VARIATIONAL HEURISTICS FOR THE MONGE PROBLEM ON COMPACT MANIFOLDS

Philippe Delanoë

VARIATIONAL HEURISTICS FOR THE MONGE PROBLEM ON COMPACT MANIFOLDS*

P. DELANOË†

Abstract. We consider the Monge optimal transport problem posed on compact manifolds (possibly with boundary) when all data are smooth and the given measures, positive. If a diffeomorphism is stationary for the total cost, it was established that it must admit a potential function. If it realizes a local minimum of the total cost, we show that the c-Hessian of its potential function must be non-negative, positive if we further assume that the cost function c is non degenerate.

Key words. optimal transport, minimizing diffeomorphism, c-potential, c-Hessian positivity

AMS subject classifications. 53A45, 58D05, 58E99

1. Introduction and statement of results. The solution of Monge's problem [Monge 1781] in optimal transportation theory, with a general cost function, has been applied to many questions in various domains tentatively listed in the survey paper [Guillen–McCann 2010], including in cosmology [Brenier *et al* 2003]. The book [Villani 2009] offers a modern account on the theory (see also [Gangbo–McCann 1996] and references therein).

In case data are smooth, manifolds compact, measures positive, maps one-to-one and the solution of Monge's problem unique, the question of the *smoothness* of that solution was addressed in the landmark paper [Ma–Trudinger–Wang 2005]. In that case, restricting Monge's problem to diffeomorphisms becomes natural. Doing so, the use of differential geometry and the calculus of variations enables one to bypass the general optimal transportation approach and figure out directly some basic features of the solution map. Such a variational heuristics goes back to [Appell 1887] and was elaborated stepwise in [Evans 1998, Urbas 1998, Brenier *et al* 2003, Trudinger 2007, Delanoë 2009] (see also [Brenier *et al* 2003]). In the present note, we take a new step in that elaboration.

Specifically, working henceforth in the C^∞ category, we are given a couple of compact connected diffeomorphic manifolds each equipped with a probability volume measure, (M,μ) and (P,ϖ), and a cost function $c:\Omega\subset M\times P\to\mathbb{R}$ defined in a domain Ω projecting onto M and P by the canonical projections. Each manifold, either has no boundary, or it is the closure of a domain contained in some larger manifold. We consider the subset $\mathrm{Diff}_{\mu,\varpi}(\Omega)$ of diffeomorphisms from M to P, pushing μ to ϖ, with graph lying in Ω (an obvious item missing in [Delanoë 2009]). The pushing condition means for a Borel map $\varphi: M\to P$ that, for any continuous function $h:P\to\mathbb{R}$, the following equality holds:

$$\int_P h\, d\varpi=\int_M (h\circ\varphi)\, d\mu, \tag{1.1}$$

a property commonly denoted by $\varphi_\#\mu=\varpi$. The existence of diffeomorphisms pushing μ to ϖ is well-known [Moser 1965, Banyaga 1974, Dacorogna–Moser 1990]. Setting $\widetilde{\Omega}=\{(p,m)\in P\times M,(m,p)\in\Omega\}$, one defines similarly the set $\mathrm{Diff}_{\varpi,\mu}(\widetilde{\Omega})$. We view $\mathrm{Diff}_{\mu,\varpi}(\Omega)$ and $\mathrm{Diff}_{\varpi,\mu}(\widetilde{\Omega})$ as manifolds respectively modeled on the Fréchet

*Work supported by the CNRS (INSMI) at UMR 6621, UNS

†Université de Nice Sophia Antipolis (UNS), Faculté des Sciences, Laboratoire J.–A. Dieudonné, Parc Valrose, 06108 Nice Cedex 2, France (Philippe.Delanoe@unice.fr)

manifolds $\mathrm{Diff}_\varpi(P) := \mathrm{Diff}_{\varpi,\varpi}(P \times P)$ and $\mathrm{Diff}_\mu(M)$ (see details in [Delanoë 2009, Ebin–Marsden 1970]). Finally, we consider the total cost functional

$$\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega) \to \mathcal{C}(\varphi) = \int_M c(m, \varphi(m))\ d\mu,$$

and its counterpart $\psi \in \mathrm{Diff}_{\varpi,\mu}(\widetilde{\Omega}) \to \widetilde{\mathcal{C}}(\psi) = \int_P c(\psi(p), p)\ d\varpi$, which satisfy the identity: $\mathcal{C}(\varphi) \equiv \widetilde{\mathcal{C}}(\varphi^{-1})$. In this setting, the *restricted Monge problem* consists naturally in minimizing the functional $\mathcal{C}$ and the variational heuristics of [Delanoë 2009], in writing down the Euler equation of that functional. It yields the following result:

PROPOSITION 1.1. *If $\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$ is stationary for $\mathcal{C}$, so is φ^{-1} for $\widetilde{\mathcal{C}}$, and there exists two functions $f : M \to \mathbb{R}$, $\tilde{f} : P \to \mathbb{R}$, defined up to addition of constants, such that each point of the graph of φ is stationary for the real function:*

$$(m, p) \in \Omega \to c(m, p) + f(m) + \tilde{f}(p).$$

Let us call the function f (resp. $\tilde{f}$) so determined (up to constant addition), the c-potential of the diffeomorphism φ (resp. φ^{-1}). In [Delanoë 2009], we further assumed on the cost function c the generating condition sometimes called (A1) [Ma–Trudinger–Wang 2005] or bi-twist [Kim–McCann 2010], under which c-potentials completely determine the diffeomorphisms which give rise to them. Let us recall it: setting $\Omega^P_m = \{p \in P, (m, p) \in \Omega\}$, $\Omega^M_p = \{m \in M, (m, p) \in \Omega\}$ and d^M (resp. d^P) for the exterior derivative with respect to the argument in the manifold M (resp. P), we will say that c is *generating* if, for each $(m_0, p_0) \in \Omega$, the maps

$$\text{(1.2)}\quad p \in \Omega^P_{m_0} \to -\mathrm{d}^M c(m_0, p) \text{ and } m \in \Omega^M_{p_0} \to -\mathrm{d}^P c(m, p_0) \text{ are } \textit{one-to-one.}$$

For the moment, we do not assume this condition and stress that it is not required for the proof of Proposition 1.1, indeed solely based on the Helmholtz lemma. For the reader's convenience, let us indicate the argument (see [Delanoë 2009] for details). We write $\delta\mathcal{C} = 0$ with $\delta\mathcal{C} = \displaystyle\int_M \mathrm{d}^P c(m, \varphi(m))(\delta\varphi(\varphi(m))\ d\mu$, where $\delta\varphi$ stands for a variation of the transporting diffeomorphism φ which keeps it on the manifold $\mathrm{Diff}_{\mu,\varpi}(\Omega)$, that is, a vector field of a special kind on P, evaluated at the image point $\varphi(m)$. Specifically, such a vector field V on P should be: first of all *tangential* to the boundary of P, if any, so that its flow send P to itself (without crossing ∂P); moreover, its flow should preserve the volume measure ϖ or, equivalently, V should satisfy: $\mathrm{div}_\varpi V = 0$. In other words, the tangent space to $\mathrm{Diff}_{\mu,\varpi}(\Omega)$ at φ is spanned by the tangential vectors that write $V \circ \varphi$ with $V \in \ker \mathrm{div}_\varpi$. Here, the symbol div_ϖ denotes the divergence operator defined by the identity: $\int_P h\, \mathrm{div}_\varpi V\ d\varpi \equiv \int_P \mathrm{d}^P h(V)\ d\varpi$ valid for each function $h : P \to \mathbb{R}$ and each vector field V on P (tangential, as said[1]). Recalling (1.1), we thus find:

$$\forall V \in \ker \mathrm{div}_\varpi,\ \int_P \mathrm{d}^P c(\varphi^{-1}(p), p)(V(p))\ d\varpi = 0.$$

Arguing likewise on $\widetilde{\mathcal{C}}$, we further get:

$$\forall U \in \ker \mathrm{div}_\mu,\ \int_M \mathrm{d}^M c(m, \varphi(m))(U(m))\ d\mu = 0.$$

[1] otherwise, a boundary integral should occur, of course

The conclusion of Proposition 1.1 now readily follows from Helmholtz lemma, which we recall:

LEMMA 1.2 (Helmholtz). *Let (N,ν) be a measured manifold as above. A 1-form α on N satisfies: $\int_N \alpha(Z)\, d\nu = 0$ for each vector field $Z \in \ker \mathrm{div}_\nu$ (tangential to ∂N if $\partial N \neq \emptyset$) if and only if α is exact.*

The variational heuristics presented so far is incomplete since no local *minimum* condition is expressed yet for the total cost $\mathcal{C}$. It is our aim in the present note to write down that minimum condition and to derive from it further properties of minimizing diffeomorphisms. Before stating our results, we require a notion of c-Hessian.

