美国数学会经典影印系列

AMS
AMERICAN MATHEMATICAL SOCIETY

Topics in Optimal Transportation

Second Edition

最优输运理论专题

第二版

Cédric Villani

高等教育出版社

美国数学会经典影印系列

AMS
AMERICAN MATHEMATICAL SOCIETY

Topics in Optimal Transportation

Second Edition

最优输运理论专题

第二版

Cédric Villani

高等教育出版社·北京

图字：01-2016-2529 号

最优输运理论专题
Zuiyou Shuyun Lilun Zhuanti

图书在版编目 (CIP) 数据

最优输运理论专题 = Topics in Optimal Transportation ：英文 / (法) 塞德里克 · 维拉尼 (Cédric Villani) 著 . —影印本. —北京 : 高等教育出版社 , 2017.1
ISBN 978-7-04-046921-9
Ⅰ. ①最… Ⅱ. ①塞… Ⅲ. ①数学规划—英文 ②蒙日—安培方程—英文 Ⅳ. ①O221 ②O175.26
中国版本图书馆 CIP 数据核字 (2016) 第 282012 号

策划编辑 赵天夫　　责任编辑 赵天夫
封面设计 张申申　　责任印制 毛斯璐

出版发行 高等教育出版社
社址 北京市西城区德外大街 4 号
邮政编码 100120
购书热线 010-58581118
咨询电话 400-810-0598
网址 http://www.hep.edu.cn
http://www.hep.com.cn
网上订购 http://www.hepmall.com.cn
http://www.hepmall.com
http://www.hepmall.cn
印刷:北京新华印刷有限公司

开本 787mm × 1092mm 1/16
印张 25.5
字数 531 千字
版次 2017 年 1 月第 1 版
印次 2017 年 1 月第 1 次印刷
定价 169.00 元

[物 料 号 46921-00]

美国数学会经典影印系列

出版者的话

近年来，我国的科学技术取得了长足进步，特别是在数学等自然科学基础领域不断涌现出一流的研究成果。与此同时，国内的科研队伍与国外的交流合作也越来越密切，越来越多的科研工作者可以熟练地阅读英文文献，并在国际顶级期刊发表英文学术文章，在国外出版社出版英文学术著作。

然而，在国内阅读海外原版英文图书仍不是非常便捷。一方面，这些原版图书主要集中在科技、教育比较发达的大中城市的大型综合图书馆以及科研院所的资料室中，普通读者借阅不甚容易；另一方面，原版书价格昂贵，动辄上百美元，购买也很不方便。这极大地限制了科技工作者对于国外先进科学技术知识的获取，间接阻碍了我国科技的发展。

高等教育出版社本着植根教育、弘扬学术的宗旨服务我国广大科技和教育工作者，同美国数学会（American Mathematical Society）合作，在征求海内外众多专家学者意见的基础上，精选该学会近年出版的数十种专业著作，组织出版了“美国数学会经典影印系列”丛书。美国数学会创建于 1888 年，是国际上极具影响力的专业学术组织，目前拥有近 30000 会员和 580 余个机构成员，出版图书 3500 多种，冯·诺依曼、莱夫谢茨、陶哲轩等世界级数学大家都是其作者。本影印系列涵盖了代数、几何、分析、方程、拓扑、概率、动力系统等所有主要数学分支以及新近发展的数学主题。

我们希望这套书的出版，能够对国内的科研工作者、教育工作者以及青年学生起到重要的学术引领作用，也希望今后能有更多的海外优秀英文著作被介绍到中国。

高 等 教 育 出 版 社

2016 年 12 月

To my optimal son, Neven

Preface of the Second Edition

Time has passed since the first edition of this book. My son, who was born during the writing of the manuscript, is now entering high school. As for the theory of optimal transport, I used to think of her as a young adult, but I now realize that in those days she was just a teenager. And she did grow up spectacularly since then, being enriched by many developments.

