



INSTITUT DE FRANCE
Académie des sciences

Comptes Rendus

Mathématique

Bamdad Hosseini and Stefan Steinerberger

Intrinsic Sparsity of Kantorovich solutions

Volume 360 (2022), p. 1173-1175

<https://doi.org/10.5802/crmath.392>



This article is licensed under the
CREATIVE COMMONS ATTRIBUTION 4.0 INTERNATIONAL LICENSE.
<http://creativecommons.org/licenses/by/4.0/>



Les Comptes Rendus. Mathématique sont membres du
Centre Mersenne pour l'édition scientifique ouverte
www.centre-mersenne.org
e-ISSN : 1778-3569



Control theory / *Théorie du contrôle*

Intrinsic Sparsity of Kantorovich solutions

Parcité intrinsèque des solutions de Kantorovich

Bamdad Hosseini^a and Stefan Steinerberger^{*, b}

^a Department of Applied Mathematics, University of Washington, Seattle, USA

^b Department of Mathematics, University of Washington, Seattle, USA

E-mails: bamdadh@uw.edu (BH), steinerb@uw.edu (SS)

Abstract. Let X, Y be two finite sets of points having $\#X = m$ and $\#Y = n$ points with $\mu = (1/m)(\delta_{x_1} + \dots + \delta_{x_m})$ and $\nu = (1/n)(\delta_{y_1} + \dots + \delta_{y_n})$ being the associated uniform probability measures. A result of Birkhoff implies that if $m = n$, then the Kantorovich problem has a solution which also solves the Monge problem: optimal transport can be realized with a bijection $\pi : X \rightarrow Y$. This is impossible when $m \neq n$. We observe that when $m \neq n$, there exists a solution of the Kantorovich problem such that the mass of each point in X is moved to at most $n/\gcd(m, n)$ different points in Y and that, conversely, each point in Y receives mass from at most $m/\gcd(m, n)$ points in X .

Résumé. Soient X et Y deux ensembles de points de cardinaux respectifs m et n ; et μ, ν les mesures de probabilité associées. Quand $m = n$, un résultat de Birkhoff assure que le problème de Kantorovitch a une solution qui résout aussi le problème de Monge : une bijection réalise le transport optimal. Quand $m \neq n$, nous montrons que le problème de Kantorovitch admet une solution telle que chaque point de X se déplace vers au plus $n/\text{pgcd}(m, n)$ points de Y , et réciproquement, chaque point de Y reçoit de la masse d'au plus $m/\text{pgcd}(m, n)$ points de X .

2020 Mathematics Subject Classification. 49Q20, 90C46.

Funding. Both authors gratefully acknowledge support from the Kantorovich Initiative. SS is partially supported by the NSF (DMS-2123224) and the Alfred P. Sloan Foundation.

Manuscript received 3 June 2022, accepted 5 June 2022.

1. Introduction

Let μ and ν be two (probability) measures. A classical question, due to Monge, is to understand the optimal way of mapping μ to ν . If we denote the cost of transporting mass from x to y by $c(x, y)$, then the Monge problem asks for

$$\inf_T \left\{ \int_X c(x, T(x)) d\mu(x) : T_*(\mu) = \nu \right\} \quad (\text{Monge}),$$

* Corresponding author.

where $T_*(\mu)$ denotes the push forward of μ by T . This problem may not be solvable because such transport maps T may simply not exist. Kantorovich proposed to relax the problem and instead try to minimize

$$\inf_{\gamma} \int_{X \times Y} c(x, y) d\gamma(x, y) \quad (\text{Kantorovich}),$$

where γ is a probability measure on $X \times Y$ having marginals μ and ν . There is a nice classical result linking these two problems in the discrete setting: if μ and ν are two uniform probability measures over two sets X and Y with n elements, then it is known that these two problems coincide.

Theorem 1 (see e.g. [2, 4, 6]). *If $\mu = (1/n) \sum_{i=1}^n \delta_{x_i}$ and $\nu = (1/n) \sum_{i=1}^n \delta_{y_i}$, then there is a solution of the Kantorovich problem which also solves the Monge problem.*

The statement is independent of the transport costs $c(x_i, y_j)$. A proof is as follows: the Kantorovich problem is, in the discrete setting, a linear program over bistochastic matrices. A theorem of Birkhoff [1] (also attributed to König [3] and von Neumann [5]) says that the bistochastic matrices are the convex hull of the permutation matrices. The minimum of a linear program in a non-empty polyhedron is attained in an extremal point.

2. Main Result

No such statement can be true when $m \neq n$: the two sets have different cardinalities and no bijection is possible. The goal of this short note is to point out that there nonetheless exists a particularly simple solution of the Kantorovich problem.

Theorem 2. *Let $\mu = (1/m) \sum_{i=1}^m \delta_{x_i}$ and $\nu = (1/n) \sum_{i=1}^n \delta_{y_i}$. There is a solution of the Kantorovich problem such that mass from each point in X is moved to at most $n/\gcd(m, n)$ different points in Y and that each point in Y receives mass from at most $m/\gcd(m, n)$ points in X .*

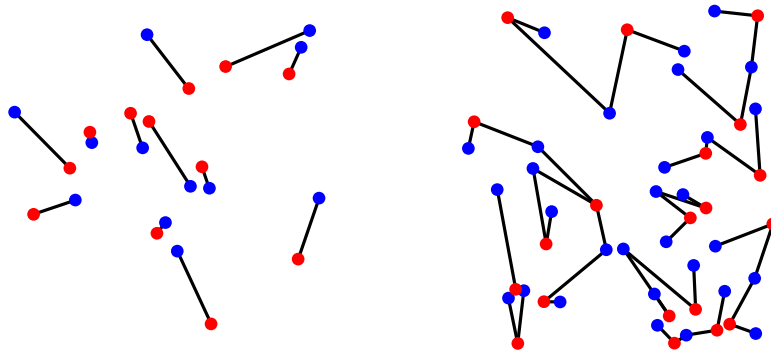


Figure 1. Left: $m = n$, the transport is a bijection. Right: $m = 20$ red points are sent to $n = 30$ blue points. Each red point is transported to at most $30/\gcd(20, 30) = 3$ blue points, each blue points receives mass from at most $20/\gcd(20, 30) = 2$ red points.

