

Cauchy problem for the Einstein equations, are not addressed at all. While it is of course unrealistic to ask for a detailed treatment of all these subjects, especially in a book for undergraduates, more steps in that direction could perhaps have been taken. Nevertheless, these small quibbles should not take away from the fact that this book is a valuable addition to the general relativity literature for mathematicians, and one which I highly recommend.

Amol Sasane, *A Mathematical Introduction to General Relativity*. World Scientific, 2021, 500 pages, Hardback ISBN 978-981-124-377-6, eBook ISBN 978-981-12-4379-0.

José Natário is a full professor at the Mathematics Department of Instituto Superior Técnico (Universidade de Lisboa). He has authored more than 50 research papers and three books, mostly on geometry and mathematical relativity.

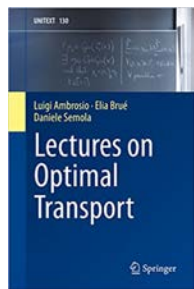
jnatar@math.tecnico.ulisboa.pt

DOI 10.4171/MAG/75

Lectures on Optimal Transport

by Luigi Ambrosio, Elia Brué and Daniele Semola

Reviewed by Filippo Santambrogio



This is the first of the two books that I am reviewing for this issue of the EMS Magazine. It is a textbook on optimal transport (in the same spirit of a book I published in 2015 [9], or of the two books by Cédric Villani [11, 12]), meant to be used by graduate students. The first author is one of the leading experts on the topic, who has been giving lectures on it for decades at SNS Pisa (by the way, it is in the course that he gave

exactly 20 years ago that I started learning about optimal transport). The second and third authors are two of the brilliant students who attended these courses in Pisa.

The book is organized into 19 chapters, each meant to correspond to a single lecture. The duration of a single lecture is not suggested explicitly, but I find the rhythm a little bit slow for graduate students, as I usually cover the material of the first 6 or 7 lectures in approximately 6 hours. Regardless, the idea of organizing the presentation according to teaching time is a very useful pedagogical tool.

The 19 lectures can be roughly divided into four series. Lectures 1 to 7 are essentially devoted to the main theory of the Monge

and Kantorovich problems, where two measures are fixed and one looks for the optimal plans or maps to transport the first measure onto the second at minimal cost. At the beginning the cost function is as general as possible, which allows to develop the whole Kantorovich theory, including existence of optimal plans and duality. Only in the last of these lectures the focus is on some precise Euclidean examples, and in particular on the quadratic cost, together with its connections with the Monge–Ampère equation (whose name is spelled correctly all along the book, except for the title of the corresponding lecture where, unfortunately, we can see an acute accent). Another very natural cost, the distance cost originally studied by Monge, is deliberately discussed for only a single page, since it is clearly the goal of the authors to move on to some notions, in connection with PDEs and differential geometry, that are more related to the quadratic cost. Some choices in the proofs or in the presentation could be debatable, for instance regarding duality: the authors do present, shortly, a proof based on rather general convex analysis (the Fenchel–Rockafellar theorem), but devote more space to a full and self-contained proof based on the c -cyclical monotonicity of the support, arguing that it is more constructive, which is absolutely true. On the other hand, this approach might suggest the wrong idea that each optimizer in the Kantorovich problem is associated with a specific maximizer of the dual (the one built from the support of this very optimizer) and this can be seen in the (absolutely classical) proof of uniqueness of optimal transport maps. This proof is based on the clever statement that if every optimal plan is induced by a map, then it is unique, but does not exploit the fact that the map corresponding to a plan can be chosen to be the same for all plans.

After the general presentation of the optimal transport problem, a second series of lectures (8–10) on the Wasserstein distances and Wasserstein spaces follows. Here the authors do a remarkable work by systematically analysing which metric properties of a metric space (X, d) are inherited by the corresponding Wasserstein space $(\mathcal{P}(X), W_2)$ (we see that the focus is explicitly on the case $p = 2$, in order to pave the way for the next part of the book): compactness, completeness, geodesics, ... Some parts require the introduction of suitable tools from analysis in metric spaces, in particular the notion of metric derivative, which are independent of optimal transport, but not always well known among graduate students in analysis.