To begin with, the connection to non-Euclidean geometry, which was only hinted at in those days, has blossomed into an ambitious, fast developing theory. This led to major progress, not only in optimal transportation but also in synthetic geometry, through a program heralded by John Lott, Karl-Theodor Sturm and myself. It also boosted nonsmooth gradient flow theory under the expert hands of Luigi Ambrosio, Nicola Gigli, Giuseppe Savaré, and many others — a book on this subject was born shortly after mine, it looked like the final word but in fact it was just a step in a broader picture. And as a side product, through the skills of the Italian school, this direction of research also led to fundamental advances in metric analysis.

Further, the connection to dynamical systems and the Aubry–Mather theory was enormously pushed through the contributions of Patrick Bernard, Albert Fathi, Wilfrid Gangbo, and an illuminating reinterpretation of the older work of John Mather. This new light had a deep influence on further presentation of the subject.

Also, the theory of fully nonlinear partial differential equations was revisited with fracas as Xiaonan Ma, Neil Trudinger, Xiu-Ja Wang and Grégoire Loeper opened the door to the regularity theory of optimal transportation in curved geometry, and essentially solved what I presented at the end of Chapter 3 as the most important open problem in the field. An ensuing revival of cut locus theory was just one of the associated amazing developments.

At the time of the first edition, I was expressing hope in the rise of certain numerical schemes, but in the end faster algorithms were developed in other ways, for instance by Quentin Mérigot and Marco Cuturi.

More generally, every feature of the theory has been revised, extended and generalized in nearly any possible way. Even more importantly, a crowd of young and talented researchers, expert in those subjects, has emerged – among many, let me quote François Bolley, Vincent Calvez, Fabio Cavalletti, Thierry Champion, Luigi De Pascale, Matthias Erbar, Max Fathi, Alessio Figalli, Nicola Gigli, Nathaël Gozlan, Young-Heon Kim, Kazumasa Kuwada, Grégoire Loeper, Jan Maas, Francesco Maggi, Emanuel Milman, Andrea Mondino, Aaron Naber, Aldo Pratelli, Tapio Rajala, Max Von Renesse, Ludovic Rifford, Filippo Santambrogio... It may be vain, but it has been an immense pride and joy to think that I participated in making this community emerge — by the writing of this book in the first place, but also by writing the next one, by lecturing on the subject all around the world, and by organizing events such as the Oberwolfach Seminar which I held in 2002 with Luigi Ambrosio.

In spite of all these developments, the present book has aged well and is still used regularly as a source of working seminars or as an introduction to the subject. Part of this longevity is certainly due to the informal and enthusiastic writing style, maybe typical of a young mathematician, and which I feel I can never recover. It is also probably due to the fact that my second book on the subject, *Optimal Transport, old and new* (Springer, 2008) has served as a reference textbook — bigger, more general, more geometric — but has not replaced the present course as an introduction to the field. For these reasons, I have decided to make the revision as modest as could be, limiting it mainly to the correction of mistakes and a few mandatory comments, and basically keeping the numbering of sections and statements unchanged. And likewise, I abandoned my longstanding desire to add a chapter on numerical simulations, preferring to leave that task to others.

Mistakes in the first edition were not so many, but not so few either. Many colleagues have helped me chase them, in particular Scott Armstrong, Eitan Bachmat, Patrick Cattiaux, Ulrik Fjordholm, Bo'az Klartag, Elliott Lieb, Sylvia Serfaty, Michel Valadier, Marcus Wunsch and Xiling Zhang. It is a pleasure to thank them all, with particularly warm feelings to Ulrik and Sylvia for the detailed reporting which came out of their working seminars.

With the passing of time it became clearer that my 1999 stay in Georgia Tech, Atlanta, which led to the writing of this course, had a huge impact on my mathematician's life; so it is only appropriate that I thank again Wilfrid Gangbo, Eric Carlen and Michael Loss for organizing it. As far as my work in optimal transportation is concerned, additional thanks are due to my research collaborators, most notably Felix Otto, Luigi Ambrosio and John Lott; to my colleagues in Lyon, in particular Albert Fathi and Étienne Ghys; and to my students, of course.

