 - x_0$ and $y_1 - x_0$ are on a same face of C and one can perform a dimensional reduction argument and proceed by induction.

In particular, this argument applies to the case of polyhedra and of arbitrary convex sets in $\mathbb{R}^2$ (since no more than a countable quantity of segments may be contained in ∂C) and provides an alternative proof for the result of [14].

Many generalizations of this last result have been performed: Bertrand and Puel [52] studied the "relativistic" cost $h(z) := 1-\sqrt{1-|z|^2}$, which is naturally endowed with the constraint $|z| \le 1$ and is a strictly convex function (but not finite valued since the constraint means $h(z) = +\infty$ outside the unit ball). They proved existence of an optimal map by adapting the arguments of [122, 197] to the case of a strictly convex function h with a strictly convex constraint C (from there, adapting to the case of a countable number of faces seems to be also possible).

Finally, Chen, Jiang, and Yang combined the arguments of [197] with those of [120], thus getting existence for the cost function

$$c(x,y) = \begin{cases} |x-y| & \text{if } x-y \in C, \\ +\infty & \text{otherwise,} \end{cases}$$

under the usual assumptions on C (strictly convex or countable number of faces). These results can be found in [123, 124] and are the first which combine lack of strict convexity and infinite values.

All these subsequent improvements paved the way to an obvious conjecture, i.e., the fact that a mixture of duality, variational approximation, and density points techniques could finally prove the existence of an optimal transport map in a very general case:

Problem (existence for every convex cost): suppose $\mu \ll \mathscr{L}^d$ and take a cost $c(x,y)$ of the form $h(x-y)$ with $h : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ convex; prove that there exists an optimal map.

This problem was open till the final redaction phases of this book, when Bardelloni and Bianchini's paper [26] appeared, containing a proof of a disintegration result implying this very same statement, at least in the case where h is finite valued but not strictly convex.

3.3.2 Concave costs (L^p, with $0 < p < 1$)

Another class of transport costs which is very reasonable for applications, rather than convex functions of the Euclidean distance, is that of concave costs[3], more

[3] We mean here costs which are concave functions of the distance $|x-y|$, not of the displacement $x-y$, as instead we considered in Remark 2.12.

precisely $c(x,y) = \ell(|x-y|)$ where $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ is a strictly concave and increasing function[4]. From the economic and modeling point of view, this is the most natural choice: moving a mass has a cost which is proportionally less if the distance increases, as everybody can note from travel fares of most transport means (railways, highway tolls, flights, etc.). All these costs are subadditive. In many practical cases, moving two masses, each on a distance d, is more expensive than moving one at distance $2d$ and keeping at rest the other. The typical example of costs with these property is given by the power cost $|x-y|^\alpha$, $\alpha < 1$.

Note that all these costs satisfy the triangle inequality and are thus distances on $\mathbb{R}^d$. Moreover, under strict concavity assumptions, these costs satisfy a strict triangle inequality. This last fact implies that the common mass between μ and ν must stay at rest, a fact first pointed out in [176].

Theorem 3.25. *Let γ be an optimal transport plan for the cost $c(x,y) = \ell(|x-y|)$ with $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ strictly concave, increasing, and such that $\ell(0) = 0$. Let $\gamma = \gamma_D + \gamma_O$, where γ_D is the restriction of γ to the diagonal $D = \{(x,x) : x \in \Omega\}$ and γ_O is the part outside the diagonal, i.e., the restriction to $D^c = (\Omega \times \Omega) \setminus D$. Then this decomposition is such that $(\pi_x)_\# \gamma_O$ and $(\pi_y)_\# \gamma_O$ are mutually singular measures.*

Proof. It is clear that γ_O is concentrated on $\mathrm{spt}(\gamma) \setminus D$ and hence $(\pi_x)_\# \gamma_O$ is concentrated on $\pi_x(\mathrm{spt}(\gamma) \setminus D)$ and $(\pi_y)_\# \gamma_O$ on $\pi_y(\mathrm{spt}(\gamma) \setminus D)$. We claim that these two sets are disjoint. Indeed suppose that a common point z belongs to both. Then, by definition, there exists y such that $(z,y) \in \mathrm{spt}(\gamma) \setminus D$ and x such that $(x,z) \in \mathrm{spt}(\gamma) \setminus D$. This means that we can apply c-cyclical monotonicity to the points (x,z) and (z,y) and get

$$\ell(|x-z|) + \ell(|z-y|) \le \ell(|x-y|) + \ell(|z-z|) = \ell(|x-y|) < \ell(|x-z|) + \ell(|z-y|),$$

where the last strict inequality gives a contradiction. □

This gives a first constraint on how to build optimal plans γ: look at μ and ν, take the common part $\mu \wedge \nu$, leave it on place, subtract it from the rest, and then build an optimal transport between the two remainders, which will have no mass in common. Note that when the cost c is linear in the Euclidean distance, then the common mass *may* stay at rest but is not forced to do so (think at the book-shifting example); on the contrary, when the cost is a strictly convex function of the Euclidean distance, in general, the common mass *does not* stay at rest (in the previous example, only the translation is optimal for $c(x,y) = |x-y|^p$, $p > 1$). Note that the fact that the common mass stays at rest implies that in general, there is no optimal map T, since whenever there is a set A with $\mu(A) > (\mu \wedge \nu)(A) = \nu(A) > 0$, then almost all the points of A must have two images: themselves and another point outside A.

[4] We will not explicitly state it every time, but this also implies that ℓ is strictly increasing.

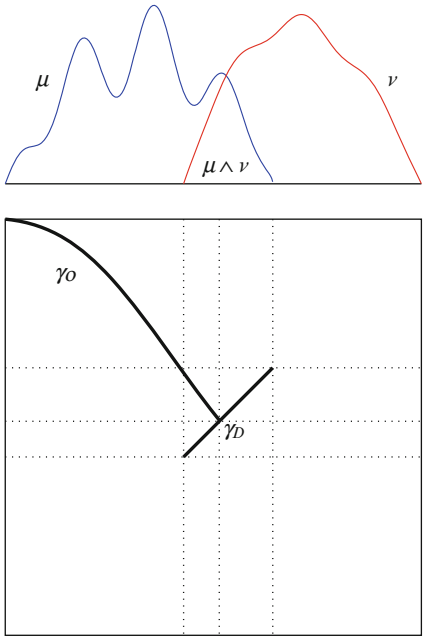


Fig. 3.5 The optimal plan in 1D in the case of common mass

A typical case is represented in Figure 3.5 in the 1D case, where the transport map (after removing the common mass) is decreasing as a consequence of Remark 2.12. Note that the precise structure of the optimal transport map for this kind of "concave" costs in 1D follows a sort of hierarchical construction, first investigated in [230] and then used in [141] for numerical purposes.

This suggests to study the case where μ and ν are mutually singular, and the best that one could do would be to prove the existence of an optimal map in this case. In particular, this allows us to avoid the singularity of the function $(x, y) \mapsto \ell(|x - y|)$ concentrated on the diagonal $\{x = y\}$ (look at the example $|x - y|^\alpha$), since when the two measures have no common mass, almost no point x is transported to $y = x$.

Yet, exploiting this fact needs some attention. The easiest case is when μ and ν have disjoint supports, since in this case, there is a lower bound on $|x - y|$ and this allows to stay away from the singularity. Yet, $\mathrm{spt}(\mu) \cap \mathrm{spt}(\nu) = \emptyset$ is too restrictive, since even in the case where μ and ν have smooth densities f and g, it may happen that, after subtracting the common mass, the two supports meet on the region $\{f = g\}$.

The problem has been solved in the classical paper by Gangbo and McCann, [176], one of the first papers about optimal transportation. In such a paper, the authors used the slightly less restrictive assumption $\mu(\mathrm{spt}(\nu)) = 0$. This assumption covers the example above of two continuous densities, but does not cover many other cases. Think at μ being the Lebesgue measure on a bounded domain Ω and ν being an atomic measure with an atom at each rational point, or other examples that one can build with fully supported absolutely continuous measures concentrated on

disjoint sets A and $\Omega \setminus A$ (see **Ex**(20)). In the present section, we show how to prove the existence of a transport map when μ and ν are singular to each other and $\mu \ll \mathscr{L}^d$.

We will prove the following result.

Theorem 3.26. *Suppose that μ and ν are two mutually singular probability measures on $\mathbb{R}^d$ such that $\mu \ll \mathscr{L}^d$, and take the cost $c(x,y) = \ell(|x-y|)$, for $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ strictly concave, C^1 on $]0,+\infty[$, and increasing. Then there exists a unique optimal transport plan γ and it is induced by a transport map.*

Proof. We will use the standard procedure explained in Chapter 1 for costs of the form $c(x,y) = h(x-y)$. In particular, we get the existence of a Kantorovich potential φ such that, if $(x_0, y_0) \in \mathrm{spt}(\gamma)$, then

$$x_0 - y_0 = (\nabla h)^{-1}(\nabla \varphi(x_0)). \tag{3.8}$$

This allows us to express y_0 as a function of x_0, thus proving that there is only one point $(x_0, y) \in \mathrm{spt}(\gamma)$ and hence that γ comes from a transport $\mathrm{T}(x) = x - (\nabla h)^{-1}(\nabla \varphi(x))$. This approach also proves uniqueness in the same way. In Chapter 1, we presented it under strict convexity assumptions on h, so that ∇h is injective and $(\nabla h)^{-1} = \nabla h^*$. But the injectivity is also true if $h(z) = \ell(|z|)$. Indeed we have $\nabla h(z) = \ell'(|z|)\frac{z}{|z|}$, and the modulus of this vector identifies the modulus $|z|$ (since ℓ' is strictly increasing), and the direction identifies the direction of z.

The main difficulty is the fact that we need to guarantee that φ is differentiable a.e. with respect to μ. Since $\mu \ll \mathscr{L}^d$, it is enough to have φ Lipschitz continuous, which is usually proven using the fact that $\varphi(x) = \inf_y h(|x-y|) - \psi(y)$. Yet, concave functions on $\mathbb{R}^+$ may have an infinite slope at 0 and be non-Lipschitz, and this could be the case for φ as well. This suggests the use of an alternate notion of differentiability.

Box 3.3. ***Important notion**: Approximate gradient*

We recall here some facts about a measure-theoretical notion replacing the gradient for less regular functions. The interested reader can find many details in [160].

Let us start from the following observation: given a function $f : \Omega \to \mathbb{R}$ and a point $x_0 \in \Omega$, we say that f is differentiable at $x_0 \in \Omega$ and that its gradient is $\mathbf{v} = \nabla f(x_0) \in \mathbb{R}^d$ if for every $\epsilon > 0$ the set

$$A(x_0, \mathbf{v}, \varepsilon) := \{x \in \Omega \,:\, |f(x) - f(x_0) - \mathbf{v}\cdot(x-x_0)| > \epsilon\,|x-x_0|\}$$

is at positive distance from x_0, i.e., if for small $\delta > 0$ we have $B(x_0,\delta) \cap A(x_0,\mathbf{v},\varepsilon) = \emptyset$. Instead of this requirement, we could ask for a weaker condition, namely, that x_0 is a zero-density point for the same set $A(\mathbf{v},\varepsilon)$ (i.e., a Lebesgue point of its complement). More precisely, if there exists a vector $\mathbf{v}$ such that

(continued)

Box 3.3. (continued)

$$\lim_{\delta \to 0} \frac{|A(x_0, \mathbf{v}, \varepsilon) \cap B(x_0, \delta)|}{|B(x_0, \delta)|} = 0$$

then we say that f is approximately differentiable at x_0 and its approximate gradient is $\mathbf{v}$. The approximate gradient will be denoted by $\nabla_{\text{app}} f(x_0)$. As one can expect, it enjoys several of the properties of the usual gradient that we list here.

- The approximate gradient, provided it exists, is unique.
- The approximate gradient is nothing but the usual gradient if f is differentiable.
- The approximate gradient shares the usual algebraic properties of gradients, in particular $\nabla_{\text{app}}(f+g)(x_0) = \nabla_{\text{app}} f(x_0) + \nabla_{\text{app}} g(x_0)$.
- If x_0 is a local minimum or local maximum for f, and if $\nabla_{\text{app}} f(x_0)$ exists, then $\nabla_{\text{app}} f(x_0) = 0$.

Another very important property is a consequence of the Rademacher theorem.

Proposition. Let $f, g : \Omega \to \mathbb{R}$ be two functions defined on a same domain Ω with g Lipschitz continuous. Let $E \subset \Omega$ be a Borel set such that $f = g$ on E. Then f is approximately differentiable almost everywhere on E and $\nabla_{\text{app}} f(x) = \nabla g(x)$ for a.e. $x \in E$.

Proof. It is enough to consider all the Lebesgue points of E where g is differentiable. These points cover almost all E. It is easy to check that the definition of approximate gradient of f at a point x_0 is satisfied if we take $\mathbf{v} = \nabla g(x_0)$.

As a consequence, it is also clear that every countably Lipschitz function is approximately differentiable a.e.

We just need to prove that φ admits an approximate gradient Lebesgue-a.e.: this would imply that Equation (3.8) is satisfied if we replace the gradient with the approximate gradient.

Recall that we may suppose

$$\varphi(x) = \varphi^{cc}(x) = \inf_{y \in \mathbb{R}^d} \ell(|x-y|) - \varphi^c(y) \, .$$

Now consider a countable family of closed balls B_i generating the topology of $\mathbb{R}^d$, and for every i, consider the function defined as

$$\varphi_i(x) := \inf_{y \in B_i} \ell(|x-y|) - \varphi^c(y)$$

for $x \in \mathbb{R}^d$. One cannot provide straight Lipschitz properties for φ_i, since a priori y is arbitrarily close to x and in general ℓ is not Lipschitz close to 0. However, φ_i is Lipschitz on every B_j such that $\text{dist}(B_i, B_j) > 0$. Indeed if $x \in B_j$, $y \in B_i$ one has $|x-y| \geq d > 0$; therefore, the Lipschitz constant of $\ell(|\cdot - y|) - \varphi^c(y)$ does not exceed $\ell'(d)$. It follows that φ_i is Lipschitz on B_j, and its Lipschitz constant does not exceed $\ell'(d)$ and the same is true for φ_i has an approximate gradient almost everywhere on $\{\varphi = \varphi_i\} \cap B_j$. By countable union, φ admits an approximate gradient a.e. on

$$\bigcup_{\substack{i,j \\ d(B_i,B_j)>0}} [\{\varphi_i = \varphi\} \cap B_j] \, .$$

In order to prove that φ has an approximate gradient μ-almost everywhere, it is enough to prove that

$$\mu\left(\bigcup_{\substack{i,j \\ d(B_i,B_j)>0}} \{\varphi_i = \varphi\} \cap B_j\right) = 1.$$

In order to do this, note that for every i and j, we have

$$\pi_x(\mathrm{spt}(\gamma) \cap (B_j \times B_i)) \subset \{\varphi = \varphi_i\} \cap B_j \, .$$

Indeed, let $(x, y) \in \mathrm{spt}(\gamma) \cap (B_j \times B_i)$. Then $\varphi(x) + \varphi^c(y) = l(|x - y|)$. It follows that

$$\varphi_i(x) = \inf_{y' \in B_i} \ell(|x - y'|) - \varphi^c(y') \leq \ell(|x - y|) - \varphi^c(y) = \varphi(x) \, .$$

On the other hand, for every $x \in \mathbb{R}^d$

$$\varphi_i(x) = \inf_{y \in B_i} \ell(|x - y|) - \varphi^c(y) \geq \inf_{y \in \mathbb{R}^d} \ell(|x - y|) - \varphi^c(y) = \varphi(x) \, .$$

As a consequence of this,

$$\begin{aligned}
\mu\left(\bigcup_{\substack{i,j \\ d(B_i,B_j)>0}} \{\varphi_i = \varphi\} \cap B_j\right) &\geq \mu\left(\bigcup_{\substack{i,j \\ d(B_i,B_j)>0}} \pi_x(\mathrm{spt}(\gamma) \cap (B_j \times B_i))\right) \\
&= \mu\left(\pi_x\left(\mathrm{spt}(\gamma) \cap \bigcup_{\substack{i,j \\ d(B_i,B_j)>0}} B_j \times B_i\right)\right) \\
&= \mu(\pi_x(\mathrm{spt}(\gamma) \setminus D)) \\
&= \gamma\left[(\pi_x)^{-1}(\pi_x(\mathrm{spt}(\gamma) \setminus D))\right] \\
&\geq \gamma(\mathrm{spt}(\gamma) \setminus D) = 1,
\end{aligned}$$

where D is the diagonal in $\Omega \times \Omega$. □

From the previous theorem, we can also easily deduce the following extension. Define $\mu \wedge \nu$ as the maximal positive measure which is both less or equal than μ and than ν, and $(\mu - \nu)_+ = \mu - \mu \wedge \nu$, so that the two measures μ and ν uniquely decompose into a common part $\mu \wedge \nu$ and two mutually singular parts $(\mu - \nu)_+$ and $(\nu - \mu)_+$.

Theorem 3.27. *Suppose that μ and ν are probability measures on Ω with $(\mu-\nu)_+ \ll \mathscr{L}^d$, and take the cost $c(x,y) = \ell(|x-y|)$, for $\ell : \mathbb{R}_+ \to \mathbb{R}_+$ strictly concave, C^1 on $]0,+\infty[$, and increasing with $\ell(0) = 0$. Then there exists a unique optimal transport plan γ, and it has the form* $(\mathrm{id},\mathrm{id})_\#(\mu \wedge \nu) + (\mathrm{id},\mathrm{T})_\#(\mu-\nu)_+$.

As we said, the above results can be extended to more general situations: differentiability of ℓ is not really necessary, and the assumption $\mu \ll \mathscr{L}^d$ can be weakened. In [250], the result is proven under the natural assumption that μ does not give mass to small sets, i.e., for every $A \subset \mathbb{R}^d$ which is $(d-1)$-rectifiable, we have $\mu(A) = 0$. The key tool is the following lemma, which is an interesting result from geometric measure theory that can be used instead of Lebesgue points-type results when we face a measure which is not absolutely continuous but "does not give mass to small sets." It states that, in such a case, for μ-a.e. point x, every cone exiting from x, even if very small, has positive μ mass. In particular, it means that we can find points of $\mathrm{spt}(\mu)$ in almost arbitrary directions close to x. This lemma is strangely more known among people in optimal transport (see also [121]) than in geometric measure theory (even if it corresponds more or less to a part of Lemma 3.3.5 in [161]).

Lemma 3.28. *Let μ be a Borel measure on $\mathbb{R}^d$, and suppose that μ does not give mass to small sets. Then μ is concentrated on the set*

$$\{x : \forall \epsilon > 0, \forall \delta > 0, \forall e \in \mathbb{S}^{d-1},\ \mu(C(x,e,\delta,\epsilon)) > 0\}\,,$$

where

$$C(x,e,\delta,\epsilon) = C(x,e,\delta) \cap \overline{B(x,\epsilon)} := \{y : \langle y-x, e\rangle \geq (1-\delta)\,|y-x|\} \cap \overline{B(x,\epsilon)}.$$

Chapter 4
Minimal flows, divergence constraints, and transport density

4.1 Eulerian and Lagrangian points of view

We review here the main languages and the main mathematical tools to describe static or dynamical transport phenomena.

4.1.1 Static and dynamical models

This section presents a very informal introduction to the physical interpretation of dynamical models in optimal transport.

In fluid mechanics, and in many other topics with similar modeling issues, it is classical to consider two complementary ways of describing motions, which are called Lagrangian and Eulerian.

When we describe a motion via Lagrangian formalism, we give "names" to particles (using either a specific label or their initial position, for instance) and then describe, for every time t and every label, what happens to that particle. "What happens" means providing its position and/or its velocity. Hence we could, for instance, look at trajectories $y_x(t)$, standing for the position at time t of particle originally located at x. As another possibility, instead of giving names, we could consider bundles of particles with the same behavior and indicate how many are they. This amounts to giving a measure on possible behaviors.

The description may be more or less refined. For instance, if one only considers two different times $t = 0$ and $t = 1$, the behavior of a particle is only given by its initial and final positions. A measure on those pairs (x, y) is exactly a transport plan. This explains why we can consider that the Kantorovich problem is expressed in Lagrangian coordinates. The Monge problem is also Lagrangian, and particles are labeled by their initial position.

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2_4

More refined models can be easily conceived: reducing a movement to the initial and final positions is embarrassingly poor! Measures on the set of paths (curves $\omega : [0,1] \to \Omega$, with possible assumptions on their regularity) have been used in many modeling and in particular in traffic issues in branched transport (see Section 4.4 for both these subjects), or in Brenier's variational formulation of the incompressible Euler equations for fluids (see Sections 1.7.4, 8.4.4, and [51, 83, 85]).

On the other hand, in the Eulerian formalism, we describe, for every time t and every point x, what happens at such a point at such a time. "What happens" usually means what are the velocity, the density, and the flow rate (both in intensity and in direction) of particles located at time t at point x.

Eulerian models may be distinguished into static and dynamical ones. In a dynamical model we usually use two variables, i.e., the density $\varrho(t,x)$ and the velocity $\mathbf{v}_t(x)$. It is possible to write the equation satisfied by the density of a family of particles moving according to the velocity field $\mathbf{v}$. This means that we prescribe the initial density ϱ_0 and that the position of the particle originally located at x will be given by the solution of the ODE

$$\begin{cases} y_x'(t) = \mathbf{v}_t(y_x(t)) \\ y_x(0) = x. \end{cases} \tag{4.1}$$

We define the map Y_t through $Y_t(x) = y_x(t)$, and we look for the measure $\varrho_t := (Y_t)_{\#}\varrho_0$. It is well known that ϱ_t and $\mathbf{v}_t$ solve together the so-called continuity equation

$$\partial_t \varrho_t + \nabla \cdot (\varrho_t \mathbf{v}_t) = 0$$

that is briefly addressed in Section 4.1.2.

The static framework is a bit harder to understand, since it is maybe not clear what "static" means when we want to describe movement. One has to think to a permanent, cyclical movement, where some mass is constantly injected into the motion at some points and constantly withdrawn somewhere else. We can also think at a time average of some dynamical model: suppose that we observe the traffic in a city and we wonder what happens at each point, but we do not want an answer depending on the hour of the day. We could, for instance, consider as a traffic intensity at every point the average traffic intensity at such a point on the whole day. In this case we usually use a unique variable $\mathbf{w}$ standing for the mass flow rate (which equals density times velocity: hence, it is rather a momentum than a velocity), and the divergence $\nabla \cdot \mathbf{w}$ stands for the excess of mass which is injected into the motion at every point. More precisely, if particles are injected into the motion according to a density μ and then exit with density ν, the vector fields $\mathbf{w}$ standing for flows connecting these two measures must satisfy

$$\nabla \cdot \mathbf{w} = \mu - \nu.$$

4.1.2 The continuity equation

This section is devoted to the equation

$$\partial_t \varrho_t + \nabla \cdot (\varrho_t \mathbf{v}_t) = 0,$$

its meaning, formulations, and uniqueness results. The remainder of the chapter will mainly deal with the static divergence equation, but we will see later that the dynamical case can also be useful (in particular, to produce transport plans associated with a given vector field). Hence, we need to develop some tools.

First, let us spend some words on the notion of solution for this equation. Here below $\Omega \subset \mathbb{R}^d$ is a bounded domain, or $\mathbb{R}^d$ itself.

Definition 4.1. We say that a family of pairs measures/vector fields $(\varrho_t, \mathbf{v}_t)$ with $\mathbf{v}_t \in L^1(\varrho_t; \mathbb{R}^d)$ and $\int_0^T ||\mathbf{v}_t||_{L^1(\varrho_t)} \mathrm{d}t = \int_0^T \int_\Omega |\mathbf{v}_t| \mathrm{d}\varrho_t \, \mathrm{d}t < +\infty$ solves the continuity equation on $]0,T[$ in the distributional sense if for any bounded and Lipschitz test function $\phi \in C^1_c(]0,T[\times\overline{\Omega})$ (note that we require the support to be far from $t = 0, 1$ but not from $\partial\Omega$, when Ω is bounded; also Ω is usually supposed to be itself closed, but we write $\overline{\Omega}$ to stress the fact that we do include its boundary), we have

$$\int_0^T \int_\Omega (\partial_t \phi) \, \mathrm{d}\varrho_t \, \mathrm{d}t + \int_0^T \int_\Omega \nabla \phi \cdot \mathbf{v}_t \, \mathrm{d}\varrho_t \, \mathrm{d}t = 0. \tag{4.2}$$

This formulation includes no-flux boundary conditions on $\partial\Omega$ for $\mathbf{v}_t$ (if Ω is not $\mathbb{R}^d$ itself, obviously), i.e., $\varrho_t \mathbf{v}_t \cdot \mathbf{n} = 0$. If we want to impose the initial and final measures, we can say that $(\varrho_t, \mathbf{v}_t)$ solves the same equation, in the sense of distribution, with initial and final data ϱ_0 and ϱ_T, respectively, if for any test function $\phi \in C^1_c([0,T] \times \overline{\Omega})$ (now we do not require the support to be far from $t = 0, 1$), we have

$$\int_0^T \int_\Omega (\partial_t \phi) \, \mathrm{d}\varrho_t \, \mathrm{d}t + \int_0^T \int_\Omega \nabla \phi \cdot \mathbf{v}_t \, \mathrm{d}\varrho_t \, \mathrm{d}t = \int_\Omega \phi(T,x) \, \mathrm{d}\varrho_T(x) - \int_\Omega \phi(0,x) \, \mathrm{d}\varrho_0(x). \tag{4.3}$$

We can also define a weak solution of the continuity equation through the following condition: we say that $(\varrho_t, \mathbf{v}_t)$ solves the continuity equation in the weak sense if for any test function $\psi \in C^1_c(\overline{\Omega})$, the function $t \mapsto \int \psi \, \mathrm{d}\varrho_t$ is absolutely continuous in t and, for a.e. t, we have

$$\frac{d}{dt} \int_\Omega \psi \, \mathrm{d}\varrho_t = \int_\Omega \nabla \psi \cdot \mathbf{v}_t \, \mathrm{d}\varrho_t.$$

Note that in this case $t \mapsto \varrho_t$ is automatically continuous for the weak convergence, and imposing the values of ϱ_0 and ϱ_1 may be done pointwisely in time.

Proposition 4.2. *The two notions of solutions are equivalent: every weak solution is actually a distributional solution and every distributional solution admits a representative (another family $\tilde{\varrho}_t = \varrho_t$ for a.e. t) which is weakly continuous and is a weak solution.*

Proof. To prove the equivalence, take a distributional solution, and test it against functions ϕ of the form $\phi(t,x) = a(t)\psi(x)$, with $\psi \in C^1_c(\overline{\Omega})$ and $a \in C^1_c(]0,1[)$. We get

$$\int_0^T a'(t) \int_\Omega \psi(x)\, \mathrm{d}\varrho_t\, \mathrm{d}t + \int_0^1 a(t) \int_\Omega \nabla\psi \cdot \mathbf{v}_t\, \mathrm{d}\varrho_t\, \mathrm{d}t = 0.$$

The arbitrariness of a shows that the distributional derivative (in time) of $\int_\Omega \psi(x)\, \mathrm{d}\varrho_t$ is $\int_\Omega \nabla\psi \cdot \mathbf{v}_t\, \mathrm{d}\varrho_t$. This last function is L^1 in time since $\int_0^T \left|\int_\Omega \nabla\psi \cdot \mathbf{v}_t\, \mathrm{d}\varrho_t\right| \mathrm{d}t \le \mathrm{Lip}\,\psi \int_0^T ||\mathbf{v}_t||_{L^1(\varrho_t)} \mathrm{d}t < +\infty$. This implies that $(\varrho, \mathbf{v})$ is a weak solution.

Conversely, the same computations show that weak solutions satisfy (4.2) for any ϕ of the form $\phi(t,x) = a(t)\psi(x)$. It is then enough to prove that finite linear combination of these functions is dense in $C^1([0,T] \times K)$ for every compact set $K \subset \mathbb{R}^d$ (this is true, but is a nontrivial exercise: **Ex**(23)). □

It is also classical that smooth functions satisfy the equation in the classical sense if and only if they are weak (or distributional) solutions. We can also check the following.

Proposition 4.3. *Suppose that ϱ is Lipschitz continuous in (t,x), that $\mathbf{v}$ is Lipschitz in x, and that the continuity equation $\partial_t\varrho + \nabla\cdot(\varrho\mathbf{v}) = 0$ is satisfied in the weak sense. Then the equation is also satisfied in the a.e. sense.*

Proof. First we note that with our assumptions both $\partial_t\varrho$ and $\nabla\cdot(\varrho\mathbf{v})$ are well defined a.e. Fix a countable set $D \subset C^1_c(\mathring{\Omega})$ which is dense for the uniform convergence in $C^0_c(\mathring{\Omega})$ (for instance, use polynomial functions with rational coefficients times suitable cut-off functions). Fix t_0 such that $(t,x) \mapsto \varrho(t,x)$ is differentiable at (t_0,x) for a.e. x, and also such that $t \mapsto \int_\Omega \phi\, \mathrm{d}\varrho_t$ is differentiable at $t = t_0$ with derivative given by $\int_\Omega \nabla\phi \cdot \mathbf{v}_t\, \mathrm{d}\varrho_t$ for all $\phi \in D$. Almost all t_0 satisfy these conditions. Then we can also write, by differentiating under the integral sign,

$$\frac{d}{dt}_{|t=t_0} \int_\Omega \phi\, \mathrm{d}\varrho_t\, \mathrm{d}x = \int_\Omega \phi\, (\partial_t\varrho)_{t_0}\, \mathrm{d}x,$$

which proves $\int_\Omega (\nabla\phi\cdot\mathbf{v}_{t_0})\varrho_{t_0}\, \mathrm{d}x = \int \phi\, (\partial_t\varrho)_{t_0}\, \mathrm{d}x$. Yet, we can write $\int_\Omega \nabla\phi\cdot\mathbf{v}_{t_0}\varrho_{t_0}\, \mathrm{d}x$ as $-\int_\Omega \phi\, \nabla\cdot(\mathbf{v}_{t_0}\varrho_{t_0})\, \mathrm{d}x$ (by integration by parts of Lipschitz functions, with no boundary term because of the compact support of ϕ), and finally we get

$$\int_\Omega \phi\, ((\partial_t\varrho)_{t_0} + \nabla\cdot(\varrho_{t_0}\mathbf{v}_{t_0}))\, \mathrm{d}x = 0 \quad \text{for all } \phi \in D,$$

which is enough to prove $(\partial_t\varrho)_{t_0} + \nabla\cdot(\varrho_{t_0}\mathbf{v}_{t_0}) = 0$ a.e. in x. □

After checking the relations between these different notions of solutions, from now on we will often say "solution" to mean, indifferently, "weak solution" or "solution in the distributional sense."

We now try to identify the solutions of the continuity equation and we want to connect them to the flow of the vector field $\mathbf{v}_t$. First, we recall some properties of the flow.

Box 4.1. Memo: *Flow of a time-dependent vector field*

We consider the ODE $y_x'(t) = \mathbf{v}_t(y_x(t))$ with initial datum $y_x(0) = x$, as in (4.1).

Proposition. If $\mathbf{v}_t$ is continuous in x for every t, then for every initial datum, there is local existence (there exists at least a solution, defined on a neighborhood of $t = 0$). If, moreover, $\mathbf{v}$ satisfies $|\mathbf{v}_t(x)| \le C_0 + C_1|x|$, then the solution is global in time. If $\mathbf{v}_t$ is Lipschitz in x and uniformly in t, then the solution is unique and defines a flow $Y_t(x) := y_x(t)$. In this case, if we set $L := \sup_t \mathrm{Lip}(\mathbf{v}_t)$, then the map $Y_t(\cdot)$ is Lipschitz in x, with constant $e^{L|t|}$. Moreover, $Y_t : \mathbb{R}^d \to \mathbb{R}^d$ is invertible and its inverse is also Lipschitz with the same constant.

Sketch of proof. We do not prove existence, which can be classically obtained by fixed-point arguments in the space of curves. As for uniqueness and for the dependence on the initial datum x, we consider two curves $y^{(1)}$ and $y^{(2)}$, solutions of $y'(t) = \mathbf{v}_t(y(t))$, and we define $E(t) = |y^{(1)}(t) - y^{(2)}(t)|^2$. We have

$$E'(t) = 2(y^{(1)}(t) - y^{(2)}(t)) \cdot (\mathbf{v}_t(y^{(1)}(t)) - \mathbf{v}_t(y^{(2)}(t))),$$

which proves $|E'(t)| \le 2LE(t)$. By Gronwall's lemma, this gives $E(0)e^{-2L|t|} \le E(t) \le E(0)e^{2L|t|}$ and provides at the same time uniqueness, injectivity, and the bi-Lipschitz behavior of Y_t (which is also a homeomorphism from $\mathbb{R}^d$ onto $\mathbb{R}^d$ since for every $x_0 \in \mathbb{R}^d$ we can solve the ODE imposing the Cauchy datum $y(t) = x_0$).

Then, we can prove the following that we state for simplicity in the case of Lipschitz and bounded vector fields.

Theorem 4.4. *Suppose that $\Omega \subset \mathbb{R}^d$ is either a bounded domain or $\mathbb{R}^d$ itself. Suppose that $\mathbf{v}_t : \Omega \to \mathbb{R}^d$ is Lipschitz continuous in x, uniformly in t, and uniformly bounded, and consider its flow Y_t. Suppose that for every $x \in \Omega$ and every $t \in [0, T]$, we have $Y_t(x) \in \Omega$ (which is obvious for $\Omega = \mathbb{R}^d$ and requires suitable Neumann conditions on $\mathbf{v}_t$ otherwise). Then, for every probability $\varrho_0 \in \mathscr{P}(\Omega)$, the measures $\varrho_t := (Y_t)_{\#}\varrho_0$ solve the continuity equation $\partial_t\varrho_t + \nabla \cdot (\varrho_t\mathbf{v}_t) = 0$ with initial datum ϱ_0. Moreover, every solution of the same equation with $\varrho_t \ll \mathscr{L}^d$ for every t is necessarily obtained as $\varrho_t = (Y_t)_{\#}\varrho_0$. In particular, the continuity equation admits a unique solution.*

Proof. First we check the validity of the equation when ϱ_t is obtained from such a flow through (4.1). We will prove that we have a weak solution. Fix a test function $\phi \in C^1(\mathbb{R}^d)$ such that both ϕ and $\nabla\phi$ are bounded, and compute

$$\frac{d}{dt}\int_\Omega \phi \,\mathrm{d}\varrho_t = \frac{d}{dt}\int \phi(y_x(t))\,\mathrm{d}\varrho_0(x) = \int_\Omega \nabla\phi(y_x(t))\cdot y_x'(t)\,\mathrm{d}\varrho_0(x)$$
$$= \int_\Omega \nabla\phi(y_x(t))\cdot \mathbf{v}_t(y_x(t))\,\mathrm{d}\varrho_0(x) = \int_\Omega \nabla\phi(y)\cdot\mathbf{v}_t(y)\,\mathrm{d}\varrho_t(y).$$

This characterizes weak solutions.

In order to prove the second part of the statement, we first observe that (4.2), i.e.,

$$\int_0^T\int_\Omega (\partial_t\phi + \mathbf{v}_t\cdot\nabla\phi)\,\mathrm{d}\varrho_t\,\mathrm{d}t = 0,$$

is also valid for all Lipschitz compactly supported test functions ϕ, whenever $\varrho_t \ll \mathscr{L}^d$ for every t. Indeed, if we fix a Lipschitz test function ϕ and we smooth it by convolution with a compactly supported kernel, we have a sequence $\phi_\varepsilon \in C_c^\infty$ such that $\nabla_{t,x}\phi_\varepsilon \to \nabla_{t,x}\phi$ a.e. and we can apply dominated convergence since $|\nabla_{t,x}\phi_\varepsilon| \le \mathrm{Lip}(\phi)$ (we need ϱ_t to be absolutely continuous because we only have Lebesgue-a.e. convergence).

Then we take a test function $\psi \in C_c^1(\mathbb{R}^d)$ and we define $\phi(t,x) := \psi((Y_t)^{-1}(x))$. If we can prove that $t \mapsto \int \phi(t,x)\,\mathrm{d}\varrho_t(x)$ is constant, then we have proven $\varrho_t = (Y_t)_\#\varrho_0$. The function ϕ is Lipschitz continuous, because the flow Y_t is bi-Lipschitz; yet, it is not compactly supported in time (it is compactly supported in space since $\mathrm{spt}\,\phi$ is compact, and if we set $M := \sup_{t,x}|\mathbf{v}_t(x)|$, we see that $\phi(t,x)$ vanishes on all points x which are at distance larger than tM from $\mathrm{spt}\,\psi$). Thus, we multiply it with a cut-off function $\chi(t)$, with $\chi \in C_c^1(]0,T[)$. We have

$$\partial_t(\chi\phi) + \mathbf{v}_t\cdot\nabla(\chi\phi) = \chi'(t)\phi(t,x) + \chi(t)\left(\partial_t\phi(t,x) + \mathbf{v}_t(x)\cdot\nabla\phi(t,x)\right).$$

We can prove that, by definition of ϕ, the term $\partial_t\phi(t,x) + \mathbf{v}_t(x)\cdot\nabla\phi(t,x)$ vanishes a.e. (this corresponds to saying that ϕ is a solution of the transport equation; see Box 6.3). This is true since we have $\phi(t,T_t(x)) = \psi(x)$, and differentiating it w.r.t. t (which is possible for a.e. (t,x)), we get $\partial_t(t,Y_t(x)) + \mathbf{v}_t(Y_t(x))\cdot\nabla\phi(t,Y_t(x)) = 0$, which means that $\partial_t\phi + \mathbf{v}_t\cdot\nabla\phi$ vanishes everywhere, as Y_t is surjective.

Hence, from the definition of distributional solution, we have

$$\int_0^T \chi'(t)\mathrm{d}t\int_{\mathbb{R}^d}\phi(t,x)\,\mathrm{d}\varrho_t(x) = \int_0^T \mathrm{d}t\int_{\mathbb{R}^d}\left(\partial_t(\chi\phi) + \mathbf{v}_t\cdot\nabla(\chi\phi)\right)\,\mathrm{d}\varrho_t(x) = 0.$$

The test function $\chi \in C_c^1(]0,T[)$ being arbitrary, we get that $t \mapsto \int_{\mathbb{R}^d}\phi(t,x)\,\mathrm{d}\varrho_t(x)$ is constant. □

4.2 Beckmann's problem

In this section we discuss a flow-minimization problem introduced by Beckmann in the 1950s [27] as a particular case of a wider class of convex optimization problem, of the form $\min\{\int H(\mathbf{w})\mathrm{d}x \ : \ \nabla\cdot\mathbf{w} = \mu-\nu\}$, for convex H. In this section we discuss the case $H(z) = |z|$, which is a limit case in the class of costs studied by Beckmann, who was indeed more interested in strictly convex costs. Strict convexity will be considered in "Discussion" (Section 4.4.1). The reason to stick to the linear cost here, and to spend more words on it than on the general case, lies in its equivalence with the Monge problem. We will provide all the tools to study this equivalence and this problem.

4.2.1 Introduction, formal equivalences, and variants

We consider here a simple version of a problem that has been proposed by Beckmann as a model for optimal transport in the 1950s. Beckmann called it *continuous transportation model* and he was not aware of Kantorovich's works and the possible links between the two theories.

Beckmann's minimal flow problem

Problem 4.5. Consider the minimization problem

$$\text{(BP)} \quad \min\left\{ \int |\mathbf{w}(x)|\,\mathrm{d}x \ : \ \mathbf{w} : \Omega \to \mathbb{R}^d,\ \nabla\cdot\mathbf{w} = \mu - \nu \right\}, \tag{4.4}$$

where the divergence condition is to be read in the weak sense, with no-flux boundary conditions, i.e., $-\int \nabla\phi\cdot\,\mathrm{d}\mathbf{w} = \int \phi\,\mathrm{d}(\mu-\nu)$ for any $\phi \in C^1(\overline{\Omega})$.

We will see now that an equivalence between (BP) and (KP) (for the cost $c(x,y) = |x-y|$) holds true. To do that, we can look at the following considerations and formal computations.

We take the problem (BP) and rewrite the constraint on $\mathbf{w}$ using

$$\sup_{\phi} \int_\Omega \nabla\phi\cdot\mathbf{w}\,\mathrm{d}x + \int_\Omega \phi\,\mathrm{d}(\mu-\nu) = \begin{cases} 0 & \text{if } \nabla\cdot\mathbf{w} = \mu-\nu, \\ +\infty & \text{otherwise.} \end{cases}$$

Hence one can write (BP) as

$$\inf_{\mathbf{w}} \left(\int_\Omega |\mathbf{w}| \, \mathrm{d}x + \sup_\phi \int_\Omega \nabla\phi \cdot \mathbf{w} \, \mathrm{d}x + \int_\Omega \phi \, \mathrm{d}(\mu - \nu) \right)$$
$$= \sup_\phi \left(\int_\Omega \phi \, \mathrm{d}(\mu - \nu) + \inf_{\mathbf{w}} \int_\Omega (|\mathbf{w}| - \nabla\phi \cdot \mathbf{w}) \, \mathrm{d}x \right), \qquad (4.5)$$

where inf and sup have been exchanged formally as in the computations of Chapter 1. Then, one notes that

$$\inf_{\mathbf{w}} \int_\Omega (|\mathbf{w}| - \nabla\phi \cdot \mathbf{w}) \, \mathrm{d}x = \begin{cases} 0 & \text{if } |\nabla\phi| \le 1 \\ -\infty & \text{otherwise} \end{cases}$$

and this leads to the dual formulation for (BP) which gives

$$\sup \left\{ \int_\Omega \phi \, \mathrm{d}(\mu - \nu) \quad : \quad |\nabla\phi| \le 1 \right\}.$$

Since this problem is exactly the same as (DP) (a consequence of the fact that Lip_1 functions are exactly those functions whose gradient is smaller than 1), this provides a formal equivalence between (BP) and (KP). We say that it is only formal because we did not prove the equality in (4.5). Note that we also need to suppose that Ω is convex; otherwise functions with gradient smaller than 1 are only Lip_1 according to the geodesic distance in Ω.

Most of the considerations above, especially those on the problem (BP), are specific to the cost equal to the distance $|x - y|$. In general, for costs of the form $h(x - y)$, h needs to be 1-homogeneous, if one wants some similar result to hold. We refer to [195, 196] for a general way to transform convex costs into convex 1-homogeneous ones (by adding one variable, corresponding to time)[1]. An interesting generalization which keeps the same formalism of the case of the Euclidean distance concerns a cost c which comes from a Riemannian distance $k(x)$.

Consider indeed

$$\min \left\{ \int k(x)|\mathbf{w}(x)| \, \mathrm{d}x \ : \ \nabla \cdot \mathbf{w} = \mu - \nu \right\}$$

[1] In this way the flow-minimization problem corresponding to costs of the form $|x - y|^p$ is transformed into the so-called Benamou-Brenier problem, which we will discuss in Chapters 5 and 6, but we do not push this analogy to further conclusions.

which corresponds, by duality with the functions u such that $|\nabla u| \leq k$, to

$$\min \left\{ \int d_k(x,y)\, \mathrm{d}\gamma(x,y) \ : \ \gamma \in \Pi(\mu,\nu) \right\},$$

where $d_k(x,y) = \inf_{\omega(0)=x,\, \omega(1)=y} L_k(\omega) := \int_0^1 k(\omega(t))|\omega'(t)|\mathrm{d}t$ is the distance associated with the Riemannian metric k.

The generalization above is suitable when we want to model a nonuniform cost for the movement (due to geographical obstacles or configurations). Such a model is not always satisfying, for instance, in urban transport, where we want to consider the fact that the metric k is usually not a priori known, but it depends on the traffic distribution itself. We will develop this aspect in the "Discussion" Section 4.4.1, together with a completely different problem which is somehow "opposite" (Section 4.4.2): instead of looking at transport problems where concentration of the mass is penalized (because of traffic congestion), we look at problems where it is indeed encouraged, because of the so-called economy of scale (i.e., the larger the mass you transport, the cheaper the individual cost).

4.2.2 Producing a minimizer for the Beckmann Problem

The first remark on problem (BP) is that a priori it is not well-posed, in the sense that there could not exist an L^1 vector field minimizing the L^1 norm under divergence constraints. This is easy to understand if we think at the direct method in calculus of variations to prove existence: we take a minimizing sequence $\mathbf{w}_n$ and we would like to extract a converging subsequence. Could we do this, from $\mathbf{w}_n \rightharpoonup \mathbf{w}$, it would be easy to prove that $\mathbf{w}$ still satisfies $\nabla \cdot \mathbf{w} = \mu - \nu$, since the relation

$$-\int \nabla\phi \cdot \mathbf{w}_n \, \mathrm{d}x = \int \phi \, \mathrm{d}(\mu - \nu)$$

would pass to the limit as $n \to \infty$. Yet, the information $\int |\mathbf{w}(x)| \, \mathrm{d}x \leq C$ is not enough to extract a converging sequence, even weakly. Indeed, the space L^1 being nonreflexive, bounded sequences are not guaranteed to have weakly converging subsequences. This is on the contrary the case for dual spaces (and for reflexive spaces, which are roughly speaking the dual of their dual).

Note that the strictly convex version that was proposed by Beckmann and that we will review in Section 4.4.1 is much better to handle: if, for instance, we minimize $\int |\mathbf{w}|^2 \, \mathrm{d}x$, then we can use weak compactness in L^2, which is much easier to have than compactness in L^1.

To avoid this difficulty, we will choose the natural setting for (BP), i.e., the framework of vector measures.

Box 4.2. Memo: *Vector measures*

Definition. A finite vector measure λ on a set Ω is a map associating with every Borel subset $A \subset \Omega$ a value $\lambda(A) \in \mathbb{R}^d$ such that for every disjoint union $A = \bigcup_i A_i$ (with $A_i \cap A_j = \emptyset$ for $i \neq j$), we have

$$\sum_i |\lambda(A_i)| < +\infty \quad \text{and} \quad \lambda(A) = \sum_i \lambda(A_i).$$

We denote by $\mathscr{M}^d(\Omega)$ the set of finite vector measures on Ω. To such measures we can associate a positive scalar measure $|\lambda| \in \mathscr{M}_+(\Omega)$ through

$$|\lambda|(A) := \sup \left\{ \sum_i |\lambda(A_i)| \ : \ A = \bigcup_i A_i \ \text{ with } A_i \cap A_j = \emptyset \text{ for } i \neq j \right\}.$$

This scalar measure is called *total variation measure* of λ. Note that for simplicity we only consider the Euclidean norm on $\mathbb{R}^d$, and write $|\lambda|$ instead of $||\lambda||$ (a notation that we keep for the total mass of the total variation measure; see below), but the same could be defined for other norms as well.

The integral of a Borel function $\xi : \Omega \to \mathbb{R}^d$ w.r.t. λ is well defined if $|\xi| \in L^1(\Omega, |\lambda|)$ is denoted $\int \xi \cdot \mathrm{d}\lambda$ and can be computed as $\sum_{i=1}^d \int \xi_i \, \mathrm{d}\lambda_i$, thus reducing to integrals of scalar functions according to scalar measures. It could also be defined as a limit of integral of piecewise constant functions.

Functional analysis facts. The quantity $||\lambda|| := |\lambda|(\Omega)$ is a norm on $\mathscr{M}^d(\Omega)$, and this normed space is the dual of $C_0(\Omega; \mathbb{R}^d)$, the space of continuous function on Ω vanishing at infinity, through the duality $(\xi, \lambda) \mapsto \int \xi \cdot \mathrm{d}\lambda$. This gives a notion of $\overset{*}{\rightharpoonup}$ convergence for which bounded sets in $\mathscr{M}^d(\Omega)$ are compact. As for scalar measures, we denote by $\rightharpoonup$ the weak convergence in duality with C_b functions.

A clarifying fact is the following.

Proposition. For every $\lambda \in \mathscr{M}^d(\Omega)$ there exists a Borel function $u : \Omega \to \mathbb{R}^d$ such that $\lambda = u \cdot |\lambda|$ and $|u| = 1$ a.e. (for the measure $|\lambda|$). In particular, $\int \xi \cdot \mathrm{d}\lambda = \int (\xi \cdot u) \, \mathrm{d}|\lambda|$.

Sketch of proof. The existence of a function u is a consequence, via Radon-Nikodym theorem, of $\lambda \ll |\lambda|$ (every A set such that $|\lambda|(A) = 0$ obviously satisfies $\lambda(A) = 0$). The condition $|u| = 1$ may be proven by considering the sets $\{|u| < 1 - \varepsilon\}$ and $\{u \cdot e > a\}$ for all hyperplanes such that the unit ball B_1 is contained in $\{x \in \mathbb{R}^d \ : \ x \cdot e \leq a\}$ (and, actually, we have $B_1 = \bigcap_{e,a} \{x \in \mathbb{R}^d \ : \ x \cdot e \leq a\}$, the intersection being reduced to a countable intersection). These sets must be negligible; otherwise we have a contradiction on the definition of $|\lambda|$.

With these definitions in mind, we can prove the following theorem. We denote by $\mathscr{M}^d_{\mathrm{div}}$ the space of vector measures with divergence which is a scalar measure.

Theorem 4.6. *Suppose that Ω is a compact convex domain in $\mathbb{R}^d$. Then, the problem*

$$\text{(BP)} \quad \min \left\{ |\mathbf{w}|(\Omega) \ : \ \mathbf{w} \in \mathscr{M}^d_{\mathrm{div}}(\Omega) \, , \ \nabla \cdot \mathbf{w} = \mu - \nu \right\}$$

(with divergence imposed in the weak sense, i.e., for every $\phi \in C^1(\overline{\Omega})$; we impose $-\int \nabla\phi \cdot \mathrm{d}\mathbf{w} = \int \phi \,\mathrm{d}(\mu - \nu)$, which also includes no-flux boundary conditions) admits a solution. Moreover, its minimal value equals the minimal value of (KP) and a solution of (BP) can be built from a solution of (KP). The two problems are hence equivalent.

Proof. The first point that we want to prove is the equality of the minimal values of (BP) and (KP). We start from min(BP)$\geq$ min(KP). In order to do so, take an arbitrary function $\phi \in C^1$ with $|\nabla\phi| \leq 1$. Consider that for any $\mathbf{w}$ with $\nabla\cdot\mathbf{w} = \mu - \nu$, we have

$$|\mathbf{w}|(\Omega) = \int_\Omega 1 \,\mathrm{d}|\mathbf{w}| \geq \int_\Omega (-\nabla\phi)\cdot \mathrm{d}\mathbf{w} = \int_\Omega \phi \,\mathrm{d}(\mu - \nu).$$

If one takes a sequence of $\mathrm{Lip}_1 \cap C^1$ functions uniformly converging to the Kantorovich potential u such that $\int_\Omega u \,\mathrm{d}(\mu - \nu) = \max(\mathrm{DP}) = \min(\mathrm{KP})$ (for instance, take convolutions $\phi_k = \eta_k * u$), then we get

$$\int_\Omega \mathrm{d}|\mathbf{w}| \geq \min(\mathrm{KP})$$

for any admissible $\mathbf{w}$, i.e., min(BP)$\geq$ min(KP).

We will show at the same time the reverse inequality and how to construct an optimal $\mathbf{w}$ from an optimal γ for (KP).

Actually, one way to produce a solution to this divergence-constrained problem is the following (see [70]): take an optimal transport plan γ and build a vector measure $\mathbf{w}_{[\gamma]}$ defined[2] through

$$\langle \mathbf{w}_{[\gamma]}, \xi\rangle := \int_{\Omega\times\Omega} \int_0^1 \omega'_{x,y}(t)\cdot\xi(\omega_{x,y}(t))\mathrm{d}t \;\mathrm{d}\gamma(x,y),$$

for every $\xi \in C^0(\Omega;\mathbb{R}^d)$, $\omega_{x,y}$ being a parametrization of the segment $[x,y]$ (it is clear that this is the point where convexity of Ω is needed). Even if for this proof it would not be important, we will fix the constant-speed parametrization, i.e., $\omega_{x,y}(t) = (1-t)x + ty$.

It is not difficult to check that this measure satisfies the divergence constraint, since if one takes $\xi = \nabla\phi$, then

$$\int_0^1 \omega'_{x,y}(t)\cdot\xi(\omega_{x,y}(t))\,\mathrm{d}t = \int_0^1 \frac{d}{dt}\big(\phi(\omega_{x,y}(t))\big)\,\mathrm{d}t = \phi(y) - \phi(x)$$

and hence $\langle \mathbf{w}_{[\gamma]}, \nabla\phi\rangle = \int \phi\, d(\nu - \mu)$ and $\mathbf{w}_{[\gamma]} \in \mathscr{M}^d_{\mathrm{div}}(\Omega)$.

[2]The strange notation $\mathbf{w}_{[\gamma]}$ is chosen so as to distinguish from the object $\mathbf{w}_Q$ that we will introduce in Section 4.2.3.

To estimate its mass, we can see that $|\mathbf{w}_{[\gamma]}| \le \sigma_\gamma$, where the scalar measure σ_γ is defined through

$$\langle \sigma_\gamma, \phi \rangle := \int_{\Omega\times\Omega} \int_0^1 |\omega'_{x,y}(t)| \phi(\omega_{x,y}(t)) \mathrm{d}t \; \mathrm{d}\gamma, \qquad \text{for all } \phi \in C^0(\Omega; \mathbb{R})$$

and it is called *transport density*. Actually, we can even say more, since we can use the Kantorovich potential u (see Chapter 3), and write[3]

$$\omega'_{x,y}(t) = -|x-y| \frac{x-y}{|x-y|} = -|x-y| \nabla u(\omega_{x,y}(t)).$$

This is valid for every $t \in]0,1[$ and every $x, y \in \mathrm{spt}(\gamma)$ (so that $\omega_{x,y}(t)$ is in the interior of the transport ray $[x, y]$, if $x \neq y$; anyway for $x = y$, both expressions vanish).

This allows to write, for every $\xi \in C^0(\Omega; \mathbb{R}^d)$,

$$\begin{aligned} \langle \mathbf{w}_{[\gamma]}, \xi \rangle &= \int_{\Omega\times\Omega} \int_0^1 -|x-y| \nabla u(\omega_{x,y}(t)) \cdot \xi(\omega_{x,y}(t)) \mathrm{d}t \; \mathrm{d}\gamma(x,y) \\ &= -\int_0^1 \mathrm{d}t \int_{\Omega\times\Omega} \nabla u(\omega_{x,y}(t)) \cdot \xi(\omega_{x,y}(t)) |x-y| \, \mathrm{d}\gamma(x,y). \end{aligned}$$

If we introduce the function $\pi_t : \Omega \times \Omega \to \Omega$ given by $\pi_t(x, y) = \omega_{x,y}(t) = (1-t)x + ty$, we get

$$\langle \mathbf{w}_{[\gamma]}, \xi \rangle = -\int_0^1 \mathrm{d}t \int_\Omega \nabla u(z) \cdot \xi(z) \, \mathrm{d}\big((\pi_t)_\#(c \cdot \gamma)\big),$$

where $c \cdot \gamma$ is the measure on $\Omega \times \Omega$ with density $c(x, y) = |x - y|$ w.r.t. γ.

Since, on the other hand, the same kind of computations gives

$$\langle \sigma_\gamma, \phi \rangle = \int_0^1 \mathrm{d}t \int_\Omega \phi(z) \, \mathrm{d}\big((\pi_t)_\#(c \cdot \gamma)\big), \tag{4.6}$$

we get $\langle \mathbf{w}_{[\gamma]}, \xi \rangle = \langle \sigma_\gamma, -\xi \cdot \nabla u \rangle$, which shows

$$\mathbf{w}_{[\gamma]} = -\nabla u \cdot \sigma_\gamma.$$

This gives the density of $\mathbf{w}_{[\gamma]}$ with respect to σ_γ and confirms $|\mathbf{w}_{[\gamma]}| \le \sigma_\gamma$.

[3] Pay attention to the use of the gradient of the Kantorovich potential u: we are using the result of Lemma 3.6 which provides differentiability of u on transport rays.

The mass of σ_γ is obviously

$$\int_\Omega \mathrm{d}\sigma_\gamma = \int_\Omega \int_0^1 |\omega'_{x,y}(t)|\mathrm{d}t \; \mathrm{d}\gamma(x,y) = \int_{\Omega\times\Omega} |x-y| \, \mathrm{d}\gamma(x,y) = \min\,(\mathrm{KP}),$$

which proves the optimality of $\mathbf{w}_{[\gamma]}$ since no other $\mathbf{w}$ may do better than this, thanks to the first part of the proof. This also proves $\min\,(\mathrm{BP}) = \min\,(\mathrm{KP})$. □

Monge-Kantorovich system and transport density

The scalar measure σ_γ that we have just defined is called *transport density*. It has been introduced in [159] and [70, 71] for different goals. In [70, 71] it is connected to some shape-optimization problems that we will not develop in this book. On the other hand, in [159] it has been used to provide one of the first solutions to the Monge problem with cost $|x-y|$. The important fact is that the measure $\sigma = \sigma_\gamma$ solves, together with the Kantorovich potential u, the so-called Monge-Kantorovich system

$$\begin{cases} \nabla\cdot(\sigma\nabla u) = \mu - \nu & \text{in } \Omega \\ |\nabla u| \le 1 & \text{in } \Omega, \\ |\nabla u| = 1 & \sigma - \text{a.e.} \end{cases} \tag{4.7}$$

This system is a priori only solved in a formal way, since the product of the L^∞ function ∇u with the measure σ has no meaning if $\sigma \notin L^1$. To overcome this difficulty there are two possibilities: either we pass through the theory of σ-tangential gradient (see, for instance, [145] or [71]), or we give conditions to have $\sigma_\gamma \in L^1$. This second choice is what we do in Section 4.3, where we also show better L^p estimates.

As we said, the solution of (4.7) has been used in [159] to find a transport map T, optimal for the cost $|x-y|$. This map was defined as the flow of a certain vector field depending on σ and ∇u, exactly in the spirit of the Dacorogna-Moser construction that we will see in Section 4.2.3. However, because of the lack of regularity of σ and ∇u, the Beckmann problem $\min\left\{\int |\mathbf{w}| \,:\, \nabla\cdot\mathbf{w} = \mu-\nu\right\}$, whose solution is $\mathbf{w} = \sigma_\gamma\nabla u$, was approximated through $\min\left\{\int |\mathbf{w}|^p \,:\, \nabla\cdot\mathbf{w} = \mu-\nu\right\}$, for $p > 1$. We will see in Section 4.4.1 that this problem, connected to traffic congestion issues, is solved by a vector field obtained from the solution of a $\Delta_{p'}$ equation and admits some regularity results. In this way, the dual of the Monge problem is approximated by p'-Laplacian equations, for $p' \to \infty$.

4.2.3 Traffic intensity and traffic flows for measures on curves

We introduce in this section some objects that generalize both $\mathbf{w}_{[\gamma]}$ and σ_γ and that will be useful for many goals. They will be used both for proving the characterization of the optimal $\mathbf{w}$ as coming from an optimal plan γ and for the modeling issues of the discussion section. As a by-product, we also present a new proof of a decomposition result by Smirnov [287] (see also [278]).

Let us introduce some notations. We define the set $\mathrm{AC}(\Omega)$ as the set of absolutely continuous curves $\omega : [0,1] \mapsto \Omega$. We suppose Ω to be compact, with nonempty interior, in the whole section. Given $\omega \in \mathrm{AC}(\Omega)$ and a continuous function ϕ, let us set

$$L_\phi(\omega) := \int_0^1 \phi(\omega(t))|\omega'(t)|\mathrm{d}t. \tag{4.8}$$

This quantity is the length of the curve, weighted with the weight ϕ. When we take $\phi = 1$, we get the usual length of ω and we denote it by $L(\omega)$ instead of $L_1(\omega)$. Note that these quantities are well defined since $\omega \in \mathrm{AC}(\Omega)$ implies that ω is differentiable a.e. and $\omega' \in L^1([0,1])$ (we recall the definition of AC curves: they are curves ω with their distributional derivative $\omega' \in L^1$ and $\omega(t_1) - \omega(t_0) = \int_{t_0}^{t_1} \omega'(t)\mathrm{d}t$ for every $t_0 < t_1$). For simplicity, from now on we will write $\mathscr{C}$ (the space of "curves") for $\mathrm{AC}(\Omega)$, when there is no ambiguity on the domain.

We consider probability measures Q on the space $\mathscr{C}$. We restrict ourselves to measures Q such that $\int L(\omega)\,\mathrm{d}Q(\omega) < +\infty$: these measures will be called *traffic plans*, according to a terminology introduced in [48]. We endow the space $\mathscr{C}$ with the uniform convergence. Note that the Ascoli-Arzelà Theorem guarantees that the sets $\{\omega \in \mathscr{C} : \mathrm{Lip}(\omega) \le \ell\}$ are compact (for the uniform convergence) for every ℓ. We will associate two measures on Ω to such a Q. The first is a scalar one, called *traffic intensity* and denoted by $i_Q \in \mathscr{M}_+(\Omega)$; it is defined by

$$\int_\Omega \phi\,\mathrm{d}i_Q := \int_{\mathscr{C}} \Big(\int_0^1 \phi(\omega(t))|\omega'(t)|\mathrm{d}t \Big)\,\mathrm{d}Q(\omega) = \int_{\mathscr{C}} L_\phi(\omega)\mathrm{d}Q(\omega),$$

for all $\phi \in C(\Omega, \mathbb{R}_+)$. This definition (taken from [112]) is a generalization of the notion of transport density. The interpretation is the following: for a subregion A, $i_Q(A)$ represents the total cumulated traffic in A induced by Q, i.e., for every path we compute "how long" it stays in A, and then we average on paths.

We also associate a vector measure $\mathbf{w}_Q$ with any traffic plan $Q \in \mathscr{P}(\mathscr{C})$ via

$$\forall \xi \in C(\Omega; \mathbb{R}^d) \quad \int_\Omega \xi \cdot \mathrm{d}\mathbf{w}_Q := \int_{\mathscr{C}} \left(\int_0^1 \xi(\omega(t)) \cdot \omega'(t)\mathrm{d}t \right) \mathrm{d}Q(\omega).$$

We will call $\mathbf{w}_Q$ *traffic flow* induced by Q. Taking a gradient field $\xi = \nabla\phi$ in the previous definition yields

$$\int_\Omega \nabla\phi\cdot \mathrm{d}\mathbf{w}_Q = \int_{\mathscr{C}} [\phi(\omega(1))-\phi(\omega(0))]\mathrm{d}Q(\omega) = \int_\Omega \phi\,\mathrm{d}((e_1)_\#Q-(e_0)_\#Q)$$

(we recall that e_t denotes the evaluation map at time t, i.e., $e_t(\omega) := \omega(t)$).

From now on, we will restrict our attention to *admissible* traffic plans Q, i.e., traffic plans such that $(e_0)_\#Q = \mu$ and $(e_1)_\#Q = \nu$, where μ and ν are two prescribed probability measures on Ω. This means that

$$\nabla\cdot\mathbf{w}_Q = \mu-\nu$$

and hence $\mathbf{w}_Q$ is an admissible flow connecting μ and ν. Note that the divergence is always considered in a weak (distributional) sense and automatically endowed with no-flux conditions, i.e., when we say $\nabla\cdot\mathbf{w}=f$, we mean $\int\nabla\phi\cdot\mathrm{d}\mathbf{w} = -\int\phi\,\mathrm{d}f$ for all $\phi\in C^1(\Omega)$, without any condition on the boundary behavior of the test function ϕ.

Coming back to $\mathbf{w}_Q$, it is easy to check that $|\mathbf{w}_Q|\le i_Q$, where $|\mathbf{w}_Q|$ is the total variation measure of the vector measure $\mathbf{w}_Q$. This last inequality is in general not an equality, since the curves of Q could produce some cancellations (imagine a non-negligible amount of curves passing through the same point with opposite directions, so that $\mathbf{w}_Q=0$ and $i_Q>0$).

We need some properties of the traffic intensity and traffic flow.

Proposition 4.7. *Both $\mathbf{w}_Q$ and i_Q are invariant under reparametrization (i.e., if $\mathrm{T}:\mathscr{C}\to\mathscr{C}$ is a map such that for every ω, the curve $\mathrm{T}(\omega)$ is just a reparametrization in time of ω, then $\mathbf{w}_{\mathrm{T}_\#Q}=\mathbf{w}_Q$ and $i_{\mathrm{T}_\#Q}=i_Q$).*

For every Q, the total mass $i_Q(\Omega)$ equals the average length of the curves according to Q, i.e., $\int_{\mathscr{C}} L(\omega)\,\mathrm{d}Q(\omega) = i_Q(\Omega)$. In particular, $\mathbf{w}_Q$ and i_Q are finite measures thanks to the definition of traffic plan.

If $Q_n\rightharpoonup Q$ and $i_{Q_n}\rightharpoonup i$, then $i\ge i_Q$.

If $Q_n\rightharpoonup Q$, $\mathbf{w}_{Q_n}\rightharpoonup\mathbf{w}$ and $i_{Q_n}\rightharpoonup i$, then $||\mathbf{w}-\mathbf{w}_Q||\le i(\Omega)-i_Q(\Omega)$. In particular, if $Q_n\rightharpoonup Q$ and $i_{Q_n}\rightharpoonup i_Q$, then $\mathbf{w}_{Q_n}\rightharpoonup\mathbf{w}_Q$.

Proof. The invariance by reparametrization comes from the invariance of both $L_\phi(\omega)$ and $\int_0^1\xi(\omega(t))\cdot\omega'(t)\mathrm{d}t$.

The formula $\int_{\mathscr{C}} L(\omega)\,\mathrm{d}Q(\omega) = i_Q(\Omega)$ is obtained from the definition of i_Q by testing against the function 1.

To check the inequality $i\ge i_Q$, fix a positive test function $\phi\in C(\Omega)$ and suppose $\phi\ge\varepsilon_0>0$. Write

$$\int_\Omega \phi\,\mathrm{d}i_{Q_n} = \int_{\mathscr{C}}\left(\int_0^1\phi(\omega(t))|\omega'(t)|\mathrm{d}t\right)\mathrm{d}Q_n(\omega). \tag{4.9}$$

Note that the function $\mathscr{C}\ni\omega\mapsto L_\phi(\omega)=\int_0^1\phi(\omega(t))|\omega'(t)|\mathrm{d}t$ is positive and lower semi-continuous w.r.t. ω. Indeed, if we take a sequence $\omega_n\to\omega$, from the bound $\phi\ge\varepsilon_0>0$, we can assume that $\int|\omega_n'(t)|\mathrm{d}t$ is bounded. Then we can

infer $\omega_n' \rightharpoonup \omega'$ weakly (as measures, or in L^1), which implies, up to subsequences, the existence of a measure $f \in \mathscr{M}_+([0,1])$ such that $f \geq |\omega'|$ and $|\omega_n'| \rightharpoonup f$. Moreover, $\phi(\omega_n(t)) \to \phi(\omega(t))$ uniformly, which gives $\int \phi(\omega_n(t))|\omega_n'(t)|\mathrm{d}t \to \int \phi(\omega(t))\mathrm{d}f(t) \geq \int \phi(\omega(t))|\omega'(t)|\mathrm{d}t$.

This allows to pass to the limit in (4.9), thus obtaining

$$\int \phi \,\mathrm{d}i = \lim_n \int \phi \,\mathrm{d}i_{Q_n} = \liminf_n \int_{\mathscr{C}} L_\phi(\omega)\mathrm{d}Q_n(\omega) \geq \int_{\mathscr{C}} L_\phi(\omega)\mathrm{d}Q(\omega) = \int \phi \,\mathrm{d}i_Q.$$

If we take an arbitrary test function ϕ (without a lower bound), add a constant ε_0 and apply the previous reasoning, we get $\int(\phi+\varepsilon_0)\mathrm{d}i \geq \int(\phi+\varepsilon_0)\mathrm{d}i_Q$. Letting $\varepsilon_0 \to 0$, as i is a finite measure and ϕ is arbitrary, we get $i \geq i_Q$.

To check the last property, fix a smooth vector test function ξ and a number $\lambda > 1$ and look at

$$\begin{aligned}\int_\Omega \xi \cdot \mathrm{d}\mathbf{w}_{Q_n} &= \int_{\mathscr{C}} \left(\int_0^1 \xi(\omega(t)) \cdot \omega'(t)\mathrm{d}t \right) \mathrm{d}Q_n(\omega) \\ &= \int_{\mathscr{C}} \left(\int_0^1 \xi(\omega(t)) \cdot \omega'(t)\mathrm{d}t + \lambda||\xi||_\infty L(\omega) \right) \mathrm{d}Q_n(\omega) - \lambda||\xi||_\infty i_{Q_n}(\Omega), \quad (4.10)\end{aligned}$$

where we just added and subtracted the total mass of i_{Q_n}, equal to the average of $L(\omega)$ according to Q_n. Now note that

$$\mathscr{C} \ni \omega \mapsto \int_0^1 \xi(\omega(t)) \cdot \omega'(t)\mathrm{d}t + \lambda||\xi||_\infty L(\omega) \geq (\lambda-1)||\xi||_\infty L(\omega)$$

is l.s.c. in ω (use the same argument as above, noting that if we take $\omega_n \to \omega$, we may assume $L(\omega_n)$ to be bounded and obtain $\omega_n' \rightharpoonup \omega'$). This means that if we pass to the limit in (4.10), we get

$$\begin{aligned}\int_\Omega \xi \cdot \mathrm{d}\mathbf{w} &= \lim_n \int_\Omega \xi \cdot \mathrm{d}\mathbf{w}_{Q_n} \\ &\geq \int_{\mathscr{C}} \left(\int_0^1 \xi(\omega(t)) \cdot \omega'(t)\mathrm{d}t + \lambda||\xi||_\infty L(\omega) \right) \mathrm{d}Q(\omega) - \lambda||\xi||_\infty i(\Omega) \\ &= \int_\Omega \xi \cdot \mathrm{d}\mathbf{w}_Q + \lambda||\xi||_\infty (i_Q(\Omega) - i(\Omega)).\end{aligned}$$

By replacing ξ with $-\xi$, we get

$$\left| \int_\Omega \xi \cdot \mathrm{d}\mathbf{w} - \int_\Omega \xi \cdot \mathrm{d}\mathbf{w}_Q \right| \leq \lambda||\xi||_\infty (i(\Omega) - i_Q(\Omega)).$$

Letting $\lambda \to 1$ and taking the sup over ξ with $||\xi||_\infty \leq 1$, we get the desired estimate $||\mathbf{w} - \mathbf{w}_Q|| \leq i(\Omega) - i_Q(\Omega)$.

The very last property is evident: indeed one can assume up to a subsequence that $\mathbf{w}_{Q_n} \rightharpoonup \mathbf{w}$ holds for a certain $\mathbf{w}$, and $i = i_Q$ implies $\mathbf{w} = \mathbf{w}_Q$ (which also implies the full convergence of the sequence). □

Box 4.3. Good to know!: *Dacorogna-Moser transport*

Let us present here a particular case of the ideas from [132] (first used in optimal transport in [159]).

Construction. Suppose that $\mathbf{w} : \Omega \to \mathbb{R}^d$ is a Lipschitz vector field with $\mathbf{w} \cdot \mathbf{n} = 0$ on $\partial\Omega$ and $\nabla \cdot \mathbf{w} = f_0 - f_1$, where f_0, f_1 are positive probability densities which are Lipschitz continuous and bounded from below. Then we can define the nonautonomous vector field $\mathbf{v}_t(x)$ via

$$\mathbf{v}_t(x) = \frac{\mathbf{w}(x)}{f_t(x)} \quad \text{where } f_t = (1-t)f_0 + tf_1$$

and consider the Cauchy problem

$$\begin{cases} y_x'(t) = \mathbf{v}_t(y_x(t)) \\ y_x(0) = x \end{cases}.$$

We define a map $Y : \Omega \to \mathscr{C}$ by associating with every x the curve $Y(x)$ given by $y_x(\cdot)$. Then, we look for the measure $Q = Y_\# f_0$ and $\varrho_t := (e_t)_\# Q := (Y_t)_\# f_0$. Thanks to the consideration in Section 4.1.2, ϱ_t solves the continuity equation $\partial_t \varrho_t + \nabla \cdot (\varrho_t \mathbf{v}_t) = 0$. Yet, it is easy to check that f_t also solves the same equation since $\partial_t f_t = f_1 - f_0$ and $\nabla \cdot (f_t \mathbf{v}_t) = \nabla \cdot \mathbf{w} = f_0 - f_1$. By the uniqueness result of Section 4.1.2, from $\varrho_0 = f_0$, we infer $\varrho_t = f_t$. In particular, $x \mapsto y_x(1)$ is a transport map from f_0 to f_1.

It is interesting to compute the traffic intensity and the traffic flow associated with the measure Q in Dacorogna-Moser construction. Fix a scalar test function ϕ:

$$\begin{aligned} \int_\Omega \phi \, \mathrm{d}i_Q &= \int_\Omega \int_0^1 \phi(y_x(t)) |\mathbf{v}_t(y_x(t))| \mathrm{d}t f_0(x) \, \mathrm{d}x \\ &= \int_0^1 \int_\Omega \varphi(y) |\mathbf{v}_t(y)| f_t(y) \mathrm{d}y \mathrm{d}t = \int_\Omega \phi(y) |\mathbf{w}(y)| \mathrm{d}y \end{aligned}$$

so that $i_Q = |\mathbf{w}|$. Analogously, fix a vector test function ξ

$$\begin{aligned} \int_\Omega \xi \cdot d\mathbf{w}_Q &= \int_\Omega \int_0^1 \xi(y_x(t)) \cdot \mathbf{v}_t(y_x(t)) \mathrm{d}t f_0(x) \, \mathrm{d}x \\ &= \int_0^1 \int_\Omega \xi(y) \cdot \mathbf{v}_t(y) f_t(y) \mathrm{d}y \mathrm{d}t = \int_\Omega \xi(y) \cdot \mathbf{w}(y) \mathrm{d}y, \end{aligned}$$

which shows $\mathbf{w}_Q = \mathbf{w}$. Note that in this case we have $|\mathbf{w}_Q| = i_Q$ and this is due to the fact that no cancellation is possible, since all the curves share the same direction at every given point.

With these tools we want to provide a proof of a decomposition result by Smirnov [287], for which we need an extra approximation result.

Lemma 4.8. *Consider two probabilities $\mu, \nu \in \mathscr{P}(\Omega)$ on a smooth compact domain Ω and a vector measure $\mathbf{w} \in \mathscr{M}^d_{\mathrm{div}}$ satisfying $\nabla \cdot \mathbf{w} = \mu - \nu$ in distributional sense (with no-flux boundary conditions). Then, for every smooth compact domain Ω' containing Ω in its interior, there exist a family of vector fields $\mathbf{w}^\varepsilon \in C^\infty(\Omega')$ with $\mathbf{w}^\varepsilon \cdot \mathbf{n}_{\Omega'} = 0$ and two families of densities $\mu^\varepsilon, \nu^\varepsilon \in C^\infty(\Omega')$, bounded from below by positive constants $k_\varepsilon > 0$, with $\nabla \cdot \mathbf{w}^\varepsilon = \mu^\varepsilon - \nu^\varepsilon$ and $\int_{\Omega'} \mu^\varepsilon = \int_{\Omega'} \nu^\varepsilon = 1$, weakly converging to $\mathbf{w}$, μ, and ν as measures, respectively, and satisfying $|\mathbf{w}^\varepsilon| \rightharpoonup |\mathbf{w}|$.*

Proof. First, extend μ, ν, and $\mathbf{w}$ at 0 out of Ω and take convolutions (in the whole space $\mathbb{R}^d$) with a Gaussian kernel η_ε, so that we get $\hat{\mathbf{w}}^\varepsilon := \mathbf{w} * \eta_\varepsilon$ and $\hat{\mu}^\varepsilon := \mu * \eta_\varepsilon$, $\hat{\nu}^\varepsilon := \nu * \eta_\varepsilon$, still satisfying $\nabla \cdot \hat{\mathbf{w}}^\varepsilon = \mu^\varepsilon - \nu^\varepsilon$. Since the Gaussian kernel is strictly positive, we also have strictly positive densities for $\hat{\mu}^\varepsilon$ and $\hat{\nu}^\varepsilon$. These convolved densities and vector field would do the job required by the theorem, but we have to take care of the support (which is not Ω') and of the boundary behavior.

Let us set $\int_{\Omega'} \hat{\mu}^\varepsilon = 1 - a_\varepsilon$ and $\int_{\Omega'} \hat{\nu}^\varepsilon = 1 - b_\varepsilon$. It is clear that $a_\varepsilon, b_\varepsilon \to 0$ as $\varepsilon \to 0$. Consider also $\hat{\mathbf{w}}^\varepsilon \cdot \mathbf{n}_{\Omega'}$: due to $d(\Omega, \partial\Omega') > 0$ and to the fact that η_ε goes uniformly to 0 locally outside the origin, we also have $|\hat{\mathbf{w}}^\varepsilon \cdot \mathbf{n}_{\Omega'}| \le c_\varepsilon$, with $c_\varepsilon \to 0$.

Consider u^ε the solution to

$$\begin{cases} \Delta u^\varepsilon = \frac{a_\varepsilon - b_\varepsilon}{|\Omega'|} & \text{inside } \Omega' \\ \frac{\partial u^\varepsilon}{\partial \mathbf{n}} = -\hat{\mathbf{w}}^\varepsilon \cdot \mathbf{n} & \text{on } \partial\Omega' \\ \int_{\Omega'} u^\varepsilon = 0 \end{cases}$$

and the vector field $\delta^\varepsilon = \nabla u^\varepsilon$. Note that a solution exists thanks to $-\int_{\partial\Omega'} \hat{\mathbf{w}}^\varepsilon \cdot \mathbf{n}_{\Omega'} = a_\varepsilon - b_\varepsilon$. Note also that an integration by parts shows

$$\int_{\Omega'} |\nabla u^\varepsilon|^2 = -\int_{\partial\Omega'} u^\varepsilon (\hat{\mathbf{w}}^\varepsilon \cdot \mathbf{n}_{\Omega'}) - \int_{\Omega'} u^\varepsilon \left(\frac{a_\varepsilon - b_\varepsilon}{|\Omega'|} \right) \le C ||\nabla u^\varepsilon||_{L^2} (c_\varepsilon + a_\varepsilon + b_\varepsilon)$$

and provides $||\nabla u^\varepsilon||_{L^2} \le C(a_\varepsilon + b_\varepsilon + c_\varepsilon) \to 0$. This shows $||\delta^\varepsilon||_{L^2} \to 0$.

Now take

$$\mu^\varepsilon = \hat{\mu}^\varepsilon \llcorner \Omega' + \frac{a_\varepsilon}{|\Omega'|}; \quad \nu^\varepsilon = \hat{\nu}^\varepsilon \llcorner \Omega' + \frac{b_\varepsilon}{|\Omega'|}; \quad \mathbf{w}^\varepsilon = \hat{\mathbf{w}}^\varepsilon \llcorner \Omega' + \delta^\varepsilon,$$

and check that all the requirements are satisfied. In particular, the last one is satisfied since $||\delta^\varepsilon||_{L^1} \to 0$ and $|\hat{\mathbf{w}}^\varepsilon| \rightharpoonup |\mathbf{w}|$ by general properties of the convolutions. □

Remark 4.9. Note that considering explicitly the dependence on Ω', it is also possible to obtain the same statement with a sequence of domains Ω'_ε converging to Ω (for instance, in the Hausdorff topology). It is just necessary to choose them so that setting $t_\varepsilon := d(\Omega, \partial\Omega'_\varepsilon)$, we have $||\eta_\varepsilon||_{L^\infty(B(0,t_\varepsilon)^c)} \to 0$. For the Gaussian kernel, this is satisfied whenever $t_\varepsilon^2/\varepsilon \to \infty$ and can be guaranteed by taking $t_\varepsilon = \varepsilon^{1/3}$.

With these tools we can now prove

Theorem 4.10. *For every finite vector measure $\mathbf{w} \in \mathscr{M}^d_{\mathrm{div}}(\Omega)$ and $\mu, \nu \in \mathscr{P}(\Omega)$ with $\nabla \cdot \mathbf{w} = \mu - \nu$, there exists a traffic plan $Q \in \mathscr{P}(\mathscr{C})$ with $(e_0)_\# Q = \mu$ and $(e_1)_\# Q = \nu$ such that $|\mathbf{w}_Q| = i_Q \le |\mathbf{w}|$, and $||\mathbf{w} - \mathbf{w}_Q|| + ||\mathbf{w}_Q|| = ||\mathbf{w} - \mathbf{w}_Q|| + i_Q(\Omega) = ||\mathbf{w}||$. In particular we have $|\mathbf{w}_Q| \neq |\mathbf{w}|$ unless $\mathbf{w}_Q = \mathbf{w}$.*

Proof. By means of Lemma 4.8 and Remark 4.9, we can produce an approximating sequence $(\mathbf{w}^\varepsilon, \mu^\varepsilon, \nu^\varepsilon) \rightharpoonup (\mathbf{w}, \mu, \nu)$ of C^∞ functions supported on domains Ω_ε converging to Ω. We apply Dacorogna-Moser's construction to this sequence of vector fields, thus obtaining a sequence of measures Q_ε. We can consider these measures as probability measures on $\mathscr{C} := \mathrm{AC}(\Omega')$ (where $\Omega \subset \Omega_\varepsilon \subset \Omega'$) which are, each, concentrated on curves valued in Ω_ε. They satisfy $i_{Q_\varepsilon} = |\mathbf{w}^\varepsilon|$ and $\mathbf{w}_{Q_\varepsilon} = \mathbf{w}^\varepsilon$. We can reparametrize by constant speed the curves on which Q_ε is supported, without changing traffic intensities and traffic flows. This means that we use curves ω such that $L(\omega) = \mathrm{Lip}(\omega)$. The equalities

$$\int_{\mathscr{C}} \mathrm{Lip}(\omega)\, \mathrm{d}Q_\varepsilon(\omega) = \int_{\mathscr{C}} L(\omega)\, \mathrm{d}Q_\varepsilon(\omega) = \int_{\Omega'} i_{Q_\varepsilon} = \int_{\Omega'} |\mathbf{w}^\varepsilon| \to |\mathbf{w}|(\Omega') = |\mathbf{w}|(\Omega)$$

show that $\int_{\mathscr{C}} \mathrm{Lip}(\omega)\, \mathrm{d}Q_\varepsilon(\omega)$ is bounded and hence Q_ε is tight (since the sets $\{\omega \in \mathscr{C} : \mathrm{Lip}(\omega) \le L\}$ are compact). Hence, up to subsequences, we can assume $Q_\varepsilon \rightharpoonup Q$. The measure Q is obviously concentrated on curves valued in Ω. The measures Q_ε were constructed so that $(e_0)_\# Q_\varepsilon = \mu^\varepsilon$ and $(e_1)_\# Q_\varepsilon = \nu^\varepsilon$, which implies, at the limit, $(e_0)_\# Q = \mu$ and $(e_1)_\# Q = \nu$. Moreover, thanks to Proposition 4.7, since $i_{Q_\varepsilon} = |\mathbf{w}_\varepsilon| \rightharpoonup |\mathbf{v}|$ and $\mathbf{w}_{Q_\varepsilon} \rightharpoonup \mathbf{w}$, we get $|\mathbf{w}| \ge i_Q \ge |\mathbf{w}_Q|$ and $||\mathbf{w} - \mathbf{w}_Q|| \le |\mathbf{w}|(\Omega) - i_Q(\Omega)$. This gives $||\mathbf{w} - \mathbf{w}_Q|| + ||\mathbf{w}_Q|| \le ||\mathbf{w} - \mathbf{w}_Q|| + i_Q(\Omega) \le ||\mathbf{w}||$ and the opposite inequality $||\mathbf{w}|| \le ||\mathbf{w} - \mathbf{w}_Q|| + ||\mathbf{w}_Q||$ is always satisfied. □

Remark 4.11. The previous statement contains a milder version of Theorem C in [287], i.e., the decomposition of any $\mathbf{w}$ into a cycle $\mathbf{w} - \mathbf{w}_Q$ (we call *cycle* all divergence-free vector measures) and a flow $\mathbf{w}_Q$ induced by a measure on paths, with $||\mathbf{w}|| = ||\mathbf{w} - \mathbf{w}_Q|| + ||\mathbf{w}_Q||$. The only difference with the theorem in [287] is the fact that it guarantees that one can choose Q concentrated on *simple* curves, which we did not take care of here (on the other hand, **Ex**(27) provides a partial solution to this issue).

Remark 4.12. It is possible to guess what happens to a cycle through this construction. Imagine the following example (as in Figure 4.1): Ω is composed of two parts Ω^+ and Ω^-, with $\mathrm{spt}(\mu) \cup \mathrm{spt}(\nu) \subset \Omega^+$ and a cycle of $\mathbf{w}$ is contained in Ω^-.

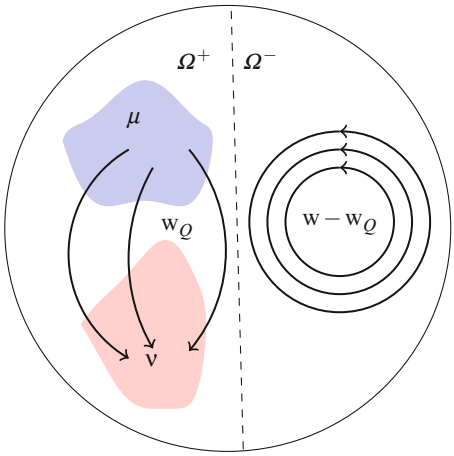


Fig. 4.1 The decomposition of $\mathbf{w}$ in the example of Remark 4.12

When we build the approximations μ^ε and ν^ε, we will have positive mass in Ω^-, but very small. Because of the denominator in the definition of $\mathbf{v}$, the curves produced by Dacorogna-Moser will follow this cycle very fast, passing many times on each point of the cycle. Hence, the flow $\mathbf{w}^\varepsilon$ in Ω^- is obtained from a very small mass which passes many times. This implies that the measure Q_ε of this set of curves will disappear at the limit $\varepsilon \to 0$. Hence, the measure Q will be concentrated on curves staying in Ω^+ and $\mathbf{w}_Q = 0$ on Ω^-. However, in this way we got rid of that particular cycle, but the same does not happen for cycles located in regions with positive masses of μ and ν. In particular nothing guarantees that $\mathbf{w}_Q$ has no cycles.

4.2.4 *Beckman problem in one dimension*

The 1D case is very easy in what concerns Beckmann formulation of the optimal transport problem, but it is interesting to analyze it: indeed, it allows to check the consistency with the Monge formulation and to use the results throughout the next sections. We will take $\Omega = [a, b] \subset \mathbb{R}$.

First of all, note that the condition $\nabla \cdot \mathbf{w} = \mu - \nu$ is much stronger in dimension one than in higher dimension. Indeed, the divergence is the trace of the Jacobian matrix, and hence prescribing it only gives one constraint on a matrix which has a priori $d \times d$ degrees of freedom. On the contrary, in dimension one there is only one partial derivative for the vector field $\mathbf{w}$ (which is actually a scalar), and this completely prescribes the behavior of $\mathbf{w}$. Indeed, the condition $\nabla \cdot \mathbf{w} = \mu - \nu$ with Neumann boundary conditions implies that $\mathbf{w}$ must be the antiderivative of $\mu - \nu$ with $\mathbf{w}(a) = 0$ (the fact that μ and ν have the same mass also implies $\mathbf{w}(b) = 0$). Note that the fact that its derivative is a measure gives $\mathbf{w} \in BV([a, b])$ (we can say that $\mathscr{M}^d_{\mathrm{div}} = BV$ when $d = 1$).

Box 4.4. Memo: *Bounded variation functions in one variable*

BV functions are defined as L^1 functions whose distributional derivatives are measures. In dimension one this has a lot of consequences. In particular these functions coincide a.e. with functions which have bounded total variation in a pointwise sense: for each $f : [a,b] \to \mathbb{R}$, define

$$TV(f;[a,b]) := \sup\left\{ \sum_{i=0}^{N-1} |f(t_{i+1}) - f(t_i)| \ : \ a = t_0 < t_1 < t_2 < \cdots < t_N = b \right\}.$$

Functions of bounded variation are defined as those f such that $TV(f;[a,b]) < \infty$. It is easy to check that BV functions form a vector space and that monotone functions are BV (indeed, if f is monotone, we have $TV(f;[a,b]) = |f(b) - f(a)|$. Lipschitz functions are also BV and $TV(f;[a,b]) \le \mathrm{Lip}(f)(b-a)$. On the other hand, continuous functions are not necessarily BV (but absolutely continuous functions are BV), nor it is the case for differentiable functions (obviously, C^1 functions, which are Lipschitz on bounded intervals, are BV). As an example one can consider

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x^2}\right) & \text{if } x \neq 0 \\ 0 & \text{for } x = 0, \end{cases}$$

which is differentiable everywhere but not BV.

On the other hand, BV functions have several properties.

Properties of BV functions in $\mathbb{R}$. If $TV(f;[a,b]) < \infty$, then f is the difference of two monotone functions (in particular we can write $f(x) = TV(f;[a,x]) - (TV(f;[a,x]) - f(x))$, both terms being nondecreasing functions); it is a bounded function and $\sup f - \inf f \le TV(f;[a,b])$; it has the same continuity and differentiability properties of monotone functions (it admits left and right limits at every point; it is continuous up to a countable set of points and differentiable a.e.).

In particular, in 1D, we have $BV \subset L^\infty$ which is not the case in higher dimension (in general, we have $BV \subset L^{d/(d-1)}$).

We finish by stressing the connections with measures: for every positive measure μ on $[a,b]$, we can build a monotone function by taking its cumulative distribution function, i.e., $F_\mu(x) = \mu([a,x])$, and the distributional derivative of this function is exactly the measure μ. Conversely, every monotone increasing function on a compact interval is the cumulative distribution function of a (unique) positive measure, and every BV function is the cumulative distribution function of a (unique) signed measure.

As a consequence, we have the following facts:

- In dimension one, there is only one competitor $\mathbf{w}$ which is given by $\mathbf{w} = F_\mu - F_\nu$ with $F_\mu(x) = \mu([a,x])$ and $F_\nu(x) = \nu([a,x])$.
- This field $\mathbf{w}$ belongs to $BV([a,b])$ and hence to every L^p space, including L^∞.
- Possible higher regularity of $\mathbf{w}$ depends on the regularity of $\mu - \nu$ (for instance, we have $\mathbf{w} \in W^{1,p}$ whenever $\mu, \nu \in L^p$).
- The minimal cost in Beckmann's problem is given by $||F_\mu - F_\nu||_{L^1}$, which is consistent with Proposition 2.17.

- The transport density σ, characterized by $\mathbf{w} = -u' \cdot \sigma$, is given by $\sigma = |\mathbf{w}|$ and shares the same summability properties of $\mathbf{w}$; it also belongs to BV as a composition of a BV function with the absolute value function.

4.2.5 *Characterization and uniqueness of the optimal* $\mathbf{w}$

In this section we will show two facts: first we prove that the optimal $\mathbf{w}$ in the (BP) always comes from an optimal transport plan γ, and then we prove that all the optimal γ give the same $\mathbf{w}_{[\gamma]}$ and the same σ_γ, provided one of the two measures is absolutely continuous.

Theorem 4.13. *Let* $\mathbf{w}$ *be optimal in (BP): then there is an optimal transport plan* γ *such that* $\mathbf{w} = \mathbf{w}_{[\gamma]}$.

Proof. Thanks to Theorem 4.10, we can find a measure $Q \in \mathscr{P}(\mathscr{C})$ with $(e_0)_\# Q = \mu$ and $(e_1)_\# Q = \nu$ such that $|\mathbf{w}| \geq |\mathbf{w}_Q|$. Yet, the optimality of $\mathbf{w}$ implies the equality $|\mathbf{w}| = |\mathbf{w}_Q|$ and the same Theorem 4.10 gives in such a case $\mathbf{w} = \mathbf{w}_Q$, as well as $|\mathbf{w}| = i_Q$. We can assume Q to be concentrated on curves parametrized by constant speed. Define $S : \Omega \times \Omega \to C$ the map associating with every pair (x, y) the segment $\omega_{x,y}$ parametrized with constant speed: $\omega_{x,y}(t) = (1-t)x + ty$. The statement is proven if we can prove that $Q = S_\# \gamma$ with γ an optimal transport plan.

Indeed, using again the optimality of $\mathbf{w}$ and Theorem 4.10, we get

$$\begin{aligned}\min(\mathrm{BP}) = |\mathbf{w}|(\Omega) = i_Q(\Omega) &= \int_{\mathscr{C}} L(\omega)\,\mathrm{d}Q(\omega) \geq \int_{\mathscr{C}} |\omega(0) - \omega(1)|\,\mathrm{d}Q(\omega) \\ &= \int_{\Omega\times\Omega} |x-y|\,\mathrm{d}((e_0, e_1)_\# Q)(x, y) \geq \min(\mathrm{KP}).\end{aligned}$$

The equality $\min(\mathrm{BP}) = \min(\mathrm{KP})$ implies that all these inequalities are equalities. In particular Q must be concentrated on curves such that $L(\omega) = |\omega(0) - \omega(1)|$, i.e., segments. Also, the measure $(e_0, e_1)_\# Q$, which belongs to $\Pi(\mu, \nu)$, must be optimal in (KP). This concludes the proof. □

The proof of the following result is essentially taken from [8].

