Acta Math., 177 (1996), 113–161
© 1996 by Institut Mittag-Leffler. All rights reserved

The geometry of optimal transportation

by

WILFRID GANGBO

Georgia Institute of Technology
Atlanta, GA, U.S.A.

and

ROBERT J. MCCANN

Brown University
Providence, RI, U.S.A.
and
Institut des Hautes Etudes Scientifiques
Bures-sur-Yvette, France

Contents

Introduction

In 1781, Monge [30] formulated a question which occurs naturally in economics: Given two sets $U, V \subset \mathbf{R}^d$ of equal volume, find the optimal volume-preserving map between them, where optimality is measured against a cost function $c(\mathbf{x},\mathbf{y}) \geqslant 0$. One views the first set as being uniformly filled with mass, and $c(\mathbf{x},\mathbf{y})$ as being the cost per unit mass for transporting material from $\mathbf{x} \in U$ to $\mathbf{y} \in V$; the optimal map minimizes the total cost of redistributing the mass of U through V. Monge took the Euclidean distance $c(\mathbf{x},\mathbf{y}) = |\mathbf{x}-\mathbf{y}|$ to be his cost function, but even for this special case, two centuries elapsed

Both authors gratefully acknowledge the support provided by postdoctoral fellowships: WG at the Mathematical Sciences Research Institute, Berkeley, CA, and RJM from the Natural Sciences and Engineering Research Council of Canada.

before Sudakov [42] showed that such a map exists. In the meantime, Monge's problem turned out to be the prototype for a class of questions arising in differential geometry, infinite-dimensional linear programming, functional analysis, mathematical economics and in probability and statistics—for references see [31], [26]; the Academy of Paris offered a prize for its solution [16], which was claimed by Appell [5], while Kantorovich received a Nobel prize for related work in economics [23].

What must have been apparent from the beginning was that the solution would not be unique [5], [21]. Even on the line the reason is clear: in order to shift a row of books one place to the right on a bookshelf, two equally efficient algorithms present themselves: (i) shift each book one place to the right; (ii) move the leftmost book to the right-hand side, leaving the remaining books fixed. More recently, two separate lines of authors—including Brenier on the one hand and Knott and Smith, Cuesta-Albertos, Matrán and Tuero-Díaz, Rüschendorf and Rachev, and Abdellaoui and Heinich on the other—have realized that for the distance squared $c(\mathbf{x},\mathbf{y})=|\mathbf{x}-\mathbf{y}|^2$, not only does an optimal map exist which is unique [7], [11], [8], [2], [12], but it is characterized as the gradient of a convex function [25], [7], [40], [38], [8]. Founded on the Kantorovich approach, their methods apply equally well to non-uniform distributions of mass throughout $\mathbf{R}^d$, as to uniform distributions on U and V; all that matters is that the total masses be equal. The novelty of this result is that, like Riemann's mapping theorem in the plane, it singles out a map with preferred geometry between U and V; a polar factorization theorem for vector fields [7] and a Brunn–Minkowski inequality for measures [28] are among its consequences. In the wake of these discoveries, many fundamental questions stand exposed: What features of the cost function determine existence and uniqueness of optimal maps? What geometrical properties characterize the maps for other costs? Can this geometry be exploited fruitfully in applications? Finally, we note that concave functions of the distance $|\mathbf{x}-\mathbf{y}|$ form the most interesting class of costs: from an economic point of view, they represent shipping costs which increase with the distance, even while the cost per mile shipped goes down.

Here these questions are resolved for costs from two important classes: $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ with h strictly convex, or $c(\mathbf{x},\mathbf{y})=l(|\mathbf{x}-\mathbf{y}|)$ with $l\geqslant 0$ strictly concave. For convex costs, a theory parallel to that for distance squared has been developed: the optimal map exists and is uniquely characterized by its geometry. This map (5) depends explicitly on the gradient of the cost, or rather on its inverse map $(\nabla h)^{-1}$, which indicates why *strict* convexity or concavity should be essential for uniqueness. Although explicit solutions are more awkward to obtain, we have no reason to believe that they should be any worse behaved than those for distance squared (see e.g. the regularity theory developed by Caffarelli [9] when $c(\mathbf{x},\mathbf{y})=|\mathbf{x}-\mathbf{y}|^2$).

For concave functions of the distance, the picture which emerges is rather different. Here the optimal maps will not be smooth, but display an intricate structure which—for us—was unexpected; it seems equally fascinating from the mathematical and the economic point of view. A separate paper explores this structure fully on the line [29], where the situation is already far from simple and our conclusions yield some striking implications. To describe one effect in economic terms: the concavity of the cost function favors a long trip and a short trip over two trips of average length; as a result, it can be efficient for two trucks carrying the same commodity to pass each other traveling opposite directions on the highway: one truck must be a local supplier, the other on a longer haul. In optimal solutions, such 'pathologies' may nest on many scales, leading to a natural hierarchy among regions of supply (and of demand). For the present we are content to prove existence and uniqueness results, both on the line and in higher dimensions, which characterize the solutions geometrically. As for convex costs, the results are obtained through constructive geometrical arguments requiring only minimal hypotheses on the mass distributions.

To state the problem more precisely requires a bit of notation. Let $\mathcal{M}(\mathbf{R}^d)$ denote the space of non-negative Borel measures on $\mathbf{R}^d$ with finite total mass, and $\mathcal{P}(\mathbf{R}^d)$ the subset of probability measures—measures for which $\mu[\mathbf{R}^d]=1$.

Definition 0.1. A measure $\mu\in\mathcal{M}(\mathbf{R}^d)$ and a Borel map $\mathbf{s}:\Omega\subset\mathbf{R}^d\to\mathbf{R}^n$ induce a (Borel) measure $\mathbf{s}_{\#}\mu$ on $\mathbf{R}^n$—called the *push-forward of* μ *through* $\mathbf{s}$—and defined by $\mathbf{s}_{\#}\mu[V]:=\mu[\mathbf{s}^{-1}(V)]$ for Borel $V\subset\mathbf{R}^n$.

One says that $\mathbf{s}$ *pushes* μ *forward* to $\mathbf{s}_{\#}\mu$. If $\mathbf{s}$ is defined μ-almost everywhere, one may also say that $\mathbf{s}$ is *measure-preserving* between μ and $\mathbf{s}_{\#}\mu$; then the push-forward $\mathbf{s}_{\#}\mu$ will be a probability measure if μ is. It is worth pointing out that $\mathbf{s}_{\#}$ maps $\mathcal{M}(\mathbf{R}^d)$ linearly to $\mathcal{M}(\mathbf{R}^n)$. For a Borel function f on $\mathbf{R}^n$, the change of variables theorem states that, when either integral is defined,

$$\int_{\mathbf{R}^n} f\,d(\mathbf{s}_{\#}\mu)=\int_{\Omega\subset\mathbf{R}^d} f(\mathbf{s}(\mathbf{x}))\,d\mu(\mathbf{x}). \tag{1}$$

Monge's problem generalizes thus: Given two measures $\mu,\nu\in\mathcal{P}(\mathbf{R}^d)$, is the infimum

$$\inf_{\mathbf{s}_{\#}\mu=\nu}\int c(\mathbf{x},\mathbf{s}(\mathbf{x}))\,d\mu(\mathbf{x}) \tag{2}$$

attained among mappings $\mathbf{s}$ which push μ forward to ν, and, if so, what is the optimal map? Here the measures μ and ν, which need not be discrete, might represent the distributions for production and consumption of some commodity. The problem would then be to decide which producer should supply each consumer for total transportation costs to be a minimum. Although Monge and his successors had deep insights into (2) (see e.g. [18]), this problem remained unsolved due to its non-linearity in $\mathbf{s}$, and intractability of the set of mappings pushing forward μ to ν.

In 1942, a breakthrough was achieved by Kantorovich [21], [22], who formulated a relaxed version of the problem as a linear optimization on a convex domain. Instead of minimizing over maps which push μ forward to ν, he considered joint measures $\gamma \in \mathcal{P}(\mathbf{R}^d \times \mathbf{R}^d)$ which have μ and ν as their *marginals*: $\mu[U]=\gamma[U\times\mathbf{R}^d]$ and $\gamma[\mathbf{R}^d\times U]=\nu[U]$ for Borel $U\subset\mathbf{R}^d$. The set of such measures, denoted $\Gamma(\mu,\nu)$, forms a convex subset of $\mathcal{P}(\mathbf{R}^d\times\mathbf{R}^d)$. Kantorovich's problem was to minimize the *transport cost*

$$\mathcal{C}(\gamma) := \int c(\mathbf{x},\mathbf{y})\,d\gamma(\mathbf{x},\mathbf{y}) \tag{3}$$

among joint measures γ in $\Gamma(\mu,\nu)$, to obtain

$$\inf_{\gamma\in\Gamma(\mu,\nu)} \mathcal{C}(\gamma). \tag{4}$$

Linearity makes the Kantorovich problem radically simpler than that of Monge; a continuity-compactness argument at least guarantees that the infimum (4) will be attained. Moreover, the Kantorovich minimum provides a lower bound for that of Monge: whenever $\mathbf{s}_{\#}\mu=\nu$, the map on $\mathbf{R}^d$ taking $\mathbf{x}$ to $(\mathbf{x},\mathbf{s}(\mathbf{x}))\in\mathbf{R}^d\times\mathbf{R}^d$ pushes μ forward to $(\mathbf{id}\times\mathbf{s})_{\#}\mu\in\Gamma(\mu,\nu)$; a change of variables (1) shows that the Kantorovich cost $\mathcal{C}((\mathbf{id}\times\mathbf{s})_{\#}\mu)$ coincides with the Monge cost of the mapping $\mathbf{s}$. Thus Kantorovich's infimum encompasses a larger class of objects than that of Monge.

Rephrasing our questions in this framework: Can a mapping $\mathbf{s}$ which solves the Monge problem be recovered from a Kantorovich solution γ—i.e., will a minimizer γ for $\mathcal{C}(\cdot)$ be of the form $(\mathbf{id}\times\mathbf{s})_{\#}\mu$? Under what conditions will solutions $\mathbf{s}$ and γ to the Monge and Kantorovich problems be unique? Can the optimal maps be characterized geometrically? Is there a qualitative (but rigorous) theory of their features?

For strictly convex cost functions $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ (satisfying a condition at infinity) our results will be as follows: Assuming that μ is absolutely continuous with respect to Lebesgue, it is true that the optimal solution γ to the Kantorovich problem is unique. Moreover $\gamma=(\mathbf{id}\times\mathbf{s})_{\#}\mu$, so the Monge problem is solved as well. The optimal map is of the form

$$\mathbf{s}(\mathbf{x}) = \mathbf{x} - \nabla h^{-1}(\nabla\psi(\mathbf{x})), \tag{5}$$

and it is uniquely characterized by a geometrical condition known as *c-concavity* of the potential $\psi: \mathbf{R}^d\to\mathbf{R}\cup\{-\infty\}$. This characterization adapts the work of Rüschendorf [34], [35, esp. (73)] from the Kantorovich setting (with general costs) to that of Monge. Discovered independently by us [20] and Caffarelli [10], it encompasses both recent progress in this direction [41], [13], [36], [37] and the earlier work of Brenier and others on the cost $h(\mathbf{x})=\frac{1}{2}|\mathbf{x}|^2$—which is special in that is has the identity map $\nabla h=\mathbf{id}$ as its gradient; the optimal map $\mathbf{s}(\mathbf{x})=\mathbf{x}-\nabla\psi(\mathbf{x})$ turns out to be pure gradient for this cost. When μ fails to be absolutely continuous but the cost is a derivative smoother, our conclusions persist as long as μ vanishes on any rectifiable set of dimension $d-1$.

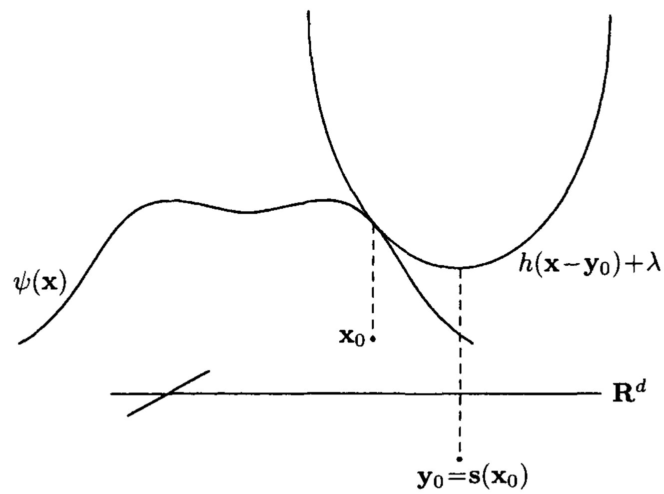
For the economically relevant costs—$c(\mathbf{x},\mathbf{y})$ a strictly concave function of the distance $|\mathbf{x}-\mathbf{y}|$—the Kantorovich minimizer γ need not be of the form $(\mathbf{id}\times\mathbf{s})_{\#}\mu$ unless the measures μ and ν are disjointly supported. Rather, because c is a metric on $\mathbf{R}^d$, the mass which is common to μ and ν must not be moved; it can be subtracted from the diagonal of γ. What remains will be a joint measure γ_o having the positive and negative parts of $\mu-\nu$ for marginals. If the mass of $\mu_o:=[\mu-\nu]_+$ and $\nu_o:=[\nu-\mu]_+$ is not too densely interwoven, and μ_o vanishes on rectifiable sets of dimension $d-1$, then γ will be unique and $\gamma_o=(\mathbf{id}\times\mathbf{s})_{\#}\mu_o$. The optimal mapping $\mathbf{s}$ can be quite complex—as a one-dimensional analysis indicates—but it is derived from a potential ψ through (5) (see Figure 1) in any case. However, a slightly stronger condition than c-concavity of ψ characterizes the solution.

Regarding the hypothesis on μ we mention the following: certainly μ cannot concentrate on sets which are too small if it is to be pushed forward to every possible measure ν. But how small is too small? For costs which norm $\mathbf{R}^d$, Sudakov proposed dimension $d-1$ as a quantitative condition to ensure existence of an optimal map [42]. When $c(\mathbf{x},\mathbf{y})=|\mathbf{x}-\mathbf{y}|^2$, McCann verified sufficiency of this condition both for existence and uniqueness of optimal maps [27]. A more precise relationship between μ and c was formulated by Cuesta-Albertos and Tuero-Díaz; it implies existence and uniqueness results for quite general costs when the target measure ν is *discrete*: $\nu:=\sum_i \lambda_i\delta_{\mathbf{x}_i}$ [14], [2], [1].

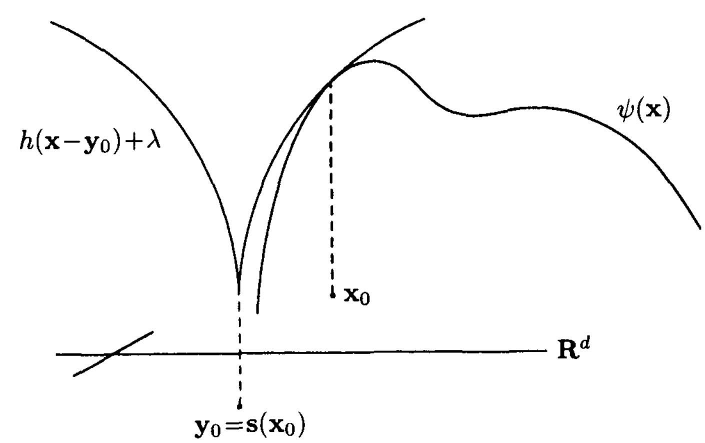
Before concluding this introduction, there are two further issues which cannot go unmentioned: our methods of proof, and the duality theory which—in the past—has been the principle tool for investigating the Monge–Kantorovich problem. The spirit of our proof can be apprehended in the context (already well understood [15], [19]) of strictly convex costs on the line. Let $\mu,\nu\in\mathcal{P}(\mathbf{R})$ be measures on the real line, the first without atoms, $\mu[\{x\}]=0$, and consider the optimal joint measure $\gamma\in\Gamma(\mu,\nu)$ corresponding to a cost $c(x,y)$. Any two points (x,y) and (x',y') from the *support* of γ, meaning the smallest closed set in $\mathbf{R}\times\mathbf{R}$ which carries the full mass of γ, will satisfy the inequality

$$c(x,y)+c(x',y')\leqslant c(x,y')+c(x',y); \tag{6}$$

otherwise it would be more efficient to move mass from x to y' and x' to y. For $c(x,y)=h(x-y)$, strict convexity of h and (6) imply $(x'-x)(y'-y)\geqslant 0$; in other words, $\operatorname{spt}\gamma$



(a) For strictly convex costs $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$.



(b) For $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})=l(|\mathbf{x}-\mathbf{y}|)$ with l strictly concave and increasing.

Fig. 1. The optimal map $\mathbf{y}_0=\mathbf{s}(\mathbf{x}_0)$ may be visualized by finding a shifted translate of $h(\mathbf{x})$ which is tangent to the potential ψ at $\mathbf{x}_0$; then $\nabla\psi(\mathbf{x}_0)=\nabla h(\mathbf{x}_0-\mathbf{y}_0)$, $D^2\psi(\mathbf{x}_0)\leqslant D^2h(\mathbf{x}_0-\mathbf{y}_0)$ and $(\mathbf{x}_0,\mathbf{y}_0)\in\partial^c\psi$. Where ψ is differentiable, strict convexity of h guarantees this translate to be unique.

will be a monotone subset of the plane. Apart from vertical segments—of which there can only be countably many—such a set is contained in the graph of a non-decreasing function $s\colon\mathbf{R}\to\mathbf{R}$. This function is the optimal map. The fact that μ has no atoms means that none of its mass concentrates under vertical segments in $\operatorname{spt}\gamma$, and is used to verify $\nu=s_{\#}\mu$. It is not hard to show that only one non-decreasing map pushes μ forward to ν, so s is uniquely determined μ-almost everywhere.

The generalization of this argument to higher dimensions was explored in [27] to sharpen results for the cost $c(\mathbf{x},\mathbf{y})=|\mathbf{x}-\mathbf{y}|^2$; our proof follows the strategy there. At the same time, we build on many ideas introduced to the transportation problem by other authors. The connection of c-concavity with mass transport was first explored by Rüschendorf [35], who used it to characterize the optimal measures γ of Kantorovich; he later constructed certain unique optimal maps for convex costs [36, §3]. The related notion of c-cyclical monotonicity is also essential; formulated by Smith and Knott [41] in analogy with a classical notion of Rockafellar [32], it supplements inequality (6). One fact that continues to amaze us is that—for the costs $c(\mathbf{x},\mathbf{y})$ we deal with—not a single desirable property of concave functions has failed to have a serviceable analog among c-concave functions. Even the kernel of Aleksandrov's uniqueness proof [4] for surfaces of prescribed integral curvature is preserved in our uniqueness argument. A non-negligible part of our effort in this paper has been devoted to developing the theory of c-concave functions as a general tool, and we hope that this theory may prove useful in other applications.

Since the literature on the Monge–Kantorovich problem is vast and fragmented [31], we have endeavoured to present a treatment which is largely self-contained. In the background section and appendices, we have therefore collected together some results which could also be found elsewhere. Absent from the discussion is any reference to the maximization problem dual to (4), discovered by Kantorovich [21] for cost functions which metrize $\mathbf{R}^d$. Subsequently developed by many authors, duality theory flourished into a powerful tool for exploring mass transport and similar problems; quite general Monge–Kantorovich duality relations were obtained by Kellerer in [24], and further references are there given. Our results are not predicated on that theory, but rather, imply duality as a result. One advantage of this approach is that the main theorems and their proofs are seen to be purely geometrical—they require few assumptions, and do not rely even on finiteness of the infimum (4). However, the potential ψ that we *construct* can generally be shown to be the maximizer for a suitable dual problem. This fact is clearer from our work in [20], where many of these results were first announced; a completely different approach, based on the Euler–Lagrange equation for the dual problem, is given there. A main conclusion, both there and here, is that for the cost functions we deal with, the potential $\psi(\mathbf{x})$—whether constructed geometrically or extracted as a solution to some dual problem—specifies both which direction and how far to move the mass of μ which sits near $\mathbf{x}$. If the cost is not *strictly* convex—so that ∇h is not one-to-one—uniqueness may fail, and further information be required to determine an optimal mapping; for radial costs $c(\mathbf{x},\mathbf{y})=l(|\mathbf{x}-\mathbf{y}|)$, the potential specifies the direction of transport but not the distance—cf. [42], [18] and Figure 1.

The remainder of this paper is organized as follows: The first section provides a summary of our main theorems, preceded by the necessary definitions and followed by a continuation of the discussion, while the second section recounts background results from the literature which apply to general cost functions and measure spaces. The narrative then splits into two parallel parts, which treat strictly convex costs and strictly concave functions of the distance separately. Each part in turn divides into two sections, which focus on the construction of a map $\mathbf{s}$ from the optimal measure γ, and the unique characterization of this map as a geometrical object. Three appendices are also provided. The first reviews some facts of life concerning Legendre transforms and conjugate costs, while the second provides a few examples of c-concave potentials. The last appendix is technical: it develops the structure and regularity properties which are required of c-concave potentials (infimal convolutions with $h(\mathbf{x})$).

It is a pleasure to express our gratitude to L. Craig Evans and Jill Pipher for their continuing encouragement and support. Fruitful discussions were also provided by Stephen Semmes and Jan-Philip Solovej, while the figures were drawn by Marie-Claude Vergne. We thank Giovanni Alberti and Ludger Rüschendorf for references, and note that this work had essentially been completed when we learned of Caffarelli's concurrent discovery [10] of similar results concerning convex costs.

1. Summary of main results

To begin, we recall the definition of c-concavity. It adapts the notion of a concave function—i.e., an infimum of affine functions—to the geometry of the cost $c(\mathbf{x},\mathbf{y})$, and will play a vital role. Except as noted, the cost functions considered here will be of the form $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ where h is continuous on $\mathbf{R}^d$.

Definition 1.1. A function $\psi:\mathbf{R}^d\to\mathbf{R}\cup\{-\infty\}$, not identically $-\infty$, is said to be *c-concave* if it is the infimum of a family of translates and shifts of $h(\mathbf{x})$: i.e., there is a set $\mathcal{A}\subset\mathbf{R}^d\times\mathbf{R}$ such that

$$\psi(\mathbf{x}):=\inf_{(\mathbf{y},\lambda)\in\mathcal{A}} c(\mathbf{x},\mathbf{y})+\lambda. \tag{7}$$

Without further structure on h, c-concavity has limited utility [6], [35], but for suitable costs it will become a powerful tool. For the quadratic cost $h(\mathbf{x})=\frac{1}{2}|\mathbf{x}|^2$, c-concavity of ψ turns out to be equivalent to convexity of $\frac{1}{2}\mathbf{x}^2-\psi(\mathbf{x})$ in the usual sense through the identity $c(\mathbf{x},\mathbf{y})=h(\mathbf{x})-\langle\mathbf{x},\mathbf{y}\rangle+h(\mathbf{y})$. More generally, we consider convex costs $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ for which

(H1) $h:\mathbf{R}^d\to[0,\infty)$ is strictly convex.

To handle measures with unbounded support, we also assume that the cost grows superlinearly at large $|\mathbf{x}|$ while the curvature of its level sets decays:

(H2) Given height $r<\infty$ and angle $\theta\in(0,\pi)$: whenever $\mathbf{p}\in\mathbf{R}^d$ is far enough from the origin, one can find a cone

$$K(r,\theta,\hat{\mathbf{z}},\mathbf{p}) := \{\mathbf{x}\in\mathbf{R}^d \mid |\mathbf{x}-\mathbf{p}|\cdot|\mathbf{z}|\cos(\tfrac{1}{2}\theta) \leqslant \langle \mathbf{z},\mathbf{x}-\mathbf{p}\rangle \leqslant r|\mathbf{z}|\} \tag{8}$$

with vertex at $\mathbf{p}$ (and $\mathbf{z}\in\mathbf{R}^d\setminus\{\mathbf{0}\}$) on which $h(\mathbf{x})$ assumes its maximum at $\mathbf{p}$;

(H3) $\lim h(\mathbf{x})/|\mathbf{x}|=+\infty$ as $|\mathbf{x}|\to\infty$.

Cost functions satisfying (H1)–(H3) include all quadratic costs $h(\mathbf{x})=\langle\mathbf{x},P\mathbf{x}\rangle$ with P positive definite, and radial costs $h(\mathbf{x})=l(|\mathbf{x}|)$ which grow faster than linearly. Occasionally, we relax strict convexity or require additional smoothness:

(H4) $h:\mathbf{R}^d\to\mathbf{R}$ is convex;

(H5) $h(\mathbf{x})$ is differentiable and its gradient is locally Lipschitz: $h\in C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$.