Last but not least, let me evoke the memory of Ralph Sizer, who was the copy-editor of the first edition. The work which he did on the manuscript was outstanding, by and large. A few years later, I tried to get his advice for another project, and was shocked to learn about his untimely passing away. Although I never met Sizer, our long exchanges of mails made me feel like I knew him. Behind those ASCII characters and handwritten scribblings lie human emotions, and the joy of working together.

I shall conclude this foreword with comments on the presentation. I have left notation and conventions unchanged from the first edition, rather than updating them in accordance with some of my later writings. In particular, in this book,

- I use the word "transportation" rather than "transport";
- I denote by $T\#\mu$ the push-forward of the measure μ by the mapping T; while alternative conventions would be $T_{\#}\mu$ (which I used in *Optimal transport, old and new*) or $T_*\mu$, or $T\mu$;
- I use the full axiom of choice in a few places, most notably Section 1.3, while I completely banned its use from more recent writings;
- I use the spelling Aleksandrov rather than Alexandrov;
- I express the Kantorovich duality with either a pair of dual semi-convex functions, or a pair of dual semi-concave functions, while the convention which I used in more recent writings is rather in the form of one semi-convex and one semi-concave functions (facilitating the connection to Mather's theory).

As a final remark, the terminology "Wasserstein distance" (Chapter 7) has been objected to, in particular because those distances were first put forward by Leonid Kantorovich himself, at least in some cases. Still, I stuck to the use of Wasserstein, partly because it has become popular in the field, partly because so many key concepts are already named after Kantorovich.

I hope that readers will enjoy these notes as much as I enjoyed discovering the subject and writing about it, more than fifteen years ago.

Cédric Villani
Paris & Lyon, July 2015

Preface of the First Edition

This set of notes grew from a graduate course that I taught at Georgia Tech, in Atlanta, during the fall of 1999, on the invitation of Wilfrid Gangbo. It is a great pleasure for me to thank Georgia Tech for its hospitality, and all the faculty members and students who attended this course, for their interest and their curiosity. Among them, I wish to express my particular gratitude to Eric Carlen, Laci Erdös, Michael Loss, and Andrzej Swiech. It was Eric and Michael who first suggested that I make a book out of the lecture notes intended for the students.

Three years passed by before I was able to complete these notes; of course, I took into account as much as I could of the mathematical progress made during those years.

Optimal mass transportation was born in France in 1781, with a very famous paper by Gaspard Monge, *Mémoire sur la théorie des déblais et des remblais.* Since then, it has become a classical subject in probability theory, economics and optimization. Very recently it gained extreme popularity, because many researchers in different areas of mathematics understood that this topic was strongly linked to their subject. Again, one can give a precise birthdate for this revival: the 1987 note by Yann Brenier, *Décomposition polaire et réarrangement des champs de vecteurs.* This paper paved the way towards a beautiful interplay between partial differential equations, fluid mechanics, geometry, probability theory and functional analysis, which has developed over the last ten years, through the contributions of a number of authors, with optimal transportation problems as a common denominator.

These notes are definitely not intended to be exhaustive, and should rather be seen as an introduction to the subject. Their reading can be complemented by some of the reference texts which have appeared recently. In particular, I should mention the two volumes of *Mass transportation problems*, by Rachev and Rüschendorf, which depict many applications of Monge-Kantorovich distances to various problems, together with the classical theory of the optimal transportation problem in a very abstract setting; the survey by Evans, which can also be considered as an introduction to the subject, and describes several applications of the L^1 theory (i.e., when the cost function is a distance) which I did not cover in these notes; the extremely clear lecture notes by Ambrosio, centered on the L^1 theory from the point of view of calculus of variations; and also the lecture notes by Urbas, which are a marvelous reference for the regularity theory of the Monge-Ampère equation arising in mass transportation. Also recommended is a very pedagogical and rather complete article recently written by Ambrosio and Pratelli, and focused on the L^1 theory, from which I extracted many remarks and examples.