Theorem 4.14. *If* $\mu \ll \mathscr{L}^d$, *then the vector field* $\mathbf{w}_{[\gamma]}$ *does not depend on the choice of the optimal plan* γ.

Proof. Let us fix a Kantorovich potential u for the transport between μ and ν. This potential does not depend on the choice of γ. It determines a partition into transport rays: Corollary 3.8 guarantees that the only points of Ω which belong to several transport rays are non-differentiability points for u and are hence Lebesgue-negligible. Let us call S the set of points which belong to several transport rays: we have $\mu(S) = 0$, but we do not suppose $\nu(S) = 0$ (ν is not supposed to be absolutely continuous). However, γ is concentrated on $(\pi_x)^{-1}(S^c)$. We can then disintegrate

(see Section 2.3) γ according to the transport ray containing the point x. More precisely, we define a map $R : \Omega \times \Omega \to \mathscr{R}$, valued in the set $\mathscr{R}$ of all transport rays, sending each pair (x, y) into the ray containing x. This is well-defined γ-a.e. and we can write $\gamma = \gamma^r \otimes \lambda$, where $\lambda = R_\# \gamma$, and we denote by r the variable related to transport rays. Note that, for a.e. $r \in \mathscr{R}$, the plan γ^r is optimal between its own marginals (otherwise we could replace it with an optimal plan, do it in a measurable way, and improve the cost of γ).

The measure $\mathbf{w}_{[\gamma]}$ may also be obtained through this disintegration, and we have $\mathbf{w}_{[\gamma]} = \mathbf{w}_{\gamma^r} \otimes \lambda$. This means that in order to prove that $\mathbf{w}_{[\gamma]}$ does not depend on γ, we just need to prove that each $\mathbf{w}_{\gamma^r}$ and the measure λ do not depend on it. For the measure λ this is easy: it has been obtained as an image measure through a map only depending on x and hence only depends on μ. Concerning $\mathbf{w}_{\gamma^r}$, note that it is obtained in the standard Beckmann way from an optimal plan, γ^r. Hence, thanks to the considerations in Section 4.2.4 about the 1D case, it uniquely depends on the marginal measures of this plan.

This means that we only need to prove that $(\pi_x)_\# \gamma^r$ and $(\pi_y)_\# \gamma^r$ do not depend on γ. Again, this is easy for $(\pi_x)_\# \gamma^r$, since it must coincide with the disintegration of μ according to the map R (by uniqueness of the disintegration). It is more delicate for the second marginal.

The second marginal $\nu^r := (\pi_y)_\# \gamma^r$ will be decomposed into two parts:

$$(\pi_y)_\#(\gamma^r_{|\Omega\times S}) \;+\; (\pi_y)_\#(\gamma^r_{|\Omega\times S^c}).$$

This second part coincides with the disintegration of $\nu_{|S^c}$, which obviously does not depend on γ (since it only depends on the set S, which is built upon u).

We need now to prove that $\nu^r_{|S} = (\pi_y)_\#(\gamma^r_{|\Omega\times S})$ does not depend on γ. Yet, this measure can only be concentrated on the two endpoints of the transport ray r, since these are the only points where different transport rays can meet. This means that this measure is purely atomic and composed by at most two Dirac masses. Not only the endpoint where u is maximal (i.e., the first one for the order on the ray; see Section 3.1) cannot contain any mass of ν: indeed the transport must follow a precise direction on each transport ray (as a consequence of $u(x)-u(y) = |x-y|$ on $\mathrm{spt}(\gamma)$), and the only way to have some mass of the target measure at the "beginning" of the transport ray would be to have an atom for the source measure as well. Yet, μ is absolutely continuous and Property N holds (see Section 3.1.4 and Theorem 3.12, which means that the set of rays r where μ^r has an atom is negligible). Hence $\nu^r_{|S}$ is a single Dirac mass. The mass equilibrium condition between μ^r and ν^r implies that the value of this mass must be equal to the difference $1 - \nu^r_{|S^c}(r)$, and this last quantity does not depend on γ but only on μ and ν.

Finally, this proves that each $\mathbf{w}_{\gamma^r}$ does not depend on the choice of γ. □

Corollary 4.15. *If $\mu \ll \mathscr{L}^d$, then the solution of (BP) is unique.*

Proof. We have seen in Theorem 4.13 that any optimal $\mathbf{w}$ is of the form $\mathbf{w}_{[\gamma]}$ and in Theorem 4.14 that all the fields $\mathbf{w}_{[\gamma]}$ coincide. □

4.3 Summability of the transport density

The analysis of the Beckmann problem performed in the previous sections was mainly made in a measure setting, and the optimal $\mathbf{w}$ and the transport density σ were just measures on Ω. We investigate here the question whether they have extra regularity properties supposing extra assumptions on μ and/or ν.

We will give summability results, proving that σ is in some cases absolutely continuous and providing L^p estimates. The proofs are essentially taken from [273] (also look at [61, 166]): previous results, through very different techniques, were first presented in [143–145]. In these papers, different estimates on the "dimension" of σ were also presented, thus giving interesting information should σ fail to be absolutely continuous. Let us also observe that [208], while applying these tools to image processing problems, contains a new[4] proof of absolute continuity.

Note that higher-order questions (such as whether σ is continuous or Lipschitz or more regular provided μ and ν have smooth densities) are completely open up to now. The only exception is a partial result in dimension 2 (see [171]), where a continuity result is given if μ and ν have Lipschitz densities on disjoint convex domains.

Open Problem (continuity of the transport density): Is it true that, supposing that μ, ν have continuous densities, the transport density σ is also continuous? Is it true that it is Lipschitz if they are Lipschitz, and/or Hölder if they are Hölder?

In all that follows Ω is a compact and convex domain in $\mathbb{R}^d$, and two probability measures μ, ν are given on it. At least one of them will be absolutely continuous, which implies uniqueness for σ (see Theorem 4.14). The first result – first proven in [162] – will concern L^1 summability (i.e., absolute continuity).

Theorem 4.16. *Suppose $\Omega \subset \mathbb{R}^d$ is a convex domain, and $\mu \ll \mathscr{L}^d$. Let σ be the transport density associated with the transport of μ onto ν. Then $\sigma \ll \mathscr{L}^d$.*

Proof. Let γ be an optimal transport plan from μ to ν and take $\sigma = \sigma_\gamma$; call μ_t the standard interpolation between the two measures: $\mu_t = (\pi_t)_\# \gamma$ where $\pi_t(x, y) = (1-t)x + ty$. We have $\mu_0 = \mu$ and $\mu_1 = \nu$.

We have already seen (go back to (4.6)) that the transport density σ may be written as

$$\sigma = \int_0^1 (\pi_t)_\#(c \cdot \gamma) \mathrm{d}t,$$

where $c : \Omega \times \Omega \to \mathbb{R}$ is the cost function $c(x, y) = |x - y|$ (hence $c \cdot \gamma$ is a positive measure on $\Omega \times \Omega$).

[4] Their proof is somehow intermediate between that of [143] and the one we present here: indeed, approximation by atomic measures is also performed in [208], as here, but on both the source and the target measure, which requires a geometric analysis of the transport rays as in [143].

Since Ω is bounded, it is evident that we have

$$\sigma \le C \int_0^1 \mu_t \, \mathrm{d}t. \tag{4.11}$$

To prove that σ is absolutely continuous, it is sufficient to prove that almost every measure μ_t is absolutely continuous, so that, whenever $|A| = 0$, we have $\sigma(A) \le C \int_0^1 \mu_t(A) \mathrm{d}t = 0$.

We will prove $\mu_t \ll \mathscr{L}^d$ for $t < 1$. First, we will suppose that ν is finitely atomic (the point $(x_i)_{i=1,\dots,N}$ being its atoms). In this case we will choose γ to be any optimal transport plan induced by a transport map T (which exists, since $\mu \ll \mathscr{L}^d$). Note that the absolute continuity of σ is an easy consequence of the behavior of the optimal transport from μ to ν (which is composed by N homotheties), but we also want to quantify this absolute continuity, in order to go on with an approximation procedure.

Remember that μ is absolutely continuous and hence there exists a correspondence $\varepsilon \mapsto \delta = \delta(\varepsilon)$ such that

$$|A| < \delta(\varepsilon) \Rightarrow \mu(A) < \varepsilon. \tag{4.12}$$

Take now a Borel set A and look at $\mu_t(A)$. The domain Ω is the disjoint union of a finite number of sets $\Omega_i = \mathrm{T}^{-1}(\{x_i\})$. We call $\Omega_i(t)$ the images of Ω_i through the map $x \mapsto (1-t)x + t\mathrm{T}(x)$ (Figure 4.2). These sets are essentially disjoint. Why? Because if a point z belongs to $\Omega_i(t)$ and $\Omega_j(t)$, then two transport rays cross at z, the one going from $x_i' \in \Omega_i$ to x_i and the one from $x_j' \in \Omega_j$ to x_j. The only possibility is that these two rays are actually the same, i.e., that the five points x_i', x_j', z, x_i, x_j are aligned. But this implies that z belongs to one of the lines connecting two atoms x_i and x_j. Since we have finitely many of these lines, this set is negligible. Note that this argument only works for $d > 1$ (we will not waste time on the case $d = 1$, since the transport density is always a BV, and hence bounded, function). Moreover, if we sticked to the ray-monotone optimal transport map, we could have actually proved that these sets are truly disjoint, with no negligible intersection.

Now we have

$$\mu_t(A) = \sum_i \mu_t(A \cap \Omega_i(t)) = \sum_i \mu_0\left(\frac{A \cap \Omega_i(t) - tx_i}{1-t}\right) = \mu_0\left(\bigcup_i \frac{A \cap \Omega_i(t) - tx_i}{1-t}\right).$$

Since for every i we have

$$\left|\frac{A \cap \Omega_i(t) - tx_i}{1-t}\right| = \frac{1}{(1-t)^d} \left|A \cap \Omega_i(t)\right|,$$

we obtain

$$\left|\bigcup_i \frac{A \cap \Omega_i(t) - tx_i}{1-t}\right| \le \frac{1}{(1-t)^d} |A|.$$

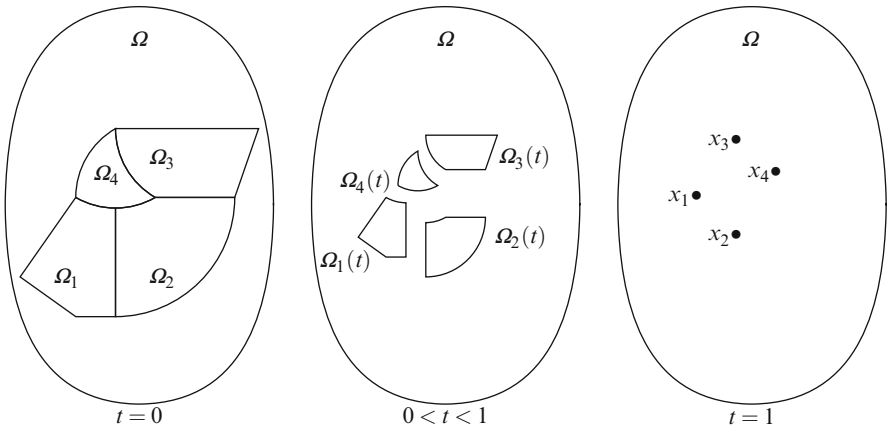


Fig. 4.2 Disjointness of the sets $\Omega_i(t)$

Hence it is sufficient to suppose $|A| < (1-t)^d\delta(\varepsilon)$ to get $\mu_t(A) < \varepsilon$. This confirms $\mu_t \ll \mathscr{L}^d$ and gives an estimate that may pass to the limit.

Take a sequence ν^n of atomic measures converging to ν. The corresponding optimal transport plans γ^n converge to an optimal transport plan γ and μ_t^n converge to the corresponding μ_t (see Theorem 1.50 in Chapter 1). Hence, to prove absolute continuity for the transport density σ associated with such a γ, it is sufficient to prove that these μ_t are absolutely continuous.

Take a set A such that $|A| < (1-t)^d\delta(\varepsilon)$. Since the Lebesgue measure is regular (see the definition in the memo Box 1.6), A is included in an open set B such that $|B| < (1-t)^d\delta(\varepsilon)$. Hence $\mu_t^n(B) < \varepsilon$. Passing to the limit, thanks to weak convergence and semi-continuity on open sets, we have

$$\mu_t(A) \le \mu_t(B) \le \liminf_n \mu_t^n(B) \le \varepsilon.$$

This proves $\mu_t \ll \mathscr{L}^d$ and hence $\sigma \ll \mathscr{L}^d$. □

Remark 4.17. Where did we use the optimality of γ? We did it when we said that the $\Omega_i(t)$ are disjoint. For a discrete measure ν, it is always true that the measures μ_t corresponding to any transport plan γ are absolutely continuous for $t < 1$, but their absolute continuity may degenerate at the limit if we allow the sets $\Omega_i(t)$ to superpose (since in this case densities sum up and the estimates may depend on the number of atoms).

Remark 4.18. Note that we strongly used the equivalence between the two different definitions of absolute continuity, i.e., the $\varepsilon \leftrightarrow \delta$ correspondence, on the one hand, and the condition on negligible sets, on the other. Indeed, to prove that the condition $\mu_t \ll \mathscr{L}^d$ passes to the limit, we need the first one, while to deduce $\sigma \ll \mathscr{L}^d$, we need the second one.

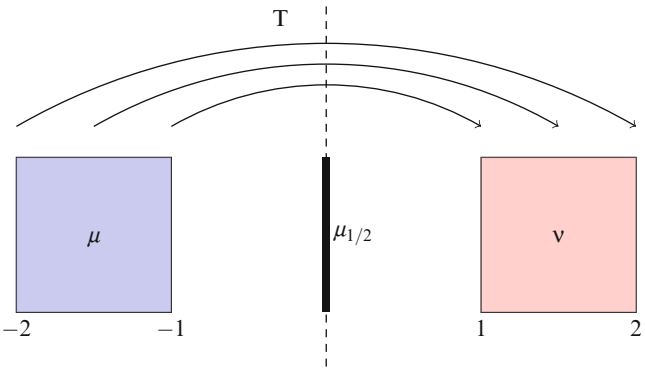


Fig. 4.3 An example of optimal transport plan that cannot be approximated with optimal plans with atomic second marginal

Remark 4.19. As a by-product of our proof, we can see that any optimal transport plan from μ to ν which is approximable through optimal transport plans from μ to atomic measures must be such that all the interpolating measures μ_t (for *every* $t \in]0,1[$) are absolutely continuous. This property is not satisfied by every optimal transport plan, since, for instance, the plan γ which sends $\mu = \mathscr{L}^2_{\lfloor[-2,-1]\times[0,1]}$ onto $\nu = \mathscr{L}^2_{\lfloor[1,2]\times[0,1]}$ moving (x,y) to $(-x,y)$ is optimal but is such that $\mu_{1/2} = \mathscr{H}^1_{\lfloor\{0\}\times[0,1]}$. Hence, this plan cannot be approximated by optimal plans sending μ onto atomic measures. On the other hand, we proved in Lemma 3.20 that the monotone optimal transport can indeed be approximated in a similar way (Figure 4.3).

From now on we will often confuse absolutely continuous measures with their densities and write $||\mu||_{L^p}$ for $||f||_{L^p(\Omega)}$ when $\mu = f \cdot \mathscr{L}^d$.

Theorem 4.20. *Suppose $\mu = f \cdot \mathscr{L}^d$, with $f \in L^p(\Omega)$, where Ω is compact and convex. Then, if $p < d' := d/(d-1)$, the unique transport density σ associated with the transport of μ onto ν belongs to $L^p(\Omega)$ as well, and if $p \geq d'$, it belongs to any space $L^q(\Omega)$ for $q < d'$.*

Proof. Start from the case $p < d'$: following the same strategy (and the same notations) as before, it is sufficient to prove that each measure μ_t (for $t \in [0,1[$) is in L^p and to estimate their L^p norm. Then we will use

$$||\sigma||_{L^p} \leq C \int_0^1 ||\mu_t||_{L^p} \mathrm{d}t,$$

(which is a consequence of (4.11) and of Minkowski inequality), the conditions on p being chosen exactly so that this integral converges.

Consider first the discrete case: we know that μ_t is absolutely continuous and that its density coincides on each set $\Omega_i(t)$ with the density of a homothetic image

of μ on Ω_i, the homothetic ratio being $(1-t)$. Hence, if f_t is the density of μ_t, we have

$$\int_\Omega f_t(x)^p\,\mathrm{d}x = \sum_i \int_{\Omega_i(t)} f_t(x)^p\,\mathrm{d}x = \sum_i \int_{\Omega_i} \left(\frac{f(x)}{(1-t)^d}\right)^p (1-t)^d\,\mathrm{d}x$$
$$= (1-t)^{d(1-p)} \sum_i \int_{\Omega_i} f(x)^p\,\mathrm{d}x = (1-t)^{d(1-p)} \int_\Omega f(x)^p\,\mathrm{d}x.$$

We get $||\mu_t||_{L^p} = (1-t)^{-d/p'}||\mu||_{L^p}$, where $p' = p/(p-1)$.

This inequality, which is true in the discrete case, stays true at the limit as well. If ν is not atomic, approximate it through a sequence μ_1^n and take optimal plans γ^n and interpolating measures μ_t^n. Up to subsequences we have $\gamma^n \rightharpoonup \gamma$ (for an optimal transport plan γ) and $\mu_t^n \rightharpoonup \mu_t$ (for the corresponding interpolation); by semi-continuity, we have

$$||\mu_t||_{L^p} \le \liminf_n ||\mu_t^n||_{L^p} \le (1-t)^{-d/p'}||\mu_0||_{L^p}$$

and we deduce

$$||\sigma||_{L^p} \le C \int_0^1 ||\mu_t||_{L^p}\mathrm{d}t \le C||\mu_0||_{L^p} \int_0^1 (1-t)^{-d/p'}\,\mathrm{d}t.$$

The last integral is finite whenever $p' > d$, i.e., $p < d' = d/(d-1)$.

The second part of the statement (the case $p \ge d'$) is straightforward once one considers that any density in L^p also belongs to any L^q space for $q < p$. □

Example 4.21. We can see an example where, even if $\mu \in L^\infty$, the singularity of ν prevents σ from being $L^{d'}$. Consider $\mu = f \cdot \mathscr{L}^d$ with $f = \mathbb{1}_A$ and $A = B(0,R) \setminus B(0,R/2)$ (with R chosen so that $\int f = 1$), and take $\nu = \delta_0$. Note that $\frac{R}{2}\gamma \le c \cdot \gamma \le R\gamma$, which implies that the summability of $\sigma = \int_0^1 (\pi_t)_\#(c \cdot \gamma)\mathrm{d}t$ is the same as that of $\int_0^1 (\pi_t)_\#(\gamma)\mathrm{d}t = \int_0^1 \mu_t\,\mathrm{d}t$. We have $\mu_t = f_t \cdot \mathscr{L}^d$, with $f_t = (1-t)^{-d}\mathbb{1}_{(1-t)A}$; hence

$$\int_0^1 f_t(x)\,\mathrm{d}t = \int_{\frac{1}{R}|x|}^{\frac{2}{R}|x|} \frac{1}{s^d}\mathrm{d}s = c|x|^{1-d} \quad \text{for } |x| \le \frac{R}{2}$$

(where we used the change of variable $s = 1-t$). This function belongs to L^p in a neighborhood of 0 in dimension d only if we have $(1-d)p + (d-1) > -1$, i.e., if $(d-1)(p-1) < 1$. This is exactly $p < d'$.

We saw in the previous theorems that the measures μ_t inherit some regularity (absolute continuity or L^p summability) from μ exactly as it happens for homotheties of ratio $1-t$. This regularity degenerates as $t \to 1$, but we saw two cases where this degeneracy produced no problem: for proving absolute continuity, where

the separate absolute continuous behavior of almost all the μ_t was sufficient, and for L^p estimates, provided the degeneracy stays integrable.

It is natural to try to exploit another strategy: suppose both μ and ν share some regularity assumption (e.g., they belong to L^p). Then we can give estimate on μ_t for $t \leq 1/2$ starting from μ and for $t \geq 1/2$ starting from ν. In this way we have no degeneracy!

This strategy works quite well, but it has an extra difficulty: in our previous estimates, we didn't know a priori that μ_t shared the same behavior of piecewise homotheties of μ; we got it as a limit from discrete approximations. And, when we pass to the limit, we do not know which optimal transport γ will be selected as a limit of the optimal plans γ^n. This was not important in the previous section, since any optimal γ induces the same transport density σ. But here we would like to glue together estimates on μ_t for $t \leq 1/2$ which have been obtained by approximating μ_1 and estimates on μ_t for $t \geq 1/2$ which come from the approximation of μ_0. Should the two approximations converge to two different transport plans, we could not put together the two estimates and deduce anything on σ.

Hence, the main technical issue that we need to consider is proving that one particular optimal transport plan, i.e., the ray-monotone one, can be approximated in both directions. Lemma 3.20 exactly does the job (and, indeed, it was proven in [273] exactly for this purpose). Yet, the transport plans γ_ε we build in the approximation are not optimal for the cost $|x-y|$ but for some costs $|x-y|+\varepsilon|x-y|^2$. We need to do this in order to force the selected limit optimal transport to be the monotone one (through a secondary variational problem, say). Anyway, this will not be an issue: these approximating optimal transport plans will share the same geometric properties that will imply disjointness for the sets $\Omega_i(t)$. In particular, we can prove the following estimate.

Lemma 4.22. *Let γ be an optimal transport plan between μ and an atomic measure ν for a transport cost $c(x,y) = h(y-x)$ where $h : \mathbb{R}^d \to \mathbb{R}$ is a strictly convex function. Set as usual $\mu_t = (\pi_t)_\# \gamma$. Then we have $||\mu_t||_{L^p} \leq (1-t)^{-d/p'} ||\mu||_{L^p}$.*

Proof. The result is exactly the same as in Theorem 4.20, where the key tool is the fact that μ_t coincides on every set $\Omega_i(t)$ with a homothety of μ_0. The only fact that must be checked again is the disjointness of the sets $\Omega_i(t)$.

To do so, take a point $x \in \Omega_i(t) \cap \Omega_j(t)$. Hence there exist x_i, x_j belonging to Ω_i and Ω_j, respectively, so that $x = (1-t)x_i + ty_i = (1-t)x_j + ty_j$, being $y_i = \mathrm{T}(x_i)$ and $y_j = \mathrm{T}(x_j)$ two atoms of ν. But this would mean that $\mathrm{T}_t := (1-t)\mathrm{id} + t\mathrm{T}$ is not injective, which is a contradiction to the following Lemma 4.23. Hence the sets $\Omega_i(t)$ are disjoint and this implies the bound on μ_t. □

Lemma 4.23. *Let γ be an optimal transport plan between μ and ν for a transport cost $c(x,y) = h(y-x)$ where $h : \mathbb{R}^d \to \mathbb{R}$ is a strictly convex function, and suppose that it is induced by a transport map* T. *Choose a representative of* T *such that $(x, \mathrm{T}(x)) \in \mathrm{spt}(\gamma)$ for all x. Then the map $x \mapsto (1-t)x + t\mathrm{T}(x)$ is injective for $t \in]0,1[$.*

Proof. Suppose that there exist $x \neq x'$ such that

$$(1-t)x + t\mathrm{T}(x) = (1-t)x' + t\mathrm{T}(x') = x_0.$$

Set $a = \mathrm{T}(x) - x$ and $b = \mathrm{T}(x') - x'$. This also means $x = x_0 - ta$ and $x' = x_0 - tb$. In particular, $x \neq x'$ implies $a \neq b$.

The c-cyclical monotonicity of the support of the optimal γ implies

$$h(a) + h(b) \leq h(\mathrm{T}(x') - x) + h(\mathrm{T}(x) - x') = h(tb + (1-t)a) + h(ta + (1-t)b).$$

Yet, $a \neq b$, and strict convexity implies

$$h(tb+(1-t)a)+h(ta+(1-t)b) < th(b)+(1-t)h(a)+th(a)+(1-t)h(b) = h(a)+h(b),$$

which is a contradiction. □

Theorem 4.24. *Suppose that μ and ν are probability measures on Ω, both belonging to $L^p(\Omega)$, and σ the unique transport density associated with the transport of μ onto ν. Then σ belongs to $L^p(\Omega)$ as well.*

Proof. Let us consider the optimal transport plan $\overline{\gamma}$ from μ to ν defined by (3.4). We know that this transport plan may be approximated by plans γ^ε which are optimal for the cost $|x-y| + \varepsilon|x-y|^2$ from μ to some discrete atomic measures ν_ε. The corresponding interpolation measures μ_t^ε satisfy the L^p estimate from Lemma 4.22 and, at the limit, we have

$$||\mu_t||_{L^p} \leq \liminf_{\varepsilon \to 0} ||\mu_t^\varepsilon||_{L^p} \leq (1-t)^{-d/p'} ||\mu||_{L^p}.$$

The same estimate may be performed from the other direction, since the very same transport plan $\overline{\gamma}$ may be approximated by optimal plans for the cost $|x-y|+\varepsilon|x-y|^2$ from atomic measures to ν. Putting together the two estimates, we have

$$||\mu_t||_{L^p} \leq \min\left\{(1-t)^{-d/p'} ||\mu||_{L^p},\, t^{-d/p'} ||\nu||_{L^p}\right\} \leq 2^{d/p'} \max\left\{||\mu||_{L^p}, ||\nu||_{L^p}\right\}.$$

Integrating these L^p norms, we get the bound on $||\sigma||_{L^p}$. □

Remark 4.25. The same result could have been obtained in a strongly different way, thanks to the displacement convexity of the functional $\mu \mapsto ||\mu||_{L^p}^p$. This functional is actually convex along geodesics in the space $W_q(\Omega)$ for $q > 1$ (see Proposition 7.29 where, unfortunately, the notation for p and q is reversed). Then we pass to the limit as $q \to 1$: this gives the result on the interpolating measures corresponding to the optimal plan which is obtained as a limit of the K_q-optimal plans. This plan is, by the way, $\overline{\gamma}$ again. And the integral estimate comes straightforward. Yet, this requires more sophisticated tools than what developed so far.

Example 4.26. We can provide examples (in dimension higher than one) where the summability of σ is no more than that of μ and ν. Take, for instance, $\mu = f \cdot \mathscr{L}^d$ and $\nu = g \cdot \mathscr{L}^d$ with $f(x_1, x_2, \dots, x_d) = f_0(x_2, \dots x_d)\mathbb{1}_{[0,1]}(x_1)$ and $g(x_1, x_2, \dots, x_d) = f_0(x_2, \dots x_d)\mathbb{1}_{[4,5]}(x_1)$. It is not difficult to check that the transport rays in this example go in the direction of the first coordinate vector and that the optimal σ is absolutely continuous with density given by $\sigma(x) = f_0(x_2, \dots x_d)s(x_1)$, where $s : \mathbb{R} \to \mathbb{R}_+$ is characterized by $s' = \mathbb{1}_{[0,1]} - \mathbb{1}_{[4,5]}$, $s(0) = s(5) = 0$ (which comes from the fact that s would be the transport density for the 1D transport between $\mathbb{1}_{[0,1]}$ and $\mathbb{1}_{[4,5]}$). Hence, from $s = 1$ on $[1, 4]$, one sees that the summability of σ is the same of that of f_0, which is also the same of those of f and g.

4.4 Discussion

4.4.1 Congested transport

As we saw in Section 4.2, Beckmann's problem can admit an easy variant if we prescribe a positive function $k : \Omega \to \mathbb{R}_+$, where $k(x)$ stands for the local cost at x per unit length of a path passing through x. This models the possibility that the metric is nonhomogeneous, due to geographical obstacles given a priori. Yet, it happens in many situations, in particular in urban traffic (as many readers probably note every day), that this metric k is indeed nonhomogeneous, but is not given a priori: it depends on the traffic itself. In Beckmann's language, we must look for a vector field $\mathbf{w}$ optimizing a transport cost depending on $\mathbf{w}$ itself!

The easiest model, chosen by Beckmann (see [27] and later [110]), is to consider the same framework as (BP) but supposing that $k(x) = g(|\mathbf{w}(x)|)$ is a function of the modulus of the vector field $\mathbf{w}$. This is quite formal for the moment (for instance, it is not meaningful if $\mathbf{w}$ is a measure, but we will see that in this case we can use better spaces than $\mathscr{M}^d_{\rm div}$). In this case we would like to solve

$$\min \left\{ \int \mathscr{H}(\mathbf{w}(x))\, \mathrm{d}x \quad : \quad \nabla \cdot \mathbf{w} = \mu - \nu \right\}, \tag{4.13}$$

where $\mathscr{H}(z) = H(|z|)$ and $H(t) = g(t)t$. Note that if H is superlinear (if $g(t)$ tends to ∞ as $t \to \infty$, i.e., if the congestion effect becomes stronger and stronger when the traffic increases), this problem is well-posed in the class of vector fields $\mathbf{w} \in L^1$. For instance, if $g(t) = t$, which is the easiest case one can imagine, we must minimize the L^2 norm under divergence constraints:

$$\min \left\{ \int |\mathbf{w}(x)|^2\, \mathrm{d}x \quad : \ \mathbf{w} \in L^2(\Omega; \mathbb{R}^d),\ \nabla \cdot \mathbf{w} = \mu - \nu \right\}.$$

This problem is easily solvable since one can see that the optimal $\mathbf{w}$ must be a gradient (we will develop this computation in a more general framework later), and

setting $\mathbf{w} = \nabla u$ one gets $\Delta u = \mu - \nu$. This is complemented with Neumann boundary conditions and allows to find u and then $\mathbf{w}$.

We want now to discuss the meaning and pertinence of this model, keeping into account the following natural questions:

- is it correct that the coefficient k depends on $|\mathbf{w}|$, or should it rather depend on other traffic quantities, such as i_Q? (note that $\mathbf{w}$ can have cancellations);
- what is the connection with equilibrium issues? In traffic congestion, typically every agent decides alone which path to choose, and the final traffic intensity is rather the output of a collection of individual choices and not the result of a global optimization made by a single planner;
- is the example $g(t) = t$ a good choice in the modeling ? This implies $g(0) = 0$, i.e., no cost where there is not traffic.

To start our analysis, we would like to present first an equilibrium model developed by Wardrop, [295], on a discrete network.

Traffic equilibria on a finite network

The main data of the model are a finite-oriented connected graph $G = (N, E)$ modeling the network, and some travel times functions g_e: $\theta \in \mathbb{R}_+ \mapsto g_e(\theta)$ giving, for each edge $e \in E$, the travel time on e when the flow on e is θ. The functions g_e are all nonnegative, continuous, and nondecreasing, and they are meant to capture the congestion effects (which may be different on the different edges, since some roads may be longer or wider and may have different responses to congestion). The last ingredient of the problem is a transport plan on pairs of nodes $(x, y) \in N^2$ interpreted as pairs of origins/destinations. We denote by $(\gamma_{x,y})_{(x,y)\in N^2}$ this transport plan: $\gamma_{x,y}$ represents the "mass" to be sent from x to y. We denote by $C_{x,y}$ the set of simple paths on the graph G connecting x to y, so that $C := \cup_{(x,y)\in N^2} C_{x,y}$ is the set of all simple paths. A generic path will be denoted by ω and we will use the notation $e \in \omega$ to indicate that the path ω uses the edge e.

The unknown of the problem is the flow configuration. The edge flows are denoted by $i = (i_e)_{e\in E}$ and the path flows are denoted by $q = (q_\omega)_{\omega\in C}$: this means that i_e is the total flow on edge e and q_ω is the mass traveling on ω. Of course the values i_e and q_ω are nonnegative and constrained by the mass conservation conditions:

$$\gamma_{x,y} = \sum_{\omega\in C_{x,y}} q_\omega, \quad \forall (x, y) \in N^2 \tag{4.14}$$

and

$$i_e = \sum_{\omega\in C\,:\, e\in\omega} q_\omega, \quad \forall e \in E, \tag{4.15}$$

which means that i is a function of q. Given the edge flows $i = (i_e)_{e\in E}$, the total cost (to be interpreted as a sort of weighted length) of the path $\omega \in C$ is

$$L_i(\omega) = \sum_{e\in\omega} g_e(i_e). \tag{4.16}$$

In [295], Wardrop defined a notion of noncooperative equilibrium that has been very popular since then among engineers working in the field of congested transport and that may be described as follows. Roughly speaking, a Wardrop equilibrium is a flow configuration such that every actually used path should be an optimal path for the cost L_i defined in (4.16). This leads to

Definition 4.27. A Wardrop equlibrium[5] is a flow configuration $q = (q_\omega)_{\omega\in C}$ satisfying $q_\omega \geq 0$ and the mass conservation constraints (4.14), such that when we compute the values i_e with (4.15), for every $(x, y) \in N^2$ and every $\omega \in C_{x,y}$ with $q_\omega > 0$, we have

$$L_i(\omega) = \min_{\omega'\in C_{x,y}} L_i(\omega').$$

A few years after the introduction of this notion by Wardrop, Beckmann, McGuire, and Winsten [29] realized that Wardrop equilibria can be characterized by the following variational principle:

Theorem 4.28. *The flow configuration $q = (q_\omega)_{\omega\in C}$ is a Wardrop equilibrium if and only if it solves the convex minimization problem*

$$\min\left\{\sum_{e\in E} H_e(i_e) \ : \ q \geq 0 \text{ satisfies (4.14)}\right\} \tag{4.17}$$

where, for each e, we define H_e to be an antiderivative of g_e: $H_e' = g_e$.

Proof. Assume that $q = (q_\omega)_{\omega\in C}$ (with associated edge flows $(i_e)_{e\in E}$) is optimal for (4.17). Then, for every admissible $\eta = (\eta_\omega)_{\omega\in C}$ with associated (through (4.15)) edge-flows $(j_e)_{e\in E}$, one has

$$\begin{aligned} 0 \leq \sum_{e\in E} H_e'(i_e)(j_e - i_e) &= \sum_{e\in E} g_e(i_e) \sum_{\omega\in C\,:\, e\in\omega} (\eta_\omega - q_\omega) \\ &= \sum_{\omega\in C} (\eta_\omega - q_\omega) \sum_{e\in\omega} g_e(i_e), \end{aligned}$$

[5]Note that this is just an example of Nash equilibrium with a continuum of players, as we will see in Section 7.4.3 and, more particularly, in Box 7.3.

so that

$$\sum_{\omega\in C} q_\omega L_i(\omega) \le \sum_{\omega\in C} \eta_\omega L_i(\omega). \tag{4.18}$$

Minimizing the right-hand side thus yields

$$\sum_{(x,y)\in N^2} \sum_{\omega\in C_{x,y}} q_\omega L_i(\omega) = \sum_{(x,y)\in N^2} \gamma_{x,y} \min_{\omega'\in C_{x,y}} L_i(\omega'),$$

which exactly says that q is a Wardrop equilibrium. To prove the converse, it is enough to see that problem (4.17) is convex so that the inequality (4.18) is indeed sufficient, and not only necessary, for a global minimum. □

The previous characterization allows to deduce for free existence results but also uniqueness for i (not for q) as soon as the functions g_e are strictly increasing (so that the H_e are strictly convex).

Remark 4.29. It would be very tempting to deduce from Theorem 4.28 that equilibria are efficient since they are minimizers of (4.17). Yet, the quantity $\sum_{e\in E} H_e(i_e)$ does not represent the natural total social cost measured by the total time lost in commuting which would rather be

$$\sum_{e\in E} i_e g_e(i_e). \tag{4.19}$$

In general, $\theta g_e(\theta) = H_e(\theta)$ and $g_e = H_e'$ are very different conditions. The efficient transport patterns are minimizers of (4.19) and thus are different from equilibria in general. Efficient and equilibria configurations coincide in the special case of power functions where $H_e(\theta) = a_e\theta^p$. Yet, we must note that this case is not realistic since, for $p > 1$, it implies $g_e(0) = 0$. This means that traveling times vanish if there is no traffic, while one should expect a residual nonzero cost even without traffic. Moreover, a famous counterexample due to Braess (see [76]) shows that it may be the case that adding an extra road with very small cost on which the traveling time is always zero leads to an equilibrium where the total commuting time is increased! This illustrates the striking difference between efficiency and equilibrium, a topic which is very well documented in the finite-dimensional network setting where it is frequently associated with the literature on the so-called price of anarchy (see [266]).

Remark 4.30. In the problem presented in this paragraph, the transport plan γ is fixed. This may be interpreted as a *short-term problem*. Instead, we could consider the *long-term* problem where only the distribution of origins μ and of destinations ν is fixed. In this case, one also obtains as an optimality condition that γ is efficient, in the sense that it minimizes among transport plans in $\Pi(\mu,\nu)$, the total cost

$$\sum \gamma_{x,y} d_i(x,y) \text{ with } d_i(x,y) := \min_{\omega \in C_{x,y}} L_i(\omega).$$

In the long-term problem where one is allowed to change the assignment as well, equilibria still are characterized by a convex minimization problem, i.e., the same above, where one also optimizes over γ.

Optimization and equilibrium in a continuous framework

If we want to generalize the previous analysis to a continuous framework[6], we will formulate the whole path-dependent transport pattern in terms of a probability measure Q on the set of paths (this is the continuous analogue of the path flows $(q_\omega)_\omega$ of the previous paragraph). Then, we will measure the traffic intensity generated by Q at every point x by using the measure i_Q of Section 4.2.3 (this is the continuous analogue of the arc flows $(i_e)_e$ of the previous paragraph). The last and main idea will be to model the congestion effects through a metric that is monotone increasing in the traffic intensity (the analogue of $g_e(i_e)$).

Avoiding technicalities and choosing an informal presentation (see [112] for more details), we take Ω a domain in $\mathbb{R}^d$ (typically, $d = 2$) and

- either two probability measures μ and ν (distribution of sources and destinations) on Ω in the case of the long-term problem,
- or a transport plan γ (joint distribution of sources and destinations, i.e., a joint probability on $\Omega \times \Omega$) in the short-term case,
- or more generally a convex and closed subset $\Gamma \subset \Pi(\mu, \nu)$, and we accept any $\gamma \in \Gamma$ (this is just a common mathematical framework for the two previous cases, where we can take $\Gamma = \{\gamma\}$ or $\Gamma = \Pi(\mu, \nu)$).

We will use the notations of Section 4.2.3 and use probability measures Q on $\mathscr{C} := \mathrm{AC}(\Omega)$, compatible with mass conservation, i.e., such that $(e_0, e_1)_\# Q \in \Gamma$. We shall denote by $\mathscr{Q}(\Gamma)$ the set of admissible Q.

The traffic intensity associated with $Q \in \mathscr{Q}(\Gamma)$ is by definition the measure i_Q defined in Section 4.2.3 and congestion effects are captured by the *metric* associated with Q: suppose $i_Q \ll \mathscr{L}^d$ and set

$$k_Q(x) := g(x, i_Q(x))$$

for a given increasing function $g(x, .)$: $\mathbb{R}_+ \to \mathbb{R}_+$. The fact that there exists at least one $Q \in \mathscr{Q}(\Gamma)$ such that $i_Q \ll \mathscr{L}^d$ is not always true and depends on Γ. We will see later how to guarantee it[7].

In order to define the notion of Wardrop equilibrium in this continuous case, let us define

[6] For a survey on the continuous framework, see also [111].

[7] Note that, even for $\Gamma = \{\gamma\}$ (which is the most restrictive case), assuming $\mu, \nu \in L^\infty$, considerations from incompressible fluid mechanics in [83] allow to build a Q such that $i_Q \in L^\infty$.

$$L_{k_Q}(\omega) = \int_0^1 g(\omega(t), i_Q(\omega(t)))|\omega'(t)|\mathrm{d}t$$

and

$$d_{k_Q}(x,y) := \inf_{\omega\in\mathscr{C},\omega(0)=x,\omega(1)=y} L_{k_Q}(\omega).$$

Paths in $\mathscr{C}$ such that $d_{k_Q}(\omega(0),\omega(1)) = L_{k_Q}(\omega)$ are called geodesics (for the metric induced by the congestion effect generated by Q).

We can define

Definition 4.31. A Wardrop equilibrium is a $Q \in \mathscr{Q}(\Gamma)$ such that

$$Q(\{\omega \ : \ L_{k_Q}(\omega) = d_{k_Q}(\omega(0),\omega(1))\}) = 1. \tag{4.20}$$

Existence, and even well-posedness (what does it mean $L_k(\omega)$ if k is only measurable and ω is a Lipschitz curve?) of these equilibria are not straightforward. However, it is possible to characterize equilibria as solutions of a minimal traffic problem.

The full proof is quite involved since it requires to take care of some regularity issues in detail. In particular, the use of the weighted length functional $L_{\bar{k}}$ and thus also the geodesic distance $d_{\bar{k}}$ require some attention when $\bar{k}$ is not continuous or at least l.s.c. In [112] a possible construction when $\bar{k}$ is just L^q is given.

Problem 4.32. Let us consider the (convex) variational problem

$$\text{(WP)} \qquad \min\left\{\int_\Omega H(x, i_Q(x))\,\mathrm{d}x \ : \ Q \in \mathscr{Q}(\Gamma)\right\} \tag{4.21}$$

where $H'(x,.) = g(x,.), H(x,0) = 0$.

Under some technical assumptions that we do not reproduce here, the main result of [112] is

Theorem 4.33. *(WP) admits at least one minimizer. Moreover, $\overline{Q} \in \mathscr{Q}(\Gamma)$ solves (WP) if and only if it is a Wardrop equilibrium and $\gamma_Q := (e_0,e_1)_\# Q$ solves the optimization problem*

$$\min\left\{\int_{\Omega\times\Omega} d_{k_Q}(x,y)\ \mathrm{d}\gamma(x,y) \ : \ \gamma \in \Gamma\right\}.$$

In particular, if Γ is a singleton, this last condition does not play any role (there is only one competitor) and we have the existence of a Wardrop equilibrium corresponding to any given transport plan γ. If, on the contrary, $\Gamma = \Pi(\mu,\nu)$, then the second condition means that γ solves a Monge-Kantorovich problem for a distance cost depending on Q itself, which is a new equilibrium condition.

As in the finite-dimensional network case, Wardrop equilibria have a variational characterization which is in principle easier to deal with than the definition. Unfortunately, the convex problem (WP) may be difficult to solve since it involves measures on sets of curves: it means two layers of infinite dimensions!

We finish this section by proposing two recent developments, with the aim of transforming the problem into simpler ones. The first development, easier to present in the short-term case, gives a dual formulation and provides interesting numerical methods. The second, only in the long-term case, gets rid of the formulation in terms of measures over curves and transforms the problem into a minimal flow problem in the spirit of Beckmann.

Duality and numerics for the short-term problem

We want here to give a tractable formulation of the variational problem for the short-term version of (WP). For every $x \in \Omega$ and $k \geq 0$, let us define $H^*(x, \cdot)$ as the Legendre transform of $H(x, \cdot)$. Let us now define the functional

$$J(k) = \int_\Omega H^*(x, k(x))\, \mathrm{d}x - \int_{\Omega\times\Omega} d_k(x, y)\, \mathrm{d}\gamma(x, y) \tag{4.22}$$

where, again, d_k is the geodesic distance associated to the metric k. Consider

$$\sup \{-J(k) : k \geq 0\}\,. \tag{4.23}$$

The above maximization problem is important because of the following duality result.

Proposition 4.34.

$$\min\ (4.21) = \max\ (4.23) \tag{4.24}$$

and k solves (4.23) if and only if $k = k_Q$ for some $Q \in \mathcal{Q}(\gamma)$ solving (WP).

Moreover, the optimization problem in k is "dimensionally smaller" than that on Q (in the purely continuous case, we optimize among scalar functions instead of measures over curves, and if we discretize in space and compute the dimensions, we really have a huge difference between the two formulations[8]).

Remark 4.35. Under reasonable assumptions, the dual problem (4.23) has a unique solution. The equilibrium metric k_Q and the equilibrium intensity of traffic i_Q are unique, although Wardrop equilibria Q might not be unique.

[8]Note that the same duality trick is also used in the discrete problem over networks, where solving the dual problem is much more convenient than the original one.

Starting from this duality result, in [42] and [43], a consistent numerical scheme to approximate the equilibrium metric k_Q was designed. The idea is to run a descent method on the dual. Hence, one needs to know how to compute (i.e., approximate) the values of d_k, which can be done in an efficient way by the *Fast Marching Algorithm*, but also to differentiate them according to k.

The functional J in (4.22) is considered by means of a discretization grid, and the values of k are considered as defined at the nodes of the grid. The first integral becomes a sum on all the points of the grid, while, for the second, one needs to replace the transport plan γ with a discretized one defined on pairs of points (x, y) on the same grid and to define $d_k(x, y)$ consequently.

To define such a distance $d_k(x_0, \cdot)$, for a fixed source x_0, as a function of the second variable, one uses the fact that it is the unique viscosity solution u of the Eikonal nonlinear PDE

$$\begin{cases} |\nabla u| = k, \\ u(x_0) = 0, \end{cases} \tag{4.25}$$

Box 4.5. Good to know!: *Discretization of the Eikonal equation and FMM*

To approximate the geodesic distance $u = d_k(x_0, \cdot)$, we discretize both the (unknown) values of u and of k on a square lattice with n points per side. Then we define

$$D_1 u^{i,j} := \max\{(u^{ij} - u^{i-1,j}), (u^{ij} - u^{i+1,j}), 0\}/h,$$
$$D_2 u^{i,j} := \max\{(u^{ij} - u^{i,j-1}), (u^{ij} - u^{i,j+1}), 0\}/h.$$

As proposed by Rouy and Tourin [267], the discrete geodesic distance $u = (u^{ij})_{i,j}$ is found as the solution of the following system that discretizes (4.25):

$$Du = k \quad \text{where} \quad Du^{i,j} = \sqrt{(D_1 u^{i,j})^2 + (D_2 u^{i,j})^2}. \tag{4.26}$$

Rouy and Tourin [267] showed that this discrete geodesic distance u converges to u_k when h tends to 0. Then, we have to choose a clever way to solve the above nonlinear system. One of the most efficient ways is the so-called Fast Marching Algorithm (see [283]): at every step of the algorithm, the values $u^{i,j}$ of a bunch of points of the grid have already been computed and the corresponding points are colored in red, and all the other values are set to $+\infty$; then, a tentative computation for all the neighbors of the red point is done, using Equation (4.26) for each of them separately (including the $+\infty$ values of the other points); then, only the smallest among the tentative values is validated, and the corresponding point is colored in red, and the algorithm goes on. This solves the system (4.26) in $O(N \log N)$ steps, where $N = n^d$ is the number of points in the grid.

Once we are able to compute the value of $J(k)$ for every discrete metric k on the grid, we want to differentiate it w.r.t. k, so as to take advantage of a gradient descent algorithm. Actually, one can see that J is not always differentiable in k, but since all the terms $c_k(x, y)$ may be proven to be concave in k, we face a convex optimization problem, and we can look at the subdifferential of J. Differentiating the equations

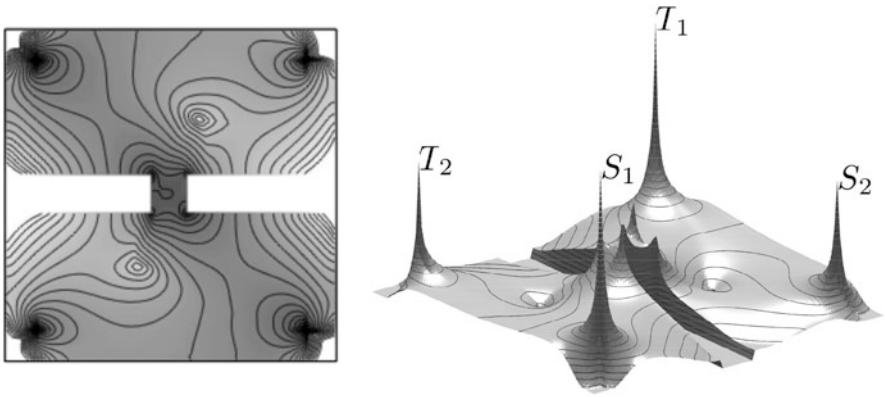


Fig. 4.4 Traffic intensity at equilibrium in a city with a river and a bridge, with two sources S_1 and S_2 and two targets T_1 and T_2. Picture taken from [43] with permission

in (4.26) (see [43]), one gets a new set of equations on the gradient $\nabla_k c_k(x, y)$. The same loop of the Fast Marching Algorithm allows to solve them in a quite efficient way, thus giving an element of the subdifferential[9]

An example of Wardrop equilibria computed with this kind of approach is given in the following Figure 4.4.

Beckmann reformulation of the long-term problem

In the long-term problem (4.21), we have one more degree of freedom since the transport plan is not fixed. This will enable us to reformulate the problem as a variational divergence constrained problem *à la* Beckmann and ultimately to reduce the equilibrium problem to solving some nonlinear PDE[10].

As we already did in Section 4.2.3, for any $Q \in \mathscr{Q}(\Gamma)$, we can take the vector field $\mathbf{w}_Q$. If we consider the scalar problem (WP), it is easy to see that its value is larger than that of the minimal flow problem à la Beckmann:

$$\text{(BP} - \text{cong)} \qquad \min \left\{ \int_\Omega \mathscr{H}(\mathbf{w}(x))\, \mathrm{d}x \,:\, \nabla \cdot \mathbf{w} = \mu - \nu \right\}, \tag{4.27}$$

where $\mathscr{H}(\mathbf{w}) = H(|\mathbf{w}|)$ and H is taken independent of x only for simplicity. The inequality is justified by two facts: minimizing over all vector fields $\mathbf{w}$ with

[9]This procedure is just a particular case of what is usually called *forward automatic differentiation*.

[10]We also observe that this reduction to a divergence-constrained convex minimization problem allows to provide alternative numerical approaches, as it is done in [35], in the same spirit of the Benamou-Brenier Augmented Lagrangian technique; see also Section 6.1

prescribed divergence gives a smaller result than minimizing over the vector fields of the form $\mathbf{w}_Q$, and then we use $|\mathbf{w}_Q| \le i_Q$ and the fact that H is increasing.

Actually, the two problems are equivalent. The main ingredient to prove it is Theorem 4.10. If we take a minimizer $\mathbf{w}$ for this minimal flow problem, then we are able to build a measure Q, and as we did in Theorem 4.13, the optimality of $\mathbf{w}$ gives $\mathbf{w} = \mathbf{w}_Q$ and $|\mathbf{w}_Q| = i_Q$, thus proving that the minimal values are the same. Also, we can build a minimizer $\mathbf{w}$ from a minimizer Q (just take $\mathbf{w} = \mathbf{w}_Q$) and conversely a minimizer Q from $\mathbf{w}$ (use Theorem 4.10).

Moreover, we also know from the proof of Theorem 4.10 that the main tool to construct the measure Q is the Dacorogna-Moser flow. If we want to use it to define a unique and canonical flow connecting the two measures μ and ν, then we would like to have some regularity of the vector field $\mathbf{v}$ defined as $\mathbf{v}_t(x) = \mathbf{w}(x)/f_t(x)$, where $f_t = (1-t)\mu + t\nu$. If $\mathbf{v}$ was Lipschitz continuous, then the Cauchy problem

$$\begin{cases} y_x'(t) = \mathbf{v}_t(y_x(t)) \\ y_x(0) = x \end{cases}$$

would have a unique solution for every initial datum x and would allow to find the measure Q (which would be unique, by the way). Obviously, we can decide to add some assumptions on μ and ν, which will be supposed to be absolutely continuous with regular densities (at least Lipschitz continuous and bounded from below).

However, one needs to prove regularity for the optimal $\mathbf{w}$ and for this one needs to look at the optimality conditions satisfied by $\mathbf{w}$ as a minimizer of (4.27). By considering perturbations of the form $\mathbf{w} + \varepsilon\tilde{\mathbf{w}}$ with $\nabla \cdot \tilde{\mathbf{w}} = 0$, one gets that the optimal $\mathbf{w}$ must satisfy $\int \nabla\mathscr{H}(\mathbf{w}) \cdot \tilde{\mathbf{w}} = 0$ for any divergence-free vector field $\tilde{\mathbf{w}}$. This means that $\nabla\mathscr{H}(\mathbf{w})$ is a gradient (see Box 6.2 in Section 6.2), which has to be adapted to an L^p setting[11]. We can call it ∇u, and by inverting $\nabla\mathscr{H}$, we get that the minimizers of (4.27) satisfy $\mathbf{w} = \nabla\mathscr{H}^*(\nabla u)$ where $\mathscr{H}^*$ is the Legendre transform of $\mathscr{H}$ and u solves the PDE:

$$\begin{cases} \nabla \cdot (\nabla\mathscr{H}^*(\nabla u)) = \mu - \nu, & \text{in } \Omega, \\ \nabla\mathscr{H}^*(\nabla u) \cdot \mathbf{n} = 0, & \text{on } \partial\Omega. \end{cases} \tag{4.28}$$

This equation turns out to be a standard Laplace equation if $\mathscr{H}$ is quadratic, or it becomes a p-Laplace equation for other power functions. In these cases, regularity results are well known, under regularity assumptions on μ and ν.

Remark 4.36. Note that the above considerations allow to answer the question "Given μ, ν, does a traffic plan Q with $i_Q \in L^p$ exist?" The answer is yes if and only if $\mu - \nu \in (W^{1,q})^*$, and the proof is left as an exercise; see **Ex**(26).

As we already mentioned, the case of power-like functions H is not suitable for modeling reasons. Recall that $H' = g$ where g is the congestion function, so it is

[11] Observe that the same result can also be directly obtained from duality arguments, as it is done in Theorem 2.1 in [80].

natural to have $g(0) > 0$: the metric should be positive even if there is no traffic! This means that the radial function $\mathscr{H}$ cannot be differentiable at 0 and then its subdifferential at 0 contains a ball. By duality, this implies $\nabla \mathscr{H}^* = 0$ on this ball which makes (4.28) very degenerate, even worse than the p-Laplacian. For instance, a reasonable model of congestion is $g(t) = 1 + t^{p-1}$ for $t \geq 0$, with $p > 1$, so that

$$\mathscr{H}(\mathbf{w}) = \frac{1}{p}|\mathbf{w}|^p + |\mathbf{w}|, \ \mathscr{H}^*(z) = \frac{1}{q}(|z| - 1)_+^q, \text{ with } q = p' = \frac{p}{p-1} \tag{4.29}$$

so that the optimal $\mathbf{w}$ is

$$\mathbf{w} = \Big(|\nabla u| - 1\Big)_+^{q-1} \frac{\nabla u}{|\nabla u|},$$

where u solves the very degenerate PDE:

$$\nabla \cdot \left(\Big(|\nabla u| - 1\Big)_+^{q-1} \frac{\nabla u}{|\nabla u|}\right) = \mu - \nu, \tag{4.30}$$

with Neumann boundary condition

$$\Big(|\nabla u| - 1\Big)_+^{q-1} \frac{\nabla u}{|\nabla u|} \cdot \mathbf{n} = 0.$$

Note that there is no uniqueness for u, but there is for $\mathbf{w}$.

For this degenerate equation, we cannot expect regularity results which go beyond Lipschitz continuity for u (indeed, since the diffusion coefficient identically vanishes in the zone where $|\nabla u| \leq 1$, every $u \in \mathrm{Lip}_1$ is the solution of the corresponding homogeneous equation). On the other hand, we can expect better regularity for $\mathbf{w}$ (roughly speaking, $\mathbf{w}$ is obtained by gluing two zones: the one where $\mathbf{w} = 0$ and the one where the equation is elliptic). However, we will not be able to obtain $\mathbf{w}$ itself to be Lipschitz. On the other hand, Sobolev regularity of $\mathbf{w}$ and Lipschitz regularity results for u are proven in [80]. This enables one to build a flow *à la* DiPerna-Lions [156] and then to justify rigorously the existence of a unique and well-defined flow for a.e. initial datum x. Interestingly, recent continuity results are also available (see [279] in dimension 2, and then [126], with a different technique in arbitrary dimension), obtained as a consequence of a fine analysis of this degenerate elliptic PDE. Besides the interest for this regularity result in itself, we also stress that continuity for $\mathbf{w}$ implies continuity for the optimal i_Q, and this exactly gives the regularity which is required in the proof of Theorem 4.33 (the main difficulty being defining $c_{\bar{k}}$ for a noncontinuous $\bar{k}$, and this is the reason why our previous arguments are only formal).

4.4.2 Branched transport

The goal of this second part of the discussion section is to present an opposite minimal flow problem, where, contrary to what happens in traffic congestion, joint transportation on a common path is indeed encouraged. This is reasonable when we consider "economy of scale" principles, and it is also something that we can observe in many different phenomena. In many cases, when a single planner needs to organize a transport (building a transportation network, for instance), the total cost happens to be *subadditive* (the cost for moving two objects cannot exceed the sum of the two costs), and in many cases it has *decreasing marginal costs* (i.e., the cost for adding a unit to a given background quantity is a decreasing function of the background, which means that the cost is actually concave).

Note that we are considering here the transport cost as a function of the mass and not of the length. The silent assumption is that the cost will be linear in the length (as it happens in all minimal flow problems of this chapter). But it will be concave in the mass, opposite to what we saw in Section 3.3.2.

Many branching structures transporting different fluids, such as road systems, communication networks, river basins, blood vessels, leaves, trees, and so on, may be easily thought of as coming from a variational principle for a cost of this kind.

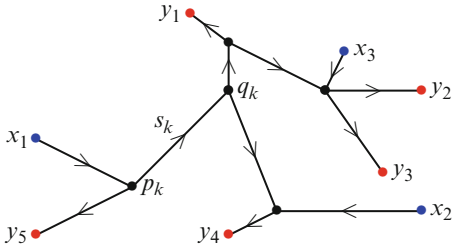
Note that, exactly as it happened for traffic congestion, modeling this kind of effects requires, either in Lagrangian or Eulerian language, to look at the paths actually followed by each particle. It could not be done with the only use of a transport plan $\gamma \in \Pi(\mu, \nu)$. Yet, once we choose the good formulation via the tools developed in this chapter, we can guess the shape of the optimal solution for this kind of problem: particles are collected at some points, move together as much as possible, and then branch toward their different destinations. This is why this class of problems is nowadays known as "branched transport."

Recently these problems received a lot of attention by mathematicians, but in fact the formalization of the discrete framework, which can be expressed in terms of Dirac masses, is very classical in optimization and operational research. As we did for congested transport, we start from the discrete framework, which was introduced by Gilbert [185]. We then move on to the continuous models, which are much more recent and were partly developed and studied by Bernot, Caselles, and Morel [46, 48–50] and Maddalena and Solimini [220, 221] for the *Lagrangian* model and by Xia [297], who gave an *Eulerian* continuous extension of the Gilbert Energy.

The discrete model

Translating into our language the work by Gilbert [185], we consider two atomic probability measures $\mu = \sum_{i=1}^{m} a_i \delta_{x_i}$ and $\nu = \sum_{j=1}^{n} b_j \delta_{y_j}$ supported in some compact subset $\Omega \subset \mathbb{R}^d$, and we want to connect them through a weighted oriented graph G. A *weighted oriented graph* in $\mathbb{R}^d$ is just a system $G = (s_k, \theta_k)_k$ where s_k is a pair of points $s_k = (p_k, q_k)$ representing an oriented segment in $\mathbb{R}^d$ and $\theta_k \in \mathbb{R}_+$ is

a weight associated with the segment s_k. We denote by $\hat{s}_k$ the orientation $\frac{q_k - p_k}{|q_k - p_k|}$ of s_k (provided it is not trivial) and $|s_k|$ its length $|q_k - p_k|$. The points p_k, q_k are called the *vertices* of the graph and the segments s_k are its *edges*.



Given μ, ν, we say that G irrigates ν from μ if it satisfies the well-known Kirchhoff law for electric circuits: for each vertex x of the graph, one must have

$$\text{incoming mass at } x = \text{outcoming mass at } x.$$

We count as *incoming* the mass conveyed by the segments $s_k = (p_k, q_k)$ reaching x, i.e., such that $x = q_k$, increased by the mass given by μ, i.e., increased by a_i if $x = x_i$. Similarly, the *outcoming mass* at x is the total weight of the segments $s_k = (p_k, q_k)$ leaving x, i.e., such that $p_k = x$, decreased by the mass taken out by ν, i.e., by b_j if $x = y_j$. The set of graphs irrigating ν from μ is denoted by $\mathrm{G}(\mu, \nu)$.

For $\alpha \in [0, 1]$, we define the cost of α-irrigation of a graph G, also called *Gilbert energy* as

$$\mathscr{E}_\alpha(G) = \sum_{k=1}^{n} \theta_k^\alpha \, |s_k| \,,$$

which means that the cost of moving a mass m along a segment of length l is $m^\alpha l$.

Given μ, ν atomic probabilities, we want to minimize this energy among all graphs sending μ to ν, which reads

$$\min \quad \{\mathscr{E}_\alpha(G) : G \in \mathrm{G}(\mu, \nu)\}.$$

This problem was introduced by Gilbert in [185, 186] as an extension of Steiner's minimal length problem which corresponds[12] to the case $\alpha = 0$, where one wants to minimize the total length of the network. The main applications that Gilbert referred to were in the field of communication networks.

Note that for $\alpha \in [0, 1]$ the function $t \mapsto t^\alpha$ is concave and subadditive; it is strictly concave for $\alpha \in]0, 1[$. In this way larger tubes bringing the mass from μ to ν are preferred to several smaller ones transporting the same total mass.

[12] To have an exact equivalence between the Steiner minimal connection problem and Gilbert problem with $\alpha = 0$, one needs to consider a measure μ composed of a unique Dirac mass, so that every point in $\mathrm{spt}(\nu)$ must be connected to it, hence getting connectedness of the graph.

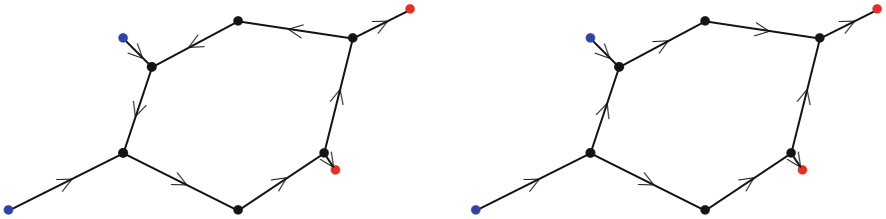


Fig. 4.5 An example of a strong cycle (left) and of a cycle which is not strong (right)

It is not evident to see why this problem admits a minimizer. Indeed, we minimize among finite objects, but we do not impose uniform bounds (on the number of edges, of vertices, etc.). We need to give an important definition.[13]

Definition 4.37. We call *cycle* (resp. *strong cycle*) of G any sequence of segments (resp. oriented segments) which are adjacent (Figure 4.5).

As the function $t \mapsto t^\alpha$ is increasing, we can easily reduce the energy of a graph by removing strong cycles. It is not difficult to check that the concavity of the same function also ensures that one reduces the energy by removing any cycle. Therefore one can take a minimizing sequence of graphs which have no cycles, i.e., which are trees, and this implies a uniform bound on the number of edges (in terms of the total number of atoms of μ and ν). This provides compactness and proves the existence of a minimizer. Any minimizer will have no strong cycles and no cycles if $\alpha \in]0,1[$.

Moreover, it is possible to prove an optimality condition on the free junctions in the graph. This means that at every node x of the graph which is not one of the points x_i or y_j, a balance of the direction must hold: the condition is

$$\sum_{k:x=p_k} \theta_k^\alpha \hat{s}_k - \sum_{k:x=q_k} \theta_k^\alpha \hat{s}_k = 0, \tag{4.31}$$

which means that the sum of the unit vectors corresponding to all the segments touching the point x, always taken as exiting from x independently of the true orientation $\hat{s}_k$, each weighted with θ_k^α, must vanish[14].

[13] Unfortunately, to cope with the language usually adopted in branched transport, we cannot be completely coherent with the rest of the chapter, where we called "cycles" what we call here "strong cycles."

[14] For triple junctions, which are the most common ones, this gives interesting results: when $\alpha = 0$, we get the well-known condition about Steiner trees, with three 120° angles, and for $\alpha = 0.5$, we have a 90° angle (see **Ex**(30)); yet, in general, the angles depend on the masses θ_k.

A relaxed continuous model (Xia)

More recently Xia, in [297], has proposed a new formalization leading to generalizations of this problem to arbitrary probability measures μ and ν. The main idea was to translate the Kirchhoff law constraint into a divergence constraint. To be precise, and since we will need this language later on, we introduce some heavy notation.

If $E \subset \Omega$ is a 1-rectifiable set[15], $\tau : E \to \mathbb{S}^{d-1}$ is a measurable vector field, and $\theta : E \to \mathbb{R}_+$ is $\mathscr{H}^1$-integrable, we define the finite vector measure $[E, \tau, \theta] \in \mathscr{M}^d(\Omega)$ through

$$\langle [E, \tau, \theta], \xi \rangle = \int_E \theta(x)\xi(x) \cdot \tau(x) \mathrm{d}\mathscr{H}^1(x),$$

for all $\xi \in C(\Omega; \mathbb{R}^d)$.

Given a finite graph $G = (s_k, \theta_k)_k$, we associate with G the measure

$$\mathbf{w}_G := \sum_k [s_k, \hat{s}_k, \theta_k] \in \mathscr{M}^d(\Omega).$$

We can express both the cost and the constraint in terms of $\mathbf{w}_G$ only. First note that $G \in \mathrm{G}(\mu, \nu)$ (i.e., it satisfies mass balance conditions for irrigating ν from μ) if and only if $\nabla \cdot \mathbf{w}_G = \mu - \nu$. Then, we also write

$$\mathscr{E}_\alpha(G) = \sum_{k=1}^n \theta_k^\alpha \, |s_k| = \int_K |\theta(x)|^\alpha \, \mathrm{d}\mathscr{H}^1(x),$$

where we called θ the density of $\mathbf{w}_G$ w.r.t. the $\mathscr{H}^1$ measure.

This suggests a continuous Eulerian formulation of Gilbert's problem by generalizing Gilbert's energy, the minimization problem becoming

$$\text{(XP)} \quad \min \quad \{M_\alpha(\mathbf{w}) : \mathbf{w} \in \mathscr{M}^d_{\mathrm{div}}(\Omega), \, \nabla \cdot \mathbf{w} = \mu - \nu\}. \tag{4.32}$$

Obviously, one needs to define the energy M_α (which is also called the "α-mass" of $\mathbf{w}$). The original idea by Xia is to define it by relaxation, i.e.,

$$M_\alpha(\mathbf{w}) := \inf \left\{ \liminf_n \mathscr{E}_\alpha(G_n) : G_n \in \mathrm{G}(\mu_n, \nu_n), \mathbf{w}_{G_n} \rightharpoonup \mathbf{w} \ \text{ in } \mathscr{M}^d_{\mathrm{div}}(\Omega) \right\},$$

where the convergence in $\mathscr{M}^d_{\mathrm{div}}(\Omega)$ means $\mathbf{w}_{G_n} \rightharpoonup \mathbf{w}$ and $\nabla \cdot \mathbf{w}_{G_n} = \mu_n - \nu_n \rightharpoonup \nabla \cdot \mathbf{w}$ as measures.

[15] We already met rectifiable sets when studying the differentiability of convex functions, in Chapter 1: 1-rectifiable sets are defined as those sets which are covered, $\mathscr{H}^1$-a.e., by a countable union of Lipschitz curves. Anyway, the reader can easily pretend that "1-rectifiable" is a synonym of "1D" and it should be enough to follow the rest of the discussion.

As a consequence of a famous rectifiability theorem by B. White [296], one can prove that $M_\alpha(\mathbf{w}) < +\infty$ implies that $\mathbf{w}$ is 1-rectifiable,[16] and with a finer analysis, we get

$$M_\alpha(\mathbf{w}) = \begin{cases} \int_E |\theta(x)|^\alpha \, \mathrm{d}\mathscr{H}^1(x) & \text{if } \mathbf{w} = [E, \tau, \theta] \\ +\infty & \text{otherwise.} \end{cases}$$

It should be proven that when μ and ν are both atomic measures, we retrieve the problem by Gilbert (this is not trivial, as we admitted lots of new competitors).

Compared to the finitely atomic case, another difference is the finiteness of the minimum in (XP). Indeed, it is always possible to connect two atomic measures with a finite graph, and the cost of finite graphs is always finite. On the contrary, it is not clear for general measures μ, ν that there exists a $\mathbf{w}$ with $M_\alpha(\mathbf{w}) < \infty$ and $\nabla \cdot \mathbf{w} = \mu - \nu$.

Irrigability and irrigation distance

For $\mu, \nu \in \mathscr{P}(\Omega)$, we set $d_\alpha(\mu, \nu) := \min(\mathrm{XP}) \in [0, +\infty]$ (which obviously depends on μ, ν). First of all we consider the question whether this value is finite. Supposing $0 \in \Omega$, it is clear that if $d_\alpha(\mu, \delta_0), d_\alpha(\nu, \delta_0) < +\infty$, then $d_\alpha(\mu, \nu) < +\infty$. Hence[17] we will concentrate on the question whether $d_\alpha(\mu, \delta_0) < +\infty$.

One of the main results of [297] is the following:

Proposition 4.38. *If $\alpha > 1 - 1/d$ and $\Omega = [-a, a]^d$, then $d_\alpha(\mu, \delta_0) < +\infty$ for every $\mu \in \mathscr{P}(\Omega)$. This implies $d_\alpha(\mu, \nu) < +\infty$ for every pair (μ, ν). The same is true in every compact domain, just by fitting it into a large cube.*

Moreover, keeping the conditions $\alpha > 1 - 1/d$ and Ω compact, d_α is actually a distance over $\mathscr{P}(\Omega)$ and it metrizes the weak convergence of probabilities.

We just prove the very first part of this statement.

Proof. The idea to prove $d_\alpha(\mu, \delta_0) < +\infty$ is the following: take a dyadic approximation of μ by partitioning the cube Ω into 2^{nd} small cubes of edge 2^{-n} and putting a Dirac mass of the corresponding measure at the center of each small cube, thus getting an atomic measure μ_n. Then connect μ_n to μ_{n+1} through a finite graph G_n (to which we associate a flow $\mathbf{w}_n$), which sends every Dirac mass of μ_n, corresponding to some small cube D, to the 2^d Dirac masses of μ_{n+1} corresponding to the sub-cubes of D, along straight lines (Figure 4.6). A simple calculation gives the estimate:

[16] We define rectifiable vector measures as those which can be expressed in the form $[E, \tau, \theta]$ with E rectifiable. This language is borrowed from that of rectifiable currents, but currents can be considered vector measures. The reader can look at [161] for more details.

[17] Indeed, even if not completely evident, it can be proven that, at least in the case where $\mathrm{spt}(\mu) \cap \mathrm{spt}(\nu) = \emptyset$, the condition $d_\alpha(\mu, \nu) < +\infty$ is actually equivalent to $d_\alpha(\mu, \delta_0), d_\alpha(\nu, \delta_0) < +\infty$.

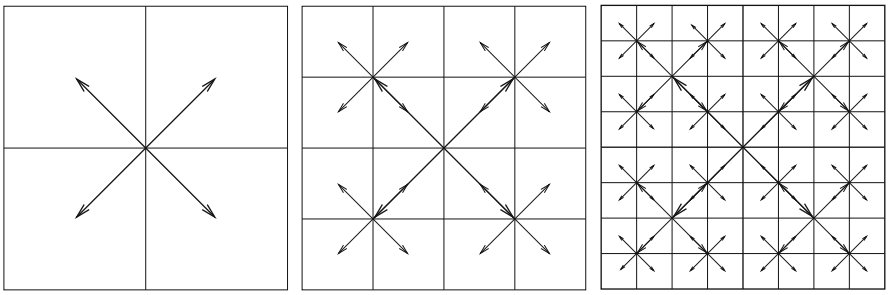


Fig. 4.6 Different steps of the construction of Proposition 4.38

$$M_\alpha(G_n) \le C \sum_{k=1}^{2^{nd}} a m_k^\alpha 2^{-n},$$

where the masses m_k are the masses contained in each of the cubes at the nth subdivision. Then we use $\sum_{k=1}^N m_k^\alpha \le N^{1-\alpha}(\sum_{k=1}^N m_k)^\alpha$ (which comes from Jensen's inequality), and get $\mathscr{E}_\alpha(G_n) \le aC2^{-n(1-d(1-\alpha))}$.

It is clear by construction that we only need $\sum_n \mathscr{E}_\alpha(G_n) < +\infty$, and this is true as soon as $1 - d(1-\alpha) > 0$. □

Some counterexamples that we do not develop here (but one can see **Ex**(29)) complement the previous result and show that the above statement is sharp in terms of α. More generally, we can think that the irrigability of a measure starting from δ_0 is a matter of its dimension, but suitable notions of dimensions (which are not simply the Hausdorff dimension of the support) have to be chosen. These matters are discussed in [153, 154].

It is interesting to note that the above distance d_α may also be obtained by relaxation (see [220]) from the functional which associates with every pair of atomic probability measures $\mu = \sum_i a_i \delta_{x_i}$, $\nu = \sum_j b_j \delta_{y_j}$ the quantity

$$\overline{d_\alpha}(\mu,\nu) := \min\left\{ \sum_{ij} |x_i - y_j| \gamma(\{(x_i,y_j)\})^\alpha \ : \ \gamma \in \Pi(\mu,\nu) \right\}.$$

This quantity, which corresponds to a sort of nonlinear version of the Kantorovich problem, is only well defined on atomic measures. Then, we can check that we have

$$d_\alpha = \sup\left\{ d : \mathscr{P}(\Omega)\times\mathscr{P}(\Omega) \to [0,+\infty] \ : \ \begin{cases} d \le \overline{d_\alpha} \text{ on atomic measures,} \\ d \text{ is l.s.c. for } \rightharpoonup, \\ d \text{ satisfies the triangle inequality.} \end{cases} \right\}. \tag{4.33}$$

The Lagrangian formulation

We want now to provide an alternative formulation of the branched transport problem, based on the use of measures on paths. This formulation[18] has been introduced in [221] and then deeply studied in [48].

As usual, we will be given two measures $\mu, \nu \in \mathscr{P}(\Omega)$ and look for a *traffic plan*[19], i.e., a measure $Q \in \mathscr{P}(\mathscr{C})$, where $\mathscr{C} := \mathrm{AC}(\Omega)$, with $\int_{\mathscr{C}} L(\omega)\,\mathrm{d}Q(\omega) < +\infty$. As usual, we will restrict our attention to those traffic plans satisfying $(e_0)_\# Q = \mu, (e_1)_\# Q = \nu$ (the case where we fix also the coupling $(e_0, e_1)_\# Q = \gamma$ will not be considered, and, by the way, the connection with the previous Eulerian model by Xia does not work if γ is fixed).

To introduce the quantity that we want to minimize, we need to define

$$\theta_Q(x) := Q(\{\omega \in \mathscr{C} \,:\, x \in \omega([0,1])\}), \quad m_Q(x) := \int_{\mathscr{C}} \#(\{t \in [0,1] \,:\, \omega(t) = x\})\,\mathrm{d}Q(\omega).$$

Proposition 4.39. *The following properties are satisfied:*

1. *We always have $0 \le \theta_Q(x) \le m_Q(x) \le 1$, with equality $\theta_Q = m_Q$ whenever Q is concentrated on loop-free curves.*
2. *The function $(x, Q) \mapsto \theta_Q(x)$ is u.s.c. w.r.t. the convergence of points $x \in \Omega$ and the weak convergence of measures $Q \in \mathscr{P}(\mathscr{C})$.*
3. *The traffic intensity i_Q is equal to $m_Q \cdot \mathscr{H}^1$ whenever there exists a set Γ, σ-finite for the $\mathscr{H}^1$ measure, with $\omega(t) \in \Gamma$ for Q-a.e. curve ω and a.e. time t.*

Then, we define the energy I_α through

$$I_\alpha(Q) := \int_{\mathscr{C}} Z_Q(\omega)\,\mathrm{d}Q(\omega),$$

where

$$Z_Q(\omega) := L_{\theta_Q^{\alpha-1}}(\omega) = \int_0^1 \theta_Q(\omega(t))^{\alpha-1} |\omega'(t)|\,\mathrm{d}t.$$

The problem that we consider is then

$$\min\{I_\alpha(Q) \,:\, Q \in \mathscr{P}(\mathscr{C}),\ (e_0)_\# Q = \mu, (e_1)_\# Q = \nu\}.$$

[18] What we provide here is just a translation into the language of this chapter of the model proposed in [48, 221], which uses "parametrized traffic plans" instead of measures on paths, but it is just a matter of language.

[19] The terminology has been introduced for this very purpose by the authors of [48].

Note that the trivial estimate $I_\alpha(Q) \geq \int_{\mathscr{C}} L(\omega)\,\mathrm{d}Q(\omega)$ provides compactness on a minimizing sequence (indeed, the energy is invariant under reparametrization of each curve, and hence length bounds turn into Lipschitz bounds, as we did in the rest of the chapter). The existence of a minimizer is then obtained by proving semi-continuity. Indeed, combining fact 2 with a simple Fatou Lemma, we obtain that the function $(Q, \omega) \mapsto Z_Q(\omega)$ is l.s.c. (for the weak convergence on Q and the uniform convergence on ω). Finally we use the following:

Proposition 4.40. *If $f : \mathscr{P}(\mathscr{C}) \times \mathscr{C} \to \mathbb{R}$ is l.s.c., then $F : Q \mapsto \int_{\mathscr{C}} f(Q, \omega)\,\mathrm{d}Q(\omega)$ is l.s.c. on $\mathscr{P}(\mathscr{C})$.*

Then, we can prove the following facts:

Proposition 4.41. *1. For every Q with $I_\alpha(Q) < +\infty$, the set $\Gamma = \{x : \theta_Q(x) > 0\}$ is σ-finite for the measure $\mathscr{H}^1$ and we have $I_\alpha(Q) = \int_\Gamma \theta_Q^{\alpha-1} m_Q\,\mathrm{d}\mathscr{H}^1$.*

2. If Q is optimal, then it is concentrated on loop-free curves and $I_\alpha(Q) = \int_\Gamma \theta_Q^\alpha\,\mathrm{d}\mathscr{H}^1$.

3. If Q is optimal, then it is cycle-free[20].

This property and the particular the expression of I_α as the integral of the α-power of the multiplicity θ_Q suggest the connection with the Eulerian model. A precise equivalence can be proven, either via the tools in [48] or via Theorem 4.10.

The single-source case

It is interesting to see the special features of the problem when we fix the source measure as a single Dirac mass δ_0. This particular case is already very rich. In this case, it is possible to study the problem with a different definition of multiplicity: indeed, we can expect that whenever two curves pass at a same point x, they had traveled together on all the path from 0 to x. In this case, it is better to change a little bit the framework, because parametrizing all the curves on a fixed finite-length integral $[0, 1]$ could be too restrictive. Also, this gives the occasion to observe the original framework presented in [221] (which can be checked to be equivalent to the general one).

We consider now the space $\Lambda = \{\omega \in \mathrm{Lip}_1(\mathbb{R}_+; \Omega),\ \omega(0) = 0\}$, endowed with the topology of uniform convergence on compact sets (recall that this topology is metrizable by some distance d and that this makes Λ a compact metric space thanks to the Ascoli-Arzelà theorem).

If $\omega \in \Lambda$, we define its *stopping time* by

$$\mathrm{Stop}(\omega) = \inf\{t \geq 0 : \omega \text{ is constant on } [t, +\infty[\}.$$

[20] We can give a precise definition in the following way: for $x, y \in \Omega$, define $[x, y] = \{\omega \in \mathscr{C} : \exists\, s < t$ such that $\omega(s) = x$ and $\omega(t) = y\}$; we say that Q is cycle-free if there are not $x_1, \ldots, x_n$ with $x_1 = x_n$ such that $Q([x_i, x_{i+1}]) > 0$ or $Q([x_{i+1}, x_i]) > 0$ for all $i = 1, \ldots, n-1$.

We denote by Λ_{stop} the set of curves ω such that $\text{Stop}(\omega) < \infty$. Then we define the following subsets of Λ_{stop} : Λ_{arc}, Λ_{inj}, Λ_{lf}, respectively, the set of curves which are parameterized by arc length on $[0, \text{Stop}(\omega)]$, which are injective on $[0, \text{Stop}(\omega)[$ and which are injective on $[0, \text{Stop}(\omega)]$ (these curves are called *loop-free*).

We call *traffic pattern* any $\eta \in \mathscr{P}(\Lambda)$ (we use the notation η to distinguish from the case $Q \in \mathscr{P}(\mathscr{C})$) satisfying

$$\int_{\Lambda} \text{Stop}(\omega)\, \mathrm{d}\eta(\omega) < +\infty.$$

We define the multiplicity and the energy according to the following variant of the previous definitions: for the multiplicity, we first define

$$[\omega]_t := \{\tilde{\omega} \in \Lambda_{\text{stop}} : \omega(s) = \tilde{\omega}(s) \quad \forall s \in [0, t]\},$$

the set of curves which coincide with ω up to time t, and for notational simplicity, we set $|\omega|_{t,\eta} := \eta([\omega]_t)$. This last quantity will play the role of the multiplicity θ. Then we define

$$\tilde{Z}_\eta(\omega) := \int_0^{\text{Stop}(\omega)} |\omega|_{t,\eta}^{\alpha-1}\, \mathrm{d}t$$

and we call η-good any curve ω with $\tilde{Z}_\eta(\omega) < +\infty$. Then, we use the energy

$$\tilde{I}_\alpha(\eta) := \int_{\Lambda} \tilde{Z}_\eta(\omega)\, \mathrm{d}\eta(\omega).$$

The problem becomes the minimization of $\tilde{I}_\alpha$ among probabilities $\eta \in \mathscr{P}(\Lambda)$ satisfying $(e_\infty)_\# \eta = \nu$, where $e_\infty(\omega) = \omega(\text{Stop}(\omega))$. Note that e_∞ is well defined on Λ_{stop}, but any finite energy traffic plan is necessarily concentrated on Λ_{stop}; hence the condition makes sense; moreover, $\eta \mapsto (e_\infty)_\# \eta$ is continuous for the weak convergence when restricted to any set $\{\eta \in \mathscr{P}(\Lambda) : \int_\Lambda \text{Stop}(\omega)\, \mathrm{d}\eta(\omega) \le C\}$.

It is not difficult to prove the existence of a minimizer in this case, and it is also possible to prove that minimizers are concentrated on $\Lambda_{\text{arc}} \cap \Lambda_{\text{inj}}$ and are even loop-free (i.e., we can prove $\eta(\Lambda_{\text{lf}}) = 1$, even if it is less evident). Note that we withdrew from the definition of the energy the term $|\omega'|$ and this forces the minimizers to saturate the constraint $|\omega'| \le 1$.

Remark 4.42. Other intermediate models may be introduced, all differing in the definition of the multiplicity of the curve ω at time t. See, for instance, [47] or [220].

The class of η-good curves is an explicitly defined set of curves which shares many of the properties satisfied η-almost everywhere. For instance, if η is optimal and Q is a measure on $\mathscr{C}$ representing η after a suitable reparametrization in time, we have $|\omega|_{t,\eta} = \theta_Q(\omega(t)) = m_Q(\omega(t))$, and this is true for every ω which is η-good.

An important point is the following: if ω_1, ω_2 are η-good and η is optimal, if $(e_\infty)(\omega_1) = (e_\infty)(\omega_2)$, then $\tilde{Z}_\eta(\omega_1) = \tilde{Z}_\eta(\omega_2)$ (see [272]). As a corollary, we get, for instance, that any η-good curve ω is in fact loop-free, i.e., it is injective on $[0, \mathrm{Stop}(\gamma)]$.

Definition 4.43 (Landscape function). We define the landscape function z_η associated with the traffic pattern η through

$$z_\eta(x) = \begin{cases} Z_\eta(\omega) & \text{if } \omega \text{ is an } \eta\text{-good curve and } x = e_\infty(\omega), \\ +\infty & \text{if there is no } \eta\text{-good curve ending at } x. \end{cases}$$

The introduction of the landscape function [272] is justified by several different considerations. On the one hand, it plays the role of the first variation of the functional

$$X_\alpha(\nu) = d_\alpha(\delta_0, \nu),$$

as we can prove (by concavity)

$$X_\alpha(\mu) \leq X_\alpha(\nu) + \alpha \int_\Omega z_\eta \, \mathrm{d}(\mu - \nu),$$

where η is any optimal traffic plan irrigating ν. Considering that we also have $X_\alpha(\nu) = \int z_\eta \, \mathrm{d}\nu$, it seems to play the same role as the Kantorovich potential plays for the classical Monge transport cost[21].

On the other hand, and this justifies its name, z_η is a natural object studied by geophysicists when dealing with river basins (see [263] for details and [272] for a brief mathematical introduction to this theory); in their model, after a first evolution phase, it is assumed that landscape erosion creates a stable configuration where the elevation $z(x)$ at each point is characterized by two properties: the river flows in the direction of maximal slope, and the slope $|\nabla z|$ is proportional to the discharge (i.e., multiplicity) coefficient θ, raised to a power close to $-1/2$ (say $\alpha - 1$, with α close to the 2D critical exponent $1/2$).

It is possible to prove few properties of the landscape function.

Theorem 4.44. *The landscape function z is lower semi-continuous. It is Hölder continuous*[22] *of exponent $\beta = d(\alpha - (1 - \frac{1}{d}))$ whenever $\alpha > 1 - \frac{1}{d}$, Ω has Lipschitz boundary and $\nu \geq c\mathscr{L}_\Omega$. Moreover, z has maximal slope in the direction of the irrigation network in the following sense: if ω is an η-good curve, $t_0 \leq \mathrm{Stop}(\omega)$, $x_0 = \omega(t_0)$, and $\theta_0 = |\omega|_{t_0,\eta}$, then for any $x \in \Omega$, we have*

$$z(x) \geq z(x_0) - \theta_0^{\alpha-1} \, |x - x_0| - o(|x - x_0|).$$

[21] But it does not seem to come from a dual problem.

[22] Other Holder results exist under different assumptions, if ν admits different dimensional lower bounds, i.e., it is Ahlfors regular of another exponent $k < d$, thus obtaining $\beta = d(\alpha - (1 - \frac{1}{k}))$; see [78].

Regularity

The regularity of the optimal networks in branched transport is a delicate issue, which has been studied both in Eulerian and Lagrangian framework. One of the first results concerns what Xia called "interior regularity": here interior means "far from the supports of μ and ν," which can be considered as the boundary of the transport network. These results, proven in [298] with a Eulerian approach and in [49] with the Lagrangian approach, state that, far from these supports, optimal networks are locally finite graphs. In particular, they satisfy the angle law (4.31) at every branching point. The question now becomes whether this angle law is satisfied at an infinitesimal level (in the sense that the directions to be considered are tangent lines and not segments) inside the support of μ and ν. [300] gives partial results in this direction (characterization of the limits up to subsequences), but a full solution is not yet available. On the other hand, in the simpler case of a single source, [241] provides the following result:

Theorem 4.45. *Let ω be an η-good curve, parametrized with constant speed, of an optimal traffic pattern irrigating a measure ν with $a\mathscr{L}_\Omega \le \nu \le b\mathscr{L}_\Omega$, for $0 < a < b < \infty$. Suppose that Ω is a Lipschitz bounded domain. Consider $t_0 < \mathrm{Stop}(\omega)$. Let θ_i, $i \in I$ the masses of all trees branching from ω in the interval $[0, t_0]$). Then $\sum_i \theta_i^\alpha < \infty$ and $\omega' \in BV([0, t_0])$. In particular, ω has two half-tangents at all points and a tangent at all points which are not branching points. Moreover, the tangent cone at every branching point x is composed of a finite (and bounded by a constant depending on α and d) number of segments whose directions and masses satisfy (4.31). More precisely, ω satisfies in the sense of distributions on $[0, t_0]$ the elliptic equation*

$$-\left(|\omega|_{t,\eta}^{\alpha}\,\omega'(t)\right)' = \sum_{i\in I} \theta_i^\alpha \delta_{\omega(t_i)} \hat{s}_i \tag{4.34}$$

where $\hat{s}_i$ is the tangent of the branch stemming from ω at the point $\omega(t_i)$ with mass θ_i. The fact that this tangent vector exists, and that the right-hand side is a vector measure with finite mass, is guaranteed by the first parts of the statement.

Counterexamples to the existence of tangent directions can be built when the assumption on ν fails (see, again, [241]). For an alternative approach to the existence of tangent directions under slightly different assumptions, one can also see the last chapter in [271].

Numerics

From the point of view of the practical search of the solutions, branched transport problems lie in between combinatorial optimization (they are similar to the Steiner minimal connection problem, which is, on a graph, known to be NP-hard) and continuous minimal flow problems. The numerical methods for finding a solution have to face at least two difficulties: concavity (and hence the high number of local

minimizers) and singularity (i.e., working with singular structures such as networks instead of using L^p or smoother functions, as in the congested traffic case).

Some numerical results, based on local and nonlocal improvements of a suitably chosen initial guess, are present in [297] and [299]. Yet, here we prefer to introduce a completely different idea, based on a Γ-convergence approach.

Box 4.6. Important notion: *Γ-convergence*

Γ-convergence is a theory, introduced by De Giorgi (see [138] and [133]), to provide approximations of minimization problems. It is also known as *epi-convergence*.

Definition. On a metric space X, let $F_n : X \to \mathbb{R} \cup \{+\infty\}$ be a sequence of functions. We define the two lower-semi-continuous functions F^- and F^+ (called $\Gamma - \liminf$ and $\Gamma - \limsup$ of this sequence, respectively) by

$$F^-(x) := \inf\{\liminf_{n\to\infty} F_n(x_n) \ : \ x_n \to x\},$$

$$F^+(x) := \inf\{\limsup_{n\to\infty} F_n(x_n) \ : \ x_n \to x\}.$$

Should F^- and F^+ coincide, then we say that F_n actually Γ-converges to the common value $F = F^- = F^+$. The definition of Γ-convergence for a continuous parameter $\varepsilon \to 0$ obviously passes through the convergence to the same limit for any subsequence $\varepsilon_n \to 0$.

Among the properties of Γ-convergence, we have the following:

- if there exists a compact set $K \subset X$ such that $\inf_X F_n = \inf_K F_n$ for any n, then F attains its minimum and $\inf F_n \to \min F$;
- if $(x_n)_n$ is a sequence of minimizers for F_n admitting a subsequence converging to x, then x minimizes F;
- if F_n is a sequence Γ-converging to F, then $F_n + G$ will Γ-converge to $F + G$ for any continuous function $G : X \to \mathbb{R} \cup \{+\infty\}$.

One of the first applications of Γ-convergence was to provide elliptic approximations of free discontinuity problems, or, more generally, variational problems with "singular energies." We will only mention the following standard example (known as "Modica-Mortola" approximation; see [237] and [133]) because of its simplicity. Define the functional F_ε on $L^1(\Omega)$ through

$$F_\varepsilon(u) = \begin{cases} \frac{1}{\varepsilon}\int W(u(x))dx + \varepsilon \int |\nabla u(x)|^2 dx & \text{if } u \in H^1(\Omega); \\ +\infty & \text{otherwise.} \end{cases}$$

Then, if $W(0) = W(1) = 0$ and $W(t) > 0$ for any $t \neq 0, 1$, we have Γ-convergence of the functionals F_ε toward the functional F given by

$$F(u) = \begin{cases} c\mathrm{Per}(S) & \text{if } u = \mathbb{1}_S \text{ and } S \text{ is a finite-perimeter set;} \\ +\infty & \text{otherwise,} \end{cases}$$

where the constant c is given by $c = 2\int_0^1 \sqrt{W(t)}dt$.

Inspired by this result, in [247] a similar result for branched transport has been proven. Unfortunately, it has only been proved for the 2D case $\Omega \subset \mathbb{R}^2$. Consider, on the space $\mathscr{M}^d_{\mathrm{div}}(\Omega)$ of vector measures on Ω with divergence which is also a measure (endowed with the weak convergence of both the measures and their divergences), the functionals

$$M^\alpha_\varepsilon(\mathbf{w}) = \varepsilon^{\alpha-1} \int_\Omega |\mathbf{w}(x)|^\beta dx + \varepsilon^{\alpha+1} \int_\Omega |\nabla \mathbf{w}(x)|^2 dx, \tag{4.35}$$

defined on $\mathbf{w} \in H^1(\Omega;\mathbb{R}^2)$ and set to $+\infty$ outside $H^1 \cap \mathscr{M}^d_{\mathrm{div}}(\Omega)$. The exponent $\beta = \frac{4\alpha-2}{\alpha+1}$ is fixed from suitable scaling computations.