Somewhat to our surprise, we were unable to find this simple but intriguing statement (illustrated in Fig. 1) in the literature. Besides its intrinsic appeal, it does seem like it could be potentially useful insofar as it guarantees the existence of “sparse” solutions of the Kantorovich problem (with sparsity depending on m, n).

Proof. Suppose that

$$\mu = \frac{1}{m} \sum_{i=1}^m \delta_{x_i} \quad \text{and} \quad \nu = \frac{1}{n} \sum_{i=1}^n \delta_{y_i}$$

are two given measures. We replace each point x_i by $n/\gcd(m, n)$ identical points $x_{i,j}$ for $1 \leq j \leq n/\gcd(m, n)$ and, likewise, we replace each point y_i by $m/\gcd(m, n)$ identical points $y_{i,j}$ where $1 \leq j \leq m/\gcd(m, n)$. This allows us to write

$$\mu = \frac{\gcd(m, n)}{mn} \sum_{i=1}^m \sum_{j=1}^{n/\gcd(m, n)} \delta_{x_{i,j}} \quad \text{and} \quad \nu = \frac{\gcd(m, n)}{mn} \sum_{i=1}^n \sum_{j=1}^{m/\gcd(m, n)} \delta_{y_{i,j}}.$$

The problem can now be interpreted as finding a transport map from $mn/\gcd(m, n)$ points of the same weight to another set of $mn/\gcd(m, n)$ points of the same weight. Applying the classical result shows that there exists bijective map between the points that realizes the optimal Kantorovich cost. This corresponds into each point in X being split into at most $n/\gcd(m, n)$ equal parts and each point in Y being split into at most $m/\gcd(m, n)$ parts. \square

We conclude by observing that the same argument also applies to linear combinations of weighted Dirac measures as long as the weights are rational. Suppose

$$\mu = \sum_{i=1}^m \frac{a_i}{b_i} \delta_{x_i} \quad \text{where} \quad \sum_{i=1}^m \frac{a_i}{b_i} = 1$$

and where $a_i, b_i \in \mathbb{N}$ are positive rational weights. We note that it is possible to equivalently represent μ as a linear combination of a number of equally weighted Dirac measures. This number will depend on the least common multiple $\text{lcm}(b_1, \dots, b_m)$ of the denominators. This can be seen by writing

$$\mu = \frac{1}{\text{lcm}(b_1, \dots, b_m)} \sum_{i=1}^m \frac{a_i \text{lcm}(b_1, \dots, b_m)}{b_i} \delta_{x_i}$$

and noting that $a_i \text{lcm}(b_1, \dots, b_m) / b_i \in \mathbb{N}$. This implies the following corollary.

Corollary 3. *Let*

$$\mu = \sum_{i=1}^m \frac{a_i}{b_i} \delta_{x_i} \quad \text{and} \quad \nu = \sum_{i=1}^n \frac{c_i}{d_i} \delta_{y_i}$$

be two probability measures with positive rational weights and let

$$B = \text{lcm}(b_1, \dots, b_m) \quad \text{and} \quad D = \text{lcm}(d_1, \dots, d_n).$$

There exists a solution of the Kantorovich problem such that mass from each point in X is moved to at most $D/\gcd(B, D)$ different points in Y and each point in Y receives mass from at most $B/\gcd(B, D)$ different points in X .

We note that this reduces to the previous Theorem when $a_i = c_i = 1$, $b_i = m$ and $d_i = n$. We also observe that it is quite possible that $D/\gcd(B, D) \gg n$ and that $B/\gcd(B, D) \gg m$ (this will usually happen when the b_i, d_i have many different prime factors). In such a case, the statement would not say anything of interest.

References

- [1] G. Birkhoff, "Tres observaciones sobre el algebra lineal", *Rev., Ser. A, Univ. Nac. Tucumán* **5** (1946), p. 147-151.
- [2] H. Brezis, "Remarks on the Monge-Kantorovich problem in the discrete setting", *C. R. Math. Acad. Sci. Paris* **356** (2018), no. 2, p. 207-213.
- [3] D. König, *Theorie der endlichen und unendlichen Graphen*, Teubner-Archiv zur Mathematik, vol. 6, Teubner, 1936.
- [4] Q. Mérigot, B. Thibert, "Optimal transport: discretization and algorithms", in *Geometric partial differential equations. Part 2*, Handbook of Numerical Analysis, vol. 22, Elsevier, 2021, p. 133-212.
- [5] J. von Neumann, "A certain zero-sum two-person game equivalent to an optimal assignment problem", in *Contributions to the theory of games. Vol. 2*, Annals of Mathematics Studies, vol. 28, Princeton University Press, 1953, p. 5-12.
- [6] G. Peyré, M. Cuturi, "Computational optimal transport: With applications to data science", *Found. Trends Mach. Learn.* **11** (2018), no. 5-6, p. 355-407.