DEFINITION 1.3. *Let $F: M \to P$ be a map whose graph lies in Ω and $h: M \to \mathbb{R}$ a function related to F by the equation $\mathrm{d}^M c(m, F(m)) + \mathrm{d}h(m) = 0$ on M. The* (c,F)-Hessian *of h is the covariant symmetric 2-tensor on M, denoted by $\mathrm{Hess}_{c,F}(f)$, intrinsically defined, in any couple of source and target charts (x,y), by:*

$$\mathrm{Hess}_{c,F}(h)(x_0) := \frac{\partial^2}{\partial x^i \partial x^j}\left[c(x,y_0) + h(x)\right] \text{ at } x = x_0,$$

where, if $x_0 = x(m_0)$, we have set $y_0 = y(F(m_0))$.

If $\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$ is stationary for the total cost $\mathcal{C}$, Proposition 1.1 shows that the couple (φ, f), with f the c-potential of φ, fulfills the assumption of Definition 1.3. In that case, for simplicity, we will simply speak of the c-Hessian of f and denote it by $\mathrm{Hess}_c(f)$. We would define likewise the c-Hessian of the c-potential $\tilde{f}$ of φ^{-1} by the local expression (stiking to the notations used in the preceding definition):

$$\mathrm{Hess}_c(\tilde{f})(y_0) := \frac{\partial^2}{\partial y^i \partial y^j}\left[c(x_0,y) + \tilde{f}(y)\right] \text{ at } y = y_0,$$

where, if $y_0 = y(p_0), x_0 = x\left(\varphi^{-1}(p_0)\right)$. We are in position to state our first result:

THEOREM 1.4. *Let $\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$ be stationary for the total cost $\mathcal{C}$. The following properties are equivalent:*

(i) *$\mathcal{C}$ admits a local minimum at φ;*

(ii) *the c-potential f of φ satisfies: $\forall U \in \ker \mathrm{div}_\mu, \int_M \mathrm{Hess}_c(f)(U,U)\, d\mu \geqslant 0$;*

(iii) *the c-potential $\tilde{f}$ of φ^{-1} satisfies: $\forall V \in \ker \mathrm{div}_\varpi, \int_P \mathrm{Hess}_c(\tilde{f})(V,V)\, d\varpi \geqslant 0$.*

Assuming (i), we will infer the *pointwise* non negativity of the c-Hessians from the conclusions (ii)(iii) of Theorem 1.4, arguing by contradiction with suitable localized divergence free vector fields. The resulting statement goes as follows:

COROLLARY 1.5. *If the total cost $\mathcal{C}$ admits a local minimum at $\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$, the c-Hessian of the c-potential of φ is non negative, and so is that of φ^{-1}.*

This corollary itself can be strengthened by making a *non degeneracy* assumption on the cost function c, called condition (A2) in [Ma–Trudinger–Wang 2005], namely:

$$\det(\mathrm{d}^M \mathrm{d}^P c) \neq 0 \text{ on } \Omega. \tag{1.3}$$

Under that condition, the implicit function theorem provides, for each $(m_0,p_0) \in \Omega$, setting $\alpha_0 = -\mathrm{d}^M c(m_0,p_0)$, neighborhoods $\mathcal{U}_0$ of p_0 in P and $\mathcal{W}_0$ of (m_0,α_0) in T^*M, and a map $\mathrm{E}_c^{(m_0,p_0)}: \mathcal{W}_0 \to \mathcal{U}_0$ such that:

$$\alpha \equiv -\mathrm{d}^M c\left(m, \mathrm{E}_c^{(m_0,p_0)}(m,\alpha)\right).$$

Following [Kim–McCann 2010], maps like $\mathrm{E}_c^{(m_0,p_0)}$ may be called local c-exponential maps. Observe that, for any (m_1,p_1) close enough to (m_0,p_0) in Ω such that, setting