Compared to the Modica-Mortola case, here the double-well potential is replaced with a concave power. Note that concave powers, in their minimization, if the average value for u is fixed in a region (which is in some sense the meaning of weak convergence), prefer either $u = 0$ or $|u|$ being as large as possible, i.e., there is sort of a double well at zero and infinity.

Theorem 4.46. *Suppose $d = 2$ and $\alpha \in]\frac{1}{2}, 1[$, then we have Γ-convergence of the functionals M^α_ε to cM^α, with respect to the convergence in $\mathscr{M}^d_{\mathrm{div}}(\Omega)$, as $\varepsilon \to 0$, where c is a finite and positive constant (the value of c is actually $c = \alpha^{-1}\left(4c_0\alpha/(1-\alpha)\right)^{1-\alpha}$, with $c_0 = \int_0^1 \sqrt{t^\beta - t}\,\mathrm{d}t$).*

As a consequence of this approximation result, É. Oudet performed efficient numerical simulations by minimizing the functionals M^α_ε (more precisely: first a large value of ε is fixed, so that the problem is more or less convex, and a minimizer

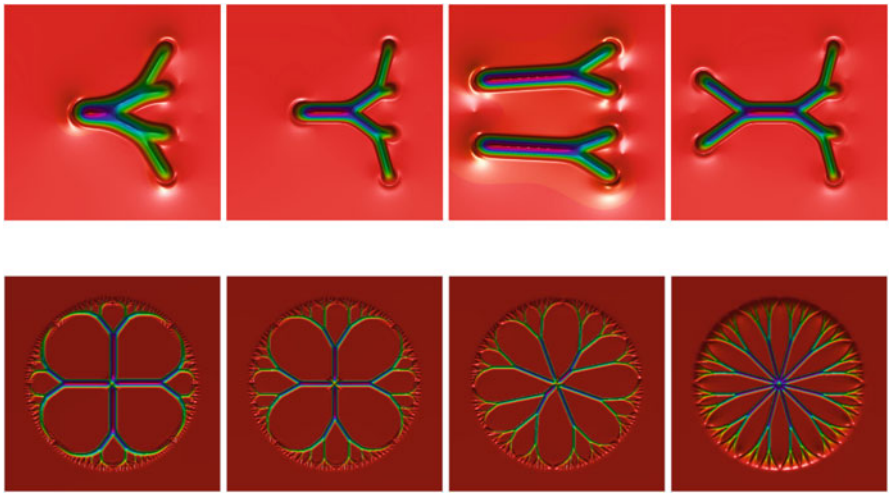


Fig. 4.7 Branched transport computed via Γ-convergence between one point and four points, two points and four points, and one point and the uniform distribution on a circle, for different values of α (increasing from left to right). Pictures partially taken from [247], with permission

is found by gradient descent, then, following a continuation method, the value of ε is reduced and at every step the gradient descent starts from the previously found minimizer). In the end we find a "well-chosen" local minimizer quite satisfactory in practice (Figure 4.7).

Chapter 5
Wasserstein distances and curves in the Wasserstein spaces

In this chapter, we will use the minimal value of transport problems between two probabilities in order to define a distance on the space of probabilities. We mainly consider costs of the form $c(x,y) = |x-y|^p$ in $\Omega \subset \mathbb{R}^d$, but the analysis can be adapted to a power of the distance in a more general metric space X. The exponent p will always be taken in $[1,+\infty[$ (we will briefly address the case $p = +\infty$ in the Discussion Section 5.5.1) in order to take advantage of the properties of the L^p norms. When Ω is unbounded, we need to restrict our analysis to the following set of probabilities:

$$\mathscr{P}_p(\Omega) := \left\{ \mu \in \mathscr{P}(\Omega) \,:\, \int_\Omega |x|^p \,\mathrm{d}\mu < +\infty \right\}.$$

In a metric space, we fix an arbitrary point $x_0 \in X$, and set $\mathscr{P}_p(X) := \{\mu \in \mathscr{P}(X) \,:\, \int d(x,x_0)^p \,\mathrm{d}\mu(x) < +\infty\}$ (the finiteness of this integral does not depend on the choice of x_0).

Without entering into details for the moment, let us only mention that the distances that we use are defined in the following way: for $\mu, \nu \in \mathscr{P}_p(X)$, we set $W_p(\mu,\nu) := \mathscr{T}_c(\mu,\nu)^{1/p}$, where $\mathscr{T}_c$ is the minimal transport cost for $c(x,y) = d(x,y)^p$ and d is the distance on X.

The distance that we obtain in this way are called Wasserstein distances.[1] They are very important in many fields of applications and they seem a natural way to

[1]The name is highly debated, in particular in Russia, since L. Vaserstein (whose name is sometimes spelled Wasserstein), a Russian scholar based in the United States, did not really play the key role that one could imagine in the introduction of these distances. Yet, this is nowadays the standard name in Western countries, probably due to the terminology used in [198, 246], and it seems impossible to change this convention, even if other names have been often suggested, such as Monge-Kantorovich distances, etc. Also note than in the most applied communities the name *Earth Mover Distance* (EMD) is very common for this distance.

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2_5

describe distances between equal amounts of mass distributed on a same space. To translate into a language more meaningful for applications,[2] one could say that Wasserstein distances are a way to take a distance onto a set X of objects (colors, words, points, etc.) and create out of it a distance on the set of all possible histograms on X (see also Section 2.5.1).

It is interesting to compare these distances to L^p distances between densities (a comparison which is meaningful when we consider absolutely measures on $\mathbb{R}^d$, for instance). A first observation is the very different behavior of these two classes of distances. We could say that, if L^p distances can be considered "vertical," Wasserstein distances are instead "horizontal." This consideration is very informal, but it is quite natural if one associates with every absolutely continuous measure the graph of its density. To compute $||f - g||_{L^p}$, we need to look, for every point x, the distance between $f(x)$ and $g(x)$, which corresponds to a vertical displacement between the two graphs, and then integrate this quantity. On the contrary, to compute $W_p(f, g)$, we need to consider the distance between a point x and a point $T(x)$ (i.e., a horizontal displacement on the graph) and then to integrate this, for a particular pairing between x and $T(x)$ which makes a coupling between f and g.

A first example where we can see the very different behavior of these two ways of computing distances is the following: take two densities f and g supported on $[0, 1]$, and define g_h as $g_h(x) = g(x - h)$. As soon as $|h| > 1$, the L^p distance between f and g_h equals $(||f||_{L^p}^p + ||g||_{L^p}^p)^{1/p}$ and does not depend on the "spatial" information consisting in $|h|$. On the contrary, the W_p distance between f and g_h is of the order of $|h|$ (for $h \to \infty$) and depends much more on the displacement than on the shapes of f and g. As another example, consider the distance between a density g and its translation g_h, for $|h| \to 0$. In this case, we have $W_p(g, g_h) = |h|$, while $||g - g_h||_{L^p}$ can be much larger than $|h|$ and is of order $|h|$ if and only if g is smooth enough (Figure 5.1).

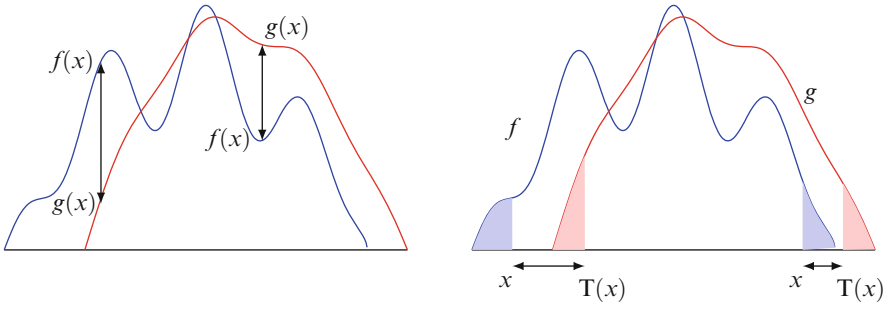


Fig. 5.1 "Vertical" vs "horizontal" distances (the transport T is calculated in the picture on the right as in the 1D case, by imposing equality between the blue and red areas under the graphs of f and g, i.e., using the CFD functions as in Chapter 2)

[2] In particular in image processing. Thanks to G. Peyré for pointing out this interpretation.

In several applications, this robustness of the Wasserstein distances w.r.t. perturbation of the densities makes them preferable to other distances, and the fact that they quantify spatial features of the measures is exactly what is required. Also, we will see (in Section 5.4 and later 6.1 and 7.3) that the geodesics for these distances provide useful interpolations between distributions of mass.

5.1 Definition and triangle inequality

In this section, we give the basic definitions of Wasserstein distances and spaces in $\Omega \subset \mathbb{R}^d$. We will see how to adapt the analysis to the case of more general metric spaces. As we said above, we restrict our analysis to the set

$$\mathscr{P}_p(\Omega) := \left\{ \mu \in \mathscr{P}(\Omega) \,:\, \int |x|^p \,\mathrm{d}\mu < +\infty \right\}.$$

Note that $p < q$ implies $\mathscr{P}_p(\Omega) \subset \mathscr{P}_q(\Omega)$ and that, whenever Ω is bounded, $\mathscr{P}_p(\Omega) = \mathscr{P}(\Omega)$.

For $\mu, \nu \in \mathscr{P}_p(\Omega)$, let us define

$$W_p(\mu,\nu) := \min \left\{ \int_{\Omega\times\Omega} |x-y|^p \,\mathrm{d}\gamma \,:\, \gamma \in \Pi(\mu,\nu) \right\}^{\frac{1}{p}},$$

i.e., the pth root of the minimal transport cost $\mathscr{T}_c$ for the cost $c(x,y) = |x-y|^p$. The assumption $\mu, \nu \in \mathscr{P}_p(\Omega)$ guarantees finiteness of this value since $|x-y|^p \le C(|x|^p + |y|^p)$, whence $W_p^p(\mu,\nu) \le C(\int |x|^p \,\mathrm{d}\mu + \int |x|^p \,\mathrm{d}\nu)$.

Note that, due to Jensen inequality, since all $\gamma \in \Pi(\mu,\nu)$ are probability measures, for $p \le q$ we have

$$\left(\int |x-y|^p \,\mathrm{d}\gamma \right)^{\frac{1}{p}} = ||x-y||_{L^p(\gamma)} \le ||x-y||_{L^q(\gamma)} = \left(\int |x-y|^q \,\mathrm{d}\gamma \right)^{1/q},$$

which implies $W_p(\mu,\nu) \le W_q(\mu,\nu)$. In particular, $W_1(\mu,\nu) \le W_p(\mu,\nu)$ for every $p \ge 1$. We will not define here W_∞, but we refer to Section 5.5.1 for a definition as a limit for $p \to \infty$.

On the other hand, for bounded Ω, an opposite inequality also holds, since

$$\left(\int |x-y|^p \,\mathrm{d}\gamma \right)^{\frac{1}{p}} \le \mathrm{diam}(\Omega)^{\frac{p-1}{p}} \left(\int |x-y| \,\mathrm{d}\gamma \right)^{\frac{1}{p}}.$$

This implies

$$W_p(\mu,\nu) \le C W_1(\mu,\nu)^{\frac{1}{p}}, \tag{5.1}$$

for $C = \mathrm{diam}(\Omega)^{(p-1)/p}$.

Proposition 5.1. *The quantity W_p defined above is a distance over $\mathscr{P}_p(\Omega)$.*

Proof. First, let us note that $W_p \geq 0$. Then, we also remark that $W_p(\mu,\nu) = 0$ implies that there exists $\gamma \in \Pi(\mu,\nu)$ such that $\int |x-y|^p \,\mathrm{d}\gamma = 0$. Such a $\gamma \in \Pi(\mu,\nu)$ is concentrated on $\{x=y\}$. This implies $\mu = \nu$ since, for any test function ϕ we have

$$\int \phi \,\mathrm{d}\mu = \int \phi(x)\,\mathrm{d}\gamma = \int \phi(y)\,\mathrm{d}\gamma = \int \phi \,\mathrm{d}\nu.$$

We need now to prove the triangle inequality. We give two different proofs of this fact. The first (Lemma 5.3) is easier and uses transport maps and approximation; the second (Lemma 5.4) is more general but requires a trickier tool, i.e., disintegrations of measures. □

Lemma 5.2. *Given $\mu,\nu \in \mathscr{P}_p(\mathbb{R}^d)$ and χ_ε any even regularizing kernel in L^1 with $\int_{\mathbb{R}^d} \chi_\varepsilon(z)\,\mathrm{d}z = 1$ and $\chi_\varepsilon(z) = \varepsilon^{-d}\chi_1(z/\varepsilon)$, then we have*

1. *the inequality $W_p(\mu * \chi_\varepsilon, \nu * \chi_\varepsilon) \leq W_p(\mu,\nu)$*
2. *the limit $\lim_{\varepsilon\to 0} W_p(\mu * \chi_\varepsilon, \nu * \chi_\varepsilon) = W_p(\mu,\nu)$.*

Proof. Take an optimal transport plan $\gamma \in \Pi(\mu,\nu)$ and define a transport plan $\gamma_\varepsilon \in \Pi(\mu * \chi_\varepsilon, \nu * \chi_\varepsilon)$ through

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} \phi(x,y)\,\mathrm{d}\gamma_\varepsilon := \int_{\mathbb{R}^d\times\mathbb{R}^d}\int_{\mathbb{R}^d} \phi(x+z,y+z)\chi_\varepsilon(z)\mathrm{d}z \;\mathrm{d}\gamma(x,y).$$

We need to check that its marginals are actually $\mu * \chi_\varepsilon$ and $\nu * \chi_\varepsilon$. Just consider

$$\begin{aligned}\int_{\mathbb{R}^d\times\mathbb{R}^d} \phi(x)\,\mathrm{d}\gamma_\varepsilon &= \int_{\mathbb{R}^d\times\mathbb{R}^d}\int_{\mathbb{R}^d} \phi(x+z)\chi_\varepsilon(z)\mathrm{d}z\,\mathrm{d}\gamma(x,y)\\ &= \int_{\mathbb{R}^d\times\mathbb{R}^d} (\phi * \chi_\varepsilon)(x)\,\mathrm{d}\gamma(x,y) = \int_{\mathbb{R}^d} (\phi * \chi_\varepsilon)(x)\,\mathrm{d}\mu(x)\\ &= \int_{\mathbb{R}^d} \phi(x)\,\mathrm{d}(\mu * \chi_\varepsilon)(x)\end{aligned}$$

(we use the fact that χ_ε is even to pass the convolution from ϕ to μ). The computation for the second marginal is the same. It is then easy to show that $\int |x-y|^p\,\mathrm{d}\gamma_\varepsilon = \int |x-y|^p\,\mathrm{d}\gamma$, since

$$\int |x-y|^p\,\mathrm{d}\gamma_\varepsilon = \int_{\mathbb{R}^d\times\mathbb{R}^d}\int_{\mathbb{R}^d} |(x+z)-(y+z)|^p\chi_\varepsilon(z)\mathrm{d}z\;\mathrm{d}\gamma(x,y) = \int |x-y|^p\,\mathrm{d}\gamma.$$

This shows the first part of the statement and also

$$\limsup_{\varepsilon\to 0}\; W_p(\mu * \chi_\varepsilon, \nu * \chi_\varepsilon)^p \leq \limsup_{\varepsilon\to 0}\; \int |x-y|^p\,\mathrm{d}\gamma_\varepsilon = \int |x-y|^p\,\mathrm{d}\gamma.$$

One can also obtain the opposite inequality by a standard semi-continuity argument. Consider $\mu_\varepsilon := \mu * \chi_\varepsilon$ and $\nu_\varepsilon := \nu * \chi_\varepsilon$, and use only the weak convergences $\mu_\varepsilon \rightharpoonup \mu$ and $\nu_\varepsilon \rightharpoonup \nu$. Call $\overline{\gamma_\varepsilon}$ the optimal plans in $\Pi(\mu_\varepsilon, \nu_\varepsilon)$. This sequence of plans is tight, since its marginals are tight (look at the proof of Theorem 1.7). First fix a sequence $\varepsilon_k \to 0$ such that $\lim_k W_p(\mu_{\varepsilon_k}, \nu_{\varepsilon_k}) = \liminf_{\varepsilon\to 0} W_p(\mu_\varepsilon, \nu_\varepsilon)$. Then extract a subsequence ε_{k_j} so as to guarantee that the optimal transport plans $\overline{\gamma_{\varepsilon_{k_j}}}$ have a weak limit $\overline{\gamma_0}$. This weak limit must belong to $\Pi(\mu, \nu)$ (the fact that the marginals of $\overline{\gamma_0}$ are μ and ν follows by the properties of composition with continuous functions of the weak convergence). Then we have

$$W_p^p(\mu, \nu) \le \int |x-y|^p \, \mathrm{d}\overline{\gamma_0} \le \liminf_j \int |x-y|^p \, \mathrm{d}\overline{\gamma_{\varepsilon_{k_j}}} = \liminf_{\varepsilon\to 0} W_p(\mu_\varepsilon, \nu_\varepsilon),$$

where the first inequality follows from the fact that $\overline{\gamma_0}$ is not necessarily optimal but is admissible and the second follows by semi-continuity (Lemma 1.6). □

Then, we can perform a proof of the triangle inequality based on the use of optimal transport maps.

Lemma 5.3. *The quantity W_p satisfies the triangle inequality (proof by transport maps and approximation).*

Proof. First consider the case where μ and ϱ are absolutely continuous and ν is arbitrary. Let T be the optimal transport from μ to ϱ and S the optimal one from ϱ to ν. Then $S\circ \mathrm{T}$ is an admissible transport from μ to ν, since $(S\circ \mathrm{T})_\# \mu = S_\#(\mathrm{T}_\# \mu) = S_\# \varrho = \nu$. We have

$$\begin{aligned} W_p(\mu, \nu) \le \left(\int |S(\mathrm{T}(x)) - x|^p \, \mathrm{d}\mu(x) \right)^{\frac{1}{p}} &= ||S\circ \mathrm{T} - \mathrm{id}||_{L^p(\mu)} \\ &\le ||S\circ \mathrm{T} - \mathrm{T}||_{L^p(\mu)} + ||\mathrm{T} - \mathrm{id}||_{L^p(\mu)}. \end{aligned}$$

Moreover,

$$||S\circ \mathrm{T} - \mathrm{T}||_{L^p(\mu)} = \left(\int |S(\mathrm{T}(x)) - \mathrm{T}(x)|^p \, \mathrm{d}\mu(x) \right)^{\frac{1}{p}} = \left(\int |S(y) - y|^p \mathrm{d}\varrho(y) \right)^{\frac{1}{p}}$$

and this last quantity equals $W_p(\varrho, \nu)$. Moreover, $||\mathrm{T} - \mathrm{id}||_{L^p(\mu)} = W_p(\mu, \varrho)$, whence

$$W_p(\mu, \nu) \le W_p(\mu, \varrho) + W_p(\varrho, \nu).$$

This gives the proof when $\mu, \varrho \ll \mathscr{L}^d$. For the general case, first write the triangle inequality for $\mu * \chi_\varepsilon$, $\varrho * \chi_\varepsilon$, and $\nu * \chi_\varepsilon$, then pass to the limit as $\varepsilon \to 0$ using Lemma 5.2. □

The proof above strongly uses the results about transport maps from Chapter 1 and is somehow specific to $\mathbb{R}^d$ (in a general metric space, the tricky part would be to

approximate arbitrary measures with measures such that the Wasserstein distance can be computed with maps instead of plans[3]). To give a more general result, we provide a different proof, which may happen to be more difficult for the reader who is not accustomed to disintegrations of measures (see Chapter 2).

Lemma 5.4. *The quantity W_p satisfies the triangle inequality (proof by disintegrations).*

Proof. Let us take μ, ϱ and $\nu \in \mathscr{P}_p(\Omega)$, $\gamma^+ \in \Pi(\mu, \varrho)$ and $\gamma^- \in \Pi(\varrho, \nu)$. We can also choose $\gamma^\pm$ to be optimal. Let us use Lemma 5.5 to say that there exists a measure $\sigma \in \mathscr{P}(\Omega \times \Omega \times \Omega)$ such that $(\pi_{x,y})_\# \sigma = \gamma^+$ and $(\pi_{y,z})_\# \sigma = \gamma^-$, where $\pi_{x,y}$ and $\pi_{y,z}$ denote the projections on the two first and two last variables, respectively. Let us take $\gamma := (\pi_{x,z})_\# \sigma$. By composition of the projections, it is easy to see that $(\pi_x)_\# \gamma = (\pi_x)_\# \sigma = (\pi_x)_\# \gamma^+ = \mu$ and, analogously, $(\pi_z)_\# \gamma = \nu$. This means $\gamma \in \Pi(\mu, \nu)$ and

$$
\begin{aligned}
W_p(\mu,\nu) &\le \left(\int |x-z|^p \,\mathrm{d}\gamma\right)^{\frac{1}{p}} = \left(\int |x-z|^p \mathrm{d}\sigma\right)^{\frac{1}{p}} = ||x-z||_{L^p(\sigma)} \\
&\le ||x-y||_{L^p(\sigma)} + ||y-z||_{L^p(\sigma)} = \left(\int |x-y|^p \mathrm{d}\sigma\right)^{\frac{1}{p}} + \left(\int |y-z|^p \mathrm{d}\sigma\right)^{\frac{1}{p}} \\
&= \left(\int |x-y|^p \,\mathrm{d}\gamma^+\right)^{\frac{1}{p}} + \left(\int |y-z|^p \,\mathrm{d}\gamma^-\right)^{\frac{1}{p}} = W_p(\mu,\varrho) + W_p(\varrho,\nu).
\end{aligned}
$$

The proof is concluded. □

The following lemma is usually known as "gluing lemma" and allows to produce a sort of "composition" of two transport plans, as if they were maps.

Lemma 5.5. *Given two measures $\gamma^+ \in \Pi(\mu, \varrho)$ and $\gamma^- \in \Pi(\varrho, \nu)$, there exists (at least) a measure $\sigma \in \mathscr{P}(\Omega \times \Omega \times \Omega)$ such that $(\pi_{x,y})_\# \sigma = \gamma^+$ and $(\pi_{y,z})_\# \sigma = \gamma^-$, where $\pi_{x,y}$ and $\pi_{y,z}$ denote the projections on the two first and two last variables, respectively.*

Proof. Start by taking γ^+ and disintegrate it w.r.t. the projection π_y (see Box 2.2 in Section 2.3). We get a family of measures $\gamma_y^+ \in \mathscr{P}(\Omega)$ (as we pointed out in the same box, we can think of them as measures over Ω, instead of viewing them as measures over $\Omega \times \{y\} \subset \Omega \times \Omega$). They satisfy (and are actually defined by)

$$
\int_{\Omega\times\Omega} \phi(x,y)\,\mathrm{d}\gamma^+(x,y) = \int_\Omega \mathrm{d}\varrho(y) \int_\Omega \phi(x,y)\ \mathrm{d}\gamma_y^+(x),
$$

[3]It is not strictly necessary to use absolutely continuous measures: we actually need atomless measures, and then we can apply Theorem 1.33.

for every measurable function ϕ of two variables. In the same way, one has a family of measures $\gamma_y^- \in \mathscr{P}(\Omega)$ such that for every ϕ, we have

$$\int_{\Omega\times\Omega} \phi(y,z)\,\mathrm{d}\gamma^-(y,z) = \int_\Omega \mathrm{d}\varrho(y) \int_\Omega \phi(y,z)\ \mathrm{d}\gamma_y^-(z).$$

For every y, take now $\gamma_y^+ \otimes \gamma_y^-$, which is a measure over $\Omega\times\Omega$. Define σ through

$$\int_{\Omega\times\Omega\times\Omega} \phi(x,y,z)\mathrm{d}\sigma(x,y,z) := \int_\Omega \mathrm{d}\varrho(y) \int_{\Omega\times\Omega} \phi(x,y,z)\, d\left(\gamma_y^+ \otimes \gamma_y^-\right)(x,z).$$

It is easy to check that, for ϕ depending only on x and y, we have

$$\begin{aligned}\int_{\Omega^3} \phi(x,y)\mathrm{d}\sigma &= \int_\Omega \mathrm{d}\varrho(y) \int_{\Omega\times\Omega} \phi(x,y)\, d\left(\gamma_y^+ \otimes \gamma_y^-\right)(x,z)\\ &= \int_\Omega \mathrm{d}\varrho(y) \int_\Omega \phi(x,y)\ \mathrm{d}\gamma_y^+(x) = \int \phi\ \mathrm{d}\gamma^+.\end{aligned}$$

This proves $(\pi_{x,y})_\#\sigma = \gamma^+$ and the proof of $(\pi_{y,z})_\#\sigma = \gamma^-$ is analogous. □

Remark 5.6. All the analysis of this section can be performed on a general Polish space X instead of a subset $\Omega \subset \mathbb{R}^d$. The only differences are the definition of $\mathscr{P}_p(X)$ (we need to use $\mathscr{P}_p(X) := \{\mu \in \mathscr{P}(X) : \int d(x,x_0)^p\,\mathrm{d}\mu(x) < +\infty\}$), and the fact that the first proof of the triangle inequality was specific to $\mathbb{R}^d$ (but the second works in general).

Definition 5.7. Given a Polish space X, for each $p \in [1,+\infty[$ we define its Wasserstein space of order p, $\mathbb{W}_p(X)$, as the space $\mathscr{P}_p(X)$, endowed with the distance W_p.

5.2 Topology induced by W_p

In this section, we want to analyze the convergence in the space $\mathbb{W}_p$ and compare it to the notion of weak convergence. First of all, we recall what we mean by "weak convergence": as we pointed out in Box 1.3 in Chapter 1, we use this term to denote the convergence in the duality with C_b, the space of bounded continuous functions (which is often referred to as narrow convergence), and we write $\mu_n \rightharpoonup \mu$ to say that μ_n converges in such a sense to μ.

We recall that another natural notion of convergence for measures is the one in duality with the space of functions vanishing at infinity

$$C_0(X) := \{\phi \in C(X) :\ \text{for all } \varepsilon > 0 \text{ there is } K \subset X \text{ compact s.t. } |\phi| < \varepsilon \text{ on } X \setminus K\}.$$

It is useful to point out some facts:

- When X is compact, the spaces $C_0(X)$, $C_c(X)$, $C(X)$, and $C_b(X)$ coincide.
- The condition defining $\phi \in C_0(X)$ means that ϕ tends to 0 "out of the compact subsets of X": this means at infinity, for $X = \mathbb{R}^d$, on $\partial\Omega$ when $X = \Omega \subset \mathbb{R}^d$ is open, and it means nothing if X is itself compact.
- For every $x_0 \in X$, the existence of functions $\phi \in C_0(X)$ with $\phi(x_0) \neq 0$ implies the existence of compact neighborhoods of x_0. In particular, the space $C_0(X)$ is interesting only for locally compact spaces X, and it only consists of the zero function whenever no point of X has compact neighborhoods (for instance, in infinite-dimensional Banach spaces). Indeed, the duality result stating that $\mathscr{M}(X)$ is the dual of $C_0(X)$ is only true for locally compact spaces (see Box 1.3 in Section 1.1).

The following lemma (which is true in locally compact spaces, but we only state it for subsets of $\mathbb{R}^d$) gives a helpful criterion for weak convergence.

Lemma 5.8. *When all measures μ_n and μ are probabilities over a subset $\Omega \subset \mathbb{R}^d$, the convergence $\mu_n \rightharpoonup \mu$ coincides with the convergence in the duality with $C_c(\Omega)$.*

Proof. We only need to show that if we take $\psi \in C_b(\Omega)$, $\mu_n, \mu \in \mathscr{P}(\Omega)$ and we assume $\int \phi \, \mathrm{d}\mu_n \to \int \phi \, \mathrm{d}\mu$ for every $\phi \in C_c(\Omega)$, then we also have $\int \psi \, \mathrm{d}\mu_n \to \int \psi \, \mathrm{d}\mu$. If all the measures are probability, then we may add "for free" a constant C to ψ, and, since ψ is bounded, we can choose C so that $\psi + C \geq 0$. Hence $\psi + C$ is the sup of an increasing family of functions $\phi_n \in C_c$ (take $\phi_n = (\psi + C)\chi_n$, χ_n being an increasing family of cutoff functions[4] with $\chi_n = 1$ on $B(0, n)$). Hence by semi-continuity, we have $\int (\psi + C) \, \mathrm{d}\mu \leq \liminf_n \int (\psi + C) \, \mathrm{d}\mu_n$, which implies $\int \psi \, \mathrm{d}\mu \leq \liminf_n \int \psi \, \mathrm{d}\mu_n$. If the same argument is performed with $-\psi$, we obtain the desired convergence of the integrals. □

Note that in the above lemma one can also use C_0 instead of C_c, as $C_c \subset C_0$. Once the weak convergence is understood, we can start from the following result.

Theorem 5.9. *If $\Omega \subset \mathbb{R}^d$ is compact, then $\mu_n \rightharpoonup \mu$ if and only if $W_1(\mu_n, \mu) \to 0$.*

Proof. Let us recall the duality formula, which gives for arbitrary $\mu, \nu \in \mathscr{P}(\Omega)$

$$W_1(\mu, \nu) = \max \left\{ \int_\Omega \varphi \, \mathrm{d}(\mu - \nu) \, : \, \varphi \in \mathrm{Lip}_1 \right\}.$$

Let us start from a sequence μ_n such that $W_1(\mu_n, \mu) \to 0$. Thanks to the duality formula, for every $\varphi \in \mathrm{Lip}_1(\Omega)$, we have $\int \varphi \, \mathrm{d}(\mu_n - \mu) \to 0$. By linearity, the same will be true for any Lipschitz function. By density, it will hold for any function in $C(\Omega)$. This shows that Wasserstein convergence implies weak convergence.

[4] The existence of cutoff functions, i.e., a sequence of continuous compactly supported functions converging pointwisely to 1, is indeed peculiar to the case of locally compact spaces.

To prove the opposite implication, let us first fix a subsequence μ_{n_k} such that $\lim_k W_1(\mu_{n_k}, \mu) = \limsup_n W_1(\mu_n, \mu)$. For every k, pick a function $\varphi_{n_k} \in \mathrm{Lip}_1(\Omega)$ such that $\int \varphi_{n_k} \, \mathrm{d}(\mu_{n_k} - \mu) = W_1(\mu_{n_k}, \mu)$. Up to adding a constant, which does not affect the integral, we can assume that the φ_{n_k} all vanish at a same point, and they are hence uniformly bounded and equicontinuous. By the Ascoli-Arzelà theorem we can extract a sub-subsequence uniformly converging to a certain $\varphi \in \mathrm{Lip}_1(\Omega)$. By replacing the original subsequence with this new one, we can avoid relabeling. We have now

$$W_1(\mu_{n_k}, \mu) = \int \varphi_{n_k} \, \mathrm{d}(\mu_{n_k} - \mu) \to \int \varphi \, \mathrm{d}(\mu - \mu) = 0,$$

where the convergence of the integral is justified by the weak convergence $\mu_{n_k} \rightharpoonup \mu$ together with the strong convergence (in $C(\Omega)$) $\varphi_{n_k} \to \varphi$. This shows that $\limsup_n W_1(\mu_n, \mu) \leq 0$ and concludes the proof. □

Theorem 5.10. *If $\Omega \subset \mathbb{R}^d$ is compact and $p \in [1, +\infty[$, in the space $\mathbb{W}_p(\Omega)$, we have $\mu_n \rightharpoonup \mu$ if and only if $W_p(\mu_n, \mu) \to 0$.*

Proof. We have already proved this equivalence for $p = 1$. For the other values of p, just use the inequalities

$$W_1(\mu, \nu) \leq W_p(\mu, \nu) \leq C W_1(\mu, \nu)^{\frac{1}{p}}$$

that give the equivalence between the convergence for W_p and for W_1. □

We can now pass to the case of unbounded domains. From the fact that the distance W_p does not really depend on the ambient domain, we can simply consider the case $\Omega = \mathbb{R}^d$.

Theorem 5.11. *In the space $\mathbb{W}_p(\mathbb{R}^d)$, we have $W_p(\mu_n, \mu) \to 0$ if and only if $\mu_n \rightharpoonup \mu$ and $\int |x|^p \, \mathrm{d}\mu_n \to \int |x|^p \, \mathrm{d}\mu$.*

Proof. Consider first a sequence μ_n which is converging to μ in $\mathbb{W}_p(\mathbb{R}^d)$. It is still true in this case that

$$\sup \left\{ \int \phi \, d(\mu_n - \mu) \, : \, \phi \in \mathrm{Lip}_1 \right\} \to 0,$$

which gives the convergence when testing against arbitrary Lipschitz functions. Note that Lipschitz functions are dense (for the uniform convergence) in the space $C_0(\Omega)$ (while it is not necessarily the case for $C_b(\Omega)$). This is enough to prove $\mu_n \to \mu$, thanks to Lemma 5.8.

To obtain the other condition, namely, $\int |x|^p \, \mathrm{d}\mu_n \to \int |x|^p \, \mathrm{d}\mu$ (which is not a consequence of the weak convergence, since $|x|^p$ is not bounded), it is sufficient to note that

$$\int |x|^p \, \mathrm{d}\mu_n = W_p^p(\mu_n, \delta_0) \to W_p^p(\mu, \delta_0) = \int |x|^p \, \mathrm{d}\mu.$$

We need now to prove the opposite implication. Consider a sequence a $\mu_n \rightharpoonup \mu$ satisfying also $\int |x|^p \, \mathrm{d}\mu_n \to \int |x|^p \, \mathrm{d}\mu$. Fix $R > 0$ and consider the function $\phi(x) := (|x| \wedge R)^p$, which is continuous and bounded. We have

$$\begin{aligned}\int (|x|^p - (|x| \wedge R)^p) \, \mathrm{d}\mu_n &= \int |x|^p \, \mathrm{d}\mu_n - \int \phi \, \mathrm{d}\mu_n \\ &\to \int |x|^p \, \mathrm{d}\mu - \int \phi \, \mathrm{d}\mu = \int (|x|^p - (|x| \wedge R)^p) \, \mathrm{d}\mu.\end{aligned}$$

Since $\int (|x|^p - (|x| \wedge R)^p) \, \mathrm{d}\mu \le \int_{B(0,R)^c} |x|^p \, \mathrm{d}\mu$ and $\int_{\mathbb{R}^d} |x|^p \, \mathrm{d}\mu < \infty$, it is possible to choose R so that

$$\int (|x|^p - (|x| \wedge R)^p) \, \mathrm{d}\mu < \varepsilon/2$$

and hence one can also guarantee that $\int (|x|^p - (|x| \wedge R)^p) \, \mathrm{d}\mu_n < \varepsilon$ for all n large enough.

We use now the inequality $(|x| - R)^p \le |x|^p - R^p = |x|^p - (|x| \wedge R)^p$ which is valid for $|x| \ge R$ (see Lemma 5.12) to get

$$\int (|x| - R)^p \, \mathrm{d}\mu_n < \varepsilon \text{ for } n \text{ large enough, and } \int (|x| - R)^p \, \mathrm{d}\mu < \varepsilon.$$

Consider now $\pi_R : \mathbb{R}^d \to \overline{B(0,R)}$ defined as the projection over $\overline{B(0,R)}$. This map is well defined and continuous and is the identity on $\overline{B(0,R)}$. Moreover, for every $x \notin \overline{B(0,R)}$, we have $|x - \pi_R(x)| = |x| - R$. We can deduce

$$\begin{aligned}W_p(\mu, (\pi_R)_\# \mu) &\le \left(\int (|x| - R)^p \, \mathrm{d}\mu \right)^{\frac{1}{p}} \le \varepsilon^{\frac{1}{p}}, \\ W_p(\mu_n, (\pi_R)_\# \mu_n) &\le \left(\int (|x| - R)^p \, \mathrm{d}\mu_n \right)^{\frac{1}{p}} \le \varepsilon^{\frac{1}{p}}.\end{aligned}$$

Note also that, due to the usual composition of the weak convergence with continuous functions, from $\mu_n \rightharpoonup \mu$, we also infer $(\pi_R)_\# \mu_n \rightharpoonup (\pi_R)_\# \mu$. Yet, these measures are all concentrated on the compact set $\overline{B(0,R)}$, and here we can use the equivalence between weak convergence and W_p convergence. Hence, we get

$$\begin{aligned}&\limsup_n W_p(\mu_n, \mu) \\ &\le \limsup_n \left(W_p(\mu_n, (\pi_R)_\# \mu_n) + W_p((\pi_R)_\# \mu_n, (\pi_R)_\# \mu) + W_p(\mu, (\pi_R)_\# \mu) \right) \\ &\le 2\varepsilon^{\frac{1}{p}} + \lim_n W_p((\pi_R)_\# \mu_n, (\pi_R)_\# \mu) = 2\varepsilon^{\frac{1}{p}}.\end{aligned}$$

As the parameter $\varepsilon > 0$ is arbitrary, we get $\limsup_n W_p(\mu_n, \mu) = 0$, and the proof is concluded. □

Lemma 5.12. *For $a, b \in \mathbb{R}_+$ and $p \geq 1$, we have $a^p + b^p \leq (a+b)^p$.*

Proof. Suppose without loss of generality that $a \geq b$. Then we can write $(a+b)^p = a^p + p\zeta^{p-1}b$, for a point $\zeta \in [a, a+b]$. Use now $p \geq 1$ and $\zeta \geq a \geq b$ to get $(a+b)^p \geq a^p + b^p$. □

Remark 5.13. How to handle more general metric spaces X instead of subsets of $\mathbb{R}^d$? First, we notice that Lemma 5.8 does not hold in general (for instance, it is trivially false in infinite-dimensional Banach spaces). Also, the problem of the density of Lipschitz functions in the space of continuous functions (even if X is compact) appears. Indeed, in $\mathbb{R}^d$ the fact that smooth functions are dense is easy to obtain by convolution, a tool which is not available in general spaces. To overcome this difficulty, one can prove the following statement, which is enough for our scopes: "if we have convergence testing against arbitrary Lipschitz functions, than we also have convergence testing against arbitrary functions in C_b" (this is left as an exercise, see **Ex**(31)). Hence, in what concerns Theorems 5.9 and 5.10, everything can be translated into the framework of Polish spaces (with no need of local compactness). On the other hand, the proof that we gave of Theorem 5.11 was somehow specific to $\mathbb{R}^d$: the key point was the existence of a continuous retraction π_R from the whole space to a compact subset (the ball $\overline{B(0,R)}$), with a bound on the distance between x and $\pi_R(x)$. The proof could be adapted to more general spaces, with ad hoc assumptions, but we prefer to refer the reader to [9, 15] for the general case.

5.3 Lipschitz curves in $\mathbb{W}_p$ and the continuity equation

In this section, we analyze some properties of Lipschitz and absolutely continuous curves in the space $\mathbb{W}_p$. In order to do that, we need some simple elements from analysis in metric spaces. The reader who wants to know more can have a look, for instance, at [12].

Box 5.1. Good to know! *Curves and speed in metric spaces*

Let us recall here some properties about Lipschitz curves in metric spaces.

A curve ω is a continuous function defined on a interval, say $[0,1]$ and valued in a metric space (X,d). As it is a map between metric spaces, it is meaningful to say whether it is Lipschitz or not, but its velocity $\omega'(t)$ has no meaning, unless X is a vector space.

Surprisingly, it is possible to give a meaning to the modulus of the velocity, $|\omega'|(t)$.

Definition: If $\omega : [0,1] \to X$ is a curve valued in the metric space (X,d), we define the metric derivative of ω at time t, denoted by $|\omega'|(t)$ through

(continued)

Box 5.1. (continued)

$$|\omega'|(t) := \lim_{h\to 0} \frac{d(\omega(t+h),\omega(t))}{|h|},$$

provided this limit exists.

The following theorem, in the spirit of Rademacher theorem (see Box 1.9 in Section 1.3), guarantees the existence of the metric derivative for Lipschitz curves.

Theorem: Suppose that $\omega : [0,1] \to X$ is Lipschitz continuous. Then the metric derivative $|\omega'|(t)$ exists for a.e. $t \in [0,1]$. Moreover, we have, for $t < s$,

$$d(\omega(t),\omega(s)) \le \int_t^s |\omega'|(\tau)\,\mathrm{d}\tau.$$

The above theorem can be proved by using the fact that every compact metric space (and $\omega([0,1])$ is the image of a compact set through a continuous map; hence it is compact) can be isometrically embedded in ℓ^∞, where one can work componentwise. For all the notions and the proofs about metric derivatives, we refer, for instance, to [12], Chapter 4.

We also need to consider more general curves, not only Lipschitz continuous.

Definition: A curve $\omega : [0,1] \to X$ is defined *absolutely continuous* whenever there exists $g \in L^1([0,1])$ such that $d(\omega(t_0),\omega(t_1)) \le \int_{t_0}^{t_1} g(s)\mathrm{d}s$ for every $t_0 < t_1$. The set of absolutely continuous curves defined on $[0,1]$ and valued in X is denoted by $\mathrm{AC}(X)$.

It is well known that every absolutely continuous curve can be reparametrized in time (through a monotone-increasing reparametrization) and become Lipschitz continuous. A possible way to achieve this goal is the following: let $G(t) := \int_0^t g(s)\mathrm{d}s$, and then set $S(t) = \varepsilon t + G(t)$ (for any $\varepsilon > 0$), which is continuous and strictly increasing and valued in an interval $[0,L]$; for $t \in [0,L]$, set $\tilde\omega(t) = \omega(S^{-1}(t))$. It is easy to check that $\tilde\omega \in \mathrm{Lip}_1$. Also, if we let $\varepsilon \to 0$, we obtain a unit-speed reparametrization of ω. If we want to have parametrizations defined on the same interval $[0,1]$, we just need to rescale by a factor L.

In particular, the above Rademacher theorem is also true for $\omega \in \mathrm{AC}(X)$ (since the reparametrization that we defined is differentiable a.e.).

The goal of this section is to identify the absolutely continuous curves in the space $\mathbb{W}_p(\Omega)$ with solutions of the continuity equation $\partial_t\mu_t + \nabla\cdot(\mathbf{v}_t\mu_t) = 0$ with L^p vector fields $\mathbf{v}_t$. Moreover, we want to connect the L^p norm of $\mathbf{v}_t$ with the metric derivative $|\mu'|(t)$.

We recall (see Section 4.1) that the continuity equation may be interpreted as the equation ruling the evolution of the density μ_t of a family of particles initially distributed according to μ_0, and each of which follows the flow

$$\begin{cases} y_x'(t) = \mathbf{v}_t(y_x(t)) \\ y_x(0) = x. \end{cases}$$

We state below the main theorem (originally proven in [15]) relating absolutely continuous curves in $\mathbb{W}_p$ with solutions of the continuity equation. For simplicity, we will only state it in the framework of a compact domain $\Omega \subset \mathbb{R}^d$. We will explain where we use the compactness assumption and how to get rid of it.

Theorem 5.14. *Let $(\mu_t)_{t\in[0,1]}$ be an absolutely continuous curve in $\mathbb{W}_p(\Omega)$ (for $p > 1$ and $\Omega \subset \mathbb{R}^d$ compact). Then for a.e. $t \in [0,1]$, there exists a vector field $\mathbf{v}_t \in L^p(\mu_t;\mathbb{R}^d)$ such that*

- *the continuity equation $\partial_t\mu_t + \nabla\cdot(\mathbf{v}_t\mu_t) = 0$ is satisfied in the weak sense (see Section 4.1.2),*
- *for a.e. t, we have $||\mathbf{v}_t||_{L^p(\mu_t)} \le |\mu'|(t)$ (where $|\mu'|(t)$ denotes the metric derivative at time t of the curve $t \mapsto \mu_t$, w.r.t. the distance W_p);*

Conversely, if $(\mu_t)_{t\in[0,1]}$ is a family of measures in $\mathscr{P}_p(\Omega)$ and for each t we have a vector field $\mathbf{v}_t \in L^p(\mu_t;\mathbb{R}^d)$ with $\int_0^1 ||\mathbf{v}_t||_{L^p(\mu_t)}\,\mathrm{d}t < +\infty$ solving $\partial_t\mu_t + \nabla\cdot(\mathbf{v}_t\mu_t) = 0$, then $(\mu_t)_t$ is absolutely continuous in $\mathbb{W}_p(\Omega)$, and for a.e. t, we have $|\mu'|(t) \le ||\mathbf{v}_t||_{L^p(\mu_t)}$.

Remark 5.15. Note that, as a consequence of the second part of the statement, the vector field $\mathbf{v}_t$ introduced in the first one must a posteriori satisfy $||\mathbf{v}_t||_{L^p(\mu_t)} = |\mu'|(t)$.

We will split the proof of this theorem into two parts, dealing with the direct and converse implications, respectively. We first introduce some tools that we need, and in particular, the following one will be needed in both parts of the statement of Theorem 5.14.

Remark 5.16. It is important to observe that, by reparametrizing in time, it is always possible to assume that the absolutely continuous curve μ_t of the first part of the statement is actually Lipschitz (and we obtain in this case a uniform bound on $||\mathbf{v}_t||_{L^p(\mu_t)}$) and that the norm $||\mathbf{v}_t||_{L^p(\mu_t)}$ in the second part is bounded (and we obtain that μ_t is Lipschitz). Hence, all the proofs will be done in this Lipschitz framework.

5.3.1 The Benamou-Brenier functional $\mathscr{B}_p$

We need to start from an easy computation, which will be useful in the sequel. We consider a pair of exponents p and q such that $\frac{1}{p} + \frac{1}{q} = 1$.

Lemma 5.17. *Set $K_q := \{(a,b) \in \mathbb{R}\times\mathbb{R}^d : a + \frac{1}{q}|b|^q \le 0\}$. Then, for $(t,x) \in \mathbb{R}\times\mathbb{R}^d$ we have*

$$\sup_{(a,b)\in K_q} (at + b\cdot x) = f_p(t,x) := \begin{cases} \frac{1}{p}\frac{|x|^p}{t^{p-1}} & \text{if } t > 0, \\ 0 & \text{if } t = 0, x = 0 \\ +\infty & \text{if } t = 0, x \neq 0, \quad \text{or } t < 0. \end{cases} \tag{5.2}$$

In particular, f_p is convex and l.s.c.

Proof. First of all, we establish (5.2). Suppose $t > 0$: then it is clear that one should take the maximal possible value of a in the sup, and hence $a = -\frac{1}{q}|b|^q$. This gives

$$\sup_b \left(-\frac{1}{q} t|b|^q + b \cdot x \right) = t \left(\sup_b -\frac{1}{q}|b|^q + b \cdot \left(\frac{x}{t}\right) \right).$$

If we recall that the Legendre transform of $p \mapsto \frac{1}{q}|p|^q$ is $x \mapsto \frac{1}{p}|x|^p$. This gives

$$\sup_b -\frac{1}{q}|b|^q + b \cdot y = \frac{1}{p}|y|^p \quad \text{for all } y, \tag{5.3}$$

and hence

$$\sup \left\{ at + b \cdot x \,:\, a \in \mathbb{R},\, b \in \mathbb{R}^d,\, a + \frac{1}{q}|b|^q \le 0 \right\} = t \frac{1}{p} \left| \frac{x}{t} \right|^p = \frac{1}{p} \frac{|x|^p}{t^{p-1}}.$$

The case $(t, x) = (0, 0)$ is straightforward. If $t = 0$ and $x \neq 0$, then it is clear that any vector b may be compensated by a sufficiently negative value of a, which gives

$$\sup \left\{ at + b \cdot x \,:\, a \in \mathbb{R},\, b \in \mathbb{R}^d,\, a + \frac{1}{q}|b|^q \le 0 \right\} = \sup_b \; b \cdot x = +\infty.$$

Finally, in the case $t < 0$, one can take a arbitrarily negative and $b = 0$ so that

$$\sup \left\{ at + b \cdot x \,:\, a \in \mathbb{R},\, b \in \mathbb{R}^d,\, a + \frac{1}{q}|b|^q \le 0 \right\} \ge \sup_{a<0} \; at = +\infty.$$

The fact that f_p is convex and l.s.c. is a consequence of the fact that it is expressed as a supremum of linear functions. □

We can note in the above result that we computed the Legendre transform of an indicator function and that we got, as expected, a 1-homogeneous convex function. We want now to use this function f_p to define a functional over measures. For $\varrho \in \mathscr{M}(X)$ and $E \in \mathscr{M}^d(X)$, we set

$$\mathscr{B}_p(\varrho, E) := \sup \left\{ \int_X a(x) \mathrm{d}\varrho + \int_X b(x) \cdot dE \,:\, (a, b) \in C_b(X; K_q) \right\},$$

where the letter $\mathscr{B}$ is chosen because of the role that this functional has in the so-called Benamou-Brenier formula (see Section 6.1).

We prove the following:

Proposition 5.18. *The functional $\mathscr{B}_p$ is convex and lower semi-continuous on the space $\mathscr{M}(X) \times \mathscr{M}^d(X)$ for the weak convergence. Moreover, the following properties hold:*

1. $\mathscr{B}_p(\varrho, E) \ge 0$,
2. $\mathscr{B}_p(\varrho, E) := \sup\{\int a(x)\mathrm{d}\varrho + \int b(x) \cdot \mathrm{d}E \,:\, (a, b) \in L^\infty(X; K_q)\}$,
3. *if both ϱ and E are absolutely continuous w.r.t. a same positive measure λ on X, we can write $\mathscr{B}_p(\varrho, E) = \int_X f_p(\varrho(x), E(x))\, \mathrm{d}\lambda(x)$, where we identify ϱ and E with their densities w.r.t. λ,*

4. *$\mathscr{B}_p(\varrho, E) < +\infty$ only if $\varrho \geq 0$ and $E \ll \varrho$,*
5. *for $\varrho \geq 0$ and $E \ll \varrho$, we have $E = \mathbf{v} \cdot \varrho$ and $\mathscr{B}_p(\varrho, E) = \int \frac{1}{p}|\mathbf{v}|^p \, \mathrm{d}\varrho$,*
6. *If $X = \mathbb{R}^d$, $\varrho^\varepsilon := \varrho * \eta_\varepsilon$ and $E^\varepsilon := E * \eta_\varepsilon$ (for standard even mollifying kernel η_ε), then we have $\mathscr{B}_p(\varrho^\varepsilon, E^\varepsilon) \leq \mathscr{B}_p(\varrho, E)$.*

Proof. The convexity and lower semi-continuity of $\mathscr{B}_p$ are a direct consequence of its definition as a supremum of linear functionals. The first property (positivity) is straightforward, as one can take $a = 0$ and $b = 0$ in the supremum.

In order to prove that the supremum can be taken over bounded functions instead of continuous ones, we just need to approximate every bounded function with continuous ones (beware that, by abuse of notation, the space $L^\infty(X; K_q)$ denotes here the space of bounded measurable functions valued in K_q, independently of the reference measure). This can be done via Lusin's theorem (see the corresponding Box 1.6 in Section 1.1), and every bounded pair (a, b) can be approximated by continuous pairs $(\tilde{a}, \tilde{b})$, such that $(\varrho + |E|)(\{(a, b) \neq (\tilde{a}, \tilde{b})\}) < \varepsilon$ and $\sup|\tilde{a}| \leq \sup|a|$ and $\sup|\tilde{b}| \leq \sup|b|$. Then, we can replace $\tilde{a}$ with $\min\{\tilde{a}, -\frac{1}{q}|\tilde{b}|^q\}$ so as to guarantee to have functions valued in K_q. The difference $\left|\int a(x)\mathrm{d}\varrho + \int b(x) \cdot \mathrm{d}E - \int \tilde{a}(x)\mathrm{d}\varrho + \int \tilde{b}(x) \cdot \mathrm{d}E\right|$ is bounded by $C\varepsilon$ where C only depends on $\sup|a|$ and $\sup|b|$.

Once we know that we can use measurable functions a, b with no need of continuity, then property 3 is a consequence of Lemma 5.2. As a consequence of Lemma 5.17, we can write

$$\mathscr{B}_p(\varrho, E) = \sup\left\{\int [a(x)\varrho(x) + b(x) \cdot E(x)]\,\mathrm{d}\lambda(x) \;:\; (a, b) \in L^\infty(X; K_q)\right\}$$
$$= \int f_p(\varrho(x), E(x))\mathrm{d}\lambda(x).$$

The fact that $\mathscr{B}_p$ is only finite if $\varrho \geq 0$ is easy: assume it to be false for contradiction, and take a set A such that $\varrho(A) < 0$, $a = -n\mathbb{1}_A$ and $b = 0$. We have $\mathscr{B}_p(\varrho, E) \geq n\varrho(A)$ and, since n is arbitrary, $\mathscr{B}_p(\varrho, E) = +\infty$. To prove the fact that $\mathscr{B}_p$ is only finite if $E \ll \varrho$, assume that there is a set A with $\varrho(A) = 0$ and $E(A) \neq 0$. Take any unit vector e and use $a = -\frac{n^q}{q}\mathbb{1}_A$ and $b = ne\mathbb{1}_A$. We have $\mathscr{B}_p(\varrho, E) \geq ne \cdot E(A)$. Since both e and n are arbitrary, we get $\mathscr{B}_p(\varrho, E) = +\infty$.

When we restrict to the case $\varrho \geq 0$ and $E = \mathbf{v} \cdot \varrho$, property 5 is a consequence of property 3. Indeed, just use $\lambda = \varrho$, thus getting

$$\mathscr{B}_p(\varrho, \mathbf{v} \cdot \varrho) = \int f_p(1, \mathbf{v}(x))\mathrm{d}\varrho(x) = \int \frac{1}{p}|\mathbf{v}|^p \, \mathrm{d}\varrho.$$

Let us now prove property 6. Take arbitrary bounded functions $a : \mathbb{R}^d \to \mathbb{R}$ and $b : \mathbb{R}^d \to \mathbb{R}^d$ satisfying $a(x) + \frac{1}{q}|b(x)|^q \leq 0$ for all x. From standard properties of convolutions, we have

$$\int a\,\mathrm{d}\varrho^\varepsilon + \int b\cdot \mathrm{d}E^\varepsilon = \int a^\varepsilon\,\mathrm{d}\varrho + \int b^\varepsilon\cdot\,\mathrm{d}E,$$

where $a^\varepsilon := a * \eta_\varepsilon$ and $b^\varepsilon := b * \eta_\varepsilon$. Note that, by Jensen inequality

$$|b^\varepsilon(y)|^q = \left|\int b(x)\eta_\varepsilon(x-y)\,\mathrm{d}x\right|^q \le \int |b(x)|^q\eta_\varepsilon(x-y)\,\mathrm{d}x$$

and hence

$$a^\varepsilon(y) + \frac{1}{q}|b^\varepsilon(y)|^q \le \int a(x) + \left(\frac{1}{q}|b(x)|^q\right)\eta_\varepsilon(x-y)\,\mathrm{d}x \le 0.$$

This proves $(a^\varepsilon, b^\varepsilon)\in C(\mathbb{R}^d; K_q)$, and hence we have

$$\int a\,\mathrm{d}\varrho^\varepsilon + \int b\cdot\mathrm{d}E^\varepsilon = \int a^\varepsilon\,\mathrm{d}\mu + \int b^\varepsilon\cdot\mathrm{d}E \le \mathscr{B}_p(\varrho, E).$$

Passing to the sup in a and b, one gets $\mathscr{B}_p(\varrho^\varepsilon, E^\varepsilon)\le \mathscr{B}_p(\varrho, E)$. □

Remark 5.19. As a convex and l.s.c. functional, $\mathscr{B}_p$ has admits a subdifferential, which one could usefully compute for optimization purposes: this computation is developed in [235], but we will not need it in this framework.

We are now ready to prove the first part of Theorem 5.14.

5.3.2 *Proof of Theorem 5.14, Part 1: μ is AC $\Rightarrow$ there exists* $\mathbf{v}$

Proof. First of all, we note that it is not restrictive to assume that μ is a Lipschitz curve (if not, we just reparametrize in time).

We will do our proof by approximation, i.e., we build a sequence μ_t^k of curves which admit the existence of suitable vector fields $\mathbf{v}_t^k$, in such a way that $\mu_t^k \rightharpoonup \mu_t$ and that we can find a limit for the vector fields also.

To define μ_t^k we proceed in the following way. For $t = i/k$, $i = 0, 1, \dots, k$, we define $\mu_t^k = \mu_t * \eta_k$, where η_k is a smooth approximation of the identity, i.e., a smooth even positive function supported on $B(0, 1/k)$ and with integral equal to 1. The only goal of this smoothing is to guarantee that the measures are absolutely continuous. The support of the measures is now contained in a slightly bigger domain $\Omega'\supset\Omega$ that we can choose convex. Then, for each $i = 0, 1, \dots, k-1$, we define $\mathrm{T}^{i,k} : \Omega'\to\Omega'$ to be the optimal transport (for the cost $c(x,y) = |x-y|^p$) from $\mu_{i/k}^k$ to $\mu_{(i+1)/k}^k$. These optimal transport maps exist since every $\mu_{i/k}^k$ is absolutely continuous. Then, for $t\in]\frac{i}{k}, \frac{i+1}{k}[$, we define

$$\mu_t^k := \big((i+1-kt)\mathrm{id} + (kt-i)\mathrm{T}^{i,k}\big)_\# \mu_{i/k}^k,$$

which means that the particle located at x at time i/k goes to the position $\mathrm{T}^{i,k}(x)$ at time $(i+1)/k$, moving with constant speed on the segment connecting these two positions in the interval $]\frac{i}{k}, \frac{i+1}{k}[$. The velocity of this movement is given by $\mathbf{v}^{i,k} := k(\mathrm{T}^{i,k}(x) - x)$. Set $\mathrm{T}_t^{i,k} := (i+1-kt)\mathrm{id} + (kt-i)\mathrm{T}^{i,k}$; this map is injective (see Lemma 4.23). Thus, we can set $\mathbf{v}_t^k := \mathbf{v}^{i,k} \circ \left(\mathrm{T}_t^{i,k}\right)^{-1}$. This last vector field exactly represents the velocity that each particle has during its movement at time t: first we detect which is the original position of such a particle (at time i/k), then we associate with this particle the velocity associated with this original position.

As a consequence of this construction and of the considerations in Section 4.1, the pair $(\mu_t^k, \mathbf{v}_t^k)$ solves the continuity equation $\partial_t \mu_t^k + \nabla \cdot (\mathbf{v}_t^k \mu_t^k) = 0$.

The important point is the computation

$$
\begin{aligned}
||\mathbf{v}_t^k||^p_{L^p(\mu_t^k)} = \int |\mathbf{v}_t^k|^p \,\mathrm{d}\mu_t^k = \int |\mathbf{v}^{i,k}|^p \circ \left(\mathrm{T}_t^{i,k}\right)^{-1} \mathrm{d}\left(\mathrm{T}_t^{i,k}\right)_{\#}\mu_{i/k}^k = \int |\mathbf{v}^{i,k}|^p \,\mathrm{d}\mu_{i/k}^k \\
= k^p W_p^p(\mu_{i/k}^k, \mu_{(i+1)/k}^k) \le k^p W_p^p(\mu_{i/k}, \mu_{(i+1)/k}) \\
\le k^p \left(\int_{i/k}^{(i+1)/k} |\mu'|(s)\mathrm{d}s \right)^p \le k \int_{i/k}^{(i+1)/k} |\mu'|^p(s)\mathrm{d}s.
\end{aligned}
$$

Here we used the fact that $W_p(\mu_{i/k} * \eta_\varepsilon, \mu_{(i+1)/k} * \eta_\varepsilon) \le W_p(\mu_{i/k}, \mu_{(i+1)/k})$ (see Lemma 5.2) and, at the end, Jensen's inequality. In particular, we also get a uniform bound on $||\mathbf{v}_t^k||_{L^p(\mu_t^k)}$ since $k \int_{i/k}^{(i+1)/k} |\mu'|^p(s)ds \le \mathrm{Lip}(\mu)$.

This implies that, for arbitrary time $0 \le a < b \le 1$ (with $i_a/k \le a \le (i_a+1)/k$ and $i_b/k \le b \le (i_b+1)/k$), we have

$$
\begin{aligned}
\int_a^b ||\mathbf{v}_t^k||^p_{L^p(\mu_t^k)} \,\mathrm{d}t &\le \sum_{i=i_a}^{i_b} \int_{i/k}^{(i+1)/k} |\mu'|^p(s)\mathrm{d}s \\
&= \int_{i_a/k}^{(i_b+1)/k} |\mu'|^p(s)\mathrm{d}s \le \int_a^b |\mu'|^p(s)\mathrm{d}s + \frac{2\,\mathrm{Lip}(\mu)}{k}. \qquad (5.4)
\end{aligned}
$$

Now, define a vector measure E^k on $\Omega' \times [0,1]$ via

$$
\int \phi(t,x) \cdot \,\mathrm{d}E^k := \int_0^1 \left(\int \phi(t,x) \cdot \mathbf{v}_t^k(x) \,\mathrm{d}\mu_t^k(x) \right) \mathrm{d}t.
$$

We can easily check that

$$
||E^k|| = \int_0^1 ||\mathbf{v}_t^k||_{L^1(\mu_t^k)} \mathrm{d}t \le \int_0^1 ||\mathbf{v}_t^k||_{L^p(\mu_t^k)} \mathrm{d}t \le \left(\int_0^1 ||\mathbf{v}_t^k||^p_{L^p(\mu_t^k)} \mathrm{d}t \right)^{\frac{1}{p}} \le C.
$$

Hence, up to a subsequence extraction, the sequence of vector measures E^k admits a weak limit E. On the other hand, it is easy to check that μ^k converges in $\mathbb{W}_p$, uniformly in time, to μ. Indeed

$$W_p(\mu_t^k, \mu_t) \le W_p(\mu_t^k, \mu_{i/k}^k) + W_p(\mu_{i/k}^k, \mu_{i/k}) + W_p(\mu_{i/k}, \mu_t).$$

The first term may be estimated using

$$W_p(\mu_t^k, \mu_{i/k}^k) \le |t - i/k| \, ||\mathbf{v}^{i,k}||_{L^p(\mu_{i/k}^k)} \le \frac{C}{k};$$

the second may be estimated by the size of the support of the kernel η_ε (which allows to obtain $\mu_{i/k}^k$ from $\mu_{i/k}$), which is $\varepsilon = 1/k$; the third is estimated by the Lipschitz constant of μ times, again, $|t - \frac{i}{k}| \le \frac{1}{k}$. Overall, we get

$$W_p(\mu_t^k, \mu_t) \le \frac{C}{k}.$$

From $\partial_t \mu^k + \nabla \cdot E^k = 0$, we can easily pass to the limit and get $\partial_t \mu + \nabla \cdot E = 0$. We want now to prove that $E_t = \mathbf{v}_t \mu_t$, with $||\mathbf{v}_t||_{L^p(\mu_t)} \le |\mu'|(t)$ a.e.

This is a consequence of Proposition 5.18 on the semi-continuity of $\mathscr{B}_p$. Indeed, using (5.4) with $a = 0, b = 1$, we have $\mathscr{B}_p(\mu^k, E^k) \le C$ (where the functional $\mathscr{B}_p$ is considered on the whole space-time domain $[0, 1] \times \Omega$). This is enough to obtain $\mathscr{B}_p(\mu, E) < +\infty$ and $E \ll \mu$, i.e., $E_t = \mathbf{v}_t \mu_t$. Then, if we apply the same lemma to the measures μ^k and E^k restricted to a closed interval $[a, b]$, we obtain

$$\int_a^b ||\mathbf{v}_t||^p_{L^p(\mu_t)} \mathrm{d}t \le \liminf_k \int_a^b ||\mathbf{v}_t^k||^p_{L^p(\mu_t^k)} \, \mathrm{d}t \le \int_a^b |\mu'|^p(t) \mathrm{d}t.$$

Since this is true for arbitrary values of a and b, on every Lebesgue point t_0 of $t \mapsto ||\mathbf{v}_t||^p_{L^p(\mu_t)}$ and $t \mapsto |\mu'|^p(t)$, we get the desired inequality. □

We want now to prove the converse implication of Theorem 5.14. Again, to be able to perform our computations rigorously, we need to regularize by convolution. The construction that we do here will be useful many times, and it deserves to be detailed and analyzed.

5.3.3 *Regularization of solutions of the continuity equation*

Suppose that (μ, E) solves the continuity equation $\partial_t \mu_t + \nabla \cdot E_t = 0$, that the supports of all measures μ_t and E_t are contained in a same compact set $\Omega \subset \mathbb{R}^d$, and that E_t is bounded in the space of measure, i.e., $||E_t|| := |E_t|(\mathbb{R}^d) \le C$. Take a C^∞ strictly positive kernel η_ε (as a function of space only) and define $\mu_t^\varepsilon := \eta_\varepsilon * \mu_t$ and $E_t^\varepsilon := \eta_\varepsilon * E_t$, and $\mathbf{v}_t^\varepsilon := E_t^\varepsilon / \mu_t^\varepsilon$. The choice is made so that $\partial_t \mu_t^\varepsilon + \nabla \cdot (\mu_t^\varepsilon \mathbf{v}_t^\varepsilon) = 0$.

Since we want to apply the characterization result of solutions of the continuity equation (Theorem 4.4), we want to check uniform global bounds on $\mathbf{v}^\varepsilon$. To do this, we need to perform a particular choice of η_ε, and in particular, we take it of the form $\eta^\varepsilon(z) = \varepsilon^{-d}\eta(|z|/\varepsilon)$, where $\eta : \mathbb{R}^+ \to \mathbb{R}^+$ is a smooth function such that $\eta(t) = Ce^{-t}$ for $t \geq 1$. In Lemma 5.20, we check that this guarantees that $\mathbf{v}^\varepsilon$ satisfies the required global bounds, and we also state regularity for μ_ε (that we will need in Section 5.3.5).

Lemma 5.20. *With the above choice of the convolution kernel η^ε, the function μ^ε is Lipschitz in (t,x) and $\mathbf{v}^\varepsilon$ is Lipschitz in x and bounded, uniformly in t (for fixed $\varepsilon > 0$), provided Ω is bounded.*

Proof. We can fix $\varepsilon = 1$ as the computations for other values of ε are similar. We set $\eta = \eta^\varepsilon$. From $\mu^\varepsilon(t,x) = \int \eta(|y-x|)\mu_t(\mathrm{d}x)$, we have a Lipschitz bound in x from standard convolution properties, but also

$$\partial_t \mu^\varepsilon(t,x) = \frac{d}{dt}\int \eta(y-x)\mathrm{d}\mu_t(x) = \int \nabla\eta(y-x)\cdot \mathrm{d}E_t(x),$$

which is bounded by $\mathrm{Lip}(\eta)|E_t|(\mathbb{R}^d)$.

As far as $\mathbf{v}^\varepsilon$ is concerned, computations are trickier since it is defined as $E^\varepsilon/\mu^\varepsilon$. From $\eta > 0$, we get $\mu^\varepsilon > 0$ which guarantees that $\mathbf{v}^\varepsilon$ is smooth as the ratio between two smooth functions, with nonvanishing denominator. Yet, we want to compute explicit and global bounds for large $|x|$. If $R_0 = \mathrm{diam}(\Omega)$ and $0 \in \Omega$, we have for sure, for $|x| = R \geq 1$, $\mu^\varepsilon(x) \geq \eta(R+R_0)$. On the other hand, we have, for $R \geq 2R_0 \vee 1$, $|E^\varepsilon| \leq C\eta(R-R_0)$, $|\nabla E^\varepsilon| \leq C|\eta'(R-R_0)|$, and, similarly, $|\nabla\mu^\varepsilon| \leq C|\eta'(R-R_0)|$. From

$$|\mathbf{v}^\varepsilon| = \frac{|E^\varepsilon|}{\mu^\varepsilon},\ |\nabla\mathbf{v}^\varepsilon| \leq \frac{|\nabla E^\varepsilon|}{\mu^\varepsilon} + \frac{|E^\varepsilon|\,|\nabla\mu^\varepsilon|}{(\mu^\varepsilon)^2},$$

using $\eta = \eta' = \eta''$ on $[1,+\infty[$ and $\eta(R+R_0) = e^{-2R_0}\eta(R-R_0)$, we get uniform bounds. □

We will see that the application of Lemma 5.20 is the only point where we needed Ω to be bounded. Without this assumption, we cannot apply the characterization result for the continuity equation that we presented in Theorem 4.4. On the other hand, we already mentioned that more general uniqueness results exist (for instance, in [15]), and we simply chose not to present them for the sake of simplicity.

Thanks to our definitions, we have

$$\partial_t \mu_t^\varepsilon + \nabla\cdot(\mathbf{v}_t^\varepsilon \mu_t^\varepsilon) = 0.$$

Moreover, applying Theorem 4.4, we know that the only solution of $\partial_t\varrho_t + \nabla\cdot(\mathbf{v}_t^\varepsilon\varrho_t) = 0$ is given by the flow of $\mathbf{v}^\varepsilon$. Hence, we get

$$\mu_t^\varepsilon = (Y_t)_\# \mu_0^\varepsilon, \quad \text{where } Y_t(x) = y_x(t), \quad \text{and } \begin{cases} y_x'(t) = \mathbf{v}_t^\varepsilon(y_x(t)), \\ y_x(0) = x. \end{cases} \tag{5.5}$$

We can summarize our construction in the following proposition.

Proposition 5.21. *Consider a pair (μ, E) solving $\partial_t \mu_t + \nabla \cdot E_t = 0$, and suppose that the supports of all measures μ_t and E_t are contained in a same compact set $\Omega \subset \mathbb{R}^d$ and that E_t is bounded in the space of measure, i.e., $|E_t|(\mathbb{R}^d) \le C$. Then, there exists a family of approximating curves $(\mu_t^\varepsilon)_t$ with a velocity vector field $\mathbf{v}_t^\varepsilon$ such that*

- *μ^ε is Lipschitz in (t,x) and $\mathbf{v}^\varepsilon$ is Lipschitz in x and bounded, uniformly in t; moreover, $\mu^\varepsilon > 0$ everywhere;*
- *μ_t^ε is obtained from μ_0^ε by following the flow of the vector field $\mathbf{v}^\varepsilon$, in the sense of (5.5);*
- *For every t, we have $\mu_t^\varepsilon \rightharpoonup \mu_t$ and for a.e. t, we have $E_t^\varepsilon \rightharpoonup E_t$.*
- *$||\mathbf{v}_t^\varepsilon||_{L^p(\mu_t^\varepsilon)} \le \mathscr{B}_p(\mu_t, E_t)$.*

Proof. We use the construction we described above. The first property is proven in Lemma 5.20, the second has been discussed above, and the third is evident by construction. The fourth is a consequence of property 5 of Lemma 5.18. Note that this means $||\mathbf{v}_t^\varepsilon||_{L^p(\mu_t^\varepsilon)} \le ||\mathbf{v}_t||_{L^p(\mu_t)}$ whenever $E_t = \mathbf{v}_t \cdot \mu_t$. □

We finish this paragraph by noting that the same construction could have been performed on the torus.

Remark 5.22. If $\mu_t \in \mathscr{P}(\mathbb{T}^d)$, one can consider an arbitrary approximation kernel $\eta^\varepsilon \in C^\infty(\mathbb{R}^d)$ of the form $\eta^\varepsilon(z) = \varepsilon^{-d}\eta(z/\varepsilon)$ (strictly positive, but with no need of imposing exponential behavior) and define the convolution on the torus. To be precise, let us stress that the periodic convolution $\eta \underline{*} \mu$ between a measure μ on $\mathbb{T}^d$ and a function η on $\mathbb{R}^d$ is defined as follows. Consider at first μ as a measure on $Q = [-\frac{1}{2}, \frac{1}{2}] \subset \mathbb{R}^d$, perform standard convolution on $\mathbb{R}^d$, and then define

$$(\eta \underline{*} \mu)(x) := \sum_{k \in \mathbb{Z}^d} (\eta * \mu)(x + k).$$

Equivalently, this corresponds to a convolution on the torus with the convolution kernel $\sum_{k \in \mathbb{Z}^d} \eta(x + k)$.

With this construction, the same conclusions as in Proposition 5.21 hold. Moreover, every compact domain Ω can be seen as a subset of a cube Q identified with the torus, and it can even fit an arbitrary small portion of Q. In this case, every optimal transport problem for the costs $|x - y|^p$ between measures on Ω "does not see" the fact that Q is identified with a torus (it is enough to have $\Omega \subset \frac{1}{2}Q$).

We stress that the uniform bounds in the case of the torus are easier to obtain, but we found conceptually easier to deal with measures on $\mathbb{R}^d$ (and in particular with

optimal transport problems in $\mathbb{R}^d$). This is why we prefer the Euclidean construction to this one. Yet, we present both of them for the use that we will do in Section 5.3.5, where compactness is crucial for our estimates.

5.3.4 *Proof of Theorem 5.14, Part 2: There exists* $\mathbf{v} \Rightarrow \boldsymbol{\mu}$ *is AC*

Proof. The starting point is a pair (μ, E) which solves the continuity equation, where $E_t = \mathbf{v}_t \cdot \mu_t$ and $\mathbf{v}_t \in L^p(\mu_t)$ for a.e. t. Again, we apply a reparametrization: with a change of variable in time, we can replace (μ_t, E_t) with a new pair where E_t is just multiplied by a scalar constant depending on t. By choosing our change of variable, we can impose a uniform bound on $||\mathbf{v}_t||_{L^p(\mu_t)}$. This also implies a bound on $||E_t|| = |E_t|(\mathbb{R}^d) = ||\mathbf{v}_t||_{L^1(\mu_t)}$.

Under this extra assumption (the uniform bound on $||E_t||$), we can use the construction provided in the previous paragraph (using the construction on $\mathbb{R}^d$ for simplicity of exposition). Thus, we can build an approximated curve μ_t^ε which is driven by the flow of the smooth (in space) vector field $\mathbf{v}^\varepsilon$. This provides a useful transport plan between μ_t^ε and μ_{t+h}^ε, taking $\gamma = (\mathrm{T}_t, \mathrm{T}_{t+h})_\# \mu_0^\varepsilon \in \Pi(\mu_t^\varepsilon, \mu_{t+h}^\varepsilon)$. We obtain

$$
\begin{aligned}
W_p(\mu_t^\varepsilon, \mu_{t+h}^\varepsilon) &\le \left(\int_{\Omega\times\Omega} |x-y|^p \,\mathrm{d}\gamma \right)^{\frac{1}{p}} = \left(\int_\Omega |\mathrm{T}_t(x) - \mathrm{T}_{t+h}(x)|^p \,\mathrm{d}\mu_0^\varepsilon \right)^{\frac{1}{p}} \\
&\le |h|^{1/q} \left(\int_\Omega \int_t^{t+h} \left| \frac{d}{ds} \mathrm{T}_s(x) \right|^p \mathrm{d}s \,\mathrm{d}\mu_0^\varepsilon \right)^{\frac{1}{p}} \\
&= |h|^{1/q} \left(\int_t^{t+h} \mathrm{d}s \int_\Omega \left| \mathbf{v}_s^\varepsilon(y_x(s)) \right|^p \,\mathrm{d}\mu_0^\varepsilon \right)^{\frac{1}{p}} \\
&= |h|^{1/q} \left(\int_t^{t+h} \mathrm{d}s \int_\Omega \left| \mathbf{v}_s^\varepsilon(y) \right|^p \,\mathrm{d}\mu_s^\varepsilon(y) \right)^{\frac{1}{p}}.
\end{aligned}
$$

We proved in Lemma 5.17 that $\int \left|\mathbf{v}_s^\varepsilon(y)\right|^p \,\mathrm{d}\mu_s^\varepsilon(y) \le ||\mathbf{v}_s||^p_{L^p(\mu_s)}$. From this, we get

$$
W_p(\mu_t^\varepsilon, \mu_{t+h}^\varepsilon) \le |h|^{1/q} \left(\int_t^{t+h} ||\mathbf{v}_s||^p_{L^p(\mu_s)} \,\mathrm{d}s \right)^{\frac{1}{p}} = |h| \left(\frac{1}{|h|} \int_t^{t+h} ||\mathbf{v}_s||^p_{L^p(\mu_t)} \,\mathrm{d}s \right)^{\frac{1}{p}}.
$$

First we can pass to the limit $\varepsilon \to 0$, thus obtaining

$$
\frac{W_p(\mu_t, \mu_{t+h})}{|h|} \le \left(\frac{1}{|h|} \int_t^{t+h} ||\mathbf{v}_s||^p_{L^p(\mu_s)} \,\mathrm{d}s \right)^{\frac{1}{p}}.
$$

This provides Lipschitz behavior for μ_t. Then, we take the limit $h \to 0$ at points t where $|\mu'|(t)$ exists and which are Lebesgue points of $t \mapsto ||\mathbf{v}_t||^p_{L^p(\mu_t)}$, we obtain

$$|\mu'|(t) \le ||\mathbf{v}_t||_{L^p(\mu_t)},$$

which provides the desired estimate. □

We remark that Theorem 5.14 has been proven thanks to an extensive use of the functional $\mathscr{B}_p$: we needed its convexity in the second part and its semi-continuity in the first part.[5]

Remark 5.23. We also note that we only defined the functional $\mathscr{B}_p$ for $p > 1$ and that the only natural extension for $p = 1$ would be given by $\mathscr{B}_1(\varrho, E) = ||E||$. It is easy to understand that the first part of Theorem 5.14 cannot stay true in the case $p = 1$. Indeed, the part of the statement which fails is the existence of a vector field $\mathbf{v}_t$, since it is no more possible to deduce $E \ll \mu$: it is not difficult to adapt the proof so as to get the existence of a family E_t of vector measures solving $\partial_t \mu_t + \nabla \cdot E_t = 0$, but E_t will be no more of the form $\mathbf{v}_t \cdot \mu_t$. An example where absolute continuity fails is the following: take $\mu_t := (1-t)\delta_{x_0} + t\delta_{x_1}$, which is Lipschitz in $\mathbb{W}_1$ (but not in $\mathbb{W}_p$, $p > 1$). In this case, one could take $E_t = \frac{x_1 - x_0}{|x_1 - x_0|} \mathscr{H}^1 \llcorner [x_0, x_1]$ (a uniform measure on the segment joining the two points, oriented as the same segment), but no measure concentrated on $\{x_0, x_1\}$ can be such that its divergence is $\delta_{x_1} - \delta_{x_0}$.

5.3.5 Derivative of W_p^p along curves of measures

We conclude this section with a useful computation which can be applied in many cases in the study of evolution PDEs. The same result is efficiently proven in [15], but we provide a different proof.

Theorem 5.24. *Let $(\varrho_t^{(i)}, \mathbf{v}_t^{(i)})$ for $i = 1, 2$ be two solutions of the continuity equation $\partial_t \varrho_t^{(i)} + \nabla \cdot (\mathbf{v}_t^{(i)} \varrho_t^{(i)}) = 0$ on a compact domain Ω (with no-flux boundary conditions), and suppose that $\varrho_t^{(i)} \ll \mathscr{L}^d$ for every t and that $\varrho^{(i)}$ are absolutely continuous curves in $\mathbb{W}_p(\Omega)$. Then we have*

$$\frac{d}{dt}\left(\frac{1}{p} W_p^p(\varrho_t^{(1)}, \varrho_t^{(2)})\right) = \int \nabla \varphi_t \cdot \mathbf{v}_t^{(1)} \varrho_t^{(1)} \, \mathrm{d}x + \int \nabla \psi_t \cdot \mathbf{v}_t^{(2)} \varrho_t^{(2)} \, \mathrm{d}x$$

[5] The proof that we gave of the second part is standard and based on the interpretation of the continuity equation that we gave in Chapter 4. The approximations that we performed are the same as in [15], but we tried to simplify them to make this exposition self-contained: in order to do that, we chose a very precise convolution kernel. Concerning the first part, we stress that a different, and much more elegant, approach can be found in [15] and does not use approximation. The approach that we presented here is more or less inspired by a paper by Lisini, [212].

for a.e. t, where (φ_t, ψ_t) is any pair of Kantorovich potentials in the transport between $\varrho_t^{(1)}$ and $\varrho_t^{(2)}$ for the cost $\frac{1}{p}|x-y|^p$.

Proof. We first prove the result under the assumption that $\varrho^{(i)}$ is Lipschitz continuous in (t,x) and $\mathbf{v}^{(i)}$ is Lipschitz in x, uniformly in t, for $i = 1, 2$. In this case, we write, for arbitrary t, t_0, the following inequality, justified by the duality formula (Theorem 1.39):

$$\frac{1}{p} W_p^p(\varrho_t^{(1)}, \varrho_t^{(2)}) - \int \varphi_{t_0} \varrho_t^{(1)} - \int \psi_{t_0} \varrho_t^{(2)} \geq 0, \tag{5.6}$$

with equality when $t = t_0$. Choose t_0 such that $t \mapsto W_p^p(\varrho_t^{(1)}, \varrho_t^{(2)})$ is differentiable at $t = t_0$ and $\varrho^{(i)}(t,x)$ is differentiable in t at (t_0, x) for a.e. x, and the derivative in time equals $-\nabla \cdot (\varrho^{(i)}(t,x)\mathbf{v}^i(t,x))$. All these conditions are satisfied for a.e. t_0 (we use Proposition 4.3). Hence, the left-hand side in (5.6) is differentiable at $t = t_0$ and the derivative must vanish, because of minimality at $t = t_0$. This proves

$$\frac{d}{dt}\left(\frac{1}{p} W_p^p\left(\varrho_t^{(1)}, \varrho_t^{(2)}\right)\right)_{|t=t_0} = \int \varphi_{t_0} \left(\partial_t \varrho^{(1)}\right)_{t_0} + \int \psi_{t_0} \left(\partial_t \varrho^{(2)}\right)_{t_0},$$

and the desired formula is obtained by using $\partial_t \varrho^{(i)} = -\nabla \cdot (\mathbf{v}^{(i)} \varrho^{(i)})$ and integrating by parts.[6]

To deal with the general case, first we reduce to the case where the curves $\varrho^{(i)}$ are Lipschitz in $\mathbb{W}_p(\Omega)$, by reparametrization in time. Then, we apply the regularization procedure of Section 5.3.3. In order to preserve compactness, we choose to apply the construction on the torus, and we suppose $\Omega \subset \frac{1}{2}Q$ (which is possible up to rescaling). Hence, we obtain two smoother families of curves $\varrho_t^{(i,\varepsilon)}$ with their velocity fields $\mathbf{v}_t^{(i,\varepsilon)}$. For $t_0 < t_1$, the first part of the proof implies

$$\frac{1}{p} W_p^p\left(\varrho_{t_1}^{(1,\varepsilon)}, \varrho_{t_1}^{(2,\varepsilon)}\right) - \frac{1}{p} W_p^p\left(\varrho_{t_0}^{(1,\varepsilon)}, \varrho_{t_0}^{(2,\varepsilon)}\right) = \int_{t_0}^{t_1} \mathrm{d}t \int_{\mathbb{T}^d} \nabla \varphi_t^\varepsilon \cdot E_t^{(1,\varepsilon)} + \nabla \psi_t^\varepsilon \cdot E_t^{(2,\varepsilon)},$$

where $E_t^{(i,\varepsilon)} = \varrho_t^{(i,\varepsilon)} \mathbf{v}_t^{(i,\varepsilon)}$ (also set $E_t^{(i)} = \varrho_t^{(i)} \mathbf{v}_t^{(i)}$) and φ_t^ε and ψ_t^ε are Kantorovich potentials in the transport from $\varrho_t^{(1,\varepsilon)}$ to $\varrho_t^{(2,\varepsilon)}$.

We just need to pass to the limit in the above equality. The left-hand side trivially converges to the corresponding Wasserstein distances between the measures $\varrho^{(i)}$ (note that we are on a compact space, and the convolutions weakly converge to the

[6] One of the difficulties in this result is that the functions $t \mapsto \int \varphi_{t_0} \varrho_t^{(1)}$ and $t \mapsto \int \psi_{t_0} \varrho_t^{(2)}$ are differentiable for a.e. t, but not necessarily at $t = t_0$ (we know that for every integrand the integral is differentiable for a.e. T, and not that for a.e. t differentiability occurs for every integrand). We propose a way to overcome this difficulty via Lipschitz solutions. As an alternative approach, one can try to find the precise set of times t such that differentiability occurs for every integrand, as in [155], based on ideas from [15].

original measure). Next, let us analyze the first part of the right-hand side (i.e., for $i = 1$, the other part being completely analogous).

For each t, we have

$$\left| \int_{\mathbb{T}^d} \nabla\varphi_t^\varepsilon \cdot E_t^{(1,\varepsilon)} \right| \le \mathrm{Lip}(\varphi_t^\varepsilon) \left\| \mathbf{v}_t^{(i,\varepsilon)} \right\|_{L^1\left(\varrho_t^{(i,\varepsilon)}\right)} \le C \left\| \mathbf{v}_t^{(i,\varepsilon)} \right\|_{L^p\left(\varrho_t^{(i,\varepsilon)}\right)} \le C \left\| \mathbf{v}_t^{(i)} \right\|_{L^p\left(\varrho_t^{(i)}\right)} \le C,$$

where we used property 5 of Proposition 5.18 and the fact that all the functions φ^ε are equi-Lipschitz (see Section 1.3.2).

It is now enough to look at the limit for fixed t, since as a function of t, we can apply dominated convergence. Let us set $\varphi_t^{\varepsilon\varepsilon} := \eta^\varepsilon * \varphi_t^\varepsilon$: we have

$$\int_{\mathbb{T}^d} \nabla\varphi_t^\varepsilon \cdot E_t^{(1,\varepsilon)} = \int_{\mathbb{T}^d} \nabla\varphi_t^\varepsilon \cdot (\eta^\varepsilon * E_t^{(1)}) = \int \nabla\varphi_t^{\varepsilon\varepsilon} \cdot E_t^{(1)},$$

because of standard properties of convolutions. Thus, we only need to prove that $\nabla\varphi_t^{\varepsilon\varepsilon} \to \nabla\varphi_t$ a.e. The functions $\varphi_t^{\varepsilon\varepsilon}$ are equi-Lipschitz and equi-semi-concave (see again Section 1.3.2) on a compact set and, up to translating their mean, we can apply the Ascoli-Arzelà theorem. Thus they converge up to subsequence to a function $\tilde{\varphi}$. Note that, because of semi-concavity, the convergence of the functions implies a.e. convergence of the gradients (see, for instance, Theorem 3.3.3 in [102]). We just need to prove $\nabla\tilde{\varphi} = \nabla\varphi_t$ (which would also imply that the convergence is true on the whole sequence).

Note that by Theorem 1.52, the limit of φ_t^ε must be a Kantorovich potential. Moreover, we have the general estimate

$$\begin{aligned} |(\eta^\varepsilon * f)(x) - f(x)| &\le \int |f(x) - f(x-y)|\, \eta\left(\frac{y}{\varepsilon}\right) \frac{dy}{\varepsilon^d} \\ &\le \mathrm{Lip}(f) \int |y| \eta\left(\frac{y}{\varepsilon}\right) \frac{dy}{\varepsilon^d} = \varepsilon\, \mathrm{Lip}(f) \int |z|\eta(z)dz, \end{aligned}$$

(where we used the change of variable $y = \varepsilon z$). The same estimate also works for periodic convolution and shows $||\varphi_t^{\varepsilon\varepsilon} - \varphi_t^\varepsilon||_{L^\infty} \le C\,\mathrm{Lip}(\varphi^\varepsilon)\varepsilon \le C\varepsilon \to 0$. This allows to conclude that the limit of $\varphi_t^{\varepsilon\varepsilon}$ must be the same as that of φ_t^ε, i.e., a Kantorovich potential. Even if the Kantorovich potential is not necessarily unique, its gradient is unique, and we deduce $\nabla\varphi_t^{\varepsilon\varepsilon} \to \nabla\varphi_t$. □

Note that in the above proof, we intensively used the compactness of the domain Ω (and of the torus: the same proof performed with regularization via measures on $\mathbb{R}^d$ seems delicate because of the lack of bounds on $\nabla\varphi_t^{\varepsilon\varepsilon}$). The assumption that the measures are absolutely continuous is used to take advantage of the a.e. convergence of the gradients to a unique limit (the gradient of the limit potential). Without this, one could have problems in recovering full convergence of the sequence.

As a corollary of the previous result, we obtain the following useful formula.

Corollary 5.25. *Under the same assumptions of Theorem 5.24, we have*

$$\frac{d}{dt}\left(\frac{1}{p}W_p^p(\varrho_t^{(1)},\varrho_t^{(2)})\right)=\int_\Omega (x-\mathrm{T}_t(x))\cdot\left(\mathbf{v}_t^{(1)}(x)-\mathbf{v}_t^{(2)}(\mathrm{T}_t(x))\right)\varrho_t^{(1)}(x)\,\mathrm{d}x,$$

where T_t *is the optimal transport map from* $\varrho_t^{(1)}$ *to* $\varrho_t^{(2)}$ *for the cost* $\frac{1}{p}|x-y|^p$.

Proof. We just need to rewrite the conclusion of Theorem 5.24, using $\nabla\varphi_t(x) = x-\mathrm{T}_t(x)$ and $\nabla\psi_t(y) = y-S_t(y)$, where S_t is the optimal transport from $\varrho_t^{(2)}$ to $\varrho_t^{(1)}$. We use

$$\int (y-S_t(y))\cdot\mathbf{v}_t^{(2)}(y)\,\mathrm{d}\varrho_t^{(2)}(y)=\int(\mathrm{T}_t(x)-x)\cdot\mathbf{v}_t^{(2)}(\mathrm{T}_t(x))\,\mathrm{d}\varrho_t^{(1)}(x),$$

which is a consequence of $(\mathrm{T}_t)_{\#}\varrho_t^{(1)}=\varrho_t^{(2)}$ and $S_t=\mathrm{T}_t^{-1}$. Hence, the formula given by Theorem 5.24 provides

$$\begin{aligned}&\frac{d}{dt}\left(\frac{1}{p}W_p^p(\varrho_t^{(1)},\varrho_t^{(2)})\right)\\&\quad=\int_\Omega(x-\mathrm{T}_t(x))\cdot\mathbf{v}_t^{(1)}(x)\,\mathrm{d}\varrho_t^{(1)}(x)+\int_\Omega(\mathrm{T}_t(x)-x)\cdot\mathbf{v}_t^{(2)}(\mathrm{T}_t(x))\,\mathrm{d}\varrho_t^{(1)}(x),\end{aligned}$$

which gives the claim.

Some examples of application of the above computation to the uniqueness and the properties of the solution to some evolution PDEs are presented in the Exercise section (**Ex**(34), **Ex**(66),**Ex**(67),**Ex**(68),**Ex**(69)).

Here, we will only give an example of application to the uniqueness for the continuity equation with given vector field.

Proposition 5.26. *Let* $\mathbf{v}_t:\Omega\to\mathbb{R}^d$ *be a family of vector fields on a compact smooth domain* Ω*, parametrized in time, satisfying a one-sided Lipschitz condition* $(\mathbf{v}_t(x)-\mathbf{v}_t(y))\cdot(x-y)\le C|x-y|^2$ *for a fixed constant* C*, not depending on* x,y*, nor* t*. Suppose also that* $\mathbf{v}_t\cdot\mathbf{n}=0$ *on* $\partial\Omega$ *and consider the equation* $\partial_t\varrho_t+\nabla\cdot(\varrho_t\mathbf{v}_t)=0$ *with given initial datum* ϱ_0*. Then there is at most one solution* ϱ_t *among AC curves in* $\mathbb{W}_2(\Omega)$ *which are absolutely continuous measures for a.e.* t*.*

Proof. Consider two solutions $\varrho_t^{(1)},\varrho_t^{(2)}$ and compute the derivative in time of $W_2^2(\varrho_t^{(1)},\varrho_t^{(2)})$. Using Corollary 5.25, we have

$$\begin{aligned}\frac{d}{dt}\left(\frac{1}{2}W_2^2(\varrho_t^{(1)},\varrho_t^{(2)})\right)&=\int_\Omega(x-\mathrm{T}(x))\cdot(\mathbf{v}_t(x)-\mathbf{v}_t(\mathrm{T}(x)))\varrho_t^{(1)}\,\mathrm{d}x\\&\le C\int_\Omega|x-\mathrm{T}(x)|^2\varrho_t^{(1)}\,\mathrm{d}x=CW_2^2(\varrho_t^{(1)},\varrho_t^{(2)}).\end{aligned}$$

This allows to apply Gronwall's lemma and deduce $W_2^2(\varrho_t^{(1)}, \varrho_t^{(2)}) = 0$ from the fact that the same distance vanishes at $t = 0$. □

Note that the above proof could also be used to prove uniqueness in the continuity equation and applied to Theorem 4.4.

5.4 Constant-speed geodesics in $\mathbb{W}_p$

We will see in this section how constant-speed geodesics in W_p are related to optimal transport maps. Before doing so, we recall the main facts about geodesics in metric spaces.

Box 5.2. Good to know! *Constant-speed geodesics in general metric spaces*

First of all, let us define the length of a curve ω in a general metric space (X, d).

Definition: For a curve $\omega : [0, 1] \to X$, let us define

$$\text{Length}(\omega) := \sup \left\{ \sum_{k=0}^{n-1} d(\omega(t_k), \omega(t_{k+1})) \, : \, n \geq 1, \, 0 = t_0 < t_1 < \cdots < t_n = 1 \right\}.$$

Note that the same definition could be given for functions defined on $[0, 1]$ and valued in X, not necessarily continuous. The functions ω with $\text{Length}(\omega) < +\infty$ are exactly those which have bounded variation in the sense of the BV functions 1D (see Box 4.4 in Section 4.2.4). It is easy to see that all curves $\omega \in \text{AC}(X)$ satisfy $\text{Length}(\omega) \leq \int_0^1 g(t)\mathrm{d}t < +\infty$.