The present volume does not go too deeply into some of the aspects which are very well treated in the above-mentioned references: in particular, the L^1 theory is just sketched, and so is the regularity theory developed by Caffarelli and by Urbas. Several topics are hardly mentioned, or not at all: the application of mass transportation to the problem of shape optimization, as developed by Bouchitté and Buttazzo; the fascinating semi-geostrophic system in meteorology, whose links with optimal transportation are now understood thanks to the amazing work of Cullen, Purser and collaborators; or applications to image processing, developed by Tannenbaum and his group. On the other hand, I hope that this text is a good elementary reference source for such topics as displacement interpolation and its applications to functional inequalities with a geometrical content, or the differential viewpoint of Otto, which has proven so successful in various contexts (like the study of rates of equilibration for certain dissipative equations). I have tried to keep proofs as simple as possible throughout the book, keeping in mind that they should be understandable by non-expert students. I have also stated many results without proofs, either to convey a better intuition, or to give an account of recent research in the field. In the end, these notes are intended to serve both as a course, and as a survey.

Though the literature on the Monge-Kantorovich problem is enormous, I did not want the bibliography to become gigantic, and therefore I did *not* try to give complete lists of references. Many authors who did valuable work on optimal transportation problems (Abdellaoui, Cuesta-Albertos, Dall'Aglio, Kellerer, Matrán, Tuero-Díaz, and many others) are not even cited within the text; I apologize for that. Much more complete lists of references on the Monge-Kantorovich problem can be found in Gangbo and McCann [**151**], and especially in Rachev and Rüschendorf [**225**]. On the other hand, I did not hesitate to give references for subjects whose relation to the optimal transportation problem is not necessarily immediate, whenever I felt that this could give the reader some insights in related fields.

At first I did not intend to consider the optimal mass transportation problem in a very general framework. But a graduate course that I taught in the fall of 2001 on the mean-field limit in statistical physics, made me realize the practical importance of handling mass transportation on infinite-dimensional spaces such as the Wiener space, or the space of probability measures on some phase space. Tools like the Kantorovich duality, or the metric properties induced by optimal transportation, happen to be very useful in such contexts — as was understood long ago by people doing research in mathematical statistics. This is why in Chapters 1 and 7 I have treated those topics under quite general assumptions, in a context of Polish spaces (which is not the most general setting that one could imagine, but which is sufficient for all the applications I am used to). Almost all the rest of the notes deals with finite-dimensional spaces. Let me mention that several researchers, in particular Üstünel and F.-Y. Wang, are currently working to extend some of the geometrical results described below to an infinite-dimensional setting allowing for the Wiener space.

A more precise overview of the contents of this book is given at the end of the Introduction, after a precise statement of the problem. I shall also summarize at the beginning the main notation used in the text; to avoid devastating confusion, note carefully the definition of a "small set" in $\mathbb{R}^n$, as a set of Hausdorff dimension at most $n-1$.

As the reader should understand, the subject is still very vivid and likely to get into new developments in the next years. Among topics which are still waiting for progress, let me only mention the numerical methods for computing optimal transportations. At the time of this writing, some noticeable progress seems to have been done on this subject by Tannenbaum and his coworkers. Even though these beautiful new schemes seem extremely promising, they need confirmation from the mathematical point of view, which is one reason why I skipped this topic (the other reason being my lack of competence). Some related results can be found in [**163**].

Also I wish to emphasize that optimal mass transportation, besides its own intrinsic interest, sometimes appears as a surprisingly effective *tool* in problems which do not a priori seem to have any relation to it. For this reason I think that getting at least superficially acquainted with it is a wise investment for any student in probability, analysis or partial differential equations.