$\alpha_1 = -\mathrm{d}^M c(m_1, p_1)$, the point (m_1, α_1) lies in $\mathcal{W}_0$, the maps $\mathrm{E}_c^{(m_0,p_0)}$ and $\mathrm{E}_c^{(m_1,p_1)}$ coincide on $\mathcal{W}_0 \cap \mathcal{W}_1$. If M is simply connected, one can thus extend uniquely the map $\mathrm{E}_c^{(m_0,p_0)}$ to a neighborhood $\mathcal{W}^{(m_0,\alpha_0)}$ of a *global* section of T^*M passing through α_0 at m_0 and call it the (m_0, p_0)-determination of the c-exponential map on M. With these notions at hand, we can state the announced strengthened corollary:

Corollary 1.6. *If the non degeneracy condition (1.3) holds and $\mathcal{C}$ admits a local minimum at $\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$, the c-Hessian of the c-potential f of φ must be positive definite, and so must be that of φ^{-1}; in particular, the local minimum of $\mathcal{C}$ must be strict. Moreover, the Jacobian equation expressing that φ lies in* $\mathrm{Diff}_{\mu,\varpi}(\Omega)$*, written locally in terms of f* via *a local exponential map, reads as a Monge–Ampère equation of* elliptic *type.*

To understand the second part of the statement, one should be aware that, if the pushing condition (1.1) is satisfied by a diffeomorphism $\varphi : M \to P$, one may use the change of variable $p = \varphi(m)$ in the left-hand integral of (1.1) and infer pointwise, in any couple (x, y) of source and target charts, keeping abusively the notation $y = \varphi(x)$ for the local expression of φ, the so-called *Jacobian equation* alluded to, namely:

$$\frac{d\varpi}{dy}(\varphi(x)) \left| \det \left(\frac{\partial \varphi}{\partial x} \right) (x) \right| = \frac{d\mu}{dx}(x), \tag{1.4}$$

where $\frac{d\mu}{dx}$ stands for the Radon–Nikodym derivative of the push-forward measure $x_{\#}\mu$ with respect to the Lebesgue measure dx of the chart x, and similarly for $\frac{d\varpi}{dy}$.

Finally, let us record a uniqueness result easily seen to hold in case the cost function c is *generating* non degenerate. If so, all local c-exponential maps coincide with the one (just denoted by E_c) globally defined by the generating condition. Summing up the main facts obtained in this note, we can state the following:

Proposition 1.7. *If the cost function satisfies (1.2)(1.3), there exists at most one local minimizer $\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$ of the total cost functional $\mathcal{C}$. If it exists, it should write $m \mapsto \varphi(m) = \mathrm{E}_c(m, df(m))$ for some c-potential f such that* $\mathrm{Hess}_c(f)$ *is positive definite on M; moreover, the local minimum $\mathcal{C}(\varphi)$ is strict.*

Proof. Indeed, if $\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$ is such a minimizer, by Proposition 1.1 it must admit a c-potential $f : M \to \mathbb{R}$ solving the Monge–Ampère equation obtained from (1.4) by setting $\varphi(m) \equiv \mathrm{E}_c(m, df(m))$. Furthermore, from Corollary 1.6, the c-Hessian of f must be positive definite, the local minimum, strict and the Monge–Ampère equation, elliptic. Being so, the latter admits at most one solution f up to addition of a constant, as a standard maximum principle argument shows □

The next three sections contain respectively the proofs of Theorem 1.4 and of its two corollaries.

2. Proof of Theorem 1.4. Let $\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$ be stationary for $\mathcal{C}$ and, for t a small real parameter, let $t \mapsto \varphi_t \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$ be an arbitrary path such that $\varphi_0 = \varphi$. We can uniquely write $\varphi_t = \xi_t \circ \varphi$ with $t \mapsto \xi_t \in \mathrm{Diff}_{\varpi}(P)$ such that $\xi_0 =$ the identity of P. In particular, $\dot{\xi}_0$ is a tangential vector field on P lying in $\ker \mathrm{div}_{\varpi}$ (setting as usual $\dot{\xi}_0(p) = \frac{\partial}{\partial t}\xi_t(p)|_{t=0}$). We will prove the equivalence $(i) \iff (ii)$ by establishing the equality:

$$\frac{d^2}{dt^2}\mathcal{C}(\varphi_t)|_{t=0} = \int_P \mathrm{Hess}_c(\tilde{f})(\dot{\xi}_0, \dot{\xi}_0)\ d\varpi. \tag{2.1}$$