Proposition: For any curve $\omega \in \text{AC}(X)$, we have

$$\text{Length}(\omega) = \int_0^1 |\omega'|(t)\mathrm{d}t.$$

We collect now some more definitions.

Definition: A curve $\omega : [0, 1] \to X$ is said to be a geodesic between x_0 and $x_1 \in X$ if it minimizes the length among all curves such that $\omega(0) = x_0$ and $\omega(1) = x_1$.

A space (X, d) is said to be a *length space* if it holds

$$d(x, y) = \inf\{\text{Length}(\omega) \, : \, \omega \in \text{AC}(X), \, \omega(0) = x, \omega(1) = y\}.$$

A space (X, d) is said to be a *geodesic space* if it holds

$$d(x, y) = \min\{\text{Length}(\omega) \, : \, \omega \in \text{AC}(X), \, \omega(0) = x, \omega(1) = y\},$$

i.e., if it is a length space and there exist geodesics between arbitrary points.

We will not enter into details here about the conditions for the existence of geodesics, but the fact that $\text{Length}(\omega)$ is defined as a sup is crucial so to establish semi-continuity results.

(continued)

Box 5.2. (continued)

Definition: In a length space, a curve $\omega : [0,1] \to X$ is said to be a *constant-speed geodesic* between $\omega(0)$ and $\omega(1) \in X$ if it satisfies

$$d(\omega(t),\omega(s)) = |t-s| d(\omega(0),\omega(1)) \quad \text{for all } t,s \in [0,1].$$

It is easy to check that a curve with this property is automatically a geodesic.

The following characterization is useful.

Proposition: Fix an exponent $p > 1$ and consider curves connecting x_0 to x_1. The three following facts are equivalent:

1. ω is a constant-speed geodesic,
2. $\omega \in \mathrm{AC}(X)$ and $|\omega'|(t) = d(\omega(0),\omega(1))$ a.e.,
3. ω solves $\min \left\{ \int_0^1 |\omega'|(t)^p \mathrm{d}t \,:\, \omega(0) = x_0, \omega(1) = x_1 \right\}$.

First, we prove that the space $\mathbb{W}_p(\Omega)$ is a length space, provided Ω is convex.

Theorem 5.27. *Suppose that Ω is convex, take $\mu, \nu \in \mathscr{P}_p(\Omega)$ and $\gamma \in \Pi(\mu,\nu)$ an optimal transport plan for the cost $c(x,y) = |x-y|^p$ ($p \geq 1$). Define $\pi_t : \Omega \times \Omega \to \Omega$ through $\pi_t(x,y) = (1-t)x + ty$. Then the curve $\mu_t := (\pi_t)_\# \gamma$ is a constant-speed geodesic in $\mathbb{W}_p$ connecting $\mu_0 = \mu$ to $\mu_1 = \nu$.*

In the particular case where μ is absolutely continuous, or in general if $\gamma = \gamma_{\mathrm{T}}$, then this very curve is obtained as $((1-t)\mathrm{id} + t\mathrm{T})_\# \mu$.

As a consequence, the space $\mathbb{W}_p(\Omega)$ is a geodesic space.

Proof. It is sufficient to prove $W_p(\mu_t,\mu_s) \leq W_p(\mu,\nu)|t-s|$. Indeed, suppose this is proven for every $s > t$, then we would have

$$W_p(\mu,\nu) \leq W_p(\mu,\mu_t) + W_p(\mu_t,\mu_s) + W_p(\mu_s,\nu) \leq W_p(\mu,\nu)(t + (s-t) + (1-s)) = W_p(\mu,\nu),$$

which implies equality everywhere. To prove the claim, take $\gamma_t^s := (\pi_t,\pi_s)_\# \gamma \in \Pi(\mu_t,\mu_s)$ and compute

$$W_p(\mu_t,\mu_s) \leq \left(\int |x-y|^p \,\mathrm{d}\gamma_t^s \right)^{\frac{1}{p}} = \left(\int |\pi_t(x,y) - \pi_s(x,y)|^p \,\mathrm{d}\gamma \right)^{\frac{1}{p}}$$
$$= |t-s| \left(\int |x-y|^p \,\mathrm{d}\gamma \right)^{\frac{1}{p}} = |t-s| W_p(\mu,\nu),$$

where we used that $|(1-t)x + ty - (1-s)x - sy| = |(t-s)(x-y)|$. □

Before analyzing the constant-speed geodesics in $\mathbb{W}_p$, we want to insist on an important consequence of the fact that we have a geodesic space and of Theorem 5.14. The idea behind this result, usually called *Benamou-Brenier formula* and considered as a dynamical version of the Kantorovich problem, will be crucial in Section 6.1 to provide a numerical characterization.

Theorem 5.28. *On a convex and compact domain Ω, given $\mu, \nu \in \mathscr{P}(\Omega)$, we have*

$$W_p^p(\mu,\nu) = \min\{\mathscr{B}_p(\varrho, E) \,:\, \partial_t \varrho + \nabla \cdot E = 0,\ \varrho_0 = \mu, \varrho_1 = \nu\},$$

where both ϱ and E have to be considered as measures on the space-time $\Omega \times [0,1]$.

Proof. From the fact that $\mathbb{W}_p$ is a length space, we have

$$W_p^p(\mu,\nu) = \left(\min\left\{\int_0^1 |\varrho'|(t)\mathrm{d}t \,:\, \varrho_0 = \mu, \varrho_1 = \nu\right\}\right)^p.$$

The minimum can be restricted to constant-speed geodesics, and, thanks to the characterization that we gave in Box 5.2, we also have

$$W_p^p(\mu,\nu) = \min\left\{\int_0^1 |\varrho'|(t)^p \mathrm{d}t \,:\, \varrho_0 = \mu, \varrho_1 = \nu\right\}.$$

Then we use Theorem 5.14 which allows to replace $|\varrho'|(t)$ with $||\mathbf{v}_t||_{L^p(\varrho_t)}$, thus obtaining

$$W_p^p(\mu,\nu) = \min\left\{\int_0^1 ||\mathbf{v}_t||^p_{L^p(\varrho_t)}\mathrm{d}t \,:\, \partial_t\varrho_t + \nabla\cdot(\varrho_t\mathbf{v}_t) = 0\ \varrho_0 = \mu, \varrho_1 = \nu\right\}.$$

Setting $E_t = \varrho_t \mathbf{v}_t$, the statement is just a rewriting of the above formula. □

Coming back to the structure of geodesics in $\mathbb{W}_p$, keeping in mind again the important result of Theorem 5.14, one can wonder what is the velocity field associated with these geodesics. Indeed, the constant-speed geodesic curves μ_t are Lipschitz curves, and hence they must admit the existence of a velocity field $\mathbf{v}_t$ (at least if $p > 1$) satisfying the continuity equation $\partial_t\mu_t + \nabla\cdot(\mu_t\mathbf{v}_t) = 0$.

In rough terms, this means: take $y \in \mathrm{spt}(\mu_t) \subset \Omega$, for $t \in]0,1[$, and try to find the velocity of the particle(s) passing at y at time t. This should be the value of $\mathbf{v}_t(y)$. It would be easier to answer this question if we had uniqueness of "the particle" passing through y at time t. To provide this uniqueness, we use the following lemma.

Lemma 5.29. *Let Ω be compact and γ be an optimal transport plan for a cost $c(x,y) = h(y-x)$ with h strictly convex, between two probabilities $\mu, \nu \in \mathscr{P}(\Omega)$ and take $t \in]0,1[$. Define $\mu_t = (\pi_t)_\#\gamma$ with $\pi_t(x,y) = (1-t)x + ty$, and take $y \in \mathrm{spt}(\mu_t)$. Then there exists a unique pair $(x,z) \in \mathrm{spt}(\gamma)$ such that $y = (1-t)x + tz$. These values of x and z will be denoted by $X_t(y)$ and $Z_t(y)$, respectively. The two maps X_t and Z_t are also continuous.*

In particular, if $\gamma = \gamma_{\mathrm{T}}$ comes from a transport map, then the map $\mathrm{T}_t := (1-t)\,\mathrm{id} + t\mathrm{T}$ is invertible and $\mathrm{T}_t^{-1} = X_t$.

Proof. The claim is essentially the same as in Lemma 4.23, and the uniqueness of $(x,z) \in \mathrm{spt}(\gamma)$ comes from c-cyclical monotonicity of $\mathrm{spt}(\gamma)$. The continuity of X_t

and Z_t is obtained by compactness. Take $y_n \to y$ and suppose (up to subsequences) that $(X_t(y_n), Z_t(y_n)) \to (X, Z)$. Since $\mathrm{spt}(\gamma)$ is closed, then $(X, Z) \in \mathrm{spt}(\gamma)$. Moreover, uniqueness implies $(X, Z) = (X_t(y), Z_t(y))$. Since any limit of converging subsequences must coincide with the value at y, and we work in a compact space, this gives continuity. □

We can now identify the velocity field of the geodesic μ_t: we know that every particle initially located at x moves on a straight line with constant speed $\mathrm{T}(x) - x$, which implies $\mathbf{v}_t(y) = (\mathrm{T} - \mathrm{id})(\mathrm{T}_t^{-1}(y))$. More generally, if γ is not induced by a map, we have $\mathbf{v}_t(y) = Z_t(y) - X_t(y)$.

Proposition 5.30. *Let $\mu_t = (\pi_t)_\# \gamma$ be the geodesic connecting μ to ν introduced above. Then the velocity field $\mathbf{v}_t := Z_t - X_t$ is well defined on $\mathrm{spt}(\mu_t)$ for each $t \in]0, 1[$ and satisfies*

$$\partial_t \mu_t + \nabla \cdot (\mu_t \mathbf{v}_t) = 0, \quad ||\mathbf{v}_t||_{L^p(\mu_t)} = |\mu'|(t) = W_p(\mu, \nu).$$

Proof. We already saw that X_t and Z_t are well defined, so we only need to check the continuity equation and the L^p estimate. To prove the continuity equation, take $\phi \in C^1$ and compute

$$\begin{aligned}\frac{d}{dt}\int \phi \, \mathrm{d}\mu_t &= \frac{d}{dt}\int \phi((1-t)x + tz)\mathrm{d}\gamma(x,z) = \int \nabla\phi((1-t)x+tz)\cdot(z-x)\mathrm{d}\gamma(x,z)\\ &= \int \nabla\phi(\pi_t(x,z))\cdot(Z_t(\pi_t(x,z)) - X_t(\pi_t(x,z)))\mathrm{d}\gamma(x,z)\\ &= \int \nabla\phi(y)\cdot(Z_t(y) - X_t(y))\mathrm{d}\mu_t(y).\end{aligned}$$

To compute the L^p norm, we have

$$\int |\mathbf{v}_t|^p \, \mathrm{d}\mu_t = \int |Z_t(y) - X_t(y)|^p \, \mathrm{d}\mu_t(y) = \int |z - x|^p \, \mathrm{d}\gamma(x,z) = W_p^p(\mu, \nu),$$

and we used in both the computations the fact that $Z_t(\pi_t(x,z)) - X_t(\pi_t(x,z)) = z - x$ for every $(x, z) \in \mathrm{spt}(\gamma)$. □

We now want to prove that, at least for $p > 1$, all the geodesics for the distance W_p have this form, i.e., they are given by $\mu_t = (\pi_t)_\# \gamma$ for an optimal γ. Non-uniqueness would stay true in case γ is not unique. The proof that we provide here is different from that of [15] and is based on the following representation theorem (in the spirit of [212]). We use the set $\mathrm{AC}(\Omega)$ of absolutely continuous curves defined on $[0, 1]$ and valued in Ω, and we denote by $K_p : \mathrm{AC}(\Omega) \to \mathbb{R} \cup \{+\infty\}$ the p-Kinetic energy, given by $K_p(\omega) := \int_0^1 |\omega'(t)|^p \mathrm{d}t$ (which is only finite if ω belongs to the Sobolev space $W^{1,p}([0, 1]) \subset \mathrm{AC}(\Omega)$). As usual, we write $\mathscr{C}$ for $\mathrm{AC}(\Omega)$, for short.

Proposition 5.31. *For every Lipschitz curve $(\mu_t)_t$ in $\mathbb{W}_p(\Omega)$, there exists a measure $Q \in \mathscr{P}(\mathscr{C})$ such that $\mu_t = (e_t)_\# Q$ and $\int_{\mathscr{C}} K_p(\omega)\,\mathrm{d}Q(\omega) \le \int_0^1 |\mu'|(t)^p\,\mathrm{d}t$.*

Proof. First we note that for a.e. t, there exists, thanks to Theorem 5.14, a vector field $\mathbf{v}_t \in L^p(\mu_t)$ such that $\partial_t \mu_t + \nabla \cdot (\mathbf{v}_t \mu_t) = 0$, and with $||\mathbf{v}_t||_{L^p(\mu_t)} \le |\mu'|(t)$. The same approximation argument (with the same convolution kernels) as in Theorem 5.14 provides a regularized curve of measures μ_t^ε and a regularized vector field $\mathbf{v}_t^\varepsilon$, with $||\mathbf{v}_t^\varepsilon||_{L^p(\mu_t^\varepsilon)} \le ||\mathbf{v}_t||_{L^p(\mu_t)}$. This vector field has a flow $Y : \Omega \to \mathrm{AC}(\Omega)$, where $Y(x)$ is the curve $y_x(\cdot)$, and we know $\mu_t^\varepsilon = (Y_t)_\# \mu_0^\varepsilon$, where $Y_t(x) = Y(x)(t) = y_x(t)$. Let us also define the measure $Q^\varepsilon := Y_\# \mu_0^\varepsilon$. We have $\mu_t^\varepsilon = (e_t)_\# Q^\varepsilon$.

Let us compute $\int_{\mathscr{C}} K_p\,\mathrm{d}Q^\varepsilon$:

$$\int_{\mathscr{C}} K_p(\omega)\,\mathrm{d}Q^\varepsilon(\omega) = \int_\Omega \int_0^1 |\mathbf{v}_t^\varepsilon(y_x(t))|^p\,\mathrm{d}t\,\mathrm{d}\mu_0^\varepsilon(x) = \int_0^1 \int_\Omega |\mathbf{v}_t^\varepsilon|^p\,\mathrm{d}\mu_t^\varepsilon$$
$$\le \int_0^1 \int_\Omega |\mathbf{v}_t|^p\,\mathrm{d}\mu_t = \int_0^1 |\mu'|(t)^p\,\mathrm{d}t.$$

Now (as we saw many times in Chapter 4), this implies that the measures Q^ε are tight, since there is a bound on $\int_{\mathscr{C}} K_p\,\mathrm{d}Q^\varepsilon$ and K_p is such that

$$\{\omega \in \mathrm{AC}(\Omega) \,:\, K_p(\omega) \le L\}$$

is compact in $C^0([0,1])$ (for the uniform convergence) for every L. Also, K_p is l.s.c. for the same convergence. Hence, we can obtain the existence of a subsequence such that $Q^\varepsilon \rightharpoonup Q$ and $\int_{\mathscr{C}} K_p \mathrm{d}Q \le \liminf_\varepsilon \int_{\mathscr{C}} K_p \mathrm{d}Q^\varepsilon$. From $\mu_t^\varepsilon = (e_t)_\# Q^\varepsilon$ we get $\mu_t = (e_t)_\# Q$ and we conclude. □

Proposition 5.32. *Let $(\mu_t)_t$ be a constant-speed geodesic in $\mathbb{W}_p(\Omega)$ between μ and ν and suppose $p > 1$. Then there exists an optimal $\gamma \in \Pi(\mu,\nu)$ for the transport cost $c(x,y) = |x-y|^p$ such that for every $t \in [0,1]$, we have $\mu_t = (\pi_t)_\# \gamma$.*

Proof. Since the curve is a constant-speed geodesic, we have $|\mu'|(t) = W_p(\mu,\nu)$ for a.e. t (the metric derivative being computed according to W_p). Applying our previous Proposition 5.31, we get the existence of a measure $Q \in \mathscr{P}(\mathscr{C})$ such that $\mu_t = (e_t)_\# Q$ and $\int_{\mathscr{C}} K_p(\omega)\,\mathrm{d}Q(\omega) \le \int_0^1 |\mu'|(t)^p\,\mathrm{d}t$.

From Jensen inequality, we have $|\omega(0) - \omega(1)|^p \le K_p(\omega)$, with equality if and only if ω is a segment parametrized with constant speed. Hence, we have

$$W_p^p(\mu,\nu) \le \int_{\mathscr{C}} |\omega(0) - \omega(1)|^p\,\mathrm{d}Q(\omega)$$
$$\le \int_{\mathscr{C}} K_p(\omega)\,\mathrm{d}Q(\omega) \le \int_0^1 |\mu'|(t)^p\,\mathrm{d}t \le W_p^p(\mu,\nu).$$

Fig. 5.2 Nine time steps of the interpolation (top, left to right, and then bottom, left to right), obtained via a $\mathbb{W}_2$ geodesic curve, between the histograms on the 2D space of chroma signals of the two images of Figure 2.4. Pictures kindly provided by G. Peyré

The first inequality comes from $\gamma := (e_0, e_1)_\# Q \in \Pi(\mu_0, \mu_1)$; the second is a consequence of the above inequality for K_p; the last comes from the properties of Q. Thus, all inequalities are equalities. Hence Q is concentrated on curves which are constant-speed segments and γ is an optimal transport plan. This proves $\mu_t = (\pi_t)_\# \gamma$. □

Note how the above proof is reminiscent of that of Theorem 4.13, with here a dynamical framework, compared to the static framework in Chapter 4.

It is not difficult to see that both Propositions 5.31 and 5.32 do not hold for $p = 1$. The counterexample is, as usual, given by $\mu_t = (1-t)\delta_{x_0} + t\delta_{x_1}$.

We finish this section insisting on the role that Wasserstein geodesics play in applications, as they allow to build an interesting interpolation between two distribution of masses (or two histograms), moving slowly the particles from one distribution to the other. For instance, Figure 5.2 shows the interpolation between the chrominance histograms of Figure 2.4. These intermediate histograms could also be used to produce intermediate images.

5.5 Discussion

5.5.1 The W_∞ distance

The case $p = \infty$ in the definition of W_p deserves special attention. It is easy to guess how to define a W_∞ distance on probability measures:

$$W_\infty(\mu, \nu) := \inf \left\{ ||x - y||_{L^\infty(\gamma)} \,:\, \gamma \in \Pi(\mu, \nu) \right\},$$

where the minimization problem in the right-hand side is exactly the one studied in Section 3.2. As we did in the rest of this chapter, the difference here is that, instead of paying attention to the optimal map, we mainly consider the optimal cost as a function of the two measures and use it as a distance on the space of probabilities.

In order to properly define $W_\infty(\mu,\nu)$, one should restrict to $\mathscr{P}_\infty(\Omega)$, which is the space of measures with bounded support. Otherwise, there could be pairs of measures such that no bounded-displacement plan γ exists. From the results in Section 3.2, we know that, if $\mu \ll \mathscr{L}^d$, then $W_\infty(\mu,\nu)$ is also equal to

$$\inf\left\{||\mathrm{T}(x)-x||_{L^\infty(\mu)} \,:\, \mathrm{T}_\#\mu=\nu\right\}.$$

It is not difficult to check the expected property

$$W_\infty(\mu,\nu)=\lim_{p\to\infty} W_p(\mu,\nu),$$

where the right-hand side is increasing in p (the proof is proposed as an exercise, **Ex**(35)). The triangle inequality can be proved by passing to the limit in p in the corresponding triangle inequality for W_p.

Note that, in dimension one, as a consequence of the results of Chapter 2, the optimal way to displace particles is the monotone increasing map, for every cost $|x-y|^p$ with $p<\infty$. Passing to the limit, this is also the optimal map for the L^∞ problem.

In general, we have a new distance on probability measures on compact sets, and it measures the minimal maximal displacement that should be done to move particles from one distribution to the other. It is interesting to see that, though the modeling which gives rise to this distance is the same as in the case $p<\infty$, the topology induced by it is not at all the same.

Indeed, if one takes two different points $x,y\in\Omega$ and two different values $s,t\in[0,1]$, one can easily check $W_\infty((1-t)\delta_x+t\delta_y,(1-s)\delta_x+s\delta_y)=|x-y|$. In particular, the probabilities $(1-t)\delta_x+t\delta_y$ do not tend to δ_x as $t\to 0$, even if there is convergence for all the distances W_p and even for the strong convergence of measures. As a matter of fact, even if Ω is compact, the space $(\mathscr{P}(\Omega),W_\infty)$ is neither compact, nor separable, as soon as the cardinality of Ω is larger than 1.

On the other hand, the characterization of Lipschitz curves via the existence of a velocity field $\mathbf{v}_t$ satisfying $||\mathbf{v}_t||_{L^\infty(\mu_t)}\le|\mu'|_{W_\infty}(t)$ stays true and can be easily deduced from Theorem 5.14 as a limit $p\to\infty$. This fact, together with an extended use of Lipschitz curves for W_∞, has been used in [79] in order to analyze an alternative model for branched transport. If the role of this model is very limited, some examples and considerations in the paper are useful to understand the structure of Wasserstein spaces.

However, the distance W_∞ appears naturally in many modeling issues and happens to be very useful to analyze very different phenomena. One of the first papers using such a distance was [232], devoted to the shape of rotating stars. In such a paper, the choice of W_∞ was motivated by the following observation (that

we will develop better in Section 7.4.3): a measure μ is a local minimizer in $\mathbb{W}_\infty$ of some functional F if and only if every point $x \in \mathrm{spt}(\mu)$ satisfies a local minimality condition for a function f (equal to the first variation of F, see Section 7.2). This equivalence fails if we replace W_∞ with W_p.

Finally, it is interesting to mention a reverse inequality between W_∞ and W_p (which could not be deduced from Equation (5.1)), of the form

$$W_\infty(\mu,\nu) \le C W_p^{\frac{p}{p+d}}(\mu,\nu),$$

where the constant C depends on p, on the dimension d, on the domain Ω, and on a lower bound on the density of μ. This inequality is a consequence of a stronger result proved in [74]: if $\mu = f \cdot \mathscr{L}^d, f \ge a > 0$ and T is the optimal transport map from μ to another measure ν, then we have

$$||\mathrm{T} - \mathrm{id}||_{L^\infty}^{p+d} \le C(\Omega,p,d) \int_\Omega |\mathrm{T}(x) - x|^p \,\mathrm{d}x \le \frac{C(\Omega,p,d)}{a} \int_\Omega |\mathrm{T}(x) - x|^p \,\mathrm{d}\mu(x).$$

This is proven by using the c-cyclical monotonicity condition on T so as to guarantee that, whenever we fix a point x with $|\mathrm{T}(x) - x| > 0$, there exists a certain region R around x, with volume proportional to $|\mathrm{T}(x) - x|^d$, such that for every $y \in R$, the distance $|\mathrm{T}(y) - y|$ is not smaller than a certain fraction of $|\mathrm{T}(x) - x|$.

5.5.2 *Wasserstein and negative Sobolev distances*

We already noted that the distance W_1 is induced by a norm, which is the dual of the Lipschitz seminorm. Indeed, if we consider the space $\mathscr{X}$ of Lipschitz functions with zero mean, endowed with the norm $||f|| := \mathrm{Lip}(f)$, the distance W_1 on $\mathscr{P}_1(\Omega)$ is exactly the distance associated with the norm of the topological dual $\mathscr{X}'$. It is interesting to note that the space of probability measures with this distance is complete, but the space of finite measures endowed with the same norm is not. More precisely, one can consider zero-mean finite scalar measures endowed with the norm $||\varrho|| := \sup\{\int f \,\mathrm{d}\varrho : f \in \mathrm{Lip}_1(\Omega)\}$ and look at its completion. Note that the sequence $\varrho_n := \sum_{i=1}^n \delta_{1/2i} - \delta_{1/(2i-1)}$ is Cauchy for this norm, but its limit is an infinite sum of Dirac masses, which is not a finite scalar measure. The completion of the set of measures for this norm has been studied in [73], where the authors prove that it amounts to the set of distributions which are the divergence of an L^1 vector field. This nice result makes an interesting link with the L^1-Beckmann problem studied in Chapter 4.

Anyway, once the role of the distance W_1 in the duality with $W^{1,\infty}$ functions is clear, one could wonder whether other distances W_p are in duality with Sobolev spaces of the form $W^{1,q}$ (for instance, with $q = p' = p/(p-1)$). The answer is obviously "not," since not all probability measures belong to the dual space $W^{-1,p}$

(in particular, Dirac masses only belong to the dual of $W^{1,q}$ when functions in $W^{1,q}$ are continuous, i.e., when $q > d$). Yet, some partial answers and some analogies exist. For instance, the following lemma was both proven in [214] and later in [225] for different purposes.

Lemma 5.33. *Assume that μ and ν are absolutely continuous measures on a convex domain Ω, whose densities are bounded by the same constant C. Then, for all function $\phi \in H^1(\Omega)$, we have the following inequality:*

$$\int_\Omega \phi \, \mathrm{d}(\mu - \nu) \;\leq\; \sqrt{C}\, ||\nabla \phi||_{L^2(\Omega)} W_2(\mu, \nu)$$

Proof. Let μ_t be the constant-speed geodesic between μ and ν, and let $\mathbf{v}_t$ be the velocity field associated with this curve by Theorem 5.14. Then $(\mu, \mathbf{v})$ satisfies the continuity equation, and $||\mathbf{v}_t||_{L^2(\mu_t)} = W_2(\mu, \nu)$. We use the fact that μ_t is absolutely continuous for all t, and its density is bounded by the same constant C, which will be proven in Chapter 7 (Prop 7.29). Therefore:

$$\begin{aligned}
\int_\Omega \phi \, \mathrm{d}(\mu - \nu) &= \int_0^1 \frac{d}{dt} \left(\int_\Omega \phi(x) \mathrm{d}\mu_t(x) \right) \mathrm{d}t \;= \int_0^1 \int_\Omega \nabla \phi \cdot \mathbf{v}_t \, \mathrm{d}\mu_t \, \mathrm{d}t \\
&\leq \left(\int_0^1 \int_\Omega |\nabla \phi|^2 \, \mathrm{d}\mu_t \, \mathrm{d}t \right)^{1/2} \left(\int_0^1 \int_\Omega |\mathbf{v}_t|^2 \, \mathrm{d}\mu_t \, \mathrm{d}t \right)^{1/2} \\
&\leq \sqrt{C}\, ||\nabla \phi||_{L^2(\Omega)} W_2(\mu, \nu),
\end{aligned}$$

and the proof is completed. □

We do not develop here the case of different exponents than $p = p' = 2$, since the quadratic case seems to be the most interesting (see **Ex**(38) for the general case). The exact negative Sobolev norm appearing in the quadratic case is the $\dot{H}^{-1}$ norm, which deserves some remarks. First let us define

$$||\mu - \nu||_{\dot{H}^{-1}(\Omega)} = \sup \left\{ \int \phi \, \mathrm{d}(\mu - \nu) \;:\; ||\nabla \phi||_{L^2} \leq 1 \right\}.$$

Note that we also have $||\mu - \nu||_{\dot{H}^{-1}(\Omega)} = ||\nabla u||_{L^2}$, where u is the solution of the Neumann problem

$$\begin{cases} -\Delta u = \mu - \nu, & \text{in } \Omega, \\ \frac{\partial u}{\partial \mathbf{n}} = 0 & \text{on } \partial\Omega. \end{cases}$$

Indeed, for this choice of u, we have $\int \phi \, \mathrm{d}(\mu - \nu) = \int \nabla \phi \cdot \nabla u \leq ||\nabla \phi||_{L^2} ||\nabla u||_{L^2}$, with equality for $\phi = u / ||\nabla u||_{L^2}$. We can also see that we have

$$||\mu-\nu||_{\dot{H}^{-1}(\Omega)} = \min\{||\mathbf{w}||_{L^2} \,:\, \nabla\cdot\mathbf{w} = \mu-\nu\},$$

since the optimal $\mathbf{w}$ in this minimization problem is a gradient (see Section 4.4.1 and the considerations on Problem (4.27)), and we have $\mathbf{w} = -\nabla u$.

The interest for the relation between the distances W_2 and $\dot{H}^{-1}$ (where $\dot{H}^{-1}$ denotes the dual of the space $\mathscr{X}$ of zero-mean H^1 functions on a connected domain, endowed with the L^2 norm of its gradient) is now very lively because these two distances are asymptotically equivalent when the involved densities are close to the value $\varrho = 1$, which is the relevant case when dealing with evolution problems where congestion imposes an L^∞ constraint $\varrho \le 1$, as in crowd motion (see Section 8.4.2). These asymptotical equivalences can be made precise through some general estimates. We cite a short note by R. Peyre [251] where some universal estimates are proven and a funny equality in dimension 1 (see **Ex**(64)). Moreover, we also prove the following result.

Theorem 5.34. *Assume that μ and ν are absolutely continuous measures on a convex domain Ω, with densities bounded from below and from above by two (same) constants a, b with $0 < a < b < +\infty$. Then, we have*

$$b^{-1/2}||\mu-\nu||_{\dot{H}^{-1}(\Omega)} \le W_2(\mu,\nu) \le a^{-1/2}||\mu-\nu||_{\dot{H}^{-1}(\Omega)}.$$

Proof. The inequality $b^{-1/2}||\mu-\nu||_{\dot{H}^{-1}(\Omega)} \le W_2(\mu,\nu)$ is a consequence of Lemma 5.33 and of the definition of the $\dot{H}^{-1}$ norm.

To prove that the opposite inequality can be obtained through the Benamou-Brenier formula of Theorem 5.28, consider any vector field $E \in \mathscr{M}^d_{\rm div}(\Omega)$ with $\nabla\cdot E = \mu-\nu$ and $\varrho_t := (1-t)\mu + t\nu$. The pair (ϱ_t, E_t), where $E_t = E$ is taken independent of time, is admissible in the Benamou-Brenier formula. If $E \in L^2(\Omega;\mathbb{R}^d)$, we have

$$W_2^2(\mu,\nu) \le \mathscr{B}_2(\varrho,E) = \int_0^1\int_\Omega \frac{|E|^2}{\varrho_t} \le \frac{1}{a}\int_\Omega |E|^2,$$

where the last inequality is justified by $\varrho_t = (1-t)\mu + t\nu \ge (1-t)a + ta = a$. Since E is arbitrary with the only constraint $\nabla\cdot E = \mu-\nu$, we obtain

$$W_2^2(\mu,\nu) \le \frac{1}{a}\min\left\{||\mathbf{w}||_{L^2}^2 \,:\, \nabla\cdot\mathbf{w} = \mu-\nu\right\} = \frac{1}{a}||\mu-\nu||^2_{\dot{H}^{-1}(\Omega)},$$

which gives the second part of the statement. □

5.5.3 Wasserstein and branched transport distances

Another class of distances which has been compared to the Wasserstein distances is that induced by branched transport. We recall that $d_\alpha(\mu,\nu)$ is defined as the minimal

value of the branched transport cost (see Section 4.4.2) between two measures μ, ν, as soon as $\alpha > 1 - \frac{1}{d}$. In this case, it is a distance, and it metrizes the weak convergence of probabilities on a bounded domain.

In branched transport, the cost per unit length to move a mass m is given by m^α, and we can assume $m \leq 1$ (because there are no cycles in the optimal networks). Hence, one immediately remarks $m^\alpha \geq m$, which implies $d_\alpha \geq d_1$, and the distance d_1 is nothing but W_1.

Since the two distances d_α and W_1 satisfy an inequality, and they induce the same topology, it was a question raised by C. Villani whether the two distances are metrically equivalent, i.e., we can bound above d_α with W_1. Inequalities of the form $d_\alpha \leq CW_1$ are necessarily false, since they do not scale in the correct way w.r.t. masses. Indeed, even if we stick to probability measures, we can consider the following situation: take $\mu, \nu \in \mathscr{P}(\Omega_1)$ and a third measure ϱ concentrated on another domain Ω_2, disjoint from Ω_1. Then take $\mu_\varepsilon = (1-\varepsilon)\varrho + \varepsilon\mu$ and $\nu_\varepsilon = (1-\varepsilon)\varrho + \varepsilon\nu$. Since both W_1 and d_α only depend on the difference between the two measures (which is false for W_p, with $p > 1$), by a scaling argument it is clear that we have $W_1(\mu_\varepsilon, \nu_\varepsilon) = c_1\varepsilon$ and $d_\alpha(\mu_\varepsilon, \nu_\varepsilon) = c_\alpha\varepsilon^\alpha$. This rules out the possibility of a linear bound but leaves the possibility of an inequality of the form $d_\alpha \leq CW_1^\beta$ for $\beta \leq \alpha$.

This question has been analyzed in [240], where the sharp exponent has been found, $\beta = d(\alpha - (1 - \frac{1}{d}))$. Indeed, a simple example shows that it cannot be improved.

Example 5.35 (Optimal exponent for $d_\alpha \leq W_1^\beta$). Divide the cube $[0,1]^d$ into n^d small cubes of edge $1/n$ and set $\mu_n = \sum_{i=1}^{n^d} \frac{1}{n^d}\delta_{x_i}$ and $\nu_n = \sum_{i=1}^{n^d} \frac{1}{n^d}\delta_{y_i}$, where each x_i is a vertex of one of the n^d cubes (let us say the vertex with minimal sum of the d coordinates) and the corresponding y_i is the center of the same cube. In this way, the point y_j closest to x_i is exactly y_i. Thus the optimal configuration both for d_α and W_1 is obtained by linking any x_i directly to the corresponding y_i, and we have

$$d_\alpha(\mu_n, \nu_n) = n^d \left(\frac{1}{n^d}\right)^\alpha \frac{c}{n} = cn^{-d(\alpha-(1-1/d))}$$

$$W_1(\mu_n, \nu_n) = n^d \frac{1}{n^d}\frac{c}{n} = cn^{-1}$$

The proof of the inequality $d_\alpha \leq CW_1^{d(\alpha-(1-1/d))}$ is a bit technical. Yet, in order to give an idea, we just sketch an alternative proof (from [81]), which provides a slightly weaker inequality, namely, $d_\alpha \leq CW_p^{d(\alpha-(1-1/d))}$ for $p = 1/\alpha$. This inequality is also optimal in some sense, as it is also possible to strengthen the lower bound $d_\alpha \geq W_1$ turning it into $d_\alpha \geq W_{1/\alpha}$ (this has been first remarked in [220] and then in [81]). The choice of the exponent p is very natural because of scaling reasons. Indeed, W_p scales as $m^{\frac{1}{p}}$ w.r.t. the mass (in the example above, if Ω_1 and Ω_2 are far enough so that the measure ϱ becomes useless, then we have $W_p(\mu_\varepsilon, \nu_\varepsilon) = c\varepsilon^{\frac{1}{p}}$), which makes the choice $p = 1/\alpha$ the good one.

Proposition 5.36. *Let μ, ν be two probabilities on $[0,1]^d$ and $\alpha > 1 - \frac{1}{d}$, $p = 1/\alpha$. Then*

$$W_p \le d_\alpha \le C W_p^{d(\alpha-(1-1/d))}.$$

Proof. First we prove the inequality $W_p \le d_\alpha$. This is a consequence of (4.33). Indeed, W_p is a distance, it is l.s.c. for $\rightharpoonup$, and we have

$$\sum_{ij} |x_i - y_j| \gamma(\{(x_i, y_j)\})^\alpha \ge \left(\sum_{ij} |x_i - y_j|^p \gamma(\{(x_i, y_j)\}) \right)^\alpha.$$

This inequality is due to the choice $p = 1/\alpha$ and to the subadditivity of $t \mapsto t^\alpha$. Passing to the minimum over $\gamma \in \Pi(\mu, \nu)$, we get $\bar{d}_\alpha(\mu, \nu) \ge W_p(\mu, \nu)$, whenever μ and ν are atomic. This implies $d_\alpha \ge W_p$.

In order to prove the other inequality, we divide the cube into n^d small disjoint cubes Q_i of edge $1/n$ and set

$$\mu_n = \sum_{i=1}^{n^d} \frac{1}{n^d} \mu(Q_i) \delta_{y_i} \quad \text{and} \quad \nu_n = \sum_{i=1}^{n^d} \nu(Q_i) \delta_{y_i},$$

where each point y_i is the center of the cube Q_i. Since the branched transport from $\mu \llcorner Q_i$ to $\mu(Q_i)\delta_{y_i}$ costs no more than $C\mu(Q_i)^\alpha \frac{1}{n}$ (see Proposition 4.38), we can estimate

$$\begin{aligned} d_\alpha(\mu, \nu) &\le d_\alpha(\mu, \mu_n) + d_\alpha(\mu_n, \nu_n) + d_\alpha(\nu_n, \nu) \\ &\le \frac{C}{n} \left(\sum_i \mu(Q_i)^\alpha + \sum_i \nu(Q_i)^\alpha \right) + d_\alpha(\mu_n, \nu_n) \le \frac{C}{n} n^{d(1-\alpha)} + d_\alpha(\mu_n, \nu_n), \end{aligned}$$

where we used the inequality $\sum_{j=1}^k a_j^\alpha \le k(\frac{1}{k})^\alpha$, which is valid whenever $\sum_{j=1}^k a_j = 1$ and $\alpha \in]0,1[$ (it is just a version of Jensen inequality).

Then, let $\gamma \in \Pi(\mu_n, \nu_n)$ be an optimal transport plan for the cost $|x-y|^p$. We have $\gamma = \sum_{i,j} \gamma_{ij} \delta_{(x_i, x_j)}$, and we can decide to use segments connecting each point x_i to each point x_j with mass γ_{ij} as a (nonoptimal) branched transport network. Its cost is $\sum_{i,j} \gamma_{ij}^\alpha |x_i - x_j|$. We denote by K the number of pairs (i,j) such that $\gamma_{ij} > 0$. Since γ is optimal in a linear programming problem, it may be taken extremal in the polyhedron of matrices $\{(\eta_{ij})_{i,j} : \eta_{ij} \ge 0, \sum_i \eta_{ij} = \nu(Q_j), \sum_j \eta_{ij} = \mu(Q_i)\}$. For these matrices, it is well known that the number of nonzero entries does not exceed the sum of the number of rows and columns (left as an exercise, **Ex**(41)), here $2n^d$. Hence we have

$$d_\alpha(\mu_n,\nu_n) \le \sum_{i,j} \gamma_{ij}^\alpha |x_i - x_j| \le \left(\sum_{i,j} \gamma_{ij}|x_i - x_j|^p\right)^{\frac{1}{p}} (2n^d)^{1-\frac{1}{p}}$$
$$= CW_p(\mu_n,\nu_n)n^{d(1-\alpha)} \le C\left(W_p(\mu,\nu) + \frac{2}{n}\right) n^{d(1-\alpha)},$$

where we used the triangle inequality on W_p together with $W_p(\mu,\mu_n), W_p(\nu,\nu_n) \le 1/n$. By putting together the estimates, one gets

$$d_\alpha(\mu,\nu) \le Cn^{d(1-\alpha)-1} + Cn^{d(1-\alpha)}W_p(\mu,\nu)$$

and we can conclude by taking $n \approx W_p(\mu,\nu)^{-1}$. □

5.5.4 The sliced Wasserstein distance

Starting from the considerations of Section 2.5.2, it is possible, and convenient in many cases, to define a distance, which is alternative to the usual W_2 distance, based on the behavior of the measures "direction by direction."

The idea is the following and is detailed in [256]: given two measures $\mu,\nu \in \mathscr{P}_2(\mathbb{R}^d)$, we define

$$SW_2(\mu,\nu) := \left(\fint_{\mathbb{S}^{d-1}} W_2^2((\pi_e)_\#\mu,(\pi_e)_\#\nu)\,\mathrm{d}\mathscr{H}^{d-1}(e)\right)^{1/2},$$

where $\pi_e : \mathbb{R}^d \to \mathbb{R}$ is the projection on the axis directed according to the unit vector e, namely, $\pi_e(x) = x \cdot e$, and $\mathscr{H}^{d-1}$ is the surface measure on $\mathbb{S}^{d-1}$. This quantity could have been called "projected Wasserstein distance" (as it is based on the behavior through projections), but since in [256] it is rather called "sliced Wasserstein distance," we prefer to keep the same terminology.

The fact that SW_2 is a distance comes from W_2 being a distance. The triangle inequality may be proven using the triangle inequality for W_2 (see Section 5.1) and for the L^2 norm. Positivity and symmetry are evident. The equality $SW_2(\mu,\nu) = 0$ implies $W_2^2((\pi_e)_\#\mu,(\pi_e)_\#\nu)$ for all $e \in \mathbb{S}^{d-1}$. This means $(\pi_e)_\#\mu = (\pi_e)_\#\nu)$ for all e and it suffices (see Box 2.4 in Section 2.5.2 on Radon and X-ray transforms) to prove $\mu = \nu$.

It is evident from its definition and from the fact that the maps π_e are 1-Lipschitz (hence $W_2^2((\pi_e)_\#\mu,(\pi_e)_\#\nu) \le W_2^2(\mu,\nu)$) that we have $SW_2(\mu,\nu) \le W_2(\mu,\nu)$. Moreover, the two distances also induce the same topology, at least on compact sets. Indeed, the identity map from $\mathbb{W}_2$ to $(\mathscr{P}(\Omega), SW_2)$ is continuous (from $SW_2 \le W_2$) and bijective. Since the space where it is defined is compact, it is also a homeomorphism.

One can also prove more, i.e., an inequality of the form $W_2 \leq C\mathrm{SW}_2^\beta$ for a suitable exponent $\beta \in]0,1[$. Chapter 5 in [65] proves this inequality[7] with $\beta = (2(d+1))^{-1}$.

The interest in the use of this distance is the fact that one has a distance on $\mathscr{P}(\Omega)$ with very similar qualitative properties as W_2, but much easier to compute, since it only depends on 1D computations (obviously, the integral over $e \in \mathbb{S}^{d-1}$ is discretized in practice and becomes an average over a large number of directions). For instance, it is useful when the Wasserstein distance is used to evaluate distances between mass distributions (or images, where each image is associated with a measure on $\mathbb{R}^5$ by associating with each pixel its two spatial coordinates and its three color components), or when one looks at the "horizontal" barycenter of several distributions, as in Section 5.5.5 here below (in the Euclidean space, the barycenter of $x_1, \ldots, x_N$ is the point x which minimizes $\sum_i |x - x_i|^2$; in the Wasserstein space, the measure minimizing $\sum_i W_2^2(\varrho, \varrho_i)^2$ is the one which is located "at the middle" of $\varrho_1, \ldots, \varrho_N$). In both cases, replacing W_2 with SW_2 simplifies a lot the computations. We remark anyway an important difference between W_2 and SW_2: the latter is not a geodesic distance. On the contrary, the geodesic distance associated with SW_2 (i.e., the minimal length to connect two measures) is exactly W_2.

We will also see in Section 8.4.2 that the gradient flow of SW_2^2 has an interest in itself, and in the very next Section 5.5.5, we see the applications of SW_2^2 to barycenter computations.

5.5.5 *Barycenters in* $\mathbb{W}_2$

An interesting problem in applications, and in particular in image processing, is that of finding a reasonable notion of barycenter between two or more probability measures. Suppose that we have two B&W images, representing similar objects, but one being very dark (almost black) and one very pale (almost white). We denote by ϱ^0 and ϱ^1 their color histograms (see Section 2.5.1): what is the middle point between these two histograms? If we answer $(\varrho^0 + \varrho^1)/2$, then we obtain an image with some black and some white, while we could prefer to have an intermediate grey image. Also, we would like to provide a definition of middle point such that the middle point of δ_{x_0} and δ_{x_1} is $\delta_{x_{1/2}}$, with $x_{1/2} = (x_0 + x_1)/2$.

The notion of geodesic in Wasserstein spaces provides what we look for: just define the middle point as $\varrho_{1/2}$, where ϱ_t is the constant-speed geodesic between ϱ^0 and ϱ^1, according to what we saw in Section 5.4. Also, we could look for a weighted barycenter, with weights $(1-t)$ and t, and the geodesic provides the correct answer. Some nonuniqueness issues subsist: geodesics are not always unique, and moreover they depend on the exponent p. We will soon fix the exponent issue by choosing $p = 2$, as we are going to see in a while.

[7] Indeed, one can define sliced Wasserstein distances associated with other exponents p, and [65] proves $W_1 \leq C\mathrm{SW}_1^{1/(d+1)}$; then, one can compose with $W_2 \leq CW_1^{1/2}$ to obtain the result for $p = 2$.

What about a barycenter between three or more measures? If we want to make use of geodesics, we could define the barycenter of $(\varrho^0, \varrho^1, \varrho^2)$, with weights $(1/3, 1/3, 1/3)$, as the barycenter with weights $(1/3, 2/3)$ between ϱ^2 and the middle point of ϱ^0 and ϱ^1. Unfortunately, in the space $\mathbb{W}_2$ in dimension larger than 1, there is no associativity, and this definition gives different results if we interchange the order of the measures ϱ^i.

Hence, it is better to switch to another characterization of the barycenter in Euclidean spaces, which we will translate into the framework of Wasserstein spaces. We note that the barycenter with weights $\lambda_i \geq 0$ between points $x_i \in \mathbb{R}^d$ is the unique point y minimizing $\sum_i \lambda_i |y - x_i|^2$. Hence, we could define the barycenter between ϱ^i with weights λ_i as any measure solving

$$\min \left\{ \sum_{i=1}^N \lambda_i W_2^2(\varrho, \varrho^i)^2 \ : \ \varrho \in \mathscr{P}(\Omega) \right\} . \tag{5.7}$$

Note that the same approach could be applied to general Riemannian manifolds, as the barycenter is otherwise not well defined on nonlinear structures. This explains also the choice of the exponent $p = 2$: using powers p of the distances, we do not recover the usual barycenter in the Euclidean space (for $p = 1$, we have a notion of median), and using the square of a W_p distance is not very convenient. The uniqueness of the barycenter will be true as soon as one of the measures ϱ^i is absolutely continuous (see Proposition 7.19).

As an interesting example, it is not difficult to analyze the 1D case. In this case, one can find the barycenter of some measures ϱ^i with weights λ_i in the following way: pick one of these measures, say ϱ^1, define the optimal transport maps T_i sending ϱ^1 onto ϱ^i (i.e., the monotone maps; note that this would work for any exponent p), and set $\varrho := (\sum_i \lambda_i \mathrm{T}_i)_\# \varrho^1$, where we set $\mathrm{T}_1 = \mathrm{id}$ (see also **Ex**(63)).

The minimization problem (5.7) has been first introduced in [1], where the authors also propose an interpretation in terms of a multi-marginal optimal transport problem. Indeed, one can choose as a variable, instead of $\varrho \in \mathscr{P}(\Omega)$, a multi-marginal transport plan $\gamma \in \mathscr{P}(\Omega^{N+1})$, with $(\pi_i)_\# \gamma = \varrho^i$ for $i = 1, \dots N$ and $(\pi_0)_\# \gamma = \varrho$ (the latter ϱ being part of the unknown). The problem becomes

$$\min \int (\sum_{i=1}^N \lambda_i |x_i - x_0|^2) \, \mathrm{d}\gamma,$$

under the constraint that the marginals of γ for $i = 1, \dots N$ are the ϱ^i. The other marginal being arbitrary, one could decide to choose x_0 as a function of the points $(x_1, \dots, x_N)$, in an optimal way, i.e., $x_0 = \sum_i \lambda_i x_i$ (suppose $\sum_i \lambda_i = 1$). In this way, we need to solve

$$\min \left\{ \int c(x_1, \dots, x_N) \, \mathrm{d}\gamma \ : \ \gamma \in \Pi(\varrho^1, \dots, \varrho^N) \right\} ,$$

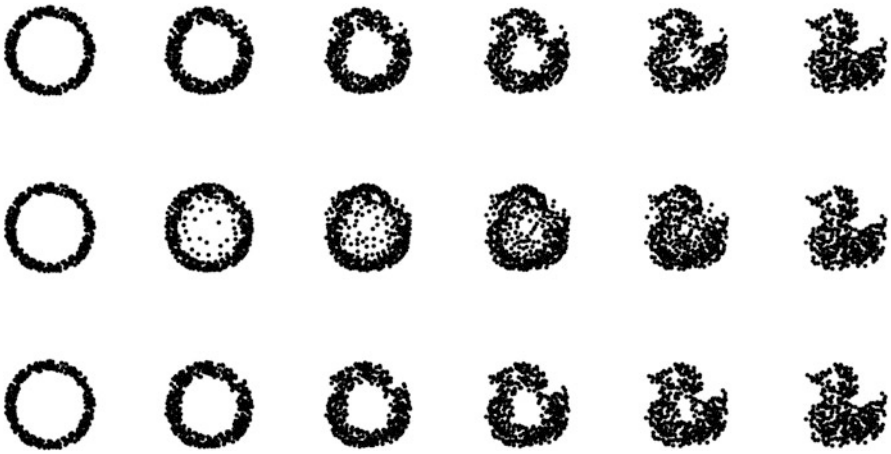


Fig. 5.3 Interpolation between a point distribution on an annulus and on a duck. On the first line, an assignment is computed as in Figure 2.5, and then points move interpolating starting and arrival positions with constant speed. On the third line, the same is done with the optimal assignment, realizing the distance W_2 (hence we have a geodesic in $\mathbb{W}_2$). On the second line, the interpolations are computed as barycenters for $\mathrm{S}W_2$, minimizing the sum of the squared distances. Pictures kindly provided by J. Rabin and adapted from [256]

with $c(x_1,\dots,x_N) := \sum_{i=1}^N \lambda_i |x_i - \bar{x}|^2$ and $\bar{x} = \sum_{i=1}^N \lambda_i x_i$. Note that, in the standard case $\lambda_i = 1/N$, we have

$$c(x_1,\dots,x_N) = \frac{1}{N}\sum_{i=1}^N |x_i - \bar{x}|^2 = -|\bar{x}|^2 + \frac{1}{N}\sum_{i=1}^N |x_i|^2 = \frac{1}{2N^2}\sum_{i,j} |x_i - x_j|^2$$

(to be convinced of the last equality, just expand the squares). This shows that the problem is equivalent to the very first multi-marginal problem studied by Gangbo and Święch in [177].

For the applications of barycenters in image processing, we refer to [256], where, for the sake of numerical computations, the distance W_2 is replaced with $\mathrm{S}W_2$ (see Section 5.5.4). An example is given in Figure 5.3. More recent numerical algorithms have also allowed to directly treat the W_2 distance, as is shown in Figure 5.4.

Fig. 5.4 Interpolation between three different shapes (a five-pointed star, a four-pointed one, and the union of two small disks). Every image represents a barycenter between the three shapes, computed in $\mathbb{W}_2$. Pictures kindly provided by G. Peyré

Chapter 6
Benamou-Brenier and other continuous numerical methods

In this chapter we present some numerical methods to solve optimal transport problems. The search for efficient numerical strategies is a very lively domain in current days, after a period where optimal transport was mainly investigated from the theoretical point of view. It would be impossible to list all the recent developments in the field, and because of the taste for calculus of variations and PDEs which haunts all the book, the choice has been done to prefer those methods which can be classified as "continuous." This means that they are especially suitable to tackle the problem of the optimal transport between two given densities. They are opposite to some more discrete methods, which will be briefly presented in "Discussion" (Section 6.4), and are concerned with the case of two finitely atomic measures or one atomic vs one density. The most famous method is for sure the one due to J. D. Benamou and Y. Brenier, which transforms the problem into a tractable convex variational problem in dimension $d+1$. This method gives the title to the chapter and justifies the position of this chapter after Chapter 5. Indeed, we will describe it strongly using the theory about Wasserstein geodesics (rather than finding the map, this method finds the geodesic curve μ_t).

Two other classical continuous methods are presented: the Angenent-Hacker-Tannenbaum method based on the fact that the optimal maps should be a gradient, and that removing non-gradient parts decreases the energy, and the Loeper-Rapetti method based on the resolution of the Monge-Ampère equation. Both require smooth and non-vanishing densities, and special domains to handle boundary data (a rectangle or, better, a torus). The problem of the resolution of the Monge Ampère equation on more general domains is much more delicate and has been addressed more recently (see the end of Section 6.3).

In this chapter we will neither be exhaustive nor present the most recent methods or compare the efficiency and advantages of all of them. It would be impossible, because of the huge progresses which are occurring on this topic. The goal is mainly to provide the reader an overview of the most natural ideas which could be translated into numerical tools.

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2_6

6.1 The Benamou-Brenier formula and its numerical applications

The results of Sections 5.3 and 5.4 allow to rewrite the optimization problem corresponding to the cost $|x-y|^p$ in a smart way, so as to transform it into a convex optimization problem. Indeed

- looking for an optimal transport for the cost $c(x,y) = |x-y|^p$ is equivalent to looking for constant-speed geodesic in $\mathbb{W}_p$ because from optimal plans we can reconstruct geodesics and from geodesics (via their velocity field) it is possible to reconstruct the optimal transport;
- constant-speed geodesics may be found by minimizing $\int_0^1 |\mu'|(t)^p \mathrm{d}t$;
- in the case of the Wasserstein spaces, we have $|\mu'|(t)^p = \int_\Omega |\mathbf{v}_t|^p \,\mathrm{d}\mu_t$, where $\mathbf{v}$ is a velocity field solving the continuity equation together with μ (this field is not unique, but the metric derivative $|\mu'|(t)$ equals the minimal value of the L^p norm of all possible fields).

As a consequence of these last considerations, for $p > 1$, solving the kinetic energy minimization problem

$$\min\left\{ \int_0^1 \int_\Omega |\mathbf{v}_t|^p \mathrm{d}\varrho_t \,\mathrm{d}t \ : \ \partial_t \varrho_t + \nabla\cdot(\mathbf{v}_t\varrho_t) = 0,\ \varrho_0 = \mu,\ \varrho_1 = \nu \right\}$$

selects constant-speed geodesics connecting μ to ν and hence allows to find the optimal transport between these two measures[1].

On the other hand, this minimization problem in the variables $(\varrho_t, \mathbf{v}_t)$ has nonlinear constraints (due to the product $\mathbf{v}_t\varrho_t$) and the functional is non-convex (since $(t,x) \mapsto t|x|^p$ is not convex). Yet, we saw with the tools of Section 5.3.1 that it is possible to transform it into a convex problem.

For this, it is sufficient to switch variables, from $(\varrho_t, \mathbf{v}_t)$ into (ϱ_t, E_t) where $E_t = \mathbf{v}_t\varrho_t$, and use the functional $\mathscr{B}_p$ in space-time. Remember $\mathscr{B}_p(\varrho, E) := \int f_p(\varrho, E)$ with $f_p : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ defined in Lemma 5.17. We recall its definition for the reader's convenience:

$$f_p(t,x) := \sup_{(a,b)\in K_q} (at + b\cdot x) = \begin{cases} \frac{1}{p}\frac{|x|^p}{t^{p-1}} & \text{if } t>0, \\ 0 & \text{if } t=0, x=0 \\ +\infty & \text{if } t=0, x\neq 0, \ \text{ or } t<0, \end{cases}$$

where $K_q := \{(a,b)\in\mathbb{R}\times\mathbb{R}^d \ : \ a + \frac{1}{q}|b|^q \le 0\}$. The problem becomes

[1] We choose to denote by ϱ the interpolating measures, in order to stress the "continuous" flavor of this method (typically, ϱ is a density, and μ a generic measure).

Problem 6.1. Solve

$$(\mathrm{B}_p\mathrm{P}) \quad \min\left\{\mathscr{B}_p(\varrho,E)\ :\ \partial_t\varrho_t+\nabla\cdot E_t=0,\ \varrho_0=\mu,\ \varrho_1=\nu\right\}.$$

Note that we can write $\mathscr{B}_p(\varrho,E)=\int_0^1\mathscr{B}_p(\varrho_t,E_t)\mathrm{d}t=\int_0^1\int_\Omega f_p(\varrho_t(x),E_t(x))\,\mathrm{d}x\mathrm{d}t$, where this third expression of the functional implicitly assumes $\varrho_t,E_t\ll\mathscr{L}^d$. Indeed, as we saw in Proposition 5.18, the functional $\mathscr{B}_p$ has an integral representation of this form as soon as ϱ_t and E_t are absolutely continuous w.r.t. a same positive measure[2].

We also want to stress that the constraints given by $\partial_t\varrho_t+\nabla\cdot E_t=0$, $\varrho_0=\mu$, $\varrho_1=\nu$ are indeed a divergence constraint in space-time (consider the vector $(\varrho,E):[0,1]\times\Omega\to\mathbb{R}^{d+1}$). The space boundary constraints are already of no-flux type, while the initial and final value of ϱ provides nonhomogeneous Neumann data on the boundaries $\{0\}\times\Omega$ and $\{1\}\times\Omega$. Indeed, the whole constraint can be read as $\nabla_{t,x}\cdot(\varrho,E)=\delta_0\otimes\mu-\delta_1\otimes\nu$ (notice that this point of view which uses derivatives in (t,x) will appear again in the section). The functional that we minimize is a 1-homogeneous functional, and in this way ($\mathrm{B}_p\mathrm{P}$) becomes a dynamical version (in space-time) of the Beckmann problem that we saw in Section 4.2. It is by the way the problem that we obtain if we apply to the cost $|x-y|^p$ the reduction suggested in [196], which transforms it into a one-homogeneous transport cost in space-time.

The constraints are now linear, and the functional is convex. Yet, the functional (and the function f_p as well) is convex, but not so much, since, as we said, it is 1-homogeneous. In particular, it is not strictly convex and not differentiable. This reduces the efficiency of any gradient descent algorithm in order to solve the problem, but some improved methods can be used.

In [34], the authors propose a numerical method, based on this convex change of variables, on duality, and on what is called "augmented Lagrangian."

Box 6.1. Memo: *Saddle points, Uzawa and Augmented Lagrangian*

Suppose that we need to minimize a convex function $f:\mathbb{R}^N\to\mathbb{R}$, subject to k linear equality constraints $Ax=b$ (with $b\in\mathbb{R}^k$ and $A\in\mathrm{M}^{k\times N}$). This problem is equivalent to

$$\min_{x\in\mathbb{R}^N} f(x)+\sup_{\lambda\in\mathbb{R}^k}\lambda\cdot(Ax-b).$$

This gives a min-max problem with a Lagrangian function $L(x,\lambda):=f(x)+\lambda\cdot(Ax-b)$. If we believe in duality, finding a minimizer for f under the constraints is the same as finding a maximizer for $g(\lambda):=\inf_{x\in\mathbb{R}^N}f(x)+\lambda\cdot(Ax-b)$. And it is the same as finding a saddle

(continued)

[2]This is typical of 1-homogeneous functionals (the fact that the result is independent of the reference measure).

Box 6.1. (continued)

point for L (a point $(\bar{x}, \bar{\lambda})$ where $L(x, \bar{\lambda}) \geq L(\bar{x}, \bar{\lambda}) \geq L(\bar{x}, \lambda)$ for every x and every λ, i.e., a point which minimizes in x and maximizes in λ).

The maximization in λ is easier, since the problem is unconstrained: it is possible to apply a gradient algorithm (see Box 6.8). We only need to compute $\nabla g(\lambda)$, but this is easily given via $\nabla g(\lambda) = Ax(\lambda) - b$, where $x(\lambda)$ is the (hopefully unique) minimizer of $x \mapsto f(x) + \lambda \cdot (Ax - b)$.

The algorithm that we obtain from this idea, called *Uzawa algorithm*[3], reads as follows: given (x_k, λ_k), set $x_{k+1} := x(\lambda_k)$ and $\lambda_{k+1} = \lambda_k + \tau \nabla g(\lambda_k) = \lambda_k + \tau(Ax_{k+1} - b)$, for a given small τ. The sequence of points x_k converges, under reasonable assumptions, to the minimizer of the original problem.

An alternative idea which makes the computation of the point $x(\lambda)$ easier and accelerates the convergence is the following: instead of using the Lagrangian $L(x, \lambda)$, use the following variant – $\tilde{L}(x, \lambda) := L(x, \lambda) + \frac{\tilde{\tau}}{2}|Ax - b|^2$, for a given value of $\tilde{\tau}$. The conditions for being a saddle point of $\tilde{L}$ and L are, respectively,

$$\text{for } \tilde{L}: \begin{cases} \nabla f(x) + A^t\lambda + \tilde{\tau}(Ax - b) = 0, \\ Ax - b = 0, \end{cases} \qquad \text{for } L: \begin{cases} \nabla f(x) + A^t\lambda = 0, \\ Ax - b = 0. \end{cases}$$

Hence, the saddle points of $\tilde{L}$ and L are the same, but using $\tilde{L}$ makes the problem more convex and tractable in x. Then, one uses the same gradient algorithm on $\tilde{g}(\lambda) := \inf_x f(x) + \lambda \cdot (Ax - b) + \frac{\tilde{\tau}}{2}|Ax - b|^2$, using a step τ. It is possible to take $\tau = \tilde{\tau}$ and iterate the following algorithm:

$$\begin{cases} x_{k+1} = \operatorname{argmin} f(x) + \lambda_k \cdot (Ax - b) + \frac{\tau}{2}|Ax - b|^2 \\ \lambda_{k+1} = \lambda_k + \tau(Ax_{k+1} - b) \end{cases}.$$

For more details and convergence result, we refer, for instance, to [170].

Here are the main steps to conceive the algorithm.

First of all, we will write the constraint in a weak form (actually, in the sense of distributions), thanks to (4.3). This means that we actually want to solve

$$\min_{\varrho, E} \quad \mathscr{B}_p(\varrho, E) + \sup_{\phi} \left(-\int_0^1 \int_\Omega ((\partial_t \phi)\, \varrho_t + \nabla \phi \cdot E_t) + G(\phi) \right),$$

where we set

$$G(\phi) := \int_\Omega \phi(1, x)\, \mathrm{d}\nu(x) - \int_\Omega \phi(0, x)\, \mathrm{d}\mu(x),$$

[3]The Uzawa algorithm is actually more general than this: it can handle inequality constraints of the form $\varphi^i(x) \leq 0$, with the Lagrangian $L(x, \lambda) = f(x) + \sum_i \lambda^i \varphi^i(x)$ and $\lambda \in (\mathbb{R}_+)^k$, but we prefer to stick to the equality constraints for simplicity of exposition.

and the sup is computed over all functions defined on $[0,1]\times\Omega$ (we do not care here about their regularity, since they will be anyway represented by functions defined on the points on a grid in $[0,1]\times\mathbb{R}^d$).

Remark 6.2. It is interesting to make a small digression to see the connection between the above problem and a Hamilton-Jacobi equation. We will consider the easiest case, $p=2$. In this case we can write the problem as

$$\min_{(E,\varrho):\varrho\geq 0}\int_0^1\int_\Omega \frac{|E|^2}{2\varrho}+\sup_\phi -\int_0^1\int_\Omega((\partial_t\phi)\,\varrho+\nabla\phi\cdot E)+G(\phi),$$

where we expressed the functional $\mathscr{B}_2$ with its integral expression, valid in the case of absolutely continuous measures, with $\varrho\geq 0$ (where $\varrho=0$, we must have $E=0$ in order to have finite energy).

If we formally interchange inf and sup, we get the following problem:

$$\sup_\phi \quad G(\phi)+\inf_{(E,\varrho):\varrho\geq 0}\int_0^1\int_\Omega\left(\frac{|E|^2}{2\varrho}-(\partial_t\phi)\,\varrho-\nabla\phi\cdot E\right).$$

We can first compute the optimal E, for fixed ϱ and ϕ, thus getting $E=\varrho\nabla\phi$. The problem becomes

$$\sup_\phi \quad G(\phi)\;+\;\inf_{\varrho\geq 0}\int_0^1\int_\Omega\left(-(\partial_t\phi)\varrho-\frac{1}{2}|\nabla\phi|^2\varrho\right).$$

The condition for the infimum to be finite (and hence to vanish) is $\partial_t\phi+\frac{1}{2}|\nabla\phi|^2\leq 0$, and at the optimum we must have equality on $\{\varrho>0\}$. This gives the Hamilton-Jacobi equation:

$$\partial_t\phi+\frac{1}{2}|\nabla\phi|^2=0\quad\varrho\text{-a.e.}$$

By the way, from the optimal ϕ, we can recover the Kantorovich potentials using $\psi(x):=\phi(1,x)$ and $\varphi(x):=-\phi(0,x)$.

The quantity $\mathscr{B}_p$ in this variational problem may be expressed as a sup and hence we get

$$\min_{\varrho,E}\sup_{(a,b)\in K_q,\phi}\int_0^1\int_\Omega(a(t,x)\mathrm{d}\varrho+b(t,x)\cdot\mathrm{d}E-\partial_t\phi\,\mathrm{d}\varrho-\nabla\phi\cdot\mathrm{d}E)+G(\phi).$$

Denote $m=(\varrho,E)$. Here $m:[0,1]\times\Omega\to\mathbb{R}^{d+1}$ is a $(d+1)$-dimensional vector field defined on a $(d+1)$-dimensional space. Again, we do not care here about m being a measure or a true function, since anyway we will work in a discretized setting, and m will be a function defined on every point of a grid in $[0,1]\times\mathbb{R}^d$. Analogously, we denote $\xi=(a,b)$. We also denote by $\nabla_{t,x}\phi$ the space-time gradient of ϕ, i.e., $\nabla_{t,x}\phi=(\partial_t\phi,\nabla\phi)$.

The problem may be rewritten as

$$\min_m \sup_{\xi,\phi\,:\,\xi\in K_q} \langle \xi - \nabla_{t,x}\phi, m\rangle + G(\phi).$$

Here comes the idea of using an augmented Lagrangian method. Indeed, the above problem recalls a Lagrangian, but in a reversed way. We must think that the dual variable should be m and the primal one is the pair (ξ,ϕ). The function $f(\xi,\phi)$ includes the term $G(\phi)$ and the constraint $\xi \in K_q$, and there is an equality constraint $\xi = \nabla_{t,x}\phi$. We do not care actually at the original constrained problem giving rise to this Lagrangian, bur we just decide to add in the optimization a term $\frac{\tau}{2}|\xi - \nabla_{t,x}\phi|^2$ (for a small step size τ).

Hence, we look for a solution of

$$\min_m \sup_{\xi,\phi\,:\,\xi\in K_q} \langle \xi - \nabla_{t,x}\phi, m\rangle + G(\phi) - \frac{\tau}{2}|\xi - \nabla_{t,x}\phi|^2.$$

The algorithm that one can consider to find the optimal m should do the following: produce a sequence m_k and find, for each k, the optimal (ξ_k,ϕ_k). Yet, instead of finding exactly the optimal (ξ_k,ϕ_k), we will optimize in two steps (first the optimal ϕ for fixed ξ, then the optimal ξ for this ϕ).

The algorithm will work in three iterate steps. Suppose we have a triplet (m_k,ξ_k,ϕ_k).

- Given m_k and ξ_k, find the optimal ϕ_{k+1}, by solving

$$\max_\phi \quad -\langle \nabla_{t,x}\phi, m_k\rangle + G(\phi) - \frac{\tau}{2}||\xi_k - \nabla_{t,x}\phi||^2,$$

  which amounts to minimizing a quadratic problem in $\nabla_{t,x}\phi$. The solution can be found as the solution of a Laplace equation $\tau\Delta_{t,x}\phi = \nabla\cdot(\tau\xi_k - m_k)$, with a space-time Laplacian and with Neumann boundary conditions. These conditions are homogeneous in space, and nonhomogeneous, due to the role of G, on $t = 0$ and $t = 1$. Most Laplace solvers can find this solution in time $O(N\log N)$, where N is the number of points in the discretization.
- Given m_k and ϕ_{k+1}, find the optimal ξ_{k+1}, by solving

$$\max_{\xi\in K_q} \quad \langle \xi, m_k\rangle - \frac{\tau}{2}|\xi - \nabla_{t,x}\phi_{k+1}|^2.$$

  By expanding the square, we see that this problem is equivalent to the projection of $\nabla_{t,x}\phi_{k+1} + \frac{1}{\tau}m_k$, and no gradient appears in the minimization. This means that the minimization may be performed pointwisely, by selecting for each (t,x) the point $\xi = (a,b)$ which is the closest to $\nabla_{t,x}\phi_{k+1}(t,x) + \frac{1}{\tau}m_k(t,x)$ in the convex set K_q. If we have a method for this projection in $\mathbb{R}^{n+1}$, requiring a constant number of operations, then the cost for this pointwise step is $O(N)$.

- Finally we update m by setting $m_{k+1} = m_k - \tau(\xi_{k+1} - \nabla_{t,x}\phi_{k+1})$.

This algorithm globally requires $O(N \log N)$ operations at each iteration (warning: N is given by the discretization in space-time). It can be proven to converge, even if, for convergence issues, the reader should rather refer to the works by K. Guittet and Benamou-Brenier-Guittet [36, 188, 189], as the original paper by Benamou and Brenier does not insist on this aspect. Also see the recent paper [194] which fixes some points in the proof by Guittet.

Compared to other algorithms, both those that we will present in the rest of the chapter and other more recent ones, the Benamou-Brenier method has some important advantages:

1. It is almost the only one for the moment which takes care with no difficulties of vanishing densities and does not require special assumptions on their supports;
2. It is not specific to the quadratic cost: we presented it for power costs $c(x, y) = |x - y|^p$ but can be adapted to other convex costs $h(x - y)$ and to all costs issued from a Lagrangian action (see Chapter 7 in [293]); it can incorporate costs depending on the position and time and is suitable for Riemannian manifolds;
3. It can be possibly adapted to take into account convex constraints on the density ϱ (for instance, lower or upper bounds);
4. It can handle different "dynamical" problems, where penalizations on the density or on the velocity are added, as it happens in the case of mean field games (see Section 8.4.4) and has been exploited, for instance, in [35];
5. It can handle multiple populations, with possibly interacting behaviors.

We refer, for instance, to [96, 248] for numerical treatments of the variants of the points 2 and 3 above and to [37] for an implementation of the multi-population case. Here we present a simple picture (Figure 6.1), obtained in the easiest case via the Benamou-Brenier method on a torus.

6.2 Angenent-Hacker-Tannenbaum

The algorithm that we see in this section comes from an interesting minimizing flow proposed and studied in [19] (see also [191]). The main idea is to start from an admissible transport map T between a given density f (which we will suppose smooth and strictly positive on a nice compact domain Ω) and a target measure ν (we do not impose any kind of regularity on ν for the moment) and then to rearrange the values of T slowly in time, by composing it with the flow of a vector field $\mathbf{v}$, i.e., considering $\mathrm{T} \circ (Y_t)^{-1}$ instead of $\mathrm{T}(x)$, where $Y_t(x) := y_x(t)$ is, as usual, the solution of

$$\begin{cases} y_x'(t) = \mathbf{v}_t(y_x(t)), \\ y_x(0) = x. \end{cases}$$

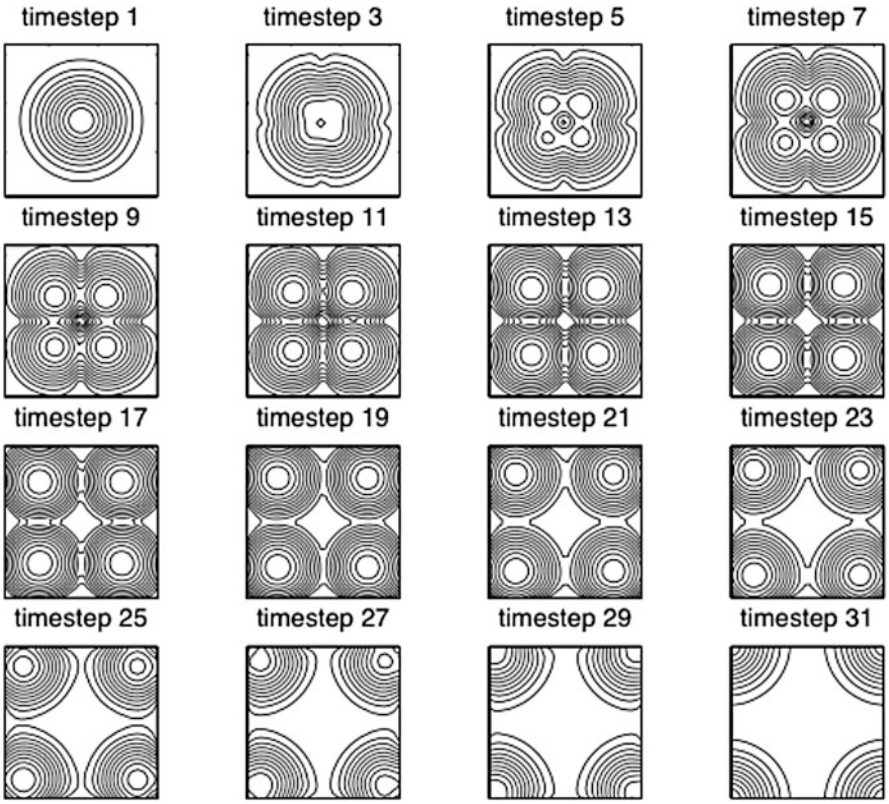


Fig. 6.1 Geodesics between a Gaussian on a 2D torus and the same Gaussian displaced of a vector $(1/2, 1/2)$. Notice that, to avoid the cut locus (see Section 1.3.2), the initial density breaks into four pieces. Picture taken from [34] with permission

The vector fields $\mathbf{v}_t$ have to be chosen so that the map Y_t, defined via $Y_t(x) := y_x(t)$, preserves the measure f and also so as to reduce the value of the Monge cost

$$M(\mathrm{T}) = \int c(x, \mathrm{T}(x)) f(x)\, \mathrm{d}x.$$

The cost c is taken smooth (say, C^1, we will see later the particular case of the quadratic cost).

The first computation to do is the derivative of the cost function: set $\mathrm{T}_t(x) := T(y_x(-t))$ where y is defined as the flow of an autonomous vector field $\mathbf{v}$. The condition for $\mathbf{v}$ to preserve the measure f is that the solution of

$$\partial_t \varrho_t + \nabla \cdot (\varrho_t \mathbf{v}) = 0 \quad \text{with } \varrho_0 = f$$

is the constant (in time) density $\varrho_t = f$. This requires $\nabla \cdot (f\mathbf{v}) = 0$. We will set $\mathbf{w} = f\mathbf{v}$ for simplicity. By choosing $\mathbf{v}$ satisfying this condition, we can write

$$\frac{d}{dt}\left(\int_\Omega c(x, \mathrm{T}(y_x(-t)))f(x)\,\mathrm{d}x\right) = \frac{d}{dt}\left(\int_\Omega c(y_x(t), \mathrm{T}(x))f(x)\,\mathrm{d}x\right)$$
$$= \int_\Omega \nabla_x c(y_x(t), \mathrm{T}(x)) \cdot y_x'(t) f(x)\,\mathrm{d}x = \int_\Omega \nabla_x c(y_x(t), \mathrm{T}(x)) \cdot \mathbf{v}(y_x(t)) f(x)\,\mathrm{d}x,$$

and, at $t = 0$, this gives a derivative equal to $\int_\Omega \nabla_x c(x, \mathrm{T}(x)) \cdot \mathbf{v}(x) f(x)\,\mathrm{d}x$. In the computation above, we used the fact that Y_{-t} preserves the measure f and that its inverse is Y_t (which is true for vector fields $\mathbf{v}$ independent of time).

We recover here an important part of the theory of optimal transport: if T is optimal, then the vector field $\nabla_x c(x, \mathrm{T}(x))$ must be orthogonal to all divergence-free vector fields $\mathbf{w}$, and hence it is a gradient (see Proposition 1.15). In the particular case $c(x, y) = \frac{1}{2}|x-y|^2$, this gives $x - \mathrm{T}(x) = \nabla\varphi(x)$ and hence T itself is a gradient. The fact that T is the gradient of a convex function, and not of any function, comes from second-order optimality conditions, at least when one has sufficient regularity (see **Ex** (44)).

On the other hand, the goal here is not to study the structure of an optimal T, but to start from a nonoptimal T and to improve it. In order to do this, we must choose a suitable vector field $\mathbf{v}$. This will be done by producing a flow T_t, and the vector field $\mathbf{v}$ will be chosen in some clever way at every instant t of time. Hence, it will be no more autonomous. Several choices for $\mathbf{v}$ are possible, but we start from the simplest one: we saw that the derivative of the cost M is given by $\int_\Omega \xi \cdot \mathbf{w}$, where $\xi(x) = \nabla_x c(x, \mathrm{T}(x))$ and $\mathbf{w} = f\mathbf{v}$. Since $\int_\Omega \xi \cdot \mathbf{w}$ is a scalar product in L^2, this identifies the gradient of M for the L^2 structure, when restricted to the set of infinitesimal displacements preserving the measure (i.e., we impose the constraint $\nabla \cdot \mathbf{w} = 0$). The optimal choice of $\mathbf{w}$ (if we want to have a bound on $||\mathbf{w}||_{L^2}$) is to take as $\mathbf{w}$ the projection of $-\xi$ onto the vector space of divergence-free vector fields. Concerning the boundary conditions to associate with this divergence constraint, we choose as usual the Neumann conditions $\mathbf{w} \cdot \mathbf{n} = 0$, which are also those required to preserve the measure without letting any mass exit Ω.

Box 6.2. Memo: *Helmholtz decomposition*

Given a vector field $\xi \in L^2(\Omega; \mathbb{R}^d)$, it is always possible to write it as the sum of a gradient and of a divergence-free vector field $\xi = \nabla u + \mathbf{w}$ with $u \in H^1(\Omega)$ and $\nabla \cdot \mathbf{w} = 0$ (in the distributional sense). The decomposition is unique if one imposes $u \in H^1_0(\Omega)$, or $\mathbf{w} \cdot \mathbf{n} = 0$ (in the weak sense, which, together with the zero-divergence condition, means $\int \mathbf{w} \cdot \nabla\phi = 0$ for all $\phi \in C^1(\Omega)$).

These decompositions have a variational origin: indeed one can solve

(continued)

Box 6.2. (continued)

$$\min\left\{\int |\xi - \nabla u|^2 \,:\, u \in H^1_0(\Omega)\right\}$$

and get the first one, or

$$\min\left\{\int |\xi - \nabla u|^2 \,:\, u \in H^1(\Omega), \int u = 0\right\}$$

and get the second one (if Ω is connected). In both cases, the zero-divergence condition on $\mathbf{w}$ comes from the optimality conditions on u.

Note that ∇u is the projection of ξ on the set of L^2 gradients (of H^1_0 functions, in the first case) and hence $\mathbf{w}$ is the projection onto the set of divergence-free vector fields (with Neumann conditions, in the second case).

The projection $\mathbf{w} = P[\xi]$ can be computed by writing $\nabla \cdot \xi = \nabla \cdot \mathbf{w} + \Delta u$, i.e., $\mathbf{w} = \xi - \nabla(\Delta^{-1}(\nabla \cdot \xi))$.

Hence, we look for a family of transport maps T_t, a flow Y_t, and a vector field $\mathbf{w}_t$ satisfying the following conditions:

- T_t is defined as $\mathrm{T}_t = \mathrm{T} \circ (Y_t)^{-1}$ at every instant t;
- the map Y_t is the flow of the time-dependent vector field $\mathbf{v}_t$, defined through $f\mathbf{v}_t = \mathbf{w}_t$;
- the vector field $\mathbf{w}_t = P[-\xi_t]$, where P is the projection onto divergence-free vector fields with $\mathbf{w} \cdot \mathbf{n} = 0$ and $\xi_t = \nabla_x c(x, \mathrm{T}_t(x))$.

We can write the PDE satisfied by T_t, which is the transport equation $\partial_t \mathrm{T}_t + \mathbf{v}_t \cdot \nabla \mathrm{T}_t = 0$.

Box 6.3. Memo: *Linear transport equation*

Given a smooth vector field $\mathbf{v}_t$, its flow $Y_t(x) = y_x(t)$ (with, as usual, $y'_x(t) = \mathbf{v}_t(y_x(t))$ and $y_x(0) = x$), and a smooth function g, we define $g_t = g \circ Y_t^{-1}$. Then, we have

$$\partial_t g_t + \mathbf{v}_t \cdot \nabla g_t = 0.$$

Indeed, we can write $g_t(y_x(t)) = g(x)$ and differentiate in time, thus getting the above equation. As soon as $\mathbf{v}_t$ stays Lipschitz continuous, its weak formulation is valid for arbitrary $g \in L^1$ since we can take a sequence of smooth approximating functions $g_n \to g$, write $\iint g_n \circ Y_t^{-1}\,(\partial_t \varphi + \mathbf{v}_t \cdot \nabla \varphi)$, use the change of variable $y = Y_t^{-1}(x)$, and pass to the limit.

This equation is called *transport equation*. It is similar to the continuity equation, but it is not in conservative form, and the mass of g is not conserved in time. On the contrary, L^∞ bounds on g are conserved, since the values of g_t are the same as those of $g = g_0$. The two equations, transport and continuity equations, are somehow dual: if a vector field $\mathbf{v}$ is fixed and we take smooth solutions ϱ_t of $\partial_t \varrho_t + \nabla \cdot (\varrho_t \mathbf{v}_t) = 0$ and g_t of $\partial_t g_t + \mathbf{v}_t \cdot \nabla g_t = 0$,

(continued)

Box 6.3. (continued)
then we have

$$\frac{d}{dt}\left(\int g_t \,\mathrm{d}\varrho_t\right) = \int (\partial_t g_t)\varrho_t + \int g_t(\partial_t \varrho_t) = -\int \varrho_t \mathbf{v}_t \cdot \nabla g_t - \int g_t \nabla \cdot (\varrho_t \mathbf{v}_t) = 0.$$

We now compute formally the derivative in time of $M(\mathrm{T}_t)$ with the above choices

$$\frac{d}{dt}M(\mathrm{T}_t) = \frac{d}{dt}\left(\int c(x,\mathrm{T}_t)f\right) = \int \nabla_y c(x,\mathrm{T}_t)(\partial_t \mathrm{T}_t)f = -\int \nabla_y c(x,\mathrm{T}_t)\cdot \nabla \mathrm{T}_t \cdot \mathbf{w}_t.$$

We go on using $\nabla[c(x,\mathrm{T}_t(x))] = \nabla_x c(x,\mathrm{T}_t(x)) + \nabla_y c(x,\mathrm{T}_t(x))\nabla \mathrm{T}_t(x)$, together with $\int \nabla[c(x,\mathrm{T}_t(x))]\cdot \mathbf{w}_t(x)\,\mathrm{d}x = 0$, which comes from $\nabla \cdot \mathbf{w}_t = 0$. Hence we get

$$\frac{d}{dt}M(\mathrm{T}_t) = \int \nabla_x c(x,\mathrm{T}_t)\cdot \mathbf{w}_t(x) = \int \xi_t \cdot \mathbf{w}_t = -\int \xi_t P[\xi_t] = -||P[\xi_t]||^2 \le 0.$$

This shows that the proposed flow makes $M(\mathrm{T}_t)$ decrease in time and that the value of $M(\mathrm{T}_t)$ is stationary if and only if $P[\xi_t] = 0$, i.e., if ξ_t is a gradient.

[19] proves the local existence of the above flow, i.e.,

$$\begin{cases} \partial_t \mathrm{T}_t + \mathbf{v}_t \cdot \nabla \mathrm{T}_t = 0, \\ \nabla \cdot (f\mathbf{v}_t) = 0, \ \ \mathbf{v}_t \cdot \mathbf{n} = 0, \\ -\nabla_x c(x,\mathrm{T}_t) = f\mathbf{v}_t + \nabla u_t, \\ \mathrm{T}_0 \ \ \text{given}, \end{cases}$$

(the second and third line can be replaced by $\mathbf{v}_t = \xi_t - \nabla(\Delta^{-1}(\nabla \cdot \xi_t))$, for $\xi_t = \nabla_x c(x,\mathrm{T}_t)$, which avoids inserting a new variable u_t, but introduces nonlocal terms in the equation) is proven. Also, it is proven that the only limits of T_t for $t \to \infty$ can be those rearrangements T of T_0 such that $\nabla_x c(x,\mathrm{T})$ is a gradient. The same is studied also in the more general case of transport plans instead of transport maps, but we avoid presenting this advanced case for the sake of simplicity. However, we note that for costs c which do not satisfy the twist condition, the framework of transport plans cannot be avoided. Indeed, in general it is not possible to converge to an optimal transport map as $t \to \infty$, simply because such a map could not exist.

In the very same paper [19], alternative choices of the vector field $\mathbf{v}$ are proposed, in particular "regularized" flows (i.e., applying regularization kernels to the above $\mathbf{w}_t$, and in this case global existence is proven). There is also a possibility of a local choice for the flow: in coordinates, we can take $\mathbf{w}^j = \xi^i_{ij} - \xi^j_{ii}$ (superscripts are components and subscripts are derivations). It can be checked that this choice as well guarantees that $M(\mathrm{T}_t)$ decreases in time.

Even if, as we noted, the method is general for any cost, it is in the case of the quadratic cost that it has been mainly employed. The idea in this case can be roughly summarized as "start from an admissible transport map, and then *gradientize* it" (i.e., transform it slowly into a gradient).

The equation becomes, for $\xi_t = x - \mathrm{T}_t$,

$$\begin{cases} \partial_t \xi_t - \mathbf{v}_t + \mathbf{v}_t \cdot \nabla \xi_t = 0, \\ \mathbf{v}_t = \frac{1}{f}\left(\xi_t - \nabla(\Delta^{-1}(\nabla \cdot \xi_t))\right), \\ \xi_0(x) = x - \mathrm{T}_0(x) \ \text{ given.} \end{cases}$$

From an implementation point of view (see [20]), the above system has been discretized and solved using an upwind scheme for the transport equation (roughly speaking, a finite difference method where the choice of the left- or right-centered derivatives is made according to the sign of the components of the advecting vector field $\mathbf{v}$) and standard solvers for the solution of the Laplacian (such as Matlab Poisson solver). The complexity of each time iteration is of the order of $N \log N$, N being the number of pixels in the space discretization.

The AHT flow has been used for numerical purposes, with some advantages and disadvantages. Even if it has been mainly used for the quadratic case, it could be used for many smooth costs, which makes it a very general approach. Yet, from the numerical point of view, the implementation involves some difficulties, and it has also the disadvantage of requiring strictly positive densities. Also, its convergence to the optimal map is not at all proven. This question can be easily understood in the quadratic case. Beware that we do not mean here rigorous convergence results: the problem is that the flow is meant to converge to a critical point for the cost M, which means (for the quadratic case) a transport T which is a gradient map. Nothing guarantees that it is the gradient of a convex map!

A priori, even the gradient of a concave function (i.e., the worst possible map for the quadratic cost, see **Ex** (2)) could be recovered. Since the flow has the property that $M(\mathrm{T}_t)$ is strictly decreasing in time as long as one does not meet a critical point, this case can be easily excluded (the only possibility is starting from T_0 being the gradient of a concave function, which is easy to avoid). But any other gradient could be found at the limit.

These issues have been discussed in [86]. It happens that a good idea to improve the results and enhance the chances of converging to the good transport map, following [19, 20], is to select a good initial guess T_0. Several choices (specific for the quadratic case) are proposed (including the Dacorogna-Moser transport map; see Box 4.3 in Section 4.2.3), but, at least at a first sight, the most suitable one is the Knothe transport (see Section 2.3). Indeed, with this choice, the initial transport map has triangular Jacobian matrix, with positive eigenvalues. Let us assume (extra than the previous assumptions) that ν is absolutely continuous with density $g(x)$ bounded from above and below by positive constants. As we know that the map T_t satisfies $(\mathrm{T}_t)_\# f = g$ for every t, the value of $|\det(D\mathrm{T}_t)|$ can never be 0 (bounded from below by $\inf f / \sup g$). If everything is smooth, this means that the sign of the determinant

is preserved and that no eigenvalue can ever be zero (this is also underlined in [86]). Unfortunately, the eigenvalues of $D\mathrm{T}_t$ are real at $t = 0$ (if we choose the Knothe map) and are real at $t = \infty$ (since, as a gradient, the Jacobian is symmetric), but nothing prevents the eigenvalues to become complex and to pass from $]0, +\infty[$ to $]-\infty, 0[$ without passing through 0 during the evolution! (even if this does not seem likely to happen, because the choice of $\mathbf{v}$ exactly forces T_t to become closer to the set of gradients, and gradients do not like complex eigenvalues).

A simple case, where there is no ambiguity if everything is smooth, is the 2D case. We can give a straightforward proposition.

Proposition 6.3. *Suppose that $(\xi_t, \mathbf{v}_t)$ is a smooth solution of the AHT system in the quadratic case, with $\mathrm{T}_t(x) = x - \xi_t(x)$ being a transport map between two given smooth densities $f, g \in \mathscr{P}(\Omega)$), where $\Omega \subset \mathbb{R}^2$ is a 2D smooth convex and bounded domain. Suppose that $\det(D\mathrm{T}_0(x))$ is positive for every x. Also suppose that $\mathrm{T}_t \to \mathrm{T}_\infty$ in $C^1(\Omega)$ as $t \to \infty$ and that the evolution is nontrivial (i.e., there exists t such that $\mathrm{T}_t \neq \mathrm{T}_0$). Then T_∞ is the gradient of a convex function and is thus optimal.*

Proof. From the above considerations, we have

$$M(T_0) - M(\mathrm{T}_t) = \int_0^t ||P[\xi_t]||_{L^2}^2 \mathrm{d}t. \tag{6.1}$$

This implies that $\int_0^\infty ||P[\xi_t]||_{L^2}^2 \mathrm{d}t < \infty$ and, at least on a sequence $t_n \to \infty$, we have $P[\xi_{t_n}] \to 0$ in $L^2(\Omega)$. Hence $P[\xi_\infty] = 0$ since $\xi_{t_n} \to \xi_\infty := x - \mathrm{T}_\infty(x)$ in C^1, and hence ξ_∞ and T_∞ are gradients. We need to prove that T_∞ is the gradient of a convex function, i.e., that the eigenvalues of $D\mathrm{T}_\infty$, which is now a symmetric matrix, are both positive (remember that we are in $\mathbb{R}^2$).

We know that $t \mapsto \det(D\mathrm{T}_t(x))$ is continuous in time, is positive at $t = 0$, and cannot vanish. From C^1 convergence, this implies $\det(D\mathrm{T}_\infty(x)) > 0$ for every x. Hence, the eigenvalues of $D\mathrm{T}_\infty(x)$ have the same sign (for each fixed x).

Let us look now at $\mathrm{Tr}(D\mathrm{T}_\infty(x))$. This is a continuous function of x, and it cannot be zero. Indeed, if it vanished somewhere, at such a point the two eigenvalues would have different signs. Hence, the sign of this trace is constant in x. Note that this only works for $t = \infty$, as for $t \in]0, \infty[$ one could a priori have complex eigenvalues.

If we exclude the case where $\mathrm{Tr}(D\mathrm{T}_\infty(x)) < 0$ everywhere, we have concluded. Yet, this case corresponds to T_∞ being the gradient of a concave function (remember again that in dimension two, positive determinant and negative trace imply negative 2D definite). And concave functions are the worse possible transports for the quadratic case (see **Ex** (2)). This is impossible since $M(\mathrm{T}_\infty) < M(\mathrm{T}_0)$ (except if the curve $t \mapsto \mathrm{T}_t$ were constant) from (6.1). □

Note that the assumptions on the initial condition of the above theorem are satisfied both by the Knothe transport and by the Dacorogna-Moser transport map.

Unfortunately, the above proof is really 2D as we characterized positive symmetric matrices as those which have positive trace and positive determinant. Also, in

practice it is almost impossible to guarantee the assumptions of the theorem above and convergence can only be observed empirically.

Open Problem (convergence of AHT): justify, even assuming smoothness and very strong convergence, that the AHT flow converges to the optimal map, if initialized at Knothe and/or Dacorogna-Moser, in dimension higher than two, and prove a rigorous result with minimal assumption in 2D.

6.3 Numerical solution of Monge-Ampère

As we saw in last section, we can use for numerical purposes the fact that solving the quadratic optimal transport problem is equivalent to finding a transport map which is the gradient of a convex function. In the last section, an iterated method was produced where a sequence of transport maps T_n (it was a continuous flow, but in practice it is discretized), with the same image measure, was considered, and they were forced to converge to a gradient. Here the approach is somehow opposite: a sequence of gradient maps is considered, and the corresponding measures are forced to converge to the desired target.

More precisely, the idea contained in [217] is the following: consider the simplified case where the domain $\Omega = \mathbb{T}^d$ is a d-dimensional torus, the source measure ϱ is a $C^{0,\alpha}$ density bounded from below and from above, and the target measure is the uniform measure with constant density 1; we look for a smooth function φ such that $D^2\varphi \leq \mathrm{I}$ (which is the condition for φ to be c-concave, with $c(x, y) = |[x - y]|^2$ on the torus; see Section 1.3.2) and

$$\det(\mathrm{I} - D^2\varphi) = \varrho.$$

We stress the fact that, in this periodic setting, it is not possible to look for a convex function u satisfying $\det(D^2u) = \varrho$. Actually, periodic convex functions do not exist (except constant ones)! The only possibility is to stick to the framework of c-concave functions, without passing to the convexity of $x \mapsto \frac{1}{2}|x|^2 - \varphi(x)$.