Similarly, the next series of lectures (11–14) is not specifically related to optimal transport: it is devoted to a detailed analysis of gradient flows in Hilbert spaces, paying attention to those notions which can be extended to metric spaces, and in particular the EVI (Evolution Variational Inequality) and the EDI (Energy Dissipation Inequality) formulations. The role played by convexity or λ -convexity is emphasized from the very beginning. A full chapter is devoted to the study of the heat flow as a gradient flow with different choices of the functional and of the Hilbert norm (the heat flow is, for instance, the gradient flow of the Dirichlet energy $u \mapsto \frac{1}{2} \int |\nabla u|^2$ in the L^2 space, but also of the simplest functional $u \mapsto \frac{1}{2} \int u^2$ in the

homogeneous H^{-1} space). This is very useful, enabling the reader to realize that one and the same equation can be seen as a gradient flow in many different ways, and that there is an interplay between the functional and the distance (one needs to change both if one is looking for the same equation).

After this detailed discussion about gradient flows, the authors come back for the last lectures (15–19) to the Wasserstein space W_2 . First, a long chapter is devoted to the study of various functionals on the space of probability measures and on their variational properties (in particular, lower semicontinuity and geodesic convexity). This recalls what I did in a chapter of [9] and I am glad to see that the authors share my feeling that, in order to address some applications of optimal transport, at some point it is absolutely necessary to clarify what we know and what we should know about the most used functionals. The next lectures are mainly devoted to curves of measures, with a detailed discussion of the continuity equation $\partial_t \mu + \nabla \cdot (\mu v) = 0$, and of the specific case of geodesic curves (for which the velocity field v is related to the gradient of the solution of a Hamilton–Jacobi equation solved by means of the Hopf–Lax semigroup). This is followed by the dynamical formulation of optimal transport proposed by Benamou and Brenier [4] and the characterization of “nice” curves in W_2 as solutions of the continuity equation with L^2 vector fields. I find it unusual to follow this order (I usually consider generic curves and then optimal ones, even if admittedly also in my presentation geodesics are used to build a velocity field), but the exposition is perfectly coherent and clear. I also note a very nice proof of the semicontinuity of the Benamou–Brenier energy based on the interpretation of the L^p norm as dual to the corresponding L^q norm. The last two lectures in this series concentrate more on the heat flow, proving that the solutions of the heat equation are metric EVI gradient flows in the Wasserstein space of the entropy functional. This is the meaning which is given to being a gradient flow in this approach, coherently with the famous book of the first author with N. Gigli and G. Savaré [2]. Since the procedure consists here in checking that an existing solution of a well-known PDE satisfies this notion, this can be seen as an interpretation tool, and not as a way to prove existence of solutions to some PDE with a gradient flow structure in W_2 (in particular, the celebrated Jordan–Kinderlehrer–Otto scheme [7] is only marginally mentioned). Finally, it is shown that the behaviour of the heat flow on Riemannian manifolds is strongly related to conditions on the curvature and to the geodesic convexity of the entropy functional, a fact that has been used as a starting point for a synthetic definition of the notion of curvature bound in metric measure spaces [8, 10] and then for a very general theory of calculus in metric measure spaces [3].