This book owes a lot to many people. I was lucky enough to learn the subject of optimal mass transportation directly from two of those researchers who have most contributed to turn it into a fascinating area: Yann Brenier and Felix Otto; it is a pleasure here to express my enormous gratitude to them. I first encountered optimal transportation in Tanaka's work about the Boltzmann equation, and my curiosity about it increased dramatically from discussions with Yann; but it was only after hearing a beautiful and enthusiastic lecture given in Paris by Craig Evans, that I made up my mind to study the subject thoroughly. My involvement in the study of functional inequalities related to mass transportation was partly triggered by interactions with Michel Ledoux, whose influence is gratefully acknowledged. The present manuscript profited a lot from numerous discussions with Luigi Ambrosio, Eric Carlen, Dario Cordero-Erausquin, Wilfrid Gangbo and Robert McCann. Both Robert and Luigi taught the material of this book, made many suggestions and pointed out numerous misprints and mistakes in the first version of these notes. The most serious one concerned the "proof" of Theorem 1.3, as given in the first version of these notes; the gap was fixed thanks to the kind help of Luigi again, and of Bernd Kirchheim, with the final result of an improved statement. Some of my students at the École normale supérieure also spotted and repaired a gap in the proof of Theorem 2.18. François Bolley, Jean-François Coulombel and Maxime Hauray should be thanked for the time they spent hunting for mistakes and misprints in various parts of the book, and testing many of the exercises and problems. Richard Dudley was kind enough to give a quick but thorough look at Chapters 1 and 7. Chapter 4 would not have existed without the explanations which I received from Luis Caffarelli and Andrzej Swiech. Most of the material in Chapter 6 was taught to me by Franck Barthe. Chapter 8 was reshaped by the exchanges which I had with Luigi Ambrosio, Nicola Gigli and Étienne Ghys during the last stages of preparation of the manuscript. Finally, Mike Cullen corrected some mistakes in the presentation of the physical model in Problem 9 of Chapter 10.

Cédric Villani
Lyon, January 2003

Notation

The identity map on an arbitrary space will be denoted by Id. Whenever X is a set, we write $1_X(x) = 1$ if $x \in X$, $1_X(x) = 0$ otherwise. The complement of a set A will be denoted by A^c.

Throughout the text, whenever we write $\mathbb{R}^n$ the dimension n is an arbitrary integer $n \geq 1$. Whenever A is a Lebesgue-measurable subset of $\mathbb{R}^n$, its n-dimensional Lebesgue measure will be denoted by $|A|$. This should not be confused with the Euclidean norm of a vector $x \in \mathbb{R}^n$, which will also be denoted by $|x|$. Whenever $x, y \in \mathbb{R}^n$ we write $x \cdot y = \langle x, y \rangle = \sum_{i=1}^n x_i y_i$.

Given some abstract measure space X, we shall denote by $P(X)$ the set of all probability measures on X, and by $M(X)$ the set of all finite signed measures on X (i.e. precisely the vector space generated by $P(X)$). The space $M(X)$ is equipped with the norm of total variation, $\|\mu\|_{TV} = \inf\{\mu_+[X] + \mu_-[X]\}$, where the infimum is taken over all nonnegative measures μ_+, μ_- such that $\mu = \mu_+ - \mu_-$. The infimum is obtained when μ_+ and μ_- are singular to each other, in which case $\mu = \mu_+ - \mu_-$ is said to be the Hahn decomposition of μ. Of course, if ν is a nonnegative measure and f a measurable map, then $\|f\|_{L^1(d\nu)} = \|f\nu\|_{TV}$. From Chapter 1 on, we shall only work in topological spaces, equipped with their Borel σ-algebra; so $P(X)$ will be the set of Borel probability measures. We shall sometimes write $w*$–$P(X)$ for $P(X)$ equipped with the weak topology.

The Dirac mass at a point x will be denoted by δ_x: $\delta_x[A] = 1$ if $x \in A$, 0 otherwise.