For these costs, c-concavity generalizes concavity in the usual sense, but we shall show that it is almost as strong a notion. In particular, except for a set of dimension $d-1$, a c-concave function ψ will be differentiable where it is finite; it will be twice differentiable almost everywhere in the sense of Aleksandrov [39, notes to §1.5].

With some final definitions, our first main theorem is stated. We say that a joint measure $\gamma\in\mathcal{P}(\mathbf{R}^d\times\mathbf{R}^d)$ is *optimal* if it minimizes $\mathcal{C}(\gamma)$ among the measures $\Gamma(\mu,\nu)$ which share its marginals, μ and ν. Since differentiability of the cost is not assumed, we define $(\nabla h)^{-1}:=\nabla h^*$ through the Legendre transform (10) in its absence. As before, $\mathbf{id}$ denotes the identity mapping $\mathbf{id}(\mathbf{x})=\mathbf{x}$ on $\mathbf{R}^d$, while $\circ$ denotes composition.

THEOREM 1.2 (for strictly convex costs). *Fix $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ with h strictly convex satisfying* (H1)–(H3), *and Borel probability measures μ and ν on $\mathbf{R}^d$. If μ is absolutely continuous with respect to Lebesgue then*

(i) *there is a c-concave function ψ on $\mathbf{R}^d$ for which the map $\mathbf{s}:=\mathbf{id}-(\nabla h)^{-1}\circ\nabla\psi$ pushes μ forward to ν;*

(ii) *this map $\mathbf{s}(\mathbf{x})$ is uniquely determined (μ-almost everywhere) by* (i);

(iii) *the joint measure $\gamma:=(\mathbf{id}\times\mathbf{s})_{\#}\mu$ is optimal;*

(iv) *γ is the only optimal measure in $\Gamma(\mu,\nu)$—except (trivially) when $\mathcal{C}(\gamma)=\infty$.*

If μ fails to be absolutely continuous with respect to Lebesgue but vanishes on rectifiable sets of dimension $d-1$, then (i)–(iv) *continue to hold provided $h\in C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$.*

Here a *rectifiable set of dimension $d-1$* refers to any Borel set $U\subset\mathbf{R}^d$ which may be covered using countably many $(d-1)$-dimensional Lipschitz submanifolds of $\mathbf{R}^d$.[1]

[1] *Remark added in proof.* Although not proven here, the theorems remain true even if one insists that the covering submanifolds be graphs of differences of convex functions: $(c-c)$-hypersurfaces in the language of Zajíček [43].

To illustrate the theorem in an elementary context, we verify the optimality of $\mathbf{t}(\mathbf{x})=\lambda\mathbf{x}-\mathbf{z}$ when μ and ν are translated dilates of each other: $\nu:=\mathbf{t}_{\#}\mu$ [13]. For $\lambda\geqslant 0$, $\mathbf{z}\in\mathbf{R}^d$ and convex costs $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$, observe the c-concavity of

$$\psi(\mathbf{x}) := (1-\lambda)^{-1}h(\mathbf{x}(1-\lambda)+\mathbf{z})$$

proved in Lemma B.1 (iv)–(vi) (if $\lambda=1$ take $\psi(\mathbf{x}):=\langle\mathbf{x},\nabla h(\mathbf{z})\rangle$). This potential ψ induces the map $\mathbf{s}=\mathbf{t}$ through (5). Since $\mathbf{t}$ pushes forward μ to ν, it must be the unique map of Theorem 1.2.

Motivated by economics, we now turn to costs of the form $c(\mathbf{x},\mathbf{y})=l(|\mathbf{x}-\mathbf{y}|)$, where $l\colon[0,\infty)\to[0,\infty)$ is strictly concave. The optimal solutions for these costs respect different symmetries. It will often be convenient to assume continuity of the cost (at the origin) and $l(0)=0$, but these additional restrictions are *not* required for Theorem 1.4. With a few caveats, our results could also be extended to strictly concave functions l which increase from $l(0)=-\infty$, but we restrict our attention to $l\geqslant 0$ instead. *For these costs, l will be strictly increasing as a consequence of its concavity.*

With this second class of costs come two new complications. Since $c(\mathbf{x},\mathbf{y})$ induces a metric on $\mathbf{R}^d$, any mass which is shared between $\mu,\nu\in\mathcal{P}(\mathbf{R}^d)$ must not be moved by a transportation plan γ that purports to be optimal. This mass is defined through the Jordan decomposition of $\mu-\nu$ into its unique representation $\mu_o-\nu_o$ as a difference of two non-negative mutually singular measures: $\mu_o:=[\mu-\nu]_+$ and $\nu_o:=[\nu-\mu]_+$. The *shared mass* $\mu\wedge\nu:=\mu-\mu_o=\nu-\nu_o$ is the maximal measure in $\mathcal{M}(\mathbf{R}^d)$ to be dominated by both μ and ν. Since one expects to find this mass on the diagonal subspace $D:=\{(\mathbf{x},\mathbf{x})\,|\,\mathbf{x}\in\mathbf{R}^d\}$ of $\mathbf{R}^d\times\mathbf{R}^d$ under γ, it is convenient to denote the *restriction of γ to the diagonal* by $\gamma_d[S]:=\gamma[S\cap D]$. The off-diagonal restriction γ_o is then given by $\gamma_o=\gamma-\gamma_d$.

The second complication is the singularity in $c(\mathbf{x},\mathbf{y})$ at $\mathbf{x}=\mathbf{y}$, which renders c-concavity too feeble to characterize the optimal map uniquely. For this reason, a refinement must be introduced to monitor the location $V\subset\mathbf{R}^d$ of the singularity:

Definition 1.3. Let $V\subset\mathbf{R}^d$. A c-concave function ψ on $\mathbf{R}^d$ is said to be the *c-transform of a function on V* if (7) holds with $\mathcal{A}\subset V\times\mathbf{R}$.

A moment's reflection reveals the existence of some function $\phi\colon V\to\mathbf{R}\cup\{-\infty\}$ for which

$$\psi(\mathbf{x}) = \inf_{\mathbf{y}\in V} c(\mathbf{x},\mathbf{y})-\phi(\mathbf{y}) \tag{9}$$

whenever Definition 1.3 is satisfied.

Finally, as with convex costs, it is a vital feature of $h(\mathbf{x})=l(|\mathbf{x}|)$ that the gradient map ∇h be invertible on its image. This follows from strict concavity of $l\geqslant 0$ since

$l'(\lambda)\geqslant 0$ will be one-to-one. Should differentiability of l fail, we define $(\nabla h)^{-1}:=\nabla h^*$ using (11) this time. The *support* $\operatorname{spt}\mu$ of a measure $\mu\in\mathcal{M}(\mathbf{R}^d)$ refers to the smallest closed set $U\subset\mathbf{R}^d$ of full mass: $\mu[U]=\mu[\mathbf{R}^d]$.

THEOREM 1.4 (for strictly concave cost as a function of distance). *Use $l:[0,\infty)\rightarrow[0,\infty)$ strictly concave to define $c(\mathbf{x},\mathbf{y}):=h(\mathbf{x}-\mathbf{y}):=l(|\mathbf{x}-\mathbf{y}|)$. Let μ and ν be Borel probability measures on $\mathbf{R}^d$ and define $\mu_o:=[\mu-\nu]_+$ and $\nu_o:=[\nu-\mu]_+$. If μ_o vanishes on $\operatorname{spt}\nu_o$ and on rectifiable sets of dimension $d-1$ then*

(i) *the c-transform $\psi:\mathbf{R}^d\rightarrow\mathbf{R}$ of some function on $\operatorname{spt}\nu_o$ induces a map $\mathbf{s}:=\mathbf{id}-(\nabla h)^{-1}\circ\nabla\psi$ which pushes μ_o forward to ν_o;*

(ii) *the map $\mathbf{s}(\mathbf{x})$ is uniquely determined μ_o-almost everywhere by* (i);

(iii) *there is a unique optimal measure γ in $\Gamma(\mu,\nu)$—except when $C(\gamma)=\infty$;*

(iv) *the restriction of γ to the diagonal is given by $\gamma_d=(\mathbf{id}\times\mathbf{id})_\#(\mu-\mu_o)$;*

(v) *the off-diagonal part of $\gamma=\gamma_d+\gamma_o$ is given by $\gamma_o=(\mathbf{id}\times\mathbf{s})_\#\mu_o$.*

The hypotheses of this theorem are satisfied when μ and ν are given by continuous densities $f,g\in C(\mathbf{R}^d)$ with respect to Lebesgue: $d\mu(\mathbf{x})=f(\mathbf{x})\,d\mathbf{x}$ and $d\nu(\mathbf{y})=g(\mathbf{y})\,d\mathbf{y}$. Alternately, if $f(\mathbf{x})=\chi_U(\mathbf{x})$ and $g(\mathbf{y})=\chi_V(\mathbf{y})$ are characteristic functions of two equal volume sets—an open set U and a closed set V—then Theorem 1.4 yields an optimal map given by $\mathbf{s}(\mathbf{x})=\mathbf{x}$ on $U\cap V$.

As for convex costs, explicit solutions may be computed to problems with appropriate symmetry. For concave functions of the distance, suitable symmetries include reflection through a sphere or through a plane (for details refer to Appendix B):

Example 1.5 (reflections). Take c and μ from Theorem 1.4. If μ is supported on the unit ball, then the spherical inversion $\mathbf{s}(\mathbf{x}):=\mathbf{x}/|\mathbf{x}|^2$ will be the optimal map between μ and $\mathbf{s}_\#\mu$. If μ vanishes on the half-space $x_1>0$ in $\mathbf{R}^d$, then reflection through the plane $x_1=0$ will be the optimal map between μ and its mirror image.

Explicit solutions may also be obtained whenever the target measure ν concentrates on finitely many points: $\operatorname{spt}\nu=\{\mathbf{y}_1,\mathbf{y}_2,...,\mathbf{y}_k\}$. The initial measure μ is arbitrary provided it vanishes on small enough sets. For convex costs, we also need Remark 4.6: the potential ψ of Theorem 1.2 may be assumed to be the c-transform of a function on $\operatorname{spt}\nu$.

Example 1.6 (target measures of finite support). Take μ, ν, c and h from Theorem 1.2 or 1.4. If $\operatorname{spt}\nu=\{\mathbf{y}_1,\mathbf{y}_2,...,\mathbf{y}_k\}\subset\mathbf{R}^d$ then the optimal map is of the form $\mathbf{s}(\mathbf{x})=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$, where ψ is the c-transform of a function on $\operatorname{spt}\nu$. In view of (9),

$$\psi(\mathbf{x})=\inf_{j=1,\ldots,k}c(\mathbf{x},\mathbf{y}_j)+\lambda_j.$$

From this family of maps, the unique solution is selected by finding any k constants $\lambda_j\in\mathbf{R}$ consistent with the mass constraints $\mu[\mathbf{s}^{-1}(\mathbf{y}_j)]=\nu[\mathbf{y}_j]$, $j=1,...,k$.

The constants λ_j should be easy to compute numerically; indeed, we speculate that flowing along the vector field $v_j(\lambda_1, ..., \lambda_k) := \mu[\mathbf{s}^{-1}(\mathbf{y}_j)] - \nu[\mathbf{y}_j]$ through $\mathbf{R}^k$ will always lead to a solution. When $k=2$, the optimal map is given by

$$\mathbf{s}(\mathbf{x}) := \begin{cases} \mathbf{y}_1, & \text{where } c(\mathbf{x}, \mathbf{y}_1) + \lambda_1 < c(\mathbf{x}, \mathbf{y}_2) + \lambda_2, \\ \mathbf{y}_2, & \text{elsewhere.} \end{cases}$$

A sketch (Figure 2) of level sets for $c(\mathbf{x}, \mathbf{y}_1) - c(\mathbf{x}, \mathbf{y}_2)$ illustrates these domains in the plane. Shading indicates the region that $\mathbf{s}(\mathbf{x})$ maps to $\mathbf{y}_1$; its size is adjusted with $\lambda_2 - \lambda_1$ to yield the right amount $\nu[\mathbf{y}_1]$ of mass for μ, and this is the only way in which the measure μ affects the optimal map. The shape of these domains plays a key role even when $\operatorname{spt}\nu$ contains more than two points: then the complicated regions $\mathbf{s}^{-1}(\mathbf{y}_j)$ of Example 1.6 arise as the intersection of $k-1$ two-point domains. Unboundedness of both domains distinguishes convex costs from strictly concave functions $l \geqslant 0$ of the distance, while half-spaces are characteristic of quadratic costs and of $\lambda_1 = \lambda_2$. Finally, Figure 2 also shows why vanishing of μ on submanifolds of dimension $d-1$ should be required to guarantee a unique map.

For both convex and concave costs $c(\mathbf{x}, \mathbf{y}) = h(\mathbf{x} - \mathbf{y})$, the inverse map to ∇h is the gradient ∇h^* of a dual function $h^*(\mathbf{y})$ known as the Legendre transform. As an example, $h(\mathbf{x}) = |\mathbf{x}|^p/p$ implies $h^*(\mathbf{y}) = |\mathbf{y}|^q/q$ with $p^{-1} + q^{-1} = 1$; here $p \in \mathbf{R}$ but $p \neq 0, 1$. More generally, if the cost is convex then $h^*: \mathbf{R}^d \to \mathbf{R} \cup \{+\infty\}$ is given by

$$h^*(\mathbf{y}) := \sup_{\mathbf{x} \in \mathbf{R}^d} \langle \mathbf{x}, \mathbf{y} \rangle - h(\mathbf{x}). \tag{10}$$

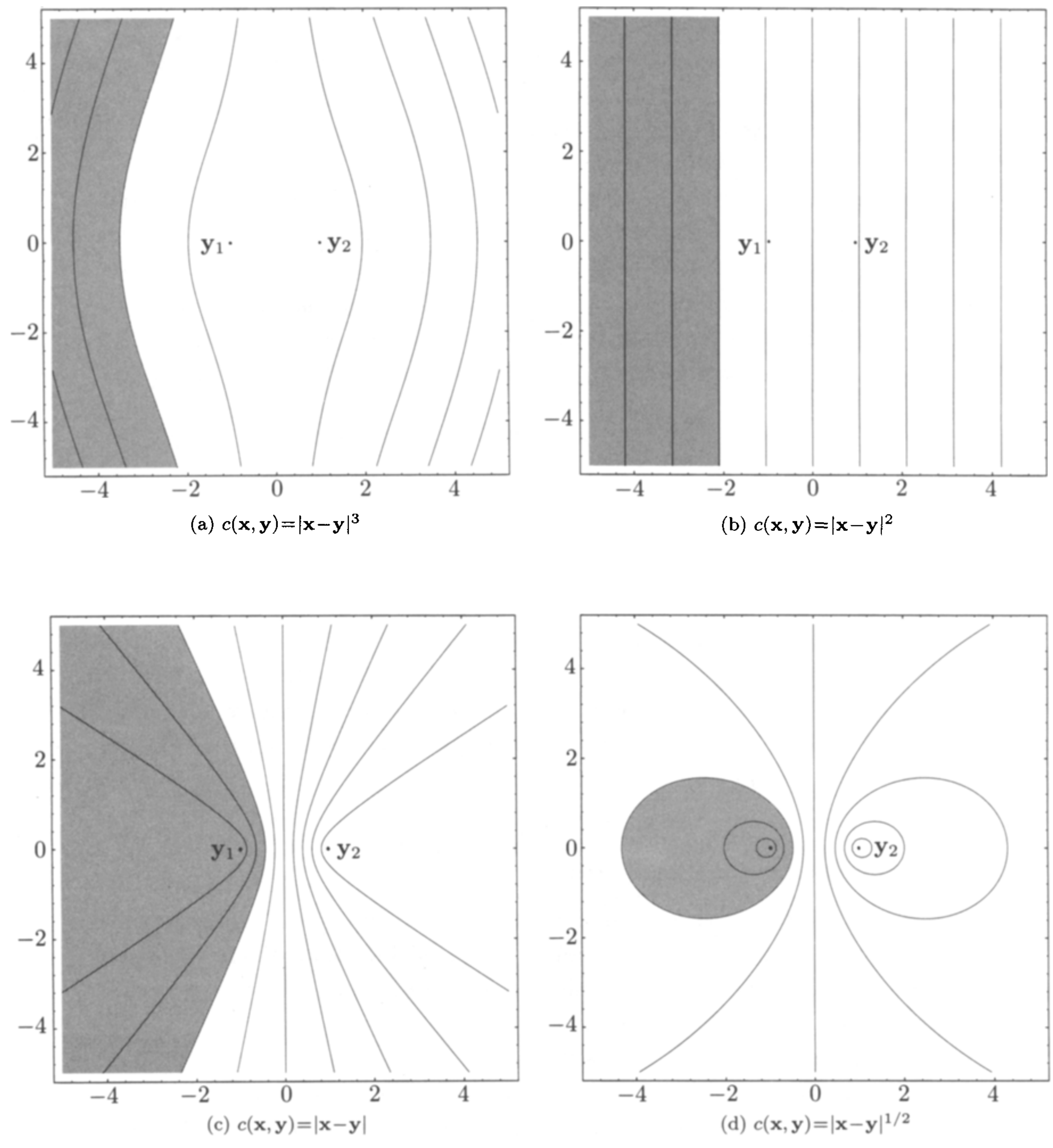
Strict convexity of $h(\mathbf{x})$ combines with (H3) to imply continuous differentiability of the convex function $h^*(\mathbf{y})$ throughout $\mathbf{R}^d$ (see Corollary A.2 of the appendix).

The dual h^* to a concave function $h(\mathbf{x}) = l(|\mathbf{x}|)$ of the distance is defined instead by

$$h^*(\mathbf{y}) := -k^*(-|\mathbf{y}|), \tag{11}$$

where the convex function $k(\lambda) = -l(\lambda)$ is extended to $\lambda < 0$ by setting $k := \infty$, before computing k^* using (10). From Proposition A.6, one has $h^*(\mathbf{y}) = -\infty$ on some ball centered at the origin, but elsewhere $h^*(\mathbf{y})$ is continuously differentiable by strict concavity of $l(\lambda)$.

For either class of cost, when (ν, μ) satisfies the same hypotheses as (μ, ν), then the map $\mathbf{s}(\mathbf{x})$ of our main theorems will be invertible. The inverse map $\mathbf{t} = \mathbf{s}^{-1}$ pushes ν forward to μ; it will be optimal with respect to the cost function $c(\mathbf{y}, \mathbf{x})$. Now, consider measures μ and ν which are absolutely continuous with respect to Lebesgue—$d\mu(\mathbf{x}) = f(\mathbf{x})\, d\mathbf{x}$ and $d\nu(\mathbf{y}) = g(\mathbf{y})\, d\mathbf{y}$. Take each to vanish on the other's support if the



(a) $c(\mathbf{x},\mathbf{y})=|\mathbf{x}-\mathbf{y}|^3$

(b) $c(\mathbf{x},\mathbf{y})=|\mathbf{x}-\mathbf{y}|^2$

(c) $c(\mathbf{x},\mathbf{y})=|\mathbf{x}-\mathbf{y}|$

(d) $c(\mathbf{x},\mathbf{y})=|\mathbf{x}-\mathbf{y}|^{1/2}$

Fig. 2. A few optimal maps to measures which concentrate at two points $\mathbf{y}_2=(1,0)=-\mathbf{y}_1$ in the plane. Shading indicates the region mapped to $\mathbf{y}_1$; its complement is mapped to $\mathbf{y}_2$.

cost is concave. Then the transformation $\mathbf{y}=\mathbf{s}(\mathbf{x})$ represents a change of variables (1) between μ and ν, so—formally at least (neglecting regularity issues)—its Jacobian determinant $\mathrm{Ds}(\mathbf{x})$ satisfies $g(\mathbf{s}(\mathbf{x}))|\det \mathrm{Ds}(\mathbf{x})|=f(\mathbf{x})$. The potential $\psi(\mathbf{x})$ satisfies the partial differential equation

$$g(\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x})))\det[\mathrm{I}-\mathrm{D}^2h^*(\nabla\psi(\mathbf{x}))\mathrm{D}^2\psi(\mathbf{x})] = \pm f(\mathbf{x}). \tag{12}$$

Our main theorems may be interpreted as providing existence and uniqueness results concerning c-concave solutions to (12) in a measure-theoretic (i.e., very weak) sense. The plus sign corresponds to convex costs, and the minus sign to concave functions $h(\mathbf{x})= l(|\mathbf{x}|)$ of the distance, reflecting the local behaviour of the optimal map: orientation-preserving in the former (convex) case and orientation-reversing in the latter case. As Caffarelli pointed out to us, this may be seen by expressing the Jacobian

$$\mathrm{Ds}(\mathbf{x}) = \mathrm{D}^2h^*(\nabla\psi(\mathbf{x}))(\mathrm{D}^2h(\mathbf{x}-\mathbf{s}(\mathbf{x}))-\mathrm{D}^2\psi(\mathbf{x}))$$

as the product of D^2h^* with a non-negative matrix. The second factor is positive semi-definite by the c-concavity(2) of ψ (see Figure 1), while the first factor D^2h^* has either no negative eigenvalues or one negative eigenvalue, depending on the convexity of h and h^*, or their concavity as *increasing* functions of $|\mathbf{x}|$. If $h(\mathbf{x})=\frac{1}{2}|\mathbf{x}|^2$, then $\mathrm{D}^2h^*=\mathrm{I}$ and equation (12) reduces to the Monge–Ampère equation [7]; Caffarelli has developed a regularity theory [9] which justifies the formal discussion in this case. However, the discontinuities in $\nabla\psi$—and points where $\nabla\psi=0$ when the cost is concave—are also of interest: they divide $\operatorname{spt}\mu$ into the regions on which one may hope for smooth transport. A summary of our notation is shown in Table 1.

2. Background on optimal measures

In this section, we review some background material germane to our further developments. Principally, this involves recounting connections between optimal mass transport, c-concave functions and c-cyclically monotone sets established in the work of Rüschendorf [34], [35], and Smith and Knott [41].

To emphasize the generality of the arguments, *this section alone* is formulated not in the Euclidean space $\mathbf{R}^d$, but on a pair of locally compact, σ-compact, second countable Hausdorff spaces $\mathbf{X}$ and $\mathbf{Y}$. The Borel probability measures on $\mathbf{X}$ are denoted by $\mathcal{P}(\mathbf{X})$, while the mass transport problem becomes: Find the measure γ which minimizes the

(2) Which implies that $(\mathbf{x},\mathbf{s}(\mathbf{x}))\in\partial^c\psi$ in Definition 2.6 through Proposition 3.4 (ii) or Proposition 6.1.

Table 1

Notation	Meaning	Definition
$\partial^{\cdot}\psi$, $\partial_{\cdot}\psi$	super- and subdifferentials	Definition 2.5
$\partial^c\psi$	c-superdifferential	Definition 2.6
$\partial^c_\circ\psi$	its off-diagonal restriction	before Lemma 5.2
$\Gamma(\mu,\nu)$	joint measures with marginals μ and ν	before (3)
(H1)–(H5)	hypotheses on convex costs	after Definition 1.1
$h^*(\mathbf{x})$	Legendre transform of the cost	(10)–(11)
$\mathbf{id}$	the identity map	before Theorem 1.2
$\mathcal{M}(\mathbf{R}^d)$	non-negative Borel measures on $\mathbf{R}^d$	before Definition 0.1
$[\mu-\nu]_+$	positive part of $\mu-\nu$	before Definition 2.8
$\mu\wedge\nu$	common mass to μ and ν	before Definition 2.8
$\mathcal{P}(\mathbf{R}^d)$	Borel probability measures on $\mathbf{R}^d$	before Definition 0.1
$\operatorname{spt}\gamma$	(closed) support of the measure γ	before Theorem 1.4
$\mathbf{s}_{\#}\mu$	the push-forward of μ through $\mathbf{s}$	Definition 0.1
$\hat{\mathbf{x}}$	the unit vector $\hat{\mathbf{x}}:=\mathbf{x}/\lvert\mathbf{x}\rvert$	before Theorem A.1

integral of a continuous cost function $c(\mathbf{x},\mathbf{y})\geqslant 0$ on $\mathbf{X}\times\mathbf{Y}$, among the joint measures $\Gamma(\mu,\nu)\subset\mathcal{P}(\mathbf{X}\times\mathbf{Y})$ with $\mu\in\mathcal{P}(\mathbf{X})$ and $\nu\in\mathcal{P}(\mathbf{Y})$ as their marginals. Definitions for the transport cost $\mathcal{C}(\gamma)$, optimal joint measures, push-forward, support, c-concavity and c-transforms must be modified in the obvious way—by replacing each occurrence of $\mathbf{R}^d$ with $\mathbf{X}$ or with $\mathbf{Y}$. Some notions from non-smooth analysis—super- and subdifferentials—are also introduced.