If we find such a function φ, then $\mathrm{T}(x) = x - \nabla\varphi(x)$ (in the periodic sense) will be the optimal map between ϱ and $\mathscr{L}^d$. Note that the choice of a uniform target measure gives a non-negligible simplification: the correct equation with a generic target measure would involve a function of $x - \nabla\varphi(x)$ at the denominator in the right-hand side. This would make the following analysis much more involved.

The idea of Loeper and Rapetti in [217] is to solve the equation through an iterative scheme, based on a Newton method.

Box 6.4. Memo: *Newton's method to solve $F(x) = 0$*

Suppose we seek solutions of the equation $F = 0$, where $F : X \to X$ is a smooth function from a vector space X into itself. Suppose a solution $\bar{x}$ exists, with $DF(\bar{x})$ which is an invertible linear operator. Given an initial point x_0, we can define a recursive sequence in this way: x_{k+1} is the only solution of the linearized equation $F(x_k)+DF(x_k)(x-x_k) = 0$. In other words, $x_{k+1} = x_k - DF(x_k)^{-1} \cdot F(x_k)$. This amounts at iterating the map $x \mapsto G(x) := x - DF(x)^{-1} \cdot F(x)$. This map is well defined, and it is a contraction in a neighborhood of $\bar{x}$. The iterations converge to $\bar{x}$ if the initial point belongs to such a suitable neighborhood. The order of convergence is quadratic, i.e., $|x_{k+1} - \bar{x}| \leq C|x_k - \bar{x}|^2$. This is a consequence of the fact that the Lipschitz constant of G in a ball $B(\bar{x}, R)$ is proportional to $\max\{|F(x)| : x \in B(\bar{x}, R)\}$ and hence decreases as $O(R)$. To see this fact, just compute

$$DG = \mathrm{I} - DF^{-1} \cdot DF + DF^{-1} \cdot D^2F \cdot DF^{-1} \cdot F = DF^{-1} \cdot D^2F \cdot DF^{-1} \cdot F.$$

An alternative and less "aggressive" method can make use of a step parameter $\tau > 0$, defining $x_{k+1} = x_k - \tau DF(x_k)^{-1} \cdot F(x_k)$. In this case the map $G(x) = x - \tau DF(x)^{-1} \cdot F(x)$ is a contraction, but its contraction factor is of the order of $1 - \tau$, which only gives slower convergence (but could converge under weaker assumptions).

Here, if we have a guess φ and look for a solution $\varphi + \psi$, the linearization of the equation reads

$$\det(\mathrm{I} - D^2\varphi) - \mathrm{Tr}[\mathrm{cof}(\mathrm{I} - D^2\varphi)D^2\psi] = \varrho,$$

where $\mathrm{cof}(A)$ denotes the cofactor matrix of a given square matrix A, satisfying $\mathrm{cof}(A) \cdot A = \det(A)I$.

Box 6.5. Memo: *Linearization of the determinant*

It is well known that $\det(\mathrm{I} + \varepsilon B) = 1 + \varepsilon \mathrm{Tr}[B] + O(\varepsilon^2)$ (this can be easily checked by expanding the determinant of $\mathrm{I} + \varepsilon B$ and looking at the terms which are first order in ε). This gives the linearization of the determinant in a neighborhood of the identity matrix. If we look for the linearization of the determinant around another matrix A and we suppose A invertible, we have

$$\begin{aligned}\det(A + \varepsilon B) = \det(A)\det(\mathrm{I} + \varepsilon A^{-1}B) &= \det(A)(1 + \varepsilon \mathrm{Tr}[A^{-1}B] + O(\varepsilon^2)) \\ &= \det(A) + \varepsilon \mathrm{Tr}[\mathrm{cof}(A)B] + O(\varepsilon^2).\end{aligned}$$

By continuity, the determinant and the map $A \mapsto \mathrm{cof}(A)$ being C^∞ functions, this last expression (avoiding the use of A^{-1}) is also valid for A non-invertible.

The iterative steps of the algorithm proposed in [217] are thus:

- Start with an arbitrary smooth function φ_0 such that $\mathrm{I} - D^2\varphi_0 > 0$ and fix a small value of $\tau > 0$. Typically, one takes $\varphi_0 = 0$.
- At every step n, compute $\varrho_n := \det(\mathrm{I} - D^2\varphi_n)$.

- Define the matrix $M_n(x) := \mathrm{cof}(\mathrm{I} - D^2\varphi_n)$. Suppose that we can guarantee that M_n is positive definite.
- Solve $\mathrm{Tr}[M_n D^2\psi] = \varrho_n - \varrho$, which is a second-order linear elliptic equation in φ, on the torus (i.e., with periodic boundary conditions). The solution is unique up to an additive constant (choose for instance the zero-mean solution) provided $M_n > 0$. Call ψ_n the solution.
- Define $\varphi_{n+1} = \varphi_n + \tau\psi_n$ and start again.

From the implementation point of view, the elliptic equation defining ψ, i.e., $\mathrm{Tr}[M_n D^2\psi] = \varrho_n - \varrho$ is solved in [217] via a finite-difference discretization of the second-order derivatives. It is worthwhile mentioning that this kind of equation, i.e., $\mathrm{Tr}[MD^2\psi] = f$, where M is the cofactor matrix of an Hessian, can also be expressed in divergence form. Indeed, they read $\sum_{i,j} M^{ij}\psi_{ij} = f$, but $\sum_{i,j} M^{ij}\psi_{ij} = \sum_i \left(M^{ij}\psi_j\right)_i$ since $\sum_i M^{ij}_i = 0$ for all j (see **Ex** (43)). This allows to give a variational structure to this elliptic equation and solve it via finite-element methods. On the other hand, the condition $\varrho_n := \det(\mathrm{I} - D^2\varphi_n)$ is more naturally interpreted in terms of second-order finite differences. Let us stress that one of the main difficulties of these equations is exactly the discretization of a nonlinear operator such as the Monge-Ampère operator (i.e., the determinant of the Hessian).

The convergence of the algorithm is guaranteed by the following theorem, under regularity assumptions:

Theorem 6.4. *Suppose $\varrho \in C^{0,\alpha}(\mathbb{T}^d)$ is a strictly positive density and that τ is smaller than a constant c_0 depending on $||\log\varrho||_{L^\infty}$ and $||\varrho||_{C^{0,\alpha}}$. Then the above algorithm provides a sequence φ_n converging in C^2 to the solution φ of* $\det(\mathrm{I} - D^2\varphi) = \varrho$.

Proof. The proof is based on a priori bounds derived from both elliptic regularity theory (see [184], for instance) and Caffarelli's regularity for Monge-Ampère (see Section 1.7.6; for the periodic boundary conditions, see also [127]).

Choosing τ small enough and suitable constants $c_1 > 0$ and C_1, C_2, C_3 large enough, one can prove by induction the following statement: at every step n, we have $\mathrm{I} - D^2\varphi_n > c_1\mathrm{I}$, $||\log(\varrho_n/\varrho)||_{L^\infty} \le C_1$, $||\varrho_n - \varrho||_{C^{0,\alpha}} \le C_2$ and $||\varphi_n||_{C^{2,\alpha}} \le C_3$. To do that, note that we have $||\psi_n||_{C^{2,\alpha}} \le C||\varrho - \varrho_n||_{C^{0,\alpha}}$ from standard elliptic regularity theory (see Schauder Estimates in Chapter 6 in [184]). We can compute

$$\begin{aligned}\varrho_{n+1} = \det(\mathrm{I} - D^2\varphi_n - \tau D^2\psi_n) &= \varrho_n - \tau\mathrm{Tr}[\mathrm{cof}(\mathrm{I} - D^2\varphi_n)D^2\psi_n] + r_n\\ &= (1-\tau)\varrho_n + \tau\varrho + r_n,\end{aligned}$$

where r_n is the second-order rest in the linearization of the determinant, i.e., it involves second derivatives of φ_n and second derivatives of ψ_n, but is of order 2 in $\tau D^2\psi_n$. Hence, from the algebra properties of the $C^{0,\alpha}$ norm, we have $||r_n||_{C^{0,\alpha}} \le C\tau^2||D^2\psi_n||^2_{C^{0,\alpha}} \le C\tau^2||\varrho_n - \varrho||^2_{C^{0,\alpha}}$. Then, from $\varrho_{n+1} - \varrho = (1-\tau)(\varrho_n - \varrho) + r_n$, we get

$$||\varrho_{n+1}-\varrho||_{C^{0,\alpha}} \le (1-\tau)||\varrho_n-\varrho||_{C^{0,\alpha}} + C\tau^2||\varrho_n-\varrho||^2_{C^{0,\alpha}} \le ||\varrho_n-\varrho||_{C^{0,\alpha}}(1-\tau+C\tau^2).$$

It is enough to choose τ small enough so that $1-\tau+C\tau^2 < 1$, to go on with the bound on $||\varrho_{n+1}-\varrho||_{C^{0,\alpha}}$. For the L^∞ bound, we use again

$$\varrho_{n+1} = (1-\tau)\varrho_n + \tau\varrho + r_n \le ((1-\tau)e^{C_1} + \tau + C\tau^2)\varrho.$$

Again, for τ small enough and $C_1 > 0$, one can guarantee $(1-\tau)e^{C_1}+\tau+C\tau^2 \le e^{C_1}$. The lower bound works in the same way.

In order to study the bound on φ_{n+1}, we write $\varphi_{n+1} = \varphi_n + \tau\psi_n$. Note that, from $\mathrm{I} - D^2\varphi_n \ge c_1\mathrm{I}$ and $|D^2\psi_n| \le C$, a suitable choice of τ allows to guarantee the convexity condition $\mathrm{I} - D^2\varphi_{n+1} > 0$. Hence, the uniform bound $||\varphi_{n+1}||_{C^{2,\alpha}} \le C_3$ simply follows from Caffarelli's theory and from the bounds on ϱ_{n+1}. And, finally, an upper bound on the Hessian of φ_{n+1} also implies a lower bound on $\mathrm{I} - D^2\varphi_{n+1}$ because of the determinant condition.

This shows that the bounds are preserved in the iterations, and the convergence comes from $||\psi_n||_{C^{2,\alpha}} \le C||\varrho - \varrho_n||_{C^{0,\alpha}}$, where the norm $||\varrho - \varrho_n||_{C^{0,\alpha}}$ goes exponentially to 0. □

We finish this section by mentioning that after this first work by Loper and Rapetti, which indeed provided interesting numerical results, many improvements have occurred. First, we underline that [280] extended to the case of a nonconstant target density (but still, Lipschitz regularity of the target density was needed). Then, we remark that another strong limitation of [217] is given by its periodic boundary conditions. When working on different domains than $\mathbb{T}^d$, the situation is much trickier. It has been successfully studied by Benamou, Froese, and Oberman in [39, 41], by using a monotone discretization of the Monge-Ampère operator. Monotonicity allows to prove the existence of a solution of the discretized problem and to prove convergence toward the viscosity solution of the limit Monge-Ampère equation. Finally, [39, 41] also attacked the problem of the boundary conditions and of the shape of the domain. With the approach by viscosity solutions, the authors are able to consider non-convex and even non-connected source domains (hence allowing vanishing densities). They handle the second boundary value problem writing the target domain as the level set of a convex function. Yet, we prefer not to enter into details of these recent results, because of their technical difficulty and of the notions about viscosity solutions that we did not develop here.

6.4 Discussion

6.4.1 Discrete numerical methods

We strongly used in Chapter 1 the linear and convex structure of the Kantorovich problem, which is indeed a linear programming problem in infinite dimension. It

is natural to look for methods which exploit this fact, after discretization. Here discretization simply means replacing μ and ν with two finitely atomic measures $\mu = \sum_{i=1}^{N} a_i \delta_{x_i}$ and $\nu = \sum_{j=1}^{M} b_j \delta_{y_j}$. For fixed N and M, the choice of "optimal" points x_i and y_j and of weights a_i, b_j so that the approximation is as precise as possible is a delicate matter that we do not want to discuss here. It is linked to quantization problems for the measures μ and ν (see [187] for a survey of the theory from a signal processing point of view and [72] for a variational asymptotical approach). We note anyway that the difficult issue is the choice of the points, since once the points are fixed, the natural choice for the weights is $a_i = \mu(V_i)$, $b_j = \nu(W_j)$, where V and W represent the Voronoi cells of the points $(x_i)_i$ and $(y_j)_j$, respectively (see below).

Box 6.6. Good to know! *Voronoi cells*

Given a finite family of distinct points $(x_i)_i \subset \mathbb{R}^d$, we define their Voronoi cells as follows:

$$V_i = \{x \in \mathbb{R}^d \,:\, |x - x_i| \le |x - x_j| \text{ for all } j\}.$$

These cells are essentially disjoint (only the boundary can be in common: $V_i \cap V_{i'} = \partial V_i \cap \partial V_{i'}$). They are convex polyhedra, possibly unbounded, and the boundary $\partial V_i \cap \partial V_{i'}$ is orthogonal to the vector $x_i - x_{i'}$. Each point x_i obviously belongs to the corresponding V_i.

Given a finite measure ϱ which does not give mass to $(d-1)$-dimensional sets, a Voronoi partition is said to be a *Centroidal Voronoi Tessellation* (CVT) if every point x_i is the barycenter (according to μ) of its cell V_i: $x_i = \frac{1}{\mu(V_i)} \int_{V_i} x \, \mathrm{d}\varrho(x)$. Note that if $\mu = \sum_{i=1}^{N} a_i \delta_{x_i}$ is a solution of $\min\{W_2(\mu, \varrho) \,:\, \#(\mathrm{spt}(\mu)) \le N\}$ (the so-called *location problem*), then the points x_i are a Centroidal Voronoi Tessellation (see **Ex** (39)). There is an efficient algorithm, called *Lloyd algorithm* which finds a CVT and a local minimizer of this optimal quantization problem by iterating a simple construction: "once we have the cells, we move each point to their barycenter" (see [213]).

Anyway, after discretization, the problem becomes finite dimensional, and the only ingredients are a family of costs c_{ij} (real numbers) and some mass constraints a_i and b_j. The problem becomes

$$\min \left\{ \sum_{i,j} c_{ij} \gamma_{ij} \;:\; \gamma_{ij} \ge 0, \sum_i \gamma_{ij} = b_j, \sum_j \gamma_{ij} = a_i \right\}. \tag{6.2}$$

This is a classical linear problem in the standard form $\min\{c \cdot x \,:\, x \ge 0,\, Cx = b\}$ (where $x \ge 0$ means $x_i \ge 0$ for every i). It can be solved, for instance, with the simplex algorithm.

Box 6.7. Good to know! *Simplex algorithm*

If $c \in \mathbb{R}^d, b \in \mathbb{R}^k, C \in \mathrm{M}^{k\times n}(\mathbb{R})$ are given, the minimum in $\min\{c \cdot x : x \in \mathbb{R}^d, x \geq 0, Cx = b\}$, if it exists, is realized by a vertex of the polyhedron $K = \{x \in (\mathbb{R}_+)^N, Cx = b\}$. A vertex is defined as an extremal point of K (i.e., a point x which cannot be written as $x = (y+z)/2$ with $y, z \in K$ and $y \neq z$). In the case of polyhedra, vertices are characterized by the property that the columns C^i of the matrix C corresponding to the indices i with $x_i > 0$ are linearly independent. Every polyhedron has a finite number of vertices.

The simplex algorithm, invented by G. B. Dantzig in 1947, considers all possible basis of $\mathbb{R}^k$ (or of the range of C) composed of columns C^i. In this way all vertices are taken into account (one vertex x can correspond to several bases: if the columns C^i with $x_i > 0$ are less than k, then they can be completed at will). The algorithms move from one basis to another, choosing the next one in a clever way. The choice of the order is based on local improvements of the objective function $c \cdot x$. In this way it can find the optimal vertex (or find that the problem has no solution, if K is unbounded and $c \cdot x$ is not bounded from below on K) without visiting all the vertices.

See [135] and [136] for an interesting history of how linear programming began. Curiously, the example provided by Dantzig to explain how linear programming opened new horizons in optimization is exactly a transport problem (maximal productivity assignment).

Among the good news with this kind of approach there is the fact that, at least in the assignment problem (i.e., the case $N = M$ and $a_i = b_j = 1/N$), one can prove that the vertices of the convex set $\Pi(\mu, \nu)$ are exactly those γ induced by a transport map (which is, in this case, a permutation of $\{1, \dots, N\}$; see **Ex** (40)).

Among the bad news, there is the computational time of the algorithm. Even if it is much cheaper than visiting all the vertices of $\Pi(\mu, \nu)$, it has in general a huge computational cost[4].

To overcome this difficulty, one possibility is to exploit the particular form of this linear programming problem, due to the particular features of the matrix C, which only has entries with values 0 and 1. For matrices of this kind, there is an interpretation in terms of optimization problems on graphs or *networks*. And in this specific case, this is quite easy to understand: the optimal transport from N atoms to M atoms (even when $N \neq M$ and their weights are different) may be seen as a network with $N + M$ nodes (the source and target points) and NM edges (a complete bipartite graph, hence), and the unknown is the flow on each edge, satisfying compatibility with the mass which is imposed at every node. The optimization involves a linear function of these flows. For linear programming problems of this form, an alternative to the simplex algorithm is the so-called *network simplex*, for which we refer to [2] and to its historical reference, [169]. These methods are widely used, even if they also lack worst-case polynomial bounds on their complexity.

[4]We sometimes hear that the simplex method is exponential: this is a worst-case estimate. Also, other solvers for linear programming problems exist, for instance, interior point methods, and their computational cost can be way better than exponential. However, in practice they are all too expensive.

Other methods have also been found, also specific to the transport problem and based on duality. Among the most well known, there is the so-called Hungarian algorithm [204], which is polynomial in time, and the Auction algorithm that we describe below.

Auction algorithm

This algorithm is also connected with the dual problem, but is not based on a sequence of improvements of the dual objective function, but on a seek for the equilibrium. Due to its connections with the economic considerations of Section 1.7.3, we give a quite detailed description of it (see also [53]).

We specialize to the assignment problem, i.e., the case where the weights a_i and b_j are all equal. We can imagine that we have N buyers, denoted by the index i, and N goods to be bought, denoted by j. We look for an assignment $j = \sigma(i), \sigma \in S_N$ which maximizes[5] $\sum_i u_{i\sigma(i)}$. The values u_{ij} are the utilities of buyer i when he buys item j. As in Section 1.7.3, we look at a price system $p = (p_j)_j$. Given a price system p and an assignment σ, we say that it is an equilibrium if for every i, we have $u_{i\sigma(i)} - p_{\sigma_i} = \max_j u_{ij} - p_j$. The buyers i satisfying this condition are said to be "happy." This only corresponds to writing the equilibrium condition that we presented in Section 1.7.3 in terms of a coupling γ induced by a permutation. It is well known that if (p, σ) is an equilibrium, then σ is an optimal assignment (and p is optimal in the dual problem).

Let us start from an arbitrary pair (p_0, σ_0).

At every step, we have a pair (p^n, σ_n). If it is an equilibrium, then we stop the algorithm. Otherwise, pick any i^* among those i such that $u_{i\sigma(i)} - p_{\sigma_i} < \max_j u_{ij} - p_j$. The selected buyer i^* implements two actions:

1. he chooses one good j^* realizing the maximum in $\max_j u_{i^*j} - p_j$ and exchanges his own good $\sigma_n(i)$ with the buyer $\sigma_n^{-1}(j^*)$ who was originally assigned to j^*
2. he increases the price of j^* to a new value which is such that he is indifferent between j^* and the second best object in $\max_j u_{i^*j} - p_j$.

Thus, the new price of j^* is $P = p^n_{j^*} + \max_j(u_{i^*j} - p^n_j) - \max_{j \neq j^*}(u_{i^*j} - p^n_j) \geq p_{j^*}$.

The new price system p^{n+1} is defined by setting $p^{n+1}_j = p^n_j$ for $j \neq j^*$ and $p^{n+1}_{j^*} = P$. The new permutation σ_{n+1} is obtained by composing with the transposition which exchanges j^* and $\sigma_n(i^*)$.

Unfortunately, the above algorithm can loop forever, because of possible ex-aequo in $\max_j u_{i^*j} - p_j$. Indeed, in such a case the prices can stay constant, and a certain subset of buyers could go on exchanging among them the same set of goods. To bypass this problem, it is useful to introduce a tolerance parameter $\varepsilon > 0$. We define a buyer i to be ε-happy whenever $u_{i\sigma(i)} - p_{\sigma_i} > \max_j(u_{ij} - p_j) - \varepsilon$. We

[5] We switch back to the language, of economists, who prefer to maximize utility rather than minimizing costs.

look now for a pair (p, σ) where all buyers are ε-happy. The algorithm evolves as before, selecting a buyer i^* who is not ε-happy, letting him exchange a good with $\sigma_n^{-1}(j^*)$, but the rule for price increase changes. Indeed, we can now set $P = p^n_{j^*} + \max_j(u_{i^*j} - p^n_j) - \max_{j \neq j^*}(u_{i^*j} - p^n_j) \geq p_{j^*} + \varepsilon$. This price is the maximal one that we can give to item j^* so that buyer i^* is ε-happy if he purchases j^* at such a price. The advantage is that we have now $P \geq p^n(j^*) + \varepsilon$ and there is a minimal increment.

It is easy to see an analogy with auction systems and that this ε-system imposes a minimal bid.

Now, we can see that every time that an item is chosen by a buyer and receives a bid, then, from that moment on, whoever owns such an item will be ε-happy. Indeed, at the very moment when a buyer gets an object by a bid, he is ε-happy, and the price of such an object cannot increase as long as he owns it, as well as the prices of the others cannot decrease. Hence, the algorithm will stop as soon as all items will have received at least a bid each. Now, suppose that there is an item which has received no bids: his price is constant, while the prices of the other increase. It is not difficult to prove that the number of possible iterations going on in the presence of an object which has never received a bid must be bounded (as soon as there is one unbidded good, all the others become too expensive after a certain number of received bids, and when they become too expensive, they receive no more bids). This means that the algorithm stops after a finite number of steps.

Once we have a pair (p, σ) which is a equilibrium in the sense that every buyer is ε-happy, then we can prove that the cost $\sum_i u_{i\sigma(i)}$ is almost optimal among permutations $\tilde{\sigma}$. Indeed, take any other permutation $\tilde{\sigma}$. We have

$$u_{i\sigma(i)} - p_{\sigma_i} \geq \max_j(u_{ij} - p_j) - \varepsilon \geq u_{i\tilde{\sigma}(i)} - p_{\tilde{\sigma}(i)} - \varepsilon.$$

Summing over i, and using $\sum_i p_{\sigma_i} = \sum_i p_{\tilde{\sigma}(i)}$, we get

$$\sum_i u_{i\sigma(i)} \geq \sum_i u_{i\tilde{\sigma}(i)} - N\varepsilon.$$

This proves that σ is optimal, up to an error of $N\varepsilon$. By the way, if the utilities u_{ij} are integer numbers and $N\varepsilon < 1$, then σ must actually be optimal, with no error.

A rough estimate of the number of iterations which are necessary to conclude the algorithm is obtained in the following way: set $C = \max |u_{ij}|$ and start with $p^0 = 0$. Till there is at least an unbidded good, no other good can receive more than C/ε bids. This means that the number of iterations is at most of the order of NC/ε. In the case of integer utilities, an exact solution is found in $O(CN^2)$ iterations[6].

[6]Observe that this assumption on the values u_{ij} being integer, corresponding to integer costs on each edge in the network interpretation, appears quite often and simplifies the complexity of other algorithms as well. For instance, in [245] a polynomial bound on the cost of a network simplex

We remark that this cost can be indeed very high because of the presence of the factor C. Indeed, if the costs are rational numbers between 0 and 1 with k digits, if we want to make them integers, we could lead to multiply up to $C = 10^k$. A way to circumvent this problem is to make an iterative ε-scaling procedure. We first find an assignment which is $N\varepsilon$-optimal for a certain value of ε, then we divide ε by two, and we go on. This procedure, introduced in [55], is also detailed in [234]. We refer to these notes and to [54] for considerations about possible accelerated variants and computational costs of this method, very efficient in many cases. We also mention that in Section 8.4.4 we will briefly describe a cosmological model which has been numerically attacked thanks to the auction algorithm.

Entropic regularization and iterated projections

We would like to discuss another, very recent, but simple to explain, approach to the discrete linear programming version of the optimal transport problem. The starting point is the following variant of the linear problem (6.2): fix $\varepsilon > 0$ and look at

$$\min\left\{\sum_{i,j}\left(c_{ij}\gamma_{ij} + \varepsilon\gamma_{ij}\log(\gamma_{ij})\right) \;:\; \gamma_{ij} \geq 0,\ \sum_i \gamma_{ij} = b_j,\ \sum_j \gamma_{ij} = a_i\right\}. \tag{6.3}$$

It is clear that, for $\varepsilon \to 0$, the above minimization problem converges (in the sense of Γ-convergence, see Box 4.6) to the minimization in (6.2). We try to describe now why this approximated problem is easier to handle.

First, we rewrite the objective function using

$$c_{ij}\gamma_{ij} + \varepsilon\gamma_{ij}\log(\gamma_{ij}) = \varepsilon\gamma_{ij}\log\left(\frac{\gamma_{ij}}{\eta_{ij}}\right), \quad \sum_{i,j}\left(c_{ij}\gamma_{ij} + \varepsilon\gamma_{ij}\log(\gamma_{ij})\right) = \varepsilon\mathrm{KL}(\gamma|\eta),$$

where $\eta_{ij} = e^{-c_{ij}/\varepsilon}$ and KL denotes the so-called *Kullback-Leibler* divergence. This is not at all a divergence in the sense of differential calculus, but rather a sort of distance based on the relative entropy:

$$\mathrm{KL}(\gamma|\eta) := \sum_{i,j} f\left(\frac{\gamma_{ij}}{\eta_{ij}}\right)\eta_{ij} \quad \text{for } f(t) = \begin{cases} t\log t & \text{if } t \geq 0,\\ +\infty & \text{if } t < 0.\end{cases}$$

Hence, the above minimization problem reads as the projection, for the Kullback-Leibler divergence, of the point $\eta \in \mathbb{R}^{N\times N}$ on the set of constraints $C = C^x \cap C^y$, where $C^x = \{\gamma \in \mathbb{R}^{N\times N} : \sum_j \gamma_{ij} = a_i\}$ and $C^y = \{\gamma \in \mathbb{R}^{N\times N} : \sum_i \gamma_{ij} = b_j\}$. It is important to notice that the positivity constraint on γ has disappeared, because it is included in the definition of the entropy function f, which acts as a sort of barrier

algorithm is given under this assumption, obtaining $O(\min(kh\log(kC), kh^2\log k)$ for a network with k nodes, h edges, and maximal cost equal to C.

for this constraint (it is not a true barrier in the usual sense because it does not tend to infinity as the variable tends to 0; yet, it is set to infinity on negative numbers, and the derivative tends to infinity at 0).

The Kullback-Leibler divergence is not a distance, but shares some of the properties of distances. In particular it can be proven that the projection onto the intersection of two linear subspaces such as C^x and C^y can be obtained by alternate projections, i.e., defining $\gamma^{2k+1} = Proj_{C^x}(\gamma^{2k})$ and $\gamma^{2k+2} = Proj_{C^y}(\gamma^{2k+1})$, with $\gamma^0 = \eta$, and looking at the limit as $k \to \infty$.

The problem becomes that of the projection on one on the two subspaces, but this happens to be doable explicitly. Given a certain $\overline{\gamma}$, we want to solve (for instance) the minimization problem

$$\min \left\{ \sum_{i,j} \gamma_{ij} \log\left(\frac{\gamma_{ij}}{\overline{\gamma}_{ij}}\right) \;:\; \sum_j \gamma_{ij} = a_i \right\}.$$

It can be seen that the problem is separable according to the variable i (i.e., that every i can be treated separately, and the optimality condition gives $\log(\gamma_{ij}) = \log(\overline{\gamma}_{ij}) + \lambda_i$, i.e., $\gamma_{ij} = p_i \overline{\gamma}_{ij}$, for some constants p_i which are determined thanks to the constraint (i.e., $p_i = a_i/(\sum_j \overline{\gamma}_{ij})$).

Analogously, it can be proven that the projection onto C^y has the form $\gamma_{ij} = q_j \overline{\gamma}_{ij}$, for $q_j = b_j/(\sum_i \overline{\gamma}_{ij})$).

As a consequence, one can solve the approximated problem by means of a sequence of iterations of the following algorithm: start from $\gamma^0 = \eta$, define $\gamma_{ij}^{2k+1} = p_i \gamma_{ij}^{2k}$ with $p_i = a_i/(\sum_j \gamma_{ij}^{2k})$, and then $\gamma_{ij}^{2k+2} = q_j \gamma_{ij}^{2k+1}$ with $q_j = b_j/(\sum_i \gamma_{ij}^{2k+1})$. The limit as $k \to \infty$ provides the desired solution of the "regularized" problem (the regularization being given by the addition of the entropy) (Figure 6.2).

This idea, called *iterative proportional fitting procedure* (IPFP), can be traced back to [142] and even [282], and it has been recently applied to many optimal transport problems in [38]. It shows that the choice of the entropic regularization has several advantages:

- it allows to get rid of the positivity constraint (which would not be the case for a quadratic regularization);
- it allows to compute explicitly the projection at every step, as soon as we only project on one subspace C^x or C^y only;
- it is not too costly in terms of memory storage, since at every step the current γ^{2k} is of the form $\gamma_{ij}^{2k} = \eta_{ij} p_i q_j$ (cumulating the product of the coefficients p_i and q_j appearing at every step), which allows to stock $2N$ data instead of N^2;
- it is not too costly at every step, as it only requires $O(N)$ operations;
- it can be easily parallelized, due to the separable structure of each step;
- it can be adapted to multi-marginal problems and other linear problems.

It is also possible to deduce, from the limit value of the coefficients p and q, the Kantorovich potentials φ and ψ, but we leave it as an exercise (see **Ex** (42)).

6.4.2 *Semidiscrete numerical methods*

After the study of two discrete measures, the case of one discrete measure and one continuous deserves a short discussion. We are speaking of the case where $\mu = f(x)\,\mathrm{d}x$ has a density and $\nu = \sum_{j=1}^{N} b_j \delta_{y_j}$ is finitely atomic.

The main ideas behind the methods that are used rely on the dual problem,[7] in the following formulation:

$$\max\left\{ F(\psi) = \sum_j \psi_j b_j + \int_\Omega \psi^c(x) f(x)\,\mathrm{d}x \right\},$$

where the numbers ψ_j represent the values of the function ψ at the points y_j and ψ^c is the c-transform of ψ. The considerations in Chapter 1 explain why solving this problem gives information on the primal problem $\min\{\int c\,\mathrm{d}\gamma \,:\, \gamma \in \Pi(\mu,\nu)\}$. In particular, once we find the optimal function ψ, the set of points x satisfying $\psi_j + \psi^c(x) = c(x, y_j)$ will provide the set of points transported onto the target point y_j. We note that $\psi^c(x) = c(x, y_j) - \psi_j$ exactly means that $c(x, y_j) - \psi_j \le c(x, y_{j'}) - \psi_{j'}$ for all j'. This gives rise to the study of some sorts of Voronoi cells, whose definition also involves some scalar coefficients ψ_j. We will analyze these cells in the simpler case where $c(x, y) = \frac{1}{2}|x - y|^2$. In this case we define

$$V_\psi(j) = \left\{ x \,:\, \frac{1}{2}|x - y_j|^2 - \psi_j \le \frac{1}{2}|x - y_{j'}|^2 - \psi_{j'} \ \text{ for all } j' \right\}.$$

By expanding the squares, we see that each $V_\psi(j)$ is defined by $N - 1$ linear inequalities of the form $x \cdot (y_{j'} - y_j) \le a(j, j') := \psi_j - \psi_{j'} + \frac{1}{2}|y_{j'}|^2 - \frac{1}{2}|y_j|^2$ and is hence a convex polyhedron. These cells are called *power cells*. They coincide with the Voronoi cells of the points y_j whenever all the values ψ_j are equal.

The problem consists in finding values ψ_j such that the masses $\int_{V_\psi(j)} f(x)\,\mathrm{d}x$ equal the prescribed values b_j. This is indeed equivalent to maximizing the quantity $F(\psi)$. The maximization of F is a finite-dimensional issue, since (thanks to ν being finitely atomic) the only values of ψ which are relevant are those taken at the points y_j. Hence, we can consider $\psi \in \mathbb{R}^N$, and look at the derivatives of F.

First, we note that F is concave in ψ. Indeed, the first part of F is linear: $\psi \mapsto \sum_j \psi_j b_j$. Then, we take the values $\psi^c(x) = \inf_j c(x, y_j) - \psi_j$; these values are infima of linear functions of ψ and thus are concave. They are weighted with a positive density f and then integrated; the result is a concave function of ψ. Let us compute the derivative of F with respect to one variable ψ_j: from

[7] See also Example 1.6 in [176] for a first genesis of these ideas.

Fig. 6.2 The optimal γ in a 1D transport problem (the starting measure is a centered Gaussian; the arrival measure is composed of two bumps), via IPFP methods with decreasing values of ε. The image on the right gives a quite precise description of the optimal (monotone) map. Pictures kindly provided by L. Nenna

$$\frac{\partial \psi^c(x)}{\partial \psi_j} = \begin{cases} -1 & \text{if } c(x,y_j) - \psi_j < c(x,y_{j'}) - \psi_{j'} \text{ for all } j' \neq j, \\ 0 & \text{if } c(x,y_j) - \psi_j > \psi^c(x), \\ \text{not defined} & \text{if } c(x,y_j) - \psi_j = \psi^c(x) = c(x,y_{j'}) - \psi_{j'} \text{ for some } j' \neq j, \end{cases}$$

and from the fact that the last of these three cases is negligible (at least in the case $c(x,y) = |x-y|^2$ where the equality case reduces to a subset of a hyperplane), we deduce that F is indeed differentiable everywhere, and[8]

$$\frac{\partial F}{\partial \psi_j} = b_j - \int_{V_\psi(j)} f(x)\,\mathrm{d}x.$$

This shows that the maximization of F is equivalent to $b_j = \int_{V_\psi(j)} f(x)\,\mathrm{d}x$ and also allows to look for the maximum of F via gradient methods. To this aim, it is also useful to check upper bounds on the second derivatives of F. We restrict again our attention to the case $c(x,y) = \frac{1}{2}|x-y|^2$. We note that moving the value of the coordinates of ψ from ψ_j to $\psi_j + h_j$ makes all the cells $V_\psi(j')$ change, but only by changing the constants $a(j,j')$ of a value $h_{j'} - h_j$. This means that the hyperplanes defining their boundaries move to a distance of the order of $\frac{|h|}{|y_j - y_{j'}|}$. If the points y_j are fixed and distinct, the measure of the symmetric difference $V_\psi(j') \Delta W_{\psi+h}(j')$ is bounded by $C|h|$. If f is L^∞, then one obtains

$$\left| \frac{\partial F}{\partial \psi_j}(\psi) - \frac{\partial F}{\partial \psi_j}(\psi + h) \right| \leq C|h|, \tag{6.4}$$

and $F \in C^{1,1}$.

[8]The computation of the gradient, and its continuity, could have been simplified by looking at the subdifferential; see [24].

Box 6.8. Memo: *Gradient algorithm*

If $F : \mathbb{R}^N \to \mathbb{R}$ is a differentiable convex function, an easy algorithm to look for the minimum of F is the following: take $x_0 \in \mathbb{R}^N$ and $\tau > 0$ a fixed size, and define

$$x_{k+1} = x_k - \tau \nabla F(x_k).$$

If F is such that $\alpha \mathrm{I} \leq D^2F(x) \leq L\mathrm{I}$ for every x and some $\alpha, L > 0$, then the above algorithm converges exponentially whenever $\tau < 2\alpha/L^2$. More precisely, F is strictly convex and has a unique minimizer $\bar{x}$, and we have $|x_k - \bar{x}| \leq C\beta^k$, where β depends on τ, α, L, and $\beta < 1$ if $\tau \in]0, 2\alpha/L^2[$. The convergence $F(x_k) - \min F$ is also exponential.

If D^2F is bounded from above but not from below (i.e., F is smooth but not enough convex), but admits a minimizer $\bar{x}$, then the rate of convergence is much worse: indeed, one can only prove $F(x_k) - F(\bar{x}) \leq C/k$, as soon as $\tau M < 1$.

Many texts present a wide introduction to gradient methods in convex optimization; see, for instance, [64] for more details. Variants where the step size τ varies with k, both in prescribed and optimized manners, also exist.

As a final point, note that the above algorithm also corresponds to an explicit Euler scheme in the discretization of the gradient flow of F (see the Memo Box 8.1).

Then, the practical implementation of a gradient algorithm of this kind depends on how one can practically compute the power cells $V_\psi(j)$ and the integral of the density f on them. This is a matter of the so-called *computational geometry*. We do not spend word on this very delicate issue, but refer, for instance, to the recent papers [209, 233] for a detailed discussion. The original implementation of these methods is proposed in [233], and the implementation has been improved in [209], where spectacular numerical simulations [9] of optimal transport between geometric shapes are also presented.

Due to the very good numerical results, which we present through Figures 6.3 and 6.4, we prefer to spend some extra words on the most efficient methods to be used in this semi-discrete setting.

Indeed, even if the description of the algorithm was easier in the case of a gradient algorithm, the convergence is much quicker if one uses instead a Newton approach. Newton's method (see Box 6.4 in Section 6.3) is suitable to find zeros of a vector function from $\mathbb{R}^N$ to $\mathbb{R}^N$, which is in this case the function associating with a dual potential $\psi = (\psi_i)_i \in \mathbb{R}^N$ the values $b_i - \int_{V_\psi(i)} f(x)\,\mathrm{d}x = \partial F/\partial\psi_i$. In order to be able to use the Newton method, one needs to be able to compute their derivatives, which are indeed the entries of the Hessian matrix of F. One also needs to justify that this Hessian matrix is nonsingular and provide lower bounds on its eigenvalues. This corresponds to uniform convexity of F.

Moreover, we know that the Newton method is only efficient when we are able to choose a suitable initialization for the algorithm, since it only converges

[9] The pictures in the next pages have been kindly provided by B. Lévy and are produced with the same algorithms as in [209].

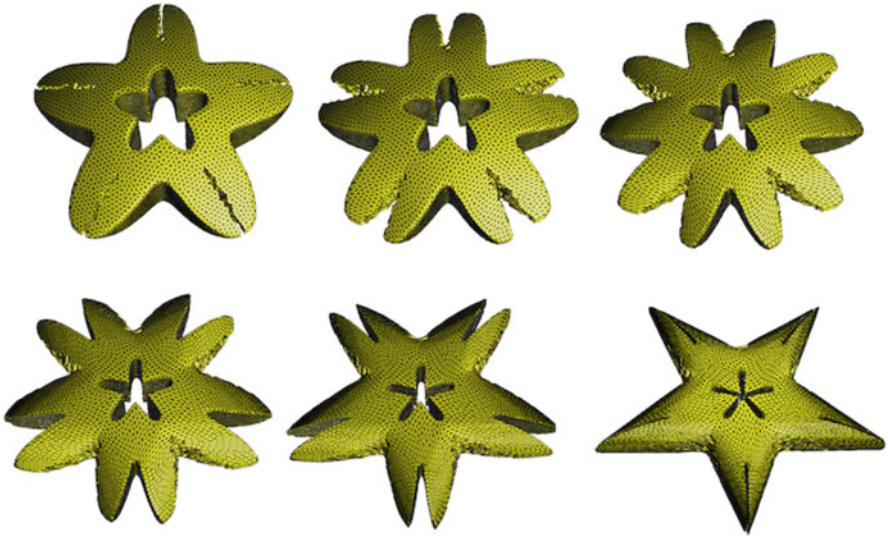


Fig. 6.3 W_2 geodesic between two different stars in 3D, computed via semi-discrete optimal transport on a simplex triangulation. Once obtained the images of every vertex x_i of the triangulation, linear interpolation between each x_i and $\mathrm{T}(x_i)$ is computed. Note the topology change and the singularities of the optimal map on every branch of the star

starting from a certain neighborhood. This is efficiently addressed in [233], where the problem is considered with a multi-resolution point of view. We associate with the family of points y_j a number of smaller families, each obtained from the previous one via a Lloyd algorithm and dividing the number of points by a fixed ratio. In this way there is a correspondence such that every point in a family has a unique image in the next one (which is smaller). We start from the computation of a potential ψ for the smallest family, which is easy to do. Then, at every step, the potential ψ computed for the ith family is used to build an initialization for the Newton method of the next step, the one for the $(i-1)$th family of points. This guarantees that we are always sufficiently close to the minimizer that we look for and that the algorithm converges in a very small number of iterations. In practice, the implementations of [209] can handle one million points in some minutes.

We come back now to the justification of uniform convexity, which is required to apply the Newton procedure. Actually, the function F is far from being strictly convex: indeed, it does not change if one adds a same constant to all the ψ_i. Yet, up to this invariance, we can expect it to be strictly convex. In order to do that, one needs to fix a value $\psi_{i_0} = 0$ and to look at the behavior w.r.t. ψ_i for $i \neq i_0$.

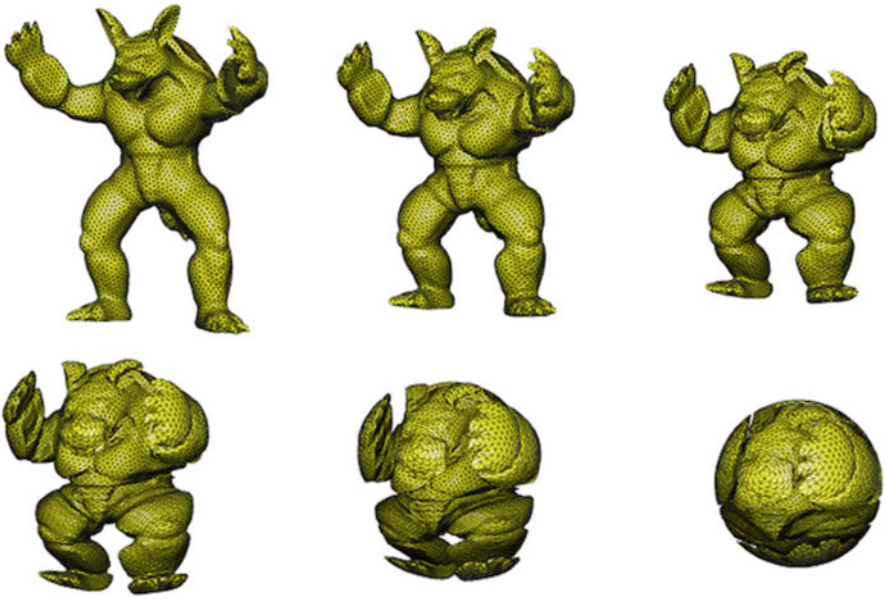


Fig. 6.4 W_2 geodesic between the uniform measure on the shapes of an "armadillo" and of a sphere (discretized with up to a million of Dirac masses)

How to compute the second derivatives of F? It is enough to improve the computations leading to (6.4): if f is a continuous density, then one obtains

$$\frac{\partial^2 F}{\partial \psi_j \partial \psi_j}(\psi) = -\frac{1}{|y_i - y_j|} \int_{\Gamma_{ij}} f(x)\, d\mathscr{H}^{d-1}(x), \qquad \text{for } i \neq j,$$

where $\Gamma_{ij} = \partial V_\psi(i) \cap \partial V_\psi(j)$ is the common boundary (if any) of the two cells corresponding to the indices i and j (which is $(d-1)$-dimensional polyhedron). For the pure derivatives, one has

$$\frac{\partial^2 F}{\partial^2 \psi_i}(\psi) = \sum_{j \neq i} \frac{1}{|y_i - y_j|} \int_{\Gamma_{ij}} f(x)\, d\mathscr{H}^{d-1}(x) = -\sum_{j \neq i} \frac{\partial^2 F}{\partial \psi_j \partial \psi_j}(\psi).$$

Note that in these sums, we also keep the index i_0. This allows to write a Hessian in N variables which is singular since the sum of every line is 0. Then, we withdraw the line and column corresponding to the index i_0, and we have a square $(N-1) \times (N-1)$ matrix M. In order to check inversibility of M, we use the following lemma:

Lemma 6.5. *Any matrix $M \in \mathrm{M}^{k \times k}$ satisfying the following properties is invertible:*

(H1) for all i, we have $M^{i,i} \geq \sum_{j \neq i} |M^{i,j}|$,
(H2) there exists i such that $M^{i,i} > \sum_{j \neq i} |M^{i,j}|$,

(H3) for any pair (i,j), there is a sequence $i_1, i_2, \dots, i_k$ with $i_1 = i$, $i_k = j$ satisfying $M^{i_h,i_{h+1}} \neq 0$.

Proof. Let $x \in \mathrm{Ker}(M)$ and let i_0 be an index such that $|x_i|$ is maximal. We may suppose for simplicity that x_{i_0} is positive. Then we have

$$0 = M^{i_0,i_0}x_{i_0} - \sum_{j\neq i_0} M^{i_0,j}x_j \geq M^{i_0,i_0}x_{i_0} - \sum_{j\neq i_0} |M^{i_0,j}|x_{i_0} = x_{i_0}\left(M^{i_0,i_0} - \sum_{j\neq i_0}|M^{i_0,j}|\right) \geq 0.$$

This implies that all inequalities are equalities and in particular $x_j = x_{i_0}$ whenever $M^{i_0,j} \neq 0$. Hence, the entries of x on all the indices which are "neighbors" of i_0 equal to x_{i_0} (and they are maximal as well). This allows to repeat the argument replacing i_0 with another maximizing index j and so on…; since any index is connected by a chain of neighbors to i_0, we get that all the entries are equal. But this implies that the vector that we selected in the kernel must be a multiple of the vector $(1, 1, \dots, 1)$. Yet, this vector is not in the kernel since the sum of the elements on each line is not zero for all lines, by assumption $(H2)$. This proves that M is invertible. □

The above lemma can be applied to the Hessian of F under some reasonable assumptions (for instance, we should require that the support of the density f is arcwise connected).

This guarantees that the Hessian of F is a strictly positive-definite matrix (on the space $\mathbb{R}^{N-1}$ of potentials with fixed $\psi_{i_0} = 0$) at every point. In order to apply a Newton's method, one should also bound from below its eigenvalues. This is more delicate, and we refer to [114] both for a counterexample under the three assumptions of the above Lemma, which are not enough for such a bound, and for a proof in the specific case where the matrix M comes from a partition into power cells.

Moreover, as we are dealing with the arguments of [114], let us underline another semi-discrete method, described in the same paper. The idea of the method comes from the considerations of our Section 2.4 about Knothe transport. Indeed, as we proved that the Knothe map is the limit of the optimal transport maps for degenerate quadratic costs corresponding to a matrix $\mathrm{diag}(1, \varepsilon, \varepsilon^2 \dots, \varepsilon^{d-1})$, we can start from such a map (Knothe) and then let ε vary. The Knothe map is easy to compute, and one can find the cells corresponding to it[10]. Since for the Knothe map the most important coordinate is the last one, these cells are horizontal strips with prescribed

[10]This is the semidiscrete method coming from this idea, implemented in [114]. There is also a continuous counterpart, i.e., a PDE on the Kantorovich potentials ψ_ε. It has been studied from the theoretical point of view in [66], where Bonnotte proved a nontrivial well-posedness result, based on the application of the Nash-Moser implicit function theorem around $\varepsilon = 0$ (once we are on $\varepsilon > 0$, the equation is more regular). Later, the same author also started numerical implementations of the time discretization of this equation in [65]. Note that his continuous method recalls in some aspects both the AHT flow described in Section 6.2 (as it starts from the Knothe map, is continuous in time but discretized for numerical purposes, and evolves by imposing prescribed image measure

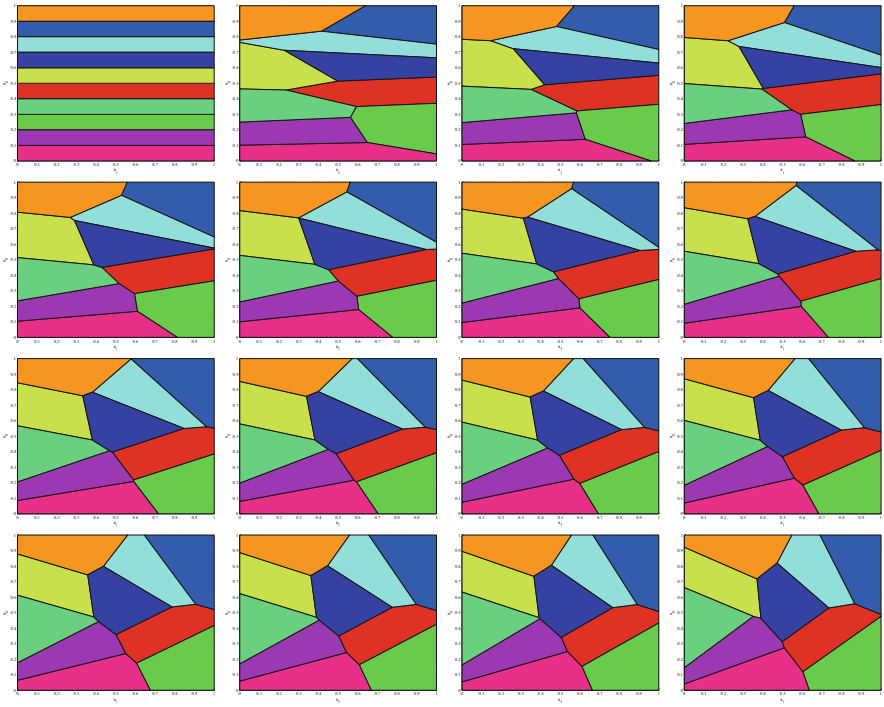


Fig. 6.5 Ten sample points: evolution of the tessellation for $\varepsilon = 0$ to $\varepsilon = 1$ (from top left to bottom right). Picture taken from the preliminary studies leading to [114]

volume. Note that we need to suppose that they have different x_d components for the method to work. After the computation of the initial condition, one runs an ODE in dimension $N-1$ (N being the number of atoms at the arrival) and finds the evolution of the Kantorovich potential $\psi(\varepsilon)$. The computations needed in order to prove that the ODE is well-posed are similar to the ones above and allows to conclude that the evolution fits into the standard existence, uniqueness, stability, and approximation theory for Cauchy problems with Lipschitz vector fields. Hence, one can apply a time discretization in the form of an explicit Euler scheme to approximate the solution. Observing the evolution of the potentials ψ_ε as ε runs from 0 to 1, we can also compute the corresponding cells, which move from rectangular strips to convex polyhedral power cells.

We show in Figure 6.5 a simple sequence of cells obtained with the above method.

at every time) and the LR Newton's iterations of Section 6.3 (as the underlying PDE is based on a linearization of the Monge-Ampère equation).

Chapter 7
Functionals on the space of probabilities

We consider in this chapter various classes of functionals on the space $\mathscr{P}(\Omega)$, which can be of interest in many variational problems, and are natural in many modeling issues. Indeed, in several applied models, we can face minimization problems where the unknown to be determined is the distribution of a certain amount of mass, and the criteria involve one or more of the following functionals:

- The integral of a given function (*potential energy*)

$$\mathscr{V}(\mu) := \int_\Omega V \, \mathrm{d}\mu.$$

- The double integral of a function on $\Omega \times \Omega$ according to the tensor product $\mu \otimes \mu$ (*interaction energy*)

$$\mathscr{W}(\mu) := \int_{\Omega\times\Omega} W(x,y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y).$$

- The Wasserstein distance W_p (or a function of it) from a fixed measure ν; for simplicity, we consider rather the pth power of W_p:

$$\mu \mapsto W_p^p(\mu,\nu).$$

  More generally, we will also consider transport cost $\mathscr{T}_c(\mu,\nu)$ defined as usual as $\min\{\int c(x,y) \, \mathrm{d}\gamma \, : \, \gamma \in \Pi(\mu,\nu)\}$, for general costs $c(x,y)$.
- The norm in a dual functional space: given a Banach space of functions on Ω, we take

$$||\mu||_{\mathscr{X}'} := \sup_{\varphi\in\mathscr{X},||\varphi||\le 1} \int \varphi \, \mathrm{d}\mu = \sup_{\varphi\in\mathscr{X}\setminus\{0\}} \frac{\int \varphi \, \mathrm{d}\mu}{||\varphi||}.$$

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2_7

- The integral of a function of the density

$$\mathscr{F}(\mu) := \begin{cases} \int_\Omega f(\varrho(x))\,\mathrm{d}x & \text{if } \mu = \varrho \cdot \mathrm{d}x, \\ +\infty & \text{otherwise.} \end{cases}$$

- The sum of a function of the masses of the atoms

$$\mathscr{G}(\mu) := \begin{cases} \sum_i g(a_i) & \text{if } \mu = \sum_i a_i \delta_{x_i}, \\ +\infty & \text{if } \mu \text{ is not purely atomic.} \end{cases}$$

The scope of this chapter is to study some properties of these functionals. The first questions that we analyze are classical variational issues (semi-continuity, convexity, first variation, etc.). Then, we also introduce and analyze a new notion, the notion of geodesic convexity. This is a natural concept in the analysis on metric spaces: is a certain functional convex *along geodesics* of the metric space? In the case of the Wasserstein spaces, and in particular of $\mathbb{W}_2$, this concept, introduced by McCann in [229], is called *displacement convexity*. It turns out to be quite different than usual convexity, but very useful in many cases.

7.1 Semi-continuity

Since the main goal is to use these functionals in variational models, in order to be able to minimize them, we first give some semi-continuity criteria. We will consider semi-continuity or continuity of these functionals with respect to weak convergence (meaning as usual the duality with C_b functions). We start from the most simple functionals.

7.1.1 Potential and interaction energies, transport costs, dual norms

We start this review from the easiest case: the potential energy.

Proposition 7.1. *If $V \in C_b(\Omega)$, then $\mathscr{V}$ is continuous for the weak convergence of probability measures. If V is l.s.c. and bounded from below, then $\mathscr{V}$ is semi-continuous. Moreover, semi-continuity of V (respectively, continuity) is necessary for the semi-continuity (continuity) of $\mathscr{V}$.*

Proof. The continuity of $\mu \mapsto \int V\,\mathrm{d}\mu$ for $V \in C_b$ is straightforward by definition of weak convergence. If V is l.s.c. and bounded from below, we know that there

is a sequence of Lipschitz and bounded functions V_k increasingly converging to V (see Box 1.5). Then, by monotone convergence, we infer that $\mathscr{V}(\mu) = \lim_k \int V_k \, d\mu = \sup_k \int V_k \, d\mu$. This allows to see $\mathscr{V}$ as a supremum of continuous functionals, and hence it is l.s.c.

The necessity part is straightforward if one considers sequences of points $x_k \to x$ and the associated Dirac masses. Since $\delta_{x_k} \rightharpoonup \delta_x$, then continuity of $\mathscr{V}$ implies $V(x_k) \to V(x)$ and semi-continuity implies $\liminf_k V(x_k) \geq V(x)$. □

We pass now to the case of the interaction functional $\mathscr{W}$.

Proposition 7.2. *If $W \in C_b(\Omega)$, then $\mathscr{W}$ is continuous for the weak convergence of probability measures. If W is l.s.c. and bounded from below, then $\mathscr{W}$ is semi-continuous.*

This result is an easy consequence of the following lemma (which we prove for subsets of $\mathbb{R}^d$ but could be adapted to general locally compact spaces).

Lemma 7.3. *Let Ω be a closed subset of $\mathbb{R}^d$, $\mu_k, \mu \in \mathscr{P}(\Omega)$ and $\mu_k \rightharpoonup \mu$, then*

$$\mu_k \otimes \mu_k \rightharpoonup \mu \otimes \mu$$

as probabilities on $\Omega \times \Omega$.

Proof. We want to prove that for any function $\phi \in C_b(\Omega \times \Omega)$, we have the convergence $\int \phi \, d\mu_k \otimes \mu_k \to \int \phi \, d\mu \otimes \mu$. First consider the case $\phi(x, y) = \chi(x)\zeta(y)$. In this case

$$\int_{\Omega\times\Omega} \phi \, d\mu_k \otimes \mu_k = \int_\Omega \chi \, d\mu_k \cdot \int_\Omega \zeta \, d\mu_k \to \int_\Omega \chi \, d\mu \cdot \int_\Omega \zeta \, d\mu = \int_{\Omega\times\Omega} \phi \, d\mu \otimes \mu.$$

This proves that the desired convergence is true for all functions $\phi \in \mathscr{A}(\Omega)$, where the class $\mathscr{A}(\Omega)$ is given by

$$\mathscr{A}(\Omega) = \left\{ \phi(x, y) = \sum_{i=1}^{N} \chi_i(x)\zeta_i(y), \quad N \in \mathbb{N}, \ \chi_i, \zeta_i \in C_b(\Omega) \right\}.$$

The class $\mathscr{A}$ is an algebra of functions which contains the constants and separates the points of $\Omega \times \Omega$. Hence, by Stone-Weierstrass's theorem (see Box 2.3 in Section 3.1), if $\Omega \times \Omega$ is compact (i.e., if Ω is compact), it is dense in $C(\Omega \times \Omega)$. It is a general fact that weak-* convergence in a dual space may be tested on a dense set, provided that the sequence we are considering is bounded. Here weak convergence corresponds in the compact case to the weak-* convergence in the duality with $C(\Omega \times \Omega)$, and we are considering a sequence of probability measures, which is bounded. This proves $\mu_k \otimes \mu_k \rightharpoonup \mu \otimes \mu$ in the compact case.

If Ω is not compact, one can use Stone-Weierstrass's theorem to prove that the convergence holds against any compactly supported function. The conclusion follows if we recall that a sequence of probability measures weakly converges (in the duality with C_b) if and only if it weakly-* converges in the duality with C_c. This was proven in Lemma 5.8. □

It is worthwhile to note that besides these two classes of functionals, one could consider higher-order interaction energies defined through

$$\mu \mapsto \int_\Omega \int_\Omega \dots \int_\Omega W(x_1, x_2, \dots, x_n)\, \mathrm{d}\mu(x_1)\, \mathrm{d}\mu(x_2) \dots \mathrm{d}\mu(x_n).$$

In particular the set of all these functional could be considered as "polynomials" on the space $\mathscr{P}(\Omega)$, and some analyses of the space of functions over measures are based on this very class of functionals[1].

We pass now to some different functionals.

Proposition 7.4. *For any $p < +\infty$, the Wasserstein distance $W_p(\cdot, \nu)$ to any fixed measure $\nu \in \mathscr{P}(\Omega)$ is continuous w.r.t. weak convergence provided Ω is compact. If Ω is not compact and $\nu \in \mathscr{P}_p(\Omega)$, then $W_p(\cdot, \nu)$ is well defined and finite valued over $\mathscr{P}_p(\Omega)$ and it is only l.s.c.*

More generally, for any continuous cost $c : \Omega \times \Omega \to \mathbb{R}$, the functional $\mathscr{T}_c(\cdot, \nu)$ is continuous if Ω is compact. Without compactness on Ω, if c is l.s.c. and bounded from below, then $\mathscr{T}_c(\cdot, \nu)$ is l.s.c.

Proof. We start from the compact case. For Wasserstein distances, continuity is straightforward since the convergence for the distance W_p exactly metrizes the weak convergence and every distance is continuous w.r.t. itself (as a consequence of triangle inequality). For general transport costs $\mathscr{T}_c$, just apply Theorem 1.51.

For the non-compact case, we give a unified proof. Consider a sequence $\mu_k \rightharpoonup \mu$ and a sequence of optimal transport plans $\gamma_k \in \Pi(\mu_k, \nu)$ for a cost c, l.s.c. and bounded from below (which includes the case $c(x, y) = |x - y|^p$). Since μ_k is tight, γ_k is also tight. First extract a subsequence such that $\lim_h \mathscr{T}_c(\mu_{k_h}, \nu) = \liminf_k \mathscr{T}_c(\mu_k, \nu)$ and then extract once more so as to guarantee $\gamma_{k_h} \rightharpoonup \gamma$ for some γ. We know that image measures through continuous functions pass to the limit, so that we obtain $\gamma \in \Pi(\mu, \nu)$. Hence,

$$\mathscr{T}_c(\mu, \nu) \le \int c(x, y) \mathrm{d}\gamma \le \liminf_h \int c(x, y) \mathrm{d}\gamma_{k_h} = \lim_h \mathscr{T}_c(\mu_{k_h}, \nu) = \liminf_k \mathscr{T}_c(\mu_k, \nu).$$

Here the first inequality comes from the fact that γ is admissible but maybe not optimal. The second from semi-continuity of the integral of c (Proposition 7.1). □

Finally, here is another class of functionals with some interesting examples.

[1] This is an approach used by P. L. Lions in his lectures at Collège de France; see [211].

Proposition 7.5. *Let $\mathscr{X}$ be a Banach space of functions over Ω such that $\mathscr{X} \cap C_b(\Omega)$ is dense in $\mathscr{X}$. Then*

$$\mu \mapsto ||\mu||_{\mathscr{X}'} := \sup_{\varphi \in \mathscr{X}, ||\varphi|| \leq 1} \int \varphi \, \mathrm{d}\mu = \sup_{\varphi \in \mathscr{X} \setminus \{0\}} \frac{\int \varphi \, \mathrm{d}\mu}{||\varphi||}$$

(a functional which is set to $+\infty$ if $\mu \notin \mathscr{X}'$) is l.s.c. for the weak convergence.

Proof. This fact is straightforward if one notes that we can write

$$||\mu||_{\mathscr{X}'} = \sup_{\varphi \in C_b(\Omega) \cap \mathscr{X}, ||\varphi|| \leq 1} \int \varphi \, \mathrm{d}\mu,$$

which expresses $||\mu||_{\mathscr{X}'}$ as a supremum of linear functionals, continuous for the weak convergence. □

It is interesting to see two examples.

Proposition 7.6. *Suppose $\mathscr{X} = H^1_0(\Omega)$ (where $\Omega \subset \mathbb{R}^d$ is a smooth compact domain, and we endow this space with the L^2 norm of the gradient) and let, for every $\mu \in H^{-1}(\Omega)$, φ_μ be the solution of*

$$\begin{cases} -\Delta \varphi = \mu & \text{in } \Omega, \\ \varphi = 0 & \text{on } \partial\Omega \end{cases}.$$

Then $||\mu||^2_{H^{-1}} = \int \varphi_\mu \, \mathrm{d}\mu = \int\int G(x,y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y)$, where $G(x,\cdot)$ is the Green function of the Dirichlet Laplacian, defined as the solution of

$$\begin{cases} -\Delta \varphi = \delta_x & \text{in } \Omega, \\ \varphi = 0 & \text{on } \partial\Omega \end{cases}.$$

In particular, if $\Omega = (-1,1) \subset \mathbb{R}$, then

$$G(x,y) = \frac{1}{2}((1+x) \wedge (1+y))((1-x) \wedge (1-y)).$$

Suppose, instead, $\mathscr{X} = H^1(\Omega)$ (endowed with the full H^1 norm) and let, for every $\mu \in (H^1(\Omega))' = \mathscr{X}'$, φ_μ be the solution of

$$\begin{cases} -\Delta \varphi + \varphi = \mu & \text{in } \Omega, \\ \frac{\partial \varphi}{\partial n} = 0 & \text{on } \partial\Omega \end{cases}.$$

Then $||\mu||^2_{\mathscr{X}'} = \int \varphi_\mu \, \mathrm{d}\mu = \int\int G(x,y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y)$, where $G(x,\cdot)$ is defined as the solution of

$$\begin{cases} -\Delta\varphi + \varphi = \delta_x & \text{in } \Omega, \\ \frac{\partial\varphi}{\partial n} = 0 & \text{on } \partial\Omega \end{cases}.$$

In particular, if $\Omega = (-1, 1) \subset \mathbb{R}$*, then*

$$G(x, y) = \cosh((1 + x) \wedge (1 + y)) \cosh((1 - x) \wedge (1 - y))/\sinh(2).$$

Proof. We first prove that $||\mu||^2_{\mathscr{X}'} = \int \varphi_\mu \, \mathrm{d}\mu$ in the two cases.

First case: H^1_0. Note that, for every function $\psi \in H^1_0$, one has

$$\int \psi \, \mathrm{d}\mu = \int \nabla\psi \cdot \nabla\varphi_\mu \leq ||\psi||_{\mathscr{X}} ||\varphi_\mu||_{\mathscr{X}}$$

and equality holds for $\psi = \varphi_\mu$ (we use the homogeneous norm $||\psi||_{\mathscr{X}} := ||\nabla\psi||_{L^2}$). This shows that the supremum is realized by a multiple of φ_μ and that $\int \varphi_\mu \, \mathrm{d}\mu = ||\varphi_\mu||^2_{\mathscr{X}}$. As a consequence, we have

$$||\mu||^2_{\mathscr{X}'} = \left(\frac{\int \varphi_\mu \, \mathrm{d}\mu}{||\varphi_\mu||_{\mathscr{X}}} \right)^2 = \int \varphi_\mu \, \mathrm{d}\mu.$$

The case of H^1 is similar, with the only exception of the first computation

$$\int \psi \, \mathrm{d}\mu = \int \nabla\psi \cdot \nabla\varphi_\mu + \psi\varphi_\mu \leq ||\psi||_{\mathscr{X}} ||\varphi_\mu||_{\mathscr{X}}$$

and, again, equality holds for $\psi = \varphi_\mu$.

Finally, expressing $\int \varphi_\mu \, \mathrm{d}\mu$ as a double integral $\iint G(x, y) \, \mathrm{d}\mu(x) \, \mathrm{d}\mu(y)$ is only a matter of expressing $\varphi_\mu(x)$ as $\int G(x, y) \, \mathrm{d}\mu(y)$. This is possible by using the theory of Green functions, and for the 1D case $\Omega = (-1, 1)$, it is enough to compute that $\int G(x, y) \, \mathrm{d}\mu(y)$ is a solution of the desired equation. □

In this way we have seen that, thanks to Green functions, we have expressed these dual norm functionals as interaction functionals.

7.1.2 Local functionals

Local functionals over measures are defined as those functionals $F : \mathscr{M}(\Omega) \to \mathbb{R}$ such that $F(\mu + \nu) = F(\mu) + F(\nu)$ whenever μ and ν are mutually singular (i.e., there exists $A, B \subset \Omega$ with $A \cup B = \Omega$, $\mu(A) = 0$ and $\nu(B) = 0$). This obviously does not make much sense for probability measures, since if μ and ν are probabilities, then $\mu + \nu$ is not a probability. Yet, we will see in this section some functionals that can be defined over probability measures but could be also seen as

the restriction to probabilities of functionals defined on more general measures. All our proofs could be easily generalized to finite positive measures; on the contrary, for simplicity we will not consider at all the case of signed or vector measures. We refer to [67] for the general theory of these functionals.

First, let us consider functionals of the form

$$\mu = \varrho \cdot \lambda \mapsto \int f(\varrho(x))\mathrm{d}\lambda(x),$$

where λ is a given positive measure over Ω (for instance, the Lebesgue measure).

Let us first see which are the natural conditions on f so as to ensure lower semi-continuity. Then, we will give a precise result.

As a first point, consider a sequence ϱ_n of densities, taking values a and b on two sets A_n and B_n chosen in the following way: we fix a partition of Ω into small cubes of diameter $\varepsilon_n \to 0$ and then we select set A_n and B_n such that, for every cube Q of this partition, we have $\lambda(A_n \cap Q) = (1-t)\lambda(Q)$ and $\lambda(B_n \cap Q) = t\lambda(Q)$. We can check that ϱ_n weakly converges to a uniform density $\varrho = (1-t)a + tb$. Semi-continuity of the functional would imply

$$f((1-t)a + tb) \leq (1-t)f(a) + tf(b),$$

i.e., convexity of f. Hence it is clear that one needs to require f to be convex.

Another requirement concerns the growth of f. For simplicity it is always possible to assume $f(0) = 0$ (up to adding a constant to f). Suppose that f satisfies $f(t) \leq Ct$ for all $t \geq 0$. Consider the functional defined through

$$\mathscr{F}(\mu) : \begin{cases} \int f(\varrho(x))\,\mathrm{d}\lambda(x) & \text{if } \mu = \varrho \cdot \lambda, \\ +\infty & \text{otherwise.} \end{cases}$$

Take a sequence μ_n of absolutely continuous probability measures weakly converging to a singular measure μ: we get $\mathscr{F}(\mu_n) \leq C$ while $\mathscr{F}(\mu) = +\infty$, thus violating the semi-continuity. This suggests that one should have $C = +\infty$, i.e., f superlinear. The following theorem gives a general result which also includes compensation for non-superlinearity:

Proposition 7.7. *Let $f : \mathbb{R}_+ \to \mathbb{R}$ be a convex l.s.c. function, and set $L := \lim_{t\to\infty} f(t)/t = \sup_{t>0} f(t)/t \in \mathbb{R} \cup \{+\infty\}$. Let λ be a fixed finite positive measure on Ω. For every measure μ, write $\mu = \varrho \cdot \lambda + \mu^s$, where $\varrho \cdot \lambda$ is the absolutely continuous part of μ and μ^s be the singular one (w.r.t. λ). Then, the functional defined through*

$$\mathscr{F}(\mu) = \int_\Omega f(\varrho(x))\,\mathrm{d}\lambda(x) + L\mu^s(\Omega)$$

(note that if $L = +\infty$, then $F(\mu) = +\infty$ whenever μ has a singular part) is l.s.c.

Proof. We will try to use the equality

$$f(t) = \sup_{a \in \mathbb{R}} \quad at - f^*(a).$$

This equality is a consequence of $f^{**} = f$, which is itself due to the fact f is convex and lower semi-continuous. Moreover, it is easy to see that for $a > L$, we have $f^*(a) = +\infty$: indeed, $f^*(a) = \sup_t at - f(t) \geq \lim_{t \to +\infty} t(a - f(t)/t)$. Hence, we can write

$$f(t) = \sup_{a \leq L} \quad at - f^*(a) = \sup_{a < L} \quad at - f^*(a)$$

(the last equality is justified by the fact that f^* is continuous on the set where it is finite).

Let us consider the following functional:

$$\tilde{\mathscr{F}}(\mu) := \sup_{a \in C_b(\Omega),\, \sup a < L} \int a(x)\, \mathrm{d}\mu(x) - \int f^*(a(x))\, \mathrm{d}\lambda(x).$$

$\tilde{\mathscr{F}}$ is obviously l.s.c. since it is the supremum of a family of affine functionals, continuous w.r.t. the weak convergence. We want to prove $\tilde{\mathscr{F}} = \mathscr{F}$.

In order to do so, first note that, thanks to Lusin's theorem (see Box 1.6 in Section 1.1), applied here to the measure $\lambda + \mu$, it is not difficult to replace bounded and continuous functions with measurable bounded functions. By abuse of notation, we denote measurable bounded functions by $L^\infty(\Omega)$ (even if we do not mean functions which are essentially bounded w.r.t. a given measure, but really bounded) and we get

$$\tilde{\mathscr{F}}(\mu) := \sup_{a \in L^\infty(\Omega),\, \sup a < L} \int a(x)\, \mathrm{d}\mu(x) - \int f^*(a(x))\, \mathrm{d}\lambda(x).$$

Then take a set A such that $\mu^s(\Omega \setminus A) = \lambda(A) = 0$: this allows to write

$$\tilde{\mathscr{F}}(\mu) := \sup_{a \in L^\infty(\Omega),\, \sup a < L} \int_{\Omega \setminus A} \left[a(x)\varrho(x) - f^*(a(x))\right] \mathrm{d}\lambda(x) + \int_A a(x)\, \mathrm{d}\mu^s(x).$$

The values of $a(x)$ may be chosen independently on A and $\Omega \setminus A$ and we can check that

$$\sup_{a \in L^\infty(\Omega),\, \sup a < L} \int_{\Omega \setminus A} \left[a(x)\varrho(x) - f^*(a(x))\right] \mathrm{d}\lambda(x) = \int f(\varrho(x))\, \mathrm{d}\lambda(x),$$

$$\sup_{a \in L^\infty(\Omega),\, \sup a < L} \int_A a(x)\, \mathrm{d}\mu^s(x) = L\mu^s(A) = L\mu^s(\Omega),$$

which allows to conclude $\tilde{\mathscr{F}} = \mathscr{F}$. □

Remark 7.8. The assumption that λ is a finite measure is necessary to avoid integrability issues for the term $\int f^*(a(x))\,\mathrm{d}\lambda(x)$. A typical case where this term could give troubles is the entropy $f(t) = t\log t$, where $f^*(s) = e^{s-1}$. Suppose that λ is the Lebesgue measure on the whole $\mathbb{R}^d$; it is easy to see that, for any L^∞ function a, the integral $\int_{\mathbb{R}^d} e^{a(x)-1}\,\mathrm{d}x$ does not converge, and this provides $\tilde{\mathscr{F}}(\mu) = -\infty$ for every μ. However, if λ is σ-finite (which is the case of the Lebesgue measure on $\mathbb{R}^d$) and $f \geq 0$, it is possible to prove the semi-continuity by approximating λ from below with finite measures[2].

Remark 7.9. The reader could be shocked by the fact that in Proposition 7.7 we did not require lower bounds on f. In particular, the same result is true for $f(t) = t$ and $f(t) = -t$, which implies continuity of the mass with respect to the weak convergence of measures! Indeed, the weak convergence that we are using is the one in duality with C_b, and it is perfectly normal that the mass (i.e., the integral against the function 1) is continuous in this convergence. If we used, instead, the weak-$*$ convergence in duality with C_c on a non-compact space, the situation would be different. In order to perform the same proof, one would need at least $L > 0$ (otherwise, the functions $a \in C_c$ do not satisfy $a > L$), or $L = 0$ and $f^*(0) < +\infty$).

If we come back to the interpretation of $\mathscr{F}$, it is not difficult to check that $\mathscr{F}$ "favors" dispersed measure: first it is only finite for absolutely continuous measures; second, due to the Jensen inequality, the value of $\mathscr{F}$ is minimal for the constant density.

If we look for functionals having an opposite behavior and favoring the concentrated part of the measure, there are at least two different choices. We can look at an interaction functional such as $\mu \mapsto \int\int |x-y|^2 \mathrm{d}\mu(x)\mathrm{d}\mu(y)$ (where the square of the distance could be replaced by any increasing function of it). This is a global and spatially dependent functional and has a different flavor than $\mathscr{F}$. Indeed, we can find in the same class of local functionals some functionals which favor concentration, by looking in particular at the atomic part of μ. It is the case of the functional

$$\mathscr{G}(\mu) := \begin{cases} \sum_i g(a_i) & \text{if } \mu = \sum_i a_i \delta_{x_i},\ x_i \neq x_j \text{ for } i \neq j, \\ +\infty & \text{if } \mu \text{ is not purely atomic.} \end{cases}$$

As before, let us first understand which are the basic properties of g so as to guarantee semi-continuity.

We also assume $g(0) = 0$ (which is by the way necessary if we want to avoid ambiguities due to zero-mass atoms). Suppose that g satisfies $g(t) \leq Ct$ for all $t > 0$ and take, similar to what we did before, a sequence $\mu_n \rightharpoonup \mu$ where μ_n is purely atomic but μ is not (take, for instance, the sequence in Proposition 4.38). Then we

[2]For the case of the entropy, the lower semi-continuity on the whole $\mathbb{R}^d$ is false, but it is true under stronger convergence assumptions; see **Ex** (45). This is the usual strategy to study on the whole space the variational problems that we will present in bounded domains in Section 8.3 for the Fokker-Planck equation.

have $\mathscr{G}(\mu_n) \le C$ but $\mathscr{G}(\mu) = +\infty$. This is a contradiction with semi-continuity and suggests that we should consider functions g such that $\lim_{t\to 0} g(t)/t = +\infty$.

Second consideration, take $\mu_n = a\delta_{x_n} + b\delta_{y_n} + \nu$ (with ν finitely atomic: we only use ν so as to come back to a total mass of 1). Suppose $x_n \to x$ and $y_n \to x$ and $x \notin \mathrm{spt}(\nu)$. Then $\mu_n \rightharpoonup \mu = (a+b)\delta_x + \nu$. In this case semi-continuity would imply

$$g(a+b) \le g(a) + g(b),$$

i.e., we need g to be subadditive.

It is useful to establish the following:

Lemma 7.10. *Suppose that $g(0) = 0$ and that g is subadditive. Suppose that $\mu_n = \sum_{i=1}^N a_{i,n}\delta_{x_{i,n}}$ is a sequence of atomic probability measures with bounded number of atoms. Then, up to a subsequence, we have $\mu_n \rightharpoonup \mu$ where μ is also atomic with at most N atoms and $\mathscr{G}(\mu) \le \liminf \mathscr{G}(\mu_n)$.*

Proof. Let us precise that we impose to the points $x_{i,n}$ to be distinct. If a measure μ_n has less than N atoms, then we choose points $x_{i,n}$ at random to complete the list of atoms and set $a_{i,n} = 0$ for those extra indices.

Now extract a subsequence (not relabeled) such that for each $i = 1, \dots, N$, one has $a_{i,n} \to a_i$ and $x_{i,n} \to x_i$. For this subsequence one has $\mu_n \rightharpoonup \mu := \sum_{i=1}^N a_i\delta_{x_i}$. It is possible that the points x_i are not distinct. If they are distinct, we have $\mathscr{G}(\mu) = \sum_{i=1}^n g(a_i)$; otherwise we have (thanks to subadditivity) $\mathscr{G}(\mu) \le \sum_{i=1}^n g(a_i)$. Anyway we have

$$\mathscr{G}(\mu) \le \sum_{i=1}^n g(a_i) = \lim_n \sum_{i=1}^n g(a_{i,n}) = \lim_n \mathscr{G}(\mu_n),$$

which proves the desired result (one can choose the subsequence so that it realizes the liminf of the whole sequence). □

It is now possible to prove the following:

Lemma 7.11. *Suppose that $g(0) = 0$, that $g(t) \ge 0$, that g is subadditive and l.s.c., and that $\lim_{t\to 0} g(t)/t = +\infty$. Then $\mathscr{G}$ is l.s.c. on $\mathscr{P}(\Omega)$.*

Proof. Assume without loss of generality that all the μ_n are purely atomic. Fix a number $M > 0$ and use $\lim_{t\to 0} g(t)/t = +\infty$: this implies that there exists ε_0 such that for all $t < \varepsilon_0$ we have $g(t) > Mt$. Consider a sequence $\mu_n \rightharpoonup \mu$, assume $\mathscr{G}(\mu_n) \le C < +\infty$, and decompose it into $\mu_n = \mu_n^s + \mu_n^b$, where $\mu_n^b = \sum_{i=1}^N a_{i_n}\delta_{x_{i,n}}$ is the sum of the atoms of μ_n with mass at least ε_0. In particular, these atoms are no more than $N := \varepsilon_0^{-1}$. The other part μ_n^s (the "small" part, μ_n^b being the "big" one) is just defined as the remaining atoms (every μ_n is purely atomic since $\mathscr{G}(\mu_n) < +\infty$).

If we write

$$C \geq \mathcal{G}(\mu_n) \geq \mathcal{G}(\mu_n^s) = \sum_i g(a_{i,n}) \geq M \sum_i a_{i,n} = M\mu_n^s(\Omega),$$

we get an estimate on the mass of the "small" part. Hence it is possible to get, up to subsequences,

$$\mu_n^b \rightharpoonup \mu^b \quad \text{and} \quad \mu_n^s \rightharpoonup \mu^s,\ \mu^s(\Omega) \leq \frac{C}{M}.$$

We can now apply Lemma 7.10 to prove that μ^b is purely atomic and that $\mathcal{G}(\mu^b) \leq \liminf_n \mathcal{G}(\mu_n^b) \leq \liminf_n \mathcal{G}(\mu_n)$.

This proves that μ must be purely atomic, since the possible atomless part of μ must be contained in μ^s, but $\mu^s(\Omega) \leq C/M$. Hence the mass of the nonatomic part of μ must be smaller than C/M for every $M > 0$, i.e., it must be zero.

We have now proven that μ is purely atomic and we have an estimate of $\mathcal{G}(\mu^b)$, where μ^b is a part of μ depending on M. If we write $(a_i)_i$ for the masses of μ and $(a_i^M)_i$ for those of μ^b, we have

$$\sum_i g(a_i^M) \leq \liminf_n \mathcal{G}(\mu_n) := \ell.$$

We want to prove $\sum_i g(a_i) \leq \ell$ and, to this aim, it is enough to let $M \to \infty$. Actually, $\mu^s(\Omega) = \sum_i (a_i - a_i^M) \leq C/M \to 0$ implies that for each i we have $a_i - a_i^M \to 0$ and thus $a_i^M \to a_i$. Using the semi-continuity of g, we have $g(a_i) \leq \liminf_{M\to\infty} g(a_i^M)$. If we fix an arbitrary number N, we get

$$\sum_{i=1}^N g(a_i) \leq \liminf_{M\to\infty} \sum_{i=1}^N g(a_i^M) \leq \ell.$$

By passing to the supremum over N, we finally get

$$\sum_{i=1}^\infty g(a_i) \leq \ell,$$

which is the claim. □

As a particular example for the functionals of type $\mathcal{F}$, we can consider the L^p norms to the power p, i.e.,

$$\mathcal{F}(\mu) = ||\mu||_{L^p}^p = \begin{cases} \int \varrho(x)^p \, \mathrm{d}x & \text{if } \mu = \varrho \cdot \mathcal{L}^d, \\ +\infty & \text{otherwise.} \end{cases}$$

For the functionals of type $\mathscr{G}$, we can instead consider the cardinality of the support, obtained for $g(t) = 1$ if $t > 0$ and $g(0) = 0$:

$$\mathscr{G}(\mu) = \#(\mathrm{spt}(\mu)).$$

Box 7.1. Good to know! *Characterization of local l.s.c. functionals on $\mathscr{M}^d(\Omega)$*

It is also useful to note that functionals of the form $\mathscr{F}$ and $\mathscr{G}$ could be mixed, obtaining local functionals accepting both absolutely continuous and atomic parts. There is a general lower semi-continuity result that comes from the general theory developed in [67–69], which also characterizes the semi-continuity. It also covers the case of vector-valued measures and can be expressed as follows.

Theorem: Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be convex and l.s.c. and $g : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be subadditive and l.s.c. For every vector $\mathbf{v} \in \mathbb{R}^d$, define

$$f^\infty(\mathbf{v}) := \lim_{t\to+\infty} \frac{f(t\mathbf{v})}{t} \quad \text{and} \quad g^0(\mathbf{v}) := \lim_{t\to 0} \frac{g(t\mathbf{v})}{t},$$

and suppose $f^\infty = g^0$. Given a finite positive measure λ, decompose every vector measure $\mu \in \mathscr{M}^d(\Omega)$ as $\mu = \varrho\cdot\lambda + \mu^c + \mu^\#$, where $\mu^c + \mu^\#$ is the singular part of μ w.r.t. λ, which is also decomposed into a purely atomic part $\mu^\# = \sum_i a_i\delta_{x_i}$ and the remainder, usually refereed to as the Cantor part; $\varrho : \Omega \to \mathbb{R}^d$ is a vector density and the $a_i = \mu(\{x_i\}) \in \mathbb{R}^d$ are also vectors; we also denote by $\mathbf{w} : \Omega \to S^{d-1}$ the density of μ^c w.r.t. its own total variation measure $|\mu^c|$. Then the functional

$$\mu \mapsto \int f(\varrho(x))\ \mathrm{d}\lambda(x) + \int f^\infty(\mathbf{w}(x))\,\mathrm{d}|\mu^c|(x) + \sum_i g(a_i)$$

is local and lower semi-continuous for the weak convergence of measures.

Conversely, every local and lower semi-continuous functional on $\mathscr{M}^d(\Omega)$ can be written in the form above, for suitable choices of f, g, and λ.

7.2 Convexity, first variations, and subdifferentials

We pass in this section to another important notion concerning these functionals. If lower semi-continuity is crucial to establish existence results for minimization problems, convexity properties (in particular, strict convexity) are crucial for uniqueness. Also, with convexity comes the notion of subdifferential; thus we come to another very natural question in calculus of variations: how to compute first variations of these functionals?

We start from the very first question: Which among these functionals are convex?

Convexity and strict convexity

- $\mathscr{V}$ is linear and hence convex but not strictly convex.
- $\mathscr{W}$ is quadratic, but not always convex. Take, for instance, $W(x,y) = |x-y|^2$ and compute

$$\begin{aligned}&\int_\Omega\int_\Omega |x-y|^2 \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\\ &= \int_\Omega\int_\Omega |x|^2 \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) + \int_\Omega\int_\Omega |y|^2 \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y) - 2\int_\Omega\int_\Omega x\cdot y \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\\ &= 2\int_\Omega |x|^2 \,\mathrm{d}\mu(x) - 2\left(\int_\Omega x\,\mathrm{d}\mu(x)\right)^2.\end{aligned}$$

  This shows that $\mathscr{W}$ has (in this case) a linear part to which we subtract the square of another linear one; it is then concave rather than convex.
- All transport costs can be expressed by duality formula as a supremum of linear functionals

$$\mathscr{T}_c(\mu,\nu) = \sup_{\varphi(x)+\psi(y)\le c(x,y)} \int \varphi \,\mathrm{d}\mu + \int \psi \,\mathrm{d}\nu$$

  and it is hence convex (but W_p is in general not). Strict convexity is discussed later (it is true for $c(x,y) = |x-y|^p$, $p>1$ and $\nu \ll \mathscr{L}^d$).
- The norm in a dual functional space is always convex since it is a norm, but is never strictly convex because it is 1-homogeneous. Note that also the square of a norm could be non-strictly convex (as it is the case for the L^∞ or the L^1 norms).
- The functionals $\mathscr{F}$ that we considered above are actually convex due to the assumptions on f. Strict convexity is true if f is strictly convex and $L = +\infty$ (for instance, if one takes $f(t) = \sqrt{1+t^2} - 1$, which is strictly convex, then $\mathscr{F}$ is not strictly convex because it is finite and linear on singular measures).
- On the contrary, the functionals $\mathscr{G}$ that we considered above are typically not convex since the typical examples of subadditive functions are concave.

First variations

We now pass to the computation of first variations of these functionals. Since many of them are only considered in the convex set $\mathscr{P}(\Omega)$, which is a proper (but convex) subset of the Banach space $\mathscr{M}(\Omega)$, we prefer to give an ad hoc, and pragmatic, definition.

Definition 7.12. Given a functional $F : \mathscr{P}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, we say that $\varrho \in \mathscr{P}(\Omega)$ is *regular* for F if $F((1-\varepsilon)\varrho + \varepsilon\tilde{\varrho}) < +\infty$ for every $\varepsilon \in [0,1]$ and every $\tilde{\varrho} \in \mathscr{P}(\Omega) \cap L^\infty_c(\Omega)$ (absolutely continuous probabilities with L^∞ density and compact support), i.e., if the convex envelop of $\{\varrho\}$ and $L^\infty_c(\Omega) \cap \mathscr{P}(\Omega)$ is contained in $\{F < +\infty\}$.

If ϱ is regular for F, we call $\frac{\delta F}{\delta \varrho}(\varrho)$, if it exists, any measurable function such that

$$\frac{d}{d\varepsilon} F(\varrho + \varepsilon\chi)_{|\varepsilon=0} = \int \frac{\delta F}{\delta \varrho}(\varrho) \mathrm{d}\chi$$

for every perturbation $\chi = \tilde{\varrho} - \varrho$ with $\tilde{\varrho} \in L^\infty_c(\Omega) \cap \mathscr{P}(\Omega)$.

Remark 7.13. From the fact that we necessarily have $\int \mathrm{d}\chi = 0$, it is clear that $\frac{\delta F}{\delta \varrho}(\varrho)$ is defined up to additive constants. On the other hand, up to this invariance, we have uniqueness.

- From $\mathscr{V}(\mu + \varepsilon\chi) = \mathscr{V}(\mu) + \varepsilon \int V \,\mathrm{d}\chi$, we infer $\frac{\delta \mathscr{V}}{\delta \varrho}(\varrho) = V$.
- As $\mathscr{W}$ is quadratic, the computation is easy:

$$\mathscr{W}(\mu + \varepsilon\chi) - \mathscr{W}(\mu) = \varepsilon \iint W(x,y) \,\mathrm{d}\mu(x) \,\mathrm{d}\chi(y) + \varepsilon \iint W(x,y) \,\mathrm{d}\mu(y) \,\mathrm{d}\chi(x) + \varepsilon^2 \iint W(x,y) \,\mathrm{d}\chi(x) \,\mathrm{d}\chi(y).$$

  This provides

$$\frac{\delta \mathscr{W}}{\delta \varrho}(\varrho)(y) = \int W(x,y) \,\mathrm{d}\mu(x) + \int W(y,y') \,\mathrm{d}\mu(y').$$

  The formula becomes simpler when W is symmetric (since the two terms are equal) and even simpler when $W(x,y) = h(x-y)$ for an even function h, in which case it takes the form of a convolution $\frac{\delta \mathscr{W}}{\delta \varrho}(\varrho) = 2h * \mu$.
- For $\mu \mapsto \mathscr{T}_c(\mu, \nu)$, the first variation is given by the Kantorovich potential φ, in the transport from μ to ν, provided it is unique; this will be discussed later.
- Analogously, the first variation of $\mu \mapsto ||\mu||_{\mathscr{X}'}$ is given by the function φ realizing the maximum in $\max\{\int \varphi \,\mathrm{d}\mu : ||\varphi||_{\mathscr{X}} \leq 1\}$, provided it exists and is unique.
- If $f \in C^1$ and f' satisfies suitable bounds (typically we need f and f' to have polynomial growth), one can see that $\frac{\delta \mathscr{F}}{\delta \varrho}(\varrho) = f'(\varrho)$. Indeed, we can take $\chi = \theta \cdot \lambda$, write $\mathscr{F}(\varrho + \varepsilon\chi) = \int f(\varrho(x) + \varepsilon\theta(x)) \,\mathrm{d}\lambda(x)$, and differentiate under the integral sign (this requires some growth conditions): we get

$$\frac{d}{d\varepsilon}\mathscr{F}(\varrho+\varepsilon\chi)_{|\varepsilon=0}=\int f'(\varrho(x))\theta(x)\,\mathrm{d}\lambda=\int f'(\varrho)\,\mathrm{d}\chi.$$

The bounds on f' are only needed to allow derivation under the integral sign.

- Note that in general the functionals $\mathscr{G}$ are not differentiable along this kind of perturbations and that in general no measure is regular for $\mathscr{G}$. Indeed, from $g'(0)=+\infty$, we can infer that $G(\mu+\varepsilon\chi)$ is usually differentiable only if χ is concentrated on the same atoms as μ, even if for χ finitely atomic one would have $\mathscr{G}(\mu+\varepsilon\chi)<+\infty$.

Remark 7.14. We stress the fact that the special role given to L^∞_c functions in the definition of first variation is completely arbitrary. However, they represent a class of densities which is at the same time dense enough in the class of probabilities and regular enough to guarantee finiteness of most interesting functionals. In particular, we can observe that for all the functionals above except $\mathscr{G}$, a probability ϱ is regular if and only if the value of the functional is finite.

7.2.1 Dual norms

We start from the nontrivial issues above, i.e., the case of the transport costs and of the dual norms. Both are characterized by the their definition as sup of linear functionals. Since they are convex functionals, we start from a general fact in convex analysis, which allows us to provide not only the first variation but also the subdifferential.