If a particular measure μ on X is singled out, for $p \in [1, \infty)$ we shall denote by $L^p(X)$ or $L^p(d\mu)$ the Lebesgue space of order p for the reference measure μ, with the usual identification of functions which coincide almost everywhere. Whenever $p \geq 1$, we shall denote by p' its conjugate exponent:

$$\frac{1}{p} + \frac{1}{p'} = 1.$$

Whenever T is a map from a measure space X, equipped with a measure μ, to an arbitrary space Y, we denote by $T\#\mu$ the image measure (or push-forward) of μ by T. Explicitly, $(T\#\mu)[B] = \mu[T^{-1}(B)]$, where $T^{-1}(B) = \{x \in X;\ T(x) \in B\}$. The set of all $T : X \to X$ such that $T\#\mu = \mu$ will be denoted by $S(X)$. We shall always use push-forward in this sense: when we write $T\#f = g$, where f and g are nonnegative functions, this means that the measure having density f is pushed foward to the measure having density g (usually the reference measure will be the Lebesgue measure).

If X is a topological space, then it will be equipped with its Borel σ-algebra. We shall denote by $C(X)$ the space of continuous functions on X; by $C_b(X)$ the space of bounded continuous functions on X; and by $C_0(X)$ the space of continuous functions on X going to 0 at infinity. Sometimes these notations will be replaced by $C(X;\mathbb{R})$, $C_b(X;\mathbb{R})$, $C_0(X;\mathbb{R})$. The space $C_b(X)$ comes with a natural norm, $\|u\|_\infty = \sup_X |u|$. Whenever $A \subset X$, we denote by $\mathrm{Int}(A)$ the largest open set contained in A, and by $\overline{A}$ the smallest closed set containing A. We set $\partial A = \overline{A} \setminus \mathrm{Int}(A)$. By definition, the support of a measure μ on X will be the smallest closed set $F \subset X$ with $\mu[X \setminus F] = 0$, and will be denoted $\mathrm{Supp}\,\mu$. On the other hand, when we say that μ is concentrated on $A \subset X$, this only means that $\mu[X \setminus A] = 0$, without A being necessarily closed.

If X is a metric space, we shall equip it with the topology induced by its distance, and denote by $B(x, r)$ the ball of radius r and center x. We shall denote by $\mathrm{Lip}(X)$ the set of all Lipschitz functions on X; we shall also denote by $P_p(X)$ the space of Borel probability measures μ on X with finite moment of order p, in the sense that $\int d(x_0, x)^p\, d\mu(x) < +\infty$ for some (and thus any) $x_0 \in X$.

When X is a Banach space and X^* its topological dual, we shall denote by $\langle \cdot, \cdot \rangle$ the duality bracket between X and X^*. A particular case of this is the scalar product in a Hilbert space.

If φ is a convex function on a Banach space X, then φ^* will stand for its dual convex function, in the sense of Legendre-Fenchel duality. The subdifferential of φ will be denoted by $\partial\varphi$, and identified with its graph, which is a subset of $X \times X^*$. Basic definitions for these objects are recalled in Chapter 2. From Chapter 3 on, we shall abbreviate "proper lower semi-continuous convex function" into just "convex function".

When X is a smooth Riemannian manifold, or a Banach space, and F is a continuous function on X, we shall denote by DF its differential map, and by $DF(x) \cdot v$ its first-order variation at some point $x \in X$, along some tangent vector v.

When X is a smooth Riemannian manifold, we shall denote by T_xX the tangent space at a point x, and by $\langle \cdot, \cdot \rangle_x$ the scalar product on T_xX defined by the Riemannian structure. We shall denote by $\mathcal{D}(X)$ the space of C^∞ functions on X with compact support, and by $\mathcal{D}'(X)$ the space of distributions on X. We define the gradient operator ∇ on $\mathcal{D}(X)$ by the identity $\langle \nabla F(x), v \rangle_x = DF(x) \cdot v$; so $\nabla F(x)$ belongs to T_xX, while $DF(x)$ lies in $(T_xX)^*$. We shall denote by $\nabla \cdot$ the divergence operator, which is the adjoint of ∇ on $\mathcal{D}(X)$. The gradient operator acts on real-valued functions, while the divergence operator acts on vector fields. We also define the Laplace operator Δ by the identity $\Delta F = \nabla \cdot \nabla F$. By duality, all these operations are extended to $\mathcal{D}'(X)$. Of course, if $X = \mathbb{R}^n$, then

$$\nabla F = \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_n} \right), \qquad \nabla \cdot u = \sum_{i=1}^n \frac{\partial u_i}{\partial x_i}, \qquad \Delta F = \sum_{i=1}^n \frac{\partial^2 F}{\partial x_i^2}.$$