For the record, our discussion begins with the standard continuity-compactness result which assures the existence of an optimal measure γ in $\Gamma(\mu,\nu)$; its well-known proof may be found e.g. in [24, Theorem 2.19]. The section closes with some results on the structure of γ when the cost is a metric on $\mathbf{X}=\mathbf{Y}$.

PROPOSITION 2.1 (existence of an optimal measure [24]). *Fix $c\geqslant 0$ lower semi-continuous on $\mathbf{X}\times\mathbf{Y}$ and measures $\mu\in\mathcal{P}(\mathbf{X})$ and $\nu\in\mathcal{P}(\mathbf{Y})$. There is at least one optimal measure $\gamma\in\mathcal{P}(\mathbf{X}\times\mathbf{Y})$ with marginals μ and ν.*

The optimal measures in $\mathcal{P}(\mathbf{X}\times\mathbf{Y})$ can be characterized [41] through Smith and Knott's notion of *c-cyclical monotonicity,* defined just below for a relation $S\subset\mathbf{X}\times\mathbf{Y}$. The ensuing theory generalizes classical results of convex analysis which pertain to the Euclidean inner product $c(\mathbf{x},\mathbf{y})=\langle\mathbf{x},\mathbf{y}\rangle$ on $\mathbf{X}=\mathbf{Y}=\mathbf{R}^d$; there c-concavity reduces to concavity in the usual sense, while after changing a sign, c-cyclical monotonicity is reduced to the cyclical monotonicity of Rockafellar by the observation that any permutation can be expressed as a product of commuting cycles [32].

Definition 2.2. A subset $S\subset \mathbf{X}\times\mathbf{Y}$ is called *c-cyclically monotone* if for any finite number of points $(\mathbf{x}_j,\mathbf{y}_j)\in S$, $j=1\ldots n$, and permutation σ on n-letters,

$$\sum_{j=1}^{n} c(\mathbf{x}_j,\mathbf{y}_j)\leqslant \sum_{j=1}^{n} c(\mathbf{x}_{\sigma(j)},\mathbf{y}_j). \tag{13}$$

For finite sets S, c-cyclical monotonicity means that the points of origin $\mathbf{x}$ and destinations $\mathbf{y}$ related by $(\mathbf{x},\mathbf{y})\in S$ have been paired so as to minimize the total transportation cost $\sum_S c(\mathbf{x},\mathbf{y})$. This intrepretation motivates the following theorem, first derived by Smith and Knott from the duality-based characterization of Rüschendorf [35]. The proof given here uses a direct argument of Abdellaoui and Heinich [2] instead; it shows that c-cyclically monotone support plays the role of an Euler–Lagrange condition for optimal measures on $\mathbf{X}\times\mathbf{Y}$.

THEOREM 2.3 (optimal measures have c-cyclically monotone support). *Fix a continuous function $c(\mathbf{x},\mathbf{y})\geqslant 0$ on $\mathbf{X}\times\mathbf{Y}$. If the measure $\gamma\in\mathcal{P}(\mathbf{X}\times\mathbf{Y})$ is optimal and $\mathcal{C}(\gamma)<\infty$ then γ has c-cyclically monotone support.*

Proof. Before beginning the proof, a useful perspective from probability theory is recalled: Given a collection of measures $\mu_j\in\mathcal{P}(\mathbf{X})$ $(j=1,\ldots,n)$, there exists a probability space $(\Omega,\mathcal{B},\eta)$ such that each μ_j can be represented as the push-forward of η through a (Borel) map $\pi_j:\Omega\to\mathbf{X}$. The demonstration is easy: let $\eta:=\mu_1\times\ldots\times\mu_n$ be product measure on the Borel subsets of $\Omega:=\mathbf{X}^n$, and take $\pi_j(\mathbf{x}_1,\ldots,\mathbf{x}_n):=\mathbf{x}_j$ to be projection onto the jth copy of $\mathbf{X}$. Also, recall that if $U\subset\mathbf{X}$ is a Borel set of mass $\lambda:=\mu[U]>0$, one can define the *normalized restriction* of μ to U: it is the probability measure assigning mass $\lambda^{-1}\mu[V\cap U]$ to $V\subset\mathbf{X}$.

Now, suppose γ is optimal; i.e., minimizes $\mathcal{C}(\cdot)$ among the measures in $\mathcal{P}(\mathbf{X}\times\mathbf{Y})$ sharing its marginals. Unless γ has c-cyclically monotone support, there is an integer n and permutation σ on n letters such that the function

$$f(\mathbf{x}_1,\ldots,\mathbf{x}_n;\mathbf{y}_1,\ldots,\mathbf{y}_n):=\sum_{j=1}^{n} c(\mathbf{x}_{\sigma(j)},\mathbf{y}_j)-c(\mathbf{x}_j,\mathbf{y}_j)$$

takes a negative value at some points $(\mathbf{x}_1,\mathbf{y}_1),\ldots,(\mathbf{x}_n,\mathbf{y}_n)\in\operatorname{spt}\gamma$. These points can be used to construct a more efficient perturbation of γ as follows. Since f is continuous, there exist (compact) neighbourhoods $U_j\subset\mathbf{X}$ of $\mathbf{x}_j$ and $V_j\subset\mathbf{Y}$ of $\mathbf{y}_j$ such that $f(\mathbf{u}_1,\ldots,\mathbf{u}_n;\mathbf{v}_1,\ldots,\mathbf{v}_n)<0$ whenever $\mathbf{u}_j\in U_j$ and $\mathbf{v}_j\in V_j$. Moreover, $\lambda:=\inf_j\gamma[U_j\times V_j]$

will be positive because $(\mathbf{x}_j,\mathbf{y}_j)\in\operatorname{spt}\gamma$. Let $\gamma_j\in\mathcal{P}(\mathbf{X}\times\mathbf{Y})$ denote the normalized restriction of γ to $U_j\times V_j$. Introducing a factor of n lest the γ_j fail to be disjointly supported, one can subtract $\sum_j \lambda\gamma_j/n$ from γ and be left with a positive measure.

For each j, choose a Borel map $\omega\to(\mathbf{u}_j(\omega),\mathbf{v}_j(\omega))$ from Ω to $\mathbf{X}\times\mathbf{Y}$ such that $\gamma_j=(\mathbf{u}_j\times\mathbf{v}_j)_\#\eta$; this map takes its values in the compact set $U_j\times V_j$. Define the positive measure

$$\gamma':=\gamma+\lambda n^{-1}\sum_{j=1}^{n}(\mathbf{u}_{\sigma(j)}\times\mathbf{v}_j)_\#\eta-(\mathbf{u}_j\times\mathbf{v}_j)_\#\eta.$$

Then $\gamma'\in\mathcal{P}(\mathbf{X}\times\mathbf{Y})$ shares the marginals of γ, while using (1) to compute its cost contradicts the optimality of γ: since the integrand f will be negative,

$$\mathcal{C}(\gamma')-\mathcal{C}(\gamma)=\lambda n^{-1}\int_\Omega\sum_{j=1}^{n}c(\mathbf{u}_{\sigma(j)},\mathbf{v}_j)-c(\mathbf{u}_j,\mathbf{v}_j)\,d\eta<0.$$

Thus γ must have c-cyclically monotone support. □

A more powerful reformulation exploits convexity to show that *all* of the optimal measures in $\Gamma(\mu,\nu)$ have support on the *same* c-cyclically monotone set.

COROLLARY 2.4. *Fix $\mu\in\mathcal{P}(\mathbf{X})$, $\nu\in\mathcal{P}(\mathbf{Y})$ and a continuous function $c(\mathbf{x},\mathbf{y})\geqslant 0$ on $\mathbf{X}\times\mathbf{Y}$. Unless $\mathcal{C}(\cdot)=\infty$ throughout $\Gamma(\mu,\nu)$, there is a c-cyclically monotone set $S\subset\mathbf{X}\times\mathbf{Y}$ containing the supports of all optimal measures γ in $\Gamma(\mu,\nu)$.*

Proof. Let $S:=\bigcup\operatorname{spt}\gamma$, where the union is over the optimal measures γ in $\Gamma(\mu,\nu)$. We shall show S to be c-cyclically monotone by verifying (13). Therefore, choose any finite number of points $(\mathbf{x}_j,\mathbf{y}_j)\in S$ indexed by $j=1,...,n$ and a permutation σ on n letters. For each j, the definition of S guarantees an optimal measure $\gamma_j\in\Gamma(\mu,\nu)$ with $(\mathbf{x}_j,\mathbf{y}_j)\in\operatorname{spt}\gamma_j$. Define the convex combination $\gamma:=(1/n)\sum_j\gamma_j$. Since $\Gamma(\mu,\nu)$ is a convex set and $\mathcal{C}(\cdot)$ is a linear functional, $\gamma\in\Gamma(\mu,\nu)$ and $\mathcal{C}(\gamma)=\mathcal{C}(\gamma_j)$; thus γ is also optimal. By Theorem 2.3, $\operatorname{spt}\gamma$ is c-cyclically monotone. But $\operatorname{spt}\gamma$ contains $\operatorname{spt}\gamma_j$ for each j, and in particular the points $(\mathbf{x}_j,\mathbf{y}_j)$, so (13) is implied. □

Rockafellar's main result in [32] exposed the connection between gradients of concave functions and cyclically monotone sets: it showed that a concave potential could be constructed from any cyclically monotone set. Smith and Knott observed that this relationship extends to c-concave functions and c-cyclically monotone sets. To state the theorem precisely requires some generalized notions of gradient, which continue to be useful throughout:

Definition 2.5. A function $\psi: \mathbf{R}^d \to \mathbf{R}\cup\{\pm\infty\}$ is *superdifferentiable* at $\mathbf{x}\in\mathbf{R}^d$ if $\psi(\mathbf{x})$ is finite and there exists $\mathbf{y}\in\mathbf{R}^d$ such that

$$\psi(\mathbf{x}+\mathbf{v}) \leqslant \psi(\mathbf{x}) + \langle \mathbf{v}, \mathbf{y}\rangle + o(|\mathbf{v}|) \tag{14}$$

for small $\mathbf{v}\in\mathbf{R}^d$; here $o(\lambda)/\lambda$ must tend to zero with λ.

A pair $(\mathbf{x}, \mathbf{y})$ belongs to the *superdifferential* $\partial^{\cdot}\psi \subset \mathbf{R}^d \times \mathbf{R}^d$ of ψ if $\psi(\mathbf{x})$ is finite and (14) holds, in which case $\mathbf{y}$ is called a *supergradient* of ψ at $\mathbf{x}$; such supergradients $\mathbf{y}$ comprise the set $\partial^{\cdot}\psi(\mathbf{x})\subset\mathbf{R}^d$, while for $V\subset\mathbf{R}^d$ we define $\partial^{\cdot}\psi(V):=\bigcup_{\mathbf{x}\in V}\partial^{\cdot}\psi(\mathbf{x})$. The analogous notions of subdifferentiability, subgradients and the subdifferential $\partial_{\cdot}\psi$ are defined by reversing inequality (14). It is not hard to see that a real-valued function will be differentiable at $\mathbf{x}$ precisely when it is both super- and subdifferentiable there; then $\partial^{\cdot}\psi(\mathbf{x}) = \partial_{\cdot}\psi(\mathbf{x}) = \{\nabla\psi(\mathbf{x})\}$.

A function $\psi: \mathbf{R}^d \to \mathbf{R}\cup\{-\infty\}$ is said to be *concave* if $\lambda\in(0,1)$ implies that

$$\psi((1-\lambda)\mathbf{x}+\lambda\mathbf{y}) \geqslant (1-\lambda)\psi(\mathbf{x}) + \lambda\psi(\mathbf{y})$$

whenever the latter is finite. The function $\psi:=-\infty$ is excluded by convention. For concave functions the error term will vanish in (14): the inequality $\psi(\mathbf{x}+\mathbf{v})\leqslant\psi(\mathbf{x})+\langle\mathbf{v},\mathbf{y}\rangle$ holds for all $(\mathbf{x},\mathbf{y})\in\partial^{\cdot}\psi$, and the supergradients of ψ parameterize supporting hyperplanes of graph(ψ) at $(\mathbf{x}, \psi(\mathbf{x}))$. To provide a notion analogous to supporting hyperplanes in the context of c-concave functions, a c-superdifferential is defined in the following way [35] (cf. Figure 1):

Definition 2.6. The *c-superdifferential* $\partial^c\psi$ of $\psi: \mathbf{X}\to\mathbf{R}\cup\{-\infty\}$ consists of the pairs $(\mathbf{x},\mathbf{y})\in\mathbf{X}\times\mathbf{Y}$ for which $\psi(\mathbf{v})\leqslant\psi(\mathbf{x})+c(\mathbf{v},\mathbf{y})-c(\mathbf{x},\mathbf{y})$ if $\mathbf{v}\in\mathbf{X}$.

Alternately, $(\mathbf{x},\mathbf{y})\in\partial^c\psi$ means that $c(\mathbf{v},\mathbf{y})-\psi(\mathbf{v})$ assumes its minimum at $\mathbf{v}=\mathbf{x}$. We define $\partial^c\psi(\mathbf{x})\subset\mathbf{Y}$ to consist of those $\mathbf{y}$ for which $(\mathbf{x},\mathbf{y})\in\partial^c\psi$, while $\partial^c\psi(V):=\bigcup_{\mathbf{x}\in V}\partial^c\psi(\mathbf{x})$ for $V\subset\mathbf{X}$.

In our applications $c(\mathbf{x},\mathbf{y})$ is continuous, so a c-concave potential ψ will be upper semi-continuous from its definition. As a consequence, $\partial^c\psi$ will be a closed subset of $\mathbf{X}\times\mathbf{Y}$—an observation which will be useful later. With this notation, Smith and Knott's generalization [41] of the Rockafellar theorem [32] can be stated. Its proof is drawn from [37].

THEOREM 2.7 (c-concave potentials from c-cyclically monotone sets). *For $S\subset\mathbf{X}\times\mathbf{Y}$ to be c-cyclically monotone, it is necessary and sufficient that $S\subset\partial^c\psi$ for some c-concave $\psi: \mathbf{X}\to\mathbf{R}\cup\{-\infty\}$.*

Proof. Sufficiency is easy: c-concavity of ψ implies that $\partial^c\psi$ is c-cyclically monotone. To see this, choose n points $(\mathbf{x}_j, \mathbf{y}_j)$ from $\partial^c\psi$ and a permutation σ on n-letters. We invoke c-concavity only to know that $\psi: \mathbf{X} \to \mathbf{R} \cup \{-\infty\}$ is finite at some $\mathbf{p} \in \mathbf{X}$; then taking $(\mathbf{x}, \mathbf{y}) := (\mathbf{x}_j, \mathbf{y}_j)$ and $\mathbf{v} := \mathbf{p}$ in Definition 2.6 implies that $\psi(\mathbf{x}_j)$ is finite, while taking $\mathbf{v} := \mathbf{x}_{\sigma(j)}$ implies $\psi(\mathbf{x}_{\sigma(j)}) - \psi(\mathbf{x}_j) \leqslant c(\mathbf{x}_{\sigma(j)}, \mathbf{y}_j) - c(\mathbf{x}_j, \mathbf{y}_j)$. Summing this inequality over $j = 1, ..., n$ yields (13), whence $\partial^c \psi$ is c-cyclically monotone.

To prove necessity, one needs to construct a suitable potential ψ from a c-cyclically monotone set $S \subset \Omega_1 \times \Omega_2 \subset \mathbf{X} \times \mathbf{Y}$. Since (13) holds true for the cycle $\sigma = (12 \ldots n)$ in particular, the construction of [37, Lemma 2.1] yields a c-concave ψ on $\Omega_1 := \mathbf{X}$ with $S \subset \partial^c \psi$. Taking $\Omega_2 = \pi'(S)$ with $\pi'(x, y) = y$ when applying this lemma forces ψ to be the c-transform of a function on Ω_2. □

We record this last observation as a corollary to the proof:

COROLLARY 2.8. *Let $S \subset \mathbf{X} \times \mathbf{Y}$ be c-cyclically monotone, and let $\pi'(S)$ denote the projection of S onto $\mathbf{Y}$ through the map $\pi'(\mathbf{x}, \mathbf{y}) := \mathbf{y}$. Then $S \subset \partial^c \psi$ for the c-transform $\psi: \mathbf{X} \to \mathbf{R} \cup \{-\infty\}$ of a function on $\pi'(S)$.*

Combining Theorems 2.3 and 2.7 makes it clear that if a measure γ solves the Kantorovich problem on $\Gamma(\mu, \nu)$ it will necessarily be supported in the c-supergradient of a c-concave potential ψ. Indeed, this fact was already appreciated by Rüschendorf, who recognized that its converse (sufficiency) also holds true [35]. Our main conclusions will be recovered from an analysis of ψ and $\partial^c \psi$ when $\mathbf{X} = \mathbf{Y} = \mathbf{R}^d$. Before embarking on that analysis, we conclude this review by casting into the present framework a few variants on well-known results which apply when $c(\mathbf{x}, \mathbf{y})$ is a metric and $\mathbf{X} = \mathbf{Y}$. We assume that $c(\mathbf{x}, \mathbf{y})$ satisfies the triangle inequality *strictly*:

$$c(\mathbf{x}, \mathbf{y}) < c(\mathbf{x}, \mathbf{p}) + c(\mathbf{p}, \mathbf{y}) \tag{15}$$

unless $\mathbf{p} = \mathbf{x}$ or $\mathbf{p} = \mathbf{y}$. In this case, any mass which is common to μ and ν will stay in its place, and can be subtracted from the diagonal of any optimal measure γ.

PROPOSITION 2.9 (any mass stays in place if it can). *Let $\mu, \nu \in \mathcal{P}(\mathbf{X})$ and denote their shared mass by $\mu \wedge \nu := \mu - [\mu - \nu]_+$. The restriction γ_d of any joint measure $\gamma \in \Gamma(\mu, \nu)$ to the diagonal $D := \{(\mathbf{x}, \mathbf{x}) \mid \mathbf{x} \in \mathbf{X}\}$ satisfies*

$$\gamma_d \leqslant (\mathbf{id} \times \mathbf{id})_{\#}(\mu \wedge \nu). \tag{16}$$

When $c(\mathbf{x}, \mathbf{y})$ is a metric on $\mathbf{X}$ satisfying the strict triangle inequality (15) *and γ has c-cyclically monotone support, then* (16) *becomes an equality.*

Proof. Let $\pi(\mathbf{x}, \mathbf{y}) := \mathbf{x}$ and $\pi'(\mathbf{x}, \mathbf{y}) := \mathbf{y}$ denote projections from $\mathbf{X} \times \mathbf{X}$ to $\mathbf{X}$, and decompose $\gamma = \gamma_d + \gamma_o$ into its diagonal and off-diagonal parts, so that γ_d is supported on D and coincides with γ there. From $\operatorname{spt}\gamma_d\subset D$ it is easily verified that the marginals of γ_d coincide: denote them by $\beta:=\pi_\#\gamma_d=\pi'_\#\gamma_d$. Moreover $\gamma_d=(\mathbf{id}\times\mathbf{id})_\#\beta$. Defining $\mu_o:=\pi_\#\gamma_o$ and $\nu_o:=\pi'_\#\gamma_o$, linearity $\pi_\#\gamma=\pi_\#\gamma_o+\pi_\#\gamma_d$ makes it clear that $\mu=\mu_o+\beta$ and $\nu=\nu_o+\beta$. These measures are all non-negative, so $\beta\leqslant\mu\wedge\nu$ is established and implies (16). Assume therefore that γ has c-cyclically monotone support. It remains to show only that μ_o and ν_o are mutually singular measures, so that $\mu_o-\nu_o$ gives the Jordan decomposition of $\mu-\nu$. Then $\mu\wedge\nu:=\mu-\mu_o=\beta$ and (16) becomes an equality.

To prove that μ_o and ν_o are mutually singular requires a set U of full measure for μ_o with zero measure for ν_o. Define $S=\operatorname{spt}\gamma\setminus D$ and take $U:=\pi(S)$; both sets are σ-compact, hence Borel. Since S is a set of full measure for γ_o, U has full measure for $\mu_o=\pi_\#\gamma_o$. Similarly, $V=\pi'(S)$ has full measure for ν_o. We argue by contradiction that U and V are disjoint, thereby establishing the proposition. Suppose $\mathbf{p}\in U\cap V$, meaning that there exist $\mathbf{x},\mathbf{y}\in\mathbf{X}$, both different from $\mathbf{p}$, such that $(\mathbf{x},\mathbf{p})$ and $(\mathbf{p},\mathbf{y})$ lie in $\operatorname{spt}\gamma$. Applying the two-point inequality ($n=2$) for c-cyclically monotonicity (13) to $\operatorname{spt}\gamma$ yields

$$c(\mathbf{x},\mathbf{p})+c(\mathbf{p},\mathbf{y})\leqslant c(\mathbf{x},\mathbf{y})+c(\mathbf{p},\mathbf{p}).$$

Since $c(\mathbf{p},\mathbf{p})=0$, the strict triangle inequality (15) is violated. The only conclusion is that U and V are disjoint and the proof is complete. □

COROLLARY 2.10 (metric costs with fixed-penalty for transport). *Fix a continuous metric $c(\mathbf{x},\mathbf{y})$ on $\mathbf{X}$ satisfying the triangle inequality strictly, and define a discontinuous cost by $\tilde c(\mathbf{x},\mathbf{y}):=c(\mathbf{x},\mathbf{y})$ for $\mathbf{x}\neq\mathbf{y}$ and $c(\mathbf{x},\mathbf{x})=-\lambda<0$. A joint measure $\gamma\in\mathcal{P}(\mathbf{X}\times\mathbf{X})$ is optimal for $\tilde c$ if and only if it is optimal for c.*

Proof. Follows easily from Theorems 2.3, 2.9 and $\tilde{\mathcal{C}}(\gamma)=\mathcal{C}(\gamma)-\lambda\gamma[D]$. □

As the last proposition suggests, when $c(\mathbf{x},\mathbf{y})$ is a metric the diagonal $D\subset\mathbf{X}\times\mathbf{X}$ plays a distinguished role among c-cyclically monotone sets. A final lemma shows that D is contained in the c-superdifferential of every c-concave function ψ. Equivalently, D can be added to any c-cyclically monotone set without spoiling the c-cyclical monotonicity. Finiteness of ψ is a useful corollary, while Kantorovich's observation that ψ will be Lipschitz continuous relative to the metric c is also deduced; cf. [21].

LEMMA 2.11 (c-concavity and the diagonal for metrics). *Let $c(\mathbf{x},\mathbf{y})$ be a metric on $\mathbf{X}$ and $\psi:\mathbf{X}\to\mathbf{R}\cup\{-\infty\}$ be c-concave. Then*

(i) *ψ is real-valued and Lipschitz with constant 1 relative to $c(\mathbf{x},\mathbf{y})$;*

(ii) *for every $\mathbf{p}\in\mathbf{X}$ one has $(\mathbf{p},\mathbf{p})\in\partial^c\psi$.*

Proof. (ii) Let $\mathbf{x},\mathbf{y},\mathbf{p}\in\mathbf{X}$ and $\lambda\in\mathbf{R}$. The triangle inequality implies that

$$c(\mathbf{x},\mathbf{y})+\lambda\leqslant c(\mathbf{x},\mathbf{p})+c(\mathbf{p},\mathbf{y})+\lambda. \tag{17}$$

Recalling the definition (7) of c-concavity, an infimum of (17) over $(\mathbf{y},\lambda)\in\mathcal{A}$ yields

$$\psi(\mathbf{x})\leqslant c(\mathbf{x},\mathbf{p})+\psi(\mathbf{p}). \tag{18}$$

Since $c(\mathbf{p},\mathbf{p})=0$ and $\mathbf{p}$ was arbitrary, $(\mathbf{p},\mathbf{p})\in\partial^c\psi$ by Definition 2.6.

(i) Since ψ is c-concave, it takes a finite value $\psi(\mathbf{x})>-\infty$ somewhere by assumption. For any $\mathbf{p}\in\mathbf{X}$ the preceding argument yields one direction (18) of the Lipschitz bound and also implies $\psi(\mathbf{p})>-\infty$. The latter observation shows that $\mathbf{x}\in\mathbf{X}$ was arbitrary, so the argument is symmetrical under interchange of $\mathbf{x}$ with $\mathbf{p}$. Thus (18) also yields $\psi(\mathbf{p})-\psi(\mathbf{x})\leqslant c(\mathbf{p},\mathbf{x})$. Since $c(\mathbf{p},\mathbf{x})=c(\mathbf{x},\mathbf{p})$ the claim $|\psi(\mathbf{x})-\psi(\mathbf{p})|\leqslant c(\mathbf{x},\mathbf{p})$ is established. □

Part I. Strictly convex costs

3. Existence and uniqueness of optimal maps

The goal of this section is to prove the existence of a solution $\mathbf{s}$ to the Monge problem for convex costs $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$. That is, given two measures μ and ν on $\mathbf{R}^d$ with the same total mass, one seeks to show that the infimum (2) is attained by some measure-preserving map $\mathbf{s}$ between μ and ν. When $h(\mathbf{x})$ is strictly convex and satisfies (H1)–(H3), this will indeed be the case provided that μ is absolutely continuous with respect to Lebesgue. Uniqueness of this solution to both the Monge and Kantorovich problems follows as a corollary to the proof. For smooth costs it is enough that no mass of μ concentrate on sets of dimension $d-1$, but this observation is relegated to Remark 4.7 for simplicity. The starting point of our analysis will be the potential function ψ of Theorem 2.7, or rather its c-superdifferential $\partial^c\psi$. Our key observation is that apart from a set of measure zero, $\partial^c\psi$—and indeed any c-cyclically monotone relation $S\subset\mathbf{R}^d\times\mathbf{R}^d$—must lie in the graph of a function $\mathbf{x}\to\mathbf{s}(\mathbf{x})$ on $\mathbf{R}^d$. This function is the optimal map.