Lemma 7.15. *Let $\mathscr{X}$ be a Banach space and $H:\mathscr{X}\to\mathbb{R}\cup\{+\infty\}$ be convex and l.s.c. Set $H^*(y)=\sup\langle x,y\rangle-H(x)$ for $y\in\mathscr{X}'$, where $\langle x,y\rangle$ denotes the duality between $\mathscr{X}$ and $\mathscr{X}'$. Then $\partial H^*(y_0)=\mathrm{argmax}_x\{\langle x,y_0\rangle-H(x)\}$.*

Proof. We know $x_0\in\partial H^*(y_0)\iff y_0\in\partial H(x_0)$ (see Box 1.12 in Section 1.6). This is equivalent to the fact that 0 belongs to the subdifferential of $x\mapsto H(x)-\langle x,y_0\rangle$ at x_0, but this is also equivalent to the fact that x_0 minimizes the same expression. □

Proposition 7.16. *Suppose that $\mathscr{X}$ is a reflexive separable Banach space of functions on Ω such that $\mathscr{X}\cap C_b(\Omega)$ is dense in $\mathscr{X}$. Let $F(\mu):=||\mu||_{\mathscr{X}'}=\sup\{\int\varphi\,\mathrm{d}\mu\ :\ ||\varphi||_{\mathscr{X}}\le 1\}$. Suppose that, for a given μ, the function $\varphi_\mu\in\mathscr{X}$ realizing the maximum in the definition of $F(\mu)$ exists and is unique. Then we have $\frac{\delta F}{\delta\varrho}(\mu)=\varphi_\mu$. Moreover, the subdifferential $\partial F(\mu)$ (where we consider F to be a functional on the Banach space $\mathscr{M}(\Omega)$, set to $+\infty$ outside $\mathscr{P}(\Omega)\cap\mathscr{X}'$) is always equal to the set (not necessarily a singleton) of maximizers.*

Proof. The second part of the statement is an easy consequence of Lemma 7.15, applied to the functional F defined as $F(\varphi)=0$ for $||\varphi||\le 1$ and $F(\varphi)=+\infty$ for $||\varphi||>1$.For the first part, take a perturbation χ such that $\mu+\varepsilon\chi$ belongs to $\mathscr{X}'$

for small ε. This implies that both μ and χ belong to $\mathscr{X}'$. From the definition of F, we can write

$$\int \varphi_\mu \, \mathrm{d}\chi \leq \frac{F(\mu + \varepsilon\chi) - F(\mu)}{\varepsilon} \leq \int \varphi_{\mu+\varepsilon\chi} \, \mathrm{d}\chi.$$

In order to pass to the limit, we just need to show that the right-hand side tends to $\int \varphi_\mu \, \mathrm{d}\chi$ as $\varepsilon \to 0$. First, note that the functions $\varphi_{\mu+\varepsilon\chi}$ belong to the unit ball of a reflexive separable Banach space. This means that, up to subsequences, we have $\varphi_{\mu+\varepsilon\chi} \rightharpoonup \varphi$, for a function φ with $||\varphi|| \leq 1$. We just need to prove $\varphi = \varphi_\mu$. From the uniqueness of the optimizer φ_μ, we just need to show that φ is also an optimizer for μ, i.e., $\int \varphi \, \mathrm{d}\mu = F(\mu)$.

To do this, note that $\mu + \varepsilon\chi$ converges strongly in $\mathscr{X}'$ to μ, and hence we have

$$\lim_{\varepsilon\to 0} F(\mu + \varepsilon\chi) = \lim_{\varepsilon\to 0} \int \varphi_{\mu+\varepsilon\chi} \mathrm{d}(\mu + \varepsilon\chi) = \int \varphi \, \mathrm{d}\mu \leq F(\mu) = \lim_{\varepsilon\to 0} F(\mu + \varepsilon\chi),$$

where the last equality comes from the continuity of F for strong convergence in $\mathscr{X}'$. Hence the inequality $\int \varphi \, \mathrm{d}\mu \leq F(\mu)$ is indeed an equality and $\varphi = \varphi_\mu$. □

7.2.2 Transport costs

We consider here the case of Wasserstein distances to a fixed measure (in particular functionals of the form $\mu \mapsto W_p^p(\mu, \nu)$) and, more generally, transport costs $\mathscr{T}_c(\mu, \nu) := \min\{\int c(x, y) \, \mathrm{d}\gamma \ : \ \gamma \in \Pi(\mu, \nu)\}$, i.e., minimal transport costs between measures and a fixed one. Note that for continuity issues it was useful to take advantage of the distance structure of W_p, but for convexity and first variation, this will not be the good approach.

For the sake of simplicity, the following result will only be given in the case of a compact domain $\Omega \subset \mathbb{R}^d$.

Proposition 7.17. *Let $\Omega \subset \mathbb{R}^d$ be compact and $c : \Omega \times \Omega \to \mathbb{R}$ be continuous. Then the functional $\mu \mapsto \mathscr{T}_c(\mu, \nu)$ is convex, and its subdifferential at μ_0 coincides with the set of Kantorovich potentials $\{\varphi \in C^0(\Omega) \ : \ \int \varphi \, \mathrm{d}\mu_0 + \int \varphi^c \, \mathrm{d}\nu = \mathscr{T}_c(\mu, \nu)\}$. Moreover, if there is a unique c-concave Kantorovich potential from μ_0 to ν up to additive constants, i.e., if the c-concave functions in the above set are all given by the same function φ_μ plus arbitrary additive constants, then we also have $\frac{\delta \mathscr{T}_c(\cdot,\nu)}{\delta \varrho}(\mu) = \varphi_\mu$.*

Proof. Let us come back to the expression

$$\mathscr{T}_c(\mu, \nu) = \sup\left\{ \int \varphi \, \mathrm{d}\mu + \int \varphi^c \, \mathrm{d}\nu \ : \ \varphi \in C^0(\Omega) \right\},$$

which allows to see $\mathscr{T}_c(\cdot,\nu)$ as a supremum of affine functionals of μ. Here ν is a fixed probability. Note that the very same supremum gives $+\infty$ if we take $\mu \in \mathscr{M}(\Omega)\setminus\mathscr{P}(\Omega)$, i.e.,

$$\sup\left\{\int \varphi\ \mathrm{d}\mu + \int \varphi^c\ \mathrm{d}\nu\ :\ \varphi\in C^0(\Omega)\right\} = \begin{cases}\mathscr{T}_c(\mu,\nu) & \text{if } \mu\in\mathscr{P}(\Omega),\\ +\infty & \text{if } \mu\in\mathscr{M}(\Omega)\setminus\mathscr{P}(\Omega).\end{cases}$$

Indeed, if μ is not a positive measure, then there exists a function $\varphi\in C^0(\Omega)$, with $\varphi\le 0$ and $\int\varphi\,\mathrm{d}\mu>0$. Since $\varphi^c(y)=\inf_y c(x,y)-\varphi(x)\ge \inf c$ (suppose c bounded from below), then $\int\varphi\,\mathrm{d}\mu+\int\varphi^c\,\mathrm{d}\nu\ge\int\varphi\,\mathrm{d}\mu+\inf c$. By replacing φ with $\lambda\varphi$ and letting $\lambda\to+\infty$, we see that the supremum is $+\infty$.

On the other hand, if $\mu\ge 0$ but $\mu(\Omega)\neq 1=\nu(\Omega)$, we also find $+\infty$ in the sup. Indeed, we can always add a constant λ to any φ, and $(\varphi+\lambda)^c=\varphi^c-\lambda$. Taking, for instance, $\varphi=0$, we get

$$\sup\left\{\int \varphi\ \mathrm{d}\mu + \int \varphi^c\ \mathrm{d}\nu\ :\ \varphi\in C^0(\Omega)\right\} \ge \int 0^c(y)\ \mathrm{d}\nu(y) + \lambda\big(\mu(\Omega)-\nu(\Omega)\big),$$

which can be made as large as we want if $\mu(\Omega)-\nu(\Omega)\neq 0$ (here 0^c is a bounded function whose precise expression depends on c).

In order to identify the subdifferential, we apply Lemma 7.15, with $\mathscr{X}=C^0(\Omega)$ (endowed with the sup norm) and $H:\mathscr{X}\to\mathbb{R}$ given by $H(\varphi):=-\int\varphi^c\,\mathrm{d}\nu$. We just need to see that this functional is convex and semi-continuous (indeed, it is continuous). Note that if we take $\varphi_0,\varphi_1\in C^0(\Omega)$, we have

$$\varphi_1^c(y)=\inf_x c(x,y)-\varphi_1(x)\le \inf_x c(x,y)-\varphi_0(x)+||\varphi_1-\varphi_0||_\infty=\varphi_0^c(y)+||\varphi_1-\varphi_0||_\infty.$$

By interchanging the roles of φ_0 and φ_1 and using the arbitrariness of y, we get $||\varphi_1^c-\varphi_0^c||_\infty\le||\varphi_1-\varphi_0||_\infty$, which implies $\left|\int\varphi_1^c\,\mathrm{d}\nu-\int\varphi_0^c\,\mathrm{d}\nu\right|\le||\varphi_1-\varphi_0||_\infty$ and hence the continuity of the functional. As far as convexity is concerned, set $\varphi_t=(1-t)\varphi_0+t\varphi_1$; we have

$$\varphi_t^c(y)=\inf_x c(x,y)-(1-t)\varphi_0(x)-t\varphi_1(x)\ge(1-t)\varphi_0^c(y)+t\varphi_1^c(y).$$

This implies the concavity $\int\varphi_t^c\,\mathrm{d}\nu\ge(1-t)\int\varphi_0^c\,\mathrm{d}\nu+t\int\varphi_1^c\,\mathrm{d}\nu$ and hence H is convex.

In order to get the first variation in the situation described in the statement (uniqueness of the c-concave Kantorovich potential), we take $\mu_\varepsilon=\mu+\varepsilon\chi$ with $\chi=\tilde\mu-\mu$ and we estimate the ratio $(\mathscr{T}_c(\mu_\varepsilon,\nu)-\mathscr{T}_c(\mu,\nu))/\varepsilon$. First, by using that $(\varphi_\mu,\varphi_\mu^c)$ is optimal in the dual formulation for μ, but not necessarily for μ_ε, we have

$$\frac{\mathscr{T}_c(\mu_\varepsilon,\nu)-\mathscr{T}_c(\mu,\nu)}{\varepsilon}\ge\frac{\int\varphi_\mu\,\mathrm{d}\mu_\varepsilon+\int\varphi_\mu^c\,\mathrm{d}\nu-\int\varphi_\mu\,\mathrm{d}\mu-\int\varphi_\mu^c\,\mathrm{d}\nu}{\varepsilon}=\int\varphi_\mu\,\mathrm{d}\chi,$$

which gives the lower bound $\liminf_{\varepsilon\to0}(\mathscr{T}_c(\mu_\varepsilon,\nu)-\mathscr{T}_c(\mu,\nu))/\varepsilon\ge\int\varphi_\mu\,\mathrm{d}\chi$.

To look at the lim sup, first fix a sequence of values of ε_k realizing the limsup, i.e., $\lim_k(\mathscr{T}_c(\mu_{\varepsilon_k},\nu)-\mathscr{T}_c(\mu,\nu))/\varepsilon_k = \limsup_{\varepsilon\to 0}(\mathscr{T}_c(\mu_\varepsilon,\nu)-\mathscr{T}_c(\mu,\nu))/\varepsilon$. Then we can estimate the same ratio using the optimality of a pair (φ_k,φ_k^c), Kantorovich potentials in the transport from μ_{ε_k} to ν (we can assume them to be c-concave functions and such that $\varphi_k(x_0)=0$ for a fixed point $x_0\in\Omega$). Now we may write

$$\frac{\mathscr{T}_c(\mu_{\varepsilon_k},\nu)-\mathscr{T}_c(\mu,\nu)}{\varepsilon_k} \le \frac{\int\varphi_k\,\mathrm{d}\mu_{\varepsilon_k}+\int\varphi_k^c\,\mathrm{d}\nu-\int\varphi_k\,\mathrm{d}\mu-\int\varphi_k^c\,\mathrm{d}\nu}{\varepsilon_k} = \int\varphi_k\,\mathrm{d}\chi. \tag{7.1}$$

As in Proposition 7.16, we need to pass to the limit in k. Theorem 1.52 shows that we have uniform convergence (up to extracting another subsequence) $(\varphi_k,\varphi_k^c)\to(\varphi,\varphi^c)$ and that (φ,φ^c) must be optimal in the duality formula for the transport between μ and ν. This implies $\varphi=\varphi_\mu$ by uniqueness. Finally, passing to the limit in (7.1), we get also $\limsup_{\varepsilon\to 0}(\mathscr{T}_c(\mu_\varepsilon,\nu)-\mathscr{T}_c(\mu,\nu))/\varepsilon\le\int\varphi_\mu\mathrm{d}\chi$. □

As Proposition 7.17 requires uniqueness of the Kantorovich potential in order to efficiently compute the first variation, we also give here a sufficient condition

Proposition 7.18. *If Ω is the closure of a bounded connected open set, c is C^1, and at least one of the measures μ or ν is supported on the whole Ω, then the c-concave Kantorovich potential in the transport from μ to ν is unique up to additive constants.*

Proof. Suppose that $\mathrm{spt}(\mu)=\Omega$. First note that $c\in C^1$ implies that c is Lipschitz on $\Omega\times\Omega$, and hence all Kantorovich potentials, which are c-concave, are Lipschitz as well and are differentiable a.e. Consider two different Kantorovich potentials φ_0 and φ_1. We use Proposition 1.15, which guarantees that their gradients must agree a.e. on Ω. Since Ω is the closure of a connected open set, this means that the difference $\varphi_0-\varphi_1$ is constant and provides the desired result. If the measure with full support is ν, just apply the same procedure to the transport from ν to μ and get the uniqueness of the potential ψ. Then, from $\varphi=\psi^c$ (here we use c-concavity), one also recovers the uniqueness of φ. □

Note that in order to apply the above result to Proposition 7.17, the uniqueness μ-a.e. would not be enough (since one integrates it also against other measures μ_ε), and, anyway, without connectedness assumptions on $\mathrm{spt}(\mu)$, it would be impossible to deduce the uniqueness of φ.

We finish this section with a remark on strict convexity. We come back for simplicity to the case $c(x,y)=|x-y|^p$, i.e., to the functional $W_p^p(\cdot,\nu)$, but the reader will see that everything works the same under the twist condition (or if $c(x,y)=h(x-y)$ with h strictly convex). Also, the assumption $\nu\ll\mathscr{L}^d$ is not sharp, at least for $p=2$ (see Section 1.3.1).

Proposition 7.19. *If $\nu\ll\mathscr{L}^d$ and $p>1$, the functional $W_p^p(\cdot,\nu)$ is strictly convex.*

Proof. Suppose by contradiction that $\mu_0\neq\mu_1$ and $t\in]0,1[$ are such that $W_p^p(\mu_t,\nu)=(1-t)W_p^p(\mu_0,\nu)+tW_p^p(\mu_1,\nu)$, where $\mu_t=(1-t)\mu_0+t\mu_1$. Let γ_0 be the optimal transport plan in the transport from ν to μ_0 (pay attention to the

direction: it is a transport map if we see it backward: *from ν to μ_0*, since $\nu \ll \mathscr{L}^d$); we write $\gamma_0 = (\mathrm{T}_0, \mathrm{id})_\# \nu$. Analogously, take $\gamma_1 = (\mathrm{T}_1, \mathrm{id})_\# \nu$ optimal from ν to μ_1. Set $\gamma_t := (1-t)\gamma_0 + t\gamma_1 \in \Pi(\mu_t, \nu)$. We have

$$(1-t)W_p^p(\mu_0,\nu) + tW_p^p(\mu_1,\nu) = W_p^p(\mu_t,\nu) \le \int |x-y|^p \, \mathrm{d}\gamma_t$$
$$= (1-t)W_p^p(\mu_0,\nu) + tW_p^p(\mu_1,\nu),$$

which implies that γ_t is actually optimal in the transport from ν to μ_t. Yet γ_t is not induced from a transport map, unless $\mathrm{T}_0 = \mathrm{T}_1$. This is a contradiction with $\mu_0 \neq \mu_1$ and proves strict convexity. □

7.2.3 *Optimality conditions*

We finish this part of the chapter devoted to first variations with an important computation, which leads to optimality conditions in the space $\mathscr{P}(\Omega)$. We do not claim this result to be sharp, but it is enough in most situations, and the reader can easily see how to adapt it to other cases.

Proposition 7.20. *Suppose that ϱ_0 minimizes a functional $F : \mathscr{P}(\Omega) \to \mathbb{R} \cup \{+\infty\}$, that ϱ_0 is regular for F (see Definition 7.12), and that there exists $\frac{\delta F}{\delta \varrho}(\varrho_0)$. Call this function g and set $\ell = \operatorname{ess\,inf} g$, the essential infimum of g (essential w.r.t. the Lebesgue measure).*

First case*: Suppose that $g \in C(\Omega)$. Then* $\mathrm{spt}(\varrho_0) \subset \operatorname{argmin} g = \{g = \ell\}$*, and we have everywhere the inequality $g \ge \ell$, with equality $g = \ell$ on* $\mathrm{spt}(\varrho_0)$.

Second case*: Suppose that g is only measurable, but $\varrho_0 \ll \mathscr{L}^d$. Then we have a.e. the inequality $g \ge \ell$, with $g = \ell$ a.e. on $\{\varrho_0 > 0\}$.*

Proof. Consider a competitor $\tilde{\varrho} \in L^\infty_c(\Omega)$. Take $\varrho_\varepsilon := (1-\varepsilon)\varrho + \varepsilon\tilde{\varrho} = \varrho + \varepsilon\chi$, with $\chi = \tilde{\varrho} - \varrho$. The measure ϱ_ε is an admissible competitor in the minimization problem as soon as $\varepsilon \in [0,1]$. We deduce that

$$\frac{d}{d\varepsilon}(F(\varrho_\varepsilon))_{|\varepsilon=0} \ge 0.$$

This means

$$\int g \, \mathrm{d}\chi \ge 0, \quad \text{i.e.,} \int g \, \mathrm{d}\tilde{\varrho} \ge \int g \, \mathrm{d}\varrho_0.$$

We now use the arbitrariness of $\tilde{\varrho}$. Take any $\ell' > \ell$. By definition of essential infimum, the set $\{g < \ell'\}$ has positive Lebesgue measure, and we can choose an L^∞ density $\tilde{\varrho}$ concentrated on it. We get $\ell' > \int g \, \mathrm{d}\varrho_0$. Letting $\ell' \to \ell$, we get $\ell \ge \int g \, \mathrm{d}\varrho_0$.

In the first case, the essential infimum coincides with the infimum, and hence we have $g \geq \ell$ everywhere. But integrating this inequality against ϱ_0, we have the opposite inequality, which implies $g = \ell$ ϱ_0-a.e. This equality extends to the support of ϱ_0 since g is continuous.

In the second case, we have $g \geq \ell$ a.e. Since ϱ_0 is supposed to be absolutely continuous, the same is true ϱ_0-a.e., and hence we have $g = \ell$ ϱ_0-a.e., which means a.e. on $\{\varrho_0 > 0\}$. □

We will see, for instance, in Sections 7.4.1 and 8.3 some applications of this criterion.

We also stress that, when F is convex, alternative arguments using subdifferentials could be used, but the above strategy is the only possible one when we do not have convexity. We can see how it works in some examples.

Example 7.21. Consider

$$\min\left\{F(\varrho) := W_2^2(\varrho,\nu) + \int V(x)\,\mathrm{d}\varrho(x) \;:\; \varrho \in \mathscr{P}(\Omega)\right\},$$

where $\Omega \subset \mathbb{R}^d$ is compact and V continuous. In such a case it is clear that a minimizer ϱ_0 exists and that it is regular for F since F is finite on the whole space $\mathscr{P}(\Omega)$. If the Kantorovich potential φ (for the cost $c(x,y) = |x-y|^2$) in the transport from ϱ_0 to ν is unique up to additive constants, then the first variation of F at ϱ_0 is given by $\varphi + V$, and we can apply Proposition 7.20 to obtain the existence of a constant C such that

$$\varphi + V \geq C \;\text{ on } \Omega, \quad \varphi + V = C \;\text{ on } \mathrm{spt}(\varrho_0).$$

This is the case if $\mathrm{spt}(\nu) = \Omega$, for instance.

Another way to obtain the optimality conditions is the following: the second term of F is a continuous and linear functional on $\mathscr{M}(\Omega)$, and the first is a convex functional defined on $\mathscr{M}(\Omega)$, with subdifferential given by the set of Kantorovich potentials. Hence, we can prove

$$\partial F(\varrho_0) = \{V + \varphi \;:\; \varphi \text{ Kantorovich potential from } \varrho_0 \text{ to } \nu\}.$$

From standard optimality conditions in convex analysis, we obtain the existence of a potential φ such that $V + \varphi = 0$. In this case the result that we get is much stronger with this convex procedure, since we get rid of the inequality case. Note, however, that the function φ that we get is not necessarily c-concave: since we know $\varphi = \varphi^{cc}$ on $\mathrm{spt}(\varrho_0)$, this means that it has been modified out of $\mathrm{spt}(\varrho_0)$ in order to guarantee the equality.

Example 7.22. Consider

$$\min\left\{F(\varrho) := W_2^2(\varrho,\nu) + \int_\Omega \frac{1}{p}\varrho^p(x)\,\mathrm{d}x \;:\; \varrho \in \mathscr{P}(\Omega), \varrho \ll \mathscr{L}^d\right\},$$

where $\Omega \subset \mathbb{R}^d$ is compact and $p > 1$. Again, a minimizer $\varrho_0 \in L^p(\Omega)$ exists and is regular (in the sense of Definition 7.12), since F is finite on $L^p(\Omega) \cap \mathscr{P}(\Omega)$ which includes for sure all convex combinations of ϱ_0 and L^∞ functions. Computing the first variation and using Proposition 7.20, we obtain the existence of a constant C such that

$$\varphi + \varrho^{p-1} \geq C \text{ on } \Omega, \quad \varphi + \varrho^{p-1} = C \text{ on } \mathrm{spt}(\varrho_0).$$

This, by the way, can be rewritten as $\varrho^{p-1} = (C - \varphi)_+$.

On the other hand, even if the functional is still convex, it is not easy to compute the subdifferential of F, viewed as a functional on $\mathscr{M}(\Omega)$ (set to $+\infty$ outside L^p probabilities). Indeed, the subdifferential of a sum is not always the sum of the subdifferentials (see, for instance, [157] for convex analysis issues in infinite dimension; however, this requires continuity of at least one of the two functionals, while here both are finite on a small subset of $\mathscr{M}(\Omega)$). Moreover, the first variation of the second term, i.e., ϱ^{p-1}, does not belong in general to the space $C(\Omega)$, i.e., to the space $\mathscr{M}(\Omega)$ is in duality with. A way of circumventing this difficulty would be to view F as a functional on $L^p(\Omega)$. Here, the second term becomes a continuous convex functional. The first one must be rewritten as $\sup\{\int \phi \, \mathrm{d}\varrho + \int \phi^c \, \mathrm{d}\nu \ : \ \phi \in L^{p'}(\Omega)\}$ (with a proper definition of the c-transform in terms of essential infima instead of infima). This would provide the optimality condition: there exists an $L^{p'}$ Kantorovich potential φ such that $\varphi + \varrho^{p-1} = 0$. Again, $\varphi = \varphi^{cc}$ ϱ_0-a.e., and $\varphi \leq \varphi^{cc}$ elsewhere. This gives again

$$\varphi^{cc} + \varrho^{p-1} \geq C \text{ on } \Omega, \quad \varphi^{cc} + \varrho^{p-1} = C \text{ on } \mathrm{spt}(\varrho_0).$$

7.3 Displacement convexity

In all the considerations of the previous section about convexity, we always considered the standard convex interpolations $[0, 1] \ni t \mapsto (1 - t)\mu_0 + t\mu_1$. Yet, another notion of convexity, more linked to the metric structure of the space $\mathbb{W}_p(\Omega)$, may be useful.

Let us start from this general definition.

Definition 7.23. In a geodesic metric space X, we define $F : X \to \mathbb{R} \cup \{+\infty\}$ to be *geodesically convex* if for every two points $x_0, x_1 \in X$ there exists a constant-speed geodesic ω connecting $\omega(0) = x_0$ to $\omega(1) = x_1$ such that $[0, 1] \ni t \mapsto F(\omega(t))$ is convex.

This obviously reduces to usual convexity in $\mathbb{R}^d$ or in any other normed vector space, where segments are the unique geodesics (note that in any normed vector space segments are geodesics but they may not be the unique ones, as it is the case in L^1 or L^∞ or any non-strictly convex space). By the way, in spaces where there is not uniqueness of the geodesics, the definition we gave could be debated, since one

could also choose to define as geodesically convex those functionals that satisfy the geodesic inequality on *every* geodesic ω. But the one we gave is the definition that is usually chosen, as it satisfies some extra stability results that would not be true with the more restrictive one requiring convexity for every geodesic.

The notion of geodesic convexity in the space $\mathbb{W}_p(\Omega)$ has been introduced by McCann in [229] and is particularly interesting since we know how to characterize geodesics in such a space. This notion of convexity is usually referred to as *displacement convexity*. Note that it could a priori depend on the exponent p.

7.3.1 Displacement convexity of $\mathscr{V}$ and $\mathscr{W}$

We recall that the geodesics for the W_p distance are of the form $\mu_t = (\pi_t)_\# \gamma$ where γ is an optimal transport for $c(x,y) = |x-y|^p$ and $\pi_t(x,y) = (1-t)x + ty$.

Proposition 7.24. *The functional $\mathscr{V}$ is displacement convex if and only if V is convex. The functional $\mathscr{W}$ is displacement convex if W is convex.*

Proof. First consider $\mathscr{V}$ and suppose that V is convex. Let us evaluate $\mathscr{V}(\mu_t)$:

$$\mathscr{V}(\mu_t) = \int V \,\mathrm{d}(\pi_t)_\#\gamma = \int V((1-t)x + ty)\,\mathrm{d}\gamma.$$

It is clear from this formula that $t \mapsto \mathscr{V}(\mu_t)$ is convex if V is convex.

On the other hand, the convexity of V is a necessary condition for $\mathscr{V}$ being convex as one can easily check by considering geodesics of the form $\mu_t = \delta_{(1-t)x+ty}$ since $\mathscr{V}(\mu_t) = V((1-t)x + ty)$.

The proof for the convexity of $\mathscr{W}$ is similar: consider

$$\begin{aligned}\mathscr{W}(\mu_t) &= \int W(x,x')\,\mathrm{d}(\pi_t)_\#\gamma(x)\mathrm{d}(\pi_t)_\#\gamma(x') \\ &= \int W((1-t)x+ty, (1-t)x'+ty')\ \mathrm{d}\gamma(x,y)\ \mathrm{d}\gamma(x',y'),\end{aligned}$$

which easily gives the condition for the convexity. □

Note that we did not state that convexity of W is a necessary condition for $\mathscr{W}$, since it is not true in general. Consider, for instance, the following 1D case.

Proposition 7.25. *Let $\Omega = (a,b)$, $\Delta^+ = \{(x,y) \in \Omega \times \Omega,\ y \geq x\}$, and $\Delta^- = \{(x,y) \in \Omega \times \Omega,\ y \leq x\}$. Then it is sufficient that W is convex when restricted to Δ^+ and Δ^-, in order to have displacement convexity of $\mathscr{W}$.*

Proof. Just consider the proof of Proposition 7.24 and check that all the segments $t \mapsto ((1-t)x + ty, (1-t)x' + ty')$ for $(x,y), (x',y') \in \mathrm{spt}(\gamma)$ are contained either in Δ^+ or in Δ^-. This is true thanks to the monotonicity properties of $\mathrm{spt}(\gamma)$ that we

showed in Section 2.2. Actually, we showed that, for every strictly convex cost, the (unique) optimal transport plan γ satisfies

$$(x, y), (x', y') \in \mathrm{spt}(\gamma), \quad x < x' \quad \Rightarrow \quad y \leq y'.$$

This means that if $x < x'$, then we have $(x, x') \in \Delta^+$ and $(y, y') \in \Delta^+$ and, by convexity, $((1-t)x + ty, (1-t)x' + ty') \in \Delta^+$. Analogously, for $x > x'$, we get that the segment is contained in Δ^-. If $x = x'$, it is enough to look at the transport plan from the point of view of y, thus getting the implication $y < y' \quad \Rightarrow \quad x \leq x'$ and concluding in the same way. The only case that stays apart is $x = x'$ and $y = y'$, but in this case the segment reduces to a point.

This proves convexity of $\mathscr{W}$ along the geodesics in $\mathbb{W}_p$ for $p > 1$. It also works for W_1 if we choose the same γ as an optimal transport plan (in this case, it will not be the unique one, but this will be enough to build a geodesic where there is convexity). □

An interesting consequence of this criterion is the fact that the squared dual norm $\mu \mapsto ||\mu||^2_{\mathscr{X}'}$ for $\mathscr{X} = H^1([-1, 1])$ is actually displacement convex, as a corollary of the characterization of Proposition 7.6 (but it does not work for $\mathscr{X} = H^1_0$). This has been pointed out and used in [62] and [110].

This is an example of displacement convex functional which involves, somehow, derivatives. We could say that, as an H^{-1} norm, it is a functional of order -1. Displacement convex functionals of order different than 0 are not so common to find. For the first-order case, the only known example is contained in [115] and it is, again, only available in 1D. Nothing is known in the multidimensional case.

7.3.2 Displacement convexity of $\mathscr{F}$

The most interesting displacement convexity result is the one for functionals depending on the density.

To consider these functionals, we need some technical facts.

The starting point is the computation of the density of an image measure, via standard change-of-variable techniques, as we saw in the Memo Box 1.14.

Then, we underline an interesting computation.

Lemma 7.26. *Let A be a $d \times d$ matrix such that its eigenvalues λ_i are all real and larger than -1 (for instance, this is the case when A is symmetric and $\mathrm{I} + A \geq 0$). Then $[0, 1] \ni t \mapsto g(t) := \det(\mathrm{I} + tA)^{1/d}$ is concave.*

Proof. We can write A in a suitable basis so that it is triangular, and we get

$$g(t)^d = \prod_{i=1}^{d} (1 + t\lambda_i).$$

Differentiating,

$$dg(t)^{d-1}g'(t) = \sum_{j=1}^{d} \lambda_j \prod_{i=1,\, i\neq j}^{d} (1+t\lambda_i) = g(t)^d \sum_{j=1}^{d} \frac{\lambda_j}{1+t\lambda_j},$$

i.e., $dg'(t) = g(t)\sum_{j=1}^{d} \frac{\lambda_j}{1+t\lambda_j}$. If we differentiate once more, we get

$$dg''(t) = g'(t)\sum_{j=1}^{d} \frac{\lambda_j}{1+t\lambda_j} - g(t)\sum_{j=1}^{d} \frac{\lambda_j^2}{(1+t\lambda_j)^2}.$$

Here we use the quadratic-arithmetic mean inequality which gives

$$\sum_{j=1}^{d} \frac{\lambda_j^2}{(1+t\lambda_j)^2} \geq \frac{1}{d}\left(\sum_{j=1}^{d} \frac{\lambda_j}{1+t\lambda_j}\right)^2 = d\left(\frac{g'(t)}{g(t)}\right)^2$$

and hence

$$dg''(t) \leq d\,\frac{g'(t)^2}{g(t)} - d\,\frac{g'(t)^2}{g(t)} = 0,$$

which proves the concavity of g. □

Remark 7.27. We observe that in dimension $d = 1$, the above lemma is straightforward and that g is actually affine.

We can now state the main theorem, which is the main result of [229].

Theorem 7.28. *Suppose that f is convex and superlinear, $f(0) = 0$, and that $s \mapsto s^{-d}f(s^d)$ is convex and decreasing. Suppose that Ω is convex and take $1 < p < \infty$. Then $\mathscr{F}$ is geodesically convex in $\mathbb{W}_p$.*

Proof. Let us consider two measures μ_0, μ_1 with $\mathscr{F}(\mu_0), \mathscr{F}(\mu_1) < +\infty$. They are absolutely continuous and hence there is a unique constant-speed geodesic μ_t between them (see Theorem 5.27 and Proposition 5.32), which has the form $\mu_t = (\mathrm{T}_t)_\# \mu_0$, where $\mathrm{T}_t = \mathrm{id} + t(\mathrm{T} - \mathrm{id})$. Note that T_t is injective because of Lemma 4.23.

We first look at the case $p = 2$, which is easier. In this case we have $\mathrm{T}_t(x) = x - t\nabla\varphi(x)$, where φ is such that $\frac{x^2}{2} - \varphi$ is convex. This implies, by Theorem 3.16, that $\nabla\varphi$ is countably Lipschitz, and so is T_t. Hence they are approximately differentiable a.e.[3] The Hessian $D^2\varphi$ (or, equivalently, the approximate gradient of $\nabla\phi$) is symmetric and $D^2\varphi \leq I$. Let us define $A(x) = -D^2\varphi(x)$. From the formula for

[3] Actually, it is also known that $D^2\varphi$ exists a.e.: indeed, convex functions are twice differentiable a.e. (see, for instance, [160], the original result being stated in [6]).

the density of the image measure, we know that μ_t is absolutely continuous and we can write its density ϱ_t as $\varrho_t(y) = \varrho(\mathrm{T}_t^{-1}(y))/\det(\mathrm{I} + tA(\mathrm{T}_t^{-1}(y)))$ and hence

$$\mathscr{F}(\mu_t) = \int f\left(\frac{\varrho(\mathrm{T}_t^{-1}(y)))}{\det(\mathrm{I} + tA(\mathrm{T}_t^{-1}(y)))}\right)\mathrm{d}y = \int f\left(\frac{\varrho(x)}{\det(\mathrm{I} + tA(x))}\right)\det(\mathrm{I}+tA(x))\,\mathrm{d}x,$$

where we used the change of variable $x = T_t^{-1}(y)$, which gives $y = T_t(x)$ and $\mathrm{d}y = \det D\mathrm{T}_t(x)\,\mathrm{d}x = \det(\mathrm{I} + tA(x))\,\mathrm{d}x$.

From Lemma 7.26 we know that $\det(\mathrm{I} + tA(x)) = g(t,x)^d$ for a function $g : [0,1] \times \Omega$ which is concave in t. It is a general fact that the composition of a convex and decreasing function with a concave one gives a convex function. This implies that

$$t \mapsto f\left(\frac{\varrho(x)}{g(t,x)^d}\right) g(t,x)^d$$

is convex (if $\varrho(x) \neq 0$, this uses the assumption on f and the fact that $t \mapsto g(t,x)/\varrho(x)^{\frac{1}{d}}$ is concave; if $\varrho(x) = 0$, then this function is simply zero). Finally, we proved convexity of $t \mapsto \mathscr{F}(\mu_t)$.

We now have to adapt to the case $p \neq 2$. In this case, setting $h(z) = \frac{1}{p}|z|^p$, we have $\mathrm{T}_t(x) = x - t\nabla h^*(\nabla\varphi(x))$. Note that both h and h^* are $C^2(\mathbb{R}^d \setminus \{0\})$. First, we note that we can decompose, up to negligible sets, Ω into two measurable parts: the set Ω' where $\nabla\varphi = 0$ and the set Ω'' where $\nabla\varphi \neq 0$. The analysis that we do is very much similar to what we did in Section 3.3.2. T is the identity on Ω', and also T_t is the identity, and so is T_t. Since T_t is injective, the density is preserved on this set, and we can apply Lemma 7.26 with $A = 0$. The set Ω'', on the contrary, can be decomposed into a countable union of sets Ω_{ij} where $x \in B_i$, $\mathrm{T}(x) \in B_j$, $(B_i)_i$ is a countable family of balls generating the topology of $\mathbb{R}^d$, and we only take pairs such that $B_i \cap B_j = \emptyset$. On this set $\varphi = \psi^c$ also coincides with φ_{ij} defined as the restriction to B_i of the function $x \mapsto \inf_{y\in B_j} h(x-y) - \psi(y)$, which is λ-concave for $\lambda = \sup\{||D^2h(z)|| : z \in B_j - B_i\}$. This proves that φ_{ij} has the same regularity of concave functions and that $\nabla\varphi$ and $\nabla h^*(\nabla\varphi)$ are countably Lipschitz on Ω''. In particular, $D^2\varphi$ exists a.e.

If we fix a point $x_0 \in \Omega''$ where $\nabla\varphi(x_0)$ and $D^2\varphi(x_0)$ exist, then we can write $\varphi(x) \leq h(x-T(x_0)) - \psi(T(x_0))$, an inequality which is true for every x, with equality at $x = x_0$. In particular we get $D^2\varphi(x_0) \leq D^2h(x_0 - T(x_0)) = D^2h(\nabla h^*(\nabla\varphi(x_0)))$. From general properties of Legendre transforms, we have

$$D^2h(\nabla h^*(z)) = [D^2h^*(z)]^{-1}$$

(just differentiate the relation $\nabla h(\nabla h^*(z)) = z$). Hence, we can apply Lemma 7.26 with $A(x) = -D^2h^*(\nabla\varphi(x))D^2\varphi(x)$, which is diagonalizable and has eigenvalues larger than -1 (see Box 7.2 below). □

Box 7.2. Memo: *Diagonalizing products of symmetric matrices*

It is well known that symmetric matrices can be diagonalized on $\mathbb{R}$. A trickier result is the following:

Theorem. If A, B are symmetric $d \times d$ matrices and A is positive definite, then AB is diagonalizable.

We do not prove it, but we just prove the following weaker result.

Proposition. If A, B are symmetric $d \times d$ matrices and A is positive definite, then AB has real eigenvalues.

Proof. Suppose that $\lambda \in \mathbb{C}$ is an eigenvalue of AB, i.e., $AB\mathbf{v} = \lambda\mathbf{v}$ for a vector $\mathbf{v} \in \mathbb{C}^n \setminus \{0\}$. Take the Hermitian product with $A^{-1}\mathbf{v}$, i.e., $\bar{\mathbf{v}}^t B\mathbf{v} = \lambda\bar{\mathbf{v}}^t A^{-1}\mathbf{v}$, where the symbol ^{t} denotes transposition in the sense of matrices. Note that, by diagonalizing independently B and A^{-1}, both the terms $\bar{\mathbf{v}}^t B\mathbf{v}$ and $\bar{\mathbf{v}}^t A^{-1}\mathbf{v}$ are real, and $\bar{\mathbf{v}}^t A^{-1}\mathbf{v} > 0$. Hence, $\lambda \in \mathbb{R}$.

Another important point that we need is the following.

Proposition. If $B \leq A^{-1}$ (in the sense of symmetric matrices, i.e., $\mathbf{v}^t(A^{-1} - B)\mathbf{v} \geq 0$ for all $\mathbf{v}$), then all the eigenvalues λ_i of AB satisfy $\lambda_i \leq 1$.

Proof. Again, write $AB\mathbf{v} = \lambda\mathbf{v}$ and deduce $\bar{\mathbf{v}}^t B\mathbf{v} = \lambda\bar{\mathbf{v}}^t A^{-1}\mathbf{v}$. This implies $\bar{\mathbf{v}}^t A^{-1}\mathbf{v} \geq \bar{\mathbf{v}}^t B\mathbf{v} = \lambda\bar{\mathbf{v}}^t A^{-1}\mathbf{v}$, which implies $\lambda \leq 1$.

Let us see some easy example of convex functions satisfying the assumptions of Theorem 7.28. For instance,

- for any $q > 1$, the function $f(t) = t^q$ satisfies these assumptions, since $s^d f(s^{-d}) = s^{-d(q-1)}$ is convex and decreasing;
- the entropy function $f(t) = t \log t$ also satisfies the assumptions since $s^d f(s^{-d}) = -d \log s$ is convex and decreasing;
- if $1 - \frac{1}{d} \leq m < 1$, the function $f(t) = -t^m$ is convex, and if we compute $s^d f(s^{-d}) = -t^{m(1-d)}$, we get a convex and decreasing function since $m(1 - d) < 1$; yet, these functions lack the superlinearity assumption, but this does not prevent us from applying the same proof of Theorem 7.28 to the case where μ_0 and μ_1 are supposed to be a priori absolutely continuous.

Let us see some consequences of Theorem 7.28 in the case $f(t) = t^q, q > 1$.

Proposition 7.29. *Consider an exponent $1 < q \leq +\infty$ and two probability measures $\mu_0, \mu_1 \in L^q(\Omega)$ (in the sense that they are absolutely continuous and their densities are L^q). Take the (unique) geodesic μ_t connecting them in $\mathbb{W}_p$ (for $p > 1$). Then the measures μ_t are also L^q and $||\mu_t||_{L^q} \leq \max\{||\mu_0||_{L^q}, ||\mu_1||_{L^q}\}$. When $q < +\infty$, we also have $||\mu_t||_{L^q}^q \leq (1-t)||\mu_0||_{L^q}^q + t||\mu_1||_{L^q}^q$.*

Proof. The case $q < +\infty$ is an easy consequence of Theorem 7.28. Actually, if we use $f(t) = t^q$, we have $\mathscr{F}(\mu) = ||\mu||_{L^q}^q$. Hence

$$||\mu_t||_{L^q}^q \leq (1-t)||\mu_0||_{L^q}^q + t||\mu_1||_{L^q}^q \leq (\max\{||\mu_0||_{L^q}, ||\mu_1||_{L^q}\})^q .$$

This allows to get the desired L^q estimate.

The case $q = +\infty$ is just obtained by taking the qth root and passing to the limit $q \to +\infty$. □

We finish this section observing that in 1D we can say something more.

Proposition 7.30. *Suppose $d = 1$, $\Omega = [a,b] \subset \mathbb{R}$ and take $q > 0$. Then the functional F defined through $F(\varrho) = \int_a^b \varrho(x)^{-q}\mathrm{d}x$ for $\rho \ll \mathscr{L}^1$ is displacement convex in $\mathbb{W}_p(\Omega)$. In particular, if $\varrho_0, \varrho_1 \geq c > 0$, then we also have $\varrho_t(x) \geq c$ for every $t \in [0,1]$, where ϱ_t is the geodesic connecting ϱ_0 to ϱ_1.*

Proof. Notice the function $f(s) = s^{-q}$ is convex for $s > 0$ and that $s \mapsto s^d f(s^{-d}) = s^{q+1}$ is convex but increasing. Yet, if we reread the proof of Theorem 7.28, we observe that we only needed it to be decreasing to ensure that the composition with g was convex, but in this case g is affine, and convexity of $s \mapsto s^d f(s^{-d})$ is enough. Moreover, if we assume $F(\varrho) < +\infty$, then we have $\varrho > 0$ a.e., which shows that it is not necessary to assume $f(0) = 0$. This proves the desired displacement convexity. To get the lower bound on geodesics, we just need to act as in Proposition 7.29, letting $q \to \infty$. □

7.3.3 Convexity on generalized geodesics

It is quite disappointing to note that the functional $\mu \mapsto W_2^2(\mu,\nu)$ is not, in general, displacement convex. This seems contrary to the intuition because usually squared distances are nice convex functions. However, we can see that this fails from the following easy example. Take $\nu = \frac{1}{2}\delta_{(1,0)} + \frac{1}{2}\delta_{(-1,0)}$ and $\mu_t = \frac{1}{2}\delta_{(t,a)} + \frac{1}{2}\delta_{(-t,-a)}$. The curve μ_t, if $a > 1$ is the geodesic between μ_{-1} and μ_1 (because the optimal transport between these measures sends $(a,-1)$ to $(a,1)$ and $(-a,1)$ to $(-a,-1)$). Yet, if we compute $W_2^2(\mu_t,\nu)$, we have

$$W_2^2(\mu_t,\nu) = a^2 + (1-t)^2 \wedge (1+t)^2.$$

But this function is not convex! (See Figure 7.1.)

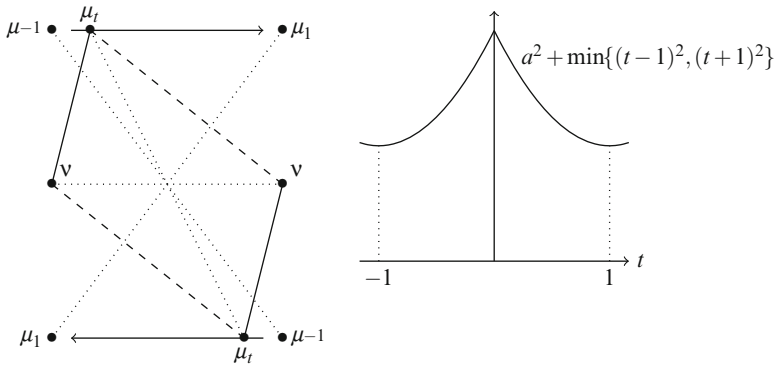


Fig. 7.1 The distance $W_2^2(\mu_t,\nu)$

The lack of geodesic convexity of this easy functional[4] is a problem for many issues (in particular for some metric approaches to gradient flows through the EVI condition; see Section 8.4.1) and an alternate notion has been proposed, namely, that of convexity along generalized geodesics.

Definition 7.31. If we fix an absolutely continuous probability $\varrho \in \mathscr{P}(\Omega)$, for every pair $\mu_0, \mu_1 \in \mathscr{P}(\Omega)$, we call *generalized geodesic* between μ_0 and μ_1 with base ϱ in $\mathbb{W}_2(\Omega)$ the curve $\mu_t = ((1-t)\mathrm{T}_0 + t\mathrm{T}_1)_{\#}\varrho$, where T_0 is the optimal transport map (for the cost $|x-y|^2$) from ϱ to μ_0 and T_1 from ϱ to μ_1.

It is clear that $t \mapsto W_2^2(\mu_t, \varrho)$ satisfies

$$\begin{aligned} W_2^2(\mu_t,\varrho) &\le \int |((1-t)\mathrm{T}_0(x) + t\mathrm{T}_1(x)) - x|^2 \mathrm{d}\varrho(x) \\ &\le (1-t)\int |\mathrm{T}_0(x)-x|^2 \mathrm{d}\varrho(x) + t\int |\mathrm{T}_1(x)-x|^2\mathrm{d}\varrho(x) \\ &= (1-t)W_2^2(\mu_0,\varrho) + tW_2^2(\mu_1,\varrho); \end{aligned}$$

and hence we have the desired convexity along this curve. Moreover, similar considerations to those we developed in this section show that all the functionals that we proved to be geodesically convex are also convex along generalized geodesics. We do not develop these proofs here, and we refer to [15] for more details: for the case of the functional $\mathscr{F}$, Lemma 7.26 has to be changed into "$t \mapsto \det((1-t)A + tB)^{1/d}$ is concave, whenever A and B are symmetric and positive-definite" (the proof is similar).

7.4 Discussion

7.4.1 *A case study:* $\min F(\varrho) + W_2^2(\varrho, \nu)$

Among variational problems involving optimal transportation and Wasserstein distances, a very recurrent one is the following:

$$\min_{\varrho\in\mathscr{P}_2(\Omega)} \frac{1}{2}W_2^2(\varrho,\nu) + \tau F(\varrho)\,, \tag{7.2}$$

where F is a given functional on probability measures, $\tau > 0$ a parameter which can possibly be small, and ν is a given probability in $\mathscr{P}_2(\Omega)$ (the space of probability measures on Ω with finite second moment $\int |x|^2 \mathrm{d}\varrho(x) < +\infty$). This very instance

[4] By the way, this functional can even be proven to be somehow geodesically concave, as it is shown in [15], Theorem 7.3.2.

of the problem is exactly the one we face in the time discretization of the gradient flow of F in $\mathbb{W}_2(\Omega)$ (see Section 8.2).

But the same problem also appears, for fixed $\tau > 0$, in other frameworks as well. For instance, in image processing, if F is a smoothing functional, this is a model to find a better (smoother) image ϱ which is not so far from the original density ν (and the choice of the distance W_2 is justified by robustness arguments); see [92, 208]. In urban planning (see [93, 270]), ν can represent the distribution of some resources and ϱ that of population, who wants to be close to ν but also to guarantee enough space to each individual. In this case the functional F favors diffuse measures (for instance, using the functional $\mathscr{F}(\varrho) = \int f(\varrho(x))\,\mathrm{d}x$ for a convex and superlinear function f with $f(0) = 0$, which gives a higher cost to high densities of ϱ and is minimized, on bounded domains, by the uniform measure). Reciprocally, ν could instead represent the distribution of population, and ϱ that of services, to be chosen so that they are close enough to ν but more concentrated. In this case F will favor concentrated measures. When F penalizes the number of points in the support of ϱ, then ϱ becomes finitely atomic and we face the so-called location problem (finding where to place a finite number of points so as to approximate a given continuous distribution), useful both in logistics and in quantization issues in signal processing. See, for instance, [72, 75] for a transport-based analysis of the asymptotics of this problem.

Note that, when F only takes value 0 and $+\infty$, the above problem becomes a projection problem. In particular (see Section 8.4.2), the projection onto the set of densities bounded above by the constant 1 has received a lot of attention because of its applications in the time discretization of evolution problems with density constraints, in particular for crowd motion (see [225, 265]), where a crowd is described as a population of particles which cannot overlap and cannot go beyond a certain threshold density.

In this section we would like to briefly discuss the characterization and regularity of the optimizer, at least in the case where $\mathscr{F}(\varrho) = \int f(\varrho)$ for a nice convex integrand $f : \mathbb{R}_+ \to \mathbb{R}$, superlinear at infinity (i.e., one of the main functionals studied in this chapter).

We are interested in the estimates that one can give on the minimizer $\hat{\varrho}$. They can be roughly divided into two categories: those which are independent of ν but depend on τ and those which are uniform in τ but refer to similar bounds on ν.

For instance, let us write down the optimality conditions for (7.2) in the case $\mathscr{F}(\varrho) = \int f(\varrho)$, using Proposition 7.20. First we suppose that $f \in C^2$, that we have $0 < c \le f'' \le C$, and that $\mathrm{spt}(\nu) = \Omega$. This allows to differentiate both terms in the functional (the assumption on ν is needed to have uniqueness of the c-concave Kantorovich potential), and we get

$$\begin{cases} \varphi + \tau f'(\hat{\varrho}) = C & \hat{\varrho} - \text{a.e.}, \\ \varphi + \tau f'(\hat{\varrho}) \ge C & \text{on } \{\hat{\varrho} = 0\}, \end{cases} \tag{7.3}$$

where φ is a suitable Kantorovich potential in the transport from $\hat{\varrho}$ to ν. We can write the above condition as

$$\hat{\varrho} = (\tau f')^{-1}\left((C-\varphi) \vee f'(0)\right). \tag{7.4}$$

This provides a Lipschitz bound on ϱ, which only depends on the Lipschitz constant of φ (which depends on $\mathrm{diam}(\Omega)$) and of $(\tau f')^{-1}$ (which is bounded by c^{-1}). In particular, if we let f and ν vary, keeping the condition $f'' \geq c > 0$, we can apply Ascoli-Arzelà's theorem to ϱ and pass to the limit (from the uniqueness of the minimizers, which is guaranteed by the strict convexity of f, the minimizers pass to the limit). Hence, we can guarantee that (7.4) holds under the only assumption $f'' \geq c > 0$. Notice that in the case $f'(0) = -\infty$, we get $\hat{\varrho} = (\tau f')^{-1}(C-\varphi)$ and that we prefer to rewrite as $\tau f'(\hat{\varrho}) = C - \varphi$.

These considerations allow to get continuity for $\hat{\varrho}$, but the bounds obviously degenerate as $\tau \to 0$. On the other hand, they do not really depend on ν.

A different type of bound that one can prove is the following:

Proposition 7.32. *If $g \in L^\infty$ is a given probability density on a compact convex and smooth domain Ω and $\hat{\varrho}$ solves*

$$\min_{\varrho \in \mathscr{P}(\Omega)} \frac{1}{2} W_2^2(\varrho, g) + \int f(\varrho(x))\,\mathrm{d}x,$$

where f is convex and superlinear, then $||\hat{\varrho}||_{L^\infty} \leq ||g||_{L^\infty}$.

This bound, which, on the contrary, is independent of τ, is proven in [110, 270]. We can provide a sketch of proof.

Proof. We start from the case where g is smooth and strictly positive, and f is smooth on $]0, +\infty[$, continuous on $\mathbb{R}_+$, with $f'' \geq c > 0$ and $f'(0) = -\infty$. From the above considerations (with $\tau = 1$), we know that the density of the optimal $\hat{\varrho}$ is a continuous function, given by $\hat{\varrho} = (f')^{-1}(C-\varphi)$. Note that $(f')^{-1}$ is Lipschitz continuous, as a consequence of the lower bound on f'', and that $\hat{\varrho} > 0$ since $\hat{\varrho} = 0$ would imply $-\varphi = -\infty$. From its continuity, $\hat{\varrho}$ admits a maximal value, $\hat{\varrho}(x_0)$. Note that a maximum point for $\hat{\varrho}$ corresponds to a minimum point for φ. Using Caffarelli's regularity theory (see Section 1.7.6), we can provide enough regularity for φ so as to justify all the following considerations (indeed, since g is smooth and $\hat{\varrho}$ is Lipschitz, and both are strictly positive, and hence bounded from below, on a convex set, one has at least $\varphi \in C^{2,\alpha}$). First, suppose that $x_0 \notin \partial\Omega$. In this case we should have $D^2\varphi(x_0) \geq 0$. But the Monge-Ampère equation gives us

$$\hat{\varrho}(x_0) = \det(\mathrm{I} - D^2\varphi(x_0))\, g(\mathrm{T}(x_0)),$$

where $\mathrm{T} = \mathrm{id} - \nabla\varphi$ is the optimal transport from $\hat{\varrho}$ to g. From $D^2\varphi(x_0) \geq 0$, we get $\det(\mathrm{I} - D^2\varphi(x_0)) \leq 1$, and hence

$$||\hat{\varrho}||_{L^\infty} = \hat{\varrho}(x_0) \le g(\mathrm{T}(x_0)) \le ||g||_{L^\infty}.$$

It remains to consider the case $x_0 \in \partial\Omega$. Since $\mathrm{T}(x_0) = x_0 - \nabla\varphi(x_0) \in \overline{\Omega}$ and Ω is convex, we have $\nabla\varphi(x_0) \cdot \mathbf{n} \ge 0$. By the optimality property of x_0 (which is a minimum point for φ), we also have the opposite inequality, and hence we obtain anyway $\nabla\varphi(x_0) \cdot \mathbf{n} = 0$. This allows to conclude that also in this case we have $D^2\varphi(x_0) \ge 0$ (indeed, second-order minimality conditions are also satisfied on the boundary if we already know that the gradient vanishes; since the tangential part always vanishes at extremal points, only the normal one has to be considered).

Finally, one can get rid of the assumptions on g and f by approximation (since we know that the minimizer of the original problem with $g \in L^\infty$ and f which is only convex and superlinear is unique). □

As a final remark, we stress that some limited higher-order similar estimates can also be obtained and are the object of very recent papers. We mention a BV estimate, namely, $||\hat{\varrho}||_{BV} \le ||g||_{BV}$, presented in [151], and based on some integration-by-parts techniques. We also mention a Lipschitz estimate which, together with some semi-concavity bounds, is proved in [207] in the case where $f(t) = t \log t$, thanks to some Pogorelov-type methods. Both estimates can be adapted to the case where a potential energy $\mathscr{V}$ is added to the functional. However, we do not develop details here.

7.4.2 Brunn-Minkowski inequality

Another interesting consequence of the displacement convexity in $\mathbb{W}_2$ is a transport-based proof of a well-known geometric inequality called Brunn-Minkowski inequality. This inequality states that for every two sets $E, F \subset \mathbb{R}^d$ (we ignore here measurability issues), we have

$$|E + F|^{1/d} \ge |E|^{1/d} + |F|^{1/d}, \quad \text{where } E + F = \{x + y : x \in E,\, y \in F\}.$$

This inequality can be proven in the following way, as pointed out in [229]. Consider the measures μ with constant density $\frac{1}{|E|}$ on E and ν with constant density $\frac{1}{|F|}$ on F. Then, take the measure $\mu_{1/2}$ obtained at time $t = 1/2$ on the constant-speed geodesic between them. Since $\mu_{1/2} = (\frac{1}{2}\mathrm{id} + \frac{1}{2}\mathrm{T})_\# \mu_0$, we know that $\mu_{1/2}$ is concentrated on $A := \frac{1}{2}E + \frac{1}{2}F$. Moreover, $\mu_{1/2}$ is absolutely continuous and we call ϱ its density (we know $\varrho \in L^\infty$ because $\mu, \nu \in L^\infty$, thanks to Proposition 7.29). Note that, for convex f, we have

$$\frac{1}{|A|} \int_A f(\varrho(x))\, \mathrm{d}x \ge f\left(\frac{\int_A \varrho(x)\, \mathrm{d}x}{|A|} \right) = f\left(|A|^{-1} \right).$$

Geodesic convexity of $\mathscr{F}$ for $f(t) = -t^{1-1/d}$ implies

$$-\frac{1}{2}|E|^{1/d} - \frac{1}{2}|F|^{1/d} = \frac{1}{2}\mathscr{F}(\mu) + \frac{1}{2}\mathscr{F}(\nu) \geq \mathscr{F}(\mu_{1/2})$$
$$= \int_A f(\varrho(x))\,\mathrm{d}x \geq |A| f\left(|A|^{-1}\right) = -|A|^{1/d}.$$

If we multiply this inequality by -2 on each side, using the scaling properties of the Lebesgue measure, we exactly obtain the Brunn-Minkowski inequality.

7.4.3 *Displacement convexity, game theory, and spatial economics: urban equilibria*

As optimal transport describes the cost of moving masses (or individuals) when their density is known, it is a very natural tool to analyze problems in spatial economics, where densities of agents and of activities are involved, and one wants to take into account geographical effects. Often, it is important to associate with this densities some "performances,"' which let the functionals that we analyzed in this chapter appear. Among the objects studied in spatial economics, one of the most intriguing, as we can see in our everyday life, is urban structure.

Many models exist in the economic literature and most of them have nothing to do with optimal transport. An interesting treatise of some spatial problems involving some notions of optimal transform (more under the form of flow minimization, such as in Chapter 4, then under the form of a Monge-Kantorovich problem) is contained in the book [28], which inspired by the way some of the works on congested transport contained in Section 4.4.1.

Yet, the goal of this section is not to give a comprehensive literature review on the subject, but to point out some recent problems in urban economics where the analysis we did in this chapter is particularly relevant. Indeed, we will see that displacement convexity can play a role in proving uniqueness of equilibria and their characterization as the minimizers of a global energy.

In particular, we will not spend words on the models for urban planning contained in [93, 94], where the goal is to find the optimal shapes of the distribution μ of the population and of the distribution ν of "services" in order to minimize a global quantity $\mathscr{F}(\mu) + \mathscr{T}_c(\mu, \nu) + \mathscr{G}(\nu)$. They are problems involving a social planner where no equilibrium issue is addressed, and optimal transport appears directly in the modeling and not as a tool. We will not spend word neither on the equilibrium model contained in [106, 107], where one looks at a distribution μ of residents, a distribution ν of jobs, a pairing γ between them, a function φ giving the price of the land, and a function ψ giving wages and at other related quantities. Here an equilibrium may be found by considering φ and ψ as Kantorovich potentials for

an appropriate cost, but the model is quite long to describe (see also [275] for a short summary of this and other models) and it uses optimal transport ideas to prove existence and not to characterize the equilibria.

We want rather to concentrate on another, simple class of models: suppose that a population ϱ is distributed in a region Ω and that every individual wants to choose his own position in order to minimize the sum of two costs:

- a residential cost, which is an increasing function of the density of population at the place where he lives: the individuals living at x "pay" a function of the form $h(\varrho(x))$, for $h : \mathbb{R}_+ \to \mathbb{R}_+$ increasing; this takes into account the fact that where more people live, the price of land is higher (or that, for the same price, they have less space);
- an interaction cost, depending on the distances with all the other individuals: people living at x "pay" a cost of the form $\int W(x-y)\varrho(y)\,\mathrm{d}y$ where W is usually an increasing function of the distance (i.e., $W(z)$ depends increasingly on $|z|$).

This means that the cost that every individual pays depends on his location x and on the whole distribution ϱ of everybody. We look for an equilibrium configuration, i.e., a density ϱ such that, for every x_0, there is no reason for people at x_0 to move to another location, since the function $f_\varrho(x) := h(\varrho(x)) + \int W(x-y)\varrho(y)\,\mathrm{d}y$ is minimal at x_0, in the spirit of Nash equilibria.

Box 7.3. Important notion: *Nash equilibria*

Definition. Consider a game where several players $i = 1, \dots, n$ must choose a strategy among a set of possibilities S_i and suppose that the pay-off of each player (i.e., how much he gains out of the game) depends on what everybody chooses, i.e., it is given by a function $p_i : S_1 \times \cdots \times S_n \to \mathbb{R}$. We say that a configuration $(s_1, \dots, s_n)$ (where $s_i \in S_i$) is an equilibrium (a *Nash equilibrium*) if, for every i, the choice s_i optimizes $S_i \ni s \mapsto f_i(s_1, \dots, s_{i-1}, s, s_{i+1}, \dots, s_n)$ (i.e., s_i is optimal for player i under the assumption that the other players freeze their choice).

Note that Nash equilibria need not exist in all situations, but the strength of Nash approach was exactly to prove (via convex analysis arguments) that they always exist when we consider the so-called mixed strategies. This means that we accept that every player instead of choosing an element $s_i \in S_i$, only chooses a probability on S_i and then randomly picks a strategy according to the law he has chosen. This allows to convexify the game in a way which is similar to what we do when we pass from the Monge to Kantorovich problems.

This notion, first introduced by J. Nash in [243, 244] in the case of a finite number of players, can be easily extended to a continuum of players where each one is negligible compared to the others (*nonatomic games*). Considering for simplicity the case of identical players, we have a common space S of possible strategies and we look for a measure $\varrho \in \mathscr{P}(S)$. This measure induces a payoff function $f_\varrho : S \to \mathbb{R}$ and we want the following condition to be satisfied: there exists $C \in \mathbb{R}$ such that $f_\varrho(x) = C$ ϱ-a.e. and $f_\varrho(x) \le C$ everywhere (if the players want to maximize the playoff, otherwise, if it is a cost to be minimized, $f_\varrho(x) \ge C$), i.e., f_ϱ must be optimal ϱ-a.e.

It is easy to see that the equilibrium condition, namely, that

$$h(\varrho(x)) + \int W(x-y)\varrho(y)\,\mathrm{d}y$$

must be minimal ϱ-a.e. is indeed the first-order optimality condition for the minimization of

$$\varrho \mapsto F(\varrho) := \int H(\varrho(x))\,\mathrm{d}x + \frac{1}{2}\int\!\!\int W(x-y)\varrho(y)\,\mathrm{d}y\varrho(x)\,\mathrm{d}x$$

among probabilities $\varrho \in \mathscr{P}(\Omega)$, where $H' = h$ (just compute the first variation with respect to perturbation of the form $\varrho_\varepsilon = (1-\varepsilon)\varrho + \varepsilon\tilde{\varrho}$). We can wonder whether this optimality condition, which is necessary for minimality, is also sufficient and if all equilibria are minimizers. This would allow to say that we face a *potential game* (i.e., a game where equilibria can be found by minimizing a global energy; see [238]). Note by the way that the energy that we minimize is not the total cost for residents, which would be given by $\int f_\varrho(x)\varrho(x)\,\mathrm{d}x$ (compare to the notion of price of anarchy in Remark 4.29). This last quantity differs from what we consider both in the $\frac{1}{2}$ factor in the interaction part and in the fact that $H(t) \neq th(t)$ in general (as we already saw in Section 4.4.1 concerning traffic congestion).

The answer to the above question would be affirmative if the functional F was convex. Yet, the first part is convex as H is convex (as it is the antiderivative of an increasing function), but the interaction energy is not, as we saw in this chapter. On the other hand, if H satisfies McCann's assumptions for displacement convexity and W is convex (which is the case for the interaction costs $|x-y|$ or $|x-y|^2$, for instance), then F is displacement convex. An interesting observation in [59] is exactly the fact that displacement convexity is also enough, as usual convexity would do, in order to guarantee that equilibria are all minimizers of F. To prove this fact, it is enough to compute the derivative of $F(\varrho_\varepsilon)$ at $\varepsilon = 0$, when we take $\varrho_\varepsilon = ((1-\varepsilon)\mathrm{id} + \varepsilon T)_\#\varrho$, where T is the optimal transport map between ϱ and any other admissible $\tilde{\varrho}$. By the way, it is not a problem to add to the equilibrium condition a geographical heterogeneity (i.e., a potential function), taking $f_\varrho(x) = h(\varrho(x)) + \int W(x-y)\varrho(y)\,\mathrm{d}y + V(x)$, and in this case we add to $F(\varrho)$ a term in $\int V(x)\varrho(x)\,\mathrm{d}x$. The convexity of V would be the good assumption to perform the same analysis.

Since we are evoking McCann's condition on the convex function H, it is worthwhile to spend some words on its meaning in economic terms. To do that, we need to check how h and H depend on the density through the formation of the price of land. This comes from an easy consideration: suppose that at every point, the agents have a certain budget to be divided into land consumption and money consumption and that they have a concave and increasing utility function U for land. This means they solve a problem of the form

$$\max\{U(L) + m \,:\, pL + m \leq 0\},$$

where p represents the price for land, L is the land consumption, m is the leftover of the money, and the budget constraint has been set to 0 for simplicity. The optimal land consumption will saturate the constraint and be such that $U'(L) = p$. This allows to compute the optimal utility, which is $U(L) - U'(L)L$, and this gives a relation between L and utility. Yet, the land consumption is the reciprocal of the density, hence $L = \frac{1}{\varrho}$, and the residential cost $h(\varrho)$, which is a cost and not a utility, can be taken to be opposite as this utility. Hence, we take

$$h(\varrho) = \frac{1}{\varrho} U'\left(\frac{1}{\varrho}\right) - U\left(\frac{1}{\varrho}\right).$$

Since the function $t \mapsto \frac{1}{t}U'(\frac{1}{t}) - U(\frac{1}{t})$ is the derivative of $-tU(\frac{1}{t})$, the correct choice for H is indeed $H(t) = -tU(\frac{1}{t})$. This makes the verification of McCann's condition for displacement convexity very easy, since $t \mapsto t^d H(t^{-d})$ becomes $t \mapsto -U(t^d)$, and we just need $t \mapsto U(t^d)$ to be concave and increasing. Monotonicity is guaranteed since U is monotone, while the fact that it stays concave when composed with the dth power is indeed an extra assumption[5].

We also stress that, in order to use displacement convexity, one needs a convex domain Ω and that [59] also presents examples of multiple equilibria when the convex domain is replaced by a 1D circle (equilibria with multiple sub-cities are found, i.e., $\mathrm{spt}(\varrho)$ can be disconnected, and the multiplicity of the equilibria is not only due to rotational invariance, but equilibria with different shapes and different number of components are found). Curiously enough, the model described in [59] strongly recalls a model by R. McCann for rotating stars, described in [232]. In such a paper, the author introduces a notion of W_∞-local minimizers which can be translated into a notion of local equilibrium: each agent only compares his cost to that of nearby agents; should the support of ϱ be disconnected (which is the case in [59] and [232]), then the constant value of their utility could change from one connected component to the others. Note the early use of the W_∞ distance and the fact that this is the only distance allowing for this kind of local phenomena. Indeed, local minimality for integral distances such as W_p, if translated via optimality conditions into the expression of an equilibrium, would not provide local minimality for each agent, but automatically global minimality. This can be observed in the example of the circle in [59]: in this example many equilibria exist, and each agent is globally satisfied, not only locally, but most of them are not global minimizers of the global energy.

An even more interesting situation is presented in [56], in the framework of the so-called *Cournot-Nash* equilibria. This name is used to denote nonatomic games with agents which are not indistinguishable. The situation is similar to the one described above, but with an extra variable, which makes the distinction between

[5]Thanks to an observation by R. McCann himself, this corresponds to the fact that the utility is a concave function not only of land "volume" but also of land "linear size," which seems reasonable since it has to be compared to "linear" quantities such as distances in the interaction term.

agents: suppose that a population of agents μ distributed over a space Z must choose, each, a location $x \in \Omega$ and that the payoff of each agent z is given by $c(z,x)+f_\varrho(x)$, where ϱ is the distribution of their choices on Ω. More precisely, we look for a measure $\gamma \in \mathscr{P}(Z \times \Omega)$ such that $(\pi_Z)_\#\gamma = \mu$, and we denote $(\pi_\Omega)_\#\gamma$ by ϱ. The goal is to find γ such that no agent z wants to change his choice, i.e., there is a function $\varphi : Z \to \mathbb{R}$ with $c(z,x)+f_\varrho(x) \geq \varphi(z)$ for all (z,x) and $c(z,x)+f_\varrho(x) = \varphi(z)$ γ-a.e. Note that the variable z represents the type of the agent, which cannot be modified by the agent itself: this is why the minimality of $c(z,x)+f_\varrho(x)$ is tested only on x and gives a minimal value depending on z.

It is not difficult to prove the following facts (see [56]): a measure γ is an equilibrium if and only if it is optimal for the cost c in $\Pi(\mu,\nu)$, the pair $(\varphi,-f_\varrho)$ is a pair of Kantorovich potentials for the same transport problems, and the c-transform φ^c satisfies $\varphi^c(x)+f_\varrho(x) \geq 0$ for all x and $\varphi^c(x)+f_\varrho(x) = 0$ ϱ-a.e. This leads to a condition on ϱ only: find ϱ such that $\psi + f_\varrho$ is minimal ϱ-a.e., where $\psi = \varphi^c$ is the Kantorovich potential in the transport from ϱ to the given measure μ, for the cost c.

The same considerations of this chapter allow to view this as the optimality condition for the minimization of $\varrho \mapsto \mathscr{T}_c(\mu,\varrho) + F(\varrho)$ and the same questions about displacement convexity hold. As we saw in Section 7.3.3, the functional $\mathscr{T}_c(\mu,\cdot)$ is not in general displacement convex, and one needs to use convexity on generalized geodesics. But the computations and the result are similar, as one can see in [56].

Finally, it is interesting to see that the same kind of models would appear if, instead of defining an equilibrium on agent type z who must choose x, one directly sets an equilibrium problem on x: agents have to choose where to locate, facing both a residential cost and an interaction cost, but also need to access a commodity z, distributed according to a distribution μ, and their access cost is $c(x,z)$. Since the distribution of the commodity is fixed, the agents cannot simply choose the z that they prefer (i.e., minimizing $z \mapsto c(x,z)$): a matching problem will appear and a price $\varphi(z)$ for each commodity. At equilibrium, according to the theory that we saw in Section 1.7.3, the commodity cost for the agents will be given by $\varphi(z) + c(x,z) = \varphi^c(y)$. This means that the agents add to their cost a new term, given by the Kantorovich potential φ^c, exactly as above.

Chapter 8
Gradient flows

In this chapter, we present one of the most spectacular applications of optimal transport and Wasserstein distances to PDEs. We will see that several evolution PDEs can be interpreted as a steepest descent movement in the space $\mathbb{W}_2$. As we already developed in the previous chapters many of the technical tools that we need, a large part of this chapter will be devoted to an informal description of the general framework.

Hence, no need to let the reader wait any more, let us immediately enter into the subject!

8.1 Gradient flows in $\mathbb{R}^d$ and in metric spaces

First of all, let us present what a gradient flow is in the simplest situation. Suppose you have a function $F : \mathbb{R}^d \to \mathbb{R}$ and a point $x_0 \in \mathbb{R}^d$. A gradient flow is an evolution stemming from x_0 and always moving in the direction where F decreases the most, thus "gradually minimizing" F, starting from x_0. More precisely, it is just the solution of the Cauchy problem

$$\begin{cases} x'(t) = -\nabla F(x(t)) & \text{for } t > 0, \\ x(0) = x_0. \end{cases}$$

This is a standard Cauchy problem which has a unique solution if ∇F is Lipschitz continuous, i.e., if $F \in C^{1,1}$. We will see that existence and uniqueness can also hold without this strong assumption, thanks to the variational structure of the equation.

A first interesting property is the following, concerning uniqueness and estimates.

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2_8

Proposition 8.1. *Suppose that F is convex and let x_1 and x_2 be two solutions of $x'(t) = -\nabla F(x(t))$ (if F is not differentiable, we can consider $x'(t) \in \partial F(x(t))$). Then we have $|x_1(t) - x_2(t)| \le |x_1(0) - x_2(0)|$ for every t. In particular this gives uniqueness of the solution of the Cauchy problem.*

Proof. Let us consider $g(t) = \frac{1}{2}|x_1(t) - x_2(t)|^2$ and differentiate it. We have

$$g'(t) = (x_1(t) - x_2(t)) \cdot (x_1'(t) - x_2'(t)) = -(x_1(t) - x_2(t)) \cdot (\nabla F(x_1(t)) - \nabla F(x_2(t))).$$

Here we use the basic property of gradient of convex functions, i.e., that for every x_1 and x_2, we have

$$(x_1 - x_2) \cdot (\nabla F(x_1) - \nabla F(x_2)) \ge 0.$$

More generally, it is also true that for every x_1, x_2 and every $p_1 \in \partial F(x_1)$, $p_2 \in \partial F(x_2)$, we have

$$(x_1 - x_2) \cdot (p_1 - p_2) \ge 0.$$

From these considerations, we obtain $g'(t) \le 0$ and $g(t) \le g(0)$. This gives the first part of the claim.

Then, if we take two different solutions of the same Cauchy problem, we have $x_1(0) = x_2(0)$, and this implies $x_1(t) = x_2(t)$ for any $t > 0$. □

Remark 8.2. From the same proof, one can also deduce uniqueness and stability estimates in the case where F is only semi-convex. We recall that semi-convex means that it is λ-convex for some $\lambda \in \mathbb{R}$, i.e., $x \mapsto F(x) - \frac{\lambda}{2}|x|^2$ is convex. Indeed, in this case we obtain $|x_1(t) - x_2(t)| \le |x_1(0) - x_2(0)|e^{-\lambda t}$, which also proves, if $\lambda > 0$, exponential convergence to the unique minimizer of F. The key point is that if F is λ-convex, it is easy, by applying the monotonicity inequalities above to $x \mapsto F(x) - \frac{\lambda}{2}|x|^2$, to get

$$(x_1 - x_2) \cdot (\nabla F(x_1) - \nabla F(x_2)) \ge \lambda |x_1(t) - x_2(t)|^2.$$

This implies $g'(t) \le -2\lambda g(t)$ and allows to conclude, by Gronwall's lemma, $g(t) \le g(0)e^{-2\lambda t}$. For the exponential convergence, if $\lambda > 0$, then F is coercive and admits a minimizer, which is unique by strict convexity. Let us call it $\bar{x}$. Take a solution $x(t)$ and compare it to the constant curve $\bar{x}$, which is a solution since $0 \in \partial F(\bar{x})$. Then we get $|x_1(t) - \bar{x}| \le e^{-\lambda t}|x_1(0) - \bar{x}|$.

Another interesting feature of those particular Cauchy problems which are gradient flows is their discretization in time. Actually, one can fix a small time step parameter $\tau > 0$ and look for a sequence of points $(x_k^\tau)_k$ defined through

$$x_{k+1}^\tau \in \operatorname{argmin}_x F(x) + \frac{|x - x_k^\tau|^2}{2\tau}.$$

We can forget now the convexity assumptions on F, which are not necessary for this part of the analysis. Indeed, very mild assumptions on F (l.s.c. and some lower bounds, for instance, $F(x) \geq C_1 - C_2|x|^2$) are sufficient to guarantee that these problems admit a solution for small τ. The case where F is λ-convex is covered by these assumptions and also provides uniqueness of the minimizers. This is evident if $\lambda > 0$ since we have strict convexity for every τ, and if λ is negative, the sum will be strictly convex for small τ.

We can interpret this sequence of points as the values of the curve $x(t)$ at times $t = 0, \tau, 2\tau, \dots, k\tau, \dots$. It happens that the optimality conditions of the recursive minimization exactly give a connection between these minimization problems and the equation, since we have

$$x^\tau_{k+1} \in \operatorname{argmin} F(x) + \frac{|x - x^\tau_k|^2}{2\tau} \Rightarrow \nabla F(x^\tau_{k+1}) + \frac{x^\tau_{k+1} - x^\tau_k}{\tau} = 0,$$

i.e.

$$\frac{x^\tau_{k+1} - x^\tau_k}{\tau} = -\nabla F(x^\tau_{k+1}).$$

This expression is exactly the discrete-time implicit Euler scheme for $x' = -\nabla F(x)$!

Box 8.1. ***Memo:*** *Explicit and implicit Euler schemes*

Given an ODE $x'(t) = \mathbf{v}(x(t))$ (that we suppose autonomous for simplicity), with given initial datum $x(0) = x_0$, Euler schemes are time discretization where derivatives are replaced by finite differences. We fix a time step $\tau > 0$ and define a sequence x^τ_k. The explicit scheme is given by

$$x^\tau_{k+1} = x^\tau_k + \tau \mathbf{v}(x^\tau_k), \quad x^\tau_0 = x_0.$$

The implicit scheme, on the contrary, is given by

$$x^\tau_{k+1} = x^\tau_k + \tau \mathbf{v}(x^\tau_{k+1}), \quad x^\tau_0 = x_0.$$

This means that x^τ_{k+1} is selected as a solution of an equation involving x^τ_k, instead of being explicitly computable from x^τ_k. The explicit scheme is obviously easier to implement, but enjoys less stability and qualitative properties than the implicit one. Suppose, for instance, $\mathbf{v} = -\nabla F$: then the quantity $F(x(t))$ decreases in t in the continuous solution, which is also the case for the implicit scheme, but not for the explicit one (which represents the iteration of the gradient method for the minimization of F).

Note that the same can be done for evolution PDEs and that solving the heat equation $\partial_t \varrho = \Delta \varrho_t$ by an explicit scheme is very dangerous: at every step, ϱ^τ_{k+1} would have two degrees of regularity less than ϱ^τ_k, since it is obtained through $\varrho^\tau_{k+1} = \varrho^\tau_k - \Delta \varrho^\tau_k$.

It is possible to prove that for $\tau \to 0$, the sequence we found, suitably interpolated, converges to the solution of the problem. It even suggests how to define solutions for functions F which are only l.s.c., with no gradient at all!

But a huge advantage of this discretized formulation is also that it can easily be adapted to metric spaces. Actually, if one has a metric space (X, d) and an l.s.c. function $F : X \to \mathbb{R} \cup \{+\infty\}$ bounded from below, one can define

$$x_{k+1}^\tau \in \operatorname{argmin}_x F(x) + \frac{d(x, x_k^\tau)^2}{2\tau} \tag{8.1}$$

and study the limit as $\tau \to 0$. Obviously the assumptions on F have to be adapted, since we need existence of the minimizers and hence a little bit of compactness. But if X is compact, then everything works for F only l.s.c.

We can consider two different interpolations of the points x_k^τ, given, respectively, by

$$x^\tau(t) = x_k^\tau, \qquad \tilde{x}^\tau(t) = \omega_{x_{k-1}^\tau, x_k^\tau}\left(\frac{t-(k-1)\tau}{\tau}\right) \quad \text{for } t \in](k-1)\tau, k\tau],$$

where $\omega_{x,y}(s)$ denotes any constant-speed geodesic connecting a point x to a point y, parametrized on the unit interval $[0, 1]$. The interpolation $\tilde{x}^\tau$ only makes sense in spaces where geodesics exist, obviously. It is in this case a continuous (locally Lipschitz) curve, which coincides with the piecewise constant interpolation x^τ at times $t = k\tau$.

De Giorgi, in [137], defined[1] the notion of *minimizing movements* (see also [7]).

Definition 8.3. A curve $x : [0, T] \to X$ is said to be a minimizing movement if there exists a sequence of time steps $\tau_j \to 0$ such that the piecewise constant interpolations x^{τ_j}, built from a sequence of solutions of the iterated minimization scheme (8.1), uniformly converge to x on $[0, T]$.

Compactness results guaranteeing the existence of minimizing movements are derived from a simple property giving a Hölder behavior for the curves x^τ: for every τ and every k, the optimality of x_{k+1}^τ provides

$$F(x_{k+1}^\tau) + \frac{d(x_{k+1}^\tau, x_k^\tau)^2}{2\tau} \le F(x_k^\tau), \tag{8.2}$$

which implies

$$d(x_{k+1}^\tau, x_k^\tau)^2 \le 2\tau \left(F(x_k^\tau) - F(x_{k+1}^\tau)\right).$$

[1] We avoid here the distinction between minimizing movements and generalized minimizing movements, which is not crucial in our analysis.

If $F(x_0)$ is finite and F is bounded from below, taking the sum over k, we have

$$\sum_{k=0}^{l} d(x_{k+1}^\tau, x_k^\tau)^2 \le 2\tau \left(F(x_0^\tau) - F(x_{l+1}^\tau)\right) \le C\tau.$$

Given $t \in [0,T]$, denote by $\ell(t)$ the unique integer such that $t \in](\ell(t)-1)\tau, \ell(t)\tau]$. For $t < s$ apply the Cauchy-Schwartz inequality to the sum between the indices $\ell(t)$ and $\ell(s)$ (note that we have $|\ell(t) - \ell(s)| \le \frac{|t-s|}{\tau} + 1$):

$$d(x^\tau(t), x^\tau(s)) \le \sum_{k=\ell(t)}^{\ell(s)-1} d(x_{k+1}^\tau, x_k^\tau) \le \left(\sum_{k=\ell(t)}^{\ell(s)-1} d(x_{k+1}^\tau, x_k^\tau)^2\right)^{\frac{1}{2}} |\ell(s) - \ell(t)|^{\frac{1}{2}}$$

$$\le C\sqrt{\tau}\frac{(|t-s|+\tau)^{\frac{1}{2}}}{\sqrt{\tau}} \le C\left(|t-s|^{1/2} + \sqrt{\tau}\right).$$

This means that the curves x^τ – if we forget for a while that they are actually discontinuous – are morally equi-Hölder continuous of exponent $1/2$. The situation is even clearer for $\tilde{x}^\tau$. Indeed, we have

$$\tau\left(\frac{d(x_{k-1}^\tau, x_k^\tau)}{\tau}\right)^2 = \int_{(k-1)\tau}^{k\tau} |(\tilde{x}^\tau)'|^2(t)\mathrm{d}t,$$

which implies, by summing up,

$$\int_0^T |(\tilde{x}^\tau)'|^2(t)\mathrm{d}t \le C.$$

This means that the curves $\tilde{x}^\tau$ are bounded in $H^1([0,T];X)$ (the space of absolutely continuous curves with square-integrable metric derivative), which also implies a $C^{0,\frac{1}{2}}$ bound by usual Sobolev embedding.

If the space X, the distance d, and the functional F are explicitly known, in some cases it is already possible to pass to the limit in the optimality conditions of each optimization problem in the time-discrete setting and to characterize the limit curve $x(t)$. It will be possible to do so in the space of probability measures, which is the topic of this chapter, but not in many other cases. Indeed, without a little bit of (differential) structure on X, it is almost impossible. If we want to develop a general theory of gradient flows in metric spaces, we need to exploit finer tools, able to characterize, with the only help of metric quantities, the fact that a curve $x(t)$ is a gradient flow.

We present here an inequality which is satisfied by gradient flows in the smooth Euclidean case and which can be used as a definition of gradient flow in the metric case, since all the quantities which are involved have their metric counterpart. Another inequality, called EVI, will be presented in the Discussion Section 8.4.1.

The first observation is the following: thanks to the Cauchy-Schwartz inequality, for every curve $x(t)$, we have

$$F(x(s)) - F(x(t)) = \int_s^t -\nabla F(x(r)) \cdot x'(r)\, dr \le \int_s^t |\nabla F(x(r))||x'(r)|\, dr$$
$$\le \int_s^t \left(\frac{1}{2}|x'(r)|^2 + \frac{1}{2}|\nabla F(x(r))|^2 \right) dr.$$

Here, the first inequality is an equality if and only if $x'(r)$ and $\nabla F(x(r))$ are vectors with opposite directions for a.e. r, and the second is an equality if and only if their norms are equal. Hence, the condition, called EDE (*Energy Dissipation Equality*)

$$F(x(s)) - F(x(t)) = \int_s^t \left(\frac{1}{2}|x'(r)|^2 + \frac{1}{2}|\nabla F(x(r))|^2 \right) \mathrm{d}r, \qquad \text{for all } s < t$$

(or even the simple inequality $F(x(s)) - F(x(t)) \ge \int_s^t \left(\frac{1}{2}|x'(r)|^2 + \frac{1}{2}|\nabla F(x(r))|^2 \right) \mathrm{d}r$) is equivalent to $x' = -\nabla F(x)$ a.e. and could be taken as a definition of gradient flow. This is what is done in a series of works by Ambrosio, Gigli, and Savaré and in particular in their book [15]. Developing this (huge) theory is not among the scopes of this book. The reader who is curious about this abstract approach but wants to find a synthesis of their work from the point of view of the author can have a look at [277]. In such a paper, the role of the different parts of the theory with respect to the possible applications is clarified as far as possible[2] (we also discuss in short some of these issues in Section 8.4.1).

8.2 Gradient flows in $\mathbb{W}_2$, derivation of the PDE

In this section we give a short and sketchy presentation of what can be done when we consider the gradient flow of a functional $F : \mathscr{P}(\Omega) \to \mathbb{R} \cup \{+\infty\}$. The functional will be supposed l.s.c. for the weak convergence of probabilities and Ω compact (which implies that F admits a minimizer and is bounded from below). In particular, we will give an heuristic on how to derive a PDE from the optimality conditions at each time step. This is obviously something that we can only do in this particular metric space and does not work in the general case of an abstract metric space.[3]

Indeed, we exploit the fact that we know, from Chapter 5, that all absolutely continuous curves in the space $\mathbb{W}_2(\Omega)$ are solution of the continuity equation $\partial_t \varrho_t +$

[2] Unfortunately, some knowledge of French is required (even if not forbidden, English is unusual in the Bourbaki seminar, since "Nicolas Bourbaki a une préférence pour le français").

[3] The original ideas which led to this theory come from [198] and [246] through what is known as "Otto's formal calculus", and have later been reconsidered in [15].

$\nabla \cdot (\varrho_t \mathbf{v}_t) = 0$, for a suitable vector field $\mathbf{v}$. The goal is to identify the vector field $\mathbf{v}_t$, which will depend on the measure ϱ_t, in a way which is ruled by the functional F.

We consider the iterated minimization scheme

$$\varrho^\tau_{(k+1)} \in \operatorname{argmin}_\varrho \; F(\varrho) + \frac{W_2^2(\varrho, \varrho^\tau_{(k)})}{2\tau}. \tag{8.3}$$

Each of these problems has a solution, by compactness of $\mathscr{P}(\Omega)$ and semi-continuity of the objective function. Exactly as we did for the Euclidean case (F defined on $\mathbb{R}^d$), we want to study the optimality conditions of these problems so as to guess the limit equation.

We recall the notation $\frac{\delta G}{\delta \varrho}(\varrho)$ for the first variation of a functional $G : \mathscr{P}(\Omega) \to \mathbb{R} \cup \{+\infty\}$. We will suppose that $\frac{\delta F}{\delta \varrho}$ is known and try to write the limit equation in terms of this operator.

Now, take an optimal measure $\hat\varrho$ for the minimization problem at step k and compute variations with respect to perturbations of the form $\varrho_\varepsilon := (1-\varepsilon)\hat\varrho + \varepsilon\tilde\varrho$, where $\tilde\varrho$ is any other probability measure. Using Proposition 7.20, we obtain

$$\frac{\delta F}{\delta \varrho}(\varrho) + \frac{\varphi}{\tau} = \textit{constant},$$

where φ is a Kantorovich potential for the transport with cost $\frac12|x-y|^2$ from $\hat\varrho$ to $\varrho^\tau_{(k)}$. Actually, the above equality holds $\hat\varrho$-a.e. and requires uniqueness of φ. For both these points proving $\varrho_\varepsilon > 0$ will be enough, but we will fix this when doing the rigorous proof.

If we combine the fact that the above sum is constant and that we have $\mathrm{T}(x) = x - \nabla\varphi(x)$ for the optimal T, we get

$$\frac{\mathrm{T}(x)-x}{\tau} = -\frac{\nabla\varphi(x)}{\tau} = \nabla\Big(\frac{\delta F}{\delta \varrho}(\varrho)\Big)(x). \tag{8.4}$$

We will denote by $-\mathbf{v}$ the ratio $\frac{\mathrm{T}(x)-x}{\tau}$. Why? Because, as a ratio between a displacement and a time step, it has the meaning of a velocity, but since it is the displacement associated with the transport from $\varrho^\tau_{(k+1)}$ to $\varrho^\tau_{(k)}$, it is better to view it rather as a backward velocity (which justifies the minus sign).

Since here we have $\mathbf{v} = -\nabla\big(\frac{\delta F}{\delta \varrho}(\varrho)\big)$, this suggests that at the limit $\tau \to 0$, we will find a solution of

$$\partial_t \varrho - \nabla \cdot \Big(\varrho\, \nabla\Big(\frac{\delta F}{\delta \varrho}(\varrho)\Big)\Big) = 0,$$

with no-flux boundary conditions on $\partial\Omega$.

Before entering into the details making the above approach rigorous (next section), we want to present some examples of this kind of equations. We will consider two functionals that we already analyzed in Chapter 7, and more precisely

$$\mathscr{F}(\varrho) = \int f(\varrho(x))\,\mathrm{d}x, \quad \text{and} \quad \mathscr{V}(\varrho) = \int V(x)\mathrm{d}\varrho.$$

In this case we already saw that we have

$$\frac{\delta\mathscr{F}}{\delta\varrho}(\varrho) = f'(\varrho), \quad \frac{\delta\mathscr{V}}{\delta\varrho}(\varrho) = V.$$

An interesting example is the case $f(t) = t\log t$. In such a case, we have $f'(t) = \log t + 1$ and $\nabla(f'(\varrho)) = \frac{\nabla\varrho}{\varrho}$: this means that the gradient flow equation associated with the functional $\mathscr{F}$ would be the *heat equation*

$$\partial_t\varrho - \Delta\varrho = 0,$$

and that for $\mathscr{F} + \mathscr{V}$ we would have the *Fokker-Planck equation*

$$\partial_t\varrho - \Delta\varrho - \nabla\cdot(\varrho\nabla V) = 0.$$

We will see a list of other gradient flow equations in Discussion section, including the so-called porous media equation, obtained for other choices of f.

Remark 8.4. One could wonder how the usual stochastic diffusion interpretation of the heat and Fokker-Planck equations is compatible with the completely deterministic framework that we present here, where the velocity of each particle is given by $-\nabla\varrho/\varrho - \nabla V$ (take $V = 0$ to get the heat equation). Indeed, the Fokker-Planck equation is also the PDE solved by the density of a bunch of particle following a stochastic equation $dX_t = -\nabla V(X_t)dt + dB_t$, where B is a standard Brownian motion (independent for each particle). The answer, as it is well pointed out in [88], is based on the idea that in this optimal transport interpretation we rearrange (i.e., relabel) the particles. This fact is quite clear in the 1D case: if at every time step particles move from x to $x - \tau\nabla V(x) + B_\tau$ (i.e., they follow the drift $-\nabla V$ and they diffuse with a Gaussian law), and then we reorder them (we always call particle 1 the one which is the most at the left, particle 2 the next one ... and particle N the one which is the most at the right), then we have a discrete family of trajectories which converge, when τ goes to 0 and the number N to infinity, to a solution of Fokker-Planck where the velocity is exactly given by $-\nabla\varrho/\varrho - \nabla V$. Similar considerations can be done in higher dimension with suitable notions of rearrangement.

Note that all these PDEs come accompanied by Neumann boundary conditions $\varrho\frac{\partial}{\partial\mathbf{n}}(\frac{\delta F}{\delta\varrho}(\varrho)) = 0$ on $\partial\Omega$, as a consequence of the no-flux conditions for the continuity equation of Section 5.3. We will see in Section 8.4.3 a case of extension to Dirichlet boundary conditions.

We finish this section with some philosophical thoughts. Why should we study some PDEs by considering them as gradient flows for the distance W_2? There are at least three reasons. The first one is that it allows to give an existence result (of a weak solution of such a PDE), with the technique that we will see in the next section. This is obviously useless when the equation is already well known, as it is the case

for the heat equation or its Fokker-Planck variant. But the same strategy could be applied to variants of the same equations (or to similar equations in stranger spaces, as it is nowadays done for the heat flow in metric measure spaces; see [18, 183]). It is also very useful when the equation is of new type (we will see the case of the crowd motion models in Section 8.4.2). The second goal could be to derive properties of the flow and of the solution, once we know that it is a gradient flow. A simple one is the fact that $t \mapsto F(\varrho_t)$ must be decreasing in time. However, this can be often deduced from the PDE in different ways and does not require in general to have a gradient flow. The third goal concerns numerics: the discretized scheme to get a gradient flow is itself a sort of algorithm to approximate the solution. If the minimization problems at each time step are discretized and attacked in a suitable way, this can provide efficient numerical tools (see, for instance, [40]).

Finally, all the procedure we presented is related to the study and the existence of weak solutions. What about uniqueness? We stress that in PDE applications the important point is to prove uniqueness for weak solutions of the continuity equation with $\mathbf{v} = -\nabla \frac{\delta F}{\delta \varrho}$ (i.e., we do not care at the metric structure and at the definitions of EVI and EDE). In some cases, this uniqueness could be studied independently of the gradient flow structure (this is the case for the heat equation, for instance). Anyway, should we use PDE approaches for weak solutions, or abstract approaches in metric spaces, it turns out that usually uniqueness is linked to some kind of convexity, or λ-convexity, and in particular to displacement convexity. This is why a large part of the theory has been developed for λ-geodesically convex functionals.

8.3 Analysis of the Fokker-Planck case

We consider here the case study of the Fokker-Planck equation, which is the gradient flow of the functional

$$J(\varrho) = \int_\Omega \varrho \log \varrho + \int_\Omega V \mathrm{d}\varrho,$$

where V is a Lipschitz function on the compact domain Ω. The initial measure $\varrho_0 \in \mathscr{P}(\Omega)$ is taken such that $J(\varrho_0) < +\infty$.