We also denote by D^2 the Hessian operator on X. Of course, if $X = \mathbb{R}^n$, then $D^2F(x)$ can be identified with the Hessian matrix $(\partial^2 F(x)/\partial x_i \, \partial x_j)$.

The space of absolutely continuous (with respect to Lebesgue measure) probability measures on $\mathbb{R}^n$ will be denoted by $P_{\mathrm{ac}}(\mathbb{R}^n)$; it can be identified with a subspace of $L^1(\mathbb{R}^n)$. The space of absolutely continuous probability measures with finite moments up to order 2 will be denoted by $P_{\mathrm{ac},2}(\mathbb{R}^n)$.

The Aleksandrov Hessian of a convex function φ on $\mathbb{R}^n$ will be denoted by $D^2_A\varphi$; it is only defined almost everywhere in the interior of the domain of φ. It should not be mistaken for the distributional Hessian of φ, denoted by $D^2_{\mathcal{D}'}\varphi$. The Hessian measure of φ will be denoted $\det_{\mathrm{H}} D^2\varphi$. All these notions will be explained within the text (see subsections 2.1.3 and 4.1.4). We shall use consistent notations for Laplace operators: the trace of $D^2_A\varphi$ (resp. $D^2_{\mathcal{D}'}\varphi$) will be denoted by $\Delta_A\varphi$ (resp. $\Delta_{\mathcal{D}'}\varphi$).

Whenever Ω is an open subset of $\mathbb{R}^n$ and $k \in \mathbb{N}$, we denote by $C^k(\Omega)$ the space of functions u which are differentiable up to order k; and, whenever $\alpha \in (0,1)$, we denote by $C^{k,\alpha}(\Omega)$ the space of functions u for which all partial derivatives at order k are Hölder-continuous with exponent α.

Whenever Ω is a smooth subset of $\mathbb{R}^n$, the group of all diffeomorphisms $s : \Omega \to \Omega$ with $\det(\nabla s) \equiv 1$ will be denoted by $G(\Omega)$.

We shall refer to a measurable set $X \subset \mathbb{R}^n$ as a *small set* if it has Hausdorff dimension at most $n - 1$.

The vector space of real $n \times n$ matrices will be denoted by $M_n(\mathbb{R})$. The trace of a matrix M will be denoted by tr M. The $n \times n$ identity matrix will be denoted by I_n. Whenever M is an element of $M_n(\mathbb{R})$, its transposed matrix will be denoted by M^T; thus $M^T = (m'_{ij})$ with $m'_{ij} = m_{ji}$. The sets of symmetric matrices ($M^T = M$), symmetric matrices with nonnegative eigenvalues ($M \geq 0$), antisymmetric matrices ($M^T = -M$) and orthogonal matrices ($MM^T = I_n$) will be respectively denoted by $S_n(\mathbb{R})$, $S_n^+(\mathbb{R})$, $A_n(\mathbb{R})$ and $O_n(\mathbb{R})$.

Finally, let us say a word about where to find the definitions of the basic objects in optimal mass transportation: the notations $I[\pi]$, $\Pi(\mu, \nu)$, $J(\varphi, \psi)$, Φ_c are defined in Theorem 1.3 of Chapter 1; $\mathcal{T}_c(\mu, \nu)$ in formula (5); $W_p(\mu, \nu)$ and $\mathcal{T}_p(\mu, \nu)$ in Theorem 7.3 of Chapter 7.