The first lemma is basic. Illustrated by Figure 1, it asserts a matching condition between the gradients of the cost and potential whenever $(\mathbf{x},\mathbf{y})\in\partial^c\psi$, cf. [35, (73)], and indicates why injectivity of ∇h determines $\mathbf{y}$ as a function of $\mathbf{x}$. The lemma is formulated for general costs of the form $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$.

LEMMA 3.1 (relating c-differentials to subdifferentials). *Let $h:\mathbf{R}^d\to\mathbf{R}$ and $\psi:\mathbf{R}^d\to\mathbf{R}\cup\{-\infty\}$. If $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ then $(\mathbf{x},\mathbf{y})\in\partial^c\psi$ implies $\partial.\psi(\mathbf{x})\subset\partial.h(\mathbf{x}-\mathbf{y})$; when h and ψ are differentiable, $\nabla\psi(\mathbf{x})=\nabla h(\mathbf{x}-\mathbf{y})$.*

Proof. Let $(\mathbf{x},\mathbf{y})\in\partial^c\psi$. Assume $\psi(\mathbf{x})>-\infty$, since otherwise $\partial.\psi(\mathbf{x})$ is empty and there is nothing to prove. If $\mathbf{z}\in\partial_\cdot\psi(\mathbf{x})$, then sub- and c-superdifferentiability of ψ yield

$$\begin{aligned}\psi(\mathbf{x})+\langle\mathbf{v},\mathbf{z}\rangle+o(|\mathbf{v}|) &\leqslant \psi(\mathbf{x}+\mathbf{v})\\ &\leqslant \psi(\mathbf{x})+h(\mathbf{x}+\mathbf{v}-\mathbf{y})-h(\mathbf{x}-\mathbf{y})\end{aligned}$$

for small $\mathbf{v}\in\mathbf{R}^d$. In other words, $\mathbf{z}\in\partial_\cdot h(\mathbf{x}-\mathbf{y})$, and so the first claim is proved. Differentiability implies the second claim because then $\partial_\cdot\psi(\mathbf{x})=\{\nabla\psi(\mathbf{x})\}$ while $\partial_\cdot h(\mathbf{x}-\mathbf{y})=\{\nabla h(\mathbf{x}-\mathbf{y})\}$. □

In view of this lemma, the business at hand is to prove some differentiability result for the potential ψ. Strict convexity of $h(\mathbf{x})$ ensures the invertibility of ∇h. The next theorem—proved in Appendix C—asserts that a c-concave potential ψ is locally Lipschitz. If the cost is a derivative smoother, then ψ satisfies a local property known as *semi-concavity*; introduced by Douglis [17] to select unique solutions for the Hamilton–Jacobi equation, it implies all the smoothness enjoyed by concave functions.

Definition 3.2. A function $\psi:\mathbf{R}^d\to\mathbf{R}\cup\{-\infty\}$ is said to be *locally semi-concave* at $\mathbf{p}\in\mathbf{R}^d$ if there is a constant $\lambda<\infty$ which makes $\psi(\mathbf{x})-\lambda\mathbf{x}^2$ concave on some (small) open ball centered at $\mathbf{p}$.

THEOREM 3.3 (regularity of c-concave potentials). *Let $\psi:\mathbf{R}^d\to\mathbf{R}\cup\{-\infty\}$ be c-concave for some convex $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ with h satisfying* (H2)–(H4). *Then there is a convex set $K\subset\mathbf{R}^d$ with interior $\Omega:=\operatorname{int}K$ such that $\Omega\subset\{\mathbf{x}\mid\psi(\mathbf{x})>-\infty\}\subset K$. Moreover, ψ is locally Lipschitz on Ω, and if $h\in C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$ then ψ will be locally semi-concave on Ω.*

Proof. Proposition C.3 yields the convex set K with interior Ω such that $\Omega\subset\{\psi>-\infty\}\subset K$. Moreover, ψ is locally bounded on Ω. Thus ψ is locally Lipschitz on Ω by Corollary C.5, and locally semi-concave if $h\in C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$. □

We use convexity of K only to know that outside of Ω, the set where ψ is finite has zero volume (indeed, is contained in a Lipschitz submanifold of dimension $d-1$). Inside Ω, Rademacher's theorem shows that the gradient $\nabla\psi$ is defined almost everywhere. When ψ is locally semi-concave, results of Zajíček [43] (or Alberti [3]) imply that the subset of Ω where differentiability fails is rectifiable of dimension $d-1$.

The next lemma and its corollary verify that any c-cyclically monotone set will lie in the graph of a map. The facts we exploit concerning the Legendre transform $h^*(\mathbf{y})$ of a convex cost (10) are summarized in Appendix A.

PROPOSITION 3.4 (c-superdifferentials lie in the graph of a map). *Fix $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ satisfying* (H1)–(H3) *and a c-concave ψ on $\mathbf{R}^d$. Let $\operatorname{dom}\psi$ and $\operatorname{dom}\nabla\psi$ denote the respective sets in $\mathbf{R}^d$ on which ψ is finite, and differentiable. Then*

(i) $\mathbf{s}(\mathbf{x}):=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ *defines a Borel map from* $\mathrm{dom}\,\nabla\psi$ *to* $\mathbf{R}^d$;

(ii) $\partial^c\psi(\mathbf{x})=\{\mathbf{s}(\mathbf{x})\}$ *whenever* $\mathbf{x}\in\mathrm{dom}\,\nabla\psi$;

(iii) $\partial^c\psi(\mathbf{x})$ *is empty unless* $\mathbf{x}\in\mathrm{dom}\,\psi$;

(iv) *the set* $\mathrm{dom}\,\psi\backslash\mathrm{dom}\,\nabla\psi$ *has Lebesgue measure zero.*

Proof. (i) Theorem 3.3 shows that ψ is continuous on the interior Ω of $\mathrm{dom}\,\psi$. Since its gradient is obtained as the pointwise limit of a sequence of continuous approximants (finite differences), $\nabla\psi$ is Borel measurable on the (Borel) subset $\mathrm{dom}\,\nabla\psi\subset\Omega$ where it can be defined. Since ∇h^* is continuous by Corollary A.2, the measurability of $\mathbf{s}(\mathbf{x})$ is established.

(ii) Since ψ is differentiable at $\mathbf{x}\in\mathrm{dom}\,\nabla\psi$ it is bounded nearby, so from Proposition C.4 we conclude that $\partial^c\psi(\mathbf{x})$ is non-empty. Choosing $\mathbf{y}\in\partial^c\psi(\mathbf{x})$, Lemma 3.1 yields $\nabla\psi(\mathbf{x})\in\partial_\cdot h(\mathbf{x}-\mathbf{y})$. Corollary A.2 then shows that $\mathbf{x}-\mathbf{y}=\nabla h^*(\nabla\psi(\mathbf{x}))$—or equivalently $\mathbf{y}=\mathbf{s}(\mathbf{x})$—and establishes $\partial^c\psi(\mathbf{x})=\{\mathbf{s}(\mathbf{x})\}$.

(iii) Part of the definition for c-concavity of $\psi:\mathbf{R}^n\to\mathbf{R}\cup\{-\infty\}$ asserts finiteness of $\psi(\mathbf{v})$ for some $\mathbf{v}\in\mathbf{R}^d$. Since $(\mathbf{x},\mathbf{y})\in\partial^c\psi$ implies $\psi(\mathbf{v})\leqslant\psi(\mathbf{x})+c(\mathbf{v},\mathbf{y})-c(\mathbf{x},\mathbf{y})$, one has $\mathbf{x}\in\mathrm{dom}\,\psi$ whenever $\partial^c\psi(\mathbf{x})$ is non-empty.

(iv) Theorem 3.3 shows ψ to be locally Lipschitz on the interior of $\mathrm{dom}\,\psi$, while the boundary of $\mathrm{dom}\,\psi$ lies in the boundary of a convex subset of $\mathbf{R}^d$ and hence has Lebesgue measure zero. In the interior, Rademacher's theorem yields ψ differentiable almost everywhere, whence $\mathrm{dom}\,\psi\backslash\mathrm{dom}\,\nabla\psi$ has Lebesgue measure zero. □

COROLLARY 3.5 (c-cyclically monotone sets lie in the graph of a map). *Let* $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ *satisfy* (H1)–(H3) *and* $S\subset\mathbf{R}^d\times\mathbf{R}^d$ *be c-cyclically monotone. Then there is a (Borel) set* $N\subset\mathbf{R}^d$ *of zero measure for which* $(\mathbf{x},\mathbf{y})$ *and* $(\mathbf{x},\mathbf{z})$ *in* S *with* $\mathbf{y}\neq\mathbf{z}$ *implies* $\mathbf{x}\in N$.

Proof. Let $S\subset\mathbf{R}^d\times\mathbf{R}^d$ be c-cyclically monotone. Theorem 2.7, due to Smith and Knott, asserts the existence of a c-concave function ψ with $S\subset\partial^c\psi$. Proposition 3.4 provides a Borel set $N:=\mathrm{dom}\,\psi\backslash\mathrm{dom}\,\nabla\psi$ of zero measure for which $(\mathbf{x},\mathbf{y})$ and $(\mathbf{x},\mathbf{z})$ in $S\subset\partial^c\psi$ but $\mathbf{x}\notin N$ imply $\mathbf{y}=\mathbf{z}=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$. □

Armed with this knowledge, we are ready to derive a measure-preserving map from existence of the corresponding potential. The argument generalizes [29, Proposition 10].

PROPOSITION 3.6 (measure-preserving maps from c-concave potentials). *Let* $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ *satisfy* (H1)–(H3), *and suppose that a joint measure* $\gamma\in\mathcal{P}(\mathbf{R}^d\times\mathbf{R}^d)$ *has support in* $\partial^c\psi$ *for some c-concave function* ψ *on* $\mathbf{R}^d$. *Let* μ *and* ν *denote the marginals of* $\gamma\in\Gamma(\mu,\nu)$. *If* ψ *is differentiable* μ*-almost everywhere, then* $\mathbf{s}(\mathbf{x}):=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ *pushes* μ *forward to* ν. *In fact,* $\gamma=(\mathbf{id}\times\mathbf{s})_\#\mu$.

Proof. To begin, we observe from Proposition 3.4 (i) that $\mathbf{s}(\mathbf{x})$ is a Borel map defined μ-almost everywhere: the (Borel) set $\operatorname{dom}\nabla\psi$ where ψ is differentiable carries the full mass of μ by hypothesis. It remains to check $(\mathbf{id}\times\mathbf{s})_{\#}\mu=\gamma$, from which $\mathbf{s}_{\#}\mu=\nu$ follows immediately.

To complete the proof, it suffices to show that the measure $(\mathbf{id}\times\mathbf{s})_{\#}\mu$ coincides with γ on products $U\times V$ of Borel sets $U, V\subset\mathbf{R}^d$; the semi-algebra of such products generates all Borel sets in $\mathbf{R}^d\times\mathbf{R}^d$. Therefore, define $S:=\{(\mathbf{x},\mathbf{y})\in\partial^c\psi \mid \mathbf{x}\in\operatorname{dom}\nabla\psi\}$. For $(\mathbf{x},\mathbf{y})\in S$, Proposition 3.4 (ii) implies $\mathbf{y}=\mathbf{s}(\mathbf{x})$, so

$$(U\times V)\cap S=((U\cap\mathbf{s}^{-1}(V))\times\mathbf{R}^d)\cap S. \tag{19}$$

Being the intersection of two sets having full measure for γ—the closed set $\partial^c\psi$ and the Borel set $\operatorname{dom}\nabla\psi\times\mathbf{R}^d$—the set S is Borel with full measure. Thus $\gamma[Z\cap S]=\gamma[Z]$ for $Z\subset\mathbf{R}^d\times\mathbf{R}^d$. Applied to (19), this yields

$$\gamma[U\times V]=\gamma[(U\cap\mathbf{s}^{-1}(V))\times\mathbf{R}^d]=\mu[U\cap\mathbf{s}^{-1}(V)]=(\mathbf{id}\times\mathbf{s})_{\#}\mu[U\times V].$$

$\gamma\in\Gamma(\mu,\nu)$ implies the second equation; Definition 0.1 implies the third. □

These two propositions combine with results of §2 to yield the existence and uniqueness of optimal solutions to the Monge and Kantorovich problems with strictly convex cost:

THEOREM 3.7 (existence and uniqueness of optimal maps). *Fix a cost* $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$, *where h strictly convex satisfies* (H1)–(H3), *and two Borel probability measures μ and ν on $\mathbf{R}^d$. If μ is absolutely continuous with respect to Lebesgue and* (4) *is finite, then there is a unique optimal measure γ in $\Gamma(\mu,\nu)$. The optimal $\gamma=(\mathbf{id}\times\mathbf{s})_{\#}\mu$ is given by a map $\mathbf{s}(\mathbf{x})=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ pushing μ forward to ν, through a c-concave potential ψ on $\mathbf{R}^d$.*

Proof. Our Corollary 2.4 to Smith and Knott's theorem yields a c-cyclically monotone set $S\subset\mathbf{R}^d\times\mathbf{R}^d$ which contains the supports of all optimal measures in $\Gamma(\mu,\nu)$. A c-concave function ψ on $\mathbf{R}^d$ with $S\subset\partial^c\psi$ is provided by Smith and Knott's next observation—Theorem 2.7. Now suppose that $\gamma\in\Gamma(\mu,\nu)$ is optimal; there is at least one such measure by Proposition 2.1. Then $\operatorname{spt}\gamma\subset\partial^c\psi$. The map $\pi(\mathbf{x},\mathbf{y})=\mathbf{x}$ on $\mathbf{R}^d\times\mathbf{R}^d$ pushes γ forward to $\mu=\pi_{\#}\gamma$, while projecting the closed set $\partial^c\psi$ to a σ-compact set of full measure for μ. Proposition 3.4 (iii)–(iv) shows that $\pi(\partial^c\psi)\subset\operatorname{dom}\psi$ and combines with absolute continuity of μ to ensure that ψ is differentiable μ-almost everywhere. Proposition 3.6 then shows that $\mathbf{s}(\mathbf{x})$ pushes μ forward to ν while γ coincides with the measure $(\mathbf{id}\times\mathbf{s})_{\#}\mu$. This measure is completely determined by μ and ψ, so it must be the only optimal measure in $\Gamma(\mu,\nu)$. □

4. Characterization of the optimal map

For cost functions $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ with h strictly convex, the last section showed that when $\gamma\in\Gamma(\mu,\nu)$ is optimal—or indeed if γ has c-cyclically monotone support—then it is determined by a map $\mathbf{s}(\mathbf{x})=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ which solves the Monge problem. The potential ψ will be c-concave and $\gamma=(\mathbf{id}\times\mathbf{s})_{\#}\mu$. The results of the present section show that only one measure in $\Gamma(\mu,\nu)$ has c-cyclically monotone support, while only one mapping $\mathbf{s}=\mathbf{id}-\nabla h^*\circ\nabla\psi$ can push μ forward to ν and also have ψ c-concave: this geometry is characteristic of γ. As in [27], the argument avoids integrability issues by relying on geometric ideas which can be traced further back to Aleksandrov's uniqueness proof for convex surfaces with prescribed Gaussian curvature [4]. The same assumptions are required that lead to existence of $\mathbf{s}$: the left marginal of γ must vanish on sets of measure zero or dimension $d-1$, depending on the smoothness of h.

The idea of the proof is that if another map $\mathbf{t}=\mathbf{id}-\nabla h^*\circ\nabla\phi$ is induced by some c-concave ϕ, then unless $\mathbf{s}=\mathbf{t}$ holds μ-almost everywere, a set $\partial^c\phi(U)$ can be constructed to have less mass for $\mathbf{s}_{\#}\mu$ than for $\mathbf{t}_{\#}\nu$. We begin with two lemmas concerning c-superdifferentials and c-concave functions which generalize Aleksandrov's observations about supporting hyperplanes for concave functions. The idea is to start by supposing that $c(\mathbf{x},\mathbf{y})+\lambda$ dominates the function $\phi(\mathbf{x})$ but fails to dominate $\psi(\mathbf{x})$, and then increase λ until $c(\mathbf{x},\mathbf{y})+\lambda$ is tangent to $\psi(\mathbf{x})$. At the point of tangency, it is obvious that ψ dominates ϕ. In what follows, $\partial^c\psi^{-1}(V)$ denotes the set of $\mathbf{x}\in\mathbf{R}^d$ for which $\partial^c\psi(\mathbf{x})$ intersects $V\subset\mathbf{R}^d$.

LEMMA 4.1. *Let $c:\mathbf{R}^d\times\mathbf{R}^d\to\mathbf{R}$, and suppose that ϕ and ψ map $\mathbf{R}^d$ into $\mathbf{R}\cup\{-\infty\}$. Define $U:=\{\mathbf{x}\mid\psi(\mathbf{x})>\phi(\mathbf{x})\}$ and $X:=\partial^c\psi^{-1}(\partial^c\phi(U))$. Then $X\subset U$.*

Proof. Let $\mathbf{x}\in X$. Then there is a $\mathbf{u}\in U$ with $\mathbf{y}\in\partial^c\phi(\mathbf{u})$ such that $(\mathbf{x},\mathbf{y})\in\partial^c\psi$. For all $\mathbf{v}\in\mathbf{R}^d$, the definition of $\partial^c\phi$ and $\partial^c\psi$ yields

$$\phi(\mathbf{v})\leqslant\phi(\mathbf{u})+c(\mathbf{v},\mathbf{y})-c(\mathbf{u},\mathbf{y}),$$
$$\psi(\mathbf{u})\leqslant\psi(\mathbf{x})+c(\mathbf{u},\mathbf{y})-c(\mathbf{x},\mathbf{y}).$$

Noting that $\phi(\mathbf{u})<\psi(\mathbf{u})$ these inequalities imply

$$\phi(\mathbf{v})<\psi(\mathbf{x})+c(\mathbf{v},\mathbf{y})-c(\mathbf{x},\mathbf{y}). \tag{20}$$

Evaluating at $\mathbf{v}=\mathbf{x}$ yields $\mathbf{x}\in U$. Since $\mathbf{x}\in X$ was arbitrary, $X\subset U$. □

Remark 4.2. If $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ satisfies (H1)–(H3) while ϕ and ψ are both differentiable and c-concave, then by Proposition 3.4 (ii), the last lemma shows that when $\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))=\mathbf{u}-\nabla h^*(\nabla\phi(\mathbf{u}))$ and $\psi(\mathbf{u})>\phi(\mathbf{u})$ then $\psi(\mathbf{x})>\phi(\mathbf{x})$.

LEMMA 4.3. *Take ϕ, ψ, U and X as in Lemma* 4.1 *while* $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ *satisfies* (H1)–(H3). *Let ψ be c-concave on $\mathbf{R}^d$ and continuous at $\mathbf{p}\in\mathbf{R}^d$ with $\psi(\mathbf{p})=\phi(\mathbf{p})=0$. If $\partial^c\psi(\mathbf{p})$ and $\partial^c\phi(\mathbf{p})$ are disjoint then $\mathbf{p}$ lies a positive distance from X.*

Proof. To produce a contradiction, suppose that a sequence $\mathbf{x}_n\in X$ converges to $\mathbf{p}$. Then there exist $\mathbf{u}_n\in U$ with $\mathbf{y}_n\in\partial^c\phi(\mathbf{u}_n)$ such that $(\mathbf{x}_n,\mathbf{y}_n)\in\partial^c\psi$. Proposition C.4 guarantees that the $\mathbf{y}_n$ are bounded since $\mathbf{x}_n\to\mathbf{p}$. A subsequence must converge to a limit point $\mathbf{y}_n\to\mathbf{y}_o$, and $(\mathbf{p},\mathbf{y}_o)$ lies in the closed set $\partial^c\psi$. The hypotheses then yield $\mathbf{y}_o\notin\partial^c\phi(\mathbf{p})$ and $\phi(\mathbf{p})=0$, so there is some $\mathbf{v}\in\mathbf{R}^d$ for which

$$\phi(\mathbf{v})>c(\mathbf{v},\mathbf{y}_o)-c(\mathbf{p},\mathbf{y}_o). \tag{21}$$

On the other hand, the same logic which led to (20) yields

$$\phi(\mathbf{v})<\psi(\mathbf{x}_n)+c(\mathbf{v},\mathbf{y}_n)-c(\mathbf{x}_n,\mathbf{y}_n).$$

Since ψ is continuous at $\psi(\mathbf{p})=0$, the large n limit $\mathbf{x}_n\to\mathbf{p}$ and $\mathbf{y}_n\to\mathbf{y}_o$ contradicts (21):

$$\phi(\mathbf{v})\leqslant c(\mathbf{v},\mathbf{y}_o)-c(\mathbf{p},\mathbf{y}_o). \qquad \square$$

THEOREM 4.4 (geometrical characterization of the optimal map). *Fix a cost $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ where h strictly convex satisfies* (H1)–(H3), *and measures $\mu,\nu\in\mathcal{P}(\mathbf{R}^d)$. If μ is absolutely continuous with respect to Lebesgue then a map $\mathbf{s}$ pushing μ forward to ν is uniquely determined μ-almost everywhere by the requirement that it be of the form $\mathbf{s}(\mathbf{x})=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ for some c-concave ψ on $\mathbf{R}^d$.*

Proof. Suppose that in addition to ψ and $\mathbf{s}$, a second c-concave potential ϕ exists for which $\mathbf{t}(\mathbf{x}):=\mathbf{x}-\nabla h^*(\nabla\phi(\mathbf{x}))$ pushes μ forward to $\mathbf{t}_\#\mu=\mathbf{s}_\#\mu=\nu$. In any case $\mathbf{t}$ and $\mathbf{s}$ are defined μ-almost everywhere, and unless they coincide there exists some $\mathbf{p}\in\mathbf{R}^d$ at which both

(i) ψ and ϕ are differentiable but $\mathbf{s}(\mathbf{p})\neq\mathbf{t}(\mathbf{p})$, and

(ii) $\mathbf{p}$ is a Lebesgue point for $d\mu(\mathbf{x})=f(\mathbf{x})\,d\mathbf{x}$ with positive density $f(\mathbf{p})>0$.

Here $f\in L^1(\mathbf{R}^d)$ is the Radon–Nikodym derivative of μ with respect to Lebesgue.

Subtracting constants from both potentials yields $\psi(\mathbf{p})=\phi(\mathbf{p})=0$ without affecting the maps $\mathbf{t}$ and $\mathbf{s}$. From (i) it is clear that $\nabla\phi(\mathbf{p})\neq\nabla\psi(\mathbf{p})$, so motivated by the lemmas we define $U:=\{\mathbf{x}\in\operatorname{int}\operatorname{dom}\psi\mid\psi(\mathbf{x})>\phi(\mathbf{x})\}$. Here $\operatorname{int}\operatorname{dom}\psi$ denotes the interior of the set on which ψ is finite; on it ψ is continuous (by Theorem 3.3) while ϕ is upper semi-continuous, being an infimum of translates and shifts of $h(\mathbf{x})$. A contradiction will be derived by showing that the push-forwards $\mathbf{s}_\#\mu$ and $\mathbf{t}_\#\mu$—alleged to coincide—must differ on $V:=\partial^c\phi(U)$:

$$\mu[\mathbf{s}^{-1}(V)]<\mu[U]\leqslant\mu[\mathbf{t}^{-1}(V)]. \tag{22}$$

This set V is Borel—in fact σ-compact—since U is open while $\partial^c\phi$ is closed.