We stress that the first term of the functional is defined as

$$\mathscr{F}(\varrho) := \begin{cases} \int_\Omega \varrho(x) \log \varrho(x)\, \mathrm{d}x & \text{if } \varrho \ll \mathscr{L}^d, \\ +\infty & \text{otherwise,} \end{cases}$$

where we identify the measure ϱ with its density, when it is absolutely continuous. This functional is l.s.c. thanks to Proposition 7.7 (which can be applied since we are on a domain with finite measure; otherwise the situation is trickier, and we refer to **Ex**(45)).

Semi-continuity is the key tool to establish the following result:

Proposition 8.5. *The functional J has a unique minimum over $\mathscr{P}(\Omega)$. In particular J is bounded from below. Moreover, for each $\tau > 0$, the following sequence of optimization problems recursively defined is well-posed:*

$$\varrho^\tau_{(k+1)} \in \operatorname{argmin}_\varrho \quad J(\varrho) + \frac{W_2^2(\varrho, \varrho^\tau_{(k)})}{2\tau}, \tag{8.5}$$

which means that there is a minimizer at every step, and this minimizer is unique.

Proof. Just apply the direct method, noting that $\mathscr{P}(\Omega)$ is compact for the weak convergence, which is the same as the convergence for the W_2 distance (again, because Ω is compact), and for this convergence $\mathscr{F}$ is l.s.c. and the other terms are continuous. This gives at the same time the existence of a minimizer for J and of a solution to each of the above minimization problems (8.5). Uniqueness comes from the fact that all the functionals are convex (in the usual sense) and $\mathscr{F}$ is strictly convex. □

Optimality conditions at each time step

We will use Propositions 7.17, 7.20, and 7.18, and to do so we first need to prove the following result:

Lemma 8.6. *Any minimizer $\hat\varrho$ in (8.5) must satisfy $\hat\varrho > 0$ a.e. and $\log\hat\varrho \in L^1$.*

Proof. Consider the measure $\tilde\varrho$ with constant positive density c in Ω (i.e., $c = |\Omega|^{-1}$). Let us define ϱ_ε as $(1-\varepsilon)\hat\varrho + \varepsilon\tilde\varrho$ and compare $\hat\varrho$ to ϱ_ε.

By optimality of $\hat\varrho$, we may write

$$\mathscr{F}(\hat\varrho) - \mathscr{F}(\varrho_\varepsilon) \le \int_\Omega V \mathrm{d}\varrho_\varepsilon - \int_\Omega V d\hat\varrho + \frac{W_2^2(\varrho_\varepsilon, \varrho^\tau_{(k)})}{2\tau} - \frac{W_2^2(\hat\varrho, \varrho^\tau_{(k)})}{2\tau}. \tag{8.6}$$

The Wasserstein term in the right-hand side may be estimated by convexity:

$$\frac{W_2^2(\varrho_\varepsilon, \varrho^\tau_{(k)})}{2\tau} \le (1-\varepsilon)\frac{W_2^2(\hat\varrho, \varrho^\tau_{(k)})}{2\tau} + \varepsilon\frac{W_2^2(\tilde\varrho, \varrho^\tau_{(k)})}{2\tau}.$$

The potential term is of order ε as well:

$$\int_\Omega V \mathrm{d}\varrho_\varepsilon - \int_\Omega V d\hat\varrho = \varepsilon \int_\Omega V \mathrm{d}(\tilde\varrho - \hat\varrho) \le C\varepsilon.$$

This shows that the whole right-hand side of (8.6) is estimated by $C\varepsilon$ and we get

$$\int_\Omega f(\hat{\varrho}) - f(\varrho_\varepsilon) \le C\varepsilon$$

where $f(t) = t \log t$ (set to 0 in $t = 0$). Write

$$A = \{x \in \Omega \,:\, \hat{\varrho}(x) > 0\}, \quad B = \{x \in \Omega \,:\, \hat{\varrho}(x) = 0\}.$$

Since f is convex, we write, for $x \in A$, $f(\hat{\varrho}(x)) - f(\varrho_\varepsilon(x)) \ge (\hat{\varrho}(x) - \varrho_\varepsilon(x)) f'(\varrho_\varepsilon(x)) = \varepsilon(\hat{\varrho}(x) - \tilde{\varrho}(x))(1 + \log \varrho_\varepsilon(x))$. For $x \in B$ we simply write $f(\hat{\varrho}(x)) - f(\varrho_\varepsilon(x)) = -\varepsilon c \log(\varepsilon c)$. This allows to write

$$-\varepsilon c \log(\varepsilon c)|B| + \varepsilon \int_A (\hat{\varrho}(x) - c)(1 + \log \varrho_\varepsilon(x))\, \mathrm{d}x \le C\varepsilon$$

and, dividing by ε,

$$-c \log(\varepsilon c)|B| + \int_A (\hat{\varrho}(x) - c)(1 + \log \varrho_\varepsilon(x))\, \mathrm{d}x \le C. \tag{8.7}$$

Note that we always have

$$(\hat{\varrho}(x) - c)(1 + \log \varrho_\varepsilon(x)) \ge (\hat{\varrho}(x) - c)(1 + \log c)$$

(just distinguish between the case $\hat{\varrho}(x) \ge c$ and $\hat{\varrho}(x) \le c$). Thus, we may write

$$-c \log(\varepsilon c)|B| + \int_A (\hat{\varrho}(x) - c)(1 + \log c)\, \mathrm{d}x \le C.$$

Letting $\varepsilon \to 0$ provides a contradiction, unless $|B| = 0$.

This proves $\hat{\varrho} > 0$ a.e. We now come back to (8.7), which is an upper bound on the integral of the functions $(\hat{\varrho}(x) - c)(1 + \log \varrho_\varepsilon(x))$. We already noted that these functions are bounded from below by $(\hat{\varrho}(x) - c)(1 + \log c)$, which is L^1 because c is a constant and $\hat{\varrho} \in L^1$. Hence, we can apply Fatou's lemma and obtain, at the limit as $\varepsilon \to 0$,

$$\int_\Omega (\hat{\varrho}(x) - c)(1 + \log \hat{\varrho}(x))\, \mathrm{d}x \le C$$

(the integral on A has been replaced by the integral on Ω, since B is negligible). Since we already know that $(\hat{\varrho} - c)(1 + \log \hat{\varrho})$ is bounded from below by an L^1 function and its integral is finite, then it is L^1. But we already know that $\hat{\varrho}, \hat{\varrho} \log \hat{\varrho} \in L^1$, and we deduce $\log \hat{\varrho} \in L^1$. □

We can now compute the first variation and give optimality conditions on the optimal $\varrho^\tau_{(k+1)}$.

Proposition 8.7. *The optimal measure $\varrho^\tau_{(k+1)}$ in (8.5) satisfies*

$$\log(\varrho^\tau_{(k+1)}) + V + \frac{\bar\varphi}{\tau} = \text{constant a.e.} \tag{8.8}$$

where $\bar\varphi$ is the (unique) Kantorovich potential from $\varrho^\tau_{(k+1)}$ to $\varrho^\tau_{(k)}$. In particular, $\log \varrho^\tau_{(k+1)}$ is Lipschitz continuous. If T^τ_{k+1} is the optimal transport from $\varrho^\tau_{(k+1)}$ to $\varrho^\tau_{(k)}$, then it satisfies

$$\mathbf{v}^\tau_{(k+1)} := \frac{\mathrm{id} - \mathrm{T}^\tau_{k+1}}{\tau} = -\nabla\left(\log(\varrho^\tau_{(k+1)}) + V\right) \text{ a.e.} \tag{8.9}$$

Proof. Take the optimal measure $\hat\varrho := \varrho^\tau_{(k+1)}$. We can say that it is regular (see Definition 7.12) for the functional J, since the only term that is not finite on all probabilities in $\mathscr{P}(\Omega)$ is the entropy term, but it is convex and finite both on the optimum $\hat\varrho$ and on L^∞ densities. We can check that J admits a first variation at $\hat\varrho$. For the linear term it is straightforward, and for the Wasserstein term, we can apply Proposition 7.17. The uniqueness of the Kantorovich potential is guaranteed by Proposition 7.18 together with Lemma 8.6. For the entropy term, consider $\varrho_\varepsilon := (1-\varepsilon)\hat\varrho + \varepsilon\tilde\varrho$, for $\tilde\varrho \in L^\infty$. Set $M = ||\tilde\varrho||_{L^\infty}$ and look at $\varepsilon \mapsto \int \varrho_\varepsilon \log \varrho_\varepsilon$. The integrand can be differentiated in ε pointwisely, thus getting $(1+\log\varrho_\varepsilon)(\tilde\varrho - \hat\varrho)$. For $\varepsilon < 1/2$, we can check that these functions are dominated by $(\hat\varrho + M)(|\log\hat\varrho| + \log M + 1)$, which is L^1 thanks to $\hat\varrho, \log\hat\varrho, \hat\varrho\log\hat\varrho \in L^1$ (from Lemma 8.6). This allows to differentiate under the integral sign and proves that the first variation of this term is $1 + \log\hat\varrho$.

The first variation of J is hence

$$\frac{\delta J}{\delta \varrho} = f'(\varrho) + V = \log(\varrho) + 1 + V.$$

Applying Proposition 7.20 we obtain Equation (8.8), which is valid a.e. since $\hat\varrho > 0$. In particular, this implies that $\varrho^\tau_{(k+1)}$ is Lipschitz continuous, since we have

$$\varrho^\tau_{(k+1)}(x) = \exp\left(C - V(x) - \frac{\bar\varphi(x)}{\tau}\right).$$

Then, one differentiates and gets the equality

$$\nabla\bar\varphi = \frac{\mathrm{id} - \mathrm{T}^\tau_{k+1}}{\tau} = -\nabla\left(\log(\varrho^\tau_{(k+1)}) + V\right) \text{ a.e.}$$

and this allows to conclude. □

Interpolation between time steps and uniform estimates

Let us collect some other tools.

Proposition 8.8. *For any $\tau > 0$, the sequence of minimizers satisfies*

$$\sum_k \frac{W_2^2(\varrho^\tau_{(k+1)}, \varrho^\tau_{(k)})}{\tau} \leq C := 2(J(\varrho_0) - \inf J).$$

Proof. This is obtained by comparing the optimizer $\varrho^\tau_{(k+1)}$ to the previous measure $\varrho^\tau_{(k)}$. We get

$$J(\varrho^\tau_{(k+1)}) + \frac{W_2^2(\varrho^\tau_{(k+1)}, \varrho^\tau_{(k)})}{2\tau} \leq J(\varrho^\tau_k), \tag{8.10}$$

which implies

$$\sum_k \frac{W_2^2(\varrho^\tau_{(k+1)}, \varrho^\tau_{(k)})}{\tau} \leq \sum_k 2(J(\varrho^\tau_k) - J(\varrho^\tau_{(k+1)})).$$

This last sum is telescopic and gives the claim. □

Let us define two interpolations between the measures $\varrho^\tau_{(k)}$.

With this time-discretized method, we have obtained, for each $\tau > 0$, a sequence $(\varrho^\tau_{(k)})_k$. We can use it to build at least two interesting curves in the space of measures:

- first we can define some piecewise constant curves, i.e., $\varrho^\tau_t := \varrho^\tau_{(k+1)}$ for $t \in]k\tau, (k+1)\tau]$; associated with this curve, we also define the velocities $\mathbf{v}^\tau_t = \mathbf{v}^\tau_{(k+1)}$ for $t \in]k\tau, (k+1)\tau]$, where $\mathbf{v}^\tau_{(k+1)}$ is defined as in (8.9): $\mathbf{v}^\tau_{(k+1)} = (\mathrm{id} - \mathrm{T}^\tau_{k+1})/\tau$, taking as T^τ_{k+1} the optimal transport from $\varrho^\tau_{(k+1)}$ to $\varrho^\tau_{(k)}$; we also define the momentum variable $E^\tau = \varrho^\tau \mathbf{v}^\tau$;
- then, we can also consider the densities $\tilde{\varrho}^\tau_t$ that interpolate the discrete values $(\varrho^\tau_{(k)})_k$ along geodesics:

$$\tilde{\varrho}^\tau_t = \left(\frac{k\tau - t}{\tau}\mathbf{v}^\tau_{(k)} + \mathrm{id}\right)_\# \varrho^\tau_{(k)}, \quad \text{for } t \in](k-1)\tau, k\tau[; \tag{8.11}$$

the velocities $\tilde{\mathbf{v}}^\tau_t$ are defined so that $(\tilde{\varrho}^\tau, \tilde{\mathbf{v}}^\tau)$ satisfy the continuity equation and $||\tilde{\mathbf{v}}^\tau_t||_{L^2(\tilde{\varrho}^\tau_t)} = |(\tilde{\varrho}^\tau)'|(t)$. To do so, we take

$$\tilde{\mathbf{v}}^\tau_t = \mathbf{v}^\tau_t \circ \left((k\tau - t)\mathbf{v}^\tau_{(k)} + \mathrm{id}\right)^{-1};$$

as before, we define a momentum variable: $\tilde{E}_\tau = \tilde{\varrho}^\tau \tilde{\mathbf{v}}^\tau$.

After these definitions we consider some a priori bounds on the curves and the velocities that we defined. We start from some estimates which are standard in the framework of minimizing movements.

Note that the speed (i.e., metric derivative) of $\tilde{\varrho}^\tau$ is constant on each interval $]k\tau, (k+1)\tau[$ and equal to

$$\frac{W_2(\varrho^\tau_{(k+1)}, \varrho^\tau_{(k)})}{\tau} = \frac{1}{\tau}\left(\int |\mathrm{id} - \mathrm{T}^\tau_{k+1}|^2 \mathrm{d}\varrho^\tau_{(k+1)}\right)^{1/2} = ||\mathbf{v}^\tau_{k+1}||_{L^2(\varrho^\tau_{(k+1)})},$$

which gives

$$||\tilde{\mathbf{v}}^\tau_t||_{L^2(\tilde{\varrho}^\tau_t)} = |(\tilde{\varrho}^\tau)'|(t) = \frac{W_2(\varrho^\tau_{(k+1)}, \varrho^\tau_{(k)})}{\tau} = ||\mathbf{v}^\tau_t||_{L^2(\varrho^\tau_t)},$$

where we used the fact that the velocity field $\tilde{\mathbf{v}}^\tau$ has been chosen so that its L^2 norm equals the metric derivative of the curve $\tilde{\varrho}^\tau$.

In particular, we can obtain

$$\begin{aligned}|E^\tau|([0,T]\times\Omega) &= \int_0^T \mathrm{d}t \int_\Omega |\mathbf{v}^\tau_t| \mathrm{d}\varrho^\tau_t = \int_0^T ||\mathbf{v}^\tau_t||_{L^1(\varrho^\tau_t)} \mathrm{d}t \le \int_0^T ||\mathbf{v}^\tau_t||_{L^2(\varrho^\tau_t)} \mathrm{d}t \\ &\le T^{1/2} \int_0^T ||\mathbf{v}^\tau_t||^2_{L^2(\varrho^\tau_t)} \mathrm{d}t = T^{1/2} \sum_k \tau \left(\frac{W_2(\varrho^\tau_{(k+1)}, \varrho^\tau_{(k)})}{\tau}\right)^2 \le C.\end{aligned}$$

The estimate on $\tilde{E}^\tau$ is completely analogous

$$\begin{aligned}|\tilde{E}^\tau|([0,T]\times\Omega) &= \int_0^T \mathrm{d}t \int_\Omega |\tilde{\mathbf{v}}^\tau_t| \mathrm{d}\tilde{\varrho}^\tau_t \le T^{1/2} \int_0^T ||\tilde{\mathbf{v}}^\tau_t||^2_{L^2(\tilde{\varrho}^\tau_t)} \\ &= T^{1/2} \sum_k \tau \left(\frac{W_2(\varrho^\tau_{(k+1)}, \varrho^\tau_{(k)})}{\tau}\right)^2 \le C.\end{aligned}$$

This gives compactness of E^τ and $\tilde{E}^\tau$ in the space of vector measures on space-time, for the weak convergence. As far as $\tilde{\varrho}^\tau$ is concerned, we can obtain more than that. Consider the following estimate, for $s < t$:

$$W_2(\tilde{\varrho}^\tau_t, \tilde{\varrho}^\tau_s) \le \int_s^t |(\tilde{\varrho}^\tau)'|(r) \mathrm{d}r \le (t-s)^{1/2} \left(\int_s^t |(\tilde{\varrho}^\tau)'|(r)^2 \mathrm{d}r\right)^{1/2}.$$

From the previous computations, we have again

$$\int_0^T |(\tilde{\varrho}^\tau)'|(r)^2 \mathrm{d}r = \sum_k \tau \left(\frac{W_2(\varrho^\tau_{(k+1)}, \varrho^\tau_{(k)})}{\tau}\right)^2 \le C,$$

and this implies

$$W_2(\tilde{\varrho}^\tau_t, \tilde{\varrho}^\tau_s) \le C(t-s)^{1/2}, \tag{8.12}$$

which means that the curves $\tilde{\varrho}^\tau$ are uniformly Hölder continuous. Since they are defined on $[0,T]$ and valued in $\mathbb{W}_2(\Omega)$ which is compact, we can apply the Ascoli-Arzelà theorem. This implies that, up to subsequences, we have

$$E^\tau \rightharpoonup E \text{ in } \mathscr{M}^d([0,T]\times\Omega),\ \tilde{E}^\tau \rightharpoonup \tilde{E} \text{ in } \mathscr{M}^d([0,T]\times\Omega);$$
$$\tilde{\varrho}^\tau \to \varrho \text{ uniformly for the } W_2 \text{ distance.}$$

The limit curve ϱ, from the uniform bounds on $\tilde{\varrho}^\tau$, is both $\frac{1}{2}$-Hölder continuous and absolutely continuous in $\mathbb{W}_2$.

As far as the curves ϱ^τ are concerned, they also converge uniformly to the same curve ϱ, since $W_2(\varrho^\tau_t, \tilde{\varrho}^\tau_t) \le C\sqrt{\tau}$ (a consequence of (8.12), of the fact that $\tilde{\varrho}^\tau = \varrho^\tau$ on the points of the form $k\tau$ and of the fact that ϱ^τ is constant on each interval $]k\tau,(k+1)\tau]$).

Let us now prove that $\tilde{E} = E$.

Lemma 8.9. *Suppose that we have two families of vector measures E^τ and $\tilde{E}^\tau$ such that*

- $\tilde{E}^\tau = \tilde{\varrho}^\tau\tilde{\mathbf{v}}^\tau$; $E^\tau = \varrho^\tau\mathbf{v}^\tau$;
- $\tilde{\mathbf{v}}^\tau_t = \mathbf{v}^\tau_t \circ \Big((k\tau - t)\mathbf{v}^\tau_{(k)} + \mathrm{id}\Big)^{-1}$; $\tilde{\varrho}^\tau = \Big((k\tau - t)\mathbf{v}^\tau_{(k)} + \mathrm{id}\Big)_\# \varrho^\tau$;
- $\iint |\mathbf{v}^\tau|^2 \mathrm{d}\varrho^\tau \le C$ *(with C independent of τ);*
- $E^\tau \rightharpoonup E$ *and* $\tilde{E}^\tau \rightharpoonup \tilde{E}$ *as* $\tau \to 0$.

Then $\tilde{E} = E$.

Proof. It is sufficient to fix a Lipschitz function $f : [0,T]\times\Omega \to \mathbb{R}^d$ and to prove $\int f\cdot \mathrm{d}E = \int f\cdot \mathrm{d}\tilde{E}$. To do that, we write

$$\int f\cdot \mathrm{d}\tilde{E}^\tau = \int_0^T \mathrm{d}t \int_\Omega f\cdot\tilde{\mathbf{v}}^\tau_t \mathrm{d}\tilde{\varrho}^\tau = \int_0^T \mathrm{d}t \int_\Omega f\circ((k\tau-t)\mathbf{v}^\tau + \mathrm{id})\cdot\mathbf{v}^\tau_t \mathrm{d}\varrho^\tau,$$

which implies

$$\left|\int f\cdot \mathrm{d}\tilde{E}^\tau - \int f\cdot \mathrm{d}E^\tau\right| \le \int_0^T \mathrm{d}t \int_\Omega |f\circ((k\tau-t)\mathbf{v}^\tau + \mathrm{id}) - f|\,|\mathbf{v}^\tau_t|\mathrm{d}\varrho^\tau$$
$$\le \mathrm{Lip}(f)\tau \int_0^T\int_\Omega |\mathbf{v}^\tau_t|^2 \mathrm{d}\varrho^\tau \le C\tau.$$

This estimate proves that the limit of $\int f\cdot \mathrm{d}\tilde{E}^\tau$ and $\int f\cdot \mathrm{d}E^\tau$ is the same, i.e., $E = \tilde{E}$. □

Relation between ϱ and E

We can obtain the following:

Proposition 8.10. *The pair (ϱ, E) satisfies, in distributional sense,*

$$\partial_t \varrho + \nabla \cdot E = 0, \quad E = -\nabla \varrho - \varrho \nabla V,$$

with no-flux boundary conditions on $\partial\Omega$. In particular we have found a solution to

$$\begin{cases} \partial_t \varrho - \Delta \varrho + \nabla \cdot (\varrho \nabla V) = 0, \\ \varrho(0) = \varrho_0, \end{cases}$$

where the initial datum is to be intended in the following sense: the curve $t \mapsto \varrho_t$ is (absolutely) continuous in $\mathbb{W}_2$, and its initial value is ϱ_0.

Proof. First, consider the weak convergence $(\tilde{\varrho}^\tau, \tilde{E}^\tau) \rightharpoonup (\varrho, E)$ (which is a consequence of $\tilde{E} = E$). The continuity equation $\partial_t \tilde{\varrho}^\tau + \nabla \cdot \tilde{E}^\tau = 0$ is satisfied (as in Section 4.1.2) in the sense of distributions by $(\tilde{\varrho}^\tau, \tilde{E}^\tau)$ (we test the equations against C^1_c function on $]0, T[\times\Omega$, which means in particular that we do not require the test functions to vanish on $\partial\Omega$) and this passes to the limit.

Hence, $\partial_t \varrho + \nabla \cdot E = 0$.

The continuity in $\mathbb{W}_2$ and the initial datum pass to the limit because of the uniform $C^{0,1/2}$ bound in (8.12).

Then, use the convergence $(\varrho^\tau, E^\tau) \rightharpoonup (\varrho, E)$. Actually, using the optimality conditions of Proposition 8.7 and the definition of $E^\tau = \mathbf{v}^\tau \varrho^\tau$, we have, for each $\tau > 0$, $E^\tau = -\nabla \varrho^\tau - \varrho^\tau \nabla V$ (again, in the sense of distributions: the term $\varrho^\tau \nabla V$ is the product of L^1 and L^∞ functions). It is not difficult to pass this condition to the limit either. Take $f \in C^1_c(]0, T[\times\Omega; \mathbb{R}^d)$ and test:

$$\int f \cdot \mathrm{d}E^\tau = -\int f \cdot \nabla \varrho^\tau - \int f \cdot \nabla V \varrho^\tau = \int \nabla \cdot f \mathrm{d}\varrho^\tau - \int f \cdot \nabla V \varrho^\tau.$$

These terms pass to the limit as $\varrho^\tau \rightharpoonup \varrho$, at least if $V \in C^1$, since all the test functions above are continuous. This would give $\int f \cdot \mathrm{d}E = \int (\nabla \cdot f) \mathrm{d}\varrho - \int f \cdot \nabla V \, \mathrm{d}\varrho$, which implies $E = -\nabla \varrho - \varrho \nabla V$ (once more, $\nabla \varrho$ is to be intended in the sense of distributions).

To handle the case where V is only Lipschitz continuous, let us note that for every τ, t we have $J(\varrho_t^\tau) \le J(\varrho_0)$ (this is a consequence of (8.10), which iterated over k, gives $J(\varrho_{(k)}^\tau)) \le J(\varrho_0)$). This gives a uniform bound on $\mathscr{F}(\varrho_t^\tau)$ and the Dunford-Pettis theorem below turns the weak convergence $\varrho_t^\tau \rightharpoonup \varrho_t$ as measures into a weak convergence in L^1. Look at the term

$$\int f \cdot \nabla V \varrho^\tau := \int_0^T \mathrm{d}t \int_\Omega f(t,x) \cdot \nabla V(x) \varrho_t^\tau(x) \, \mathrm{d}x.$$

The integrand in t is uniformly bounded by $||f||_{L^\infty}||\nabla V||_{L^\infty}$. Hence, we can compute its limit as $\tau \to 0$ looking at the limit for each fixed t because it is dominated. And, for fixed t, we have weak convergence in L^1 of ϱ_t^τ to ϱ_t, which, multiplied with a fixed L^∞ function, i.e., $f \cdot \nabla V$, gives the desired limit of the integral. □

Box 8.2. ***Memo:*** *Equi-integrability and the Dunford-Pettis theorem*

Definition. A sequence of function ϱ_n is said to be equi-integrable if for every $\varepsilon > 0$ there exists $\delta > 0$ such that on every set A with $|A| \le \delta$, we have $\int_A \varrho_n(x)\,\mathrm{d}x \le \varepsilon$ for any n.

We can check that equi-integrability is equivalent to a bound $\int f(\varrho_n(x))\,\mathrm{d}x \le C < +\infty$ for any superlinear function f. Let us only prove that such a bound implies equi-integrability. Indeed, fix $\varepsilon > 0$ and take M such that for all t with $t > M$, we have $f(t)/t > 2C/\varepsilon$ (this is possible thanks to the superlinearity of f). Then we estimate

$$\int_A \varrho_n = \int_{A\cap\{\varrho_n \le M\}} \varrho_n + \int_{A\cap\{\varrho_n > M\}} \frac{\varrho_n}{f(\varrho_n)} f(\varrho_n)\,\mathrm{d}x \le M|A| + \frac{\varepsilon}{2C}C.$$

It is now enough to take $\delta = \frac{\varepsilon}{2M}$ and we get $\int_A \varrho_n(x)\,\mathrm{d}x < \varepsilon$.

Theorem. Suppose that ϱ_n is a sequence of probability densities weakly converging as measures to $\varrho \in \mathscr{P}(\Omega)$. Suppose that $\int f(\varrho_n)$ is bounded for some superlinear function f. Then ϱ is also absolutely continuous and the weak convergence also holds in L^1 (i.e., in duality with L^∞ functions and not only with continuous ones).

Proof. The absolute continuity of ϱ is a consequence of the lower semi-continuity of $\mathscr{F}$ defined as $\mathscr{F}(\varrho) = \int_\Omega f(\varrho(x))\,\mathrm{d}x$ if $\varrho \ll \mathscr{L}^d$ ($+\infty$ otherwise).

Moreover, the bound on $\int f(\varrho_n(x))\,\mathrm{d}x$ implies equi-integrability. Fix a test function $\varphi \in L^\infty$ and use the fact that for every $\varepsilon > 0$, there is a $\delta > 0$ with $|A| \le \delta \Rightarrow \int_A \varrho_n(x)\,\mathrm{d}x \le \varepsilon$ and that, by Lusin's theorem, for this $\delta > 0$, there exists a continuous function $\tilde\varphi$ with $||\tilde\varphi||_{L\infty} \le ||\varphi||_{L\infty}$ and $\{\varphi \ne \tilde\varphi\} \subset A$, where A is open with $|A| < \delta$. We have

$$\left|\int \varphi(\varrho_n - \varrho)\right| \le \left|\int \tilde\varphi(\varrho_n - \varrho)\right| + 2||\varphi||_{L^\infty}\int_A (\varrho_n + \varrho).$$

Since A is open and $\varrho_n \rightharpoonup \varrho$, we have $\int_A \varrho(x)\,\mathrm{d}x \le \liminf \int_A \varrho_n(x)\,\mathrm{d}x \le \varepsilon$; hence

$$\limsup \left|\int \varphi(x)\varrho_n(x)\,\mathrm{d}x - \int \varphi(x)\varrho(x)\,\mathrm{d}x\right| \le 0 + 2||\varphi||_{L^\infty}\varepsilon,$$

which implies, ε being arbitrary, $\int \varphi(x)\varrho_n(x)\,\mathrm{d}x \to \int \varphi(x)\varrho(x)\,\mathrm{d}x$.

8.4 Discussion

8.4.1 EVI, uniqueness, and geodesic convexity

In the general theory of gradient flows [15] in metric spaces, another characterization, different from the EDE, is proposed in order to cope with uniqueness and

stability results. It is based on the following observation: if $F : \mathbb{R}^d \to \mathbb{R}$ is convex, then the inequality

$$F(y) \geq F(x) + p \cdot (y - x) \quad \text{for all } y \in \mathbb{R}^d$$

characterizes (by definition) the vectors $p \in \partial F(x)$, and if $F \in C^1$, it is only satisfied for $p = \nabla F(x)$. Analogously, if F is λ-convex, the inequality that characterizes the gradient is

$$F(y) \geq F(x) + \frac{\lambda}{2}|x - y|^2 + p \cdot (y - x) \quad \text{for all } y \in \mathbb{R}^d.$$

Hence, we can pick a curve $x(t)$ and a point y and compute

$$\frac{d}{dt}\frac{1}{2}|x(t) - y|^2 = (y - x(t)) \cdot (-x'(t)).$$

Consequently, imposing

$$\frac{d}{dt}\frac{1}{2}|x(t) - y|^2 \leq F(y) - F(x(t)) - \frac{\lambda}{2}|x(t) - y|^2,$$

for all y, will be equivalent to $-x'(t) = \nabla F(x(t))$. This will provide a second characterization (called EVI, *evolution variational inequality*) of gradient flows in a metric environment. Indeed, all the terms appearing in the above inequality have a metric counterpart (only squared distances and derivatives w.r.t. time appear). Even if we often forget the dependence on λ, it should be noted that the condition EVI should actually be written as EVI_λ, since it involves a parameter λ, which is a priori arbitrary. Actually, λ-convexity of F is not necessary to define the EVI_λ property, but it will be necessary in order to guarantee the existence of curves which satisfy such a condition. The notion of λ-convexity will hence be crucial also in metric spaces, where it will be rather "λ-geodesic-convexity."

The role of the EVI condition in the uniqueness and stability of gradient flows is quite easy to guess. Take two curves, which we call $x(t)$ and $y(s)$, and compute

$$\frac{d}{dt}\frac{1}{2}d(x(t), y(s))^2 \leq F(y(s)) - F(x(t)) - \frac{\lambda}{2}d(x(t), y(s))^2, \tag{8.13}$$

$$\frac{d}{ds}\frac{1}{2}d(x(t), y(s))^2 \leq F(x(t)) - F(y(s)) - \frac{\lambda}{2}d(x(t), y(s))^2. \tag{8.14}$$

If one wants to estimate $E(t) = \frac{1}{2}d(x(t), y(t))^2$, summing up the two above inequalities, after a chain-rule argument for the composition of the function of two variables[4] $(t, s) \mapsto \frac{1}{2}d(x(t), y(s))^2$ and of the curve $t \mapsto (t, t)$, gives

[4]Note that a similar argument, based on doubling the variables, is also often performed for the differentiation of W_2^2 along curves in $\mathbb{W}_2$, which we did differently in Section 5.3.5.

$$\frac{d}{dt}E(t) \le -2\lambda E(t).$$

By Gronwall's lemma, this provides uniqueness (when $x(0) = y(0)$) and stability.

The situation concerning these two different notions of gradient flows (EVI and EDE) in abstract metric spaces has been clarified by Savaré (in an unpublished note, but the proof can also be found in [9]), who showed that

- All curves which are gradient flows in the EVI sense also satisfy the EDE condition.
- The EDE condition is not in general enough to guarantee uniqueness of the gradient flow. As a simple example, take $X = \mathbb{R}^2$ with the ℓ^∞ distance

$$d((x_1, x_2), (y_1, y_2)) = |x_1 - y_1| \vee |x_2 - y_2|,$$

  and take $F(x_1, x_2) = x_1$; we can check that any curve $(x_1(t), x_2(t))$ with $x_1'(t) = -1$ and $|x_2'(t)| \le 1$ satisfies EDE.
- On the other hand, the existence of a gradient flow in the EDE sense is quite easy to get and provable under very mild assumption.
- The EVI condition is in general too strong in order to get existence (in the example above of the ℓ^∞ norm, no EVI gradient flow would exist), but always guarantees uniqueness and stability (w.r.t. initial data).

Also, the existence of EVI gradient flows is itself very restricting on the function F: indeed, it is proven in [134] that if F is such that from every starting point x_0 there is an EVI_λ gradient flow, then F is necessarily λ-geodesically convex.

We need to stress that all these abstract considerations on EDE and EVI are mandatory when dealing with a general metric framework, but can often be skipped when dealing with precise applications. For the sake of the applications to PDE that are at the core of this chapter, the point of view which seems to be more concrete is the following: there is an evolution model translated into a PDE, and we want to study this equation (existence; uniqueness and stability; qualitative, regularity, and asymptotical properties; numerical simulations). If this PDE happens to be the gradient flow of a certain functional for the W_2 distance, then much information can be obtained by the gradient flow approach (more specifically, the minimizing movement approach via discretized variational problems). But the notion of solution that one should use is a priori the one which came with the PDE in the modeling of the problem (in general, the notion of weak solution of a continuity equation). This is why, even if (nontrivial) results exist proving equivalence between some abstract (EVI or EDE) notions in the metric space $\mathbb{W}_2$ and weak solutions of evolution equations (see [15]), we preferred not to enter into these details here in this book.

On the other hand, we will see in a while (next section) that, as a last remark in favor of the general theory developed in [15], it is not always easy to prove existence via the minimizing movement method when we face nonlinear terms and that exploiting the characterization of the gradient flows in $\mathbb{W}_2$ developed in [15] could be sometimes easier.

8.4.2 Other gradient flow PDEs

We saw in Section 8.3 the example, and a detailed analysis, of the Fokker-Planck equation. The main reason to choose such example is its simplicity, because it is a linear equation. This allows easily to pass to the limit all the terms in the relation between E and ϱ. Yet, many other important equations can be obtained as gradient flows in $\mathbb{W}_2$, choosing other functionals. We will see some of them here, without entering into details of the proof. We stress that handling nonlinear terms is often difficult and requires ad hoc estimates. We will discuss some of the ideas that one should apply. Note, on the other hand, that if one uses the abstract theory of [15], there is no need to produce these ad hoc estimates: after developing a general (and hard) theory for general metric spaces, the second part of [15] explains which are the curves that one finds as gradient flows in $\mathbb{W}_2$, with the relation between the velocity field $\mathbf{v}$ and the derivatives (with an ad hoc notion of subdifferential in the Wasserstein space) of the functional F. This automatically gives the desired result as a part of a larger theory.

We will discuss six classes of PDEs in this section: the *porous media equation*; the *Keller-Segel equation*; more general *aggregation models*; more involved equations with *diffusion, advection, and aggregation*; a model for *crowd motion* with density constraints; and the *flow of the squared distance* $\mathrm{S}W_2^2$.

Porous media equation

This equation models the diffusion of a substance into a material whose properties are different than the void and which slows down the diffusion. If one considers the case of particles which are advected by a potential and subject to this kind of diffusion, the PDE reads

$$\partial_t \varrho - \Delta(\varrho^m) - \nabla \cdot (\varrho \nabla V) = 0,$$

for an exponent $m > 1$. One can formally check that this is the equation of the gradient flow of the energy

$$F(\varrho) = \frac{1}{m-1} \int \varrho^m(x)\,\mathrm{d}x + \int V(x)\varrho(x)\,\mathrm{d}x$$

(set to $+\infty$ for $\varrho \notin L^m$). Indeed, the first variation of the first part of the functional is $\frac{m}{m-1}\varrho^{m-1}$, and $\varrho\nabla\left(\frac{m}{m-1}\varrho^{m-1}\right) = m\varrho\cdot\varrho^{m-2}\nabla\varrho = \nabla(\varrho^m)$.

Note that, in the discrete step $\min_\varrho F(\varrho) + \frac{W_2^2(\varrho,\varrho_0)}{2\tau}$, the solution ϱ satisfies

$$\begin{cases} \frac{m}{m-1}\varrho^{m-1} + V + \frac{\varphi}{\tau} = C & \varrho - \text{a.e.} \\ \frac{m}{m-1}\varrho^{m-1} + V + \frac{\varphi}{\tau} \geq C & \text{on } \{\varrho = 0\}. \end{cases}$$

This allows to express $\varrho^{m-1} = \frac{m-1}{m}(C - V - \varphi/\tau)_+$ (see Section 7.2.3). This implies that ϱ is compactly supported if ϱ_0 is compactly supported, as soon as V has some growth conditions (see **Ex**(52)). This fact contradicts the usual infinite propagation speed that one finds in linear diffusion models (Heat and Fokker-Planck equation).

The above analysis works in the case $m > 1$; the fact that the usual Fokker-Planck equation can be obtained for $m \to 1$ can be seen in the following way: nothing changes if we define F via $F(\varrho) = \frac{1}{m-1}\int(\varrho^m(x) - \varrho(x))\,\mathrm{d}x + \int V(x)\varrho(x)\,\mathrm{d}x$, since the mass $\int \varrho(x)\,\mathrm{d}x = 1$ is a given constant. Yet, then it is easy to guess the limit, since

$$\lim_{m\to 1} \frac{\varrho^m - \varrho}{m-1} = \varrho \log \varrho,$$

which provides the entropy that we already used in Section 8.3.

It is also interesting to consider the case $m < 1$: the function $\varrho^m - \varrho$ is no longer convex, but it is concave and the negative coefficient $1/(m-1)$ makes it a convex function. Unfortunately, it is not superlinear at infinity, which makes it more difficult to handle. But for $m \geq 1 - 1/d$, the functional F is still displacement convex. The PDE that we get as a gradient flow is called *fast diffusion* equation, and it has different (and opposite) properties in terms of diffusion rate than the porous media one.

From a technical point of view, proving compactness of the minimizing movement scheme for these equations is not very easy, since one needs to pass to the limit the nonlinear term $\Delta(\varrho^m)$, which means proving strong convergence on ϱ instead of weak convergence. The main ingredient is a sort of H^1 bound in space, which comes from the fact that we have

$$\int_0^T \int_\Omega |\nabla(\varrho^{m-1})|^2\,\mathrm{d}x\mathrm{d}t \approx \int_0^T \int_\Omega \frac{|\nabla\varphi|^2}{\tau^2}\,\mathrm{d}x\mathrm{d}t = \int_0^T |\varrho'|(t)^2\mathrm{d}t \leq C \tag{8.15}$$

(but one has to deal with the fact that this is not a full H^1 bound, and the behavior in time has to be controlled).

Keller-Segel

An interesting model in mathematical biology (see [201, 202] for the original modeling) is the following: a population ϱ of bacteria evolves in time, following diffusion and advection by a potential. The potential is given by the concentration u of a chemoattractant nutrient substance, produced by the bacteria themselves. This kind of phenomenon is also known under the name of *chemotaxis*. More precisely, bacteria move (with diffusion) in the direction where they find more nutrient, i.e., in the direction of ∇u, where the distribution of u depends on their density ϱ. The easiest model uses linear diffusion and supposes that the distribution of u is related to ϱ by the condition $-\Delta u = \varrho$, with Dirichlet boundary conditions $u = 0$ on $\partial\Omega$. This gives the system

$$\begin{cases} \partial_t \varrho + \alpha \nabla \cdot (\varrho \nabla u) - \Delta \varrho = 0, \\ -\Delta u = \varrho, \\ u = 0 \text{ on } \partial\Omega,\ \varrho(0,\cdot) = \varrho_0,\ \partial_{\mathbf{n}}\varrho - \alpha\varrho\partial_{\mathbf{n}} u = 0 \text{ on } \partial\Omega. \end{cases}$$

The parameter α stands for the attraction intensity of bacteria toward the chemo-attractant. By scaling, instead of using probability measures $\varrho \in \mathscr{P}(\Omega)$, one can set $\alpha = 1$ and play on the mass of ϱ (indeed, the nonlinearity is only in the term $\varrho\nabla u$, which is quadratic in ϱ).

Alternative equations can be considered for u, such as $-\Delta u + u = \varrho$ with Neumann boundary conditions. On the contrary, the boundary conditions on ϱ must be of Neumann type, to guarantee conservation of the mass (see next Section 8.4.3). This system can also be set in the whole space, with suitable decay conditions at infinity. Note also that often the PDE condition defining u as the solution of a Poisson equation is replaced, when $\Omega = \mathbb{R}^2$, by the explicit formula

$$u(x) = -\frac{1}{2\pi} \int_{\mathbb{R}^2} \log(|x-y|)\varrho(y)\, \mathrm{d}y. \tag{8.16}$$

There is some confusion in higher dimension, as the very same formula does not hold for the Poisson equation (the logarithmic kernel should indeed be replaced by the corresponding Green function), and there are two alternatives: either keep the fact that u solves $-\Delta u = \varrho$ or the fact that it derives from ϱ through (8.16). We prefer keeping the PDE definition, which sounds more justified.

One can see that this equation is the gradient flow of the functional

$$F(\varrho) = \int_\Omega \varrho \log \varrho - \frac{1}{2}\int_\Omega |\nabla u_\varrho|^2, \quad \text{where } u_\varrho \in H^1_0(\Omega) \text{ solves } -\Delta u_\varrho = \varrho.$$

Indeed, the only nonstandard computation is that of the first variation of the Dirichlet term $-\frac{1}{2}\int |\nabla u_\varrho|^2$. Suppose $\varrho_\varepsilon = \varrho + \varepsilon\chi$ and set $u_{\varrho+\varepsilon\chi} = u_\varrho + \varepsilon u_\chi$. Then

$$\frac{d}{d\varepsilon}\left(-\frac{1}{2}\int |\nabla u_{\varrho+\varepsilon\chi}|^2\right)_{|\varepsilon=0} = -\int \nabla u_\varrho \cdot \nabla u_\chi = \int u_\varrho \Delta u_\chi = -\int u_\varrho \chi.$$

It is interesting to note that this Dirichlet term is indeed (up to the coefficient $-1/2$) the square of the H^{-1} norm of ϱ, since $||u||_{H^1_0} = ||\nabla u||_{L^2} = ||\varrho||_{H^{-1}}$. We will call it the H^{-1} term.

It is also possible to replace linear diffusion with nonlinear diffusion of porous media type, replacing the entropy $\int \varrho \log \varrho$ with a power-like energy $\int \varrho^m$.

Note that the variational problem $\min F(\varrho) + \frac{W_2^2(\varrho,\varrho_0)}{2\tau}$ requires some assumption to admit the existence of minimizers, as unfortunately the Dirichlet term has the wrong sign. In particular, it would be possible that the infimum is $-\infty$ or that the energy is not l.s.c. (the H^1 part being u.s.c. because of the negative sign).

When we use nonlinear diffusion with $m > 2$, the existence of a solution is quite easy (see **Ex**(53)). Sophisticated functional inequalities allow to handle smaller exponents, and even the linear diffusion case in dimension 2, provided $\alpha \leq 8\pi$. We refer to [57] and to the references therein for details on the analysis of this equation. Some of the technical difficulties are similar to those of the porous media equation, when passing to the limit nonlinear terms. In [57], the H^{-1} term is treated in terms of its logarithmic kernel, and ad hoc variables symmetrization tricks are used. Note however that the nonlinear diffusion case is easier, as L^m bounds on ϱ translate into $W^{2,m}$ bounds on u and hence strong compactness for ∇u.

We also remark that the above model, coupling a parabolic equation on ϱ and an elliptic one on u, implicitly assumes that the configuration of the chemoattractant instantaneously follows that of ϱ. More sophisticated models can be expressed in terms of the so-called *parabolic-parabolic Keller-Segel* equation, in the form

$$\begin{cases} \partial_t \varrho + \alpha \nabla \cdot (\varrho \nabla u) - \Delta \varrho = 0, \\ \partial_t u - \Delta u = \varrho, \\ u = 0 \text{ on } \partial\Omega,\ \varrho(0,\cdot) = \varrho_0,\ \varrho(\partial_{\mathbf{n}}\varrho - \partial_{\mathbf{n}} u) = 0 \text{ on } \partial\Omega, \end{cases}$$

or other variants with different boundary conditions. This equation can also be studied as a gradient flow in two variables, using distance W_2 on ϱ and L^2 on u; see [58].

Aggregation models

Consider a more general case where the movement is advected by a potential determined by the superposition of many potentials, each created by one particle. For instance, given a function $W : \mathbb{R}^d \to \mathbb{R}$, the particle located at x produces a potential $W(\cdot - x)$ and, globally, the potential is given by $V(y) = \int W(y - x)\mathrm{d}\varrho(x)$, i.e., $V = W * \varrho$. The equation, if every particle follows $-\nabla V$, is

$$\partial_t \varrho - \nabla \cdot (\varrho\, ((\nabla W) * \varrho)) = 0,$$

where we used $\nabla(W * \varrho) = (\nabla W) * \varrho$. If W is even (i.e., the interaction between x and y is the same as between y and x), then this is the gradient flow of the functional

$$F(\varrho) = \frac{1}{2} \iint W(x - y)\mathrm{d}\varrho(x)\mathrm{d}\varrho(y).$$

When W is convex, for instance, in the quadratic case $\iint |x - y|^2 \mathrm{d}\varrho(x)\mathrm{d}\varrho(y)$, this gives rise to a general aggregation behavior of the particle, and as $t \to \infty$ one expects $\varrho_t \rightharpoonup \delta_{x_0}$ (the point x_0 depending on the initial datum ϱ_0: in the quadratic

example above, it is the barycenter of ϱ_0; see **Ex**(51)). If W is not smooth enough, the aggregation into a unique point can also occur in finite time; see [118].

Obviously, more general interaction terms can be considered, of the form $W(x, y)$. As we saw in Chapter 7, some other energies (as the H^{-1} term in Keller-Segel) can also be expressed in this way, using their Green function (but in general they lose convexity of the kernel). Note that these equations are both nonlinear (the term in the divergence is quadratic in ϱ) and nonlocal. It is rare to see these nonlocal aggregation terms alone in the equation, as they are often coupled with diffusion or other terms. This is why we do not provide specific references except [118]. We also note that from the technical point of view, this kind of nonlinearity is much more compact than the previous ones, since the convolution operator transforms weak convergence into strong one, provided W is regular enough (the difficulties with the kernel in Keller-Segel exactly come from its singularity).

Diffusion, advection, and aggregation: the full picture

Very often, the aggregation energy of the previous paragraph is considered together with an internal energy and a confining potential energy, which brings to the functional

$$F(\varrho) = \int f(\varrho(x))\,\mathrm{d}x + \int V(x)\,\mathrm{d}\varrho(x) + \frac{1}{2}\iint W(x-y)\mathrm{d}\varrho(x)\mathrm{d}\varrho(y).$$

This gives the equation

$$\partial_t \varrho - \nabla\cdot\left(\varrho\left[\nabla(f'(\varrho)) + \nabla V + (\nabla W) * \varrho\right]\right) = 0.$$

Among the mathematical interest for this family of equations, we stress that they are those where more results (in terms of stability and convergence to equilibrium) can be proven, due to the fact that conditions to guarantee that F is displacement convex are well known (Section 7.3). See in particular [116, 117] for physical considerations and convergence results on this equation.

Crowd motion

The theory of Wasserstein gradient flows has interestingly been applied to the study of a continuous model of crowd movement under density constraints.

Let us explain the modeling, starting from the discrete case. Suppose that we have a population of particles such that each of them, if alone, would follow its own velocity u (which could a priori depend on time, position, on the particle itself, etc.). Yet, there is a constraint: these particles are modeled by rigid disks that cannot overlap; hence, the actual velocity cannot always be u, in particular if u tends to concentrate the masses. We will call $\mathbf{v}$ the actual velocity of each particle, and the main assumption of the model is that $\mathbf{v} = P_{\mathrm{adm}(q)}(u)$, where q is the particle

configuration, $\mathrm{adm}(q)$ is the set of velocities that do not induce (for an infinitesimal time) overlapping starting from the configuration q, and $P_{\mathrm{adm}(q)}$ is the projection on this set.

The simplest example is the one where every particle is a disk with the same radius R and center located at q_i. In this case we define the admissible set of configurations K through

$$K := \{q = (q_i)_i \in \Omega^N \ : \ |q_i - q_j| \geq 2R \ \text{ for all } i \neq j\}.$$

In this way the set of admissible velocities is easily seen to be

$$\mathrm{adm}(q) = \{\mathbf{v} = (\mathbf{v}_i)_i \ : \ (\mathbf{v}_i - \mathbf{v}_j) \cdot (q_i - q_j) \geq 0 \ \text{ for all } (i,j) \text{ with } |q_i - q_j| = 2R\}.$$

The evolution equation which has to be solved to follow the motion of q is then

$$q'(t) = P_{\mathrm{adm}(q(t))} u(t) \tag{8.17}$$

(with $q(0)$ given). Equation (8.17), not easy from a mathematical point of view, was studied by Maury and Venel in [223, 224].

We are now interested in the simplest continuous counterpart of this microscopic model (without pretending that it is any kind of homogenized limit of the discrete case, but only an easy reformulation in a density setting). In this case the particle population will be described by a probability density $\varrho \in \mathscr{P}(\Omega)$, the constraint becomes a density constraint $\varrho \leq 1$ (we define the set $K = \{\varrho \in \mathscr{P}(\Omega) \ : \ \varrho \leq 1\}$), the set of admissible velocities will be described by the sign of the divergence on the saturated region $\{\varrho = 1\}$: $\mathrm{adm}(\varrho) = \left\{\mathbf{v} : \Omega \to \mathbb{R}^d \ : \ \nabla \cdot \mathbf{v} \geq 0 \ \text{ on } \{\varrho = 1\}\right\}$, and we will consider a projection P, which will be either the projection in $L^2(\mathscr{L}^d)$ or in $L^2(\varrho)$ (this will turn out to be the same, since the only relevant zone is $\{\varrho = 1\}$). Finally, we solve the equation

$$\partial_t \varrho_t + \nabla \cdot \left(\varrho_t \left(P_{\mathrm{adm}(\varrho_t)} u_t\right)\right) = 0. \tag{8.18}$$

The main difficulty is the fact that the vector field $\mathbf{v} = P_{\mathrm{adm}(\varrho_t)} u_t$ is neither regular (since it is obtained as an L^2 projection and may only be expected to be L^2 a priori), nor it depends regularly on ϱ (it is very sensitive to small changes in the values of ϱ: for instance, passing from a density 1 to a density $1 - \varepsilon$ completely modifies the saturated zone and hence the admissible set of velocities and the projection onto it).

In [225] these difficulties have been overtaken in the case $u = -\nabla D$ (where $D : \Omega \to \mathbb{R}$ is a given Lipschitz function), and the existence of a solution (with numerical simulations) is proven via a gradient flow method. Indeed, (8.18) turns out to be the gradient flow in $\mathbb{W}_2$ of the energy

$$F(\varrho) = \begin{cases} \int D \, \mathrm{d}\varrho & \text{if } \varrho \in K; \\ +\infty & \text{if } \varrho \notin K. \end{cases}$$

We do not enter into the details of the study of this equation, but we just make a little bit more precise the definitions above. Actually, instead of considering the divergence of vector fields which are only supposed to be L^2, it is more convenient to give a better description of $\mathrm{adm}(\varrho)$ by duality :

$$\mathrm{adm}(\varrho) = \left\{ \mathbf{v} \in L^2(\varrho) \ : \ \int \mathbf{v} \cdot \nabla p \leq 0 \quad \forall p \in H^1(\Omega) \ : \ p \geq 0, \, p(1-\varrho) = 0 \right\}.$$

In this way we characterize $\mathbf{v} = P_{\mathrm{adm}(\varrho)}(u)$ through

$$u = \mathbf{v} + \nabla p, \quad \mathbf{v} \in \mathrm{adm}(\varrho), \quad \int \mathbf{v} \cdot \nabla p = 0,$$
$$p \in \mathrm{press}(\varrho) := \{ p \in H^1(\Omega), \, p \geq 0, \, p(1-\varrho) = 0 \},$$

where $\mathrm{press}(\varrho)$ is the space of functions p used as test functions in the dual definition of $\mathrm{adm}(\varrho)$, which play the role of a pressure affecting the movement. The two cones $\nabla\mathrm{press}(\varrho)$ (defined as the set of gradients of elements of $\mathrm{press}(\varrho)$) and $\mathrm{adm}(\varrho)$ are in duality for the L^2 scalar product (i.e., one is defined as the set of vectors which make a negative scalar product with all the elements of the other). This allows for an orthogonal decomposition $u_t = \mathbf{v}_t + \nabla p_t$ and gives the alternative expression of Equation (8.18), i.e.,

$$\begin{cases} \partial_t \varrho_t + \nabla \cdot \big(\varrho_t (u_t - \nabla p_t) \big) = 0, \\ 0 \leq \varrho \leq 1, \, p \geq 0, \, p(1-\varrho) = 0. \end{cases} \tag{8.19}$$

More details can be found in [225, 226, 265]. In particular, in [225] it is explained how to handle the nonlinearities when passing to the limit. Two sources of nonlinearity are observed: the term $\varrho \nabla p$ is easy to consider, since it is actually equal to ∇p (from $p = 0$ on $\{\varrho \neq 1\}$); on the other hand, we need to deal with the equality $p(1-\varrho) = 0$ and pass it to the limit. This is done by obtaining strong compactness on p, from a bound on $\int_0^T \int_\Omega |\nabla p|^2$, obtained similarly to (8.15).

Sliced Wasserstein distance

We already mentioned in Section 2.5.2 the idea by M. Bernot which consisted in moving every particle following the vector field $\mathbf{v}$ obtained in the following way: given two measures $\varrho, \nu \in \mathscr{P}(\mathbb{R}^d)$, we project them onto any 1D direction $e \in \mathbb{S}^{d-1}$ via the map $\pi_e : \mathbb{R}^d \to \mathbb{R}$ given by $\pi_e(x) = x \cdot e$; call $\mathrm{T}_e : \mathbb{R} \to \mathbb{R}$ the monotone optimal transport between the two image measures $(\pi_e)_\# \varrho$ and $(\pi_e)_\# \nu$. Then we define $\mathbf{v}_e(x) := (\mathrm{T}_e(\pi_e(x)) - \pi_e(x))e$ and $\mathbf{v}(x) = \fint_{S^{d-1}} \mathbf{v}_e(x) \, \mathrm{d}\mathscr{H}^{d-1}(e)$, where $\mathscr{H}^{d-1}$ is the uniform measure on the sphere.

For numerical approaches, Bernot proposed to iterate a construction with a time step $\tau > 0$, but a natural continuous counterpart exists: simply define, for every

absolutely continuous measure $\varrho \in \mathscr{P}(\mathbb{R}^d)$, the vector field $\mathbf{v}_{(\varrho)}$ defined above (absolute continuity is just required to avoid atoms in the projections). Then, we solve the equation

$$\partial_t \varrho_t + \nabla \cdot (\varrho_t \mathbf{v}_{(\varrho_t)}) = 0.$$

It happens that this equation has a gradient flow structure: it is indeed the gradient flow in $\mathbb{W}_2$ (i.e., for the distance W_2) of the functional

$$F(\varrho) := \mathrm{S} W_2^2(\varrho, \nu).$$

Existence and estimates on the solution of this equation are proven in [65], and the nonlinearity of $\mathbf{v}_{(\varrho)}$ is quite easy to deal with. On the other hand, many, natural and useful, questions are still open: Is it true that $\varrho_t \rightharpoonup \nu$ as $t \to \infty$? Can we define (at least under regularity assumptions on the initial data) the flow of the vector field $\mathbf{v}_{(\varrho_t)}$, and what is the limit of this flow as $t \to \infty$? The idea is that it should be a transport map between ϱ_0 and ν and, if not the optimal transport map, at least a "good" one. This has been tested in the discrete case (where both ϱ_0 and ν are a finite number of equal Dirac masses) with satisfactory numerical results, which we presented briefly in Section 2.5.2.

8.4.3 Dirichlet boundary conditions

For sure, the attentive reader has already noted that all the equations that have been identified as gradient flows for the distance W_2 on a bounded domain Ω are always accompanied by Neumann boundary conditions. This should not be surprising. Wasserstein distances express the movement of masses when passing from a configuration to another, and the equation represents the conservation of mass. It means that we are describing the movement of a collection ϱ of particles, bound to stay inside a given domain Ω, and selecting their individual velocity $\mathbf{v}$ in a way that is linked to the global value of a certain functional $F(\varrho)$. It is natural in this case to have boundary conditions which write down the fact that particles do not exit the domain. The pointwise value of the density ϱ at the points of the boundary $\partial\Omega$ is not particularly relevant in this analysis. Note that "do not exit" does not mean "those on the boundary stay on the boundary," which is what happens when solutions are smooth and the velocity field $\mathbf{v}$ satisfies $\mathbf{v} \cdot \mathbf{n} = 0$. Yet, the correct no-flux condition here is rather $\varrho \mathbf{v} \cdot \mathbf{n} = 0$ a.e., which means that particles could enter from $\partial\Omega$ into the interior $\mathring{\Omega}$, but immediately after it happens, there will be (locally) no mass on the boundary, and the condition is not violated, hence. On the contrary, should some mass go from $\mathring{\Omega}$ to outside Ω, then we would have a violation of the no-flux condition, since there would be (intuitively) some mass $\varrho > 0$ on the boundary with velocity directed outward.

Anyway, we see that Dirichlet conditions do not find their translation into W_2 gradient flows!

To cope with Dirichlet boundary conditions, Figalli and Gigli defined in [165] a sort of modified Wasserstein distance, with a special role played by the boundary $\partial\Omega$, in order to study the heat equation $\partial_t\varrho = \Delta\varrho$ with Dirichlet b.c. $\varrho = 1$ on $\partial\Omega$.

Their definition is as follows: given two finite positive measures $\mu, \nu \in \mathscr{M}_+(\mathring{\Omega})$ (not necessarily probabilities, not necessarily with the same mass), we define

$$\Pi b(\mu,\nu) = \{\gamma \in \mathscr{M}_+(\overline{\Omega}\times\overline{\Omega}) \ : \ (\pi_x)_\#\gamma \llcorner \mathring{\Omega} = \mu,\ (\pi_y)_\#\gamma \llcorner \mathring{\Omega} = \nu\}.$$

Then, we define

$$Wb_2(\mu,\nu) := \sqrt{\inf\left\{\int_{\overline{\Omega}\times\overline{\Omega}} |x-y|^2\,\mathrm{d}\gamma,\ \gamma\in\Pi b(\mu,\nu)\right\}}.$$

The index b stands for the special role played by the boundary. Informally, this means that the transport from μ to ν may be done usually (with a part of γ concentrated on $\mathring{\Omega}\times\mathring{\Omega}$), by moving some mass from μ to $\partial\Omega$ (using $\gamma\llcorner(\mathring{\Omega}\times\partial\Omega)$), then moving from one point of the boundary to another point of the boundary (this should be done by using $\gamma\llcorner(\partial\Omega\times\partial\Omega)$, but since this part of γ does not appear in the constraints, then we can forget about it, and the transport is free on $\partial\Omega$), and finally moving from $\partial\Omega$ to ν (using $\gamma\llcorner(\partial\Omega\times\mathring{\Omega})$).

In [165] the authors prove that Wb_2 is a distance, and that the space $\mathscr{M}_+(\mathring{\Omega})$ is always a geodesic space, independently of convexity or connectedness properties of Ω (differently from what happens with Ω, since here the transport is allowed to "teleport" from one part of the boundary to another, either to pass from one connected component to another or to follow a shorter path going out of Ω), and they study the gradient flow, for this distance, of the functional $F(\varrho) = \int(\varrho\log\varrho - \varrho)\,\mathrm{d}x$. Note that in the usual study of the entropy on $\mathscr{P}(\Omega)$, one can decide to forget the term $-\int\varrho$, which is anyway a constant because the total mass is fixed. Here this term becomes important (if the function $f(t) = t\log t - t$ is usually preferred to $t\log t$, it is because its derivative is simpler, $f'(t) = \log t$, without changing its main properties).

With this choice of the functional and of the distance, the gradient flow that Figalli and Gigli obtain is the heat equation with the particular boundary condition $\varrho = 1$ on $\partial\Omega$. One could wonder where the constant 1 comes from and a reasonable explanation is the following: if the transport on the boundary is free of charge, then automatically the solution selects the value which is the most performant for the functional, i.e., the constant t which minimizes $f(t) = t\log t - t$. In this way, changing the linear part and using $F(\varrho) = \int(\varrho\log\varrho - c\varrho)\,\mathrm{d}x$ could change the constant on the boundary, but the constant 0 is forbidden for the moment. It would be interesting to see how far one could go with this approach and which Dirichlet

conditions and which equations could be studied in this way, but this does not seem to be done at the moment.

Moreover, the authors explain that, due to the lack of geodesic convexity of the entropy w.r.t. Wb_2, the standard abstract theory of gradient flows is not able to provide uniqueness results (the lack of convexity is due in some sense to the possible concentration of mass on the boundary, in a way similar to what happened in [225] when dealing with the door on $\partial\Omega$). On the other hand, standard Hilbertian results on the heat equation can provide uniqueness for this equation, as the authors underlined in [165].

We observe that these kinds of distances with free transport on the boundary were already present in [70, 71], but in the case of the Wasserstein distance W_1, and the analysis in those papers was not made for applications to gradient flows, which are less natural to study with $p = 1$. We can also point out a nice duality formula:

$$Wb_1(\mu,\nu) := \min\left\{\int |x-y|\,\mathrm{d}\gamma \;:\; \gamma \in \Pi b(\mu,\nu)\right\}$$
$$= \sup\left\{\int u\,\mathrm{d}(\mu-\nu) \;:\; u \in \mathrm{Lip}_1(\Omega),\ u = 0 \text{ on } \partial\Omega\right\}.$$

Note that in general we also have the (easy) equality

$$Wb_p(\mu,\nu) = \inf\left\{W_p(\mu+\chi^+,\nu+\chi^-) \;:\; \mathrm{spt}(\chi^\pm) \subset \partial\Omega\right\},$$

where the Wasserstein distance on the right is defined, by abuse of notation, for every pair of positive measures with the same mass, not only for probabilities. In the special case $p = 1$ and $\mu(\mathring{\Omega}) = \nu(\mathring{\Omega})$, we also obtain

$$Wb_1(\mu,\nu) = \mathscr{T}_c(\mu,\nu), \qquad \text{for } c(x,y) = \min\{|x-y|, d(x,\partial\Omega) + d(y,\partial\Omega)\}.$$

The cost c is a pseudo-distance on Ω where moving on the boundary is free. This kind of distance has also been used in [95] (inspired by [70]) to model free transport costs on other lower dimensional sets and not only the boundary (with the goal to model, for instance, transportation networks and optimize their shape). It is interesting to see the same kind of ideas appear for so different goals.

8.4.4 Evolution PDEs: not only gradient flows

We finish this chapter (and this book) with some other applications and connections of optimal transport with evolution PDEs (note that we already saw connections with static PDEs: the Monge-Ampère equation, in Section 1.7.6, and the Monge-Kantorovich system, in Section 4.2.2).

Incompressible Euler equation

Let us consider an incompressible and homogeneous fluid moving inside a smooth domain Ω. The Euler equation for incompressible fluids describes the evolution of its velocity field $\mathbf{v}$ in terms of the pressure field p:

$$\partial_t \mathbf{v}_t + (\mathbf{v}_t \cdot \nabla)\mathbf{v}_t + \nabla p_t = 0, \qquad \nabla \cdot \mathbf{v}_t = 0.$$

To clarify the notation $(\mathbf{v} \cdot \nabla)\mathbf{v}$, we could write componentwise as $\partial_t \mathbf{v}^i + \mathbf{v}^j \mathbf{v}^i_j = -p_i$ (we ignore here the subscripts t for time, and use superscripts for components and subscripts for derivatives and Einstein's convention for summation over repeated indices). Note that the term $\partial_t \mathbf{v}_t + (\mathbf{v}_t \cdot \nabla)\mathbf{v}_t$ is nothing but the acceleration in Lagrangian coordinates. Indeed, the trajectory of a particle initially at position x is obtained by following the vector field $\mathbf{v}_t$, i.e., we have, as usual, the flow $y_x'(t) = \mathbf{v}_t(y_x(t))$ with $y_x(0) = x$. If we denote by Y_t the map $x \mapsto y_x(t)$, for each time t the map Y_t is a measure-preserving diffeomorphism of Ω, and for each x we have

$$y_x''(t) = -\nabla p_t(y_x(t)).$$

In [22], Arnold interpreted the equation above as a geodesic equation on the space $\mathrm{SDiff}(\Omega)$ of measure-preserving diffeomorphism of Ω, viewed as an infinite-dimensional manifold with the metric inherited from the embedding in $L^2(\Omega)$ (the tangent space corresponding to the divergence-free vector fields). According to this interpretation, one can find solutions by minimizing

$$\int_0^T \int_D \frac{1}{2} |\dot{g}(t,x)|^2 \, \mathrm{d}x \, \mathrm{d}t \tag{8.20}$$

among all paths $g(t,\cdot) : [0,T] \to \mathrm{SDiff}(\Omega)$ with $g(0,\cdot) = g_0$ and $g(T,\cdot) = g_T$ prescribed (typically, by right invariance, g_0 is taken as the identity map). In this way, the pressure field arises as a Lagrange multiplier from the incompressibility constraint.

Shnirelman proved in [284, 285] that when $d \geq 3$ the infimum is not attained in general and that when $d = 2$ there exists $h \in \mathrm{SDiff}(D)$ which cannot be connected to the identity by a path with finite action. These "negative" results motivate the study of relaxed versions of Arnold's problem.

The first relaxed version of Arnold's minimization problem was introduced by Brenier in [83] and it is a sort of transport problem. Its relation to Arnold's problem is the same that Kantorovich problem has w.r.t. Monge. Brenier considered probability measures Q on the space $\mathrm{AC}(\Omega)$ of absolutely continuous paths (here below simply denoted $\mathscr{C}$) $\omega : [0,T] \to \Omega$ and looked at the variational problem.

Problem 8.11. Solve

$$\text{(EP)} \quad \min \left\{ \int_{\mathscr{C}} K_2(\omega)\, \mathrm{d}Q(\omega) \ : \ Q \in \mathscr{P}(\mathscr{C}), (e_0, e_T)_\# Q = \gamma, \ (e_t)_\# Q = \mathscr{L}_\Omega \ \forall\, t \right\}, \tag{8.21}$$

where $K_2(\omega) := \int_0^T \frac{1}{2}|\omega'(t)|^2\, \mathrm{d}t$ is the kinetic energy of a curve ω and the transport plan $\gamma \in \mathscr{P}(\mathscr{L}_\Omega, \mathscr{L}_\Omega)$ is fixed, typically of the form $\gamma = (\mathrm{id}, h)_\# \mathscr{L}_\Omega$.

The existence of a minimizer Q is standard, provided that there exists at least a Q with finite action (see [83]), which is true when Ω is smooth enough.

We observe that any sufficiently regular path $g(t, \cdot) : [0, 1] \to \mathrm{SDiff}(\Omega)$ induces a flow (a traffic plan, in our language; see Section 4.2) Q, but the converse is far from being true: particles starting from different points are allowed to cross at a later time, and particles starting from the same point are allowed to split, which is of course forbidden by classical flows. These solutions are called nondeterministic, and [83] proposes indeed an example in the easy case $\mathrm{T} = \pi$, $\Omega = B(0, 1)$ and $h(x) = -x$ where they can be optimal.

On the other hand, in [83], a consistency result was proved: smooth solutions to the Euler equation are optimal even in the larger class of admissible Q provided the pressure field p satisfies

$$\sup_{t\in[0,T]} \sup_{x\in\Omega} D^2 p(t, x) \leq \frac{\pi^2}{T^2} \mathrm{I} \tag{8.22}$$

and are the unique ones if the above inequality is strict.

Uniqueness of the solution of the minimal action problem is an interesting matter: in [51] the situation is analyzed in 1D (when a general uniqueness result is proven assuming the pressure to be smooth) and 2D (where, on the contrary, [51] extends and characterizes the nondeterministic solutions of [83] finding a very rich family of nondeterministic optimizers).

Optimal transport appears in this framework in two aspects. First, the action-minimization problem by Brenier is indeed a multi-marginal Kantorovich problem, with marginal constraints for each time t (see Section 1.7.4). Second, when we discretize in time this problem, natural strategies for numerics involve the following procedure: suppose we have determined the positions of the particles at time t_{i-1} and t_{i+1} (with a uniform partition $t_i = T\frac{i}{n}$) and that they are represented by two functions X_{i-1} and X_{i+1}, both with image measure equal to $\mathscr{L}_\Omega$; without the incompressibility constraint, the intermediate value $X_i = (X_{i-1} + X_{i+1})/2$ would be optimal in order to minimize the kinetic energy, but its image measure is not, in general, $\mathscr{L}_\Omega$. Hence, we need to project it onto the set of maps X with prescribed image measure, which is a classical transport problem, in the polar factorization formulation (see Section 1.7.2).

Coupling HJ and continuity equations, mean field games

We saw in Remark 6.2 in Chapter 6 that the Benamou-Brenier method provides a solution of the following PDE system

$$\begin{cases} \partial_t \varrho + \nabla \cdot (\varrho \nabla \varphi) = 0 \\ \partial \varphi + \frac{1}{2}|\nabla \varphi|^2 = 0, \ \varrho_0 = \mu, \varrho_1 = \nu, \end{cases}$$

where the second equation is satisfied ϱ-a.e., with inequality $\partial_t \varphi + \frac{1}{2}|\nabla \varphi|^2 \leq 0$ everywhere.

This kind of coupled systems with Hamilton-Jacobi and continuity equations also arises in other situations and in particular in the newborn theory of mean field games (MFG).

Let us give a quick and informal presentation of this theory. We will only consider the deterministic case (i.e., there will be no stochastic part in the evolution of the agents and we will use deterministic optimal control theory and not stochastic control). The reader may refer to [205, 206] and [105] to have a wider introduction to this theory.

Suppose that a population of agents may evolve in time, each agent following trajectories of the controlled equation

$$y'(t) = f(t, y(t), \alpha(t)), \quad t \in [0, T] \tag{8.23}$$

$\alpha : [0, T] \to \mathbb{R}^d$ being a control that every agent may choose. At each time t, the goal of each agent is to maximize the payoff

$$-\int_t^T \left(\frac{|\alpha(s)|^2}{2} + g(\varrho(s, y(s))) \right) ds + \Phi(y(T)), \tag{8.24}$$

where g is a given increasing function. This means that α is the effort that every agent makes to move in the desired direction, and he pays for it (for simplicity, we take a quadratic cost), that its position depends on α through Equation (8.23), and that he tries to optimize the final payoff Φ, but he also pays for the densities of the regions he passes by. In particular agents would like to avoid overcrowded areas. At this first step, the density $\varrho(t, x)$ is supposed to be a given function. Yet, the MFG problem is an equilibrium problem: given the initial density ϱ_0, find a time-dependent family of densities ϱ_t such that when every agent selects its optimal trajectory so as to optimize (8.24), the density realized by these optimal trajectories at time t is exactly ϱ_t.

One can study the optimal control problem given by (8.23) and (8.24) by means of its value function.

Box 8.3. ***Memo:*** *Optimal control and Hamilton-Jacobi*

Consider the optimization problem

$$\max\left\{\int_0^T L(t,y(t),\alpha(t))\mathrm{d}t+\Phi(y(T))\ :\ y'=f(t,y,\alpha),\ y(0)=x_0\right\},$$

where α is the control that we can use to affect the solution of the state equation on the trajectory x. Define the so-called *value function*

$$\varphi(t,x)=\sup\left\{\int_t^T L(s,y(s),\alpha(s))ds+\Phi(y(T))\ :\ y(t)=x,\ y'=f(s,y,\alpha)\right\}.$$

It is well known from optimal control theory (see, for instance, Section 10.3 in [158]) that φ satisfies the Hamilton-Jacobi-Bellmann equation (in the viscosity sense)

$$\partial_t\varphi(t,x)+H(t,x,\nabla\varphi(t,x))=0,\quad \varphi(T,x)=\Phi(x),$$

where the Hamiltonian H is defined through

$$H(t,x,\xi)=\sup_\alpha\ \xi\cdot f(t,x,\alpha)+L(t,x,\alpha). \tag{8.25}$$

Moreover, it is also well known that in the control problem, for every (t,x), the optimal choice of $\alpha(t)$ so as to solve the optimization problem starting from x at time t is the control α which maximizes in the definition of H given in (8.25) for $\xi=\nabla\varphi(t,x)$, i.e., which maximizes $\nabla\varphi(t,x)\cdot f(t,x,\alpha)+L(t,x,\alpha)$.

In the remarkable case where $f(t,x,\alpha)=\alpha$ and $L(t,x,\alpha)=-\frac{1}{2}|\alpha|^2-g(t,x)$, we get $H(t,x,\xi)=\frac{1}{2}|\xi|^2-g(t,x)$, thus $\partial_t\varphi+\frac{1}{2}|\nabla\varphi|^2=g$, and the optimal curves follow $y'=\nabla\varphi(t,y)$.

This gives a system of two coupled equations, since φ solves an HJB equation where ϱ appears in the Hamiltonian, and ϱ evolves according to a continuity equation $\partial_t\varrho+\nabla\cdot(\varrho\mathbf{v})=0$. To be more precise, we can give an explicit example, in the case $f(t,x,\alpha)=\alpha$. In such a case, the vector field to be put in the continuity equation is exactly $\nabla\varphi$. This gives the system

$$\begin{cases}\partial_t\varphi+\frac{|\nabla\varphi|^2}{2}-g(\varrho)=0,\\ \partial_t\varrho+\nabla\cdot(\varrho\nabla\varphi)=0,\\ \varphi(T,x)=\Phi(x),\quad \varrho(0,x)=\varrho_0(x),\end{cases} \tag{8.26}$$

where the first equation is satisfied in the viscosity sense and the second in the distributional sense.

In this case, it is also known that a solution of this system (i.e., an equilibrium in the mean field game) may be obtained by minimizing a suitable global functional (obviously, up to a change of sign, we could also express it as a maximization). Actually, one can solve

$$\min \quad \int_0^T \int_\Omega \left(\frac{1}{2} |\alpha(t,x)|^2 \varrho(t,x) + G(\varrho(t,x)) \right) \mathrm{d}x\mathrm{d}t - \int_\Omega \Phi(x)\varrho(T,x)\,\mathrm{d}x$$

among solutions (ϱ, α) of the continuity equation $\partial_t \varrho + \nabla \cdot (\varrho\alpha) = 0$ with initial datum $\varrho(0,x) = \varrho_0(x)$. When the function G is chosen as the antiderivative G of g (i.e., $G' = g$, and in particular G is convex), it happens that the minimizers of this global functional are equilibria in the sense explained above (the trajectories that the agents chose according to these densities are such that they exactly realize these densities). These functionals are a modification of the functional proposed by Benamou and Brenier [34] and studied in Chapters 5 and 6, and the minimization corresponds to finding a good curve in $\mathbb{W}_2$, which avoids over-congested measures. This kind of problems was also studied, without equilibrium issues, by Buttazzo, Jimenez, and Oudet in [96]. Also note the analogy with the congested problems of Section 4.4.1. The main difference here is the crucial role of time, while the movement in Section 4.4.1 was stationary.

Semi-geostrophic equation

One of the first applications of optimal transport has been its role in this equation modeling large-scale atmospheric behavior in a certain asymptotical regime. The main ingredients of the physical description are the Coriolis force, the gravity, and an incompressibility constraint. We suppose that the particles could have different weights, thus leading to different impact of gravity, but their number per unit of volume should be constant in time and space. This consists in imposing that the density (mass per unit of volume) is constant in Lagrangian coordinates, but not necessarily in time and space.

By ignoring, in a very shallow atmosphere (consider that the radius of the Earth is much larger than the width of the atmosphere), the vertical effects of the Coriolis force and using local coordinates where the vertical axis is orthogonal to the Earth surface, we can write an equation in Lagrangian formulation:

$$X''(t,a) + JX'(t,a) + \nabla p_t(X(t,a)) = f(a),$$

where p_t is the pressure at time t, a is a Lagrangian label, and the matrix $J \in \mathrm{M}3 \times 3$ is given by

$$J = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The force f stands for gravity, is vertical, and only depends on the label a. The pressure p_t appears in order to take care of the incompressibility constraint, which corresponds to $a \mapsto X(t,a)$ being measure preserving for every t. For details on this theory, we refer the reader to the book by M. Cullen [130].

In Eulerian coordinate the above equation is written

$$\begin{cases} \partial_t \mathbf{v}_t + (\mathbf{v}_t \cdot \nabla)\mathbf{v}_t + J\mathbf{v} + \nabla p_t = f_t, \\ \partial_t f_t + \mathbf{v}_t \cdot \nabla f_t = 0, \\ \nabla \cdot \mathbf{v}_t = 0, \end{cases}$$

where the second equation represents the fact that the force is transported by the flow itself (hence it "follows" the particles) and the third stands for incompressibility.

The time scale regime of the model is quasi-stationary, in the sense that we suppose that the configuration at time t is stable. This can be interpreted in several ways. A first one, with a variational flavor, is the following: at each time t, we have a distribution of particles, each with its velocity. Suppose that these particles want to move, each from a position x to x'. The new distribution of positions must be incompressible and, among these new redistributions, we need to impose that the original one minimizes the energy of the system, which is given by kinetic energy and potential energy. Consider that because of Coriolis force, when moving from x to x', the velocity moves from $\mathbf{v}(x)$ to $\mathbf{v}(x) + J(x' - x)$. Hence the energy to be minimized is of the form $\int \frac{1}{2}|J(x' - x)|^2 + \ell(x, x' - x)$, where $\ell(x, \cdot)$ is affine for every x (the gravity part is included in ℓ). Consider the pressure, which takes care of the incompressibility constraint. This means that minimizing such an energy under prescribed image measure (which is by the way an optimal transport problem with a cost equivalent to the quadratic one) is the same (see also the part on incompressible Euler) as minimizing $\int \frac{1}{2}|J(x' - x)|^2 + p(x') + \ell(x, x' - x)$ under no constraints. It is clear that each position is stable if and only if it minimizes $x' \mapsto \frac{1}{2}|J(x' - x)|^2 + p(x') + \ell(x, x' - x)$. This imposes first-order conditions, but also the fact that $x' \mapsto \frac{1}{2}|Jx'|^2 + p(x')$ is convex.

Another way of seeing the same fact is the following. Fix a time t_0 and a point x_0, where particles have a velocity $\mathbf{v}_0$. Suppose that some particles around x_0 have a small perturbation in their velocity $\mathbf{v} \approx \mathbf{v}_0$. We want to see what happens to these particles: do they stay close to x_0 or do they move far away? We freeze the pressure in time and approximate it around x_0 with its Taylor expansion at x_0. Consider a particle starting from x_0 with velocity $\mathbf{v}(s)$: here the variable s denotes time, but at a shorter time scale than that of the global evolution in t (which justifies the fact that we freeze $t = t_0$). The equation on $\mathbf{v}$, if we differentiate the acceleration law, is of the form $\mathbf{v}'' + J\mathbf{v}' + D^2 p(x_0)\mathbf{v} = 0$ (here we take a second-order Taylor expansion for p). By replacing $\mathbf{v}'$ with $-J\mathbf{v} + const$ (taking, for this lower-order condition, a first-order Taylor expansion for p), we get

$$\mathbf{v}'' + (D^2 p(x_0) - J^2)\mathbf{v} = const.$$

For stability, we want to avoid the existence of solutions of this equation which diverge exponentially. If only bounded solutions exist, this means that small perturbations in the velocity produce small displacements, and hence the profile at (t_0, x_0) is stable.

In both interpretations (the variational one above and the one related to linear stability), the condition that we find is that $D^2p(x_0) - J^2$ (which is a symmetric matrix) only has positive eigenvalues. This means

$$D^2p + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \geq 0$$

in the sense of positive-definite symmetric matrices. Equivalently, the function $x \mapsto P_t(x) := p_t(x) + \frac{1}{2}(|x_1|^2 + |x_2|^2)$ must be convex.

This convexity fact is known as the Cullen-Purser stability condition; see [130] or [131, 286].

We now come back to the acceleration law, we differentiate it as above, and replacing X'' using the equation, we get

$$X'''(t,a) + [\nabla_t p(X(t,a))]' - J^2X'(t,a) = J\nabla p_t(X(t,a)).$$

If we decide, on this time scale, to ignore the third derivative (yep!) and we set $Y_t(a) = \nabla P_t(X(t,a))$, we have

$$Y'(t,a) = J(Y(t,a) - X(t,a)),$$

and this is an evolution equation on Y where X is defined from Y in the following way: for each time t, $X(t,\cdot)$ is measure preserving, and Y is decomposed into the composition of the gradient of a convex function P and X. Here is where optimal transport, and in particular polar factorization (see Section 1.7.2) comes into play! We refer to [33] for the first applications of optimal transport in this field.

This equation is also studied in the so-called dual variables, i.e., Y and its image measure ϱ (which has nothing to do with the density of the fluid). They solve

$$\begin{cases} \partial_t\varrho_t + \nabla \cdot (\varrho_t J(y - \nabla P_t^*(y))) = 0, \\ \det(D^2P_t^*) = \varrho_t, \ P_t^* \text{ is convex.} \end{cases}$$

This equation has been mathematically studied in several papers, where recent regularity improvements on the Hessian of the Kantorovich potential (the $W^{2,1}$ result that we briefly addressed in Section 1.7.6) have allowed for rigorous results (see, for instance, [16, 17]).

Reconstruction of the early universe

This is another surprising domain where optimal transport for a quadratic cost played an important role. The problem considers the evolution of the density ϱ of the

matter in the universe, subject to the gravity field created by ϱ itself. In the so-called *co-moving coordinates* the equations read as follows:

$$\begin{cases} \partial_t \mathbf{v} + (\mathbf{v}\cdot\nabla)\mathbf{v} = -\frac{3}{2}\frac{\nabla\varphi+\mathbf{v}}{t}, \\ \partial_t \varrho + \nabla\cdot(\varrho\mathbf{v}) = 0, \\ \Delta\varphi = \frac{\varrho-1}{t}. \end{cases} \tag{8.27}$$

The first equation gives the acceleration of a particle in terms of the gradient of the potential, the second is conservation of mass, and the third is the relation of the potential with the density itself. To understand where the precise expressions of the first and third equations come from, the reader is invited to look at the very first pages of [45]. He will easily see that the main point is given by the change of variable imposed in the choice of the co-moving coordinates, i.e., coordinates which follow the expansion of the universe. In these coordinates we use $x' = a(t')x$ as a space variable, where x is the physical space variable and a is the expansion coefficient, which grows in time. We change the time scale from t to t', via $\mathrm{d}t = a(t')\mathrm{d}t'$. After the change of variables, we drop the superscripts in the new variables. The expansion factor a needs to satisfy precise equations, called Friedmann equations. Note that also the friction term in the first equation (i.e., the negative velocity at the r.h.s.) exactly comes from this change of variable.

One sees from these equations that the density ϱ at time 0 must be uniform. It corresponds to the distribution of matter in the very first instants of the universe. On the contrary, the density of the universe today is far from being uniform and is well approximated by a number of Dirac masses corresponding each to a galaxy cluster (i.e., each atom represents several galaxies, very close to each other). The question arises whether we can guess the evolution of the universe in the past by only knowing the current density ϱ_T (T being the final time, i.e., now, but we will set $T = 1$ for simplicity), i.e., without knowing the initial velocity. This question has been analyzed by U. Frisch and collaborators, involving among others Y. Brenier because of the connections this problem has with optimal transport. See [89] and [172]. By the way, Appendix A in [89] explains in detail the derivation of the above equations in co-moving coordinates presenting the set of choices and simplifications that have been made compared to the general framework in [45].

The connections of these equations and of the problem of the reconstruction from the only knowledge of ϱ_T with optimal transport theory are twofold. First, let us start from the so-called Zel'dovich approximation. This approximation consists in assuming that, for $t > 0$, all the particles move in straight lines with constant speed only depending on the initial data. Let us be more precise: in Lagrangian coordinates the first equation of (8.27) reads

$$\frac{2t}{3}X''(t,x_0) + X'(t,x_0) = -\nabla\varphi_t(X(t,x_0)).$$

Here the potential φ_t is required to satisfy $1 + t\Delta\varphi_t = \varrho_t$, where ϱ_t is the image measure of $X(t, \cdot)$. The approximation consists in ignoring the second derivative and freezing the r.h.s. at its value at $t = 0$, i.e., $X(t) = x_0 - t\nabla\varphi_0(x_0)$.

It seems a rude approximation, but it gives good results in practice, and it has an optimal transport justification, as pointed out by Y. Brenier. Indeed, choose as ϱ_t the geodesic in $\mathbb{W}_2$ between $\varrho_0 = 1$ and ϱ_1. If $-\varphi_t$ is the Kantorovich potential between ϱ_t and ϱ_0, we know that it solves the Monge-Ampère equation

$$\det(\mathrm{I} + tD^2\varphi_t) = \varrho_t.$$

Moreover, if T is the optimal transport from ϱ_0 to ϱ_1, the Lagrangian map $X(t, x_0) = (1 - t)x_0 + t\mathrm{T}(x_0)$ is the optimal transport from ϱ_0 to ϱ_t. This means that particles should arrive at time t with velocity equal to $-\nabla\varphi_t(X(t, x_0))$ and proves that $-\nabla\varphi_t(X(t, x_0)) = \mathrm{T}(x_0) - x_0$ does not depend on time. In particular, it is also equal to its initial value $-\nabla\varphi_0(x_0)$. In this way, we have constructed a Lagrangian map compatible with the Zel'dovich approximation (i.e., particles move with $X'' = 0$), with the only exception that it solves $\det(\mathrm{I} + tD^2\varphi_t) = \varrho_t$ instead of $1 + t\Delta\varphi_t = \varrho_t$. In this way, the Zel'dovich approximation provides approximate solution for the original equation but exact solutions for this new, modified equation. Note the use of a nonlinear equation to approximate a linear one (the linear one being the linearization of the nonlinear around $t = 0$), quite peculiar in this case.

From the above observations, the authors of [89] concentrated on the search for an optimal map T in the transport between the uniform density ϱ_0 and the current density ϱ_1 and obtained satisfactory reconstruction results. Numerically, they discretized also the source measure and used an auction algorithm.

On the other hand, [89] also shows (in Section 6 and Appendix D) that System (8.27) (without approximation or replacing the Poisson equation with the Monge-Ampère equation) can be obtained as the optimality condition of a variational problem, namely,

$$\min\left\{\int_0^1\int\left(\frac{t^{3/2}}{2}\varrho|\mathbf{v}|^2 + \frac{3t^{-1/2}}{4}|\nabla\varphi|^2\right) \;:\; \begin{cases}\partial_t\varrho + \nabla\cdot(\varrho\mathbf{v}) = 0,\\ \Delta\varphi = \varrho - 1,\\ \varrho_0 = 1, \varrho_1 = \text{given}\end{cases}\right\}. \tag{8.28}$$

This problem is a sort of modified Benamou-Brenier problem, with a time scale factor, and the addendum of a term $F(\varrho)$ which penalizes (with a time scale coefficient as well) the H^{-1} norm of $\varrho - 1$. Note that this term is the same which appears in chemotaxis gradient flows (see Section 8.4.2), but with the good sign, and also note the same formalism as the variational formulation of mean field games. Moreover, the important fact is that this minimization problem is strictly convex, which implies uniqueness of the solution of (8.27).

We leave the proof of the fact that the optimality conditions of the above minimization lead to (8.27) as an exercise, **Ex**(54).

We also stress that the existence of a minimizer is not straightforward. Indeed, the coefficient $t^{3/2}$ in front of the kinetic term does not allow to preserve the condition at $t = 0$ along with minimizing sequence, and the H^{-1} term is multiplied with a coefficient $t^{-1/2}$ which is integrable, hence not enough to compensate for it. We are in a situation analogous to the following 1D problem:

$$\min \left\{ \int_0^1 t^{1+\alpha} |x'(t)|^2 + t^{-1+\alpha} G(x(t)) \ : \ x(1) = x_1 \right\}$$

where G is a convex functional, with quadratic growth, minimal at x_0. We omit the constraint $x(0) = 0$ since this condition would not be stable under convergence. In particular, there are finite-energy curves $x(t)$ with $x(0) \neq x_0$ (compare to **Ex**(56)). On the other hand, one can prove that the minimizers of this problem must satisfy an extra condition at $t = 0$, because of optimality (like Neumann conditions, which are not meaningful here since they only give $0 = 0$ because of the coefficient $t^{3/2}$). It is possible to see, as in **Ex**(55), that this condition is $x(0) = x_0$.

The minimal action problem (8.28) has also been studied, in a simplified setting, by Loeper in [215]. We also mention a variant problem proposed by Brenier, which is meant to give a better insight in the formation of concentration [87]. Indeed, it is quite disappointing to think that in the optimal transport problem between ϱ_0 and ϱ_1, Dirac masses only appear at the final time (coherently with the analysis of the intermediate measures μ_t that we did in Section 4.3). This means that they only appear now, i.e. today ... what should be the conclusion if we studied again the same question tomorrow?

Finally, it is challenging to think that both the "static" version of this cosmological problem (i.e., the optimal transport between ϱ_0 and ϱ_T) and the "dynamical" one (8.28) still have many secrets to reveal, if treated with the new numerical methods and approaches (see Chapter 6, in particular for semi-discrete methods which are exactly suitable for $\varrho_0 = 1$ and atomic ϱ_T) which were not available at the time of [89].

Exercises

The following is a short list of possible exercises related to the subject of the book. The level of difficulty is strongly heterogeneous, and the most difficult exercises are indicated with one * or two **. The symbol ♠ is associated to those exercises which are more relevant for the theory. Some (very) short hints follow.

Exercises from the topics of Chapter 1

Exercise 1. Find an example of a sequence of functions $f_n : [0,1] \to [0,1]$ such that $(f_n)_\#(\mathscr{L}^1 \llcorner [0,1]) = \mathscr{L}^1 \llcorner [0,1]$ but $f_n \rightharpoonup \frac{1}{2}$. Can these functions be taken C^1?

Exercise 2. ♠ Consider the problem

$$\max \left\{ \int |x - \mathrm{T}(x)|^2 \, \mathrm{d}\mu \quad : \quad \mathrm{T}_\#\mu = \nu \right\}.$$

Prove that, if μ is absolutely continuous, this problem admits a solution and the optimal T is the gradient of a concave function.

Exercise 3. Find the optimal transport map for the quadratic cost $c(x,y) = |x-y|^2$ between $\mu = f \cdot \mathscr{L}^2$ and $\nu = g \cdot \mathscr{L}^2$ in 2D, where $f(x) = \frac{1}{\pi} \mathbb{1}_{B(0,1)}(x)$ and $g(x) = \frac{1}{8\pi}(4 - |x|^2)$.

Exercise 4. Let $R : \mathbb{R}^d \to \mathbb{R}^d$ be given by $R(x) = -x$. Characterize the probabilities $\mu \in \mathscr{P}_2(\mathbb{R}^d)$ such that R is an optimal transport map between μ and $R_\#\mu$ for the quadratic cost.

Exercise 5. Let $S : \mathbb{R}^d \to \mathbb{R}^d$ be given by $S(x_1, x_2, \dots, x_d) = (x_1, x_2, \dots, -x_d)$. Prove that S cannot be the optimal transport map between μ and $S_\#\mu$ for the quadratic cost if $\mu \ll \mathscr{L}^d$. Prove that if S is optimal in the above sense, then μ is concentrated on a Lipschitz hypersurface of the form $\{x_d = f(x_1, x_2, \dots, d_{d-1})\}$, with $f \in \mathrm{Lip}_1(\mathbb{R}^{d-1})$.

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2

Exercise 6. Consider the multi-marginal problem

$$\min\left\{\int c(x_1,\dots,x_N)\mathrm{d}\gamma \,:\, \gamma\in\mathscr{P}((\mathbb{R}^d)^N)\,(\pi_i)_\#\gamma=\mu_i\right\},$$

where $\mu_i\in\mathscr{P}(\mathbb{R}^d)$ are given compactly supported measures. Suppose that γ is concentrated on a set $\{(x_1,\dots,x_N)\in(\mathbb{R}^d)^N \,:\, x_1+\cdots+x_N=\textit{constant}\}$. Prove that γ is optimal whenever the cost c is given by $c(x_1,\dots,x_N)=h(x_1+\cdots+x_N)$ for a convex function $h:\mathbb{R}^d\to\mathbb{R}$ and also in the case $c(x_1,\dots,x_N)=-\sum_{i\neq j}|x_i-x_j|^2$.

Exercises from the topics of Chapter 2

Exercise 7. Consider two probability measures μ and ν on $\mathbb{R}$ that we will suppose atomless and with full support for simplicity. Let F_μ and F_ν be their cumulative distribution functions: $F_\mu(x)=\mu((-\infty,x])$ and $F_\nu(x)=\nu((-\infty,x])$. Prove that the following facts are equivalents:

1. $F_\mu\geq F_\nu$,
2. there exists $\gamma\in\Pi(\mu,\nu)$ concentrated on $\{(x,y)\,:\,y\geq x\}$,
3. there exists $\mathrm{T}:\mathbb{R}\to\mathbb{R}$ such that $\mathrm{T}_\#\mu=\nu$ and $\mathrm{T}(x)\geq x$ for every x,
4. the monotone increasing transport T from μ to ν satisfies $\mathrm{T}(x)\geq x$ for every x.