Using arbitrary source and target charts x, y, and the Einstein summation convention, one routinely finds: $\frac{d}{dt}\mathcal{C}(\varphi_t) = \int_M \frac{\partial c}{\partial y^i}(x, \varphi_t(x)) \frac{\partial \varphi_t^i}{\partial t}(x)\ d\mu(x)$ where $d\mu(x) = \frac{d\mu}{dx}(x)dx$, then the second variation expression:

$$\frac{d^2}{dt^2}\mathcal{C}(\varphi_t) = \int_M \left\{ \frac{\partial c}{\partial y^i}(x,\varphi_t(x))\frac{\partial^2 \varphi_t^i}{\partial t^2}(x) + \frac{\partial^2 c}{\partial y^i \partial y^j}(x,\varphi_t(x))\frac{\partial \varphi_t^i}{\partial t}(x)\frac{\partial \varphi_t^i}{\partial t}(x) \right\} d\mu(x).$$

Here, to spare the unfamiliar reader, we did not use a global linear connection (on P) to compute the second derivatives which occur in the integrand. Doing so, we must be careful, as explained in the following remark.

Remark 1. Taken separately, the local scalar terms $\frac{\partial c}{\partial y^i}(x,\varphi_t(x))\frac{\partial^2 \varphi_t^i}{\partial t^2}(x)$ and $\frac{\partial^2 c}{\partial y^i \partial y^j}(x,\varphi_t(x))\frac{\partial \varphi_t^i}{\partial t}(x)\frac{\partial \varphi_t^i}{\partial t}(x)$ are *not* invariant under a change of charts, unlike their sum, indeed equal to the global real function: $m \in M \to \frac{\partial^2}{\partial t^2}c(m,\varphi_t(m))$. Therefore these terms *cannot be integrated separately* (unless the deformation ξ_t is supported in the domain of a single chart of P, of course). Still, in each couple of source and target charts, splitting the local expression $\frac{\partial^2}{\partial t^2}c(x,\varphi_t(x))$ of the global function into the above non invariant terms can be done in a *unique* way, by using the canonical flat connection of the P chart. Anytime we will have to integrate on a manifold a global real function splitting uniquely into a sum of non invariant local terms, we will stress that the integral should not be splitted by putting the sum between braces, as done above. As long as they are written in the *same charts*, the addition of two such sums $\{a_1 + a_2\} + \{b_1 + b_2\}$ may, of course, be written between a single pair of braces $\{a_1 + a_2 + b_1 + b_2\}$.

Back to our proof, using $\varphi_t = \xi_t \circ \varphi$, recalling (1.1) and setting $t = 0$, we infer:

$$\frac{d^2}{dt^2}\mathcal{C}(\varphi_t)|_{t=0} = \int_P \left\{ \ddot{\xi}_0^i \frac{\partial c}{\partial y^i}(\varphi^{-1}(y),y) + \dot{\xi}_0^i\dot{\xi}_0^j \frac{\partial^2 c}{\partial y^i \partial y^j}(\varphi^{-1}(y),y) \right\} d\varpi(y), \tag{2.2}$$

where $d\varpi(y) = \frac{d\varpi}{dy}(y)dy$ and $\ddot{\xi}_0^i = \frac{\partial^2}{\partial t^2}\xi_t^i(y)|_{t=0}$. To proceed further, we note that the integral $\int_P \tilde{f}(\xi_t(y))\ d\varpi(y)$ is independent of t. Differentiating it twice with respect to t at $t = 0$, we get:

$$0 = \int_P \left\{ \ddot{\xi}_0^i \frac{\partial \tilde{f}}{\partial y^i}(y) + \dot{\xi}_0^i\dot{\xi}_0^j \frac{\partial^2 \tilde{f}}{\partial y^i \partial y^j}(y) \right\} d\varpi(y).$$

Adding this vanishing integral to the right-hand side of (2.2), applying the last part of Remark 1 and recalling the stationary point equation:

$$\frac{\partial c}{\partial y^i}(\varphi^{-1}(y),y) + \frac{\partial \tilde{f}}{\partial y^i}(y) \equiv 0,$$

derived at once from Proposition 1.1, we obtain (2.1) as desired.