The second inequality is easy: $\mathbf{t}(\mathbf{x})$ is defined for μ-almost every $\mathbf{x}\in U$, while Proposition 3.4 (ii) implies $\{\mathbf{t}(\mathbf{x})\}=\partial^c\phi(\mathbf{x})\subset V$, or equivalently $\mathbf{x}\in\mathbf{t}^{-1}(V)$. Thus

$$\mu[U]\leqslant\mu[\mathbf{t}^{-1}(V)].$$

To prove the first inequality, observe that $\mathbf{s}^{-1}(V)\subset\partial^c\psi^{-1}(V)$ follows from the same proposition. Now $\mathbf{s}^{-1}(V)\subset\operatorname{int}\operatorname{dom}\psi$ combines with Lemma 4.1 to imply $\mathbf{s}^{-1}(V)\subset U$, whence

$$\mu[\mathbf{s}^{-1}(V)]\leqslant\mu[U].$$

Strict inequality is not yet apparent, but it will be derived from Lemma 4.3. Indeed $\partial^c\psi(\mathbf{p})=\{\mathbf{s}(\mathbf{p})\}\neq\{\mathbf{t}(\mathbf{p})\}=\partial^c\phi(\mathbf{p})$, so the lemma provides a neighbourhood Ω of $\mathbf{p}$ that is disjoint from $\mathbf{s}^{-1}(V)\subset\partial^c\psi^{-1}(V)$. It remains to verify that a little bit of the mass of μ in U lies in Ω, which will imply strict inequality in (22) and complete the proof.

This follows from our choice of $\mathbf{p}$. Translate μ, ψ and ϕ so that $\mathbf{p}=\mathbf{0}$ and consider the cone $C:=\{\mathbf{x}\,|\,\langle\mathbf{x},\nabla\psi(\mathbf{p})-\nabla\phi(\mathbf{p})\rangle\geqslant\frac{1}{2}|\mathbf{x}|\}$. Differentiability (i) of ψ and ϕ at $\mathbf{p}=0$ yields

$$\psi(\mathbf{x})-\phi(\mathbf{x})=\langle\mathbf{x},\nabla\psi(\mathbf{p})-\nabla\phi(\mathbf{p})\rangle+o(|\mathbf{x}|).$$

Thus $\mathbf{x}\in C$ sufficiently small implies $\mathbf{x}\in U$. Since $\mathbf{p}$ is a Lebesgue point (ii), the average value of $f(\mathbf{x})$ over $C\cap B_r(\mathbf{p})$ must converge to $f(\mathbf{p})>0$ as r shrinks to zero. For small r, this set lies both in U and in Ω, so $\mu[U\cap\Omega]>0$ and (22) is established. □

Summarizing our results for convex costs:

MAIN THEOREM 4.5 (strictly convex costs). *Fix $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ where h strictly convex satisfies* (H1)–(H3), *and Borel probability measures μ and ν on $\mathbf{R}^d$. If μ is absolutely continuous with respect to Lebesgue, then*

(i) *there is a c-concave function ψ on $\mathbf{R}^d$ such that the map $\mathbf{s}(\mathbf{x}):=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ pushes μ forward to ν;*

(ii) *the map $\mathbf{s}(\mathbf{x})$ is unique—up to a set of measure zero for μ;*

(iii) *$\gamma:=(\mathbf{id}\times\mathbf{s})_\#\mu$ is the only measure in $\Gamma(\mu,\nu)$ with c-cyclically monotone support. If the target measure ν satisfies the same hypothesis as μ, then*

(iv) *$\gamma=(\mathbf{t}\times\mathbf{id})_\#\nu$ for some inverse map $\mathbf{t}:\mathbf{R}^d\to\mathbf{R}^d$, and*

(v) *$\mathbf{t}(\mathbf{s}(\mathbf{x}))=\mathbf{x}$ μ-almost everywhere, while $\mathbf{s}(\mathbf{t}(\mathbf{y}))=\mathbf{y}$ ν-almost everywhere.*

Proof. (i) For any $\mu,\nu\in\mathcal{P}(\mathbf{R}^d)$, there is a joint measure $\gamma\in\Gamma(\mu,\nu)$ with c-cyclically monotone support: when (4) is finite this follows from Theorem 2.3 with Proposition 2.1, while more generally γ may be constructed as in [29, Theorem 12]. Smith and Knott's result—Theorem 2.7—guarantees the existence of a c-concave potential ψ on $\mathbf{R}^d$ with

$\partial^c\psi \supset \operatorname{spt}\gamma$. The map $\pi(\mathbf{x},\mathbf{y})=\mathbf{x}$ on $\mathbf{R}^d\times\mathbf{R}^d$ pushes γ forward to $\mu=\pi_\#\gamma$, while projecting the closed set $\partial^c\psi$ to a σ-compact set $\pi(\partial^c\psi)$ of full measure for μ. Proposition 3.4 shows that ψ is differentiable μ-almost everywhere on $\pi(\partial^c\psi)$, while Proposition 3.6 shows that $\mathbf{s}(\mathbf{x}):=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ pushes μ forward to ν; it also expresses γ in terms of $\mathbf{s}$ and μ.

(ii)–(iii) There is only one such map $\mathbf{s}$ by Theorem 4.4. Thus $\gamma=(\mathbf{id}\times\mathbf{s})_\#\mu$ is uniquely determined by μ, ν and c-cyclical monotonicity of its support.

(iv) Define $\tilde c(\mathbf{x},\mathbf{y}):=c(\mathbf{y},\mathbf{x})$ and $\tilde\gamma\in\Gamma(\nu,\mu)$ by $\tilde\gamma[V\times U]:=\gamma[U\times V]$. Then $\tilde\gamma$ has $\tilde c$-cyclically monotone support. The result (i) just established provides a map $\mathbf{t}(\mathbf{y})$ such that $\tilde\gamma=(\mathbf{id}\times\mathbf{t})_\#\nu$, or equivalently $\gamma=(\mathbf{t}\times\mathbf{id})_\#\nu$.

(v) Since $\gamma=(\mathbf{id}\times\mathbf{s})_\#\mu=(\mathbf{t}\times\mathbf{id})_\#\nu$, there are sets $U,V\subset\mathbf{R}^d$ of full mass $\mu[U]=1$ and $\nu[V]=1$ such that $\mathbf{x}\in U$ implies $(\mathbf{x},\mathbf{s}(\mathbf{x}))\in\operatorname{spt}\gamma$ while $\mathbf{y}\in V$ implies $(\mathbf{t}(\mathbf{y}),\mathbf{y})\in\operatorname{spt}\gamma$. Moreover, $\operatorname{spt}\gamma$ is c-cyclically monotone: Corollary 3.5 yields a set $N\subset\mathbf{R}^d$ of zero measure for μ such that $(\mathbf{x},\mathbf{y})$ and $(\mathbf{x},\mathbf{z})$ in $\operatorname{spt}\gamma$ with $\mathbf{y}\neq\mathbf{z}$ imply $\mathbf{x}\in N$. Choose $\mathbf{y}$ from the set $V\cap\mathbf{t}^{-1}(U\setminus N)$ which has full mass for ν. On one hand $(\mathbf{t}(\mathbf{y}),\mathbf{y})\in\operatorname{spt}\gamma$, while on the other $(\mathbf{t}(\mathbf{y}),\mathbf{s}(\mathbf{t}(\mathbf{y})))\in\operatorname{spt}\gamma$. Since $\mathbf{t}(\mathbf{y})\notin N$ one concludes that $\mathbf{s}(\mathbf{t}(\mathbf{y}))=\mathbf{y}$. By symmetry $\mathbf{t}(\mathbf{s}(\mathbf{x}))=\mathbf{x}$ holds on a set of full measure for μ. □

Remark 4.6. In fact, one may even assume the potential ψ of the theorem to be the c-transform (9) of a function on $\operatorname{spt}\nu$. This is clear from the proof of part (i), where we can appeal to Corollary 2.8 instead of Theorem 2.7.

Remark 4.7 (results for more concentrated measures). If the convex cost $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ is a derivative smoother than Lipschitz, $h\in C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$, then all our results—Theorems 3.7, 4.4 and 4.5—extend to measures which fail to be absolutely continuous with respect to Lebesgue, provided μ vanishes on Lipschitz submanifolds of dimension $d-1$ and hence on rectifiable sets. Of course, the cost must still satisfy (H1)–(H3).

The existence part of this assertion is clear: in the proof of Theorem 3.7 absolute continuity was used only to know that finiteness implies differentiability μ-almost everywhere for c-concave potentials ψ on $\mathbf{R}^d$:

$$\mu[\operatorname{dom}\psi\setminus\operatorname{dom}\nabla\psi]=0 \tag{23}$$

was a consequence of Proposition 3.4 (iv). Now if $h\in C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$, then Theorem 3.3 shows ψ to be locally semi-concave on the interior of $\operatorname{dom}\psi$, where by Proposition C.6 its differentiability can fail only on a rectifiable set of dimension $d-1$. The same theorem shows that the boundary of $\operatorname{dom}\psi$ is contained in the boundary of a convex set—hence a Lipschitz submanifold of dimension $d-1$. Thus (23) holds provided μ vanishes on rectifiable sets of dimension $d-1$.

On the other hand, to prove the uniqueness result of Theorem 4.4, it was necessary to find a point where (i) both ϕ and ψ are differentiable, with $\phi(\mathbf{p})=\psi(\mathbf{p})$ but $\mathbf{s}(\mathbf{p})\neq\mathbf{t}(\mathbf{p})$, and (ii) each neighbourhood of $\mathbf{p}$ intersects $\{\phi\neq\psi\}$ in a set carrying positive mass under μ. When μ was absolutely continuous with respect to Lebesgue, the second condition was fulfilled by choosing a Lebesgue point of μ with positive density. However, when ϕ and ψ are locally semi-concave and μ vanishes on sets of dimension $d-1$, then (ii) follows from (i) for any $\mathbf{p}\in\operatorname{spt}\mu$. This can be deduced from a version of the implicit function theorem [29, Theorem 17] which shows that near a point where $\nabla\psi(\mathbf{p})\neq\nabla\phi(\mathbf{p})$, the set $\{\phi=\psi\}$ is given by a Lipschitz function of $d-1$ variables. The full argument may be found in the proof of Theorem 6.3.

Proof of Theorem 1.2. Our argument uses results proved for μ absolutely continuous with respect to Lebesgue, which Remark 4.7 extends to the case where μ merely vanishes on rectifiable sets of dimension $d-1$ but $h\in C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$.

(i)–(ii) are already asserted by Theorem 4.5.

(iii)–(iv) By Proposition 2.1, there is a measure γ which minimizes $\mathcal{C}(\cdot)$ on $\Gamma(\mu,\nu)$. If $\mathcal{C}(\gamma)=\infty$, there is nothing further to prove. If $\mathcal{C}(\gamma)<\infty$ then γ has c-cyclically monotone support by Abdellaoui and Heinich's argument in Theorem 2.3. Thus γ coincides with the unique measure of Theorem 4.5, and the results (iii)–(iv) follow immediately.

Two points deserve further comment. A standard measure-theoretic argument shows that if $(\mathbf{id}\times\mathbf{s})_{\#}\mu=(\mathbf{id}\times\mathbf{t})_{\#}\mu$, then $\mathbf{s}(\mathbf{x})=\mathbf{t}(\mathbf{x})$ holds μ-almost everywhere. Thus the optimal map is unique. Finally, when $h(\mathbf{x})$ is differentiable, Corollary A.2 establishes the identity $(\nabla h)^{-1}=\nabla h^*$. □

Part II. Costs which are strictly concave as a function of distance

5. The role of optimal maps

At this point, we return to the economically natural costs $c(\mathbf{x},\mathbf{y})=l(|\mathbf{x}-\mathbf{y}|)$ given by strictly concave functions $l\geqslant 0$ of the distance. For these costs, an optimal measure γ for Kantorovich's problem does not generally lead to a solution $\mathbf{s}$ of the Monge problem unless its marginals $\mu,\nu\in\mathcal{P}(\mathbf{R}^d)$ are disjointly supported. The main difference stems from the fact that the cost gives a metric on $\mathbf{R}^d$, which satisfies the strict triangle inequality (15). The results summarized in §2 therefore imply that any mass which is common to μ and ν will stay in its place; it can be subtracted from the diagonal of γ. After doing so, what remains will be a measure of the form $(\mathbf{id}\times\mathbf{s})_{\#}[\mu-\nu]_+$ under suitable hypotheses on $\mu-\nu$. The map $\mathbf{s}$ is given by $\mathbf{s}(\mathbf{x})=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ where the potential ψ is the c-transform of a function on $\operatorname{spt}[\nu-\mu]_+$. The main goal of this section is to confirm this description of γ by demonstrating the existence of $\mathbf{s}$.

We begin by verifying that $c(\mathbf{x},\mathbf{y})$ is a metric on $\mathbf{R}^d$ and satisfies the triangle inequality strictly. This elementary lemma combines with the results of §2 to put fundamental limitations on the geometry of c-cyclically monotone sets.

LEMMA 5.1 (concave costs metrize $\mathbf{R}^d$). *If* $l\colon[0,\infty)\to[0,\infty)$ *is strictly concave and* $l(0)=0$, *then* $c(\mathbf{x},\mathbf{y}):=l(|\mathbf{x}-\mathbf{y}|)$ *defines a metric on* $\mathbf{R}^d$ *and* $c(\mathbf{x},\mathbf{y})<c(\mathbf{x},\mathbf{p})+c(\mathbf{p},\mathbf{y})$ *unless* $\mathbf{p}=\mathbf{x}$ *or* $\mathbf{p}=\mathbf{y}$.

Proof. Since $l(\lambda)$ is strictly concave on $[0,\infty)$ and yet remains positive, it must increase strictly. Thus $c(\mathbf{x},\mathbf{y})=0$ precisely when $\mathbf{x}=\mathbf{y}$. Symmetry is obvious under $\mathbf{x}\leftrightarrow\mathbf{y}$, so the only thing to verify is the strict triangle inequality. Therefore, let $\mathbf{x},\mathbf{y},\mathbf{p}\in\mathbf{R}^d$ with $\mathbf{p}$ different from both $\mathbf{x}$ and $\mathbf{y}$. Define $\lambda:=|\mathbf{x}-\mathbf{p}|+|\mathbf{p}-\mathbf{y}|$. Then $|\mathbf{x}-\mathbf{p}|=(1-t)\lambda$ and $|\mathbf{p}-\mathbf{y}|=t\lambda$ for some $t\in(0,1)$. Since $\lambda\neq0$, invoking strict concavity of l together with $l(0)=0$ yields $c(\mathbf{x},\mathbf{p})=l((1-t)\lambda+t0)>(1-t)l(\lambda)$ and $c(\mathbf{p},\mathbf{y})=l((1-t)0+t\lambda)>tl(\lambda)$. These inequalities sum to $c(\mathbf{x},\mathbf{p})+c(\mathbf{p},\mathbf{y})>l(\lambda)$. On the other hand, the usual triangle inequality states that $\lambda\geqslant|\mathbf{x}-\mathbf{y}|$, so monotonicity of l implies $l(\lambda)\geqslant l(|\mathbf{x}-\mathbf{y}|)=c(\mathbf{x},\mathbf{y})$. □

For any optimal measure $\gamma\in\Gamma(\mu,\nu)$, Proposition 2.9 can now be invoked to conclude that any mass common to $\mu,\nu\in\mathcal{P}(\mathbf{R}^d)$ will be located on the diagonal $D:=\{(\mathbf{x},\mathbf{x})\}$ in $\mathbf{R}^d\times\mathbf{R}^d$. Here we proceed by assuming that μ and ν have no mass in common, to develop a theory which parallels the convex case, before returning to full generality in our main theorem. Since $D\subset\partial^c\psi$ whenever ψ is c-concave (Lemma 2.11), it will be convenient to restrict our attention to the off-diagonal part $\partial^c_\circ\psi:=\{(\mathbf{x},\mathbf{y})\in\partial^c\psi\mid\mathbf{x}\neq\mathbf{y}\}$ of the c-superdifferential; $\partial^c_\circ\psi(\mathbf{x}):=\partial^c\psi(\mathbf{x})\setminus\{\mathbf{x}\}$ and $\partial^c_\circ\psi(V):=\bigcup_{\mathbf{x}\in V}\partial^c_\circ\psi(\mathbf{x})$ are defined in the obvious way. As for convex costs, a lemma will be required relating differentiability to c-superdifferentiability through the conjugate cost h^* from (11).

LEMMA 5.2 (the c-superdifferential lies in the graph of a map). *Let* $c(\mathbf{x},\mathbf{y}):=h(\mathbf{x}-\mathbf{y}):=l(|\mathbf{x}-\mathbf{y}|)$ *be continuous with* $l(\lambda)\geqslant0$ *strictly concave, and suppose that* $\psi\colon\mathbf{R}^d\to\mathbf{R}$ *is differentiable at some* $\mathbf{x}\in\mathbf{R}^d$. *Then* $\mathbf{y}\in\partial^c_\circ\psi(\mathbf{x})$ *implies that* h^* *is differentiable at* $\nabla\psi(\mathbf{x})$ *and that* $\mathbf{y}=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$.

Proof. Let $\mathbf{y}\in\partial^c_\circ\psi(\mathbf{x})$. Then Lemma 3.1 yields the subgradient $\nabla\psi(\mathbf{x})\in\partial_\cdot h(\mathbf{x}-\mathbf{y})$. Since $\mathbf{x}\neq\mathbf{y}$, the cost h is also *super*differentiable at $\mathbf{x}-\mathbf{y}$ by Corollary A.5, hence differentiable with $\nabla h(\mathbf{x}-\mathbf{y})=\nabla\psi(\mathbf{x})$. This gradient does not vanish since $h(\mathbf{x})=l(|\mathbf{x}|)$ with $l(\lambda)\geqslant0$ strictly concave and hence strictly increasing. Proposition A.6 (ii)–(iii) implies both $(\nabla\psi(\mathbf{x}),\mathbf{x}-\mathbf{y})\in\partial^\cdot h^*$ and differentiability of h^* at $\nabla\psi(\mathbf{x})$, whence $\nabla h^*(\nabla\psi(\mathbf{x}))=\mathbf{x}-\mathbf{y}$. □

For the c-transform ψ of a function on a closed set $V\subset\mathbf{R}^d$, a converse is provided in the next section. Our present priority is a regularity result for ψ outside of V:

PROPOSITION 5.3 (local semi-concavity for c-transforms). *Take $V\subset\mathbf{R}^d$ to be closed and define $c(\mathbf{x},\mathbf{y}):=h(\mathbf{x}-\mathbf{y}):=l(|\mathbf{x}-\mathbf{y}|)$ on $\mathbf{R}^d$, where $l:[0,\infty)\to\mathbf{R}$ is concave non-decreasing. Then the c-transform ψ of any function on V will be locally semi-concave on $\mathbf{R}^d\setminus V$.*

Proof. Let $\mathbf{p}\in\mathbf{R}^d$ be separated from V by a distance greater than $\varepsilon>0$. We shall show local semi-concavity of ψ at $\mathbf{p}$. Define $\xi\geqslant 0$ using the right-derivative $2\varepsilon\xi:=l'(\varepsilon^+)$ of l at ε. Then the function $l_\varepsilon(\lambda)=l(\lambda)-\xi\lambda^2$ is concave on $[\varepsilon,\infty)$, and non-increasing since $l'_\varepsilon(\varepsilon^+)=0$. Extend this function to $\lambda\leqslant\varepsilon$ by making $l_\varepsilon(\lambda)$ constant-valued there. Then $h_\varepsilon(\mathbf{x}):=l_\varepsilon(|\mathbf{x}|)$ will be concave on $\mathbf{R}^d$: taking $\mathbf{x},\mathbf{y}\in\mathbf{R}^d$ and $0<t<1$ implies

$$\begin{aligned} h_\varepsilon((1-t)\mathbf{x}+t\mathbf{y}) &\geqslant l_\varepsilon((1-t)|\mathbf{x}|+t|\mathbf{y}|) \\ &\geqslant (1-t)h_\varepsilon(\mathbf{x})+th_\varepsilon(\mathbf{y}). \end{aligned}$$

Note that $h(\mathbf{x})=h_\varepsilon(\mathbf{x})+\xi\mathbf{x}^2$ whenever $|\mathbf{x}|\geqslant\varepsilon$. For a small enough ball U around $\mathbf{p}$, taking $\mathbf{x}\in U$ and $\mathbf{y}\in V$ implies $|\mathbf{x}-\mathbf{y}|>\varepsilon$. Then (7) yields

$$\psi(\mathbf{x})-\xi\mathbf{x}^2=\inf_{(\mathbf{y},\alpha)\in\mathcal{A}} h_\varepsilon(\mathbf{x}-\mathbf{y})-2\xi\langle\mathbf{x},\mathbf{y}\rangle+\xi\mathbf{y}^2+\alpha,$$

where $\mathcal{A}\subset V\times\mathbf{R}$ since ψ is the c-transform of a function on V. Thus $\psi(\mathbf{x})-\xi\mathbf{x}^2$ is manifestly concave on U: it is the infimum of a family of concave functions of $\mathbf{x}\in U$. Local semi-concavity of ψ is established at $\mathbf{p}$. □

PROPOSITION 5.4 (a map between marginals with disjoint support). *Fix $c(\mathbf{x},\mathbf{y}):=h(\mathbf{x}-\mathbf{y}):=l(|\mathbf{x}-\mathbf{y}|)$ continuous with $l(\lambda)\geqslant 0$ strictly concave, and measures $\mu,\nu\in\mathcal{P}(\mathbf{R}^d)$. Suppose that a joint measure $\gamma\in\Gamma(\mu,\nu)$ is supported on $\partial^c\psi\supset\operatorname{spt}\gamma$, where $\psi:\mathbf{R}^d\to\mathbf{R}$ is the c-transform of a function on $\operatorname{spt}\nu$. If μ vanishes on $\operatorname{spt}\nu$ and on rectifiable sets of dimension $d-1$, then the map $\mathbf{s}(\mathbf{x}):=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ pushes μ forward to ν. In fact, $\gamma=(\mathbf{id}\times\mathbf{s})_\#\mu$.*

Proof. To begin, one would like to know that the map $\mathbf{s}(\mathbf{x})$ is Borel and defined μ-almost everywhere. Proposition 5.3 shows ψ to be locally semi-concave on the open set $\Omega:=\mathbf{R}^d\setminus\operatorname{spt}\nu$, so differentiability of ψ can only fail on a rectifiable set of dimension $d-1$ in Ω by Proposition C.6. The hypotheses on μ ensure $\mu[\Omega]=1$, and that the map $\nabla\psi$ is defined μ-almost everywhere. Moreover, $\Omega\times\operatorname{spt}\nu$ is a set of full measure for γ. Since it is disjoint from the diagonal $D\subset\mathbf{R}^d\times\mathbf{R}^d$, one obtains $\gamma[\partial^c_\circ\psi]=1$ because $\operatorname{spt}\gamma$ is contained in the closed set $\partial^c\psi$. Therefore, define $S:=\{(\mathbf{x},\mathbf{y})\in\partial^c_\circ\psi\mid\mathbf{x}\in\operatorname{dom}\nabla\psi\}$, where

$\operatorname{dom}\nabla\psi$ denotes the subset of Ω on which ψ is differentiable. For $(\mathbf{x},\mathbf{y})\in S$, Lemma 5.2 implies that $\mathbf{s}$ is defined at $\mathbf{x}$ and $\mathbf{y}=\mathbf{s}(\mathbf{x})$. Thus $\mathbf{s}$ is defined on the projection of S onto $\mathbf{R}^d$ by $\pi(\mathbf{x},\mathbf{y})=\mathbf{x}$; it is a Borel map since Propositions C.6 and A.6 show that $\nabla\psi$ and ∇h^* are. Moreover, the set $\pi(S)$ is Borel and has full measure for μ: both $\partial^c_\circ\psi$ and $\pi(\partial^c_\circ\psi)$ are σ-compact, so $\pi(S)=\pi(\partial^c_\circ\psi)\cap\operatorname{dom}\nabla\psi$ is the intersection of two Borel sets with full measure.

The verification that $(\mathbf{id}\times\mathbf{s})_\#\mu=\gamma$ and $\mathbf{s}_\#\mu=\nu$ proceeds as in the proof of Proposition 3.6: we have just seen that $\mathbf{y}=\mathbf{s}(\mathbf{x})$ if $(\mathbf{x},\mathbf{y})\in S$, from which (19) is immediate; the remainder of the proof is identical. □

At this point, an argument parallel to the proof of Theorem 3.7 would lead to the analogous results for costs $c(\mathbf{x},\mathbf{y})=l(|\mathbf{x}-\mathbf{y}|)$ given by concave functions $l\geqslant 0$ of the distance. Since existence and uniqueness of optimal maps for the Monge problem follow from our main theorem in any case, we proceed toward its demonstration.

6. Uniqueness of optimal solutions

The goal of this final section is to prove the uniqueness of measures γ on $\mathbf{R}^d\times\mathbf{R}^d$ with fixed marginals $\mu,\nu\in\mathcal{P}(\mathbf{R}^d)$ and c-cyclically monotone support. Here the cost $c(\mathbf{x},\mathbf{y})=l(|\mathbf{x}-\mathbf{y}|)$ is given by a strictly concave function $l\geqslant 0$ of the distance. Preceding developments reduce this problem to the case in which μ and ν are mutually singular, and one would then like to know that there is a unique map $\mathbf{s}(\mathbf{x})=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ pushing μ forward to ν derived from the c-transform ψ of a function on $\operatorname{spt}\nu$. As it turns out, this will be the case provided that μ concentrates no mass on the closed set $\operatorname{spt}\nu$, nor on sets of dimension $d-1$.