Exercise 8. Prove that an optimal $\gamma\in\Pi(\mu,\mu)$ for the cost $c(x,y)=\frac{1}{|x-y|}$ exists and that it is concentrated on at most two increasing graphs, one above and one below the diagonal. Prove that these two graphs do not superpose and that the plan comes indeed from a transport map. Is it increasing?

Exercise 9. * In the case $\mu=\mathscr{L}^1\llcorner[0,1]$, find the optimal transport map from μ to μ for the cost $c(x,y)=\frac{1}{|x-y|}$.

Exercise 10. ♠ Prove the optimality of γ_{mon} in 1D under the assumption that the cost $c:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ is C^2 with $\frac{\partial^2}{\partial x\partial y}c(x,y)<0$ for every $(x,y)\in\mathbb{R}\times\mathbb{R}$. Deduce the existence of an optimal map if μ is atomless.

Exercise 11. Let A and B two symmetric positive definite matrices and $L_A,L_B:\mathbb{R}^d\to\mathbb{R}^d$ be affine functions defined through $L_A(x)=Ax+a$, $L_B(x)=Bx+b$, for given $a,b\in\mathbb{R}^d$. Take a reference measure ϱ which is invariant under rotations (i.e., $R_\#\varrho=\varrho$ for every matrix R s.t. $R\cdot R^t=\mathrm{I}$) and define $\mu:=(L_A)_\#\varrho$ and $\nu:=(L_B)_\#\varrho$. Find the optimal transport map from μ to ν for the quadratic cost.

Exercises from the topics of Chapter 3

Exercise 12. Consider a domain $\Omega=[0,3]\times[0,2]$ and two absolutely continuous measures on it, μ_0 and μ_1, with density ϱ_0 and ϱ_1, respectively, given by

$$\varrho_0(x,y)=\frac{2}{9}\text{ if }x\leq 1\text{ or }x\in]1,2[\text{ and }y\leq 1,$$

$$\varrho_0(x,y)=\frac{1}{9}\text{ if }x\geq 2\text{ or }x\in]1,2[\text{ and }y>1,$$

$$\varrho_1(x,y) = \frac{2}{9} \text{ if } x \geq 2 \text{ or } x \in]1,2[\text{ and } y \leq 1,$$

$$\varrho_1(x,y) = \frac{1}{9} \text{ if } x \leq 1 \text{ or } x \in]1,2[\text{ and } y \geq 1.$$

Represent graphically these two densities and find an optimal transport, as well as a Kantorovich potential, between μ_0 and μ_1 for the cost $c(x,y) = |x-y|$. Is the optimal transport unique? And the potential?

Exercise 13. Find the optimal transport map for the cost $c(x,y) = |x-y|$ between $\mu = f \cdot \mathscr{L}^2$ and $\nu = g \cdot \mathscr{L}^2$ in 2D, where $f(x) = \frac{1}{\pi} \mathbb{1}_{B(0,1)}(x)$ and $g(x) = \frac{1}{8\pi}(4 - |x|^2)$.

Exercise 14. Consider the cost $c(x,y) = |x-y|^\alpha$ for a given exponent $0 < \alpha < 1$. Prove that, for every pair $\mu, \nu \in \mathscr{P}(\Omega)$, we have

$$\mathscr{T}_c(\mu,\nu) = \sup \left\{ \int u \, \mathrm{d}(\mu - \nu) \ : \ u \in C^{0,\alpha}(\Omega),\ [u]_{0,\alpha} \leq 1 \right\},$$

where $[u]_{0,\alpha}$ is the $C^{0,\alpha}$ seminorm of u: $[u]_{0,\alpha} := \sup_{x \neq y} \frac{|u(x)-u(y)|}{|x-y|^\alpha}$.

Exercise 15. Consider the cost $c(x,y) = \arctan(|x-y|)$ and two measures $\mu, \nu \in \mathscr{P}(\mathbb{R}^d)$ with compact and disjoint supports.

1. What can we say about c-concave functions for this c? Are they (uniformly) continuous, Lipschitz?
2. Do the problems $\min \left\{ \int c(x,y) \, \mathrm{d}\gamma \ : \ \gamma \in \Pi(\mu,\nu) \right\}$ admit a solution?
3. Write the dual problem in terms of functions u having a given modulus of continuity.
4. What is the connection between the support of an optimal γ and the optimal u in the dual problem?
5. Supposing that μ is absolutely continuous, prove that the optimal γ is unique and derives from a transport. Give its expression in terms of ∇u.

Exercise 16. Consider $h : \mathbb{R}^d \to \mathbb{R}$ defined through $h(z) = (|z|-1)_+^2$. Prove that, for every $\mu, \nu \in \mathscr{P}(\Omega)$ with $\mu \ll \mathscr{L}^d$ and Ω compact, there exists an optimal transport map for the cost $c(x,y) = h(x-y)$.

Exercise 17. Consider the cost $c(x,y) = \lfloor |x-y| \rfloor$ (the integer part of the distance: $\lfloor a \rfloor = \max\{n \in \mathbb{Z} \ : \ n \leq a\}$. Prove that, for every $\mu, \nu \in \mathscr{P}(\Omega)$ with $\mu \ll \mathscr{L}^d$ and Ω compact, there exists an optimal transport plan for such a cost of the form $\gamma = \sum_{i=1}^N \gamma_i$ where each γ_i is induced by a transport map and $N \leq \mathrm{diam}(\Omega) + 1$.

Exercise 18. For $\varepsilon > 0$, consider the cost $c_\varepsilon(x,y) = \varepsilon \left\lfloor \frac{|x-y|}{\varepsilon} \right\rfloor$ and let γ_ε be an optimal transport plan for such a cost between two measures $\mu, \nu \in \mathscr{P}(\Omega)$ (consider Ω compact for simplicity). Prove that, up to subsequences, $\gamma_\varepsilon \rightharpoonup \gamma$, where γ is an optimal transport plan between the same measures for the cost $c(x,y) = |x-y|$.

Exercise 19. ♠ Consider the cost $c_\varepsilon(x,y) = \sqrt{\varepsilon^2 + |x-y|^2}$ and let γ_ε be optimal for the cost c_ε in $\Pi(\mu,\nu)$. Let $\varepsilon \to 0$. Prove that, up to subsequences, we have $\gamma_\varepsilon \rightharpoonup \gamma$ where γ is optimal for the cost $|x-y|$. Supposing $\mu \ll \mathscr{L}^d$ and $\mathrm{spt}(\mu) \cap \mathrm{spt}(\nu) = \emptyset$, prove that we have full convergence to the ray-monotone transport plan. What can we say about the limit $\varepsilon \to \infty$?

Exercise 20. * Find a Borel set $A \subset [0,1]$ such that for every interval $I \subset [0,1]$, we have $0 < \mathscr{L}^1(I \cap A) < \mathscr{L}^1(I)$.

Exercise 21. Let us define a distance, called *SNCF distance*, on a set X in the following way:

$$\mathrm{SNCF}(x,y) := \begin{cases} 0 & \text{if } x = y, \\ P(x) + P(y) & \text{if } x \neq y, \end{cases}$$

where $P : X \to \mathbb{R}$ is a bounded function, to be interpreted as the distance to a fixed connecting hub: imagine that every point is connected by a direct train to this hub and to no other location.[1] Prove that $\mathrm{SNCF}(x,y)$ is a distance on X and that, for $c(x,y) = \mathrm{SNCF}(x,y)$ and $\mu,\nu \in \mathscr{P}(X)$, we have $\mathscr{T}_c(\mu,\nu) = \int P(x)\,\mathrm{d}|\mu-\nu|(x)$.

Exercises from the topics of Chapter 4

Exercise 22. Let y be a solution of the ODE $y'(t) = \mathbf{v}_t(y(t))$. Prove that $\rho_t := \delta_{y(t)}$ is a solution of the continuity equation $\partial_t \rho_t + \nabla \cdot (\rho_t \mathbf{v}_t) = 0$.

Exercise 23. * Prove that the set of polynomial functions is dense for the C^1 norm in the set $C^1([0,1]^d)$.

Exercise 24. Find the transport density σ (see Section 4.3) between the measures $\mu = f \cdot \mathscr{L}^2$ and $\nu = g \cdot \mathscr{L}^2$ in 2D, where $f = \frac{1}{\pi}\mathbb{1}_{B(0,1)}$ and $g(x) = \frac{1}{4\pi}\mathbb{1}_{B(0,2)}$.

Exercise 25. Suppose $\mu \in L^p(\Omega)$ and $\nu \in L^q(\Omega)$, with $p > q$. Prove that, if $p < d/(d-1)$, the transport density σ in the transport from μ to ν belongs to L^p and, if $p \geq d/(d-1)$, it belongs to $L^r(\Omega)$ for all the exponents r satisfying

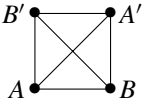
$$r < r(p,q,d) := \frac{dq(p-1)}{d(p-1)-(p-q)}.$$

Exercise 26. ♠ Given $\mu,\nu \in \mathscr{P}(\Omega)$, prove that there exists $Q \in \mathscr{P}(\mathscr{C})$ with $(e_0)_\# Q = \mu$, $(e_1)_\# Q = \nu$ and $i_Q \in L^p(\Omega)$ (we recall that i_Q is the traffic intensity defined in Section 4.2.3 and that the e_t are the evaluation maps) if and only if $\mu - \nu \in (W^{1,p'})^*$.

[1] The common hub could be called Paris, for instance. However, no restriction on the space X is imposed, and in particular, there is no need to assume that it has hexagonal shape.

Exercise 27. Given $Q \in \mathscr{P}(\mathscr{C})$ with $(e_0)_\# Q = \mu, (e_1)_\# Q = \nu$, prove that there exists $\tilde{Q} \in \mathscr{P}(\mathscr{C})$ concentrated on injective curves and such that $(e_0)_\# \tilde{Q} = \mu, (e_1)_\# \tilde{Q} = \nu$ and $i_{\tilde{Q}} \le i_Q$.

Exercise 28. Consider the simplified network on the right, where the cost for a commuter to move along a segment of length ℓ with traffic intensity i is given by $i\ell$. Do not distinguish between different traffic directions on a same segment (i.e., i is the sum of the number of commuters going in one direction plus the number going in the other direction). Consider the case where a quantity 1 of commuters must go from A to A' and another quantity 1 from B to B'. Find the Wardrop equilibrium.



Exercise 29. Prove that $d_\alpha(\mathscr{L}_\Omega, \delta_0) = +\infty$ for every compact domain $\Omega \subset \mathbb{R}^d$ with nonempty interior and $|\Omega| = 1$ and every $\alpha < 1 - 1/d$.

Exercise 30. Consider the branched transport for $\alpha = 1/2$ between two atomic measures $\mu = \sum_i a_i \delta_{x_i}$ and $\nu = \sum_j b_j \delta_{y_j}$. Consider a triple junction at a point z which is not one of the points x_i or y_j: among the three branches at z, either two are incoming and one outcoming or two outcoming and one incoming. Prove that the two with the same orientation must make an angle of 90°.

Exercises from the topics of Chapter 5

Exercise 31. ♠ Let X be a Polish metric space and $\mu_n, \mu \in \mathscr{P}_1(X)$ be probability measures on X such that $\int \phi \, \mathrm{d}\mu_n \to \int \phi \, \mathrm{d}\mu$ for every $\phi \in \mathrm{Lip}_1(X)$. Prove $\int \phi \, \mathrm{d}\mu_n \to \int \phi \, \mathrm{d}\mu$ for every $\phi \in C_b(X)$.

Exercise 32. Prove the following equalities, for every compact set $\Sigma \subset \Omega$ and every density $f \in L^1(\Omega)$ with $f \ge 0$ and $\int_\Omega f(x) \, \mathrm{d}x = 1$,

$$\int d(x, \Sigma) f(x) \, \mathrm{d}x = \min\{W_1(f, \nu) \,:\, \nu \in \mathscr{P}(\Sigma)\}$$
$$= \min\{|\mathbf{w}|(\Omega) \,:\, \mathbf{w} \in \mathscr{M}^d_{\mathrm{div}}(\Omega),\ \mathrm{spt}(\nabla \cdot \mathbf{w} - f) \subset \Sigma\}.$$

What could we say if we considered, instead, $\int d(x, \Sigma)^p f(x) \, \mathrm{d}x$, for $p > 1$?

Exercise 33. Let $\gamma \in \Pi(\mu, \nu)$ be a transport plan between two probabilities on two compact spaces X and Y, respectively. Let $\mu_n \rightharpoonup \mu$ and $\nu_n \rightharpoonup \nu$ be two weakly converging sequences of probabilities. Prove that there exists a sequence $\gamma_n \in \Pi(\mu_n, \nu_n)$ with $\gamma_n \rightharpoonup \gamma$. Deduce that for any sequence of costs c_n converging uniformly to c, we have $\mathscr{T}_{c_n}(\mu_n, \nu_n) \to \mathscr{T}_c(\mu, \nu)$.

Exercise 34. Let $\Omega \subset \mathbb{R}^d$ be smooth and compact, and V be an operator which associates with every probability $\varrho \in \mathscr{P}(\Omega)$ a vector field $V[\varrho] : \Omega \to \mathbb{R}^d$ with the following properties: for every ϱ, $V[\varrho]$ is Lipschitz continuous and satisfies $\mathrm{Lip}(V[\varrho]) \le C$ (for a constant C not depending on ϱ) and $V[\varrho] \cdot \mathbf{n} = 0$ on $\partial\Omega$; the map $\varrho \mapsto V[\varrho]$ is Lipschitz continuous in the following sense: $||V[\varrho] - V[\mu]||_{L^\infty} \le C W_2(\varrho, \mu)$ for every ϱ, μ. Consider the equation $\partial_t \varrho + \nabla \cdot (\varrho V[\varrho]) = 0$ with initial

Cauchy datum $\varrho_0 \ll \mathscr{L}^d$. Prove that any solution ϱ_t stays absolutely continuous for $t > 0$ and prove uniqueness of the solution.

Exercise 35. ♠ Prove the equality $W_\infty(\mu,\nu) = \lim_{p\to\infty} W_p(\mu,\nu)$ for μ,ν compactly supported. Prove also that for any sequence $p_n \to \infty$, the optimal plans γ_n for the cost $|x-y|^{p_n}$ converge, up to subsequences, to an optimal plan for the problem defining the W_∞ distance.

Exercise 36. Prove the inequality $d_H(\mathrm{spt}(\mu),\mathrm{spt}(\nu)) \le W_\infty(\mu,\nu)$, where d_H denotes the Hausdorff distance.

Exercise 37. * Is it true that for $1 \le p \le q \le \infty$, if the same map T is optimal in the definition of W_p and W_q, then it is also optimal in the definition of W_r for every $r \in [p,q]$?

Exercise 38. Prove that if $f \in W^{1,p}(\Omega)$, Ω is a convex set, $\mu,\nu \in \mathscr{P}(\Omega) \cap L^r(\Omega)$ and $||\mu||_{L^r}$, $||\nu||_{L^r} \le C$, then, for $\frac{1}{p}+\frac{1}{q}+\frac{1}{r} = 1+\frac{1}{qr}$, one has

$$\int_\Omega f \,\mathrm{d}(\mu-\nu) \le C^{\frac{1}{q'}} ||\nabla f||_{L^p(\Omega)} W_q(\mu,\nu).$$

Exercises from the topics of Chapter 6

Exercise 39. Let $\Sigma = \{x_1,\dots x_N\} \subset \Omega$ be a finite subset of a compact convex domain Ω. For every i, let us set $V_i := \{x \in \Omega : |x-x_i| \le |x-x_j| \text{ for all } j\}$. Suppose that Σ solves the problem $\min\{\int_\Omega d(x,\Sigma)^2\,\mathrm{d}x : \#\Sigma \le N\}$. Prove that, for every i, the point x_i is the barycenter of V_i.

Exercise 40. ♠ For $A \subset \mathbb{R}^d$ we define the extremal points of A as those points $x \in A$ such that $y,z \in A, x = \frac{y+z}{2} \Rightarrow y = z$. Let $K \subset \mathscr{M}_n(\mathbb{R})$ be the set of bistochastic matrices (those matrices $M = (m_{ij})$ with $m_{ij} \ge 0$, $\sum_j m_{ij} = \sum_i m_{ij} = 1$ for all i,j). Prove that the extremal points of k are the permutation matrices.

Exercise 41. For $a \in (\mathbb{R}_+)^n$ and $b \in (\mathbb{R}_+)^m$ consider the set $\Pi(a,b)$ defined through $\Pi = \{M \in \mathscr{M}_{n,m}(\mathbb{R}) : m_{ij} \ge 0, \sum_j m_{ij} = a_i, \sum_i m_{ij} = b_j\}$. Prove that every extremal matrix M in $\Pi(a,b)$ is such that $\#\{(i,j) : m_{ij} > 0\} \le n+m$.

Exercise 42. Let $\gamma^\varepsilon_{ij} := p^\varepsilon_i q^\varepsilon_j \eta_{ij}$ be a solution of the approximated problem in (6.3) (see Section 6.4.1). Suppose that $\varepsilon\log(p^\varepsilon_i)$, $\varepsilon\log(q^\varepsilon_j)$, and γ^ε_{ij} converge as $\varepsilon \to 0$. Prove that the limit of $\varepsilon\log(p^\varepsilon_i)$ and $\varepsilon\log(q^\varepsilon_j)$ are Kantorovich potentials in the limit problem for $\varepsilon = 0$.

Exercise 43. Let u be a smooth function on Ω and $M(x) := \mathrm{cof}(D^2u(x))$. Let M^{ij} be the (i,j)-entry of M and M^{ij}_k its derivative w.r.t. x_k. Prove that for every index j, we have $\sum_i M^{ij}_i(x) = 0$.

Exercise 44. * ♠ Consider a smooth transport map $\mathrm{T} : \Omega \to \mathbb{R}^d$ of the form $\mathrm{T} = \nabla u$ on a smooth domain Ω. Suppose $\mathrm{T}_\# f = \nu$, where f is a smooth positive density on Ω. For every vector field $\mathbf{v} : \Omega \to \mathbb{R}^d$ with $\nabla\cdot(f\mathbf{v}) = 0$ and $\mathbf{v}\cdot\mathbf{n} = 0$

on $\partial\Omega$, consider the flow of the autonomous vector field $\mathbf{v}$, given by $Y_t(x) = y_x(t)$ (where $y_x'(t) = \mathbf{v}(y_x(t))$ and $y_x(0) = x$), and define $\mathrm{T}_t = \mathrm{T} \circ Y_t$. Let $M(S) = \int |S(x) - x|^2 f(x)\,\mathrm{d}x$. Prove that the function $j(t) = M(\mathrm{T}_t)$ satisfies $j''(0) \geq 0$ for every $\mathbf{v}$ if and only if u is convex.

Exercises from the topics of Chapter 7

Exercise 45. * ♠ Prove that the entropy functional

$$\mathscr{F}(\varrho) := \begin{cases} \int_{\mathbb{R}^d} \varrho(x) \log \varrho(x)\,\mathrm{d}x & \text{if } \varrho \ll \mathscr{L}^d, \\ +\infty & \text{otherwise.} \end{cases}$$

is l.s.c. in $\mathbb{W}_1(\mathbb{R}^d)$. Also prove that if $\varrho_n \in \mathscr{P}_2(\mathbb{R}^d)$ and $\varrho_n \rightharpoonup \varrho$, $\int |x|^2 \mathrm{d}\varrho_n(x) \leq C$, then $\mathscr{F}(\varrho) \leq \liminf_n \mathscr{F}(\varrho_n)$.

Exercise 46. Prove that the problem

$$\min \left\{ \int_0^1 \int_\Omega \varrho |\mathbf{v}|^2 \mathrm{d}x\mathrm{d}t \ : \ \partial_t \varrho + \nabla\cdot(\varrho\mathbf{v}) = 0,\ \varrho_t \in \mathscr{P}(\Omega),\ \varrho(0) = \varrho_0,\ \varrho(1) = \varrho_1 \right\}$$

is equivalent to the same problem with the additional constraint $\varrho_t \leq 1$, provided that we have $\varrho_0, \varrho_1 \leq 1$ and Ω is convex. Also, imagine a non-convex (but connected) domain Ω where this is no more the case.

Exercise 47. Given a compact and convex domain $\Omega \subset \mathbb{R}^d$ and $\nu \in \mathscr{P}(\Omega)$, prove that the following problem admits a solution:

$$\min \left\{ W_1(\varrho, \nu) + \frac{1}{2} \int \varrho^2(x)\,\mathrm{d}x \ : \ \varrho \in L^2(\Omega),\ \varrho \geq 0,\ \int \varrho(x)\,\mathrm{d}x = 1 \right\}.$$

Is this solution unique? Write the optimality conditions satisfied by the optimal ϱ, prove that $\varrho \in \mathrm{Lip}_1$, and find explicitly the solution when $\Omega = \overline{B(0,5)}$ and $\nu = \delta_0$.

Exercise 48. ** Given $p > 1$, consider $g \in \mathscr{P}(\Omega)$ an absolutely continuous probability measure, and look at the problem $\min\{W_p(\varrho, g) \ : \ \varrho \leq f\}$, where $f \in L^1(\Omega)$ is a given positive function. Prove that a solution exists, that it is unique, and that it has the form $\varrho = \mathbb{1}_A g + \mathbb{1}_{A^c} f$.

Exercises from the topics of Chapter 8

Exercise 49. Consider the fourth-order PDE

$$\partial_t \varrho + \nabla \cdot (\varrho \nabla(\Delta\varrho)) + \nabla \cdot \left(\frac{\nabla \varrho}{\sqrt{\varrho}} \right) = 0.$$

This PDE is the gradient flow in $\mathbb{W}_2$ of an energy: which one? Also prove that the variational problems to be solved at every time step (see Equation (8.3)) in the corresponding minimizing movement scheme admit a solution.

Exercise 50. Let $V : \mathbb{R}^d \to \mathbb{R}$ be $C^{1,1}$ and let $\mathscr{G} : \mathscr{P}(\mathbb{R}^d) \to \mathbb{R}$ be defined through $G(\varrho) = \#(\mathrm{spt}(\varrho))$ and let $\varrho_0 \in \mathscr{P}(\mathbb{R}^d)$ be a purely atomic probability measure $\varrho_0 = \sum_{i=1}^N a_i \delta_{x_i^0}$ with finitely many atoms. Prove that for small τ (depending on what?), the solution of the following minimization problem

$$\min\left\{\mathscr{G}(\varrho) + \int V \mathrm{d}\varrho + \frac{W_2^2(\varrho^0, \varrho)}{\tau} \ : \ \varrho \in \mathscr{P}_2(\mathbb{R}^d)\right\}$$

exists, is unique, and has the form $\varrho_1 = \sum_{i=1}^N a_i \delta_{x_i^1}$ with $\mathscr{G}(\varrho_1) = \mathscr{G}(\varrho_0)$. Deduce that the minimizing movement scheme for $\mathscr{G} + \mathscr{V}$ gives the same limit curve as that for $\mathscr{V}$, and find the equation of this curve.

Exercise 51. Let $\varrho_0 \in \mathscr{P}_2(\mathbb{R}^d)$, x_0 its barycenter, and ϱ a solution of

$$\partial_t \varrho_t - \nabla \cdot (\varrho_t \mathbf{v}_t) = 0$$

where $\mathbf{v}_t(x) = \int (x-y) \mathrm{d}\varrho_t(y)$. Let $E(t) := \int |x - x_0|^2 \mathrm{d}\varrho_t(x)$. Prove $E'(t) \leq -E(t)$. What can we conclude on the asymptotical behavior of ϱ_t as $t \to \infty$?

Exercise 52. Let $V : \mathbb{R}^d \to \mathbb{R}$ be a continuous function such that $|V(x)| \leq A + B|x|^2$ for suitable constants A, B, and let $f : \mathbb{R}_+ \to \mathbb{R}$ a convex function with $f'(0) > -\infty$. Define

$$F(\varrho) = \int f(\varrho(x))\, \mathrm{d}x + \int V(x)\varrho(x)\, \mathrm{d}x$$

and consider, for $\varrho_0 \in \mathscr{P}_2(\mathbb{R}^d)$,

$$\min\left\{F(\varrho) + \frac{W_2^2(\varrho, \varrho_0)}{2\tau} \ : \ \varrho \in \mathscr{P}_2(\mathbb{R}^d)\right\}.$$

Prove that, for small τ, a solution exists, is unique, and is compactly supported if ϱ_0 is compactly supported.

Exercise 53. Consider $m > 2$ and define

$$F(\varrho) = \int_\Omega \varrho^m - \frac{1}{2}\int_\Omega |\nabla u_\varrho|^2, \quad \text{where } u_\varrho \in H_0^1(\Omega) \text{ solves } -\Delta u_\varrho = \varrho.$$

Prove that

$$\min\left\{F(\varrho) + W_2^2(\varrho, \varrho_0) \ : \ \varrho \in \mathscr{P}_2(\mathbb{R}^d)\right\}$$

admits a solution.

Exercise 54. Prove that the optimality conditions of

$$\min\left\{\int_0^T\int\left(\frac{t^{3/2}}{2}\varrho|\mathbf{v}|^2+\frac{3t^{-1/2}}{4}|\nabla\varphi|^2\right)\ :\ \begin{cases}\partial_t\varrho+\nabla\cdot(\varrho\mathbf{v})=0,\\ \Delta\varphi=\varrho-1,\\ \varrho_0=1,\varrho_T=\text{given}\end{cases}\right\}$$

impose, at least formally, the system of Equation (8.27).

Exercise 55. Given $\alpha\in]0,1[$, prove that a minimizer for

$$\min\left\{\int_0^1 t^{1+\alpha}|x'(t)|^2+ct^{-1+\alpha}|x(t)-x_0|^2\ :\ x(1)=x_1\right\}$$

exists, is unique, and satisfies $x(0)=x_0$. Prove that a minimizer for the same problem adding the condition $x(0)=\tilde{x}_0$, with $\tilde{x}_0\neq x_0$, does not exist.

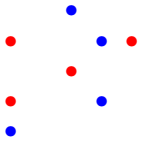
Exercise 56. Consider a triplet $(\varrho_t,\mathbf{v}_t,\varphi_t)$ on the torus $\mathbb{T}^d$ (to avoid boundary issues) satisfying

$$\int_0^1\int\left(\frac{t^{3/2}}{2}\varrho|\mathbf{v}|^2+\frac{3t^{-3/2}}{2}|\nabla\varphi|^2\right)<+\infty\ \text{with}\ \begin{cases}\partial_t\varrho_t+\nabla\cdot(\varrho_t\mathbf{v}_t)=0,\\ \Delta\varphi_t=\varrho_t-1\end{cases}$$

with $||\varrho_t||_{L^\infty}\leq C$ and $t\mapsto\varrho_t$ continuous for W_2 on $]0,1]$. Prove that we necessarily have $W_2(\varrho_t,1)\to 0$ as $t\to 0$.

Miscellaneous

Exercise 57. Find the optimal transport map for the costs of the form $|x-y|^p$, $p>0$, between the two measures sketched below:



$$\mu=\tfrac{1}{4}(\delta_{(2,3)}+\delta_{(3,2)}+\delta_{(3,0)}+\delta_{(0,-1)})$$
$$\nu=\tfrac{1}{4}(\delta_{(0,0)}+\delta_{(0,2)}+\delta_{(2,1)}+\delta_{(4,2)})$$

Exercise 58. Find an example of a pair of compactly supported absolutely continuous probability measures μ,ν on $\mathbb{R}^d$ and a continuous cost $c:\mathbb{R}^d\times\mathbb{R}^d\to\mathbb{R}_+$ such that no optimal transport map exists in the corresponding Monge problem.

Exercise 59. Find the optimal transport plan for the cost $c(x,y)=(x^2-y^2)^2$ in 1D, when $\mu=\mathscr{L}^1\llcorner[0,1]$ and $\nu=\frac{1}{2}\mathscr{L}^1\llcorner[-1,1]$.

Exercise 60. Find an example of bounded l.s.c. (but not continuous) cost $c:\mathbb{R}^d\to\mathbb{R}$, together with two atomless compactly supported measures $\mu,\nu\in\mathscr{P}(\mathbb{R}^d)$ such that inf(MP)> min(KP).

Exercise 61. * Prove that there exists a continuous map $\mathrm{T} : [0,1] \to [0,1]^2$ such that $\mathrm{T}_\#(\mathscr{L}^1 \llcorner [0,1]) = \mathscr{L}^2 \llcorner [0,1]^2$.

Exercise 62. Consider a translation $S : \mathbb{R}^d \to \mathbb{R}^d$ given by $S(x) = x + a$.

1. Prove that S is the optimal transport for the cost $c(x,y) = |x-y|^2$ between any measure μ and $S_\#\mu$.
2. Prove the same result for $c(x,y) = |x-y|$.
3. If T is the optimal transport between μ and ν, can we say that $S \circ T$ is the optimal transport between μ and $S_\#\nu$? Answer this question both for $c(x,y) = |x-y|^2$ and $c(x,y) = |x-y|$.
4. More generally, is it true that the composition of two optimal transports is still optimal?

Exercise 63. ♠ Prove that, if $J \subset \mathbb{R}$ is a (possibly unbounded) given interval, then $\mathbb{W}_p(J)$ is isometric to a closed convex subset of $L^p([0,1])$.

Exercise 64. Given $g \in L^1([0,1]) \cap \mathscr{P}([0,1])$, prove $W_2(1,g) = ||1-g||_{\dot{H}^{-1}([0,1])}$.

Exercise 65. Consider $\mu, \nu \in \mathscr{P}_1(\mathbb{R}^d)$ with $\int x \,\mathrm{d}\mu(x) = \int y \,\mathrm{d}\nu(y) = 0$. Let γ be an optimal transport plan for the cost $c(x,y) = -x \cdot y$ between these two measures and $\mathscr{T}_c(\mu,\nu)$ the optimal value of this transport cost. Prove $\mathscr{T}_c(\mu,\nu) \leq 0$ and $\mathrm{spt}(\gamma) \subset \{(x,y) \,:\, x \cdot y \geq \mathscr{T}_c(\mu,\nu)\}$. Deduce that if $\mu_n \rightharpoonup \mu$ and $\int x \,\mathrm{d}\mu_n(x) = \int y \,\mathrm{d}\nu(y) = 0$, then $\mathscr{T}_c(\mu,\nu) \geq \limsup_n \mathscr{T}_c(\mu_n,\nu)$.

Exercise 66. Let f, g be two smooth probability densities on a convex domain Ω and φ, ψ the Kantorovich potentials for the transport from f to g for the cost $c(x,y) = \frac{1}{2}|x-y|^2$. Prove $\int \nabla f \cdot \nabla\varphi + \nabla g \cdot \nabla\psi \geq 0$. Conclude that the same is true for $f, g \in H^1(\Omega)$.

Exercise 67. Prove the W_2 distance between two (smooth) solutions to the heat equation $\partial_t \varrho = \Delta\varrho$ decreases in time:

1. in a bounded domain Ω with homogeneous Neumann boundary conditions,
2. in $\mathbb{R}^d$, with initial data in $\mathscr{P}_2(\mathbb{R}^d)$.

Exercise 68. Let $\mathbf{v} : [0,T] \times \Omega \to \mathbb{R}^d$ be Lipschitz continuous in space, uniformly in time. Prove the uniqueness of the solution to the Fokker-Planck equation $\partial_t\varrho - \Delta\varrho + \nabla \cdot (\varrho \mathbf{v}) = 0$ (with no-flux boundary conditions) in the class of solutions which are $H^1(\Omega) \cap \mathscr{P}(\Omega)$ for every t, by differentiating W_2^2 between two solutions. Why is H^1 a reasonable assumption for the regularity of the solution?

Exercise 69. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a C^2 strictly convex function with $D^2 f \geq \alpha \mathrm{I}$ (for $\alpha > 0$) on a smooth bounded domain Ω. Prove that $\varrho_\infty = ce^{-f}$ (for $c = (\int e^{-f})^{-1}$) is a solution of $\Delta\varrho + \nabla \cdot (\varrho \nabla f) = 0$. Consider now any smooth solution of $\partial_t\varrho - \Delta\varrho - \nabla \cdot (\varrho \nabla f) = 0$ with Neumann boundary conditions on Ω, and prove $W_2(\varrho_t, \varrho_\infty) \to 0$ exponentially as $t \to \infty$.

Hints

Hint to Ex. 1. Use piecewise affine oscillating functions. With C^1 it is impossible, because C^1 functions f with image measure with bounded density should satisfy $f' \neq 0$ and be diffeomorphisms. Yet, only $f(x) = x$ and $f(x) = 1 - x$ preserve the measure among diffeomorphisms.

Hint to Ex. 2. Write the Kantorovich problem and use a change of variable $y \mapsto -y$ in the arrival measure.

Hint to Ex. 3. Find a monotone radial transport. The answer is

$$\mathrm{T}(x) = 4\sqrt{1 - \sqrt{1 - |x|^2}}\,\frac{x}{|x|}.$$

Hint to Ex. 4. Use c-cyclical monotonicity on two points (i.e., monotonicity).

Hint to Ex. 5. For the first question, use the characterization in terms of gradients of convex functions. For the second, use again c-cyclical monotonicity.

Hint to Ex. 6. For the first cost, use Jensen inequality. For the second, express it as $(x_1 + \ldots x_N)^2 +$ functions of the x_i separately.

Hint to Ex. 7. You can (partially) use the transport problem with cost $h(y-x)$ with $h(z) = +\infty$ for $z < 0$ and $h(z) = 0$ for $z \geq 0$.

Hint to Ex. 8. Decompose γ above and below the diagonal and use results on convex costs in 1D. Then prove that the superposition is not optimal.

Hint to Ex. 9. Check that T defined as $T(x) = x + \frac{1}{2}$ for $x < \frac{1}{2}$ and $T(x) = x - \frac{1}{2}$ for $x \geq \frac{1}{2}$ is optimal by using the functions $\phi(x) = 2 - 4|x - \frac{1}{2}|$ and $\psi(y) = 2 - 4|y - \frac{1}{2}|$ as a guess for the Kantorovich potentials.

Hint to Ex. 10. If two pairs $(x_0, y_0), (x_1, y_1) \in \operatorname{spt}\gamma$ exist with $x_0 < x_1$ and $y_0 > y_1$ exist, integrate $\frac{\partial^2 c}{\partial x \partial y}$ on $[x_0, x_1] \times [y_1, y_0]$ and find a contradiction.

Hint to Ex. 11. Write $A^{-1}B$ as the product of a symmetric positive definite matrix and a rotation.

Hint to Ex. 12. $u(x) = -x$ is a potential.

Hint to Ex. 13. Same idea as in **Ex**(3). Use the potential $u(x) = -|x|$ to prove that you have found an optimal map.

Hint to Ex. 14. The cost $c(x, y) = |x - y|^\alpha$ is a distance.

Hint to Ex. 15. Use the same analysis as in Section 3.3.2.

Hint to Ex. 16. Decomposition strategy of Section 3.3.1.

Hint to Ex. 17. Use the results about L^∞ optimal transport.

Hint to Ex. 18. The cost c_ε uniformly converges to the linear cost.

Hint to Ex. 19. Consider the Taylor expansion of the cost in ε to guess the lower-order term.

Hint to Ex. 20. Start from classical examples of fattened rationals.

Hint to Ex. 21. Prove that all functions $u : X \to \mathbb{R}$ with $|u(x)| \leq P(x)$ belong to Lip_1 for the distance SNCF, and apply the duality formula to have a lower bound on $\mathscr{T}_c$. For the opposite bound, just leave the common mass at rest.

Hint to Ex. 22. Just test against a C^1_c function $\psi(x)$.

Hint to Ex. 23. Two methods: induction on the dimension and use a polynomial convolution (with something like $x^n(1-x)^n$).

Hint to Ex. 24. By uniqueness, σ is radial. The Kantorovich potential is $u(x) = -|x|$. Use the Monge-Kantorovich equation.

Hint to Ex. 25. Use the same strategy as in Theorem 4.24. See also [273].

Hint to Ex. 26. It is equivalent to the existence of $\mathbf{w} \in L^p$ with $\nabla \cdot \mathbf{w} = \mu - \nu$.

Hint to Ex. 27. Solve the minimization problem

$$\min \left\{ i_{\tilde{Q}}(\Omega) \,:\, (e_0)_\# \tilde{Q} = \mu,\ (e_1)_\# \tilde{Q} = \nu,\ i_{\tilde{Q}} \leq i_Q \right\}.$$

Hint to Ex. 28. Make the ansatz that the mass from A to A' will be split into three parts, α passing through B, α through B', and β directly to A', and the same for B to B'. Write a 2×2 system for the equilibrium condition.

Hint to Ex. 29. Take a regular grid on a cube contained in Ω. Count how much does it cost to move the mass of the central half of each cube out of the cube, and sum up.

Hint to Ex. 30. Use the angle law (4.31).

Hint to Ex. 31. Suppose $\phi \geq 0$, and use the k-Lipschitz functions ϕ_k defined in the memo 1.5 on l.s.c. functions.

Hint to Ex. 32. Prove that the optimal ν is the image of f through the projection onto Σ. Use Beckmann's formulation.

Hint to Ex. 33. Use Lemma 5.5 to modify the marginals of γ.

Hint to Ex. 34. Differentiate W_2^2.

Hint to Ex. 35. Prove Γ-convergence.

Hint to Ex. 36. Take a point $x_0 \in \mathrm{spt}(\mu)$ realizing $\max d(x, \mathrm{spt}(\nu))$: it must be transported somewhere.

Hint to Ex. 37. Build a counterexample with a map T which is a gradient but such that $\nabla h(\mathrm{T}(x) - x)$ is not a gradient for $h(z) = |z|^4$.

Hint to Ex. 38. Adapt the proof of Lemma 5.33.

Hint to Ex. 39. Write optimality conditions when we move one point.

Hint to Ex. 40. Consider the entries of an extremal matrix which are in $]0,1[$. Prove that if there is one such an entry, then there is a "cycle" and find a contradiction.

Hint to Ex. 41. Same strategy as above.

Hint to Ex. 42. From $\log \gamma_{ij}^\varepsilon = \log p_i^\varepsilon + \log q_j^\varepsilon + \log \eta_{ij}$ deduce $\varepsilon \log p_i^\varepsilon + \varepsilon \log q_j^\varepsilon \le c_{ij}$, with equality if at the limit $\gamma_{ij} > 0$.

Hint to Ex. 43. Write $\mathrm{cof}(D^2u)D^2u = \det(D^2u)\mathrm{I}$ and differentiate.

Hint to Ex. 44. Compute the second derivative and integrate by parts to get $\int u_{ij} v^i v^j f$.

Hint to Ex. 45. First part: add and subtract suitable functionals, continuous for W_1, to make the integrand positive, and restrict it to bounded domains. Second part: prove that a bound on the second moment implies W_1 convergence.

Hint to Ex. 46. Use the fact that the L^∞ norm is bounded by 1 on a geodesic whenever this is the case on the starting and arrival measures.

Hint to Ex. 47. Use the direct method for existence, strict convexity for uniqueness, and let the potential appear in optimality conditions.

Hint to Ex. 48. Difficult: prove a sort of interior ball condition, i.e., if $\varrho(x_0) > g(x_0)$, then $\varrho = f$ on $B(T(x_0), |x_0 - T(x_0)|)$, and use Lebesgue points. See [151].

Hint to Ex. 49. Use $\int \frac{1}{2}|\nabla \varrho|^2 + 4\sqrt{\varrho}$.

Hint to Ex. 50. Consider the optimal γ between ϱ_0 and ϱ: if ϱ has more than N atoms, then replacing each $\gamma \llcorner \big(\{x_i\} \times \mathbb{R}^d\big)$ with its barycenter gives a better result.

Hint to Ex. 51. Prove that the barycenter is preserved and use the equation to compute E'.

Hint to Ex. 52. Write the optimality conditions in terms of the potential (direct method for existence, strict convexity for uniqueness).

Hint to Ex. 53. Estimate the ∇u_ϱ part in terms of $||\varrho||_{L^m}^2$, and use $m > 2$ to see that these norms must be bounded on a minimizing sequence.

Hint to Ex. 54. Write the problem as a min-max, introducing Lagrange multiplier for the constraints, and deduce the conditions to have a saddle point.

Hint to Ex. 55. Write and solve the Euler-Lagrange equation.

Hint to Ex. 56. Prove that $t \mapsto ||\varrho_t - 1||_{H^{-1}}^2$ is BV, by using Young inequality and Lemma 5.33.

Hint to Ex. 57. Check that moving every atom of ν on the (a) closest atom of ν is possible.

Hint to Ex. 58. Use concave costs and measures with a common mass.

Hint to Ex. 59. Try to realize the cost 0.

Hint to Ex. 60. Use μ and ν as in the counterexample of Section 1.4, and choose $c(x, y)$ as a discontinuous function of $|x - y|$.

Hint to Ex. 61. Adapt the Peano construction for a surjective map onto the square, defining a map as a fixed point of a contracting operator.

Hint to Ex. 62. Use the characterization of optimal maps as gradients of convex functions (quadratic). Find a nonoptimal transport between two measures with compact support on $\mathbb{R}$ and translate them far away (linear cost). Find a composition of two gradients which is not a gradient.

Hint to Ex. 63. Use the properties of $F_\mu^{[-1]}$.

Hint to Ex. 64. Write the optimal transport as $x - \varphi'(x)$, check $||1 - g||_{\dot{H}^{-1}} = ||\varphi'||_{L^2}$, and integrate by parts.

Hint to Ex. 65. For $\mathscr{T}_c \geq 0$, use the plan $\gamma = \mu \otimes \nu$ to have a bound. For the support of γ, take (x, y) and (x', y') in $\mathrm{spt}(\gamma)$, write c-cyclical monotonicity, and integrate w.r.t. (x', y'). This gives an upper bound that allows to prove upper semi-continuity.

Hint to Ex. 66. Use displacement convexity of the entropy, compare the derivative of the entropy at $t = 0$ and $t = 1$ on a geodesic.

Hint to Ex. 67. For the first question, differentiate W_2^2 and use the **Ex**(66). For the second, use the inequality in Lemma 5.2 and the semigroup properties of the heat equation.

Hint to Ex. 68. After differentiating W_2^2, use Gronwall. H^1 is reasonable because the H^1 norm is controlled in time.

Hint to Ex. 69. Differentiate W_2^2 between the solution of the evolution equation and the static solution.

References

1. M. Agueh, G. Carlier, Barycenters in the Wasserstein space. SIAM J. Math. Ann. **43**(2), 904–924 (2011)
2. R.K. Ahuja, T. Magnanti, J. Orlin, *Network Flows: Theory, Algorithms, and Applications* (Prentice Hall, Upper Saddle River, 1993)
3. G. Alberti, On the structure of singular sets of convex functions. Calc.Var. Part. Differ. Equat. **2**, 17–27 (1994)
4. G. Alberti, L. Ambrosio, A geometrical approach to monotone functions in $\mathbb{R}^d$. Math. Z. **230**, 259–316 (1999)
5. G. Alberti, L. Ambrosio, P. Cannarsa, On the singularities of convex functions. Manuscripta Math. **76**, 421–435 (1992)
6. A.D. Aleksandrov, Almost everywhere existence of the second differential of a convex functions and related properties of convex surfaces. Uchenye Zapisky Leningrad. Gos. Univ. Math. Ser. **37**, 3–35 (1939) (in Russian)
7. L. Ambrosio, Movimenti minimizzanti. Rend. Accad. Naz. Sci. XL Mem. Mat. Sci. Fis. Natur. **113**, 191–246 (1995)
8. L. Ambrosio, Lecture notes on optimal transport problems, in *Mathematical Aspects of Evolving Interfaces*. Lecture Notes in Mathematics (1812) (Springer, New York, 2003), pp. 1–52
9. L. Ambrosio, N. Gigli, A user's guide to optimal transport, in *Modelling and Optimisation of Flows on Networks*. Lecture Notes in Mathematics (2013), Springer Berlin Heidelberg, pp. 1–155
10. L. Ambrosio, A. Pratelli, Existence and stability results in the L^1 theory of optimal transportation, in *Optimal Transportation and Applications*, ed. by L.A. Caffarelli, S. Salsa. Lecture Notes in Mathematics (CIME Series, Martina Franca, 2001) 1813 (2003), Springer Berlin Heidelberg, pp. 123–160
11. L. Ambrosio, G. Savaré, Gradient flows of probability measures, *Handbook of Differential Equations, Evolutionary Equations*, ed. by C.M. Dafermos, E. Feireisl, vol. 3 (Elsevier, Amsterdam, 2007)
12. L. Ambrosio, P. Tilli, *Topics on Analysis in Metric Spaces*. Oxford Lecture Series in Mathematics and its Applications, vol. 25 (Oxford University Press, Oxford, 2004)
13. L. Ambrosio, N. Fusco, D. Pallara, *Functions of Bounded Variation and Free Discontinuity Problems* (Oxford University Press, Oxford, 2000)

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2

14. L. Ambrosio, B. Kirchheim, A. Pratelli, Existence of optimal transport maps for crystalline norms, Duke Math. J. **125**, 207–241 (2004)
15. L. Ambrosio, N. Gigli, G. Savaré, *Gradient Flows in Metric Spaces and in the Spaces of Probability Measures.* Lectures in Mathematics, ETH Zurich (Birkhäuser, Basel, 2005)
16. L. Ambrosio, M. Colombo, G. De Philippis, A. Figalli, Existence of Eulerian solutions to the semigeostrophic equations in physical space: the 2-dimensional periodic case Commun. Part. Differ. Equat. **37**(12), 2209–2227 (2012)
17. L. Ambrosio, M. Colombo, G. De Philippis, A. Figalli, A global existence result for the semigeostrophic equations in three dimensional convex domains. Discr. Contin. Dyn. Syst. **34**(4), 1251–1268 (2013)
18. L. Ambrosio, N. Gigli, G. Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below Inv. Math. **195**(2), 289–391 (2014)
19. S. Angenent, S. Haker, A. Tannenbaum, Minimizing flows for the Monge–Kantorovich problem. SIAM J. Math. Ann. **35**(1), 61–97 (2003)
20. S. Angenent, S. Haker, A. Tannenbaum, L. Zhu, Optimal transport for registration and warping. Int. J. Comput. Vis. **60**(3), 225–240 (2004)
21. P. Appell, Mémoire sur les déblais et les remblais de systèmes continus ou discontinus, *Mémoires présentés par divers savants à l'Académie royale des sciences de l'Institut de France. Sciences mathématiques et physiques.* 1827–1914 (2e s. I-XXXV) (1887)
22. V. Arnold, Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. (French) Ann. Inst. Fourier (Grenoble) **16**(1), 319–361 (1996)
23. H. Attouch, G. Buttazzo, G. Michaille, *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization* (SIAM, Philadelphia, 2006)
24. F. Aurenhammer, F. Hoffmann, B. Aronov, Minkowski-type theorems and least-squares clustering. Algorithmica **20**, 61–76 (1998)
25. R. Balka, Y. Peres, Restrictions of Brownian motion, preprint. C. R. Math. Acad. Sci. Paris **352**(12), 1057–1061 (2014)
26. M. Bardelloni, S. Bianchini, The decomposition of optimal transportation problems with convex cost. (2014) http://arxiv.org/pdf/1409.0515
27. M. Beckmann, A continuous model of transportation. Econometrica **20**, 643–660 (1952)
28. M. Beckmann, T. Puu, *Spatial Economics: Density, Potential and Flow* (North-Holland, Amsterdam, 1985)
29. M. Beckmann, C. McGuire, C. Winsten, *Studies in Economics of Transportation* (Yale University Press, New Haven, 1956)
30. M. Beiglböck, Cyclical monotonicity and the ergodic theorem. Ergodic Theory Dyn. Syst. **35**(3), 710–713 (2015)
31. M. Beiglböck, N. Juillet, On a problem of optimal transport under marginal martingale constraints. Ann. Probab. (2012) http://arxiv.org/abs/1208.1509
32. M. Beiglböck, P. Henry-Labordère, F. Penkner, Model-independent bounds for option prices—a mass transport approach. Fin. Stoch. **17**(3), 477–501 (2013)
33. J.-D. Benamou, Y. Brenier, Weak existence for the semigeostrophic equations formulated as a coupled Monge-Ampère/transport problem. SIAM J. Appl. Math. **58**, 1450–1461 (1998)
34. J.-D. Benamou, Y. Brenier, A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numer. Math. **84**, 375–393 (2000)
35. J.-D. Benamou, G. Carlier, Augmented Lagrangian methods for transport optimization, mean-field games and degenerate PDEs. (2014) https://hal.inria.fr/hal-01073143
36. J.-D. Benamou, Y. Brenier, K. Guittet, The Monge-Kantorovich mass transfer and its computational fluid mechanics formulation. Int. J. Numer. Methods Fluids **40**(1–2), 21–30 (2002)
37. J.-D. Benamou, Y. Brenier, K. Guittet, Numerical analysis of a multi-phasic mass transport problem. Contemp. Math. **353**, 1–18 (2004)

38. J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, G. Peyré, Iterative Bregman projections for regularized transportation problems. SIAM J. Sci. Comput. **37**(2), A1111–A1138 (2015).
39. J.-D. Benamou, B. Froese, A. Oberman, Two numerical methods for the elliptic Monge-Ampere equation. ESAIM: Math. Model. Numer. Ann. **44**(4), 737–758 (2010)
40. J.-D. Benamou, G. Carlier, Q. Mérigot, É. Oudet, Discretization of functionals involving the Monge-Ampère operator. (2014) http://arxiv.org/abs/1408.4536
41. J.-D. Benamou, B. Froese, A. Oberman, Numerical solution of the Optimal Transportation problem using the Monge–Ampère equation. J. Comput. Phys. **260**, 107–126 (2014)
42. F. Benmansour, G. Carlier, G. Peyré, F. Santambrogio, Numerical approximation of continuous traffic congestion equilibria. Net. Het. Media **4**(3), 605–623 (2009)
43. F. Benmansour, G. Carlier, G. Peyré, F. Santambrogio, Fast marching derivatives with respect to metrics and applications. Numer. Math. **116**(3), 357–381 (2010)
44. P. Bernard, B. Buffoni, The Monge problem for supercritical Mané potentials on compact manifolds. Adv. Math. **207**(2), 691–706 (2006)
45. F. Bernardeau, S. Colombi, E. Gaztanaga, R. Scoccimarro, *Large-Scale Structure of the Universe and Cosmological Perturbation Theory*. Phys. Rep. **367**, 1–248 (2002)
46. M. Bernot, Optimal transport and irrigation. Ph.D. Thesis, ENS Cachan (2005) http://perso.crans.org/bernot
47. M. Bernot, A. Figalli, Synchronized traffic plans and stability of optima. ESAIM Control Optim. Calc. Var. **14**, 864–878 (2008)
48. M. Bernot, V. Caselles, J.-M. Morel, Traffic plans. Publ. Math. **49**(2), 417–451 (2005)
49. M. Bernot, V. Caselles, J.-M. Morel, The structure of branched transportation networks. Calc. Var. Part. Differ. Equat. **32**(3), 279–317 (2008)
50. M. Bernot, V. Caselles, J.-M. Morel, *Optimal Transportation Networks, Models and Theory*. Lecture Notes in Mathematics, vol. 1955 (Springer, New York, 2008)
51. M. Bernot, A. Figalli, F. Santambrogio, Generalized solutions for the Euler equations in one and two dimensions. J. Math. Pures et Appl. **91**(2), 137–155 (2009)
52. J. Bertrand, M. Puel, The optimal mass transport problem for relativistic costs. Calc. Var. PDE **46**(1–2), 353–374 (2013)
53. D.P. Bertsekas, A distributed algorithm for the assignment problem, in *Lab. for Information and Decision Systems Working Paper* (MIT, Cambridge, 1979)
54. D.P. Bertsekas, Auction Algorithms, Encyclopedia of Optimization, Kluwer, (2001)
55. D.P. Bertsekas, J. Eckstein, Dual coordinate step methods for linear network flow problems. Math. Program. **42**(1–3), 203–243 (1988)
56. A. Blanchet, G. Carlier, Optimal transport and Cournot-Nash equilibria. Math. Oper. Res. (to appear) https://www.ceremade.dauphine.fr/~carlier/publis.html
57. A. Blanchet, V. Calvez, J.A. Carrillo, Convergence of the mass-transport steepest descent scheme for the subcritical Patlak-Keller-Segel model. SIAM J. Numer. Anal. **46**(2), 691–721 (2008)
58. A. Blanchet, J.-A. Carrillo, D. Kinderlehrer, M. Kowalczyk, P. Laurençot, S. Lisini, A hybrid variational principle for the Keller-Segel system in $\mathbb{R}^2$. ESAIM M2AN (2015). http://www.math.cmu.edu/math/faculty/kinderlehrer.php
59. A. Blanchet, P. Mossay, F. Santambrogio, Existence and uniqueness of equilibrium for a spatial model of social interactions. Int. Econ. Rev. (2014). http://cvgmt.sns.it/paper/1998/
60. S. Bianchini, L. Caravenna, On optimality of c-cyclically monotone transference plans. C. R. Math. Acad. Sci. Paris **348**(11–12), 613–618 (2010)
61. S. Bianchini, M. Gloyer, On the Euler-Lagrange equation for a variational problem: the general case II. Math. Zeit. **265**(4), 889–923 (2009)
62. G. Blower, Displacement convexity for the generalized orthogonal ensemble. J. Statist. Phys. **116**(5-6), 1359–1387 (2004)
63. V.I. Bogachev, A.V. Kolesnikov, The Monge–Kantorovich problem: achievements, connections, and perspectives. Russ. Math. Surv. **67**(5), 785–890 (2012)

64. J.F. Bonnans, J.-C. Gilbert, C. Lemarechal, C. Sagastizábal, *Numerical Optimization: Theoretical and Practical Aspects*, 2nd edn. (Springer, New York, 2006)
65. N. Bonnotte, Unidimensional and evolution methods for optimal transportation. Ph.D. Thesis, Université Paris-Sud, 2013
66. N. Bonnotte, From Knothe's rearrangement to Brenier's optimal transport map. SIAM J. Math. Anal. **45**(1), 64–87 (2013)
67. G. Bouchitté, G. Buttazzo, New lower semi-continuity results for nonconvex functionals defined on measures. Nonlinear Anal. **15**, 679–692 (1990)
68. G. Bouchitté, G. Buttazzo, Integral representation of nonconvex functionals defined on measures. Ann. Inst. H. Poincaré Anal. Non Linéaire **9**(1), 101–117 (1992)
69. G. Bouchitté, G. Buttazzo, Relaxation for a class of nonconvex functionals defined on measures. Ann. Inst. H. Poincaré Anal. Non Linéaire **10**(3), 345–361 (1993)
70. G. Bouchitté, G. Buttazzo, Characterization of optimal shapes and masses through Monge-Kantorovich equation. J. Eur. Math. Soc. **3**(2), 139–168 (2001)
71. G. Bouchitté, G. Buttazzo, P. Seppecher, Shape optimization solutions via Monge-Kantorovich equation. C. R. Acad. Sci. Paris Sér. I Math. **324**(10), 1185–1191 (1997)
72. G. Bouchitté, C. Jimenez, M. Rajesh, Asymptotique d'un problème de positionnement optimal. C. R. Acad. Sci. Paris Ser. I **335**, 1–6 (2002)
73. G. Bouchitté, T. Champion, C. Jimenez, Completion of the space of measures in the Kantorovich norm, proc. of "Trends in the Calculus of Variations", Parma, 2004, E.D. Acerbi and G.R. Mingione Editors. Rivi. Mat. Univ. Parma Ser. **7**(4), 127–139 (2005)
74. G. Bouchitté, C. Jimenez, M. Rajesh, A new L^∞ estimate in optimal mass transport. Proc. Am. Math. Soc. **135**, 3525–3535 (2007)
75. G. Bouchitté, C. Jimenez, M. Rajesh, Asymptotic analysis of a class of optimal location problems. J. Math. Pures Appl. **95**(4), 382–419 (2011)
76. D. Braess, Über ein Paradoxon aus der Verkehrsplanung. Unternehmensforschung **12**, 258–268 (1969)
77. A. Braides, *Γ-Convergence for Beginners* (Oxford University Press, Oxford, 2002)
78. A. Brancolini, S. Solimini, On the holder regularity of the landscape function. Interfaces Free Boundaries **13**(2), 191–222 (2011)
79. L. Brasco, F. Santambrogio, An equivalent path functional formulation of branched transportation problems. Discr. Contin. Dyn. Syst. A **29**(3), 845–871 (2011)
80. L. Brasco, G. Carlier, F. Santambrogio, Congested traffic dynamics, weak flows and very degenerate elliptic equations. J. Math. Pures et Appl. **93**(6), 652–671 (2010)
81. L. Brasco, G. Buttazzo, F. Santambrogio, A Benamou-Brenier approach to branched transportation. SIAM J. Math. Ann. **43**(2), 1023–1040 (2011)
82. Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs. (French) C. R. Acad. Sci. Paris Sér. I Math. **305**(19), 805–808 (1987)
83. Y. Brenier, The least action principle and the related concept of generalized flows for incompressible perfect fluids. J. Am. Mat. Soc. **2**, 225–255 (1989)
84. Y. Brenier, Polar factorization and monotone rearrangement of vector-valued functions. Commun. Pure Appl. Math. **44**, 375–417 (1991)
85. Y. Brenier, Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations. Commun. Pure Appl. Math. **52**, 411–452 (1999)
86. Y. Brenier, Optimal transport, convection, magnetic relaxation and generalized Boussinesq equations. J. Nonlinear Sci. **19**(5), 547–570 (2009)
87. Y. Brenier, A modified least action principle allowing mass concentrations for the early universe reconstruction problem. Confluentes Mathematici **3**(3), 361–385 (2011)
88. Y. Brenier, Rearrangement, convection, convexity and entropy. Philos. Trans. R. Soc. A **371**, 20120343 (2013)
89. Y. Brenier, U. Frisch, M. Hénon, G. Loeper, S. Matarrese, R. Mohayaee, A. Sobolevskii, Reconstruction of the early Universe as a convex optimization problem. Mon. Not. R. Astron. Soc. **346**, 501–524 (2003)
90. H. Brezis, *Analyse fonctionnelle, Théorie et applications* (Masson, Paris, 1983)

91. H. Brezis, Liquid crystals and energy estimates for $\mathbb{S}^2$-valued maps, in *Theory and Applications of Liquid Crystals*. The IMA Volumes in Mathematics and its Applications, vol. 5 (Springer, New York, 1987), pp. 31–52
92. M. Burger, M. Franek, C.-B. Schönlieb, Regularized regression and density estimation based on optimal transport. Appl. Math. Res. Express **2012**(2), 209–253 (2012)
93. G. Buttazzo, F. Santambrogio, A model for the optimal planning of an urban area. SIAM J. Math. Anal. **37**(2), 514–530 (2005)
94. G. Buttazzo, F. Santambrogio, A mass transportation model for the optimal planning of an Urban region. SIAM Rev. **51**(3), 593–610 (2009)
95. G. Buttazzo, É. Oudet, E. Stepanov, Optimal transportation problems with free Dirichlet regions, in *Variational Methods for Discontinuous Structures*. PNLDE, vol. 51 (Birkhäuser, Basel, 2002), pp. 41–65
96. G. Buttazzo, C. Jimenez, É. Oudet, An optimization problem for mass transportation with congested dynamics. SIAM J. Control Optim. **48**, 1961–1976 (2010)
97. G. Buttazzo, L. De Pascale, P. Gori-Giorgi, Optimal transport meets electronic density functional theory. Phys. Rev. A **85**(6), 062502 (2012)
98. L. Caffarelli, A localization property of viscosity solutions to the Monge-Ampère equation and their strict convexity. Ann. Math. **131**(1), 129–134 (1990)
99. L. Caffarelli, Interior $W^{2,p}$ estimates for solutions of the Monge-Ampère equation. Ann. Math. **131**(1), 135–150 (1990)
100. L. Caffarelli, Some regularity properties of solutions of Monge Ampère equation. Commun. Pure Appl. Math. **44**(8–9), 965–969 (1991)
101. L. Caffarelli, M. Feldman, R. McCann, Constructing optimal maps for Monge's transport problem as a limit of strictly convex costs. J. Am. Math. Soc. **15**, 1–26 (2002)
102. P. Cannarsa, C. Sinestrari, *Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control* (Birkhäuser, Basel, 2004)
103. L. Caravenna, A proof of Sudakov theorem with strictly convex norms. Math. Z. **268**, 371–407 (2011)
104. L. Caravenna, S. Daneri, The disintegration of the Lebesgue measure on the faces of a convex function. J. Funct. Anal. **258**(11), 3604–3661 (2010)
105. P. Cardaliaguet, Notes on mean field games (from P.-L. Lions' lectures at Collège de France). (2013) https://www.ceremade.dauphine.fr/~cardalia/
106. G. Carlier, I. Ekeland, The structure of cities. J. Global Optim. **29**, 371–376 (2004)
107. G. Carlier, I. Ekeland, Equilibrium structure of a bidimensional asymmetric city. Nonlinear Anal. Real World Appl. **8**(3), 725–748 (2007)
108. G. Carlier, I. Ekeland, Matching for teams. Econ. Theory **42**(2), 397–418 (2010)
109. G. Carlier, B. Nazaret, Optimal transportation for the determinant. ESAIM Control Optim. Calc. Var. **14**(4), 678–698 (2008)
110. G. Carlier, F. Santambrogio, A variational model for urban planning with traffic congestion. ESAIM Control Optim. Calc. Var. **11**(4), 595–613 (2005)
111. G. Carlier, F. Santambrogio, A continuous theory of traffic congestion and Wardrop equilibria, proceedings of the conference, proceedings of optimization and stochastic methods for spatially distributed information, St Petersburg, 2010, published (English version). J. Math. Sci. **181**(6), 792–804 (2012)
112. G. Carlier, C. Jimenez, F. Santambrogio, Optimal transportation with traffic congestion and Wardrop equilibria. SIAM J. Control Optim. **47**, 1330–1350 (2008)
113. G. Carlier, L. De Pascale, F. Santambrogio, A strategy for non-strictly convex transport costs and the example of $||x - y||^p$ in $\mathbb{R}^2$. Commun. Math. Sci. **8**(4), 931–941 (2010)
114. G. Carlier, A. Galichon, F. Santambrogio, From Knothe's transport to Brenier's map and a continuation method for optimal transport. SIAM J. Math. Ann. **41**(6), 2554–2576 (2010)
115. J.-A. Carrillo, D. Slepčev, Example of a displacement convex functional of first order Calc. Var. Part. Differ. Equat. **36**(4), 547–564 (2009)

116. J.-A. Carrillo, R.J. McCann, C. Villani, Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. Rev. Math. Iberoam. **19**, 1–48 (2003)
117. J.-A. Carrillo, R.J. McCann, C. Villani, Contractions in the 2-Wasserstein length space and thermalization of granular media. Arch. Ration. Mech. Ann. **179**, 217–263 (2006)
118. J.-A. Carrillo, M. Di Francesco, A. Figalli, T. Laurent, D. Slepčev, Global-in-time weak measure solutions and finite-time aggregation for nonlocal interaction equations. Duke Math. J. **156**(2), 229–271 (2011)
119. T. Champion, L. De Pascale, The Monge problem for strictly convex norms in $\mathbb{R}^d$. J. Eur. Math. Soc. **12**(6), 1355–1369 (2010)
120. T. Champion, L. De Pascale, The Monge problem in $\mathbb{R}^d$. Duke Math. J. **157**(3), 551–572 (2011)
121. T. Champion, L. De Pascale, On the twist condition and c-monotone transport plans. Discr. Contin. Dyn. Syst. **34**(4), 1339–1353 (2014)
122. T. Champion, L. De Pascale, P. Juutinen, The ∞-Wasserstein distance: local solutions and existence of optimal transport maps. SIAM J. Math. Ann. **40**(1), 1–20 (2008)
123. P. Chen, F. Jiang, X.-P. Yang, Two dimensional optimal transportation for a distance cost with a convex constraint. ESAIM: COCV **19**(4), 1064–1075 (2013)
124. P. Chen, F. Jiang, X.-P. Yang, Optimal transportation in $\mathbb{R}^d$ for a distance cost with convex constraint. Zeitschrift fuer Angewandte Mathematik und Physik, **66**(3), 587–606 (2015)
125. P.-A. Chiappori, R.J. McCann, L.P. Nesheim, Hedonic price equilibria, stable matching, and optimal transport: equivalence, topology, and uniqueness. Econ. Theory **42**(2), 317–354 (2010)
126. M. Colombo, A. Figalli, Regularity results for very degenerate elliptic equations. J. Math. Pures Appl. **101**(1), 94–117 (2014)
127. D. Cordero-Erausquin, Sur le transport de mesures périodiques. C. R. Acad. Sci. Paris Sér. I Math. **329**(3), 199–202 (1999)
128. C. Cotar, G. Friesecke, C. Klüppelberg, Density functional theory and optimal transportation with coulomb cost. Commun. Pure Appl. Math. **66**(4), 548–599 (2013)
129. C. Cotar, G. Friesecke, C. Klüppelberg, Smoothing of transport plans with fixed marginals and rigorous semiclassical limit of the Hohenberg-Kohn functional (in preparation)
130. M.J.P. Cullen, *A Mathematical Theory of Large-Scale Atmosphere/Ocean Flow* (Imperial College Press, London, 2006)
131. M. J. P. Cullen, W. Gangbo, A variational approach for the 2-dimensional semi-geostrophic shallow water equations. Arch. Ration. Mech. Ann. **156**(3), 241–273 (2001)
132. B. Dacorogna, J. Moser, On a partial differential equation involving the Jacobian determinant. Ann. Inst. H. Poincaré Anal. Non Linéaire **7**(1), 1–26 (1990)
133. G. Dal Maso, *An Introduction to Γ-Convergence* (Birkhauser, Basel, 1992)
134. S. Daneri, G. Savaré, Eulerian calculus for the displacement convexity in the Wasserstein distance. SIAM J. Math. Ann. **40**, 1104–1122 (2008)
135. G.B. Dantzig, Maximization of a linear function of variables subject to linear inequalities, in *Activity Analysis of Production and Allocation* (Wiley, New York, 1951), pp. 339–347
136. G.B. Dantzig, Linear programming, in *History of Mathematical Programming: A Collection of Personal Reminiscences*, ed. by J.K. Lenstra, A.H.G. Rinnooy Kan, A. Schrijver (CWI, Amsterdam, 1991)
137. E. De Giorgi, New problems on minimizing movements, *Boundary Value Problems for PDE and Applications*, ed. by C. Baiocchi, J.L. Lions (Masson, Paris, 1993), pp. 81–98
138. E. De Giorgi, T. Franzoni, Su un tipo di convergenza variazionale. Atti Acc. Naz. Lincei Rend. **58**(8), 842–850 (1975)
139. F. Delbaen, W. Schachermayer, What is... a free lunch? Not. Am. Math. Soc. **51**(5), 526–528 (2004)
140. J. Delon, J. Salomon, A. Sobolevskii, Fast transport optimization for Monge costs on the circle. SIAM J. Appl. Math. **70**(7), 2239–2258 (2010)

141. J. Delon, J. Salomon, A. Sobolevskii, Local matching indicators for transport problems with concave costs. SIAM J. Discr. Math. **26**(2), 801–827 (2012)
142. W.E. Deming, F.F. Stephan, On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. Ann. Math. Stat. **11**(4), 427–444 (1940)
143. L. De Pascale, A. Pratelli, Regularity properties for monge transport density and for solutions of some shape optimization problem. Calc. Var. Part. Differ. Equat. **14**(3), 249–274 (2002)
144. L. De Pascale, A. Pratelli, Sharp summability for Monge transport density via interpolation. ESAIM Control Optim. Calc. Var. **10**(4), 549–552 (2004)
145. L. De Pascale, L.C. Evans, A. Pratelli, Integral estimates for transport densities. Bull. Lond. Math. Soc. **36**(3), 383–385 (2004)
146. L. De Pascale, J. Louet, F. Santambrogio, The Monge problem with vanishing gradient penalization: vortices and asymptotical profile. (2015) http://cvgmt.sns.it/paper/2449/
147. G. De Philippis, A. Figalli, $W^{2,1}$ regularity for solutions of the Monge-Ampère equation. Inv. Math. **192**(1), 55–69 (2013)
148. G. De Philippis, A. Figalli, Sobolev regularity for Monge-Ampère type equations. SIAM J. Math. Ann. **45**(3), 1812–1824 (2013)
149. G. De Philippis, A Figalli, The Monge–Ampère equation and its link to optimal transportation. Bull. Am. Math. Soc. **51**(4), 527–580 (2014)
150. G. De Philippis, A. Figalli, O. Savin, A note on interior $W^{2,1+\varepsilon}$ estimates for the Monge–Ampère equation. Math. Ann. **357**(1), 11–22 (2013)
151. G. De Philippis, A. Mészáros, F. Santambrogio, B. Velichkov, BV estimates in optimal transportation and applications. To appear in Archives of Rational Mechanics and Analysis (2015) http://cvgmt.sns.it/paper/2559/
152. C. Dellacherie, P.-A. Meyer, *Probabilities and Potential*. North-Holland Mathematics Studies, vol. 29 (North-Holland, Amsterdam/New York, 1978)
153. G. Devillanova, S. Solimini, On the dimension of an irrigable measure. Rend. Semin. Mat. Univ. Padova **117**, 1–49 (2007)
154. G. Devillanova, S. Solimini, Elementary properties of optimal irrigation patterns. Calc. Var. Part. Differ. Equat. **28**(3), 317–349 (2007)
155. S. Di Marino, A. Mészáros, Uniqueness issues for evolutive equations with density constraints (2015) http://cvgmt.sns.it/paper/2734/
156. R.J. DiPerna, P.-L. Lions, Ordinary differential equations, transport theory and Sobolev spaces. Invent. Math. **98**, 511–547 (1989)
157. I. Ekeland, R. Temam, *Convex Analysis and Variational Problems*. Classics in Applied Mathematics, vol. 28 (SIAM, Philadelphia, 1999)
158. L.C. Evans, *Partial Differential Equations*. Graduate Studies in Mathematics (American Mathematical Society, Providence, 2010)
159. L.C. Evans, W. Gangbo, Differential equations methods for the Monge-Kantorovich mass transfer problem. Mem. Am. Math. Soc. **137**(653) (1999)
160. L.C. Evans, R.F. Gariepy, *Measure Theory and Fine Properties of Functions*. Studies in Advanced Mathematics (CRC, Boca Raton, 1992)
161. H. Federer, *Geometric Measure Theory*. Classics in Mathematics (Springer, New York, 1996 (reprint of the 1st edn. Berlin, Heidelberg, New York 1969 edition)
162. M. Feldman, R. McCann, Uniqueness and transport density in Monge's mass transportation problem. Calc. Var. Par. Differ. Equat. **15**(1), 81–113 (2002)
163. S. Ferradans, N. Papadakis, G. Peyré, J-F. Aujol, Regularized discrete optimal transport. SIAM J. Imag. Sci. **7**(3), 1853–1882 (2014)
164. A. Figalli, Regularity of optimal transport maps (after Ma-Trudinger-Wang and Loeper), in *Proceedings of the Bourbaki Seminar* (2009)
165. A. Figalli, N. Gigli, A new transportation distance between non-negative measures, with applications to gradients flows with Dirichlet boundary conditions. J. Math. Pures et Appl. **94**(2), 107–130 (2010)
166. A. Figalli, N. Juillet, Absolute continuity of Wasserstein geodesics in the Heisenberg group. J. Funct. Anal. **255**(1), 133–141 (2008)

167. A. Figalli, F. Maggi, A. Pratelli, A mass transportation approach to quantitative isoperimetric inequalities. Invent. Math. **182**(1), 167–211 (2010)
168. A. Figalli, Y.-H. Kim, R.J. McCann, When is multidimensional screening a convex program? J. Econ. Theory **146**(2), 454–478 (2011)
169. L.R. Ford Jr., D.R. Fulkerson, *Flows in Networks* (Princeton University Press, Princeton, 1962)
170. M. Fortin, R. Glowinski, *Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems*. Studies in Mathematics and its Applications (North-Holland, Amsterdam, 1983)
171. I. Fragalà, M.S. Gelli, A. Pratelli, Continuity of an optimal transport in Monge problem. J. Math. Pures Appl. **84**(9), 1261–1294 (2005)
172. U. Frisch, S. Matarrese, R. Mohayaee, A. Sobolevski, A reconstruction of the initial conditions of the Universe by optimal mass transportation. Nature **417**, 260–262 (2002)
173. A. Galichon, P. Henry-Labordère, N. Touzi, A stochastic control approach to no-arbitrage bounds given marginals, with an application to lookback options. Ann. Appl. Probab. **24**(1), 312–336 (2014)
174. W. Gangbo, An elementary proof of the polar factorization of vector-valued functions. Arch. Ration. Mech. Anal. **128**, 381–399 (1994)
175. W. Gangbo, The Monge mass transfer problem and its applications. Contemp. Math. **226**, 79–104 (1999)
176. W. Gangbo, R. McCann, The geometry of optimal transportation. Acta Math. **177**, 113–161 (1996)
177. W. Gangbo, A. Święch, Optimal maps for the multidimensional Monge-Kantorovich problem. Commun. Pure Appl. Math. **51**(1), 23–45 (1998)
178. M. Ghisi, M. Gobbino, The monopolist's problem: existence, relaxation and approximation. Calc. Var. Part. Differ. Equat. **24**(1), 111–129 (2005)
179. N. Ghoussoub, B. Maurey, Remarks on multi-marginal symmetric Monge-Kantorovich problems. Discr. Contin. Dyn. Syst. **34**(4), 1465–1480 (2014)
180. N. Ghoussoub, A. Moameni, Symmetric Monge-Kantorovich problems and polar decompositions of vector fields. Geom. Funct. Ann. **24**(4), 1129–1166 (2014)
181. É. Ghys, Gaspard Monge, Images des Mathématiques, CNRS (2012), http://images.math.cnrs.fr/Gaspard-Monge,1094.html
182. N. Gigli, On the inverse implication of Brenier-McCann theorems and the structure of $(P_2(M), W_2)$. Methods Appl. Anal. **18**(2), 127–158 (2011)
183. N. Gigli, K. Kuwada, S. Ohta, Heat flow on Alexandrov spaces. Commun. Pure Appl. Math. **66**(3), 307–33 (2013)
184. D. Gilbarg, N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order*. Grundlehren der Mathematischen Wissenschaften, vol. 224 (Springer, Berlin, 1977)
185. E.N. Gilbert, Minimum cost communication networks. Bell Syst. Tech. J. **46**, 2209–2227 (1967)
186. E.N. Gilbert, H.O. Pollak, Steiner minimal trees. SIAM J. Appl. Math. **16**, 1–29 (1968)
187. S. Graf, H. Luschgy, *Foundations of Quantization for Probability Distributions* (Springer, Berlin, 2000)
188. K. Guittet, Contributions à la résolution numérique de problèmes de transport optimal de masse. Ph.D. thesis, University of Paris 6, 2003
189. K. Guittet, On the time-continuous mass transport problem and its approximation by augmented lagrangian techniques. SIAM J. Numer. Ann. **41**(1), 382–399 (2004)
190. C. Gutiérrez, *The Monge-Ampère Equation* (Birkhaüser, Basel, 2001)
191. S. Haker, L. Zhu, A. Tannenbaum, S. Angenent, Optimal mass transport for registration and warping. Int. J. Comput. Vis. **60**(3), 225–240 (2004)
192. S. Helgason, *The Radon Transform*. Progress in Mathematics (Springer, New York, 1999)
193. P. Henry-Labordère, N. Touzi, An explicit martingale version of Brenier's theorem. (2015) http://www.cmap.polytechnique.fr/~touzi/

194. R. Hug, N. Papadakis, E. Maitre On the convergence of augmented Lagrangian method for optimal transport between nonnegative densities (2015) https://hal.archives-ouvertes.fr/hal-01128793
195. C. Jimenez, Optimisation de Problèmes de Transport. Ph.D. thesis of Université du Sud-Toulon-Var, 2005
196. C. Jimenez, Dynamic formulation of optimal transport problems. J. Convex Anal. **15**(3), 593–622 (2008)
197. C. Jimenez, F. Santambrogio, Optimal transportation in the quadratic case with a convex constraint. J. Math. Pures Appl. **98**(1), 103–113 (2012)
198. R. Jordan, D. Kinderlehrer, F. Otto, The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. **29**(1), 1–17 (1998)
199. N. Juillet, On displacement interpolation of measures involved in Brenier's Theorem. Proc. Am. Math. Soc. **139**(10), 3623–3632 (2011)
200. L. Kantorovich, On the transfer of masses. Dokl. Acad. Nauk. USSR **37**, 7–8 (1942)
201. E.F. Keller, L.A. Segel, Initiation of slide mold aggregation viewed as an instability. J. Theor. Biol. **26**, 399–415 (1970)
202. E.F. Keller, L.A. Segel, Model for chemotaxis. J. Theor. Biol. **30**, 225–234 (1971)
203. H. Knothe, Contributions to the theory of convex bodies. Mich. Math. J. **4**, 39–52 (1957)
204. H.W. Kuhn, The Hungarian method for the assignment problem. Naval Res. Log. Q. **2**, 83–97 (1955)
205. J.-M. Lasry, P.-L. Lions, Jeux à champ moyen. II. Horizon fini et contrôle optimal. C. R. Math. Acad. Sci. Paris **343**(10), 679–684 (2006)
206. J.-M. Lasry, P.-L. Lions, Mean-field games. Jpn. J. Math. **2**, 229–260 (2007)
207. P.W.Y. Lee, On the Jordan-Kinderlehrer-Otto scheme. J. Math. Anal. Appl. **429**(1), 131–142 (2015)
208. J. Lellmann, D.A. Lorenz, C. Schönlieb, T. Valkonen, Imaging with Kantorovich-Rubinstein discrepancy. SIAM J. Imag. Sci. **7**(4), 2833–2859 (2014)
209. B. Lévy, A numerical algorithm for L^2 semi-discrete optimal transport in 3D. ESAIM M2AN (2015). http://www.loria.fr/~levy/M2AN/transport.pdf
210. Q.R. Li, F. Santambrogio, X.J. Wang, Regularity in Monge's mass transfer problem. J. Math. Pures Appl. **102**(6), 1015–1040 (2014)
211. P.-L. Lions, Series of lectures on mean filed games, Collège de France, Paris, 2006–2012, video-recorderd and available at the web page http://www.college-de-france.fr/site/audio-video/
212. S. Lisini, Characterization of absolutely continuous curves in Wasserstein spaces. Calc. Var. Part. Differ. Equat. **28**, 85–120 (2007)
213. S.P. Lloyd, Least squares quantization in PCM. IEEE Trans. Inf. Theory **IT-28**, 129–137 (1982)
214. G. Loeper, Uniqueness of the solution to the Vlasov-Poisson system with bounded density. J. Math. Pures Appl. **86**(1), 68–79 (2006)
215. G. Loeper, The reconstruction problem for the Euler-Poisson system in cosmology. Arch. Ration. Mech. Anal. **179**(2), 153–216 (2006)
216. G. Loeper, On the regularity of solutions of optimal transportation problems. Acta Math. **202**(2), 241–283 (2009)
217. G. Loeper, F. Rapetti, Numerical solution of the Monge-Ampere equation by a Newton's method. C. R. Acad. Sci. Paris Ser. 1 **340**(4), 319–324 (2005)
218. J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport. Ann. Math. **169**, 903–991 (2009)
219. X.-N. Ma, N.S. Trudinger, X.-J. Wang, Regularity of potential functions of the optimal transportation problem. Arch. Ration. Mech. Anal. **177**(2), 151–183 (2005)
220. F. Maddalena, S. Solimini, Transport distances and irrigation models. J. Conv. Ann. **16**(1), 121–152 (2009)
221. F. Maddalena, S. Solimini, J.-M. Morel, A variational model of irrigation patterns. Interfaces Free Boundaries **5**, 391–416 (2003)

222. J. Malý, L. Zajíček, Approximate differentiation: Jarník points. Fund. Math. **140**(1), 87–97 (1991)
223. B. Maury, J. Venel, Handling of contacts in crowd motion simulations. Traffic Granular Flow **07**, 171–180 (2007)
224. B. Maury, J. Venel, A discrete contact model for crowd motion. ESAIM: M2AN **45**(1), 145–168 (2011)
225. B. Maury, A. Roudneff-Chupin, F. Santambrogio, A macroscopic crowd motion model of gradient flow type. Math. Models Methods Appl. Sci. **20**(10), 1787–1821 (2010)
226. B. Maury, A. Roudneff-Chupin, F. Santambrogio, J. Venel, Handling congestion in crowd motion modeling. Net. Het. Media **6**(3), 485–519 (2011)
227. R.J. McCann, Existence and uniqueness of monotone measure preserving maps. Duke Math. J. **80**, 309–323 (1995)
228. R.J. McCann, N. Guillen, Five lectures on optimal transportation: geometry, regularity and applications. In Analysis and Geometry of Metric Measure Spaces: Lecture Notes of the Seminaire de Mathematiques Superieure (SMS) Montreal 2011. G. Dafni et al, eds. Providence: Amer. Math. Soc. 145–180 (2013)
229. R.J. McCann, A convexity principle for interacting gases. Adv. Math. **128**(1), 153–159 (1997)
230. R.J. McCann, Exact solutions to the transportation problem on the line. Proc. R. Soc. Lond. Ser. A **455**, 1341–1380 (1999)
231. R.J. McCann, Polar factorization of maps on Riemannian manifolds. Geom. Funct. Anal. **11**(3), 589–608 (2001)
232. R. J. McCann, Stable rotating binary stars and fluid in a tube. Houst. J. Math. **32**(2), 603–631 (2006)
233. Q. Mérigot, A multiscale approach to optimal transport. Comput. Graph. Forum 30, 1583–1592 (2011)
234. Q. Mérigot, A comparison of two dual methods for discrete optimal transport, in *Geometric Science of Information*. Lecture Notes in Computer Science, vol. 8085 Springer Berlin Heidelberg (2013), 389–396
235. A. Mészáros, F. Silva, A variational approach to second order Mean Field Games with density constraints: the stationary case. J. Math. Pures Appl. (2015) http://cvgmt.sns.it/paper/2630/
236. V.D. Milman, G. Schechtman, *Asymptotic Theory of Finite-Dimensional Normed Spaces*, with an appendix by M. Gromov, Lecture notes in Mathematics, vol. 1200 (Springer, Berlin, 1986)
237. L. Modica, S. Mortola, Il limite nella Γ-convergenza di una famiglia di funzionali ellittici. Boll. Un. Mat. Ital. A (5) **14**(3), 526–529 (1977)
238. D. Monderer, L.S. Shapley, Potential games. Games Econ. Behav. **14**, 124–143 (1996)
239. G. Monge, Mémoire sur la théorie des déblais et des remblais. Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année, 666–704, (1781)
240. J.-M. Morel, F. Santambrogio, Comparison of distances between measures. Appl. Math. Lett. **20**(4), 427–432 (2007)
241. J.-M. Morel, F. Santambrogio, The regularity of optimal irrigation patterns. Arch. Ration. Mech. Ann. **195**(2), 499–531 (2010)
242. J. Morovic, P. L. Sun, Accurate 3d image colour histogram transformation. Pattern Recogn. Lett. **24**, 1725–1735 (2003)
243. J. Nash, Equilibrium points in n-person games. Proc. Natl. Acad. Sci. **36**(1), 48–49 (1950)
244. J. Nash, Non-cooperative games. Ann. Math. **54**(2), 286–295 (1951)
245. J. B. Orlin, A polynomial time primal network simplex algorithm for minimum cost flows. J. Math. Prog. **78**(2), 109–129 (1997)
246. F. Otto, The geometry of dissipative evolution equations: the porous medium equation. Commun. Part. Differ. Equat. **26**, 101–174 (2011)
247. É. Oudet, F. Santambrogio, A Modica-Mortola approximation for branched transport and applications. Arch. Ration. Mech. Ann. **201**(1), 115–142 (2011)
248. N. Papadakis, G. Peyré, É. Oudet, Optimal transport with proximal splitting. SIAM J. Imag. Sci. **7**(1), 212–238 (2014)

249. B. Pass, Multi-marginal optimal transport: theory and applications. ESAIM: M2AN (2014) http://arxiv.org/abs/1406.0026
250. P. Pegon, D. Piazzoli, F. Santambrogio, Full characterization of optimal transport plans for concave costs. Discr. Contin. Dyn. Syst. – Series A (DCDS-A) **35**(12), 6113–6132 (2015)
251. R. Peyre, Non-asymptotic equivalence between W_2 distance and $\dot{H}^{-1}$ norm. Available at http://arxiv.org/pdf/1104.4631v1.pdf
252. F. Pitié, A. C. Kokaram, R. Dahyot, Automated colour grading using colour distribution transfer. Comput. Vis. Image Underst. **107**(1–2), 123–137 (2007)
253. A. Pratelli, On the equality between Monge's infimum and Kantorovich's minimum in optimal mass transportation. Ann. Inst. H. Poincare' Probab. Stat. **43**(1), 1–13 (2007)
254. A. Pratelli, On the sufficiency of c-cyclical monotonicity for optimality of transport plans. Math. Z. **258**(3), 677–690 (2008)
255. J. Rabin, J. Delon, Y. Gousseau, Regularization of transportation maps for color and contrast transfer, in *Proceedings of IEEE ICIP'10, 1933–1936* (2010)
256. J. Rabin, G. Peyré, J. Delon, M. Bernot, Wasserstein Barycenter and its application to texture mixing, in *Scale Space and Variational Methods in Computer Vision*, ed. by A.M. Bruckstein, B.M. Haar Romeny, A.M. Bronstein, M.M. Bronstein. Lecture Notes in Computer Science, vol. 6667 (Springer, Berlin/Heidelberg, 2012), pp. 435–446
257. S.T. Rachev, L. Rüschendorf, *Mass Transportation Problems. Vol. I: Theory* (Springer, New York, 1998)
258. S.T. Rachev, L. Rüschendorf, *Mass Transportation Problems. Vol. II: Applications* (Springer, New York, 1998)
259. E. Reinhard, M. Adhikhmin, B. Gooch, P. Shirley, Color transfer between images. IEEE Trans. Comput. Graph. Appl. **21**, 34–41 (2001)
260. R.T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, 1970)
261. J.-C. Rochet, A necessary and sufficient condition for rationalizability in a quasilinear context. J. Math. Econ. **16**(2), 191–200 (1987)
262. J.-C. Rochet, P. Choné, Ironing, sweeping, and multidimensional screening. Econometrica **66**(4), 783–826 (1998)
263. I. Rodríguez-Iturbe, A. Rinaldo, *Fractal River Basins: Chance and Self-Organization* (Cambridge University Press, Cambridge, 2001)
264. M. Rosenblatt, Remarks on a multivariate transformation. Ann. Math. Stat. **23**, 470–472 (1952)
265. A. Roudneff-Chupin, Modélisation macroscopique de mouvements de foule. Ph.D. Thesis, Université Paris-Sud (2011). Available at www.math.u-psud.fr/ roudneff/Images/these_roudneff.pdf
266. T. Roughgarden, *Selfish Routing and the Price of Anarchy* (MIT, Cambridge, 2005)
267. E. Rouy, A. Tourin, A viscosity solution approach to shape from shading. SIAM J. Numer. Anal. **29**, 867–884 (1992)
268. W. Rudin, *Real and Complex Analysis*, 3rd edn. (McGraw-Hill, New York, 1986)
269. L. Rüschendorf, On *c*-optimal random variables. Stat. Probab. Lett. **27**, 267–270 (1996)
270. F. Santambrogio, Transport and concentration problems with interaction effects. J. Global Optim. **38**(1), 129–141 (2007)
271. F. Santambrogio, Variational problems in transport theory with mass concentration. Ph.D. thesis, Edizioni della Normale, Birkhäuser, 2007
272. F. Santambrogio, Optimal channel networks, landscape function and branched transport. Interfaces and Free Boundaries **9**, 149–169 (2007)
273. F. Santambrogio, Absolute continuity and summability of transport densities: simpler proofs and new estimates. Calc. Var. Par. Differ. Equat. **36**(3), 343–354 (2009)
274. F. Santambrogio, Introduction to optimal transport theory, in *Optimal Transportation, Theory and Applications*, ed. by Y. Ollivier, H. Pajot, C. Villani. The London Mathematical Society, 2014, Lecture notes for a Summer School held in Cambridge University Press, Grenoble (2009) pp. 3–21

275. F. Santambrogio, Models and applications of optimal transport theory, in *Optimal Transportation, Theory and Applications*, ed. by Y. Ollivier, H. Pajot, C. Villani. The London Mathematical Society, 2014, Lecture notes for a Summer School held in Grenoble, Cambridge University Press (2009), pp. 22–40
276. F. Santambrogio, Inégalités Isopérimétriques quantitatives via le transport optimal, (d'après A. Figalli, F. Maggi et A. Pratelli), in *Proceedings of the Bourbaki Seminar*, 2011 (in French)
277. F. Santambrogio, Flots de gradient dans les espaces métriques et leurs applications (d'après Ambrosio-Gigli-Savaré), in *Proceedings of the Bourbaki Seminar*, 2013 (in French)
278. F. Santambrogio, A Dacorogna-Moser approach to flow decomposition and minimal flow problems. ESAIM: Proc. Surv. (SMAI 2013) **45**, 265–174 (2014)
279. F. Santambrogio, V. Vespri, Continuity for a very degenerate elliptic equation in two dimensions. Nonlinear Anal.: Theory Methods Appl. **73**, 3832–3841 (2010)
280. L.-P. Saumier, M. Agueh, B. Khouider, An efficient numerical algorithm for the L^2 optimal transport problem with periodic densities. IMA J. Appl. Math. **80**(1), 135–157 (2015)
281. T. Schmidt, $W^{2,1+\varepsilon}$ estimates for the Monge-Ampère equation. Adv. Math. **240**, 672–689 (2013)
282. E. Schrödinger, Uber die umkehrung der naturgesetze. Sitzungsberichte Preuss. Akad. Wiss. Berlin. Phys. Math. **144**, 144–153 (1931)
283. J.A. Sethian, *Level Set Methods and Fast Marching Methods* (Cambridge University Press, Cambridge, 1999)
284. A.I. Shnirelman, The geometry of the group of diffeomorphisms and the dynamics of an ideal incompressible fluid. (Russian) Mat. Sb. (N.S.) **128**(170), 82–109 (1985)
285. A.I. Shnirelman, Generalized fluid flows, their approximation and applications. Geom. Funct. Anal. **4**(5), 586–620 (1994)
286. G.J. Shutts, M.J.P. Cullen, Parcel stability and its relation to semigeostrophic theory. J. Atmos. Sci. **44**, 1318–1330 (1987)
287. S.K. Smirnov, Decomposition of solenoidal vector charges into elementary solenoids, and the structure of normal one-dimensional flows. Algebra i Analiz **5**(4), 206–238 (1993). Later translated into English in St. Petersburg Math. J. **5**(4), 841–867 (1994)
288. K.-T. Sturm, On the geometry of metric measure spaces. I. Acta Math. **196**, 65–131 (2006)
289. K.-T. Sturm, On the geometry of metric measure spaces. II. Acta Math. **196**, 133–177 (2006)
290. V.N. Sudakov, Geometric problems in the theory of infinite-dimensional probability distributions. Cover to cover translation of Trudy Mat. Inst. Steklov 141 (1976). Proc. Steklov Inst. Math. **2**(i–v), 1–178 (1979)
291. N. Trudinger, X.-J. Wang, On the Monge mass transfer problem. Calc. Var. Part. Differ. Equat. **13**, 19–31 (2001)
292. C. Villani, *Topics in Optimal Transportation*. Graduate Studies in Mathematics (American Mathematical Society, Providence, 2003)
293. C. Villani, *Optimal Transport: Old and New* (Springer, New York, 2008)
294. C. Villani, Regularity of optimal transport and cut locus: from nonsmooth analysis to geometry to smooth analysis. Discr. Contin. Dyn. Syst. A **30**(2), 559–571 (2011)
295. J.G. Wardrop, Some theoretical aspects of road traffic research. Proc. Inst. Civ. Eng. **2**, 325–378 (1952)
296. B. White, Rectifiability of flat chains. Ann. Math. (2) **150**(1), 165–184 (1999)
297. Q. Xia, Optimal paths related to transport problems. Commun. Contemp. Math. **5**(2), 251–279 (2003)
298. Q. Xia, Interior regularity of optimal transport paths. Calc. Var. Part. Differ. Equat. **20**(3), 283–299 (2004)
299. Q. Xia, Numerical simulation of optimal transport paths, in *Proceedings of the Second International Conference on Computer Modeling and Simulation* (ICCMS 2010), vol. 1 (2010), pp. 521–525
300. Q. Xia, Boundary regularity of optimal transport paths. Adv. Calc. Var. **4**(2), 153–174 (2011)

Index

F. Santambrogio, *Optimal Transport for Applied Mathematicians*,
Progress in Nonlinear Differential Equations and Their Applications 87,
DOI 10.1007/978-3-319-20828-2