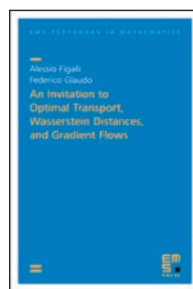
Even if some points in the exposition differ from what I would have done – and I had fun in pointing this out – there is no doubt in my mind that this very well written 250-page book can be an extremely useful tool to teach optimal transport classes or, for a more experienced researcher from a related but different field, to access

the subject. It contains both heuristic discussions which could be completed by external reading and rigorous proofs, and covers a large part of the existing theory. Covering the entire theory was of course impossible, and the choice was made to focus on topics that better prepare the reader to deal with some mathematical applications, in particular in connection with differential geometry and partial differential equations from an abstract point of view, and the book is clearly meant for a public of analysts or geometers. And it contains many clever ideas and tricks that the reader will appreciate and re-use in his own work!

Luigi Ambrosio, Elia Brué and Daniele Semola, *Lectures on Optimal Transport*. Springer, 2021, 259 pages, Paperback ISBN 978-3-030-72161-9, eBook ISBN 978-3-030-72162-6.

An Invitation to Optimal Transport, Wasserstein Distances, and Gradient Flows by Alessio Figalli and Federico Glaudo

Reviewed by Filippo Santambrogio



We now move on to the second book I am reviewing for this issue. It is also a text on optimal transport, and it is also produced by researchers from the same school (i.e., the Italian school on calculus of variations centred at SNS Pisa). This book aims at being a self-contained introduction to optimal transport and some of its applications, with a quite explicit goal to acquaint the reader with the theory and invite her/him

to start working on and with it, possibly looking for more detailed developments or proofs elsewhere.

Optimal transport is a very active field and every new monograph that can attract colleagues from related disciplines or can help researchers who need to understand it after a first encounter, is more than welcome. One of the key advantages of this short monograph is the authorship, since one of the authors is well known in the whole mathematical world because of the Fields medal he was awarded exactly for his work on optimal transport: this is very likely to attract more readers than any other manuscript on the topic.

The first important point to notice is the book’s size, approximately 130 pages. It is much shorter than other references on the topic, and in this respect it is difficult to compare it to the book by Ambrosio, Brué and Semola or to other classical references [9, 11, 12]. The closest manuscript that one can use as a comparison should probably be *A user’s guide to optimal transport* by Ambrosio and Gigli [1], even if inviting readers to discover a topic is not exactly the same as the claimed goal to offer a guide to under-

stand, or at least use, it. In comparing the two texts one has to note that Ambrosio and Gigli's work is in the end only partially about optimal transport, being instead heavily oriented towards metric measure spaces. By contrast, Figalli and Glaudo's work remains focused on optimal transport, which allows for a more complete and useful presentation, despite the small length.

In comparing to classical books, the reference is always the first book by Villani [11], since my own text [9], published 12 years later, aimed at adding developments of the theory that did not exist in 2003, and Villani's second book [12], includes several hundreds of pages on new extensions and connections, in particular in the direction of differential geometry. From this point of view, the present manuscript does not aim at covering new material, as the core of its exposition concentrates on topics already present in [11], and new developments are only worth a few lines in the further reading part. This is a very legitimate choice if one wants to keep the presentation short as well as reasonably self-contained. On the other hand, I would say that, despite aiming at a slightly more pure-math oriented audience, this book shares a "concrete" flavour with [9].

As a short introductory text, the book is composed of only five chapters, and the last one, called "further reading" honestly discusses the other existing references on the topic, and some extensions or connections. Chapter 1 also plays a different role than the others, including some examples of transport maps, some applications (for instance, how to prove the isoperimetric inequality using the Knothe map), and some preliminary background material. The core of the book thus consists of Chapters 2, 3, and 4.

Chapter 2 is devoted to the already classical theory of optimal transport, following more or less the same structure as that of the book by Ambrosio, Brué and Semola (including very similar approaches to duality and to the uniqueness of optimal maps), even if an important role is given to the cost $c(x, y) = -x \cdot y$, which is equivalent to the quadratic cost $\frac{1}{2}|x - y|^2$ and allows for a direct use of convex analysis without the need to introduce c -convexity (or c -concavity). If general costs arrive first in what concerns existence of optimal plans, they appear later in what concerns Kantorovich duality. In this same chapter we can also praise the detailed discussion of the various connections of optimal transport with the incompressible Euler equation (which also allows to underline the multiple roles Yann Brenier played in the theory of optimal transport, see [4–6]).