A similar argument would yield:

$$\frac{d^2}{dt^2}\widetilde{\mathcal{C}}(\varphi_t^{-1})|_{t=0} = \int_M \mathrm{Hess}_c(f)(\dot{\zeta}_0,\dot{\zeta}_0)\ d\mu,$$

with the vector field $\dot{\zeta}_0 \in \ker \mathrm{div}_\mu$ obtained by writing $\varphi_t^{-1} = \zeta_t \circ \varphi^{-1}$ for a unique path $t \mapsto \zeta_t \in \mathrm{Diff}_\mu(M)$. It would imply the other equivalence $(i) \iff (iii)$, since $\widetilde{\mathcal{C}}(\varphi^{-1}) \equiv \mathcal{C}(\varphi)$ is a local minimum of $\widetilde{\mathcal{C}}$ as well. The proof of Theorem 1.4 is complete.

3. Proof of Corollary 1.5.

Strategy. Let $\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$ realize a local minimum of the total cost $\mathcal{C}$ and let $f : M \to \mathbb{R}$ denote its c-potential, as provided by Proposition 1.1. Arguing by contradiction, we suppose the existence of a point $m_0 \in M$ such that the quadratic form associated to the symmetric bilinear one $\mathrm{Hess}_c(f)(m_0) : T_{m_0}M \times T_{m_0}M \to \mathbb{R}$ can take *negative* values. We will contradict property (*ii*) of Theorem 1.4 by constructing a vector field $U \in \ker \mathrm{div}_\mu$ supported near m_0 such that $\int_M \mathrm{Hess}_c(f)(U,U)\, d\mu < 0$. A similar argument would hold for φ^{-1}, of course.
We will proceed stepwise, choosing a good chart at m_0, constructing the vector field U in that chart and evaluating the above integral.

Choice of a chart. We pick any chart y of P at $\varphi(m_0)$ but a special chart x of M centered at m_0, namely a chart which pushes the measure $d\mu$ to the canonical Lebesgue measure dx. The existence of such μ-adapted charts, to call them so, is well-known [Banyaga 1974, Dacorogna–Moser 1990] and timely, here, to transform the div_μ operator on M into the usual div operator of $\mathbb{R}^n$ (up to sign) that is, the divergence operator associated to the measure dx (simply denoted by div below). Since the orthogonal group $O(n)$ preserves the measure dx, we may further choose the chart x such that the matrix $H_{ij}(0)$ of $\mathrm{Hess}_c(f)(m_0)$ is diagonal, with eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$ (each repeated with its multiplicity). Under our assumption: $\lambda_1 < 0$. Since the unimodular group $SL(n,\mathbb{R})$ preserves the measure dx, we may rescale the chart x in order to have: $\lambda_1 \leqslant -3$ and, $\forall i \in \{2,\ldots,n\}$, $\lambda_i \leqslant \frac{1}{2}$. Let the chart x be fixed so and let $x \mapsto H_{ij}(f)(x)$ denote the local expression of the map $m \mapsto \mathrm{Hess}_c(f)(m)$. The inequality: $\forall v \in \mathbb{R}^n$, $H_{ij}(f)(0)v^iv^j \leqslant -3(v^1)^2 + \frac{1}{2}\sum_{i=2}^n (v^i)^2$, combined with the continuity of the map $(x,v) \mapsto H_{ij}(f)(x)v^iv^j$ as (x,v) varies near $x = 0$ with v of length 1 (say), implies the existence of a real $\varepsilon > 0$ such that:

$$\forall (x,v) \in \mathbb{R}^n \times \mathbb{R}^n,\ \max_{1\leqslant i\leqslant n} |x^i| \leqslant \varepsilon \Rightarrow H_{ij}(f)(x)v^iv^j \leqslant -2(v^1)^2 + \sum_{i=2}^n (v^i)^2. \tag{3.1}$$