The proof parallels the development for convex costs in §4, but this time attention is focused on the off-diagonal part $\partial^c_\circ\psi:=\{(\mathbf{x},\mathbf{y})\in\partial^c\psi\mid\mathbf{x}\neq\mathbf{y}\}$ of the c-superdifferential of ψ. For $V\subset\mathbf{R}^d$, we define $\partial^c_\circ\psi^{-1}(V)\subset\mathbf{R}^d$ to consist of those $\mathbf{x}$ which are related to some $\mathbf{y}\in V$ by $(\mathbf{x},\mathbf{y})\in\partial^c_\circ\psi$. For c-concave ψ, Lemma 2.11 (and Lemma 5.1) shows that the diagonal part of $\partial^c\psi$ carries no information about ψ. The next proposition characterizes $\partial^c_\circ\psi$ at points where $\mathbf{s}(\mathbf{x})$ is defined. It provides a converse to Lemma 5.2.

Proposition 6.1 (c-superdifferentiability of c-transforms). *Fix $c(\mathbf{x},\mathbf{y}):=h(\mathbf{x}-\mathbf{y}):=l(|\mathbf{x}-\mathbf{y}|)$ continuous with $l(\lambda)\geqslant 0$ strictly concave, and a closed set $V\subset\mathbf{R}^d$. Let $\psi:\mathbf{R}^d\to\mathbf{R}$ be the c-transform of a function on V and suppose that $\mathbf{s}(\mathbf{x}):=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ can be defined at some $\mathbf{p}\in\mathbf{R}^d\setminus V$; (i.e., ψ is differentiable at $\mathbf{p}$, and h^* at $\nabla\psi(\mathbf{p})$). Then $\partial^c_\circ\psi(\mathbf{p})=\{\mathbf{s}(\mathbf{p})\}$.*

Proof. From Lemma 5.2 it is already clear that $\partial^c_\circ\psi(\mathbf{p})\subset\{\mathbf{s}(\mathbf{p})\}$. One need only prove that $\partial^c_\circ\psi(\mathbf{p})$ is non-empty. Therefore, assume that $\mathbf{s}(\mathbf{p})$ is defined at some $\mathbf{p}\in\mathbf{R}^d\setminus V$. Because ψ is the c-transform of a function on V, there is a sequence $(\mathbf{y}_n,\alpha_n)\in\mathcal{A}\subset V\times\mathbf{R}$ such that

$$\psi(\mathbf{p}) = \lim_n c(\mathbf{p},\mathbf{y}_n)+\alpha_n. \tag{24}$$

As is shown below, the $|\mathbf{y}_n|$ must be bounded. We first assume this bound to complete the proof. Since the $|\mathbf{y}_n|$ are bounded, a subsequence must converge to a limit $\mathbf{y}_n\to\mathbf{y}$ in the closed set V. Certainly $\mathbf{y}\neq\mathbf{p}$ since $\mathbf{p}$ is outside of V. On the other hand, $\mathbf{y}\in\partial^c_\circ\psi(\mathbf{p})$, since for all $\mathbf{x}\in\mathbf{R}^d$ (7) and (24) imply

$$\psi(\mathbf{x}) \leqslant \inf_n c(\mathbf{x},\mathbf{y}_n)+\alpha_n \tag{25}$$
$$\leqslant c(\mathbf{x},\mathbf{y})+\psi(\mathbf{p})-c(\mathbf{p},\mathbf{y}). \tag{26}$$

It remains only to bound the $|\mathbf{y}_n|$. To produce a contradiction, suppose that a subsequence diverges in a direction $\hat{\mathbf{y}}_n\to\hat{\mathbf{y}}$. Then $|\mathbf{p}-\mathbf{y}_n|$ is bounded away from zero by $\delta>0$. Corollary A.5 shows that there is a supergradient $\mathbf{w}_n\in\partial^\cdot c(\mathbf{p}-\mathbf{y}_n)$ for each n, while Lemma A.4 shows that the $\mathbf{w}_n$ must be bounded: the right-derivative of l is positive decreasing, so $l'(\delta^+)\geqslant|\mathbf{w}_n|$. The $|\mathbf{y}_n|$ can only diverge if $|\mathbf{w}_n|$ tends to $l'(\infty):=\inf_\lambda l'(\lambda^+)$. Taking a subsequence if necessary ensures that the $\mathbf{w}_n$ converge to a limit $\mathbf{w}\in\mathbf{R}^d$. The uniform superdifferentiability of h in Corollary A.5 gives

$$h(\mathbf{p}-\mathbf{y}_n+\mathbf{x})\leqslant h(\mathbf{p}-\mathbf{y}_n)+\langle\mathbf{x},\mathbf{w}_n\rangle+O_\delta(\mathbf{x}^2)$$

for arbitrary $\mathbf{x}\in\mathbf{R}^d$ and $O_\delta(\mathbf{x}^2)$ independent of n. Combined with (25) this yields

$$\psi(\mathbf{p}+\mathbf{x})\leqslant\psi(\mathbf{p})+\langle\mathbf{x},\mathbf{w}\rangle+O_\delta(\mathbf{x}^2),$$

where the large n limit has been taken using (24). Thus $\mathbf{w}\in\partial^\cdot\psi(\mathbf{p})$. On the other hand, differentiability of ψ at $\mathbf{p}$ implies $\partial^\cdot\psi(\mathbf{p})=\{\nabla\psi(\mathbf{p})\}$, whence $\mathbf{w}=\nabla\psi(\mathbf{p})$. Now $(\mathbf{w},\mathbf{p}-\mathbf{s}(\mathbf{p}))\in\partial^\cdot h^*$ follows from the definition of $\mathbf{s}(\mathbf{p})$. Since $|\mathbf{w}|\leqslant l'(\delta^+)<\sup_{\lambda>0}l'(\lambda)$ one cannot have $\mathbf{s}(\mathbf{p})=\mathbf{p}$ without contradicting Proposition A.6 (iv). Thus $\mathbf{s}(\mathbf{p})\neq\mathbf{p}$ and the same proposition yields $(\mathbf{p}-\mathbf{s}(\mathbf{p}),\mathbf{w})\in\partial^\cdot h$. Lemma A.4 gives $(|\mathbf{p}-\mathbf{s}(\mathbf{p})|,|\mathbf{w}|)\in\partial^\cdot l$. Since $l(\lambda)$ is strictly concave, $|\mathbf{w}|>l'(\infty)$ whence the $\mathbf{y}_n$ are bounded. □

LEMMA 6.2. *Let $c(\mathbf{x},\mathbf{y}):=h(\mathbf{x}-\mathbf{y}):=l(|\mathbf{x}-\mathbf{y}|)$ be continuous with $l(\lambda)\geqslant0$ strictly concave. Take both $\phi,\psi:\mathbf{R}^d\to\mathbf{R}$ to vanish at a point $\mathbf{p}\in\mathbf{R}^d$ where ψ is locally semi-concave. If $U:=\{\mathbf{x}\mid\psi(\mathbf{x})>\phi(\mathbf{x})\}$ and $X:=\partial^c_\circ\psi^{-1}(\partial^c\phi(U))$ then $X\subset U$. Moreover, if $\mathbf{p}-\nabla h^*(\nabla\psi(\mathbf{p}))$ is defined but is not in $\partial^c\phi(\mathbf{p})$, then $\mathbf{p}$ lies a positive distance from X.*

Proof. Lemma 4.1, which was proved for all costs, yields $X\subset U$ immediately from $\partial^c_\circ\psi\subset\partial^c\psi$; the only thing to prove is that $\mathbf{p}$ is not from the closure of X. To produce a contradiction, suppose that $\mathbf{x}_n\in X$ converges to $\mathbf{p}$. Then there exist $\mathbf{u}_n\in U$ with

$\mathbf{y}_n \in \partial^c \phi(\mathbf{u}_n)$ such that $(\mathbf{x}_n, \mathbf{y}_n) \in \partial^c_\circ \psi$. Corollary C.8 implies that the $\mathbf{y}_n$ converge to $\mathbf{y}_o := \mathbf{p} - \nabla h^*(\nabla \psi(\mathbf{p}))$. Since $\mathbf{y}_o \notin \partial^c \phi(\mathbf{p})$ and $\phi(\mathbf{p}) = 0$, (21) holds for some $\mathbf{v} \in \mathbf{R}^d$. A contradiction is derived by the remainder of the argument which proved Lemma 4.3. □

THEOREM 6.3 (uniqueness of optimal maps). *Fix $c(\mathbf{x}, \mathbf{y}) := h(\mathbf{x} - \mathbf{y}) := l(|\mathbf{x} - \mathbf{y}|)$ continuous with $l(\lambda) \geqslant 0$ strictly concave, and measures $\mu, \nu \in \mathcal{P}(\mathbf{R}^d)$. Let μ vanish on a closed set $Y \supseteq \operatorname{spt} \nu$ and on rectifiable sets of dimension $d-1$. A map of the form $\mathbf{s}(\mathbf{x}) = \mathbf{x} - \nabla h^*(\nabla \psi(\mathbf{x}))$, with $\psi: \mathbf{R}^d \to \mathbf{R}$ the c-transform of a function on Y, is uniquely determined μ-almost everywhere by the requirement that $\mathbf{s}_\# \mu = \nu$.*

Proof. Suppose that $\phi: \mathbf{R}^d \to \mathbf{R}$ satisfies the same hypotheses as ψ, and that both $\mathbf{t}(\mathbf{x}) := \mathbf{x} - \nabla h^*(\nabla \phi(\mathbf{x}))$ and $\mathbf{s}(\mathbf{x})$ push μ forward to ν. Then both $\mathbf{t}$ and $\mathbf{s}$ must be defined μ-almost everywhere, while Proposition 5.3 shows ϕ and ψ to be locally semi-concave on the open set $\Omega := \mathbf{R}^d \setminus Y$. This set has full mass $\mu[\Omega] = 1$ by hypothesis. As in the proof of Theorem 4.4, a contradiction will be derived if there is any point $\mathbf{p} \in \Omega \cap \operatorname{spt} \mu$ at which both $\mathbf{t}(\mathbf{p})$ and $\mathbf{s}(\mathbf{p})$ are defined but do not coincide. Then $\nabla \phi(\mathbf{p}) \neq \nabla \psi(\mathbf{p})$ though both gradients exist, while subtracting constants from each potential yields $\phi(\mathbf{p}) = \psi(\mathbf{p}) = 0$ without changing the map $\mathbf{t}$ or $\mathbf{s}$. By the local semi-concavity at $\mathbf{p}$, one can express $\phi - \psi$ as the difference of two concave functions $\phi(\mathbf{x}) - \lambda \mathbf{x}^2$ and $\psi(\mathbf{x}) - \lambda \mathbf{x}^2$ near $\mathbf{p}$. Then a non-smooth implicit function theorem [29, Theorem 17] applies: since $\phi - \psi$ vanishes at $\mathbf{p}$ but has non-zero gradient, there is a neighbourhood of $\mathbf{p}$ on which $\phi = \psi$ occurs precisely on the graph of a Lipschitz function of $d-1$ variables. This set has zero measure for μ. On the other hand, all of the neighbourhoods of $\mathbf{p} \in \operatorname{spt} \mu$ must have positive measure for μ. Exchanging the roles of ψ and ϕ if necessary, $U := \{\mathbf{x} \in \mathbf{R}^d \mid \psi(\mathbf{x}) > \phi(\mathbf{x})\}$ intersects each such neighbourhood B in a set with positive μ-measure.

The continuity of ϕ and ψ shown in Lemma 2.11 ensures that U is open and $\partial^c \phi$ is closed, whence $V := \partial^c \phi(U)$ is σ-compact. As before, the contradiction is obtained by showing that the push-forwards $\mathbf{s}_\# \mu$ and $\mathbf{t}_\# \mu$ disagree on V:

$$\mu[\mathbf{s}^{-1}(V)] < \mu[U] \leqslant \mu[\mathbf{t}^{-1}(V)]. \tag{27}$$

This is derived from $\mathbf{t}(\mathbf{p}) \neq \mathbf{s}(\mathbf{p})$ in the following way:

One knows that $\mathbf{s}(\mathbf{p}) \neq \mathbf{p}$ and also that $\mathbf{s}(\mathbf{p}) \notin \partial^c_\circ \phi(\mathbf{p}) = \{\mathbf{t}(\mathbf{p})\}$ from Proposition 6.1. Thus the hypotheses of Lemma 6.2 are satisfied. As a consequence, $\partial^c_\circ \psi^{-1}(V) \subset U$ but excludes a neighbourhood B of $\mathbf{p} \in \operatorname{spt} \mu$. Now $\mu[B \cap U] > 0$ by construction, whence $\mu[\partial^c_\circ \psi^{-1}(V)] < \mu[U]$. The strict inequality (27) is derived by noting that $\mu[\mathbf{s}^{-1}(V)] = \mu[\Omega \cap \mathbf{s}^{-1}(V)]$, while $\Omega \cap \mathbf{s}^{-1}(V) \subset \partial^c_\circ \psi^{-1}(V)$ follows from Proposition 6.1.

The second inequality in (27) is established by observing that U and $\{\mathbf{x} \in U \cap \Omega \mid \mathbf{t}(\mathbf{x}) \text{ is defined}\}$ have the same mass for μ; the latter set is seen to be contained in $\mathbf{t}^{-1}(V)$ by applying Proposition 6.1 to ϕ. □

Our conclusions for costs which are strictly concave as a function of distance are summarized by the following theorem. It assumes continuity of the cost function $c(\mathbf{x},\mathbf{y})$—but this assumption can be relaxed through Corollary 2.10 to allow a discontinuous drop at the origin. Such a drop represents a fixed penalty (per unit mass) for initiating motion—a "loading cost" in economics.

MAIN THEOREM 6.4 (strictly concave cost as a function of distance). *Fix* $c(\mathbf{x},\mathbf{y}):=h(\mathbf{x}-\mathbf{y}):=l(|\mathbf{x}-\mathbf{y}|)$ *continuous with* $l(\lambda)\geqslant 0$ *strictly concave and* $l(0)=0$. *Given two measures* $\mu,\nu\in\mathcal{P}(\mathbf{R}^d)$, *define* $\mu_o:=[\mu-\nu]_+$ *and* $\nu_o:=[\nu-\mu]_+$, *and assume that* μ_o *vanishes on* $\operatorname{spt}\nu_o$ *and on rectifiable sets of dimension* $d-1$. *Then*

(i) *there is a unique measure* $\gamma\in\Gamma(\mu,\nu)$ *with c-cyclically monotone support;*

(ii) *its restriction to the diagonal is given by* $\gamma_d=(\mathbf{id}\times\mathbf{id})_\#(\mu-\mu_o)$;

(iii) *the c-transform* $\psi:\mathbf{R}^d\to\mathbf{R}$ *of some function on* $\operatorname{spt}\nu_o$ *induces a map* $\mathbf{s}:=\mathbf{id}-(\nabla h)^*\circ\nabla\psi$ *which pushes* μ_o *forward to* ν_o;

(iv) *the map* $\mathbf{s}(\mathbf{x})$ *of* (iii) *is unique—up to a set of zero measure for* μ_o;

(v) *the off-diagonal part of* $\gamma=\gamma_d+\gamma_o$ *is given by* $\gamma_o=(\mathbf{id}\times\mathbf{s})_\#\mu_o$.

If ν_o *also vanishes on* $\operatorname{spt}\mu_o$ *and on rectifiable sets of dimension* $d-1$, *then*

(vi) *there exists an inverse map* $\mathbf{t}(\mathbf{y})$ *such that* $\gamma_o=(\mathbf{t}\times\mathbf{id})_\#\nu_o$, *and*

(vii) $\mathbf{t}(\mathbf{s}(\mathbf{x}))=\mathbf{x}$ *almost everywhere with respect to* μ_o *while* $\mathbf{s}(\mathbf{t}(\mathbf{y}))=\mathbf{y}$ *almost everywhere with respect to* ν_o.

Proof. (i) Once again, the existence of a joint measure $\gamma\in\Gamma(\mu,\nu)$ with c-cyclically monotone support follows either from Theorem 2.3 and Proposition 2.1 (when $\mathcal{C}(\gamma)<\infty$) or from [29, Theorem 12] otherwise. Denote the restriction of γ to the diagonal by γ_d, and the off-diagonal remainder by $\gamma_o=\gamma-\gamma_d$. If we succeed in establishing the rest of this theorem, uniqueness of γ is an immediate corollary: the off-diagonal part $\gamma_o=(\mathbf{id}\times\mathbf{s})_\#\mu_o$ will be uniquely determined by μ and ν through (iii)–(v), while (ii) gives the restriction of γ to the diagonal.

(ii) Propositions 2.9 and 5.1 verify that $\gamma_d=(\mathbf{id}\times\mathbf{id})_\#(\mu-\mu_o)$.

(iii)–(v) On the other hand $\gamma_o\in\Gamma(\mu_o,\nu_o)$ and, like γ, has c-cyclically monotone support. Noting $\operatorname{spt}\gamma_o\subset\operatorname{spt}\mu_o\times\operatorname{spt}\nu_o$, Corollary 2.8 yields $\operatorname{spt}\gamma_o\subset\partial^c\psi$, where $\psi:\mathbf{R}^d\to\mathbf{R}\cup\{-\infty\}$ is the c-transform of a function on $\operatorname{spt}\nu_o$. In fact, ψ is finite-valued by Lemma 2.11. In general, γ_o will have mass $\lambda^{-1}:=\gamma_o[\mathbf{R}^d\times\mathbf{R}^d]$ less than one, but $\lambda\gamma_o$ is a probability measure with the same support as γ_o; otherwise $\gamma_o=0$ and there is nothing to prove. Moreover, $\lambda\gamma_o\in\Gamma(\lambda\mu_o,\lambda\nu_o)$ so Proposition 5.4 shows that $\lambda\gamma_o=(\mathbf{id}\times\mathbf{s})_\#\lambda\mu_o$. Linearity of $(\mathbf{id}\times\mathbf{s})_\#$ completes the proof of (v) and of (iii).

(iv) Since the map $\mathbf{s}$ in (iii) pushes $\lambda\mu_o$ forward to $\lambda\nu_o$, it is uniquely determined μ_o-almost everywhere in view of Theorem 6.3.

(vi) Let $\tilde{\gamma}\in\Gamma(\nu,\mu)$ denote the measure defined by $\tilde{\gamma}[U\times V]=\gamma[V\times U]$. Then $\tilde{\gamma}$ has c-cyclically monotone support. If ν_o vanishes on $\operatorname{spt}\mu_o$ and on rectifiable sets of dimension $d-1$, (iii)–(v) guarantee the c-transform ϕ on a function on $\operatorname{spt}\mu_o$ for which the map $\mathbf{t}(\mathbf{y}):=\mathbf{y}-\nabla h^*(\nabla\phi)$ yields $\tilde{\gamma}_o=(\mathbf{id}\times\mathbf{t})_{\#}\nu_o$. This is equivalent to (vi).

(vii) Since $\mu_o[\operatorname{spt}\nu_o]=\nu_o[\operatorname{spt}\mu_o]=0$, (vi) implies a set $V\subset\mathbf{R}^d\setminus\operatorname{spt}\mu_o$ of full measure for ν_o, on which $\mathbf{y}\in V$ implies $\mathbf{t}(\mathbf{y})\notin\operatorname{spt}\nu_o$ but $(\mathbf{y},\mathbf{t}(\mathbf{y}))\in\operatorname{spt}\tilde{\gamma}_o$. Let $U\subset\mathbf{R}^d$ be the (Borel) set on which $\mathbf{s}(\mathbf{x})$ is defined. Then $\mu_o[U]=1$, which with $\mu_o=\mathbf{t}_{\#}\nu_o$ implies that $\mathbf{t}^{-1}(U)$ must carry the full mass of ν_o. Now assume $\mathbf{y}\in\mathbf{t}^{-1}(U)\cap V$. Then $\mathbf{s}$ is defined at $\mathbf{t}(\mathbf{y})\in\mathbf{R}^d\setminus\operatorname{spt}\nu_o$ while $(\mathbf{t}(\mathbf{y}),\mathbf{y})\in\operatorname{spt}\gamma_o\subset\partial^c\psi$. Proposition 6.1 yields $\mathbf{t}(\mathbf{y})\neq\mathbf{y}$ when applied to ϕ, and $\mathbf{y}=\mathbf{s}(\mathbf{t}(\mathbf{y}))$ when applied to ψ. Thus $\mathbf{y}=\mathbf{s}(\mathbf{t}(\mathbf{y}))$ holds on a set of full measure for ν_o. The other half of (vii) follows by symmetry. □

Proof of Theorem 1.4. Assume first that the cost $c(\mathbf{x},\mathbf{y})$ is continuous and vanishes when $\mathbf{x}=\mathbf{y}$. Then (i)–(ii) are asserted by Theorem 6.4.

(iii)–(v) Proposition 2.1 yields an optimal measure γ in $\Gamma(\mu,\nu)$. If $\mathcal{C}(\gamma)=\infty$, there is nothing more to prove. Otherwise, γ has c-cyclically monotone support by Theorem 2.3 so it coincides with the unique measure of Theorem 6.4. Results (iii)–(v) follow immediately.

Of course, the fact that $c(\mathbf{0})=0$ is completely irrelevant: none of the assertions in the theorem are sensitive to the addition of an overall constant to $c(\mathbf{x},\mathbf{y})$; for probability measures γ the only effect is to shift $\mathcal{C}(\gamma)$ by the same constant, while the class of c-concave functions is not modified. We therefore proceed to the case of discontinuous costs.

Any strictly concave function $l(\lambda)\geqslant0$ of $\lambda\geqslant0$ must increase with λ; it must also be continuous except at $\lambda=0$. Thus there is a continuous function $\tilde{c}(\mathbf{x},\mathbf{y})$ which agrees with $c(\mathbf{x},\mathbf{y}):=l(|\mathbf{x}-\mathbf{y}|)$ except that $\tilde{c}(\mathbf{0})\geqslant c(\mathbf{0})$. Apart from an irrelevant additive constant, $\tilde{c}$ is a continuous metric on $\mathbf{R}^d$ which by Lemma 5.1 satisfies the triangle inequality strictly. Corollary 2.10 then asserts that the optimal measures for c and $\tilde{c}$ coincide. Thus conclusions (i)–(v) are implied for the discontinuous cost c by the statements already proved for $\tilde{c}$, and the observation that the $\tilde{c}$-transform and c-transform of a function on $\operatorname{spt}\nu_o$ coincide on $\mathbf{R}^d\setminus\operatorname{spt}\nu_o$. □

Part III. Appendices

A. Legendre transforms and conjugate costs

This appendix begins by recalling basic properties of the Legendre transform for convex functions, and proceeds to deduce the corresponding properties of the conjugate $h^*(\mathbf{y})$

to a cost function $h(\mathbf{x})=l(|\mathbf{x}|)$ given by a concave function $l \geqslant 0$ of the distance. As usual, $\hat{\mathbf{x}}$ denotes the unit vector in the direction of $\mathbf{x} \in \mathbf{R}^d \setminus \{\mathbf{0}\}$.

The first theorem summarizes Theorems 12.2, 26.1, 26.3 and Corollaries 23.5.1, 25.5.1 of Rockafellar's text [33]; in his language, h is assumed to be *closed*—meaning lower semi-continuous—and *proper*—meaning finite somewhere, while the assertion of (iii) is that h^* be *essentially smooth.* By convention, we exclude $h := \infty$ from the class of convex functions. Implications for strictly convex costs $h(\mathbf{x})$ which grow superlinearly (H3) are summarized as a corollary.

THEOREM A.1 (Legendre transforms [33]). *Let* $h: \mathbf{R}^d \to \mathbf{R} \cup \{+\infty\}$ *be lower semi-continuous and convex, and define its Legendre transform* $h^*(\mathbf{y})$ *by* (10). *Then* h^* *satisfies the same assumptions as* h, *while*

(i) $(\mathbf{x}, \mathbf{y}) \in \partial_\cdot h$ *if and only if* $(\mathbf{y}, \mathbf{x}) \in \partial_\cdot h^*$;

(ii) *the Legendre transform of* h^* *is* h, *that is,* $h = h^{**}$;

(iii) *strict convexity of* h *implies* h^* *differentiable where it is subdifferentiable*;

(iv) *differentiability of* $h(\mathbf{x})$ *on an open set* $\Omega \subset \mathbf{R}^d$ *implies* $h \in C^1(\Omega)$.