Chapter 3 and Chapter 4 both include at the same time the second and the third key concepts evoked by the title: Wasserstein distances and gradient flows. Chapter 3 is more metric in nature: it introduces the Wasserstein distances W_p (for every p) and after a short (a few pages, not a few chapters) digression on Hilbertian gradient flows, moves on to the Jordan–Kinderlehrer–Otto scheme (JKO [7]), thus attacking gradient flows via a sequence of iterated minimization problems involving the W_2 distance. This chapter treats only the case of the heat equation, which is the simplest

one, but has the drawback of being also a gradient flow in L^2 , differently from the Fokker–Planck equation with a potential V which is dealt with in Chapter 4. The main theorem here states that the limit as the time step τ of the JKO scheme tends to 0 is a distributional solution of the heat equation, and in this sense I consider this presentation as more "concrete": it finds a solution – in a very standard sense – to a PDE, and not to a metric condition such as the EVI or the EDI definitions of gradient flows, which could sound unnatural as a definition when PDE tools are available.

Chapter 4 then goes on with the differential and Riemannian structure of the Wasserstein space, discussing geodesic curves and the Benamou–Brenier formula, introducing Otto's calculus in order to endow this space with a formal notion of tangent space and make it a sort of infinite-dimensional Riemannian manifold, and then presenting the notion of geodesic convexity. In a way that I strongly approve, geodesic convexity is shown to be crucial for the study of gradient flows in what concerns finding properties of their solutions, but it is not at all evoked when it comes to proving existence for some PDEs. As an example, the authors then concentrate on the Fokker–Planck equation, adding a convex potential to the heat flow, and prove a series of inequalities which allow to obtain well-known rates of convergence to the steady state for strongly convex confining potentials. They finish the chapter proving convergence to the same steady state in the strong L^1 sense (and not only in the Wasserstein sense, which means weak convergence), providing a very nice proof of a suitable functional inequality (the Csiszár–Kullback–Pinsker inequality; my only criticism here is that the authors claim that a certain step of the proof, establishing that a certain function is negative on the boundary in order to apply later a sort of maximum principle, is "easy", while it required me some work to reach this conclusion).

Besides Chapters 1–5, the book also contains two appendices, both including exercises. Appendix B aims at providing a proof of the disintegration theorem in measure theory via a series of guided exercises, while Appendix A is a collection of 11 fully solved exercises (though I think many readers would have liked to see more exercises than these 11).

Book reviews are not meant for authors, but for potential readers, but in case the authors will read this review they will probably realize, in view of the similarity of some comments and sentences, that I also refereed their manuscript before publication, which means that I had more than one occasion to look at their work (further occasions also include a student whom I supervised who decided to build up her knowledge of optimal transport on this very book). I must say that I get more and more convinced every time I take a closer look: yes, I like this book. It does the job of inviting readers to the field, and it does it well.

Alessio Figalli and Federico Glaudo, *An Invitation to Optimal Transport, Wasserstein Distances, and Gradient Flows*. EMS Press, 2021, 144 pages, Hardback ISBN 978-3-98547-010-5, eBook ISBN 978-3-98547-510-0.