Construction of a divergence free vector field. The vector field on $\mathbb{R}^n$ given by $w(x) = x^1\dfrac{\partial}{\partial x^2} - x^2\dfrac{\partial}{\partial x^1}$ satisfies $\mathrm{div}(w) = 0$. The flow of w preserves any function $h : \mathbb{R}^n \to \mathbb{R}$ factoring through a function $H : [0,\infty) \times \mathbb{R}^{n-2} \to \mathbb{R}$ as: $h(x) = H(\sqrt{(x^1)^2 + (x^2)^2}, x^3, \ldots, x^n)$. Any such function h thus satisfies $\mathrm{div}(hw) = 0$.
Let us fix a cut-off function $\alpha : [0,\infty) \to [0,1]$ equal to 1 on $[0,\varepsilon/2]$, vanishing on $[\varepsilon/\sqrt{2},\infty)$, decreasing in-between, and consider the function $h : \mathbb{R}^n \to \mathbb{R}$ given by:

$$h(x) = \alpha\left(\sqrt{(x^1)^2 + (x^2)^2}\right)\prod_{i=3}^n \alpha\left(|x^i|\right).$$

We know that $\mathrm{div}(hw) = 0$ and that the vector field hw is supported in the open box $B^n_\varepsilon = \left\{x \in \mathbb{R}^n, \max_{1\leqslant i\leqslant n}|x^i| < \varepsilon\right\}$. Since the chart x is μ-adapted, we may view hw as the expression in that chart of a div_μ free vector field U in M supported in the inverse image $x^{-1}(B^n_\varepsilon) \subset M$.

Calculation of an integral. Let us consider the integral $\int_M \mathrm{Hess}_c(f)(U,U)d\mu$ which is equal to: $\int_{B^n_\varepsilon} h^2(x)H_{ij}(f)(x)w^iw^j\,dx$. From (3.1), it is bounded above by:

$$\left(2\int_0^\infty \alpha^2(\rho)d\rho\right)^{n-2}\int_{B^2_\varepsilon}\alpha^2\left(\sqrt{(x^1)^2+(x^2)^2}\right)\left(-2(x^2)^2 + (x^1)^2\right)dx^1dx^2.$$

Note that the function $(x^1, x^2) \to \alpha\left(\sqrt{(x^1)^2 + (x^2)^2}\right)$ vanishes outside the Euclidean ball of radius ε centered at 0; using polar coordinates (r, θ) in $\mathbb{R}^2 \setminus \{0\}$, we thus find that the last integral is equal to:

$$\int_0^\varepsilon \alpha^2(r) r^3 dr \times \int_0^{2\pi} \left(\cos^2\theta - 2\sin^2\theta\right) d\theta \equiv -\pi \int_0^\varepsilon \alpha^2(r) r^3 dr.$$

We conclude that $\int_M \mathrm{Hess}_c(f)(U,U) d\mu$ is bounded above by a *negative* real, namely by $-\pi\left(2\int_0^\infty \alpha^2(\rho) d\rho\right)^{n-2} \int_0^\varepsilon \alpha^2(r) r^3 dr$, contradicting property (ii) of Theorem 1.4, as desired.

4. Proof of Corollary 1.6. Let again $\varphi \in \mathrm{Diff}_{\mu,\varpi}(\Omega)$ realize a local minimum of the total cost $\mathcal{C}$. Fix an arbitrary point $m_0 \in M$ and take a μ-adapted chart x of M centered at m_0 and a ϖ-adapted chart y of P centered at $p_0 = \varphi(m_0)$. From (1.3) and the definition of $\mathrm{E}_c^{(m_0,p_0)}$, the map $x \mapsto y = \varphi(x)$ is defined near $x = 0$ by the equation:

$$\frac{\partial f}{\partial x^i}(x) + \left(\frac{\partial c}{\partial x^i}(x,y)\right)_{y=\varphi(x)} = 0. \tag{4.1}$$

Differentiating the latter yields (sticking to the notation $\mathrm{Hess}_c(f) = H_{ij}(f) dx^i dx^j$):

$$H_{ij}(f)(x) = -\left(\frac{\partial^2 c}{\partial y^k \partial x^i}(x,y)\right)_{y=\varphi(x)} \frac{\partial \varphi^k}{\partial x^i}(x).$$

Taking determinants, recalling that the symmetric matrix $H_{ij}(f)(x)$ is non negative (by Corollary 1.5) and using (1.4), we obtain a local Monge–Ampère equation satisfied by f, namely:

$$\det\left(H_{ij}(f)(x)\right) = \left|\det\left(\frac{\partial^2 c}{\partial y^k \partial x^i}(x,y)\right)_{y=\varphi(x)}\right|,$$

where $y = \varphi(x)$ is given by (4.1). From (1.3), the right-hand side of this equation nowhere vanishes; so the matrix $H_{ij}(f)(x)$ must be positive definite. So must be the c-Hessian of f throughout the manifold M, since the point m_0 is arbitrary. Finally, as well-known, the positive definiteness just obtained implies, indeed, the ellipticity of the Monge–Ampère equation. A similar argument would hold for φ^{-1}.