COROLLARY A.2 (inverting the gradient of a strictly convex cost). *If* $h(\mathbf{x})$ *strictly convex satisfies* (H1) *and* (H3), *then its Legendre transform* $h^*(\mathbf{y})$ *will be continuously differentiable on* $\mathbf{R}^d$. *Moreover,* $\mathbf{x} = \nabla h^*(\mathbf{y})$ *if and only if* $\mathbf{y} \in \partial_\cdot h(\mathbf{x})$.

Proof. The function $h(\mathbf{x})$ was assumed to take non-negative real values throughout $\mathbf{R}^d$ and be strictly convex by (H1); it is therefore continuous and Theorem A.1 applies. Thus $h^*: \mathbf{R}^d \to \mathbf{R} \cup \{+\infty\}$ is convex, and claim (i) combines with claim (iii) to show equivalence of $(\mathbf{x}, \mathbf{y}) \in \partial_\cdot h$ with $\mathbf{x} = \nabla h^*(\mathbf{y})$. It remains only to prove $h^*(\mathbf{y}) < \infty$ on $\mathbf{R}^d$: then $h^*(\mathbf{y})$ will be subdifferentiable by its convexity, differentiable everywhere in view of claim (iii), and continuously differentiable by claim (iv).

Suppose not: let $h^*(\mathbf{y}) = \infty$. Then (10) yields a sequence $\mathbf{x}_n \in \mathbf{R}^d$ for which

$$0 \leqslant \langle \mathbf{x}_n, \mathbf{y} \rangle - h(\mathbf{x}_n) \tag{28}$$

increases without bound. Since $h \geqslant 0$, the $\mathbf{x}_n$ diverge to infinity, but a subsequence can be extracted whose direction vectors $\hat{\mathbf{x}}_n$ converge to a limit $\hat{\mathbf{x}}$ on the unit sphere. From (28), $\limsup h(\mathbf{x}_n)/|\mathbf{x}_n| \leqslant \langle \hat{\mathbf{x}}, \mathbf{y} \rangle < \infty$. This contradicts (H3) and completes the proof. □

To address concave functions of the distance, we restate Theorem A.1 after changing a sign, adding a remark about monotonicity.

THEOREM A.3 (concave Legendre transforms on the line). *Let* $l: \mathbf{R} \to \mathbf{R} \cup \{-\infty\}$ *be upper semi-continuous and concave. Define its dual function* $l^\circ(\xi) := -k^*(-\xi)$ *through the Legendre transform* (10) *of* $k := -l$. *Then* l° *satisfies the same hypotheses as* l, *and*

(i) $(\lambda,\xi)\in\partial^{\cdot} l$ *if and only if* $(\xi,\lambda)\in\partial^{\cdot} l^{\circ}$;

(ii) *the dual function of* l° *is* l, *that is*, $l=l^{\circ\circ}$;

(iii) *strict concavity of* l *implies* l° *differentiable where it is superdifferentiable*;

(iv) $l^{\circ}(\xi)$ *is non-decreasing if and only if* $l(\lambda)=-\infty$ *for all* $\lambda<0$.

Proof. From its definition, one verifies l° to be the *concave* Legendre transform of l:

$$l^{\circ}(\xi)=\inf_{\lambda\in\mathbf{R}}\xi\lambda-l(\lambda).$$

Then (i)–(iii) follow by a change of sign from the corresponding statements in Theorem A.1. Assertion (iv) is easily proved: to verify the *only if* implication suppose that $l(-\lambda)$ is finite at some $\lambda>0$; we shall show that l° decreases somewhere. Being concave, l must be superdifferentiable at $-\lambda$ (or some nearby point): $(-\lambda,\xi)\in\partial^{\cdot} l$. Then (i) implies that l° is finite at ξ and decreasing to its right: $l^{\circ}(\xi+\varepsilon)\leqslant l^{\circ}(\xi)-\lambda\varepsilon$.

To prove the converse, suppose that l° decreases somewhere. Then one has $(\xi,-\lambda)\in\partial^{\cdot} l^{\circ}$ for some $\xi\in\mathbf{R}$ and $\lambda>0$. Invoking (i) once again yields $(-\lambda,\xi)\in\partial^{\cdot} l$, from which one concludes finiteness of $l(-\lambda)$. □

An elementary lemma relates the superdifferential of $h(\mathbf{x}):=l(|\mathbf{x}|)$ to that of $l(\lambda)$.

LEMMA A.4 (the superdifferential of the cost). *Let* $l(\lambda)$ *be concave non-decreasing on* $\lambda\geqslant 0$ *and define* $h(\mathbf{x}):=l(|\mathbf{x}|)$ *on* $\mathbf{R}^d$. *Unless* h *is a constant*: $(\mathbf{x},\mathbf{y})\in\partial^{\cdot} h$ *if and only if* $(|\mathbf{x}|,|\mathbf{y}|)\in\partial^{\cdot} l$ *with* $\mathbf{y}=|\mathbf{y}|\hat{\mathbf{x}}$ *and* $\mathbf{x}\neq\mathbf{0}$.

Proof. Fix $\mathbf{x}\in\mathbf{R}^d\setminus\{\mathbf{0}\}$ and suppose $l(\lambda)$ admits ξ as a *superderivative* at $|\mathbf{x}|$: $(|\mathbf{x}|,\xi)\in\partial^{\cdot} l$. Since l is concave non-decreasing, $\xi\geqslant 0$, while for $\varepsilon\in\mathbf{R}$,

$$l(|\mathbf{x}|+\varepsilon)\leqslant l(|\mathbf{x}|)+\varepsilon\xi. \tag{29}$$

Now $h(\mathbf{x}+\mathbf{v})=l(|\mathbf{x}|+\varepsilon)$ where $\varepsilon=\langle\mathbf{v},\hat{\mathbf{x}}\rangle+o(|\mathbf{v}|)$, cf. (32). It follows immediately from (29) that h is superdifferentiable at $\mathbf{x}$ with $(\mathbf{x},\xi\hat{\mathbf{x}})\in\partial^{\cdot} h$. On the other hand, since $l(\lambda)$ is concave, non-decreasing and non-constant, h cannot be superdifferentiable at the origin: it grows linearly in every direction, or $h(\mathbf{0})=-\infty$.

Now let $(\mathbf{x},\mathbf{y})\in\partial^{\cdot} h$, so $\mathbf{x}\neq 0$, while for small $\mathbf{v}\in\mathbf{R}^d$,

$$h(\mathbf{x}+\mathbf{v})\leqslant h(\mathbf{x})+\langle\mathbf{v},\mathbf{y}\rangle+o(|\mathbf{v}|).$$

Spherical symmetry of h forces $\mathbf{y}$ to be parallel to $\mathbf{x}$: otherwise a slight rotation $\mathbf{x}+\mathbf{v}:=\mathbf{x}\cos\theta-\hat{\mathbf{z}}|\mathbf{x}|\sin\theta$ of $\mathbf{x}$ in the direction $\mathbf{z}:=\mathbf{y}-(\langle\mathbf{y},\hat{\mathbf{x}}\rangle)\hat{\mathbf{x}}$ would contradict $h(\mathbf{x}+\mathbf{v})=h(\mathbf{x})$ for θ sufficiently small. Moreover, taking $\mathbf{v}:=\varepsilon\hat{\mathbf{x}}$ yields (29), with $\xi:=\langle\hat{\mathbf{x}},\mathbf{y}\rangle+o(1)$. Thus $(|\mathbf{x}|,\langle\hat{\mathbf{x}},\mathbf{y}\rangle)\in\partial^{\cdot} l$, which concludes the lemma: $|\mathbf{y}|=\pm\langle\hat{\mathbf{x}},\mathbf{y}\rangle$ holds with a plus sign since l cannot decrease. □

COROLLARY A.5 (uniform superdifferentiability of the cost). *Let l and h be real-valued in the lemma above. Then $h(\mathbf{x})$ is superdifferentiable on $\mathbf{R}^d\setminus\{\mathbf{0}\}$. Moreover, for $\delta>0$, there is a real function $O_\delta(\lambda)$ tending to zero linearly with $|\lambda|$, such that $|\mathbf{x}|>\delta$, $\mathbf{y}\in\partial^{\cdot}h(\mathbf{x})$ and $\mathbf{v}\in\mathbf{R}^d$ imply*

$$h(\mathbf{x}+\mathbf{v})\leqslant h(\mathbf{x})+\langle\mathbf{v},\mathbf{y}\rangle+O_\delta(\mathbf{v}^2). \tag{30}$$

Proof. For $\lambda>0$, the concave function l admits a supergradient $\xi\in\partial^{\cdot}h(\lambda)$: for example, take its left derivative $\xi=l'(\lambda^-)$. If $|\mathbf{x}|=\lambda$, the lemma implies $(\mathbf{x},\xi\hat{\mathbf{x}})\in\partial^{\cdot}h$, so $h(\mathbf{x})$ is superdifferentiable at $\mathbf{x}$.

Now suppose $(\mathbf{x},\mathbf{y})\in\partial^{\cdot}h$. The opposite implication of the lemma yields $\mathbf{y}=\xi\hat{\mathbf{x}}$ with $(|\mathbf{x}|,\xi)\in\partial^{\cdot}l$ so (29) holds. Moreover, $\xi\geqslant0$. If $\mathbf{v}\in\mathbf{R}^d$, then $h(\mathbf{x}+\mathbf{v})=l(|\mathbf{x}|+\varepsilon)$ where

$$\varepsilon:=\sqrt{\mathbf{x}^2+2\langle\mathbf{x},\mathbf{v}\rangle+\mathbf{v}^2}-|\mathbf{x}| \tag{31}$$

$$\leqslant\langle\hat{\mathbf{x}},\mathbf{v}\rangle+\mathbf{v}^2/|2\mathbf{x}|; \tag{32}$$

the inequality follows from $\sqrt{1+\lambda}\leqslant1+\frac{1}{2}\lambda$. By concavity of l, its left derivative $l'(\lambda^-)$ is a non-increasing function of λ. Assume $|\mathbf{x}|>\delta$ so that $\xi\leqslant l'(\delta^-)$. Together with (29) and (32), this assumption yields (30):

$$h(\mathbf{x}+\mathbf{v})\leqslant h(\mathbf{x})+\langle\xi\hat{\mathbf{x}},\mathbf{v}\rangle+\mathbf{v}^2l'(\delta^-)/2\delta.$$ □

PROPOSITION A.6 (the conjugate cost). *Let $h(\mathbf{x}):=l(|\mathbf{x}|)$ be continuous on $\mathbf{R}^d$ with $l(\lambda)$ strictly concave increasing on $\lambda\geqslant0$. Define the dual function $h^*:\mathbf{R}^d\to\mathbf{R}\cup\{-\infty\}$ via (11). For some $R\geqslant0$,*

(i) *$h^*(\mathbf{y})$ is continuously differentiable on $|\mathbf{y}|>R$ while $h^*=-\infty$ on $|\mathbf{y}|<R$;*

(ii) *$(\mathbf{y},\mathbf{x})\in\partial^{\cdot}h^*$ with $\mathbf{x}\neq\mathbf{0}$ if and only if $(\mathbf{x},\mathbf{y})\in\partial^{\cdot}h$ with $\mathbf{y}\neq\mathbf{0}$;*

(iii) *if $(\mathbf{y},\mathbf{x})\in\partial^{\cdot}h^*$ then $\mathbf{x}=\nabla h^*(\mathbf{y})$;*

(iv) *if $(\mathbf{y},\mathbf{0})\in\partial^{\cdot}h^*$ then $|\mathbf{y}|\geqslant\sup_{\lambda>0}l'(\lambda)$.*

Proof. Extend l to $\mathbf{R}$ by defining $l(\lambda)=-\infty$ for $\lambda<0$. If l° is defined as in Theorem A.3 then $h^*(\mathbf{y})=l^\circ(|\mathbf{y}|)$; moreover, $l^\circ:\mathbf{R}\to\mathbf{R}\cup\{-\infty\}$ is itself upper semi-continuous and concave non-decreasing with $l^\circ(\xi)=-\infty$ where $\xi<0$.

(ii) Let $(\mathbf{x},\mathbf{y})\in\partial^{\cdot}h$ with $\mathbf{y}\neq0$. Then h cannot be constant. Lemma A.4 yields $\mathbf{x}\neq0$, but $\mathbf{x}=|\mathbf{x}|\hat{\mathbf{y}}$ and $(|\mathbf{x}|,|\mathbf{y}|)\in\partial^{\cdot}l$. Then Theorem A.3 (i) implies $(|\mathbf{y}|,|\mathbf{x}|)\in\partial^{\cdot}l^\circ$. Since $|\mathbf{x}|$ and $|\mathbf{y}|$ do not vanish, h^* cannot be constant and the the reverse implication of Lemma A.4 yields $(\mathbf{y},\mathbf{x})\in\partial^{\cdot}h^*$. This proves the *if* part of the claim. Since we have not used *strict* concavity of $l(\lambda)$, the *only if* statement follows immediately from the duality between l and l° expressed in Theorem A.3 (ii).

(i) Since $l^\circ(\lambda)$ is non-decreasing and not identically $-\infty$, there is some $R\in\mathbf{R}$ such that $l^\circ(\lambda)=-\infty$ for $\lambda<R$ while l° is finite-valued for $\lambda>R$. By concavity, l° is continuous and superdifferentiable on $\lambda>R$; Theorem A.3 (iii) shows that l° is differentiable where superdifferentiable, which combines with Theorem A.1 (iv) to yield continuous differentiability on $\lambda>R$. Thus $h^*(\mathbf{y})$ is continuously differentiable on $|\mathbf{y}|>R$ while $h^*(\mathbf{y})=\infty$ on $|\mathbf{y}|<R$.

(iii) As has just been noted, l° is differentiable where superdifferentiable. The same holds true for $h^*(\mathbf{y})=l^\circ(|\mathbf{y}|)$ in view of Lemma A.4.

(iv) Finally, assume $(\mathbf{y},\mathbf{0})\in\partial^\cdot h^*$. If h^* is non-constant, Lemma A.4 yields $(|\mathbf{y}|,0)\in\partial^\cdot l^\circ$—a result which is obvious when h^* is constant. Thus $(0,|\mathbf{y}|)\in\partial^\cdot l$ by Theorem A.3 (i). Since the derivative of l cannot increase, $|\mathbf{y}|\geqslant l'(\lambda)$ whenever l is differentiable at $\lambda>0$. □

B. Examples of c-concave potentials

In this appendix we do nothing more than present a few examples of c-concave potentials. For strictly convex costs or concave functions $l\geqslant 0$ of the distance, they verify the claims made in §1 about the respective optimality of translations and dilations, or reflections, of $\mathbf{R}^d$.

LEMMA B.1 (examples of c-concave potentials). *Let $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ convex satisfy* (H3)–(H4). *Fix $\mathbf{z}\in\mathbf{R}^d$, $r\in\mathbf{R}$ and a c-concave function $\phi(\mathbf{x})$ on $\mathbf{R}^d$. Then the following functions ψ are also c-concave on $\mathbf{R}^d$:*

(i) $\psi(\mathbf{x})=h(\mathbf{x})$;

(ii) *the infimum ψ of a family of c-concave functions (except $\psi:=-\infty$);*

(iii) *the shifted translate $\psi(\mathbf{x}):=\phi(\mathbf{x}-\mathbf{z})+r$;*

(iv) *the dilation $\psi(\mathbf{x}):=r\phi(\mathbf{x}/r)$ by a factor $r\geqslant 1$;*

(v) *the linear function $\psi(\mathbf{x}):=\langle\mathbf{x},\mathbf{z}\rangle$;*

(vi) *any (upper semi-)continuous concave function $\psi:\mathbf{R}^d\to\mathbf{R}\cup\{-\infty\}$.*

Proof. (i)–(iii) The first three claims are apparent from the definition (7) of c-concavity; they require no special features of the cost function $h(\mathbf{x})$.

(iv) First suppose $\phi=h$ and let $\lambda:=r^{-1}$. Then $0<\lambda\leqslant 1$, so for $\mathbf{x},\mathbf{y}\in\mathbf{R}^d$ convexity of the cost h implies

$$h(\lambda\mathbf{x})\leqslant\lambda h(\mathbf{x}-(1-\lambda)\mathbf{y})+(1-\lambda)h(\lambda\mathbf{y}).$$

Equality holds if $\mathbf{x}=\mathbf{y}$. Thus

$$h(\lambda\mathbf{x})/\lambda=\inf_{y\in\mathbf{R}^d}c(\mathbf{x},(1-\lambda)\mathbf{y})+(1-\lambda)\lambda^{-1}h(\lambda\mathbf{y}) \tag{33}$$

is manifestly c-concave. For a general c-concave ϕ, one obtains

$$\phi(\lambda\mathbf{x})/\lambda = \inf_{(\lambda\mathbf{y},\alpha)\in\mathcal{A}} \lambda^{-1}h(\lambda(\mathbf{x}-\mathbf{y}))+\lambda^{-1}\alpha \tag{34}$$

from (7). The c-concavity of $\phi(\lambda\mathbf{x})/\lambda$ follows from (ii)–(iii) and (33).

(v) In view of (H3), the continuous function $h(\mathbf{x})-\langle\mathbf{x},\mathbf{z}\rangle$ assumes its minimum at some $\mathbf{x}=\mathbf{p}$ in $\mathbf{R}^d$:

$$h(\mathbf{p})-\langle\mathbf{p},\mathbf{z}\rangle \leqslant h(\mathbf{x}-\mathbf{y})-\langle\mathbf{x}-\mathbf{y},\mathbf{z}\rangle$$

for $\mathbf{x},\mathbf{y}\in\mathbf{R}^d$, with equality when $\mathbf{x}-\mathbf{y}=\mathbf{p}$. Thus

$$\langle\mathbf{x},\mathbf{z}\rangle = \inf_{\mathbf{y}\in\mathbf{R}^d} c(\mathbf{x},\mathbf{y})+\langle\mathbf{y}+\mathbf{p},\mathbf{z}\rangle-h(\mathbf{p}); \tag{35}$$

its c-concavity (7) as a function of $\mathbf{x}$ is manifest.

(vi) Any upper semi-continuous concave function ψ can be represented as an infimum of affine functions (as in Theorem A.1 (ii) for example); its c-concavity therefore follows from (ii)–(iii) and (v). □

For strictly convex costs, this lemma was invoked to check optimality of translations and dilations on $\mathbf{R}^d$. In this context claim (vi) is equivalent to an observation of Smith and Knott [41]; see also Rüschendorf [36], [37].

To verify optimality for the reflections of Example 1.5 when the cost is a strictly concave function $l\geqslant 0$ of the distance, our argument will be less direct. It relies on a simple observation about the transportation problem on the line [28]: if the full mass of $\mu\in\mathcal{P}(\mathbf{R})$ lies to the left of $\operatorname{spt}\nu$, then the optimal map of μ onto ν will be orientation-reversing. Indeed, it will be the unique non-increasing map $s\colon\mathbf{R}\to\mathbf{R}$ pushing μ forward to $\nu\in\mathcal{P}(\mathbf{R})$, which exists whenever μ is free from point masses. Taking Lebesgue measure on $[0,1]$ for μ, and its image under the inversion $s(x)=1/x$ for ν, one concludes that the map s is optimal between μ and $s_{\#}\mu$. (Better yet, replace Lebesgue measure by $d\mu(x):=\frac{1}{2}x\,dx$ to avoid infinite transport cost.) In view of Theorem 1.4 this means that s can be expressed in the form $s(x)=x-(l')^{-1}(\phi'(x))$, where ϕ is the c-transform of a function on $[1,\infty)$. (Here $l(\lambda):=l(|\lambda|)$ for $\lambda<0$.) Defining $\psi(\mathbf{x}):=\phi(|\mathbf{x}|)$ on $\mathbf{R}^d$, it follows that ψ is the c-transform of a function on the complement $\mathbf{R}^d\setminus B$ of the unit ball. Invoking Theorem 1.4 again with $h(\mathbf{x}):=l(|\mathbf{x}|)$ establishes Example 1.5: $\mathbf{s}(\mathbf{x}):=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))=\mathbf{x}/|\mathbf{x}|^2$ must be the optimal map between any measure μ and its spherical reflection $\mathbf{s}_{\#}\mu$ provided $\operatorname{spt}\mu=B$. If $\operatorname{spt}\mu\subset B$, a slight refinement is required: the optimal map will still be given by the c-transform of a function on $\mathbf{R}^d\setminus B$, and coincides with $\mathbf{s}(\mathbf{x})=\mathbf{x}/|\mathbf{x}|^2$ in view of Theorem 6.3.

The same analysis adapts easily to the case of reflection through a hyperplane instead of a sphere. Instead of Lebesgue measure on the unit interval, one considers the reflection

$s(x):=-x$ of some measure μ which has a first moment, and is given by a non-vanishing density throughout $\operatorname{spt}\mu=(-\infty,0]$.

C. Regularity of c-concave potentials

This appendix explores the extent to which a c-concave potential ψ inherits structure and smoothness from a convex cost $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$. Its primary purpose is to assemble the necessary machinery to prove Theorem 3.3, which was central to the analysis in §§ 3 and 4. When $h\in C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$ the potential will be locally semi-concave, and therefore share all the regularity enjoyed by concave functions—e.g. two derivatives almost everywhere—as a consequence. Otherwise ψ will be locally Lipschitz where finite. The proof is divided into three main propositions; it is here that the technical restrictions (H2)–(H4) on convex costs play a role.

We begin by recalling a standard estimate showing that the c-transform $\psi:\mathbf{R}^d\to\mathbf{R}\cup\{-\infty\}$ of any function on a bounded set V is locally Lipschitz throughout $\mathbf{R}^d$.

LEMMA C.1 (locally Lipschitz). *Suppose $c:\mathbf{R}^d\times\mathbf{R}^d\to\mathbf{R}$ to be locally Lipschitz. Then the c-transform ψ of any function on a bounded set $V\subset\mathbf{R}^d$ will be locally Lipschitz on $\mathbf{R}^d$.*

Remark on proof. Fixing any ball $U\subset\mathbf{R}^d$, $c(\mathbf{x},\mathbf{y})$ satisfies a global Lipschitz bound on $U\times V$. Thus $\psi(\mathbf{x})$ is an infimum (9) of functions $c(\mathbf{x},\mathbf{y})-\phi(\mathbf{y})$ indexed by $\mathbf{y}\in V$ and satisfying a uniform Lipschitz bound on U. By assumption ψ is finite somewhere, and it is then well known that ψ satisfies the same Lipschitz condition on U. □

When the cost is a derivative smoother—$c(\mathbf{x},\mathbf{y})$ in $C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d\times\mathbf{R}^d)$—a more novel estimate yields local semi-concavity of ψ. For notational simplicity only, the proof here is restricted to costs taking the form $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$.

PROPOSITION C.2 (locally semi-concave). *Let $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ with $h\in C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$. Then the c-transform ψ of any function on a bounded set $V\subset\mathbf{R}^d$ will be locally semi-concave on $\mathbf{R}^d$.*

Proof. We first check that the cost $h(\mathbf{x})$ itself is semi-concave on any open ball $\Omega\subset\mathbf{R}^d$: that is, for $\lambda<\infty$ sufficiently large, the function $h_\lambda(\mathbf{x}):=h(\mathbf{x})-\lambda\mathbf{x}^2$ should be concave on Ω. To see that this is true, let 2λ be the Lipschitz bound for ∇h on Ω. Since $\mathbf{x},\mathbf{y}\in\Omega$ imply $|\nabla h(\mathbf{x})-\nabla h(\mathbf{y})|\leqslant 2\lambda|\mathbf{x}-\mathbf{y}|$, one obtains

$$0\geqslant\langle\nabla h(\mathbf{x})-\nabla h(\mathbf{y}),\mathbf{x}-\mathbf{y}\rangle-2\lambda|\mathbf{x}-\mathbf{y}|^2=\langle\nabla h_\lambda(\mathbf{x})-\nabla h_\lambda(\mathbf{y}),\mathbf{x}-\mathbf{y}\rangle.$$

This monotonicity relation and differentiability of $h_\lambda(\mathbf{x})$ imply that $h_\lambda(\mathbf{x})$ is concave on Ω [39, Theorem 1.5.9].

Now, consider the c-transform ψ of a function on V: it will be of the form (7) with $\mathcal{A}\subset V\times\mathbf{R}$. Let U be an open ball around $\mathbf{x}\in\mathbf{R}^d$, and let Ω be large enough to contain $U-V:=\{\mathbf{x}-\mathbf{y}\,|\,\mathbf{x}\in U, \mathbf{y}\in V\}$. Taking λ large enough to ensure h_λ concave on Ω, one has

$$\psi(\mathbf{x})-\lambda\mathbf{x}^2=\inf_{(\mathbf{y},\xi)\in\mathcal{A}} h_\lambda(\mathbf{x}-\mathbf{y})-2\lambda\langle\mathbf{x},\mathbf{y}\rangle+\lambda\mathbf{y}^2+\xi.$$

The infimum is over a family of concave functions of $\mathbf{x}\in U$, whence $\psi(\mathbf{x})-\frac{1}{2}\lambda\mathbf{x}^2$ itself is concave on U. Thus local semi-concavity of ψ is established at arbitrary $\mathbf{x}\in\mathbf{R}^d$. □

As the first lemma shows, the c-transform ψ of a function on a bounded set $V\subset\mathbf{R}^d$ will be finite throughout $\mathbf{R}^d$. Thus the smoothness results alone imply Theorem 3.3 in this case. The remainder of this appendix exploits (H2)–(H3) to extend the theorem to the c-concave potentials arising when $\operatorname{spt}\nu$ is unbounded. The intuitions derive from Figure 1.