References

- [1] L. Ambrosio and N. Gigli, A user's guide to optimal transport. In *Modelling and optimisation of flows on networks*, Lecture Notes in Math. 2062, Springer, Heidelberg, 1–155 (2013)
- [2] L. Ambrosio, N. Gigli and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics, ETH Zürich, Birkhäuser, Basel (2005)
- [3] L. Ambrosio, N. Gigli and G. Savaré, Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below. *Invent. Math.* **195**, 289–391 (2014)
- [4] J.-D. Benamou and Y. Brenier, A computational fluid mechanics solution to the Monge–Kantorovich mass transfer problem. *Numer. Math.* **84**, 375–393 (2000)
- [5] Y. Brenier, Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.* **305**, 805–808 (1987)
- [6] Y. Brenier, The least action principle and the related concept of generalized flows for incompressible perfect fluids. *J. Amer. Math. Soc.* **2**, 225–255 (1989)
- [7] R. Jordan, D. Kinderlehrer and F. Otto, The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* **29**, 1–17 (1998)
- [8] J. Lott and C. Villani, Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)* **169**, 903–991 (2009)
- [9] F. Santambrogio, *Optimal transport for applied mathematicians*. Progress in Nonlinear Differential Equations and their Applications 87, Birkhäuser/Springer, Cham (2015)
- [10] K.-T. Sturm, On the geometry of metric measure spaces. I and II. *Acta Math.* **196**, 65–131 and 133–177 (2006)
- [11] C. Villani, *Topics in optimal transportation*. Grad. Stud. Math. 58, AMS, Providence (2003)
- [12] C. Villani, *Optimal transport*. Grundlehren Math. Wiss. 338, Springer, Berlin (2009)

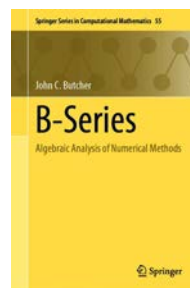
Filippo Santambrogio is professor at Université Claude Bernard Lyon 1, where he moved after spending twelve years in the Paris area (in Dauphine and Orsay), and after his studies at SNS Pisa. His research focuses on optimal transport and calculus of variations, both in what concerns the general theory and in their applications to the modeling of traffic systems, crowd motion and optimal location of resources. santambrogio@math.univ-lyon1.fr

DOI 10.4171/MAG/88

B-Series: Algebraic Analysis of Numerical Methods

by John C. Butcher

Reviewed by Raffaele D'Ambrosio



The algebraic analysis of numerical methods is driven along a bridge matching numerical analysis, graph theory, group theory, differentiation of vector fields, and so on. The result of the wise union of these aspects is an effective and elegant theory that is useful for investigating meaningful features of numerical methods for differential equations. One of the pioneering sources of this theory is the brilliant and ingenious work of

John C. Butcher. He has authored a remarkable number of seminal contributions to the founding, establishment and development of the modern theory of Runge–Kutta methods and its subsequent extensions.

A building block for the algebraic analysis of numerical methods and, in particular, of their order of convergence, is the connection between rooted trees and the differentiation of vector fields. Such a link lies at the basis of Butcher's theory and its precursors¹, such as Arthur Cayley (1821–1895) and Robin H. Merson (1921–1992), a scientist at the Royal Aircraft Establishment, who became popular for his involvement in the computations of an accurate orbit for Sputnik 1, launched in 1957. In the same year, Butcher attended the talk by Merson at the conference "Data Processing and Automatic Computing Machines" held in Salisbury, South Australia, where Merson described the one-to-one correspondence between derivatives and rooted trees. The full theory will only be provided later by Butcher, but it is worth mentioning that Cayley introduced trees with the same purpose as in Butcher theory (namely, to understand and effectively represent the interaction of vector fields repeatedly applied to one another), and then for one century this aspect was totally forgotten in the literature, and reconsidered with effectiveness only when the theory of numerical methods was established with rigor.

Other two building blocks for the algebraic analysis of numerical methods, allowing to detect and elegantly prove properties of numerical methods for differential problems, are well described in the book under review, namely

- Butcher series (in short, B-series), allowing to represent both exact and numerical solutions to a differential problem in terms of series expansions whose coefficients are functions of rooted trees of a prescribed order;

¹ Historical aspects have been described in detail, for instance, in the paper by R. McLachlan et al., *Asia Pacific Math. Newsletter* 7(1), 1–11 (2017).