Acknowledgments. The author got the idea of this note soon after returning from a visit at the Fields Institute during the "Concentration period on Partial Differential Equations and Geometric Analysis" (October 25-29, 2010). He wishes to thank the Institute for partial support and Robert McCann for his extremely nice invitation and a stimulating discussion.

REFERENCES

[Appell 1887] P. Appell, *Mémoire sur les déblais et remblais des systèmes continus ou discontinus*, Mémoires Acad. Sci. Inst. France Paris, 29 (1887), pp. 1-208.

[Banyaga 1974] A. Banyaga, *Formes volume sur les variétés à bord*, Enseignement Math., 20 (1974), pp. 127-131.

[Brenier *et al* 2003] Y. Brenier, U. Frisch, M. Hénon, G. Loeper, S. Matarrese, R. Mohayaee, and A. Sobolevskiĭ, *Reconstruction of the early Universe as a convex optimization problem*, Mon. Not. R. Astron. Soc., 346 (2003), pp. 501-524.

[Dacorogna–Moser 1990] B. Dacorogna and J. Moser, *On a partial differential equation involving the Jacobian determinant*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 7 (1990), pp. 1-26.

[Delanoë 2009] P. Delanoë, *Variational heuristics for optimal transportation maps on compact manifolds*, Analysis, 29 (2009), pp. 221-228.

[Delanoë–Ge 2010] P. Delanoë and Y. Ge, *Regularity of optimal transport on compact, locally nearly spherical, manifolds*, J. reine angew. Math., 646 (2010), 65–115.

[Ebin–Marsden 1970] D. G. Ebin and J. Marsden, *Groups of diffeomorphisms and the motion of an incompressible fluid*, Ann. Math., 92 (1970), pp. 102-163.

[Evans 1998] L. C. Evans, *Partial differential equations and Monge–Kantorovichmass transfer*, in: Current Developments in Mathematics 1997, International Press, Boston Mass. (1998), pp. 65-126.

[Gangbo–McCann 1996] W. Gangbo and R. J. McCann, *The geometry of optimal transportation*, Acta Math., 177 (1996), pp.113-161.

[Guillen–McCann 2010] N. Guillen and R. J. McCann, *Five lectures on optimal transportation: geometry, regularity and applications*, Preprint (2010), presently downloadable at: `http://www.math.toronto.edu/~mccann/publications`

[Kim–McCann 2010] Y.–H. Kim and R. J. McCann, *Continuity, curvature, and the general covariance of optimal transportation*, J. Eur. Math. Soc., 12 (2010), pp. 1009-1040.

[Ma–Trudinger–Wang 2005] X.–N. Ma, N. S. Trudinger and X.–J. Wang, *Regularity of potential functions of the optimal transportation problem*, Arch. Rat. Mech. Anal., 177 (2005), pp. 151-183.

[Monge 1781] G. Monge, *Mémoire sur la théorie des déblais et remblais*, Mémoires Acad. Royale Sci. Paris (1781).

[Moser 1965] J. Moser, *On the volume elements on a manifold*, Transac. Amer. Math. Soc., 120 (1965), pp. 286-294.

[Trudinger 2007] N. S. Trudinger, *Optimal transportation and nonlinear partial differential equations* (slides), 26th Brazilian Mathematical Colloquium, August 2007, presently downloadable at: `http://maths.anu.edu.au/~neilt/RecentPapers.html`

[Urbas 1998] J. Urbas, *Mass transfer problems*, Univ. Bonn Lecture Notes (1998).

[Villani 2009] C. Villani, *Optimal Transport, Old and New*, Grund. math. Wiss. 338, Springer–Verlag berlin Heidelberg (2009).