Recall that ψ is *locally bounded* at $\mathbf{p}\in\mathbf{R}^d$ if there exists $R<\infty$ such that $|\psi(\mathbf{x})|\leqslant R$ holds on a neighbourhood of $\mathbf{p}$.

PROPOSITION C.3 (locally bounded on a convex domain). *Let* $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ *continuous satisfy* (H2) *and fix a c-concave* ψ *on* $\mathbf{R}^d$. *Define the convex hull* K *of the set where* ψ *is finite. Then* ψ *is locally bounded throughout the interior of* K.

Proof. Suppose that ψ fails to be locally bounded at $\mathbf{p}\in\mathbf{R}^d$. We shall show that $\mathbf{p}$ lies on the boundary of an open half-space $H_{\hat{\mathbf{z}}}(\mathbf{p}):=\{\mathbf{x}\,|\,\langle\hat{\mathbf{z}},\mathbf{x}-\mathbf{p}\rangle>0\}$ in which $\psi(\mathbf{x})=-\infty$. Then $\{\mathbf{x}\,|\,\psi(\mathbf{x})>-\infty\}$ will be disjoint from $H_{\hat{\mathbf{z}}}(\mathbf{p})$, so its convex hull K will be disjoint as well. Since $\mathbf{p}$ cannot lie in the interior of K, the proposition will have been established.

To prove $\psi=-\infty$ on some open half-space, recall that any c-concave ψ must be finite at some $\mathbf{v}\in\mathbf{R}^d$. Thus $\mathcal{A}\subset\mathbf{R}^d\times\mathbf{R}$ is non-empty in (7), and it follows immediately that ψ is bounded above by a shifted translate of the continuous function $h(\mathbf{x})$ on $\mathbf{R}^d$. On the other hand, ψ can certainly fail to be bounded below in each neighbourhood of $\mathbf{p}$. In this case there is a sequence $\mathbf{p}_n\to\mathbf{p}$ with $\psi(\mathbf{p}_n)<-n$. Recalling the definition (7) of $\psi(\mathbf{p}_n)$, there is a sequence $(\mathbf{y}_n,\lambda_n)\in\mathcal{A}$ such that

$$c(\mathbf{p}_n,\mathbf{y}_n)+\lambda_n\leqslant -n. \tag{36}$$

Applied at $\mathbf{v}$, where ψ is finite, the same definition couples with (36) to yield

$$\psi(\mathbf{v})\leqslant c(\mathbf{v},\mathbf{y}_n)-c(\mathbf{p}_n,\mathbf{y}_n)-n.$$

Since $c(\mathbf{x},\mathbf{y})$ is continuous and $\mathbf{p}_n\to\mathbf{p}$, certainly $|\mathbf{y}_n|\to\infty$ to avoid contradicting $\psi(\mathbf{v})>-\infty$. For each n, choose the height r_n and direction $\hat{\mathbf{z}}_n$ of the largest cone (8) with vertex $\mathbf{p}_n-\mathbf{y}_n$ such that $K(r_n,\pi/(1+r_n^{-1}),\hat{\mathbf{z}}_n,\mathbf{p}_n-\mathbf{y}_n)\subset\{\mathbf{x}\,|\,h(\mathbf{x})\leqslant h(\mathbf{p}_n-\mathbf{y}_n)\}$; we allow $0\leqslant r_n\leqslant\infty$. Since $|\mathbf{p}_n-\mathbf{y}_n|$ diverges with n, the curvature condition (H2) on level sets of h implies $r_n\to\infty$ with n. Extracting a subsequence if necessary ensures that the unit vectors $\hat{\mathbf{z}}_n$ converge to a limit $\hat{\mathbf{z}}\in\mathbf{R}^d$ on the unit sphere.

Now, suppose that $\mathbf{x}\in H_{\hat{\mathbf{z}}}(\mathbf{p})$ so that $\langle\hat{\mathbf{z}},\mathbf{x}-\mathbf{p}\rangle>0$. Taking n sufficiently large ensures

$$|\mathbf{x}-\mathbf{p}_n|\cos\left(\frac{\pi}{2}\cdot\frac{1}{1+1/r_n}\right)<\langle\hat{\mathbf{z}}_n,\mathbf{x}-\mathbf{p}_n\rangle<r_n,$$

since the left and right bounds have zero and infinity as their limits. Thus

$$\mathbf{x}\in K\left(r_n,\frac{\pi}{1+1/r_n},\hat{\mathbf{z}}_n,\mathbf{p}_n\right)$$

follows from (8), and $c(\mathbf{x},\mathbf{y}_n)\leqslant c(\mathbf{p}_n,\mathbf{y}_n)$ from our construction. Combining (7) with (36) yields

$$\psi(\mathbf{x})\leqslant c(\mathbf{x},\mathbf{y}_n)+\lambda_n\leqslant -n.$$

Since n can be arbitrarily large, $\psi(\mathbf{x})=-\infty$. Because $\mathbf{x}\in H_{\hat{\mathbf{z}}}(\mathbf{p})$ was arbitrary, the proposition is proved. □

PROPOSITION C.4 (local boundedness of c-superdifferentials). *Let $c(\mathbf{x},\mathbf{y})=h(\mathbf{x}-\mathbf{y})$ convex satisfy* (H3)–(H4), *and fix a c-concave ψ on $\mathbf{R}^d$. If ψ is bounded on some neighbourhood of a compact, non-empty set $U\subset\mathbf{R}^d$, then $\partial^c\psi(U)$ is bounded and non-empty.*

Proof. The proposition consists of two claims to be established in parallel: c-superdifferentiability of ψ on U and boundedness of $\partial^c\psi(U)$. Since ψ is bounded on a neighbourhood of the compact set U, there is a $0<\delta<1$ and $R<\infty$ such that $|\psi(\mathbf{x})|<R$ whenever $|\mathbf{x}-\mathbf{p}|<\delta$ and $\mathbf{p}\in U$. Suppose that a sequence $(\mathbf{y}_n,\lambda_n)\in\mathbf{R}^d\times\mathbf{R}$ satisfies

$$\psi(\mathbf{x})\leqslant c(\mathbf{x},\mathbf{y}_n)+\lambda_n, \tag{37}$$

for all $\mathbf{x}\in\mathbf{R}^d$, while

$$c(\mathbf{x}_n,\mathbf{y}_n)+\lambda_n<R \tag{38}$$

holds for each n and some $\mathbf{x}_n\in U$. The last paragraph shows that the $|\mathbf{y}_n|$ are bounded; for the moment, we assume this bound to complete the proof.

Fix $\mathbf{p}\in U$. By the c-concavity (7) of ψ, there is a sequence $(\mathbf{y}_n,\lambda_n)$ such that $c(\mathbf{p},\mathbf{y}_n)+\lambda_n$ converges to $\psi(\mathbf{p})<R$; it may be taken to satisfy (37)–(38) with $\mathbf{x}_n:=\mathbf{p}$. Since the $\mathbf{y}_n$ are bounded, a limit point $\mathbf{y}_n\to\mathbf{y}$ may be extracted after replacing the $(\mathbf{y}_n,\lambda_n)$ with a subsequence. The λ_n will converge to $\lambda:=\psi(\mathbf{p})-c(\mathbf{p}-\mathbf{y})$. The large n limit of (37) shows that $\mathbf{y}$ is a c-supergradient of ψ at $\mathbf{p}\in U$. Thus $\partial^c\psi(U)\supset\{\mathbf{y}\}$ cannot

be empty. On the other hand, any sequence $\mathbf{y}_n \in \partial^c\psi(\mathbf{x}_n)$ for which $\mathbf{x}_n \in U$ satisfies (37)–(38) with $\lambda_n = \psi(\mathbf{x}_n) - c(\mathbf{x}_n, \mathbf{y}_n)$ by definition. The bound on $|\mathbf{y}_n|$ therefore shows that $\partial^c\psi(U)$ must be bounded.

It remains to show that (37)–(38) imply a bound on the $\mathbf{y}_n \in \mathbf{R}^d$. If not, some subsequence $|\mathbf{y}_n| \to \infty$ escapes to infinity; setting $\mathbf{v}_n := \mathbf{x}_n - \mathbf{y}_n$, we may assume $|\mathbf{v}_n| > 1$ since all the $\mathbf{x}_n$ lie in a bounded set U. Use the $\delta > 0$ above to define a sequence $\xi_n := 1 - \delta|\mathbf{v}_n|^{-1}$ which converges to 1. Evaluated at $\mathbf{x} = \mathbf{x}_n + (\xi_n - 1)\mathbf{v}_n$, (37)–(38) combine with the lower bound $-R < \psi(\mathbf{x})$ to yield

$$2R \geqslant h(\mathbf{v}_n) - h(\xi_n \mathbf{v}_n).$$

Since h is convex, this difference may be bounded using a subgradient $\mathbf{z}_n \in \partial_\cdot h(\xi_n \mathbf{v}_n)$:

$$2R \geqslant \langle (1-\xi_n)\mathbf{v}_n, \mathbf{z}_n \rangle \tag{39}$$

$$= \delta \langle \mathbf{v}_n / |\mathbf{v}_n|, \mathbf{z}_n \rangle. \tag{40}$$

On the other hand, being a subgradient also implies

$$h(\mathbf{0}) \geqslant h(\xi_n \mathbf{v}_n) + \langle \mathbf{z}_n, \mathbf{0} - \xi_n \mathbf{v}_n \rangle.$$

Since the $\xi_n > 1 - \delta > 0$ are bounded away from zero and $|\mathbf{v}_n| \to \infty$, dividing by $|\xi_n \mathbf{v}_n| \to \infty$ yields $\liminf \langle \mathbf{z}_n, \mathbf{v}_n \rangle / |\mathbf{v}_n| \geqslant \liminf h(\xi \mathbf{v}_n)/|\xi \mathbf{v}_n|$. Assumption (H3) ensures that both of these limits diverge, yielding a contradiction with(40). The only conclusion must be that the $\mathbf{y}_n$ were bounded. □

COROLLARY C.5. *Let $c(\mathbf{x}, \mathbf{y}) = h(\mathbf{x} - \mathbf{y})$ convex satisfy* (H3)–(H4), *and ψ be c-concave on $\mathbf{R}^d$. Then ψ is locally Lipschitz wherever it is locally bounded. Moreover, if $h \in C^{1,1}_{\mathrm{loc}}(\mathbf{R}^d)$, then ψ is locally semi-concave wherever it is locally bounded.*

Proof. If ψ is locally bounded at $\mathbf{p} \in \mathbf{R}^d$, it is possible to choose a compact neighbourhood U of $\mathbf{p}$ with ψ bounded in a neighbourhood of U. Since a single point $\mathbf{u}$ forms a bounded set by itself, Proposition C.4 implies c-superdifferentiability of ψ at $\mathbf{u} \in U$. It follows that equality holds in

$$\psi(\mathbf{x}) \leqslant \inf_{\substack{\mathbf{y} \in \partial^c \psi(\mathbf{u}) \\ \mathbf{u} \in U}} c(\mathbf{x}, \mathbf{y}) - c(\mathbf{u}, \mathbf{y}) + \psi(\mathbf{u}) \tag{41}$$

for all $\mathbf{x} \in U$. This infimum is manifestly the c-transform of a function on $\partial^c\psi(U)$. Proposition C.4 implies that $\partial^c\psi(U)$ is bounded. Since any convex function h will be locally Lipschitz, Lemma C.1 implies the infimum in (41) to be a locally Lipschitz function of

$\mathbf{x}\in\mathbf{R}^d$. If $h\in C^{1,1}_{\rm loc}(\mathbf{R}^d)$ then Proposition C.2 implies this infimum to be locally semi-concave on $\mathbf{R}^d$. Since ψ coincides with this infimum throughout U, it is locally Lipschitz or semi-concave at $\mathbf{p}$, according to the smoothness of h. □

As a proposition without proof, we summarize the differentiability properties of semi-concave potentials. Such potentials differ from concave functions locally by something smooth (Definition 3.2), so they immediately inherit all the (i) continuity [39, §1.5], (ii) differentiability [43, Theorem 1] or [3], (iii) continuous differentiability [33, §24.5], and (iv) second differentiability [39, notes to §1.5] of concave functions. Measurability of $\nabla\psi$ follows from continuity of ψ as in Proposition 3.4 (i).

PROPOSITION C.6 (differentiability of semi-concave potentials). *Let $\psi:\Omega\to\mathbf{R}$ be locally semi-concave on an open set $\Omega\subset\mathbf{R}^d$. Then*

(i) *ψ is continuous on Ω, so $\nabla\psi$ is a Borel map on the set where it can be defined;*

(ii) *differentiability of ψ fails only on a rectifiable set of dimension $d-1$;*

(iii) *if $(\mathbf{x}_n,\mathbf{y}_n)\in\partial^{\cdot}\psi$ is a sequence with $\mathbf{x}_n\to\mathbf{p}$ in Ω, then the $\mathbf{y}_n$ accumulate on $\partial^{\cdot}\psi(\mathbf{p})$; in particular, if ψ is differentiable at $\mathbf{x}$ then $\mathbf{y}_n\to\nabla\psi(\mathbf{x})$;*

(iv) *the map $\nabla\psi$ is differentiable almost everywhere on Ω in the sense of Aleksandrov* [39, §1.5].

Finally, to close the circle of ideas, a companion lemma to Lemma 3.1 is provided. It allows us to derive a c-differential continuity result for c-transforms which facilitates the uniqueness proof.

LEMMA C.7 (relating c-differentials to superdifferentials). *Let both h and ψ map $\mathbf{R}^d$ to $\mathbf{R}$ while $c(\mathbf{x},\mathbf{y}):=h(\mathbf{x}-\mathbf{y})$. If $(\mathbf{x},\mathbf{y})\in\partial^c\psi$ then $\partial^{\cdot}h(\mathbf{x}-\mathbf{y})\subset\partial^{\cdot}\psi(\mathbf{x})$.*

Proof. Let $(\mathbf{x},\mathbf{y})\in\partial^c\psi$. If h fails to be superdifferentiable at $\mathbf{x}-\mathbf{y}$, there is nothing to prove. Therefore, assume $\mathbf{z}\in\partial^{\cdot}h(\mathbf{x}-\mathbf{y})$. Combined with c-superdifferentiability of ψ this yields

$$\begin{aligned}\psi(\mathbf{x}+\mathbf{v}) &\leqslant \psi(\mathbf{x})+h(\mathbf{x}+\mathbf{v}-\mathbf{y})-h(\mathbf{x}-\mathbf{y})\\ &\leqslant \psi(\mathbf{x})+\langle\mathbf{v},\mathbf{z}\rangle+o(|\mathbf{v}|)\end{aligned}$$

for small $\mathbf{v}\in\mathbf{R}^d$. Thus $\mathbf{z}\in\partial^{\cdot}\psi(\mathbf{x})$. □

COROLLARY C.8 (c-differential continuity). *Fix $c(\mathbf{x},\mathbf{y}):=h(\mathbf{x}-\mathbf{y}):=l(|\mathbf{x}-\mathbf{y}|)$ continuous with $l(\lambda)\geqslant 0$ strictly concave, and let $\psi:\mathbf{R}^d\to\mathbf{R}$ be locally semi-concave at $p\in\mathbf{R}^d$. Assume that $\mathbf{s}(\mathbf{x}):=\mathbf{x}-\nabla h^*(\nabla\psi(\mathbf{x}))$ is defined at $\mathbf{p}$. Then $(\mathbf{x}_n,\mathbf{y}_n)\in\partial^c_\circ\psi$ with $\mathbf{x}_n\to\mathbf{p}$ implies $\mathbf{y}_n\to\mathbf{s}(\mathbf{p})$.*

Proof. Let $(\mathbf{x}_n,\mathbf{y}_n)\in\partial^c_\circ\psi$ with $\mathbf{x}_n\to\mathbf{p}$. Since $\mathbf{x}_n\neq\mathbf{y}_n$, Corollary A.5 provides super-gradients $\mathbf{w}_n\in\partial^{\cdot}h(\mathbf{x}_n-\mathbf{y}_n)$, which by Lemma C.7 also lie in $\mathbf{w}_n\in\partial^{\cdot}\psi(\mathbf{x}_n)$. Since ψ

was assumed to be differentiable at $\mathbf{p}=\lim \mathbf{x}_n$ and locally semi-concave, Proposition C.6 yields $\mathbf{w}_n \to \nabla\psi(\mathbf{p})$. On the other hand, Proposition A.6 provides a conjugate cost h^*, continuously differentiable at $\nabla\psi(\mathbf{p})$, for which $\mathbf{x}_n - \mathbf{y}_n = \nabla h^*(\mathbf{w}_n)$. Thus the $\mathbf{y}_n$ converge to $\mathbf{p} - \nabla h^*(\nabla\psi(\mathbf{p})) = \mathbf{s}(\mathbf{p})$. □

References

[1] ABDELLAOUI, T., Détermination d'un couple optimal du problème de Monge–Kantorovitch. *C. R. Acad. Sci. Paris Sér. I Math.*, 319 (1994), 981–984.

[2] ABDELLAOUI, T. & HEINICH, H., Sur la distance de deux lois dans le cas vectoriel. *C. R. Acad. Sci. Paris Sér. I Math.*, 319 (1994), 397–400.

[3] ALBERTI, G., On the structure of singular sets of convex functions. *Calc. Var. Partial Differential Equations*, 2 (1994), 17–27.

[4] ALEKSANDROV, A. D., Existence and uniqueness of a convex surface with a given integral curvature. *C. R. (Doklady) Acad. Sci. URSS (N. S.)*, 35 (1942), 131–134.

[5] APPELL, P., Mémoire sur les déblais et les remblais des systèmes continues ou discontinues. *Mémoires présentés par divers Savants à l'Académie des Sciences de l'Institut de France, Paris, I. N.*, 29 (1887), 1–208.

[6] BALDER, E. J., An extension of duality-stability relations to non-convex optimization problems. *SIAM J. Control Optim.*, 15 (1977), 329–343.

[7] BRENIER, Y., Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Paris Sér. I Math.*, 305 (1987), 805–808.

[8] — Polar factorization and monotone rearrangement of vector-valued functions. *Comm. Pure Appl. Math.*, 44 (1991), 375–417.

[9] CAFFARELLI, L., The regularity of mappings with a convex potential. *J. Amer. Math. Soc.*, 5 (1992), 99–104.

[10] — Allocation maps with general cost functions, in *Partial Differential Equations and Applications* (P. Marcellini, G. Talenti and E. Vesintini, eds.), pp. 29–35. Lecture Notes in Pure and Appl. Math., 177. Dekker, New York, 1996.

[11] CUESTA-ALBERTOS, J. A. & MATRÁN, C., Notes on the Wasserstein metric in Hilbert spaces. *Ann. Probab.*, 17 (1989), 1264–1276.

[12] CUESTA-ALBERTOS, J. A., MATRÁN, C. & TUERO-DÍAZ, A., Properties of the optimal maps for the L^2-Monge–Kantorovich transportation problem. Preprint.

[13] CUESTA-ALBERTOS, J. A., RÜSCHENDORF, L. & TUERO-DÍAZ, A., Optimal coupling of multivariate distributions and stochastic processes. *J. Multivariate Anal.*, 46 (1993), 335–361.

[14] CUESTA-ALBERTOS, J. A. & TUERO-DÍAZ, A., A characterization for the solution of the Monge–Kantorovich mass transference problem. *Statist. Probab. Lett.*, 16 (1993), 147–152.

[15] DALL'AGLIO, G., Sugli estremi dei momenti delle funzioni di ripartizione-doppia. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)*, 10 (1956), 35–74.

[16] DARBOUX, G., Prix Bordin (géométrie). *C. R. Acad. Sci. Paris*, 101 (1885), 1312–1316.

[17] DOUGLIS, A., Solutions in the large for multi-dimensional non-linear partial differential equations of first order. *Ann. Inst. Fourier (Grenoble)*, 15:2 (1965), 1–35.

[18] EVANS, L. C. & GANGBO, W., Differential equations methods for the Monge–Kantorovich mass transfer problem. Preprint.

[19] FRÉCHET, M., **Sur la distance de deux lois de probabilité.** *C. R. Acad. Sci. Paris*, 244 (1957), 689–692.

[20] GANGBO, W. & MCCANN, R. J., Optimal maps in Monge's mass transport problem. *C. R. Acad. Sci. Paris Sér. I Math.*, 321 (1995), 1653–1658.

[21] KANTOROVICH, L., On the translocation of masses. *C. R. (Doklady) Acad. Sci. URSS (N. S.)*, 37 (1942), 199–201.

[22] — On a problem of Monge. *Uspekhi Mat. Nauk.*, 3 (1948), 225–226 (in Russian).

[23] KATZ, B. S. (editor), *Nobel Laureates in Economic Sciences: A Biographical Dictionary.* Garland Publishing Inc., New York, 1989.

[24] KELLERER, H. G., Duality theorems for marginal problems. *Z. Wahrsch. Verw. Gebiete*, 67 (1984), 399–432.

[25] KNOTT, M. & SMITH, C. S., On the optimal mapping of distributions. *J. Optim. Theory Appl.*, 43 (1984), 39–49.

[26] LEVIN, V. L., General Monge–Kantorovich problem and its applications in measure theory and mathematical economics, in *Functional Analysis, Optimization, and Mathematical Economics* (L. J. Leifman, ed.), pp. 141–176. Oxford Univ. Press, New York, 1990.

[27] MCCANN, R. J., Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80 (1995), 309–323.

[28] — A convexity principle for interacting gases. To appear in *Adv. in Math.*

[29] — Exact solutions to the transportation problem on the line. To appear.

[30] MONGE, G., Mémoire sur la théorie des déblais et de remblais. *Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année*, 1781, pp. 666–704.

[31] RACHEV, S. T., The Monge–Kantorovich mass transference problem and its stochastic applications. *Theory Probab. Appl.*, 29 (1984), 647–676.

[32] ROCKAFELLAR, R. T., Characterization of the subdifferentials of convex functions. *Pacific J. Math.*, 17 (1966), 497–510.

[33] — *Convex Analysis.* Princeton Univ. Press, Princeton, NJ, 1972.

[34] RÜSCHENDORF, L., Bounds for distributions with multivariate marginals, in *Stochastic Orders and Decision Under Risk* (K. Mosler and M. Scarsini, eds.), pp. 285–310. IMS Lecture Notes—Monograph Series. Inst. of Math. Statist., Hayward, CA, 1991.

[35] — Fréchet bounds and their applications, in *Advances in Probability Distributions with Given Marginals* (G. Dall'Aglio et al., eds.), pp. 151–187. Math. Appl., 67. Kluwer Acad. Publ., Dordrecht, 1991.

[36] — Optimal solutions of multivariate coupling problems. *Appl. Math. (Warszaw)*, 23 (1995), 325–338.

[37] — On c-optimal random variables. *Statist. Probab. Lett.*, 27 (1996), 267–270.

[38] RÜSCHENDORF, L. & RACHEV, S. T., A characterization of random variables with minimum L^2-distance. *J. Multivariate Anal.*, 32 (1990), 48–54.

[39] SCHNEIDER, R., *Convex Bodies: The Brunn–Minkowski Theory.* Cambridge Univ. Press, Cambridge, 1993.

[40] SMITH, C. & KNOTT, M., On the optimal transportation of distributions. *J. Optim. Theory Appl.*, 52 (1987), 323–329.

[41] — On Hoeffding–Fréchet bounds and cyclic monotone relations. *J. Multivariate Anal.*, 40 (1992), 328–334.

[42] SUDAKOV, V. N., Geometric problems in the theory of infinite-dimensional probability distributions. *Proc. Steklov Inst. Math.*, 141 (1979), 1–178.

[43] ZAJÍČEK, L., On the differentiation of convex functions in finite and infinite dimensional spaces. *Czechoslovak Math. J.*, 29 (104) (1979), 340–348.

WILFRID GANGBO
School of Mathematics
Georgia Institute of Technology
Atlanta, GA 30332
U.S.A.
gangbo@math.gatech.edu

ROBERT J. MCCANN
Department of Mathematics
Brown University
Providence, RI 02912
U.S.A.
mccann@math.brown.edu
and
Institut des Hautes Etudes Scientifiques
91440 Bures-sur-Yvette
France

Received May 7, 1996